



UNIVERSITY OF
THESSALY

UNIVERSITY OF THESSALY
SCHOOL OF AGRICULTURAL SCIENCES
DEPARTMENT OF AGRICULTURE CROP PRODUCTION AND RURAL ENVIRONMENT
LABORATORY OF BIOMETRY



PhD Thesis

Statistical methods and experimental design for agreement and similarity studies between open-source software and hardware devices, and corresponding agriculture technology (agritech) market devices.

by
Evmorfia P. Bataka

Under the supervision of
Christos Nakas, Professor

«The implementation of the doctoral thesis was co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the Act "Enhancing Human Resources Research Potential by undertaking a Doctoral Research" Sub-action 2: IKY Scholarship Programme for PhD candidates in the Greek Universities».

Volos, 2023

Statistical methods and experimental design for agreement and similarity studies between open-source software and hardware devices, and corresponding agriculture technology (agritech) market devices.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΓΕΩΠΟΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΓΕΩΠΟΝΙΑΣ ΦΥΤΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΑΓΡΟΤΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ
ΠΡΟΓΡΑΜΜΑ ΔΙΔΑΚΤΟΡΙΚΩΝ ΣΠΟΥΔΩΝ



Διδακτορική Διατριβή

**Στατιστικές μέθοδοι και πειραματικός σχεδιασμός ερευνών
συμφωνίας και ομοιότητας συσκευών ανοιχτού κώδικα και υλικού
με αντίστοιχες εμπορικές για εφαρμογές στη γεωπονία.**

Ευμορφία Π. Μπατάκα

Επιβλέπων:
Χρήστος Νάκας, Καθηγητής

ΒΟΛΟΣ 2023

Supervisor:

Christos Nakas, Professor, University of Thessaly

Advisory Committee:

Christos Nakas, Professor, University of Thessaly

Nikolaos Papadopoulos, Professor, University of Thessaly

Nikolaos Katsoulas, Professor, University of Thessaly

Examining Committee:

Christos Nakas, Professor, University of Thessaly

Nikolaos Papadopoulos, Professor, University of Thessaly

Nikolaos Katsoulas, Professor, University of Thessaly

Aris Kyparissis, Associate Professor, University of Thessaly

Persefoni Maletsika, Assistant Professor, University of Thessaly

Georgios Stamoulis, Professor, University of Thessaly

Georgios Papadopoulos, Associate Professor, University of Athens

Evmorfia P. Bataka

Statistical methods and experimental design for agreement and similarity studies between open-source software and hardware devices, and corresponding agriculture technology (agritech) market devices.

Evmorfia P. Bataka, 2023: Statistical methods and experimental design for agreement and similarity studies between open-source software and hardware devices, and corresponding agriculture technology (agritech) market devices.

Number of introductory pages: 10

Total number of pages of the thesis: 257

Number of Figures: 87

Number of Tables: 52

Number of bibliographic references: 328

To Navis and Odysseas

Acknowledgements

First and foremost, I am extremely grateful to my esteemed mentor, Professor Christos Nakas for his keen involvement, continuous support, and guidance throughout the entire period of my research. This dissertation could not have been possible without his constant help. These five years were a journey full of great experiences and challenges and I feel proud being a member of the laboratory of Biometry. I will be forever grateful to him for his trust, patience, and guidance throughout all the stages of this thesis.

I am also thankful to the other two members of my Advisory committee Professor Nikolaos Papadopoulos and Professor Nikolaos Katsoulas, for their trust, guidance, and valuable comments.

Moreover, I would like to thank my examining committee Associate Professor Aris Kyparissis for his guidance, advice and comments, Assistant Professor Persefoni Maletsika for her valuable help, collaboration and comments, Professor Georgios Stamoulis and Associate Professor Georgios Papadopoulos for their help and comments.

Special thanks to Associate Professor Rania Pavli for all the encouragement, great support, collaboration, and aid through all these years. Thanks, should also go to Dr. Natasa Aggelaki for her aid and advice, our discussions and her great taste in music.

I would also like to thank my parents Vasiliki Bataka and Panagiotis Batakas. This whole journey could never have been completed without my mother. Her love and support through all the years of my studies and especially the last 5 years were priceless. Thank you, mom, for everything. Thanks, should also go to Virginia Chatzikyriakou for her support and encouragement.

I am also grateful to my friend and colleague Anastasia Mourantian for our discussions and her support all those days that seemed never-ending and tough. You are a great friend, and I am happy we met. Moreover, I would like to thank my friends and colleagues Eva Tsoumalakou, Evi Stathopoulou, Konstantinos Vlachopoulos, Leonidas Kyrgiakos, Chrysa Foti, Chrysanna Chatzipetrou, Vasilisa Tsipni, George Dimakopoulos, Kleopatra Moraiti, Stella Papanastasiou, Georgia Baliota and Vaso Aslanoglou for playing a major part of this journey.

I would also like to thank Vanina, my sister, even if you were miles away you were always close. Thank you so much for these 2 to 3-hour video calls that gave me courage and patience to go on. Moreover, I would like to thank my friend Anca, who always supported me regardless of the distance between us.

My dear Navis and Odysseas, this is all dedicated to you for all these hours I was away, but my mind and heart was always on you.

Finally, I would like to express my profound gratitude to my beloved spouse Kostas Chatzikyriakou who continuously supported me, and always believed in me.

Περίληψη

Η παρούσα διατριβή έχει ως στόχο να συνεισφέρει στις μεθοδολογικές προσεγγίσεις που αφορούν την αξιολόγηση της συμφωνίας και ομοιότητας μεταξύ συσκευών ανοιχτού κώδικα και υλικού σε σχέση με αντίστοιχες εμπορικές με τη χρήση στατιστικών μεθόδων και πειραματικού σχεδιασμού. Οι συσκευές ανοιχτού κώδικα και υλικού χρησιμοποιούνται σε μεγάλο αριθμό εφαρμογών σε τομείς υγείας, εκπαίδευσης, γεωπονίας, αθλημάτων, βιολογίας, έρευνας και άλλων. Η παρούσα διατριβή εστιάζει στους τομείς της γεωπονίας και έρευνας.

Το Κεφάλαιο 1 παρουσιάζει μία σύντομη εισαγωγή για τους παράγοντες που έχουν συνεισφέρει στην διάδοση των συσκευών ανοιχτού κώδικα και υλικού. Επιπλέον παραθέτει ορισμούς για τον ανοιχτό κώδικα και το υλικό όπως και την περιγραφή των εφαρμογών τους στη γεωργία ακριβείας, έξυπνη γεωργία και έρευνα. Ακόμα επισημαίνει τη φιλοσοφία του ανοιχτού κώδικα και υλικού όπως και την ελεύθερη διάδοση της γνώσης για κάθε άνθρωπο που επιθυμεί να αντιγράψει, τροποποιήσει ή δημιουργήσει συσκευές της συγκεκριμένης φιλοσοφίας ενώ ταυτόχρονα μοιράζεται τη διαδικασία και όλα τα υλικά και μεθόδους που χρησιμοποίησε. Παρουσιάζονται παραδείγματα δημοσιεύσεων που αφορούν τη συγκεκριμένη φιλοσοφία όπως και τα πλεονεκτήματα χρήσης τους ενώ αναγνωρίζεται η ανάγκη ανάπτυξης πλαισίου για την αξιολόγηση των λειτουργιών και της αξιοπιστίας των μετρήσεων των συγκεκριμένων συσκευών. Συγκεκριμένα, ορίζεται η έννοια των ερευνών συμφωνίας και ομοιότητας και η συνεισφορά τους στην αξιολόγηση και στον θετικό ή αρνητικό προσδιορισμό της χρήσης των δύο συσκευών εναλλακτικά χωρίς να επηρεάζει κάποιο αποτέλεσμα.

Το Κεφάλαιο 2 περιγράφει τις μεθοδολογίες που χρησιμοποιήθηκαν για την αξιολόγηση της συμφωνίας και ομοιότητας μεταξύ συσκευών ανοιχτού κώδικα και των αντίστοιχων εμπορικών. Συγκεκριμένα, παρέχει μία εκτεταμένη συζήτηση για τη χρήση διαφορετικών μοντέλων ανάλογα με τα δεδομένα που παράγονται από συσκευές υπό διερεύνηση και την αξιολόγηση δεικτών και γραφημάτων που ερμηνεύονται εύκολα ακόμα και από άτομα χωρίς ιδιαίτερη εμπειρία στη στατιστική. Επιπλέον, σε περίπτωση που το συμπέρασμα της έρευνας καταλήγει σε μη αποδεκτή συμφωνία, πιθανές πηγές ασυμφωνίας μπορούν να εντοπιστούν και μέθοδοι διόρθωσης της συμφωνίας μπορούν να εφαρμοστούν. Ακόμα, αναφέρονται τα πλεονεκτήματα και μειονεκτήματα κάθε προσέγγισης για την αξιολόγηση της συμφωνίας και ομοιότητας. Περαιτέρω παρουσιάζεται ο πειραματικός σχεδιασμός και η στατιστική ανάλυση που ακολουθείται ανάλογα με τη φύση της έρευνας καθώς και τα βήματα για τον έγκυρο σχεδιασμό της. Επιπλέον παρουσιάζονται λανθασμένες ή ελλείψεις μέθοδοι αξιολόγηση της συμφωνίας και ομοιότητας μεταξύ δύο συσκευών. Τέλος, αναφέρονται συνοπτικά οι μεθοδολογίες όπως το υλικό και το λογισμικό που χρησιμοποιήθηκε για το σχεδιασμό των συσκευών για την ολοκλήρωση της διατριβής.

Στο Κεφάλαιο 3 παρουσιάζεται ο σχεδιασμός και η ανάπτυξη ενός ασύρματου μετεωρολογικού σταθμού ανοιχτού κώδικα και υλικού που μετρά τη θερμοκρασία και υγρασία του περιβάλλοντος. Αρχικά, αξιολογείται η λειτουργία του σταθμού παρακολουθώντας τις μετρήσεις μέσω ασύρματης τεχνολογίας Wi-Fi για επτά μέρες. Έπειτα, σχεδιάζεται πείραμα αξιολόγησης συμφωνίας και ομοιότητας του σταθμού με αντίστοιχο εμπορικό με αισθητήρα Thygro και τα αντίστοιχα περιφερειακά από την εταιρεία Symmetron. Αναφέρεται ο πειραματικός σχεδιασμός καθώς και η χρήση μοντέλων μεικτών επιδράσεων μαζί με τη συνεισφορά τους στη σύγκριση των δύο συσκευών. Για το πείραμα κατασκευάστηκαν τριάντα συσκευές ανοιχτού κώδικα και υλικού και χρησιμοποιήθηκαν τρεις αντίστοιχοι εμπορικοί σταθμοί. Για τον λόγο αυτό δημιουργήθηκαν τρεις ομάδες. Κάθε ομάδα περιέχει 10 σταθμούς ανοιχτού κώδικα και υλικού και μία εμπορική. Κάθε πειραματική μονάδα παράγει ένα σύμπλεγμα μετρήσεων σε συνάρτηση του χρόνου. Το πείραμα τερματίστηκε μετά από επτά μέρες συνεχόμενων μετρήσεων. Καθώς οι

μετρήσεις ανά πειραματική μονάδα ήταν πάνω από 220 για κάποιες μέρες (αφού υπήρχαν απώλειες δεδομένων λόγω κακής σύνδεσης στο διαδίκτυο) οι μετρήσεις συγχωνεύτηκαν σε χρονικές ζώνες, με τη μέθοδο που χρησιμοποιεί το Εθνικό Αστεροσκοπείο Αθénas. Μόνο δύο από τις τρεις μεθοδολογίες συμφωνίας και ομοιότητας υλοποιήθηκαν με επιτυχία και εκτιμήθηκε ένας δείκτης συμφωνίας. Η συμφωνία μεταξύ των συσκευών θεωρήθηκε ανεπαρκής και ως πιθανή πηγή ασυμφωνίας θεωρήθηκε η ακρίβεια των αισθητήρων. Περεταίρω έρευνα μπορεί να υλοποιηθεί για να συμπεριληφθούν όλα τα σημεία των μετρήσεων όπως και επιπλέον παράγοντες που πιθανόν να επηρεάζουν τη συμφωνία, όπως η υγρασία του αέρα ή η τοποθεσία κάθε σταθμού. Τέλος, εναλλακτικές μορφές ασύρματης επικοινωνίας όπως η LoRa μπορούν να προστεθούν στο σταθμό όπως και επιπλέον αισθητήρες. Τα πλεονεκτήματα χρήσης της νέας συσκευής στηρίζονται στο χαμηλό κόστος παραγωγής του όπως και στην εύκολη μορφοποίηση και προσθήκη επιπλέον αισθητήρων.

Στο Κεφάλαιο 4 παρουσιάζεται ο σχεδιασμός και η κατασκευή ενός ασύρματου καταγραφέα LoRa ανοιχτού κώδικα και υλικού για τη χρήση σε πειράματα ενυδρείου σε σάκους υποστρώματος περλίτη. Συγκεκριμένα, η νέα συσκευή είναι συμβατή με αισθητήρες περιεκτικότητας νερού, με πρωτόκολλο επικοινωνίας RS-485. Επιπλέον, ο αισθητήρας που χρησιμοποιήθηκε για το συγκεκριμένο πείραμα μετρά τη θερμοκρασία του σάκου υποστρώματος, την ηλεκτρική αγωγιμότητα καθώς και το pH. Μετά τον έλεγχο καλής λειτουργίας του καταγραφέα, σχεδιάστηκε το πείραμα αξιολόγησης συμφωνίας και ομοιότητας της νέας συσκευής με αντίστοιχη εμπορική. Συγκεκριμένα η εμπορική συσκευή που χρησιμοποιήθηκε για το πείραμα είναι η Handheld Meter (HH150) and SM150 Soil Moisture Sensor. Το μέγεθος δείγματος ήταν 100 σάκοι υποστρώματος περλίτη. Επιπλέον, για την ανάλυση χρησιμοποιήθηκε μοντέλο μέτρησης-σφαλμάτων για την εκτίμηση της διαφοράς στις μετρήσεις σε σχέση με τον καλύτερο γραμμικό αμερόληπτο παράγοντα πρόβλεψης της πραγματικής τιμής της μέτρησης που παράχθηκαν από την εμπορική συσκευή. Τα αποτελέσματα ανέφεραν σημαντική απόκλιση από την εκτίμηση της πραγματικής τιμής των μετρήσεων παραθέτοντας το σταθερό και αναλογικό σφάλμα. Η συμφωνία των δύο οργάνων θεωρήθηκε ικανοποιητική για συγκεκριμένα διαστήματα μετρήσεων. Τα πλεονεκτήματα χρήσης της νέας συσκευής συγκριτικά με της εμπορικής είναι το χαμηλό κόστος παραγωγής και η ευκολία μορφοποίησης της ανάλογα με τις ανάγκες των πειραμάτων. Ο σάκος υποστρώματος περλίτη έχει αρκετή παραλλακτικότητα ως προς την περιεκτικότητα του σε νερό. Επομένως η χρήση πολλαπλών αισθητήρων σε διαφορετικά σημεία του σάκου είναι θεμελιώδης και αποτελεί βασική απαίτηση ως εισροή πληροφορίας σε αλγόριθμους πρόβλεψης περιεκτικότητας νερού στο σάκο.

Στο Κεφάλαιο 5 παρουσιάζεται ο σχεδιασμός και η κατασκευή καταγραφέα ανοιχτού κώδικα και υλικού που μετρά pH διαλυμάτων για τη χρήση σε πειράματα Δενδροκομίας. Για τη βαθμονόμηση της συσκευής χρησιμοποιήθηκαν ρυθμιστικά διαλύματα pH τιμών 4,01 και 7,01. Έπειτα αξιολογήθηκε η καλή λειτουργία της συσκευής μετρώντας το pH καρπών εσπεριδοειδών. Στη συσκευή προστέθηκε αδιάβροχος αισθητήρας θερμοκρασίας για την εξισορρόπηση της τιμής pH. Επιπλέον, σχεδιάστηκε μελέτη αξιολόγησης της συμφωνίας και ομοιότητας της νέας συσκευής με αντίστοιχη εμπορική. Συγκεκριμένα, η εμπορική συσκευή που επιλέχθηκε είναι η Hanna HI9024 Waterproof pHMeter. Για την ανάλυση χρησιμοποιήθηκαν μοντέλα μεικτών επιδράσεων και παράχθηκαν δείκτες και χρήσιμα γραφήματα υλοποιώντας παραπάνω από μία μεθοδολογίες. Έπειτα, αναφέρθηκαν τα πλεονεκτήματα και τα μειονεκτήματα κάθε μεθοδολογίας. Τα αποτελέσματα έδειξαν μη αποδεκτή συμφωνία μεταξύ των δύο οργάνων. Η ανάλυση ομοιότητας ανέφερε σταθερή απόκλιση -0.2215 μονάδων pH. Μετά τη χρήση μεθόδων αναβαθμονόμησης η συμφωνία θεωρείται άριστη. Η συσκευή είναι φορητή και μπορεί να χρησιμοποιηθεί για “επι τόπου” μετρήσεις στον αγρό ή στο θερμοκήπιο. Τα πλεονεκτήματα χρήσης της συσκευής είναι το χαμηλό κόστος παραγωγής, περίπου 4 φορές μικρότερο συγκριτικά με τη Hanna HI9024. Περαιτέρω έρευνα μπορεί να υλοποιηθεί για τη χρήση της συσκευής για τη μέτρηση pH εδάφους ή θρεπτικών υποστρωμάτων. Τέλος, επιπλέον αισθητήρες μπορούν να τοποθετηθούν εύκολα στη

συσκευή για την ταυτόχρονη άντληση περισσότερων πληροφοριών για τα χαρακτηριστικά του δείγματος.

Στο Κεφάλαιο 6 περιγράφεται η κατασκευή συσκευής ανοιχτού κώδικα και υλικού για τις ανάγκες πειραμάτων που αφορούν τη μύγα της Μεσογείου (*Ceratitis Capitata*). Η συσκευή καταγράφει τη θερμοκρασία στην επιφάνεια και τον πυρήνα καρπών όπως και την εξωτερική θερμοκρασία. Επιπλέον, αξιολογήθηκε η λειτουργία της στο εργαστήριο αλλά και στον αγρό. Για τα πειράματα στον αγρό η συσκευή τοποθετήθηκε σε συγκεκριμένα σημεία σε ποικιλία δέντρων. Το σχήμα αλλά και το μέγεθος της συσκευής καθιστά ιδανική την τοποθέτηση του σε οποιοδήποτε σημείο του δέντρου επιθυμεί ο ερευνητής. Η συσκευή είναι φορητή και λειτουργεί με μπαταρία. Η αυτονομία της μπορεί να φτάσει στις 7 έως 30 μέρες ανάλογα με το χρονικό διάστημα καταγραφής. Έπειτα σχεδιάστηκε μελέτη συμφωνίας και ομοιότητας για τη σύγκριση της θερμοκρασίας του πυρήνα και της επιφάνειας του καρπού καθώς ο στόχος του πειράματος ήταν η ποσοτικοποίηση της απόκλισης των θερμοκρασιών των δύο σημείων σε διαφορετικές συνθήκες. Για την ανάλυση χρησιμοποιήθηκαν μοντέλα μέτρησης-σφαλμάτων. Τα αποτελέσματα αναφέρουν σταθερή απόκλιση $0,331^{\circ}\text{C}$ καθώς και αναλογική απόκλιση 0,982. Επιπλέον, η σύγκριση ακρίβειας των μετρήσεων στα δύο σημεία υποδηλώνει ότι ο πυρήνας διατηρεί τη θερμοκρασία περισσότερο συγκριτικά με την επιφάνεια σε διαφορετικές συνθήκες. Τέλος, σε μελλοντικές μελέτες, ασύρματη τεχνολογία LoRa μπορεί να προστεθεί στη συσκευή, ηλιακοί συλλέκτες για περισσότερη αυτονομία καθώς και ρολόι ακριβείας για την καλύτερη λειτουργία των κύκλων αδρανοποίησης της συσκευής. Με τις συγκεκριμένες τροποποιήσεις η συσκευή μπορεί να λειτουργεί για μήνες λόγω των λειτουργιών χαμηλής κατανάλωσης.

Το κεφάλαιο επτά αποτελεί μία γενική συζήτηση για τα προηγούμενα κεφάλαια καθώς και τα ευρήματα κάθε μελέτης. Τέλος αναφέρονται μελλοντικές έρευνες για την τροποποίηση των συσκευών καθώς και για τη βελτίωση των μεθόδων συμφωνίας και ομοιότητας.

Το κεφάλαιο οκτώ συμπεριλαμβάνει τα συμπεράσματα αλλά και τη συνεισφορά και καινοτομία της διατριβής.

Λέξεις-Κλειδιά: ανοιχτός κώδικας, ανοιχτό υλικό, συμφωνία, ομοιότητα, παρακολούθηση μετεωρολογικών παραμέτρων, εργαστήρια ανοιχτού κώδικα και υλικού, καταγραφείς ανοιχτού κώδικα, μύγα της μεσογείου, έξυπνη γεωργία, περιεκτικότητα νερού υποστρωμάτων περλίτη, pH.

Abstract

The present PhD thesis aims to contribute to methodological approaches regarding the assessment of agreement and similarity between open-source devices and their corresponding commercials using statistical methods and experimental design. Open-source devices can be found in a vast number of fields like medicine, education, agriculture, sports, biology, research, and others. This thesis focuses on the development of a methodology applied in four open-source devices designed and developed to assess experiments in agricultural applications and research.

Chapter 1 provides a brief introduction on the factors that contributed to the emergence of open-source devices. Moreover, open software and hardware are defined, and their applications in precision agriculture, smart farming and research are summarized. The open-source philosophy and the free distribution of knowledge for any individual interested in replicating, modifying, or creating such devices while sharing all the materials needed for its design is emphasized. Examples of publications and advantages of open-source devices are presented while acknowledging the need to develop a framework to assess the functionality and the reliability of their outputs. Specifically, the meanings of method comparison studies, agreement and similarity are defined and their contribution in assessing if two devices can be used interchangeably.

Chapter 2 outlines the methodology used to assess the agreement and similarity between open-source and their corresponding industrial devices. Specifically, it provides an extended discussion for the usage of various models for fitting data produced by the devices under investigation and the evaluation of indices and graphs that are interpreted easily even by non-statisticians. Moreover, in case of concluding nonacceptable agreement between the device's measurements, sources of disagreement can be identified, and even corrective actions can be applied to improve the agreement. Advantages and disadvantages of the approaches for agreement and similarity assessment are presented. This chapter also discusses the analysis followed according to the nature of the research and proper formulation of the experimental design along with key steps for the development of valid method comparison studies. Moreover, misleading, or incomplete methods to evaluate agreement and similarity between two devices are presented.

Last, a summary of the methodology and the open-source hardware and software used to design and develop the devices is presented.

In Chapter 3 an open-source wireless weather station that measures air temperature and humidity was designed and developed. First, its functionality was evaluated by monitoring its readings via Wi-Fi for seven days. Then, a method comparison study was designed between the device and a corresponding commercial weather station using a Thygro sensor from Symmetron. The experimental design and three methodologies that use mixed-effects models were presented along with their provision to assess the two device's agreement. Thirty open-source devices were constructed and three commercial weather stations were used for the experiment. Three clusters were created with a ratio of ten open-source devices for one industrial. Each subject produced a trajectory of measurements over time. The experiment was terminated after seven days of continuous measurements. Since, the measurements per subject were more than 220 for some days (loss of data due to inappropriate connectivity) they were summarized in eight times zones according to the National Observatory of Athens. Two of the methodologies were applied successfully and an index was produced. The agreement is considered low, and a possible source of disagreement is the precision of the two instruments. Further research needs to be performed to include all the data occasions per day and not just the summary. Moreover, covariates such as air humidity or location of the cluster must be included in the study. Due to its modular nature, other sensors can be added and alternative wireless communication such as LoRa can be established in the open-source

weather station. The advantages of using open-source weather stations are their low cost for production and modularity.

In Chapter 4 a wireless LoRa open-source logger was designed and developed to aid aquaponics experiments with cultivations in perlite slabs. Specifically, the open-source logger is compatible with RS-485 soil/ soilless water content sensors. Moreover, the sensor used in the experiment can also record temperature, electrical conductivity, and pH of the slab. After testing the device and assessing its proper functionality, a method comparison study was designed to assess the agreement between the open-source device and a commercial, Handheld Meter (HH150) and SM150 Soil Moisture Sensor. The sample size was 100 perlite slabs. A measurement-error model was used to estimate the bias of the open-source device compared to the best unbiased linear predictor of the water content's true value using the measurements taken from the commercial device. The results of the analysis reported significant differential and proportional bias but acceptable agreement for specific intervals of the measurement range, depending on the application. The advantages of using open-source devices compared to commercials is the low cost and the increased configurability of the device's functions. Since perlite's water content is not homogeneous along the slab, information on the water content for more than one point per slab is essential as inputs for precise water content estimation using specific algorithms.

In Chapter 5 an open-source pH logger was designed and developed to assist experiments regarding pomology topics. Specifically, the device was calibrated manually using pH buffers for values of 4.01 and 7.01. Then it was tested by measuring pH from juice of citrus fruits. A waterproof temperature sensor was added to the device for temperature compensation when measuring pH. A method comparison study between the open-source device and a Hanna HI9024 Waterproof pHMeter was designed to assess their agreement. Using mixed-effects models, indices and useful graphs were produced using different methodologies. Their advantages and disadvantages for interpreting agreement were discussed. The indices reported mediocre agreement and the similarity analysis reported a fixed bias of -0.2215 pH units as source of disagreement. After using recalibration methods, the agreement improved and was considered excellent. The device is portable and can be used for in situ measurements. Moreover, the cost was substantially lower (more than 4 times lower) compared to Hanna HI9024. Future research can assess the measurements validity of the open-source device on different soil types and soilless substrates. Last, more sensors can be added to the device and more information can be retrieved from samples simultaneously while measuring for pH.

In Chapter 6 an inexpensive custom-made device was designed and developed to aid experiments regarding the Mediterranean fruit fly (medfly), *Ceratitis capitata*. The open-source device records core, surface and external temperature of a fruit and was tested in laboratory conditions and further in field experiments of interest. For the field experiments, the device was placed on specific locations on different types of trees. Its shape and size make the device ideal to be placed in any setting the researcher wishes to investigate. The device is battery-operated and can last from 7 to 30 days depending on the logging interval. A method comparison study was designed to compare the temperature values captured by the core and the surface sensor in different conditions. Since the goal of this research was to quantify the bias between the temperature of the two locations, the differential and proportional bias was estimated using measurement-error models. The results reported a differential bias of 0.331°C and a proportional bias of 0.982 . Moreover, the precision comparison of the two sensors suggests that the core sensor maintains each temperature for more time compared to the surface sensor. Last, wireless technology such as LoRa can be added to the device with no hardware alteration. Solar panels can make the device energy-efficient and the use of a precise RTC to improve the sleep modes allows the device to function for months without replacing the battery due to its low-power consumption functionality.

Chapter 7 concludes all the previous chapters along with all the findings and future work.

Chapter 8 summarizes the contribution of this thesis.

Keywords: open-source hardware, open-source software, agreement, similarity, meteorological parameter monitoring, open-source laboratories, smart agriculture, open-source logger, fruit fly hosts, perlite substrate water content, pH.

Table of Contents

Table of Tables.....	i
Table of Figures	v
Table of Abbreviations.....	xiv
1. Introduction.....	1
1.1. Open-source devices in Precision Agriculture.....	3
1.2. Open-source devices in research.....	5
1.3. Evaluation of open-source devices and procedures using statistical methods.....	6
1.4. Goals of the thesis.....	9
2. Materials and Methods.....	11
2.1. Agreement and Similarity Techniques	11
2.1.1. Simple Linear Model.....	11
2.1.2. Models for fitting the method comparison study data.	13
2.1.3. Some useful definitions	22
2.1.4. Graphical methods for assessing Agreement and Similarity.....	25
2.1.5. Indices for assessing Agreement and Similarity	39
2.1.6. Assessing Similarity.....	50
2.1.7. Assessing Repeatability	51
2.1.8. R code to assess Agreement and Similarity.....	52
2.1.9. Examples using Agreement and Similarity techniques.	53
2.2. Experimental Design and Statistical Analysis	54
2.2.1. Inappropriate methods to assess Agreement.	56
2.3. Designing and Using the Hardware.....	59
2.3.1. Microcontrollers.....	60
2.3.2. Development Boards	62
2.3.3. Arduino IDE	64
2.3.4. Electrical Connections and Communication Protocols	64
2.3.5. Wireless Communication.....	70
2.3.6. Designing the Hardware	77
2.3.7. Loggers for Environmental Monitoring and Sensor Working Principles.....	77
2.3.8. Internet of Things platforms.	79
3. A method comparison study between open-source and industrial weather stations.	83
3.1. Goal of the Project	83

3.2.	Software.....	83
3.3.	Cost of the device.....	85
3.4.	Industrial Weather Station	85
3.5.	Hardware	85
3.5.1.	Real Time Clock	86
3.5.2.	Temperature and Humidity Sensor	86
3.6.	Experimental Design	89
3.7.	Method Comparison Study Techniques	91
3.8.	Results	92
3.8.1.	Exploratory Analysis	92
3.8.2.	Agreement Analysis	98
3.9.	Discussion	100
4.	A method comparison study between an open-source and an industrial system for measuring water content from perlite slabs.....	101
4.1.	Goal of the study.....	101
4.2.	Design of the Device.....	102
4.2.1.	Hardware	103
4.2.2.	Software	107
4.2.3.	Cost of the Device.....	109
4.3.	The reference device	110
4.4.	Experimental Design	110
4.5.	Agreement and Similarity Techniques	113
4.6.	Results	113
4.6.1.	Exploratory data analysis.....	113
4.6.2.	Agreement Evaluation	117
4.7.	Discussion	122
5.	A method comparison study between open-source and industrial pH meter logger and sensors.	123
5.1.	Goal of the Study.....	123
5.2.	Design of the Device.....	123
5.2.1.	Hardware	123
5.2.2.	Software	126
5.2.3.	Calibration Method	127
5.2.4.	Cost of the Device.....	128

5.3.	The reference device	129
5.4.	Experimental Design	129
5.5.	Method Comparison Study Techniques	130
5.6.	Results	131
5.6.1.	Data Exploration	131
5.6.2.	Choosing the proper model	136
5.6.3.	Similarity Assessment	138
5.6.4.	Evaluation of Repeatability	139
5.6.5.	Agreement Assessment	140
5.6.6.	Recalibrating the open-source device	147
5.7.	Discussion	148
6.	FruiTTemp: Design, Implementation and Analysis for an Open-Source Temperature Logger Applied to Fruit Fly Host Experimentation	150
6.1.	Goal of the Study	150
6.2.	Design of the device	151
6.2.1.	Hardware	152
6.2.2.	Software	154
6.2.3.	Calibration Method	157
6.2.4.	Power Consumption and Power Saving Techniques	158
6.2.5.	Cost of the Device	158
6.2.6.	Bias Removing Techniques, Accuracy and Measurement Range	159
6.3.	Design of the Method Comparison Study: Measuring the Agreement Between the Core and the Surface Thermistor.	159
6.4.	Agreement Techniques	160
6.5.	Design of the Field Experiment	161
6.6.	Results	162
6.7.	Improving Fruitemp – Future Work	165
6.8.	Discussion	167
7.	General Discussion	169
8.	Conclusions	172
	References	173
	Appendix A	196
	Appendix B	205
	Appendix C	207

Appendix D	211
D1. Business Models	211
D2. Licenses, copyrights, patents and trademarks.....	211
Appendix E: Extended Greek Abstract.....	213

Table of Tables

Table 2.1. RF Power-Consumption specifications for 802.11 (WiFi). The power consumption measurements are taken with a 3.3 V supply at 25°C of ambient temperature at the RF port. All transmitters' measurements are based on a 50% duty cycle. (source: ESPRESSIF, 2023)	72
Table 2.2. Spreading factors, Bandwidth, Data Rate and Bit Rate combinations for LoRa's transmission in Europe.	73
Table 2.3. Main current consumption details on LoRa/LoRaWan common transceivers from different manufacturers. (source: Ibanez et al., 2017).....	74
Table 3.1. Libraries used in the code and their functionality.	83
Table 3.2. Detailed cost of every component for the construction of the open-source weather station. 86	
Table 3.3. Technical specifications for the Adafruit SHTC3 temperature and humidity sensor.	88
Table 3.4. Common occasions for both devices recorded per day.....	90
Table 3.5. The experimental design. The experimental unit is defined as the sensor and logger system. The data are considered longitudinal. The design is balanced.	90
Table 3.6. Summary of time zones used to report daily temperature from the National Observatory of Athens.	91
Table 3.7. Model formulation for the mixed-effects model per methodology for CCC evaluation.....	92
Table 3.8. Time-zones per day. Days 1,2,3,6,7 include measurements for all the time-zones. However, day 4 and 5 include 5 and 6 time-zones respectively.	95
Table 3.9. CCC estimate and its corresponding confidence intervals. The agreement is considered low. 98	
Table 4.1. Technical Specifications for the Industrial Soil Moisture, Temperature and Electrical Conductivity sensor.	104
Table 4.2. Summary of the cost for the construction of the open-source device.....	109

Table 4.3. Summary of the cost for the construction of the LoRa gateway.	109
Table 4.4. The experimental design details summarized. The experimental unit is the perlite slab. There were 3 repeated measurements for the open-source device and 10 for SM150 per slab. The data are considered unlinked since the repeated measurements are not paired. The sample size is 100 and the data are unbalanced.	111
Table 4.5. A summary for the water content (%) measurements. The table presents the minimum, 1 st , 2 nd and 3 rd quartile, the mean and the maximum value of the measurements.	114
Table 4.6. The differential and proportional bias of the open-source device compared to SM1500 (reference).	120
Table 5.1. Technical Specifications of pH sensor SEN0169.	125
Table 5.2. Waterproof DS18B20 digital temperature sensor specifications.	125
Table 5.3. ADS1115 16bit ADC technical specifications.	126
Table 5.4. Temperature of the pH buffer during the calibration procedure. Eight measurements were taken. Four for 4.01 pH buffer and four for 7.01 pH buffer.	127
Table 5.5. Device cost in 2023 local vendor prices and international vendor Mouser.	129
Table 5.6. Summary of the experimental design information. The sample size is 30, since each fruit is considered one experimental unit. There were ten repeated measurements, and the design was balanced. The data are considered linked.	130
Table 5.7. The minimum, 1st quartile, median, mean 3rd quartile, and maximum pH values per device. The measurement range lies between 2.78 and 3.7.	136
Table 5.8. AIC, BIC, Log-Likelihood and degrees of freedom for the homoscedastic and the heteroscedastic model. The heteroscedastic model is slightly better than the homoscedastic model. Either can be chosen. In this case the heteroscedastic is chosen for illustrative purposes. The analysis for the homoscedastic model can be found in Appendix C.	136

Table 5.9. AIC, BIC, log-likelihood and degrees of freedom for model selection. The model which does not include covariates is slightly better than the one which includes the covariates. Thus, the model with no covariates is selected for further analysis and indices calculation.....	136
Table 5.10. Model counterparts. The fixed bias is – 0.13. Recalibration actions can be taken to improve the agreement by adding the constant. The standard errors seem reasonable for the model’s validity.	138
Table 5.11. Precision ratio estimates as a function of magnitude of measurement. The open-source device is twice to three times more precise than Hanna. The fixed bias is -0.22 units for the open-source device compared to Hanna. The open-source device underestimates the pH measurement 0.22 units since the entire interval is below zero.....	138
Table 5.12. Summary for CCC and TDI(0.9) for the intra- method agreement.	140
Table 5.13. 95% limits of agreement for inter-method Agreement for pH data as a function of magnitude of measurement. The inter-method limits, based on the distribution of D, are centered at -0.22. For lower pH values the LOA are narrower comparing to the higher pH values and range from [-0.3464, -0.3237] for lower LOA and –0.1193, –0.966 for upper LOA.	141
Table 5.14. 95% limits of agreement for intra-method Agreement for pH data as a function of magnitude of measurement. The open-source device has narrower LOA compared to Hanna pH meter. Thus, the open-source device has better repeatability compared to Hanna. Based on the distributions of D_1 and D_2 , the intra-method limits are centered at zero.	141
Table 5.15. CCC and TDI estimates with their corresponding lower and upper confidence bounds throughout the pH measurement range.	144
Table 5.16. Summary of TDI (0.9) and CCC with their corresponding bounds using Escaramis and Carrasco’s methodology respectively.....	147
Table 5.17. CCC and TDI estimates after recalibration. Both indices improve significantly and report excellent agreement between the devices.....	147
Table 5.18. TDI (0.9) from Escaramis and CCC from Carrasco after recalibration. The agreement improved significantly and is now considered excellent.....	148
Table 6.1. FruiTemp individual parts’ cost.....	159

Table 6.2. The means and standard deviations of the reference temperatures measured by the external temperature sensor per fixed condition.	160
Table 6.3. Summary of the experimental design for the laboratory experiment.	162
Table 6.4. Differential and Proportional bias for the measurement at the surface relative to the measurement at the core of the fruit and associated 95% confidence intervals.	166
Table A1. ATmega32u4 pin mapping (source: Arduino (2013))	196
Table A2. ESP 32 variations characteristics.	198
Table A3. Raspberry Pi 3, model B Specifications	199
Table A4. Specifications of the thermistors.	199
Table A5. <i>The coefficients of the Hoge-2 equation based on the values of the fraction R_t/R_{25}, which is the value of the resistance measured using Ohm's Law over the value of the reference resistance (10 KOhm in our case).</i>	201
Table A6. <i>Formulas used to calculate the resistance of the thermistor (R_t), the temperature reading (Hoge-2) and the ADC value.</i>	201
Table A7. Adafruit feather M0 RFM95 Lora Radio technical specifications.....	201
Table A8. SM150 soil moisture sensor and HH150 meter technical specifications.	202
Table A9. Technical specifications for Hanna Instruments HI9024 Waterproof pH meter.	203
Table A10. Summary of the LCC, LPC and LA for the assessment of agreement between the industrial and open-source device.....	203
Table B1. Chemical composition of the substrate during the Chapter 4 experiment.	206
Table B2. Physical properties of the perlite substrate during the Chapter 4 experiment.....	206

Table of Figures

Figure 2.1. Representation of Precision and Accuracy through the darts and target analogy. (a). Accurate and precise measurement. (b). Inaccurate but precise measurement. (c). Accurate and imprecise measurement. (d). Inaccurate and imprecise measurement.	23
Figure 2.2 Bland-Altman Plot. The limits of agreement are [-0.363, 13.605]. Thus, 95 % of the differences are expected to lie between -0.363 and 13.605.	26
Figure 2.3. Probability Agreement Plot for comparing two methods. (Source: Stevens et al., 2017)	32
Figure 2.4. Bias plot. The differential bias is 11.581 while the proportional bias is 0.463 (data from Chapter 4).	34
Figure 2.5. Precision plot. The reference method is more precise compared to the new method (data from Chapter 4).....	36
Figure 2.6. Compare plot. The methods agreement improves after recalibration (data from Chapter 4).	37
Figure 2.7. Deming Regression for unreplicated data using “mcr” R package (data from Chapter 4).....	38
Figure 2.8. Passing-Bablok for unreplicated using "mcr" R package (data from Chapter 4).	39
Figure 2.9. The scatterplot of two sets of measurements with a fixed bias of 2 and a proportional bias of 2.5. The dashed line shows the set of measurements, and the solid line shows the line of equity.	58
Figure 2.10. Lineplots. The x-axis presents the time, while on the y-axis a percentage of cloudiness index is presented. To differentiate between the two devices the authors use green color for the reference device and blue for their prototype. The disagreement intervals are not obvious neither the sources of bias (fixed or proportional). (source: Laerhoven et al. (2022))	59
Figure 2.11. (a) Atmega 32u4 on the left (source: Microchip Technology Inc). (b) Pinout ATmega 32u4 on the right (source: Arduino documentation).	61
Figure 2.12. The 10-bit resolution of the Analog to Digital converter. The 5 Volts are divided into 1023 increments of 4.8 milli Volts.	61

Figure 2.13. (a) ESP-32 WROOM-32E, with Internal antenna, on the left (source: espressif, esp32 documentation) . (b) ESP32-WROOM-32UE, with surface-mount uFL connector for external antennas, on the right (source: espressif, esp32 documentation)	62
Figure 2.14. (a) Arduino Uno development board on the left (source: Arduino official web page) . (b) Adafruit feather 32u4 proto stacked with an adalogger on the right (source: Adafruit official web page).	63
Figure 2.15. Raspberry pi 3, model B stacked with a LoraWan Gateway with RAK2245 Pi Hat. This is the gateway where all the LoRa nodes send their data. The Raspberry Pi connects via WiFi or 4G (dongle) to upload the data to an IOT server in case no internet connection is available.	63
Figure 2.16. Arduino IDE 2. The interface where users can write, compile and upload sketches to the Arduino and Arduino compatible development boards.	65
Figure 2.17. A simplified analog as an example of communication protocols. Person A wants to initiate conversation with person B. The first action is to say “hi”. If person B replies, then person A will start the conversation. If person B is not replying, then person A will either repeat his first attempt by saying hi or stop the action.....	65
Figure 2.18. UART communication protocol. The two signals are the Transmitter (Tx) and the Receiver (Rx). The devices share a common ground.....	67
Figure 2.19. The I2C communication protocol with one master and two slaves. A 4.7 K Ω resistor is placed between VDD and the two wires (SCL, SDA).	68
Figure 2.20. SPI communication with one master and many slaves.	69
Figure 2.21. Pin diagram for SPI communication.....	70
Figure 2.22. (a) Adafruit HUZZAH32 – ESP32 feather board (source: Adafruit official web site). (b) Adafruit WINC1500 Wi-Fi shield for Arduino Uno (source: Arduino official web site).	71
Figure 2.23. (a) OEM LoRa development board embedded with an OLED screen (source: Wirelesslan.gr) (b) Adafruit LoRa radio FeatherWing. A shield for the Adafruit feather development board series (source: Adafruit official web site).....	75

Figure 2.24. Benefits of using LoRaWAN in agriculture. Up to 1000 end-nodes can establish connection to the gateway. (source: LoRa alliance web site).....	75
Figure 2.25. LoRaWAN infrastructure. End-node device route to the application server (source: LoRa alliance web site)	76
Figure 2.26. (a) Fritzing interface. A tool for designing electronics and producing documentation, on the left (source: Fritzing official web site). (b) Kicad user interface, on the right. Kicad is an open-source Electronics Design Automation Suite (source: Kicad official website).....	78
Figure 2.27. (a) A voltage divider with input voltage 5 Volts. The output is connected to an analog input of the Arduino Uno development board. (source: created using Fritzing) (b). SEN0262 Current-to-voltage module. The module converts current that ranges from 4 to 20 mA to voltage (source: DFrobot official web site).	79
Figure 2.28. ThingSpeak menu. Creating a channel.....	81
Figure 2.29. ThingSpeak data visualization from Field 5 and 6. The x-axis is the timeline while the y-axis displays the value of the characteristic the sensor measures.....	81
Figure 2.30. The Things Network settings for the LoRa gateway used in the present research. The frequency plan is chosen as EU_863_870_TTN for Europe.....	82
Figure 3.1. Flowchart that summarizes the sequence of the weather station's functions.	84
Figure 3.2. Temperature and humidity sensor enclosed in a radiation shield of Thygro SDI-12.....	85
Figure 3.3. (a) Adafruit Sensirion SHTC3 temperature/humidity sensor. (b) ESP-WROOM-32 module, for functionalities coordination and wireless connection. (c) Real Time Clock SD3231 to acquire the timestamp. (d) SD card module to store the data	87
Figure 3.4. Finalized PCB, which does not include the sensor. All the parts were soldered, and the device was tested for 1 week in the lab to make sure everything will work as planned during the field experiment	88
Figure 3.5. The final product for a single system. The box that houses the system is an IP66 plastic box and the body of the station is a camera tripod embedded with aluminum corners that contain mounting holes, which are useful for placing the box in the desirable height.	89

Figure 3.6. The topology of the experiment. Every industrial weather station creates a cluster of ten open-source weather stations.	89
Figure 3.7. Trajectories of the differences per subject. Occasion 11 represents the first day and first time zone. Occasion 24 represents the fourth time zone of the second day. There are 51 occasions in total. A trend throughout the days is obvious. The differences start with around -3 units, increase, and reach a peak of approximately 4 units. Thus, the open-source device underestimates then overestimates the temperature.	93
Figure 3.8. Scatterplot of Industrial vs open-source by time-zone. The open-source device underestimates the temperature for time-zones 1, 2, 8 and overestimates for time-zones 5 and 6. For time-zones 3,4 and 7 both effects are noticed.....	94
Figure 3.9. Bland-Altman plot without the limits of agreement per time-zone. A trend is visible for time-zones 3,7 and 8.....	95
Figure 3.10. Boxplot of the temperature differences per time zone.....	96
Figure 3.11. Boxplot of difference by occasion. A cubic polynomial provides a reasonable fit.	97
Figure 3.12. (a) Estimates and 95% confidence intervals for the longitudinal concordance correlation. LCC starts from 0.2236 for occasion 11 and reaches 0.1146. As the occasions increase LCC reduces. (b) Estimates and 95% confidence intervals for the longitudinal Pearson correlation. (c) Estimates and 95% confidence intervals for the longitudinal accuracy. The points on the plots represent the concordance correlation coefficient considering independent measurements taken over time.	99
Figure 4.1. The open-source logger with the water content, temperature, electrical conductivity and pH sensor.....	102
Figure 4.2. Soil moisture, temperature, electrical conductivity and pH Sensor by ComWinTop. (source: ComWinTop official webpage).	103
Figure 4.3. Sensor's interface. The program is available by ComWinTop company (source: ComWinTop official webpage).	104
Figure 4.4. The development board, Adafruit feather M0 RFM95 LoRa Radio (900 MHz). (source: Adafruit official website)	105

Figure 4.5. Uart serial port to RS485 Converter module and its connection diagram with the MCU (source: fruugo (2023)).	106
Figure 4.6. Raspberry pi 3 B+ combined with a RAK2245 Pi Hat for LoRa gateway deployment. To connect to the server a 4G/LTE dongle is added with a SIM card.	108
Figure 4.7. The example of a response frame provided by the manufacturer of the sensor. The fourth and fifth byte corresponds to the water content. The sixth and seventh to the temperature, the eighth and ninth to the electrical conductivity and the tenth and eleventh corresponds to pH.	108
Figure 4.8. HH150 Handheld Meter and SM150 Soil Moisture Sensor used as the industrial device for the method comparison study.	110
Figure 4.9. A circle of 5 cm diameter was drawn on every slab. Each circle was 10 cm away from the planting area. All the measurements were taken inside the circle to minimize the water content variability.	112
Figure 4.10. HH150 handheld meter and SM150 soil water content sensor. All the measurements are taken with the sensor inserted inside the circle.	112
Figure 4.11. Water content sensor connected to the open-source device. The measurement is taken inside the circle, next to the one taken from SM150. Its measurement marks are obvious on the left side of the circle.	113
Figure 4.12. Boxplot for the water content (%) measurements per device. SM150's measurements are significantly higher than the open-source's	114
Figure 4.13. Trellis plot for the soil water content measurements. The vertical axis is divided into rows and each row displays all the repeated measurements for both devices using method-specific colors. Pink color represents the measurements for the open-source device while blue represents the measurements for the reference device. The open-source device underestimates the soil water content for most of the subjects by a non-constant quantity. As the measurement range increases the difference decreases revealing dependance of the difference on the magnitude of the measurements.	115
Figure 4.14. Scatterplot for the water content (%) measurements of the reference device versus the open-source. Each subject is represented by a unique id number and the repeated values share the same id subject number symbol. The open-source device underestimates the water content for all subjects apart from subject 6, 28 and 47. The scatterplot reveals low correlation between the methods.	116

Figure 4.15. Extended 95% Bland-Altman limits of agreement. The differences depend on the magnitude of the water content since they increase when it increases. Moreover, the variability is not constant across the measurement range since it increases when the measurement range increases. The points in the plot are not centered at zero, suggesting a difference in the fixed bias. 118

Figure 4.16. Compare plot. This plot allows the visualization of the bias-corrected values of the open-source device. The blue solid line presents the open-source device regression line before recalibration and its corresponding measurements are presented with a blue solid circle. The red dashed line presents the open-source device regression line after recalibration and its corresponding measurements are presented by red solid rhombus symbols. The black solid line presents the SM150 regression line, and the corresponding measurements are presented with a circle..... 120

Figure 4.17. Precision plot. This plot allows the visual comparison of the precision of the open-source device with that of SM150. A scatter plot is created for the estimated standard deviations against the BLUP of the true unobserved value. The standard deviations are estimated after the open-source device is recalibrated. The open-source device's precision decreases as the measurement magnitude increases. The SM150 has higher precision compared to the open-source device. Specifically, the SM150 device is almost 6 times more precise than the open-source. 121

Figure 5.1. DFRobot PH meter (SEN0169) (source: DFRobot official website). 124

Figure 5.2. Waterproof DS18B20 digital temperature sensor for Arduino and Arduino compatible microcontrollers. (source: DFRobot official website). 124

Figure 5.3. Callibration curves for 7.5°C, 13.5°C, 17.5°C and 22.4°C..... 128

Figure 5.4. (a) The hand squeezing procedure. (b) Hanna pH meter. (c) DFRobot pH meter probe. 130

Figure 5.5. (a) Scatterplot for the pH measurements of Hanna pH meter versus the open-source. Each subject is represented by a unique id number and the repeated values share the same id subject number symbol. A systematic underestimate of the open-source device for pH measurements is clear in the graph. There are two clusters formed in the data. The lower left corresponds to the grapefruit pH while the upper right corresponds to the juice oranges. (b) Bland-Altman plot (averages vs differences) without the limits of agreement. For higher values of pH the differences seem to have slightly higher spread compared to lower values of pH. This is a sign of possible heteroscedastic errors. There is no obvious trend suggesting a common scale for the assays, verifying the common scale assumption for the mixed effects model. 132

Figure 5.6. Trellis plot for the pH measurements. The vertical axis is divided into rows and each row displays all the repeated measurements for one subject using method-specific colors. Blue color represents the measurements for the Hanna pH meter while pink represents the measurements for the

open-source device. The open-source device underestimates pH measurements for most of the subjects by approximately 0.22 units. Moreover, the within subject-variation for the open-source device is less than Hanna's. The between-subject variability ranges between 2.78 and 3.7..... 133

Figure 5.7. (a) Interaction plot between method and subjects. There is evidence of subject x method interaction since the lines intersect. (b) Interaction plot between subjects and time. A few of the lines intersect, providing evidence of possible, but not strong interaction between subjects and time. 134

Figure 5.8. Extended Bland-Altman plot and LOA. There is slight evidence of heteroskedastic errors. No trend is apparent; thus, a common scale is assumed for the assays..... 135

Figure 5.9. Precision ratio with the corresponding confidence intervals. The open-source device is more precise than Hanna, especially while the magnitude of the measurements increases. The open-source device is twice to three times more precise than Hanna..... 139

Figure 5.10. 95% Limits of inter- and intra- method agreement. The intervals reveal a systematic underestimation of the pH measurements from the open-source device. For the 95% limits of the intra-method agreement the open-source device has narrower LOA compared to Hanna pH meter; thus, the open-source device has better repeatability compared to Hanna. For lower pH values the LOA are narrower compared to the higher pH values. 142

Figure 5.11. Bland - Altman plot using "blandr" package. Apart from the limits of agreement and the mean difference, their corresponding confidence intervals are plotted. 143

Figure 5.12. (a) One-sided 95% pointwise confidence bands for CCC as a function of the magnitude of the measurements. The solid line represents CCC lower confidence bound for the inter-method agreement and ranges between 0.4776 and 0.4839. The agreement is considered insufficient. The dashed and dotted line represents the intra-method one-sided 95% pointwise confidence band for Hanna and open-source respectively. For Hanna the band ranges between 0.9979 and 0.9987 and for the open-source between 0.9986 and 0.9997. The intra-method agreement for both devices is considered excellent. (b) One-sided 95% pointwise upper confidence bands for intra- method versions of TDI (0.9) and their reflections over the horizontal line at zero. The open-source device has higher intra- method agreement compared to Hanna. 145

Figure 5.13. Bias-plot. The proportional bias is bias is 0.965 (95% CI [0.9352, 0.9938]) and the fixed bias - 0.1052 (95% CI [-0.2013, -0.091]). The fixed biases estimate is different compared to Choudhary's and Nagaraja's method. However, the red solid line which corresponds to the total bias confirms the findings in Table 5.11. 146

Figure 6.1. Initially the device was designed according to specifications based on the nature of the experiment. Then, the laboratory experiment was designed, and data collection followed. Data analysis involved statistical methods to assess the Agreement. The experiment's findings confirmed the initial hypothesis thus designing the field experiment.	151
Figure 6.2. The device consists of two precise medical sensors inserted in the fruit and an environmental temperature probe. There is also an On/Off waterproof switch. The box is waterproof (IP 54) and flanged.	151
Figure 6.3. (a) The data logger interior consists of an Adafruit 32u4 proto, an Adafruit FeatherWing logger, an ADS1115 ADC with a gain amplifier and an LM4040 Voltage reference breakout. (b) The case of the system is an IP 54 flanged box. (c) The MC65 precise medical sensors. (d) The JS8746A environment temperature probe with its enclosure.	153
Figure 6.4. The calibration procedure. First, the ADC value is measured. Second, R_t is retrieved using Ohm's Law and a voltage divider. Last, proper coefficients from the manufacturer are chosen as an input to Hoge-2 equation.	157
Figure 6.5. (a) The proposed device set during the field experiment. (b) The proposed device set in the Elvem CLP 600 chamber during the laboratory experiment.	162
Figure 6.6. Scatterplot of measurement pairs including the identity line.	163
Figure 6.7. Bland–Altman LOA plot. The blue dashed lines are the limits of agreement, and the solid blue line is the regression line.	163
Figure 6.8. Precision plot showing the precision of each measurement method. The blue dots correspond to the precision of the surface readings along the temperature and the black dots to the precision of the core.	164
Figure 6.9. The Bias Plot. The bias is displayed as the red line and reveals the bias throughout the range of measurements. For the value of 30°C the bias is approximately -0.2°C . The black horizontal line is the zero reference for bias, the dashed blue and green line is regression of the Surface and Core temperature readings, respectively. The solid black line is the identity line.	164
Figure 6.10. LoRa gateway consists of an Adafruit feather M0 with RFM95 LoRa module connected via I2C with an Adafruit Feather Fona which sends the data to ThingSpeak via GPRS.	166

Figure 6.11. ThingSpeak interface. The core, surface and air temperature are visualized using linegraphs. Red dots represent the measurements.....	167
Figure A1. The circuit sketch describes the connections among the electronic parts. The sketch was created using Fritzing.....	200
Figure C1. Diagnostics for the model. There is evidence of significant deviation from normality.	207
Figure C2. (a) Standardized residuals versus fitted values of the homoscedastic model for Hanna pH meter and the open-source device. (b) Absolute values of standardized residuals versus fitted values of the homoscedastic model for Hanna pH meter and the open-source device. There is no obvious sign of a fan shape thus a formal criterion (AIC, BIC and log-likelihood) will be used to compare the homoscedastic and the heteroscedastic model.	208
Figure C3. Standardized residuals on the horizontal axis vs the quantiles of the standard normal distribution. The plot reveals a slight deviation from the normal distribution.....	210

Table of Abbreviations

ADC – Analogue to Digital Converter

CAD – Clinically Accepted Difference

CC – Creative Commons

CCC – Concordance Correlation Coefficient

CSS – Chirp Spread Spectrum

DAC – Digital to Analogue Converter

DC – Direct Current

DevEUI – Device Extended Unique Identifier

DIY – Do It Yourself

EMI – Electro Magnetic Interference

FDA – Food and Drug Administration

FSK – Frequency Shift Keying

GPRS – General Packet Radio Service

GPS – Geographical Positioning System

GSM – Global System for Mobile Telecommunications

HQ – High Quality

I2C – Inter-Integrated Circuit

IDE – Integrated Development Environment

ISO – International Organization of Standardization

LOA – Limits of Agreement

LoRa – Long Range

LoRaWAN – Long Range Wide Area Networking

LTE – Long-Term Evolution

MCU – Microcontroller

NDVI – Normalized Difference Vegetative Index

OS – Open Source

PA – Precision Agriculture

PCB – Printed Circuit Board

PF – Precision Farming

RF – Radio Frequency

RTC – Real Time Clock

SA – Smart Agriculture

SPI – Serial Peripheral Interface

TDI – Total Deviation Index

TSS – Total Soluble Solids

TTN – The Things Network

UART – Universal Asynchronous Receiver/Transmitter

WSN – Wireless Sensor Network

1. Introduction

Open-source devices are becoming very popular and, in some cases, essential in education (according to the systematic mapping study approach) (Heradio et al., 2018), agriculture (Coleman et al., 2022; Mesas-Carrascosa et al., 2015), medicine (Niezen et al., 2016), biology (Jolles, 2021) and an increased number of fields. There is an extensive list of factors that incite people to take an active part in prototyping and innovation. The Maker Movement (Papavlasopoulou et al., 2017) that unfolded after the resurgence of the participatory Web 2.0 (Atkinson et al., 2012), the interfusion of Open Source, the decreased cost of electronic parts and other social influences are a few of the examples that contributed to the flourishing of prototype development. Furthermore, this phenomenon was boosted after the launch of development boards like Arduino (Louis, 2018; Arduino, 2015) and Raspberry-pi (Jolles, 2021) that tend to simplify intricate electronic assemblies by using software and programming. These companies provide straightforward tutorials, accessible to everyone. The tutorials simplify the technical parts, provide visual aids for the wiring, issue the code in each case, and allow users with basic or no experience in electronics and coding, to replicate or customize projects according to their needs. Major contributors to this trend are also the OSHWA (2023) which was established around 1997 and the Open-Source Initiative (2023) which was founded in February 1998.

As defined by the official open-source forum (Open-Source Community, 2023; Open-Source Hardware Association, 2023) “open hardware, or open-source hardware, refers to the design specifications of a physical object which is licensed in such a way that said object can be studied, modified, created, and distributed by anyone”. Also, according to Open-Source Initiative (2023), “open hardware is a set of design principles and legal practices, not a specific type of object”. It can be any object such as an automobile, a computer, a robot, a building, or a house. For the open-source software, as defined by the Open-Source Initiative (2023), “Open-source software is software that can be freely accessed, used, changed, and shared (in modified or unmodified form) by anyone”. Open-source software is made by many people and distributed under licenses that comply with the Open-Source definition. Moreover, “the rise of the participative web or web 2.0 did not only democratize knowledge production but also the means of design and invention by industrializing the Do it Yourself (DIY spirit)” (Gershenfeld, 2005; Voigt et al., 2016). Thus, the public is encouraged to participate in the manufacturing process of prototypes and final products.

The radical shift in approach, when it comes to open source, is that all the material that will aid someone to build, modify or enhance the hardware, such as schematics, blueprints, logic designs, Computer Aided Design (CAD) or any other file, is open and available under permissive licenses. Features can be added according to the needs of the user, or bugs can be fixed in case they are encountered. In any situation, the users can tailor these resources according to their needs and even share their version. This idea contradicts the closed hardware, where the design documentation is tightly sealed and inaccessible to the public.

Open-source hardware refers to sharing and Do-it-yourself solutions. Specifically, with the use of development boards and compatible sensors. The most popular and user-friendly boards are based on Arduino hardware (Arduino, 2023a) and software (Arduino, 2023b) which was established in 2005 by Massimo Banzi and David Cuartielles in Ivrea, Italy. The Arduino development boards are the most popular since they are easy to program and build. The Arduino Integrated Development Environment (IDE) is an interface based on C, C++ and some special functions to control electronics. In 2013 Circuit Python programming language for microcontrollers makes its debut as an alternative to Arduino IDE, which was

founded by Damien George. As its name proclaims, Circuit Python is based on Python programming language. Its benefits compared to Arduino IDE are discussed in Chapter 2. More details about the development boards and programming interfaces are discussed in Chapter 2.

Open-source projects can be used in research and industry. Low acquisition cost and easy customization are two of the most important advantages of using such devices compared to industrial. Many articles that focus on open-source devices aid in research and industry have been published and widely used. Coleman et al. (2022) refers to an open-source, low-cost device for fallow weed detection, to maximize crop yield potential in moisture limited agricultural environments. The technology used in Coleman et al. (2022) is based on image analysis. The hardware used is a Raspberry pi and an HQ camera combined with other essential electronics. This device focuses on the low cost, small form factor and a large-scale deployment on crop systems. Another system developed for precision agriculture is developed by Mesas-Carrascosa et al. (2015). The authors developed a device which remotely monitors soil temperature and humidity and air temperature and humidity. The advantage of this system compared to its corresponding industrial ones is the low cost and its scalability. Aliagas et al. (2022) proposed a low-cost and DIY device for pumping monitoring in deep aquifers. The authors designed a wireless, LoRa based (LoRa Alliance, 2023a), device which applies edge computing and eliminates the need to transfer big loads of data. Instead, the preprocess and computation of the data are conducted in the sensor-nodes that acquire the data, minimizing time delay and the need for bandwidth, to inform the users about the predictions for water level evolution. This open-source wireless system based on a raspberry pi microcomputer and Arduino microcontroller, has three key advantages compared to industrial. First, it is low-cost and “do-it-yourself”, thus reachable by any researcher to implement and customize to any project. Second, the computation of an approximated model in the device for predicting water-level change due to pumping, instead of exact. Last, all the hardware and software used to build and program the device is available to everyone, paving a way of knowledge share. These key points result to the increased repeatability of experiments, since the sample size increases. Ji et al. (2019) developed a LoRa-based visual monitoring scheme for agriculture IoT. The authors managed to transmit image data via LoRa with respect to bandwidth limitations. Since images use a significant amount of bandwidth and LoRa technology is used mostly for a data transmission rate in the magnitude of Kbps, a technique was developed that splits image to grid patches and transmits only the modified area of an image based on their dissimilarity measure. Ji et al (2019) amplified the importance of customizability which is not applicable to closed hardware and software. Regarding the medical field, Niezen and Thimbleby (2016) compare 10 open-source healthcare projects in terms of how easy it is to obtain the materials needed to build the devices. The authors describe devices that relate to myoelectric prostheses, CT scanners, infusion pumps, physiological monitoring, and diabetes related projects.

This thesis is elaborated in three parts. First, the proposal of open-source solutions for designing and building agriculture and research loggers. Second, the evaluation of the open-source solutions using formal experimental designs and assessing their agreement with their corresponding industrial equivalents using formal statistical methods. Third, the development of a user-friendly R script to automate the evaluation procedure. This thesis follows the open-source spirit, thus, all the material used to build the devices and the methods to evaluate them are available for the readers and is protected under a Creative Commons license (CC).

1.1. Open-source devices in Precision Agriculture

The need to embed open-source philosophy in agriculture and specifically precision agriculture (PA) comes from the extend of small family farms to huge profitable businesses comprised of many partners.

The concept of precision farming (PF) or precision agriculture (PA) first emerged in the United States in the late 1980s. It did not take long to spread to the developed countries that invested in technology and progress in farming. Precision farming can be defined as “the application of technologies and principles to manage spatial and temporal variability associated with all aspects of agricultural production for the purpose of improving crop performance and environmental quality” (Pierce & Nowak, 1999). From sensors to robots and drones, PA became vital to farmers that started cultivating farms of only a few acres, to hundreds or thousands. When the farms were small, they were easy to control and supervise. Since their increasing in size, farms were hard to manage. Farmers then turned to technology and the promising PA.

PA offers many benefits such as the increase in yield, saving of nutrients and scarce goods, reducing labor time and use decision-support systems that can increase profitability on the farm and reduce the negative environmental impact (Pedersen & Lind, 2017). Furthermore, due to social and political trends and phenomena such as consumer demands, high security fruits and vegetables and monitoring of the impact of crops cultivation to the environment, the technology offered by PA is required more than ever. There are four elements that PA consists of. First, the geographical positioning system (GPS). Second, gathering information. Third, decision-support and last, variable-rate treatment. Smart sensors fall into the element of information gathering and smart agriculture. Smart agriculture (SA) differentiates from precision agriculture by focusing on capturing data and interpreting them using computers to make farm procedures predictable and efficient rather than capturing precise measurements. SA consists of proper farming management on water usage, fertilizers, crops, animal tracking for security etc. Remote sensing can be embedded in SA using various wireless technologies or their combination to transmit data to private or public servers. These wireless technologies can be 2.4 GHz RF, LoRaWAN, GPRS, LTE and others. More information on these wireless technologies can be found in Section 2.3.5. Remote sensing reduces labor substantially by cutting down the hours a worker needs to be present to collect data from sensors that don't use remote sensing. Furthermore, servers like thingspeak.com (MathWorks, 2022) or Adafruit IO (Adafruit, 2023) grant the user the opportunity to store and directly use interactive graphs, functions and formulas to process the data. More details about public Internet of Things servers are provided in Section 2.3.8.

Smart sensors were introduced in the late 1990s and were considered a breakthrough for crop control and supervision. To manage the farms efficiently with all the benefits mentioned above, the climate, weather, soil conditions, solar radiation, pH solution and other features need to be measured. Specifically, leaf wetness duration, soil moisture, soil temperature, ambient temperature, and relative humidity measurements are factors associated with plant diseases development. Patle et al. (2022) developed an IoT (Internet of Things) sensor system measuring these characteristics. The goal of this system was to gather data that was used to feed a Long Short-Term Memory (LSTM) network to predict powdery mildew, anthracnose, and root rot disease on the mango plant. As mentioned by the authors, the crop loss due to diseases is about 36%. Thus, the early prediction knowledge gained through the sensor-system and the predictive network grant the farmers enough time to manage the disease on time. This solution is affordable, according to the authors, and will enhance crops production and guarantee profitability and

sustainability. The long-standing experience of the farmer cannot be taken for granted since environmental conditions and other factors might change as the years pass and the scale of the farm becomes larger. Thus, instruments that can measure these characteristics with high or even mediocre precision grant farmers access to valuable information for predicting crop disease development or even yield. Li et al. (2023) developed an accurate and portable Hyperspectral Imaging device for dicot leaves. The device collects an image of a dicot leaf within 20 seconds. The data obtained allows users to utilize the NDVI (Normalized Difference Vegetative Index) spatial distribution heatmap of the entire leaf to predict the nitrogen content of a soybean plant. The device design is based on the open-source Raspberry pi 4 microprocessor and an Arduino Nano microcontroller.

Although smart sensors and automated systems thrive nowadays, there are two main reasons that prevent farmers from implementing smart sensor systems in their fields.

First, the cost to buy the smart sensor systems and to install them was and still is considered prohibitive to middle-class farmers. The present cost, according to Greek vendors, for a system that covers the full needs of a farmer and is capable for remote monitoring, ranges from around five to some hundred thousand of euros, depending on the hectare coverage. Specifically, a weather station (Wireless Vantage Pro2 Plus) that measures air temperature, air humidity, rainfall, wind speed, wind direction, solar radiation intensity, atmospheric pressure and is capable of remote sensing via Wi-Fi, providing an application for data monitoring, costs approximately 1650 euros according to Meteoshop (2023). A soil moisture station with 2 soil moisture sensors and its corresponding logger with no remote sensing capability costs around 570 euros according to Analytika (2023). To cover a vast amount of hectares of land, the cost of the sensors and their corresponding wireless loggers increases at a formidable rate and prohibits the implementation of smart agricultural solutions.

Second, the lack of interest in the farmer's profession by younger generations. On average of the 27 member states of the European Union, 55% of the private farmers is over 55 years old (Rovný, 2016). According to Dhraief et al. (2019), the authors investigated the factors affecting innovative technologies adoption by livestock holders in the arid area of Tunisia. Empirical findings show that farmer's age and experience had a significant negative effect on IT adoption. Furthermore, Huang et al. (2020) report aging and off-farm employment as significant factors on the negative impact of farmers' adoption behavior of soil and water conservation technology in the Loess Plateau in China. The negative reputation of agricultural professions, as tedious and challenging work in combination with relatively low income, tends to lead younger people to uncertainty in choosing agriculture as their main income (GmbH & Temmen, 2021). Precision agriculture offers a less laborious profession and guarantees a controlled income since the automated environment targets specific goals in production. Younger people (40 years old and lower) are more familiar with IT and almost all the population has access to the internet. The numerous sources of knowledge and their easy access entices the younger generation to see the agricultural field from a different perspective. Open-source hardware/software offers solutions to implement precision agriculture with low cost and high customizability. Two open-source systems for precision agriculture are proposed in Chapter 3 and 4.

First a DIY weather station equipped with air temperature and humidity sensors. The weather stations were designed and constructed using Arduino based development boards and programmed via Arduino IDE. The weather station sends the data via Wi-Fi to Adafruit IO. The user can assess the measurement via graphs and even download the data in multiple formats. Apart from the online storage of the

measurements, all the data including a time stamp are stored in a micro-SD card. Last, the users can perform statistical analysis or feed models directly to the server and create interfaces that suit them best. The device is modular. This means that the users can customize it easily and add more sensors according to their needs. The cost of the device is 55.6 euros (2020 market prices). The experiment was performed in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece.

Second, a system that is equipped with a probe that measures soil or soilless substrates' water content, temperature, electrical conductivity and pH. LoRa wireless technology is embedded to the logger and thingspeak.com server is used to upload the data. The user can add more probes in the same logger. The system is highly customizable. Similar to the weather station, the device was designed and constructed using Arduino based development boards and programmed via Arduino IDE. The cost of the device was around 380 euros, including the DIY LoRa gateway. This solution allows the user to add hundreds of such devices using a single LoRa gateway. The experiment was performed in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. The measurements were taken from a perlite slab with cucumber plants.

The cost to build these systems is substantially lower than the corresponding industrial ones and all the blueprints, sketches, designs, and code are accessible under creative commons (CC) license. The vast number of tutorials allows the user to easily configure the device hardware and software-wise. Moreover, there are many tools like DeepPCB (InstaDeep, 2019) that offer automated electronic circuit production using Reinforcement Learning (RL). Circuit Tree (2016) is an Indian company which uses automation for printed circuit board (PCB) design, using GUI based online tools to input the required data. Other companies deployed similar automated options. ChatGPT (OpenAI, 2023) was used to generate Arduino code. Specifically, Fahad (2023) described the modules and the wiring connections via text and the AI tool developed an Arduino sketch. Apart from a few errors that were easily highlighted and corrected, the sketch performed as expected.

1.2. Open-source devices in research

At first, open-source electronics (hardware) were introduced to attract hobbyists and Do-It-Yourself creators. Open electronics are increasingly taken up by broader public audiences linking all age groups, further triggered by the rise of the Internet of Things (Ibrahim, 2015). The scientific community started to adapt to the open-source philosophy, and it didn't take long to pay off in many fields. Oellermann et al. (2022) highlighted three points where open-source electronics aided the scientific community. First, open-source devices help individual researchers by increasing the customization, the efficiency and scalability of the experiments, while increasing data quantity and improving data quality. Second, they assist institutes since the open access to customizable high-end technologies increased the interdisciplinary collaborative networks potential. Third, they succor the scientific community by improving transparency and reproducibility. Also, they help detach research capacity from funding and escalate innovation.

Most of the labs worldwide do not have access to vital funding to keep up with the state-of-the-art lab equipment. Open-source devices contribute to the rapid supply of equipment for the labs with low cost and high level of customizability.

This thesis proposes two open-source devices to aid research in entomology and pomology.

First, a highly accurate temperature logger (Chapter 6) that was designed and developed to study the response of endophytic insect species such as fruit flies to variable temperature conditions including the controlled laboratory and fluctuating field settings. It consists of two fine medical and one harsh environment thermistors that measure temperature. The logger is battery powered and its flanged box allows the researcher to place it anywhere. Specifically, the device was placed on apple and orange trees. During the experiments the logger was capturing measurements every five minutes. The battery lasted for almost ten days. Depending on the logging interval the device battery can last months. The device was designed using documentation available online. The device's cost is less than 110 euros (end of 2020 market prices) and is designed and built using Arduino compatible development boards and programmed using Arduino IDE. Its modular nature allows the researchers to add more sensors and even embed wireless technology for remote monitoring.

Second, a PH logger (Chapter 5) was designed and built to measure pH and temperature during various experiments performed in the laboratory of Pomology, University of Thessaly. The sensor used for this device is a DFRobot Gravity Analog Industrial pH Sensor. The detailed instructions to build the logger for the sensor can be found on the manufacturers' page (DFRobot, 2022). The device is designed and built using Arduino compatible development boards and programmed using Arduino IDE. The total cost for the device is less than 130 Euros (summer of 2023 market prices).

1.3. Evaluation of open-source devices and procedures using statistical methods

Open-source devices are limited only by the imagination of the makers. They have countless uses, based on the creativity and inspiration of their creators. From devices that produce art or even contribute to art, to devices that are used for health monitoring, robotics, augmented reality, environmental monitoring, and others. When it comes to environmental monitoring, which is the topic that this thesis engages, there is a need for a reference device to be compared with the novel open-source device for validation since it is equipped with sensors that measure environmental traits. The comparison is not limited to open-source and industrial devices but extends to the comparison of methods, or two different outputs of measurements generated by open-source devices. Thus, there is need for an official method or protocol for validation of these novel devices outputs against references proven to be trustworthy, accurate and widely used. In other words, a protocol that assesses their agreement. Agreement measures the "closeness" between readings. Thus, agreement is a comprehensive term that contains both accuracy and precision. If one of the devices/methods is treated as the reference, the agreement concerns the degree to which evidence and theory support the interpretation of measurement. This procedure can be covered by a method or measurement comparison study (MCS) which assesses the agreement and similarity between two devices/methods. The goal of this comparison may vary in emphasis by context. Four goals were highlighted by Dunn (2004). First, calibration problems, which study the relationship between the new and the reference device. Second, comparison problems, which assess the level of agreement between two measurement systems whose measurements are on the same scale. Third, conversion problems, which handle comparison of two systems whose measurements are not in the same scale. Fourth, gold-standard comparison problems, which compare two different systems where one of them is known to take measurements without error. In this thesis we focus on the first two goals since the case-studies refer to devices that measure traits on the same scale and no device is considered as a gold standard.

According to Choudhary and Nagaraja (2017a), method comparison studies are designed to compare two competing methods of measurement of the same quantity, having a common unit of measurement. It is assumed that neither device/method (the reference or the new device) measures the true value of a characteristic without error. The measurements are taken by each method on every subject in the study, and there may or may not be repeated measurements (defined in Section 2.1). Repeated measurements are defined as measurements taken by the same device on the same subject under identical conditions. In case the true value changes over time and identical conditions are impossible to achieve, the measurements' dependency is taken into account in the data modeling procedure via longitudinal data analysis techniques.

There are two goals for a method comparison study. First, is the quantification of the extent of agreement between the two devices/methods and the determination whether the agreement is adequate to use the devices/methods interchangeably for a specific task. Thus, regardless of which devices will be used for a specific purpose, there will be no difference in the practical use of the measurement. Second, the comparison of characteristics of the devices/methods such as their biases, accuracy, sensitivity, reliability, precision (defined in Section 2.1.3.) etc. This is defined as the evaluation of similarity. By studying the device's/method's characteristics, the researcher is aided with tools to assess possible sources of disagreement. Furthermore, evaluation of similarity may also suggest corrective actions that may ameliorate the agreement. As an example, the agreement between two devices/methods can be improved by either adding a constant or rescaling one of the methods. The characteristics mentioned are defined in Section 2.1.3.

The two devices/methods may have remarkably similar characteristics but may not agree well. If the methods agree well usually, they will also be similar. For example, two temperature sensors have similar characteristics, but they strongly disagree. This is an example of low precision. Both sensors can have the same (low) precision, so when they measure the same quantity, they might report completely different values. Thus, the precision of the sensors is considered as a source of disagreement. If the new device and the reference have the same precision, which is low, they will certainly not agree well. On the other hand, if the devices agree to a high extent, then certainly they will share similar characteristics.

Similarly, the questions asked during a method comparison study are encapsulated in two different categories (Altman & Bland, 1983):

1. When it comes to the properties of each method, how replicable are the measurements? To put it simply, how much do two or more consecutive measurements of the same subject differ, assuming the quantity does not change during the procedure? Is there any difference between two or more measurements taken by two different operators of the same device?
2. When it comes to the comparison of the two or more devices/methods, do they measure the same thing on average? Is there any relative bias? What additional variability is there which is not explained by the predefined factors stated at the beginning of the research? The additional variability might include different kinds of errors. For example, errors due to user/subject interaction, time/subject interaction etc. When only specific factors are mentioned in method comparison studies, the rest of the possible sources of error sum up to the total error.

These sources of variability can be standardized, thus not considering their effects. However, accounting for all, or most of them might be tricky and confusing either to their modeling or interpretation.

A researcher may want to study if the two devices/methods/observers agree. If not, sources of disagreement would be of interest. Examples for sources of disagreement can be a systematic shift or random error. Thus, a systematic shift can be fixed easily, unlike random error which is more laborious due to variation-reduction needs. The degree of agreement for all possible ranges of measurements can be reported either by using indices or easily interpreting graphs.

There are many reasons one would conduct a method comparison study. Some of them are listed below.

Cost of the reference device.

A new device that measures a specific attribute is designed to replace an expensive device that is widely used and is reliable. In this case the expensive device is the reference of the study and the agreement/disagreement compared to the new device will be calculated for all the possible ranges of measurements. The researcher/individual can then decide if the agreement is sufficient to use the new method interchangeably. Furthermore, the two devices may have acceptable agreement (subjective to the researcher's/individuals' needs) in specific intervals of the measurement range, depending on the application. For example, Luiz et al. (2003) compared the measurements of two methods, computerized tomography versus urography for measuring the inferior pelvic infundibular angle (IPIA) for 52 kidneys. The financial cost of tomography is much higher than measuring the inferior pelvic infundibular angle (IPIA) with urography. Another example refers to the temperature of an oven or a greenhouse, where we expect a limited measurement temperature range (probably 25°C to 400°C for the oven). The researcher/individual may decide to use the new device, saving a great amount of money or even adding more sensors to increase repeatability since they are cheaper than the reference sensors.

Less intrusive new method

A new device is designed to replace a reference device which is intrusive and difficult to use. For example, a study was conducted by Bland and Altman (1986) where percent saturation of hemoglobin with oxygen is measured in 72 adult patients receiving general anesthesia or intensive care with two instruments. The reference is the oxygen saturation monitor (OSM) and is considered intrusive since it uses samples of arterial blood. The new, under investigation, method is the pulse oximetry screener (POS) which is easy to use and non-intrusive.

Improvement in the measurement technique

A new device might be more precise than the reference device. For example, a new digital scale for measuring weight might be more precise than the reference analog device. A new method for diagnosing a plant's disease status using machine learning might be more precise than the reference method which is the observation of the phenotype.

Less time consuming in application.

For instruments used in the field during experiments or environmental condition assessment for crop health, fast readings are important, especially when the conditions around the procedure are harsh. The study of eggshell structure is important for economic, evolutionary-ecological and ecotoxicological purposes. Igic et al. (2010) compared a micrometer and a scanning electron microscopy for measuring eggshell thickness. The first method can be used in the field since the instrument is portable and its working principle does not include complex electronic systems. For thicker eggshells (Kiwi and Palaeognathae eggs) the two methods had reasonable agreement unlike thinner eggshells (cuckoo eggs).

The authors conclude that the more time- and cost-effective method of using a micrometer is no less reliable than using the scanning electron microscope.

Less labor-intensive

Lucas et al. (1987) introduced an automated enzymic micro method for the measurement of fat in human milk. The method is less labor-intensive than the standard Gerber method since it only requires 10 to 50 microliters of milk, thus presenting a technique to work with small mammals. Bland and Altman (1999) demonstrate their method of agreement using the data from Lucas et al. (1987).

1.4. Goals of the thesis

The novelty of this research is based on the development of proper experimental design techniques and formal statistical tools to assess the agreement and similarity using a methodology applied in medical research (Barnhart et al., 2007; Chinchilli et al., 1996; Bland & Altman, 1999; Carrasco & Jover, 2003; Choudhary & Tony, 2006; Barnhart et al., 2005), geography (Cox, 2006), social sciences (Chen et al., 2018), agriculture (Bataka et al., 2021a,b) and other (Astrua et al., 2007). By evaluating agreement between two or more devices/methods/observers we can decide if they can be used interchangeably. Observers can be the operators of a device that measures a specific quantity or a doctor making a diagnosis. For example, two or more operators of a system of sensors might get slightly different readings, depending on the way they hold the sensor (the angle formed to the normal). Methods can be two or more ways to distribute questionnaires to people. According to Chen et al. (2018) the responses from the Epworth Sleepiness Scale, which was originally developed as a paper questionnaire, were compared with the electronic version. Devices may refer to an environmental logger which is expensive and a cheaper alternative (Bataka et al., 2021a) or a medical instrument which is intrusive and unpleasant but is a reference and an alternative which is more passive and less intrusive (Khan et al., 2021).

There are two main and two secondary goals in this thesis. First, the construction of four open-source devices for assisting laboratory experiments and environmental monitoring. Second, the evaluation of open-source devices using statistical methods to assess agreement and similarity with their industrial equivalents. This goal is covered by developing a protocol that discusses the experimental design of a method comparison study and implementing agreement and similarity techniques. Specifically, the agreement and similarity methodologies are used to evaluate four devices designed and built using mostly open-source or DIY material. These devices are used in applications related to agriculture and research. The suggested protocol describes a proper and valid experimental design. For this reason, it contains a detailed guide that describes scenarios that can be adapted for the most common experiments for agricultural instruments that measure environmental characteristics. After choosing the proper experimental design formal statistical analysis is presented, and user-friendly graphs are produced to be interpreted even from non-specialists in the field of statistics. Apart from the graphs, indices that quantify the agreement and similarity of the two or more devices/methods are available and presented embedded with their advantages and disadvantages when it comes to interpretation. Implementation of inferences via formal statistical methods about various characteristics of the method comparison study data complete the assessment of the agreement and similarity between the two devices under investigation. Both goals are implemented via R version 4.2.2. (The R Foundation for Statistical Computing, Vienna, Austria) and Arduino version 1.8.19 IDE. The secondary goals of this thesis cover two main topics. First, the accumulation of all the Arduino IDE sketches, schematics, the material list, and a step-by-step guide for constructing the open-source devices. Second, access to all the R-scripts used to perform the

exploratory and inferential analysis to assess agreement and similarity between two devices/methods. All the data produced by the current research are available on GitHub (Bataka, 2023).

Henceforth, this thesis is organized as follows:

Chapter 2: Section 2.1 describes the statistical tools to assess the agreement and similarity between two methods/devices. Moreover, it describes the statistical models, the indices and the graphs used in this thesis and depicts improper methods to assess agreement with examples. Section 2.2 describes proper experimental design and provides key steps to plan for a method comparison study. Section 2.3 describes the hardware and software this thesis was based on to develop the four open-source devices.

Chapter 3: Section 3.1 contains the goal of the study. Section 3.2 describes the software and code used to develop the device. Section 3.3 discusses the cost of the device. Section 3.4 describes the reference device. Section 3.4 describes the hardware used to develop the device. Section 3.6 describes the experimental design. Section 3.7 describes the analysis. Section 3.8 discusses the results and Section 3.9 concludes the case study's findings with discussions on future directions.

Chapter 4: Presents the design and development of a wireless soil moisture sensor logger. Section 4.1 describes the goal of the study. Section 4.2 describes the design of the device, including the software, hardware, and cost of the device. Section 4.3 describes the reference device. Section 4.4 describes the experimental design. Section 4.5 describes the analysis. Section 4.6 discusses the results and Section 4.7 concludes the case study's findings with discussions on future directions.

Chapter 5: Presents the design and development of a soil and substrate pH logger. Section 5.1 describes the goal of the study. Section 5.2 describes the design of the device, including the software, hardware, calibration method and cost of the device. Section 5.3 describes the reference device. Section 5.4 describes the experimental design. Section 5.5 describes the analysis. Section 5.6 discusses the results and Section 5.7 concludes the case study's findings with discussions on future directions.

Chapter 6: Presents the design and development of FruiTTemp, an open-source temperature logger applied to fruit fly host experimentation. Section 6.1 describes the goal of the study. Section 6.2 describes the design of the device, including the software, hardware, calibration method, power consumption and power saving techniques, cost of the device, bias removing techniques, its accuracy and measurement range. Section 6.3 describes the design of the method comparison study. Section 6.4 describes the Agreement techniques. Section 6.5 discusses the design of the field experiment. Section 6.6 discusses the results. Section 6.7 and 6.8 discuss future work and concludes the case study's findings with discussions on future directions.

Chapter 7: Concludes with a general discussion.

Chapter 8: Discusses the contribution of the current thesis.

2. Materials and Methods

2.1. Agreement and Similarity Techniques

Accurate measurements are the building block of any kind of research. The measurements can be continuous, for example temperature, humidity, weight, speed, pressure etc. or categorical, for example presence or absence of water, color of crop etc. The measurements taken by various devices/individuals are prone to error. Thus, statistical tools are developed to study and report the relationship between measurements containing errors and their corresponding true values.

The structure of the measurements, which is based on the experimental design (discussed in Section 2.2) followed by the researcher falls under two different categories. The **unreplicated** and **repeated measurements**. The unreplicated are very common in research but lack of information needed to evaluate repeatability of the device/method and possible interactions between the methods and the subject. However, repeated measurements provide essential information to assess vital attributes of the method comparison study.

There are three kinds of repeated measurements. Carstensen et al. (2008) introduce the unlinked and linked data. The last case is the longitudinal data (Fitzmaurice et al., 2011):

Repeated measurements taken from:

- **Unlinked data** refer to repeated measurements that are not paired in the sense that the measurements of the two methods are obtained separately. Thus, unlinked data are not necessarily measured concurrently. There is no need for the methods to have the same number of repeated measurements. Typically, these repeated measurements are identically distributed. For example, multiple measurements are obtained in quick succession (such that the true value is not changing) or when a sample's homogeneous specimen is subsampled.
- **Linked data**, in which each subject may experience consecutive measurements over time and are paired. Unlike the unlinked data the devices/methods need to have an equal number of paired repeated measurements per one subject but may vary between different subjects. The true value does not need to stay constant over time but there is no systematic effect of time on the paired trajectories beyond the dependence induced in them by the common measurement time. For example, measurements taken together on different days.
- **Longitudinal data**. Measurements from the two methods are made over a period of time. They are similar to the linked data, with the difference that there is a systematic fixed effect of time or a time-dependent covariate of measurements. The contribution of each subject consists of two trajectories over time, but these may or may not be paired over the measurement occasions. The true value of the subject necessarily changes. Time can be a discrete or continuous quantity. Longitudinal data predicate modeling the measurement trajectories as functions of time, ensuring a treatment of its own.

2.1.1. Simple Linear Model

Fleiss (1999) and Dunn (2004) describe the simple linear model that relates the observed value of a measurement with the true unobserved value. An example of this relationship can be found in agriculture. If Y is the observed value of soil humidity measured by a sensor in a pot, then b is the true unobserved value.

$$Y = \beta_0 + \beta_1 b + e \quad (2.1)$$

Under this model the conditional mean of the method's measurement Y given the true value b is:

$$E(Y|b) = \beta_0 + \beta_1 b$$

And the conditional variance of the method's measurement Y given the true value b is:

$$\text{var}(Y|b) = \sigma_e^2$$

These quantities are interpreted as the average and the variance of an infinitely large number of replications taken by the method, for the same subject, under identical conditions.

The marginal mean of Y is the following:

$$E(Y) = \beta_0 + \beta_1 \mu_b$$

The marginal variance of Y is the following:

$$\text{var}(Y) = \beta_1^2 \sigma_b^2 + \sigma_e^2$$

These quantities represent the average and variance of the measurements in the population from which the subject is being sampled.

An important characteristic of the conditional mean under model (2.1) is that it represents the error-free measurement for the method. However, it should not be confused with the bias-free measurement (which is the true value b). To have bias-free measurements ideally $\beta_0 = 0$ and $\beta_1 = 1$.

- a. The parameters β_0 and β_1 are considered fixed. Their values depend on the measurement method. The parameter β_0 is defined as the **fixed bias** (or differential bias) of the method, while β_1 is defined as the **proportional bias**. The fixed bias is the added constant that the measurement method adds to the true value for every measurement. The proportional bias is based on the measured quantity and is the slope of (2.1). The true value is multiplied by the proportional bias and is interpreted as the amount of change in the measurement method if the true value changes by 1 unit. Both coefficients add some error to the measurement of the method which results in the deviation from the true value. Since the coefficients are fixed the error produced by them is called systematic.
- b. The true value b has a probability distribution over the population of subjects. The distribution's variance is defined as the **between-subject variance** which represents the variability between the subjects and the true values. In simple terms, it represents the heterogeneity in the population.
- c. The quantity e is defined as the error and its distribution is over replications of the same underlying measurement by the same method on the same subject under identical conditions. It has a probability distribution with mean zero. Its variance is defined as the **within-subject variance** and many reliability indices are calculated in terms of this quantity. Usually, we include in the error term all the effects of uncontrolled factors that affect the measurement.
- d. An important assumption for this model is that the within-subject variation remains constant while the measurements' magnitude changes. This assumption is violated most of the time in real-world experiments. Thus, the model is modified to account for this heterogeneity. These methods are described in Section 2.1.2.

A clarification that the reader needs to consider is the discrimination between the true value of the subject and the true value of the measurement method. The true value of the subject can almost never be observed and is considered as a latent trait. The true value of the measurement method is defined as the error-free value of the quantity $\beta_0 + \beta_1 b$.

2.1.2. Models for fitting the method comparison study data.

Many published papers highlight the importance of modeling the data instead of using their sample statistics to assess agreement. This is to account for certain situations, such as heteroscedasticity in the errors, or considering the effect of time on the agreement. Rathnayake and Choudhary (2017), Carrasco and Jover (2003), Carrasco et al. (2009) used mixed-effects models to assess agreement between two or more methods. Moreover, Taffé (2018) used measurement-error models to evaluate bias and precision of two measurement methods. These two models including their extensions for repeated measurements, heteroscedastic errors and time covariates are described below.

To design a method comparison study usually one of the two devices/methods is the reference and the other is a novel method/device under investigation. To assess the agreement between two different methods/devices two different practices are used to produce indices and intervals. First, the use of sample statistics estimated from data like the mean, the standard deviation, or the covariance. This approach is preferred since it is not complicated to perform by individuals with no background in statistics. Second, a model-based approach. Two modeling frameworks are very popular for this application. Mixed-effects and measurement-error models (Cheng & Ness, 1999) fit to the data and their estimated coefficients and variance components are used to produce agreement and similarity indices. Compared to the first practice, model-based approaches offer better understanding of possible sources of disagreement in case mediocre or poor agreement is evaluated. Moreover, special cases like unbalanced experimental designs, or clustered data can be elaborated easily using the latter approach.

Measurement-error model

Following Buonaccorsi (2010) and Choudhary and Nagaraja (2017a), assuming the simple linear model (2.1), we retrieve the following setup for the bivariate data since we need to link the observed (Y_1, Y_2) to the unobserved true value b and the measurement errors associated with the methods:

$$Y_j = \beta_{0j} + \beta_{1j}b + e_j \quad (2.2)$$

For $j = 1, 2$ we define method 1 and method 2.

β_{0j} , β_{1j} , e_j and b are the fixed bias, proportional bias, measurement error and true measurement of the subject respectively.

The assumptions for model (2.2) regarding b are the same as for model (2.1). For the rest of the terms, the following are assumed:

- $e_j \sim N(0, \sigma_{e_j}^2)$, $j = 1, 2$
- The errors are mutually independent of each other and of the true value.

In a method comparison study, researchers are not interested in determining how accurate the individual methods devices are, but rather in comparing them and evaluating their agreement and similarity. Thus, we assume that $\beta_{01} = 0$ and $\beta_{11} = 1$. Method 1 can be either the reference or the new method/device. The

true value is redefined as what method 1 measures on average. However, we cannot claim if this method is accurate or not. Thus, model (2.2) is modified as follows:

$$Y_1 = b + e_1, Y_2 = \beta_0 + \beta_1 b + e_2 \quad (2.3)$$

In the concept of measurement-error models, the modified model (2.3) is a structural model, as opposed to a functional model since the true value b is considered random.

Assuming (Y_1, Y_2) follow a bivariate normal distribution the model-based moments are the following:

Mean:

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mu_b \\ \beta_0 + \beta_1 \mu_b \end{pmatrix} \quad (2.4)$$

Covariance matrix:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_b^2 + \sigma_{e_1}^2 & \beta_1 \sigma_b^2 \\ \beta_1 \sigma_b^2 & \beta_1^2 \sigma_b^2 + \sigma_{e_2}^2 \end{pmatrix} \quad (2.5)$$

From (2.4) and (2.5) the mean d and the variance v^2 for the difference $D = Y_2 - Y_1$ are the following:

$$d = \beta_0 + (\beta_1 - 1) \mu_b$$

$$v^2 = (\beta_1 - 1)^2 \sigma_b^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2$$

Using model (2.3) for the paired measurements (Y_{i1}, Y_{i2}) the model can be written as follows:

$$Y_{i1} = b_i + e_{i1}, Y_{i2} = \beta_0 + \beta_1 b_i + e_{i2} \quad (2.6)$$

A basic Mixed-effects model

Being a special case of the measurement error model obtained by substituting $\beta_1 = 1$ to (2.6), assuming both methods/devices have the same measurement scale and having the same assumptions as (2.6) regarding the model's terms Choudhary and Nagaraja (2017a) describe the mixed-effects model using the following formula (Gelman & Hill, 2006):

$$Y_{i1} = b_i + e_{i1}, Y_{i2} = \beta_0 + b_i + e_{i2} \quad (2.7)$$

- $i = 1, \dots, n$, where n is the number of subjects.
- b_i is a subject specific random effect.
- β_0 is a fixed effect.

The model-based moments are the following:

Mean:

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mu_b \\ \beta_0 + \mu_b \end{pmatrix} \quad (2.8)$$

Covariance matrix:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_b^2 + \sigma_{e_1}^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma_{e_2}^2 \end{pmatrix} \quad (2.9)$$

The measurement differences:

$$D_i = \beta_0 + e_{i2} - e_{i1}$$

from (2.8) and (2.9) the measurement differences D_i 's are i.i.d with $D \sim N_1(d, v^2)$ where,

$$d = \beta_0$$

$$v^2 = \sigma_{e_1}^2 + \sigma_{e_2}^2$$

Building the proper model

Depending on the experimental design of the method comparison study more factors should be added to the mixed-effects model. When the experimental design is based on unreplicated measurements the data will not have adequate information for all the model parameters estimation. Thus, the data collected with repeated measurements are more informative. Moreover, by obtaining repeated measurements certain interactions can be incorporated in the analysis that are present in the data. For example, subject \times method or subject \times time interactions. Furthermore, even though the methods are allowed to have different variances in model (2.6) and (2.7), we assume that they remain constant over the range of measurements. In this case, the errors are considered **homoscedastic**. In case the error variance depends on the magnitude of measurement then the errors are **heteroscedastic**. There are two methods to fix heteroscedastic errors. First, a variance stabilizing transformation of the data is possible by using the log transformation for easier interpretation, but this action will not work most of the times. If this is not the case then modelling heteroscedasticity is the proper way to deal with this trait of the errors (Carroll & Rupert, 1988). The mixed-effects model is extended by letting the variances be functions of a suitable variance covariate parameter (Pinheiro & Bates, 2000). We define u_i , to be the variance covariate parameter for the i^{th} subject. The measures of agreement will be functions of this covariate, allowing us to evaluate the extent of agreement change with the magnitude of measurement.

Last, in the case of longitudinal data, modifications to the mixed-effects model should be made to account for the effect of time such as the mean per method depends on the time covariate.

The mixed-effect model for **unreplicated, paired data** is formed as in (2.7).

The mixed-effects model that includes repeated measurements, possible interactions and homoscedastic variances for each method is formed as follows:

- **Unlinked data:**

$$Y_{i1k} = b_i + b_{i1} + e_{i1k}, Y_{i2k} = \beta_0 + b_i + b_{i2} + e_{i2k} \quad (2.10)$$

- $k = 1, \dots, m_{ij}$, are the repeated measurements

- $i = 1, \dots, n$, is the subject's number

- $j = 1, 2$, is the method's number

- b_{ij} follow independent $N_1(0, \psi^2)$ distributions. This is actually an interaction term. One interpretation for b_{ij} is the effect of method j on subject i . These interactions are subject-specific biases of the methods. They are a characteristic of the method-subject combination that remains stable during the measurement period.

- b_i follow independent $N_1(\mu_b, \sigma_b^2)$ distributions.

- e_{ijk} follow independent $N_1(0, \sigma_{e_j}^2)$ distributions.

- b_i , b_{ij} and e_{ijk} are mutually independent.

To examine the measures of similarity and agreement we must retrieve the parameters of the assumed model (2.10) which produces a bivariate distribution for (Y_1, Y_2) . By dropping the subscripts for the sake of simplicity, we have:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{e_1}^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \psi^2 + \sigma_{e_2}^2 \end{pmatrix} \right)$$

Thus, $D = Y_2 - Y_1 \sim N_1(\beta_0, 2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2)$

Then the model has a total of 6 unknown parameters $(\beta_0, \mu_b, \sigma_b^2, \psi^2, \sigma_{e_1}^2, \sigma_{e_2}^2)$.

- **Linked data:**

Linked data are modeled same as model (2.10) except from the addition of the term b_{ik}^* which represents the random effect of the common time k on the measurements.

$$Y_{i1k} = b_i + b_{i1} + b_{ik}^* + e_{i1k}, Y_{i2k} = \beta_0 + b_i + b_{i2} + b_{ik}^* + e_{i2k} \quad (2.11)$$

- b_{ik}^* follow independent $N_1(0, \sigma_{b^*}^2)$ distributions and they are mutually independent of b_i, b_{ij} and e_{ijk} , which continue to follow the assumptions of model (2.10). The distribution of b_{ik}^* does not depend on either method or time. The term b_{ik}^* can also be interpreted as a subject x time interaction or a subject-specific bias that gets introduced in the measurements at time k .

To examine the measures of similarity and agreement we must retrieve the parameters of the assumed model (2.11) which produces a distribution for (Y_1, Y_2) . By dropping the subscripts, we have:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 & \sigma_b^2 + \sigma_{b^*}^2 \\ \sigma_b^2 + \sigma_{b^*}^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 \end{pmatrix} \right)$$

Thus, $D = Y_2 - Y_1 \sim N_1(\beta_0, 2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2)$

Then the model has a total of 7 unknown parameters $(\beta_0, \mu_b, \sigma_b^2, \psi^2, \sigma_{b^*}^2, \sigma_{e_1}^2, \sigma_{e_2}^2)$.

- **Longitudinal data:**

Often, data collected for method comparison studies follows a cohort of subjects over a period of time. The within-subject errors of the methods may be correlated over time. The time effect is captured primarily by letting the means of the methods depend on it. A trajectory is created for each method on every subject. The trajectories of the observations on the same subject are dependent. Thus, both measures of similarity and agreement depend on time. Rathnayake and Choudhary (2017) proposed modelling the trajectories nonparametrically through penalized regression splines within the framework of mixed-effects models. The model parameters are estimated using maximum likelihood (ML) instead of restricted maximum likelihood (REML) since it only estimates the variance-covariance parameters and not the fixed-effects parameters. REML does not jointly estimate all model parameters, thus, it does not provide a joint covariance matrix of all the parameter estimates which is needed for inference on agreement measures because they are functions of both fixed-effects and variance-covariance parameters. Using simulations, Rathnayake and Choudhary (2017) indicated that the proposed method performs reasonably well for 30 or more subjects. Choudhary and Nagaraja (2017a) proposed a similar method but without the use of penalized splines to model the mean functions. Instead, they used polynomials.

Measurements are to be taken at specific time points $1, \dots, m$, called measurement occasions in case the time points are discrete. In case of the time variable being continuous authors use functions over time. Repeated measurements per subject may be balanced or unbalanced. Meaning that for one subject the repeated measurements length may differ. More than two methods/devices can be used to assess their agreement and similarity. Assuming the errors are homoscedastic, model (2.11) is modified as follows:

$$Y_{ijk} = \mu_j(t_{ijk}) + b_i + b_{ij} + b_{iv_{ijk}}^* + e_{ijk}, \quad (2.12)$$

where

- $i = 1, \dots, n$ denote the subjects.
- $j = 1, \dots, J$ denote the methods.
- The study plans for m visits to collect data.
- t_{ijk} are values of the time covariate $t \in T$, which is treated as continuous variable. In case t is discrete, then $t \in T = \{1, \dots, m\}$ which are the prespecified time points and are called measurement occasions.

- $Y_{ijk}(t_{ijk})$ denote the observation by the j^{th} method on the i^{th} subject at time t_{ijk} , $k = 1, \dots, m_{ij}$, $j = 1, 2$, $i = 1, \dots, n$.
- $k = 1, \dots, m_{ij}$, denotes the k^{th} repeated value on subject i by method j . It denotes the order in which measurements are obtained in a specific visit.
- v_{ijk} is the visit number during which the measurement $Y_{ijk}(t_{ijk})$ taken, where v_{ijk} is the value of a discrete variable $v \in \{1, \dots, m\}$. By definition $m_{ij} \leq m$. It is primarily used to link the observations taken during the same visit.
- $N_i = \sum_{j=1}^J m_{ij}$ is the total number of observations of the i^{th} subject.
- $N = \sum_{i=1}^n N_i$ is the total number of observations in the data.
- $\mu_j(t)$ is the mean function of the j^{th} method.
- b_i is the subjects' random effect which follow independent $N_1(0, \sigma_b^2)$ distributions.
- b_{ij} is the random subject x method interaction term which follows independent $N_1(0, \psi^2)$.
- $b_{iv_{ijk}}^*$ is the subject x visit number interaction term which follows independent $N_1(0, \sigma_{b^*}^2)$.
- e_{ijk} is the error term which follows $N_1(0, \sigma_{e_j}^2)$.
- errors associated with different subjects or methods are independent, but a method's errors on the same subject may be dependent following the correlation structure $cor(e_{ijk}, e_{ijl}) = \rho^{|t_{ijk} - t_{ijl}|}$, where $0 \leq \rho < 1$. This autocorrelation structure is a continuous time analog of the autoregressive structure of order 1 (AR(1)) which can be easily replaced by other correlation models (Pinheiro & Bates, 2000). More about time series data the reader can follow Brockwell and Davis (2016) and for spatial data Cressie (1993).
- for b_{ij} the same assumptions apply as model (2.10).
- $b_i, b_{ij}, b_{iv_{ijk}}^*$ and e_{ijk} are mutually independent.
- In case t is discrete then $t_{ijk} = v_{ijk}$ and $T = \{1, \dots, m\}$. If t is continuous then $t_{ijk} \neq v_{ijk}$ and T may be taken as the observed range of t in the data.
- k is different from v_{ijk} since the third repeated measurement may be taken during the fourth visit, thus $k = 3$ and $v_{ij3} = 4$. When observations are available during every visit the k and v_{ijk} are identical.

Since the means are modeled as functions of time, the within-subject errors of methods are allowed to be correlated over time. The available classes of correlation structures are available in Pinheiro and Bates (2000).

To examine the measures of agreement we must retrieve the counterparts of the assumed model (2.12) which produces a distribution for $(Y_1(t), Y_2(t))$. By dropping the subscripts and assuming two methods/devices, we have:

$$\begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_1(t) \\ \mu_2(t) \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 & \sigma_b^2 + \sigma_{b^*}^2 \\ \sigma_b^2 + \sigma_{b^*}^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 \end{pmatrix} \right)$$

Thus, for the differences $D(t) = Y_2(t) - Y_1(t) \sim N_1(\mu_2(t) - \mu_1(t), 2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2)$

For discrete t : $\mu_j(t) = \beta_{jt}, j = 1, 2$

For continuous t , following Choudhary and Nagaraja (2017a):

$$\mu_j(t) = \beta_{0j} + \beta_{1j}t + \dots + \beta_{qj}t^q, j = 1, 2 \quad (2.13)$$

modeled as a polynomial of a given degree q .

The two distributions depend on t only through the mean functions.

For continuous t , following Rathnayake and Choudhary (2017):

$$\mu_j(t) = \beta_{0j} + \beta_{1j}t + \dots + \beta_{qj}t^q + \sum_{q=1}^Q U_{qj}(t - c_q)_+^p, j = 1, \dots, J,$$

Where Q is the number of knots, $c_1 < \dots < c_Q$ are the locations of knots and $u_{1j} < \dots < u_{Qj}$ are coefficients of the truncated polynomial basis functions $(t - c_1)_+^p, \dots, (t - c_Q)_+^p$ with $(t - c_q)_+^p = (\max\{0, t - c_q\})^p$. The same set of knots are used for both mean functions. The spline coefficients u_{qj} are treated as independent $N_1(0, \sigma_{uj}^2)$ random variables.

Heteroscedastic data

In case the errors of models (2.10, 2.11, 2.12) are heteroscedastic, $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$ are replaced with $\sigma_{e_1}^2 g_1^2(u, \delta_1)$ and $\sigma_{e_2}^2 g_2^2(u, \delta_2)$. For a given u_i denoted as \tilde{u}_i (\tilde{u} for subject i) function g is the variance function and δ is a vector of heteroscedasticity parameters, which for $\delta = 0$ corresponds to homoscedasticity. Variance covariate u is defined in advance (\tilde{u}) and accounts for heteroscedasticity of measurements. Choudhary and Nagaraja (2017a) set $\tilde{u}_i = h(\bar{y}_{i1}, \bar{y}_{i2}) = \bar{y}_{i1}$ if method 1 is the reference and $\tilde{u}_i = h(\bar{y}_{i1}, \bar{y}_{i2}) = \frac{\bar{y}_{i1} + \bar{y}_{i2}}{2}$ otherwise. For the variance function g , two simple models are introduced. First, the power model, where $g(u, \delta) = |u|^\delta$. Second, the exponential model, $g(u, \delta) = \exp(u\delta)$. The parameters δ_j can be estimated while fitting the model using ML and the “nlme” package (Pinheiro et al., 2022). More details on the choice of the variance function g can be found in (Pinheiro & Bates, 2000, Chapter 5). AIC and BIC can be compared to distinguish between different candidate models for the variance functions.

The distribution of (Y_1, Y_2) for unlinked data is the following:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} | \tilde{u} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{e_1}^2 g_1^2(\tilde{u}, \delta_1) & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \psi^2 + \sigma_{e_2}^2 g_2^2(\tilde{u}, \delta_2) \end{pmatrix} \right)$$

For the linked data:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} | \tilde{u} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 g_1^2(\tilde{u}, \delta_1) & \sigma_b^2 + \sigma_{b^*}^2 \\ \sigma_b^2 + \sigma_{b^*}^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 g_2^2(\tilde{u}, \delta_2) \end{pmatrix} \right)$$

And for the longitudinal data:

$$\begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} | \tilde{u} \sim N_2 \left(\begin{pmatrix} \mu_1(t) \\ \mu_2(t) \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 g_1^2(\tilde{u}, \delta_1) & \sigma_b^2 + \sigma_{b^*}^2 \\ \sigma_b^2 + \sigma_{b^*}^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 g_2^2(\tilde{u}, \delta_2) \end{pmatrix} \right)$$

Based on the model's parameters (2.10, 2.11) the heteroscedastic difference distribution is the following for the unlinked and linked data:

$$D | \tilde{u} \sim N_1(\beta_0, 2\psi^2 + \sigma_{e_1}^2 g_1^2(\tilde{u}, \delta_1) + \sigma_{e_2}^2 g_2^2(\tilde{u}, \delta_2))$$

For the longitudinal data:

$$D | \tilde{u} \sim N_1(\mu_2(t) - \mu_1(t), 2\psi^2 + \sigma_{e_1}^2 g_1^2(\tilde{u}, \delta_1) + \sigma_{e_2}^2 g_2^2(\tilde{u}, \delta_2))$$

Models with covariates

So far, the measurements for both methods were considered the only variables vital for the analysis. However, other factors might affect the agreement between the two methods. Such covariates can be temperature, gender, fruit type and others. Igic et al. (2010) reported that the egg type affected significantly the agreement between the two methods under investigation. Thus, a model that includes covariates (categorical or continuous) is presented by Choudhary and Nagaraja (2017a). The covariates might affect the means of the methods (mean covariates), explaining part of the variability in the measurements. The covariates might also interact with the method or affect the error variance (variance covariates). In any case, the extend of the agreement is affected by the covariates. The methodology can be extended for repeated, paired or unreplicated measurements.

For repeated measurements the model is the following:

For unlinked data:

$$Y_{ijk} = \mu_j(x_{1i}, \dots, x_{ri}) + v_i + b_{ij} + e_{ijk} \quad (2.14)$$

For linked data:

$$Y_{ijk} = \mu_j(x_{1i}, \dots, x_{ri}) + v_i + b_{ik}^* + b_{ij} + e_{ijk}$$

- x_1, \dots, x_r are the mean covariates.
- $v_i \sim$ independent $N_1(0, \sigma_b^2)$ and is defined as $v_i = b_i - \mu_b$

- $e_{ijk} \sim$ independent $N_1(0, \sigma_{e_j}^2 g_j^2(\tilde{u}_i, \delta_j))$, accounts for possible heteroskedasticity
- $b_{ij} \sim$ independent $N_1(0, \psi^2)$
- $b_{ik}^* \sim$ independent $N_1(0, \sigma_b^{*2})$

Choudhary and Nagaraha (2017a) describe the detailed methodology for defining mean and variance covariates (Chapter 8).

Diagnostics for model validity

To check if the model assumptions are verified, there are three assumptions that need to be inspected. First, the equal scale assumption. This can be examined by a Bland-Altman plot. A linear trend in the plot suggests a correlation between averages and measurements. However, this trend might also occur due to differences in precisions. Choudhary and Nagaraja (2017a) report that such a trend “is more likely to be visible due to proportional bias since the between-subject variance σ_b^2 dominates the error variances” (p. 25).

Second, the assumption of homoscedasticity, by plotting the standardized residuals against fitted values or the Bland-Altman plot for the data before modeling. The most common shape of heteroscedasticity in the Bland-Altman plot is the fan. When the variation increases as the magnitude of the measurement increases. Further, as Choudhary and Nagaraja (2017a) describe, the plot should be supplemented by a hypothesis test (Chapter 6), where

$$H_o: \delta_1 = \delta_2 = 0 \text{ versus } H_a: \delta_j \neq 0 \text{ for at least one } j = 1, 2.$$

The test can be performed by using a likelihood ratio test by fitting the heteroscedastic (L_2) and homoscedastic (L_1) model, that results when the null hypothesis is true. Then the test statistic is:

$$2\{\log(L_2) - \log(L_1)\}$$

The p-value for the test can be approximated by a χ^2 distribution with two degrees of freedom.

Last, the AIC and BIC of the heteroscedastic and homoscedastic models can be compared.

Third, the normality assumptions, by using Q-Q plots of standardized residuals and predicted random effects. Ideally, each method should be inspected separately to reveal possible method-specific structures.

For the estimates of the parameters, it is possible that one of the estimated variance components is close to zero or has an abnormally large standard error. The cure to this issue is to drop the interaction term b_{ij} or drop b_i and assume (b_{i1}, b_{i2}) are drawn from a bivariate normal distribution and an unstructured covariance matrix.

2.1.3. Some useful definitions

Bias

Following International Organization of Standardization's (ISO, 2023) definition, bias is the difference between the expectation of a test result or measurement result and a true value. Bias is a total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the true value is reflected by a larger bias value. Following model (2.3), β_0 , which is the intercept of the equation for the new method, represents the difference in fixed biases of the same methods/devices. The slope β_1 , of the equation for the new methods, is defined as the proportional bias.

Accuracy (of a measurement method)

A measurement method is said to be accurate if it has no bias in estimating the true value b . According to the US Food and Drug Administration (FDA) (1999) guidelines on bioanalytical method validation accuracy is defined as the closeness of mean test results obtained by the method to the true value of the analyte. The deviation of the mean from the true value, i.e., systematic bias, serves as the measure of accuracy. ISO used accuracy to measure both systematic bias and random error. This thesis will follow FDA's definition.

Accuracy (of a sensor)

The deviation of a measurement obtained using a sensor, from the true value of the subject's trait. For example, the DS18B20 (Maxim integrated products, 2019) temperature sensor has an accuracy of $\pm 0.5^\circ\text{C}$. This means that the deviation from the true temperature is either -0.5°C or 0.5°C . In terms of model (2.1) the method is considered accurate if $\beta_0=0$ and $\beta_1=1$. When it comes to the DS18B20 temperature sensor, the accuracy of the sensor is Y minus the true value. This accuracy though is defined for the measurement range of -10°C to 85°C , according to the manual.

Precision

FDA (1999) defined precision as the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Precision can only be compared under identical conditions.

This is a measure defined by the variability of a set of readings of a sensor for the same input. Under model (2.1) the method is considered fully precise if $\sigma_e^2 = 0$. Thus, the method has no measurement error, which means that its observed value equals its error-free value. A condition that can almost never occur in real life, since there is always some measurement error that can be added due to various sources. For example, an individual is measuring the soil humidity for one hundred rockwool-substrate aquaponic slabs using a handheld sensor. The insertion angle will be slightly different while measuring the slabs. This might affect the measurement and by extension, this is a source of error.

A measurement method may appear highly precise only because it is too rough to distinguish small changes in the true value (Mandel, 1978). The difference in the precisions of the methods can be measured by the precision ratio (Choudhary & Nagaraja, 2017a, p. 12), under model (2.3)

$$\lambda = \frac{\sigma_{e_1}^2}{\sigma_{e_2}^2} \quad (2.15)$$

If $\lambda=1$, the methods are equally precise.

- If $\lambda < 1$, method 1 is more precise than method 2.
- If $\lambda > 1$, method 2 is more precise than method 1.

When we refer to sensor's measurements, we can use another definition for precision, which is the degree to which values cluster around the mean of the distribution of values. Based on this definition, precision and accuracy can be connected by the following figure using an analogy of darts and a target (Figure 2.1).

The sensors measure the same unit under the same conditions five times. The 'x' is one of the five measurements. The solid red circle is the true value. The top left picture suggests that the sensor is accurate and precise since it hits the red solid circle, and all the measurements are very close to each other. Thus, the variation is small. The bottom right picture suggests that the sensor is inaccurate and not precise since the measurements deviate significantly from the red solid circle and the variation is high.

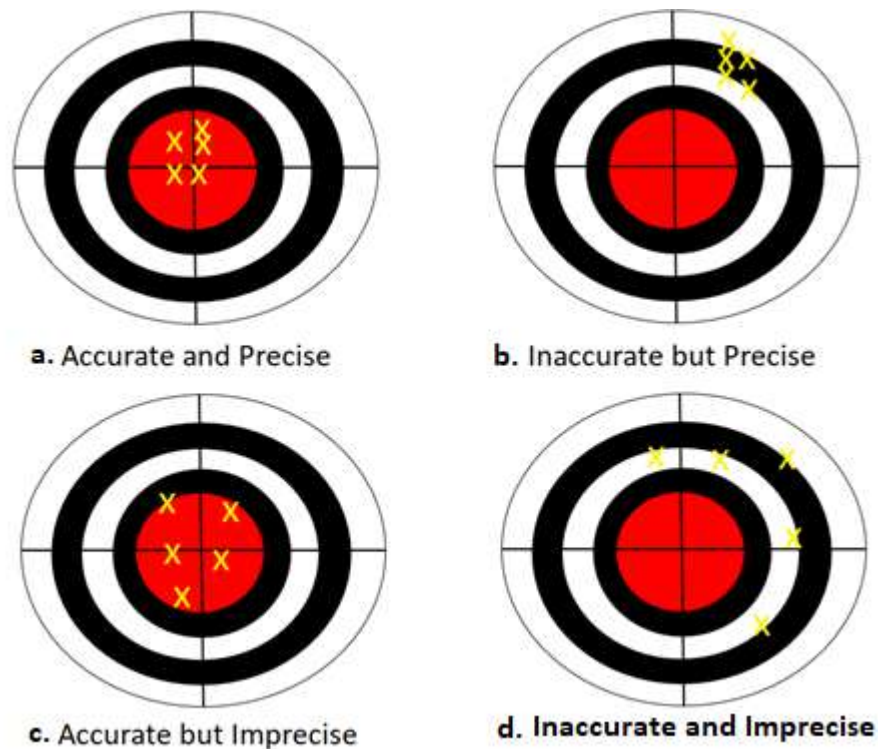


Figure 2.1. Representation of Precision and Accuracy through the darts and target analogy. (a). Accurate and precise measurement. (b). Inaccurate but precise measurement. (c). Accurate and imprecise measurement. (d). Inaccurate and imprecise measurement.

Sensitivity

The ability of a method to discern small changes. More generally, according to National Instruments Corp. (2023) it is defined as the minimum input of physical parameter that will create a detectable output change. This is considered as a relative measure of precision. The notion of sensitivity combines the rate of change and the precision of a measurement method in a single index. For model (2.1), it is given by $\frac{|\beta_1|}{\sigma_e}$.

If β_1 is large, a small change in the true value b will cause a comparatively large change in its measured value, resulting in increased sensitivity. The larger the sensitivity, the more effective is the measurement method. Under model (2.3) the sensitivities of two measurement methods can be compared using the **squared sensitivity ratio**:

$$\gamma^2 = \frac{(\text{sensitivity of method 2})^2}{(\text{sensitivity of method 1})^2} = \frac{\frac{\beta_1^2}{\sigma_{e_2}^2}}{\frac{1}{\sigma_{e_1}^2}} = \beta_1^2 \lambda \quad (2.16)$$

- If $\gamma^2 = 1$, the methods yield the same β_1 (proportional bias) and precision.
- If $\gamma^2 < 1$, method 1 is more sensitive than method 2.
- If $\gamma^2 > 1$, method 2 is more sensitive than method 1.

Reliability (of a measurement method)

The reliability of a measurement method is defined as the proportion of variation in observed measurements that is not explained by the error variation inherent in the method (Choudhary & Nagaraja, 2017a, p. 4). Considering model (2.1), reliability is defined as follows:

$$1 - \frac{\text{var}(e)}{\text{var}(Y)} = 1 - \frac{\sigma_e^2}{\beta_1^2 \sigma_b^2 + \sigma_e^2} = \frac{\beta_1^2 \sigma_b^2}{\beta_1^2 \sigma_b^2 + \sigma_e^2} = \frac{1}{1 + \frac{\sigma_e^2}{\beta_1^2 \sigma_b^2}}$$

The interpretation of reliability is the correlation between two independent replications of the same underlying measurement. It ranges between 0 and 1. A high value for the reliability of the method indicates that the error variation is small compared to the variation in the error-free values. On the other hand, a drawback of this index is that it will increase as the population heterogeneity increases even if the precision of the method does not change. While agreement parameters are based on measurement error, reliability parameters are highly dependent on the heterogeneity of the study sample (Vet et al., 2006).

Resolution (of a sensor)

The smallest portion of the signal that can be observed. For example, the analog to digital converter of the development board Arduino Uno is 10 bits. This means that it can detect $2^{10} = 1024$ values. Since Arduino Uno has a reference voltage of 5V, then it can detect $\frac{5}{1024} = 0.00489$ Volts or 4.89 mV. If a soil humidity sensor output is 2mV for 1% and 4mV for 2% then the board cannot distinguish the two values since it can only read increments of 4.89mV. A solution to this problem is given in Chapters 5 and 6.

Repeatability (of a measurement method)

The ISO (2023) defined repeatability as precision under repeatability conditions. Repeatability conditions are the observation conditions where independent test/measurement results are obtained with the same method on identical test/measurement items in the same test or measuring facility by the same operator using the same equipment within short intervals of time. To achieve repeatability conditions, the following conditions must be valid. First, the same measurement procedure or test procedure must be followed. Second, the same operator should take the measurements. Third, the same measuring or test equipment should be used under the same conditions. Fourth, the measurements should be taken in the same location. Fifth, the measurements should be repeated over a brief period. In agreement studies, repeatability of a method refers to its agreement with itself. This is also known as the **intra-method** agreement. It is very important to assess the repeatability of a method since, if it does not agree with itself, it will certainly not agree with another method. To assess repeatability, repeated measurements must be obtained per experimental unit (discussed in Section 2.1).

2.1.4. Graphical methods for assessing Agreement and Similarity

Bland Altman Plot and 95% Limits of Agreement

Bland and Altman (1983) proposed a graphical approach to compare two different methods in medicine. The method has been cited over 40,000 times since it is very straightforward to the non-statistical expert audiences and is easily interpretable. To compare two different methods/devices, paired measurements are taken by the two methods/devices per subject. It is important that the paired measurements come from a unique subject, to validate the assumption of independence. The paired differences (y-axis) are plotted against the averages of the pairs (x-axis). The averages are the estimates of the true value. An interval called **limits of agreement** is calculated using the mean of the paired differences \bar{d} of the two methods/devices and their standard deviation s_d . They are interpreted as the interval within which 95% of differences between measurements made by the two methods/devices are expected to lie. The upper and lower limit is then superimposed to the plot as horizontal lines, along with the mean of the differences. If the paired differences are normally distributed, then we would expect 95% of them to lie between $\bar{D} - 1.96 \cdot s_d$ and $\bar{D} + 1.96 \cdot s_d$. Specifically, when the researcher is interested to compare two methods with no repeated values, paired samples are acquired for each experimental unit, and the graph is constructed as follows:

1. On the x-axis we plot the averages of the paired measurements.
2. On the y-axis we plot the averages of the differences of the paired measurements.
3. The mean and the limits of agreement are plotted.
4. The horizontal line $y = 0$ can also be plotted to contrast the sample mean with the best-case scenario.

Figure 2.2 shows a Bland-Altman Plot (Chapter 4 data). The limits of agreement are [-0.363, 13.605]. By using only the statistics we cannot infer if the agreement between the two methods is sufficient. The clinically acceptable difference (CAD) must be stated by the researcher before the investigation. CAD is defined (Altman & Bland, 1983; Bland & Altman, 1999) as the maximum allowable difference between two measurements that would adversely affect clinical decisions. In some applications this agreement can be acceptable and in others not.

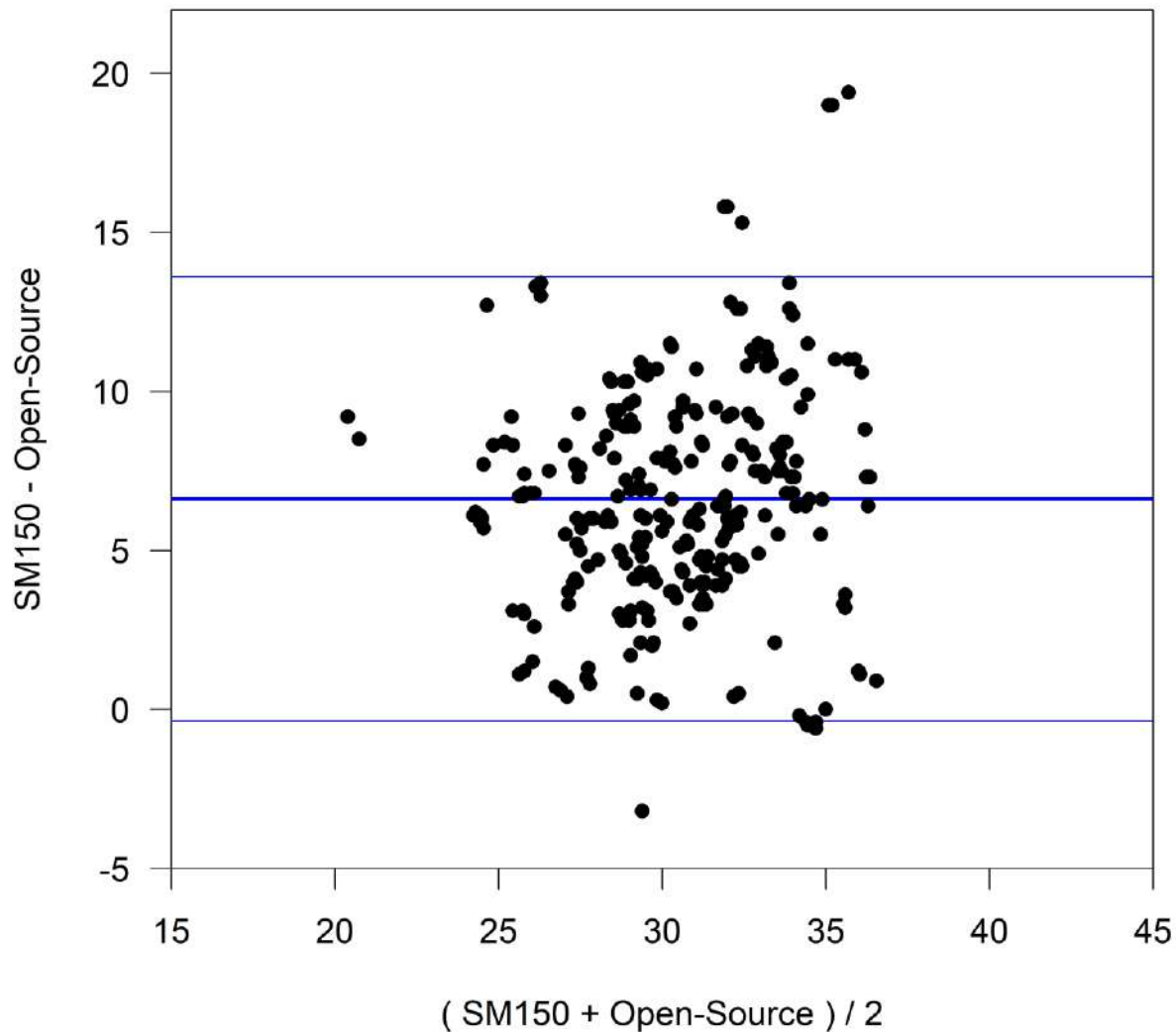


Figure 2.2 Bland-Altman Plot. The limits of agreement are $[-0.363, 13.605]$. Thus, 95 % of the differences are expected to lie between -0.363 and 13.605 .

There may be a consistent tendency for one method to exceed the other. This is defined as the bias and can be estimated by calculating the mean difference. The variation about the bias can be estimated by the standard deviation of the differences.

Despite the easy implementation of the method, its assumptions are usually violated when using real world situations and the estimates for the limits of agreement may not be meaningful. For this reason, further analysis is required to fix these issues that leads to a more complex situation. The assumptions and the cure to their violations according to Bland and Altman are described below.

Assumption 1: The bias is uniform throughout the range of measurements.

In many cases, the bias (or mean of the differences) might change, depending on the magnitude of the measurement. The differences may tend to be in one direction for lower values of the trait being measured and in the other direction for higher values. Non-uniform bias might also be present if the differences of the measurements are not normally distributed. The assumption of normality may not be valid in case the distribution is skewed or has exceptionally long tails. A simple cure for this violation is the logarithmic transformation of both measurements before the analysis. The interpretation will remain simple by back transforming the data and interpreting the limits for the ratio of the original measurements. When the logarithmic transformation fails to fix the problem, then the authors (Altman & Bland, 1983) suggest regressing the difference between the methods measurements on the average of the two methods.

$$D = b_0 + b_1 A \quad (2.17)$$

- D is the difference between the methods.
- A is the average for the two methods.

If b_1 is significantly different from zero, then the estimated difference between the methods is obtained from equation (2.17) for any true value of the measurement, estimated by A .

Assumption 2: The variability is uniform throughout the range of measurement.

If this assumption is violated and the scatter around the line of equation (2.17) is not uniform, then the scatter of the residuals from model (2.17) needs to be modeled as the function of the size of the measurement (estimated by A). Following Bland and Altman (1999), the residuals are assumed to have a normal distribution regardless of the size of measurement, which is a natural extension of the assumption we make in such analyses. We regress the absolute values of the residuals, on A :

$$R = c_0 + c_1 A \quad (2.18)$$

R is the absolute value of the residuals.

The limits of agreement are obtained by combining the two regression equations (2.18,2.19) (Altman, 1993):

$$b_0 + b_1 A \pm 2.46\{c_0 + c_1 A\} \quad (2.19)$$

The assumptions and method followed to form equation (2.19) are explained in detail in Bland and Altman (1999).

Both assumptions can be checked by plotting the average of the paired measurements (x-axis) vs the difference of the paired measurements (y-axis).

In case the normality assumption is violated, the authors provide a non-parametric solution.

Another important concept in method comparison studies that affects agreement is the repeatability of each method. Initially when Altman and Bland (1983) introduced the limits of agreement, replicate measurements were not taken into consideration since they were rarely made in method comparison studies. Having only one measurement, it was impossible to tell which method is more precise. If one method has poor repeatability, then the agreement is bound to be poor. Lack of agreement in unreplicated studies may suggest that the new method cannot be used interchangeably with the reference. However, this might be caused by poor repeatability of the reference method.

To quantify the repeatability of a method, Bland and Altman (1986) suggested to correct the standard

deviation for the LOA by calculating $s_c = \sqrt{s_D^2 + \frac{s_1^2}{4} + \frac{s_2^2}{4}}$. Furthermore, Bland and Altman (2007)

suggested using one-way analysis of variance to retrieve two types of variances. The one for repeated differences between the two methods on the same subject and that for the differences between the averages for the two methods across subjects. For the model, subjects will be the random factor and the within-subject standard deviation can be estimated from the square root of the residual mean square. To assess the precision of each method, their standard deviations can be compared.

Limits of agreement can be calculated using the standard deviations of the methods. Two readings by the same method will be within $1.96\sqrt{2}s_w$, where s_w is the within-subject standard deviation. This value is called the **repeatability coefficient**. It is interpreted as the value within which any two readings by one method/device would lie for 95% of subjects. In case the 95% limits of agreement are similar to the 95% repeatability coefficients, then the lack of agreement can be explained by the lack of repeatability. If the limits of agreement are wider compared to the repeatability would indicate, then another source of disagreement is present. The authors give two different cases for studying repeatability. First, repeated measurements are taken from a subject, but the true value changes. Second, repeated measurements are taken from a subject, but the true value stays constant.

Bland and Altman (1986, 1999, 2007) discuss the estimation of the LOA using repeated measurements. Furthermore, they discuss intra-method agreement by introducing the repeatability coefficient. Carstensen et al. (2008) focus on explicit modeling of the data to estimate LOA. Lai and Shiao (2005) model the differences instead of the individual measurements. Choudhary and Nagaraja (2017a) discuss the estimation of LOA for assessing the repeatability of each method (intra-method agreement) using model's parameters and variance components.

Specifically, for unlinked data under model (2.10), the 95% LOA for method j is evaluated using the following:

$$\pm 1.96\sqrt{2}\sigma_{e_j}, \quad j=1,2 \quad (2.20)$$

For linked data under model (2.11):

$$\pm 1.96\sqrt{2\left(\sigma_{b^*}^2 + \sigma_{e_j}^2\right)}, \quad j=1,2$$

Moreover, for the 95% limits of agreement (inter-method agreement) for unlinked, linked and longitudinal data using model (2.10, 2.11 and 2.12):

For homoscedastic errors (linked and unlinked data).

$$LOA = \beta_0 \pm 1.96 \cdot \sqrt{2\hat{\psi}^2 + \hat{\sigma}_{e_1}^2 + \hat{\sigma}_{e_2}^2}, j=1,2 \quad (2.21)$$

For homoscedastic errors (longitudinal data).

$$LOA = \mu_2(t) - \mu_1(t) \pm 1.96 \cdot \sqrt{2\hat{\psi}^2 + \hat{\sigma}_{e_1}^2 + \hat{\sigma}_{e_2}^2}, j=1,2$$

For heteroscedastic data we replace $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$ with $\sigma_{e_1}^2 g_1^2(u_i, \delta_1)$ and $\sigma_{e_2}^2 g_2^2(u_i, \delta_2)$ as described in Section 2.1.2.

Nonparametric methods for LOA estimation

When the differences of the measurements for the two methods are not normally distributed, Bland and Altman (1999) describe a nonparametric approach to compare two methods/devices. The difference versus average plot is constructed and then the proportion of differences that is greater than the CAD is calculated. The reference values are indicated on the plot. Alternatively, they calculate the proportion of values (fixing a percent) outside the CAD by ordering the observations and use the range of values remaining after defining a fixed percentage by the researcher. The centiles can be superimposed on the plot. Confidence intervals are provided by using the standard method for binomial proportions or by using the standard error of a centile.

In cases where the two assumptions for the LOA are not met even after applying the remedies Bland and Altman (1986, 1999) suggested, then (Gerke, 2020; Frey et al., 2020) reviewed nonparametric quantile estimators for the 2.5% and 97.5% quantiles of the differences by using simulations.

Inference for LOA

It is important to calculate the 95% confidence intervals for the LOA. The estimates of the LOA alone cannot guarantee coverage on the range of possible differences between two measurements, nor do they guarantee that future differences of the two measurements will fall within the LOA. Thus, they should not be used directly for inference (Hamilton & Stamey, 2007). The $100(1-\alpha)\%$ confidence interval of the $100(1-\gamma)\%$ LOA lower and upper confidence bounds are calculated as follows:

$$LCB = \bar{D} - z_{1-\frac{\gamma}{2}} s_D - t_{n-1, 1-\frac{\alpha}{2}} s_D \sqrt{\frac{1}{n} + \frac{z_{1-\frac{\gamma}{2}}^2}{2(n-1)}}$$

$$UCB = \bar{D} + z_{1-\frac{\gamma}{2}} s_D - t_{n-1, 1-\frac{\alpha}{2}} s_D \sqrt{\frac{1}{n} + \frac{z_{1-\frac{\gamma}{2}}^2}{2(n-1)}}$$

Where $z_{1-\frac{\gamma}{2}}$ is the cumulative $100(1-\gamma)\%$ percentile of a standard normal distribution and $t_{n-1, 1-\frac{\alpha}{2}}$ as the cumulative $100(1-\alpha)\%$ percentile of t distribution with $n-1$ degrees of freedom. Generally, α and γ are set as 0.05.

According to Bland and Altman (1986), these bounds issue an expression of the uncertainty in the estimated LOAs. If these bounds fall within the CAD then the two methods agree sufficiently. Simultaneous hypothesis can be formed for the upper and lower confidence bound:

$$H_{01}: \text{Lower bound} < -\delta, H_{a1}: \text{Lower bound} \geq -\delta$$

$$H_{02}: \text{Upper bound} > \delta, H_{a2}: \text{Upper bound} \leq \delta$$

Carkeet (2015) proposed exact parametric confidence intervals for Upper or Lower LOAs considered individually or as pairs.

The author aids the researchers by submitting a table with coefficients calculated as k values corresponding to γ values of 0.025 and 0.975 for a range of degrees of freedom. For the upper 95% LOA he discusses Bland's and Altman's approximate method, which assumes that the calculation of the LOAs is performed separately and not at the same time.

Jan and Shied (2018) proposed an exact interval procedure for LOA and provide a R-script in the supplementary files of their paper.

Problems with the limits of agreement

The limits of agreement by Bland and Altman (1986, 1999) are very popular and easy to implement. However, there are many problems applying this method as Stevens et al. (2017) report. Specifically, no model is assumed for the relationship between the measurements of the two methods. Replicate measurements are not considered and the parameters of their equation cannot be estimated separately. Thus, without separate estimates proportional, fixed bias and σ_1, σ_2 (precisions of the two methods) cannot be distinguished. For this reason, there is no way of telling which method is more precise and there is risk to reject a new method due to imprecision of the reference. Further, by using the difference plot to visualize the agreement, confounding biases cannot be disentangled. Without the additional information added by repeated measurements Bland-Altman plot can be misleading, and the precision cannot be assessed. Another issue for this method is false correlation. When using the average on the x-axis (which serves as an estimate of the true value measured), a significant relationship between the differences and true values might show while it is not there. This can be avoided if repeated values are present. Even though Bland and Altman (1999, 2007) suggested an extension for the limits of agreement for repeated measurements, their recommendation for averaging them on a single subject by a particular measurement system produces LOA that are too narrow.

Stevens et al. (2017) suggested an alternative method for estimating the limits of agreement called probability of agreement plot.

Probability of Agreement Plot

Stevens et al. (2017) proposed a single metric, called probability of agreement and a plot to quantify two measurement system's agreement. This method attempts to overcome Bland and Altman's dearth. The authors suggest quantifying the probability that the observed differences are small enough to be considered clinically acceptable. They suggest the following mixed-effects model:

$$Y_{ilk} = S_i + M_{ilk} \quad (2.22)$$

$$Y_{i2k} = a + \beta S_i + M_{i2k} \quad (2.23)$$

where

- $i = 1, \dots, n$ indexes the subjects.
- the first equation corresponds to the reference method, while the second to the new method.
- $k = 1, \dots, r$, denotes the repeated measurements for each subject.
- S_i is a random variable that represents the unknown true value of the measurement for subject i where $S_i \sim N(\mu, \sigma_s^2)$.
- M_{ijk} is a random variable which represents the measurement error model of system $j = 1, 2$ where $M_{ijk} \sim N(0, \sigma_j^2)$

Thus, if the researchers specify in advance the clinically accepted difference, or $CAD = (-c, c)$ then the probability of agreement based on models (2.22, 2.23) is defined as:

$$\theta(s) = P(|Y_{i2} - Y_{i1}| \leq c \mid S_i = s)$$

And can be rewritten as:

$$\theta(s) = \Phi\left(\frac{c - \alpha - (\beta - 1)s}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) - \Phi\left(\frac{-c - \alpha - (\beta - 1)s}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right), \quad (2.24)$$

Where $\Phi(\cdot)$ is the standard normal cumulative distribution function evaluated at x .

Thus, using probabilities of this form, a probability plot can be constructed using the estimated probability of agreement across a range of plausible values for s (Figure 2.3). Maximum likelihood estimation is used to obtain the following estimates: $(\mu, \alpha, \beta, \sigma_1, \sigma_2, \sigma_s)$ from models (2.22, 2.23) and (2.24) and are used to estimate $\theta(s)$. The interpretation of $\theta(s)$ is very simple, even for non-statisticians. Even if we extend models (2.22, 2.23) to include more terms or account for heteroscedasticity the interpretation will not change.

To estimate and interpret $\theta(s)$, the authors suggest setting a predefined value of $\theta(s) \geq 0.95$, which is similar to the typical limits of agreement. In case the probability of agreement plot indicates low $\theta(s)$ in the range of interest for s , then we can investigate possible sources of disagreement by checking $(\mu, \alpha, \beta, \sigma_1, \sigma_2, \sigma_s)$ in other words assess the similarity. If the researcher wishes to assess the overall agreement of the measurement methods by summarizing it into a single index, the parameter estimates of models (2.22, 2.23) can be used, and the unconditional probability of agreement is:

$$\theta = P(|Y_{i2} - Y_{i1}| \leq c) = \Phi\left(\frac{c - \alpha - (\beta - 1)\mu}{\sqrt{(\beta - 1)^2 \sigma_s^2 + \sigma_1^2 + \sigma_2^2}}\right) - \Phi\left(\frac{-c - \alpha - (\beta - 1)\mu}{\sqrt{(\beta - 1)^2 \sigma_s^2 + \sigma_1^2 + \sigma_2^2}}\right)$$

This summary is appropriate when the probability of agreement is similar for all values of s .

Overall, we can investigate the probability of agreement plot and if s is close to the mean throughout the range of the measurements of interest, then we can summarize the agreement into a single index θ . Even though Bland and Altman (1999) propose a non-parametric approach similar to the probability of agreement, its lack of modeling does not allow the use of all the information retrieved from the data.

In case the CAD is not known in advance, the authors suggest repeating the analysis for various c to investigate the sensitivity of the conclusion to this value. For a specific value of s , the analysis can be summarized with a plot of $\theta(s)$ versus c . An alternative way to choose c , is to use Lin's TDI (discussed in Section 2.1.5) and according to the index value the researcher can choose the proper c covered by the desired population.

To assess the validity of the method two diagnostics are proposed by the authors. First, a QQ-plot of the average repeated values per subject to check the normality of the data. The QQ-plot can be extended by superimposing the quantiles of 50 simulated normal datasets with mean and variance equal to the sample mean and sample variance of the n averages. To check whether the variation across the range of measurements for every method is constant a repeatability plot is used. It is an individual values plot of

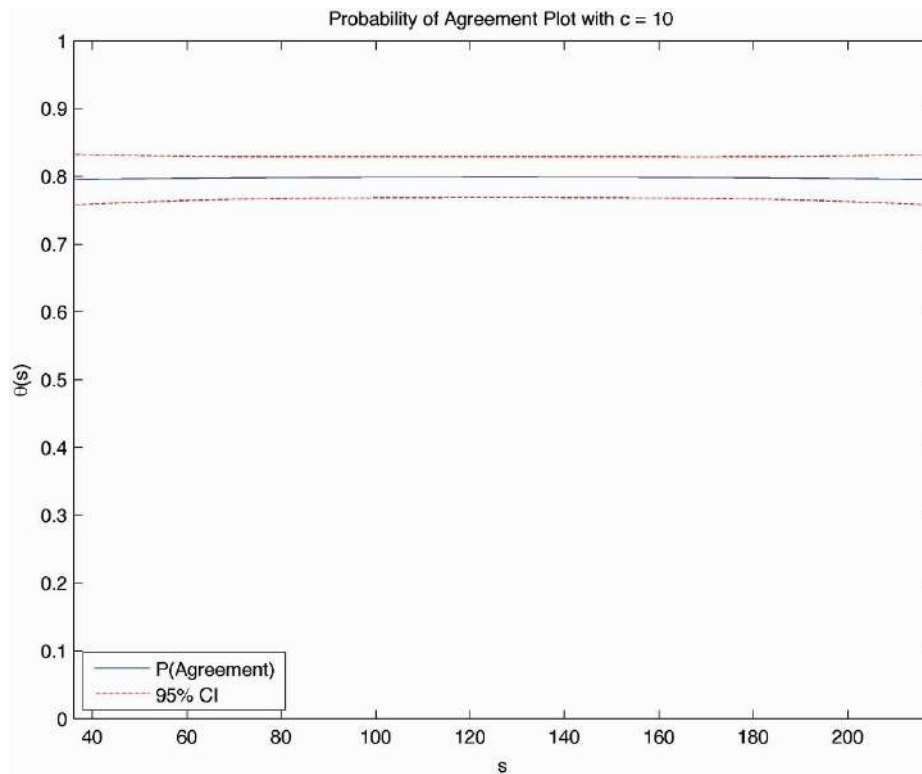


Figure 2.3. Probability Agreement Plot for comparing two methods. (Source: Stevens et al., 2017)

the residuals of the replicate measurements on each method versus the average of those replicate measurements, ordered by size. If there is no relation between residuals and averages, then the variance is homoscedastic.

Bias and Precision Plot

Taffé (2018) proposed various plots to assess bias and precision of measurement methods in method comparison studies. He pointed out Bland's and Altman's limits of agreement deficiencies when the methods have unequal variances. Specifically, for unequal method variances a zero slope is possible in the presence of differential bias. Moreover, Bland and Altman's methodology allows the estimation for the precision of the differences but not of each method separately (Taffé et al., 2020). Last, a restrictive assumption of the Bland and Altman method, is that the ratio of the two variances of measurement error must be strictly equal to the proportional bias. This assumption is unlikely to hold in practice, thus, biased estimates are provided.

For the estimation of the parameters used for the study, Taffé (2018) followed the measurement error model proposed by Nawarathna and Choudhary (2015) but proposed his own estimation procedure. Taffé's estimation method is based on an empirical Bayes approach, which the author considers simpler to implement compared to Nawarathna's and Choudhary's estimation method. This method considers the possible heteroskedastic variances in the model. He also investigated and reported different cases where Bland's and Altman's plot is misleading by using simulated data. Taffé's method identifies and quantifies the amount of fixed and proportional bias. Furthermore, he proposed a recalibration method to increase the agreement between the new and the reference method. Last, he proposed two new plots, **bias, precision and compare plot** as a visual aid to the researcher. To use this method several measurements with the reference standard and possibly one with the new method for each individual is required. There is no need for individuals to have equal number of repeated measurements. The measurement error model used by Taffé is the following:

$$Y_{1ij} = \alpha_1 + \beta_1 x_{ij} + \varepsilon_{1ij}, \quad \varepsilon_{1ij} | x_{ij} \sim N(0, \sigma_{\varepsilon_1}^2(x_{ij}, \theta_1)) \quad (2.25)$$

$$Y_{2ij} = \alpha_2 + \beta_2 x_{ij} + \varepsilon_{2ij}, \quad \varepsilon_{2ij} | x_{ij} \sim N(0, \sigma_{\varepsilon_2}^2(x_{ij}, \theta_2)) \quad (2.26)$$

$$x_{ij} \sim f_x(\mu_x, \sigma_x^2) \quad (2.27)$$

- Y_{kij} is the j th replicate measurement by method k on individual i , $j = 1, \dots, n_i$ $i = 1, \dots, N$, $k = 1, 2$,
- n_i denotes the number of repeated measurements per subject.
- x_{ij} is a latent trait variable with density f_x representing the true unknown trait.
- ε_{kij} represents measurement errors by method k .
- The variances of these methods $\sigma_{\varepsilon_k}^2(x_{ij}, \theta_k)$ are heteroskedastic and increase with the level of the true latent trait x_{ij} in a way to be precisely specified by Taffé (2018), which depends on the vectors of unknown parameters θ_k .

- The mean value of the latent trait is μ_x and the variance is σ_x^2 .
- It is assumed that the latent variable represents the true unknown but constant value of the trait for individual i and therefore $x_{ij} \equiv x_i$.

Since we are not interested to measure the accuracy of the reference method rather than measuring the agreement between the new and the reference method, we consider the reference method as method 2 and set $\alpha_2 = 0$ and $\beta_2 = 1$.

Even though Choudhary and Nagaraja (2017a) suggested a model that includes method by subject interactions (pp. 19, 20), Taffé did not include any since they are considered to be absorbed into the measurement error terms. This is desirable since his method focuses on the identification of fixed and proportional bias and the recalibration of the new method.

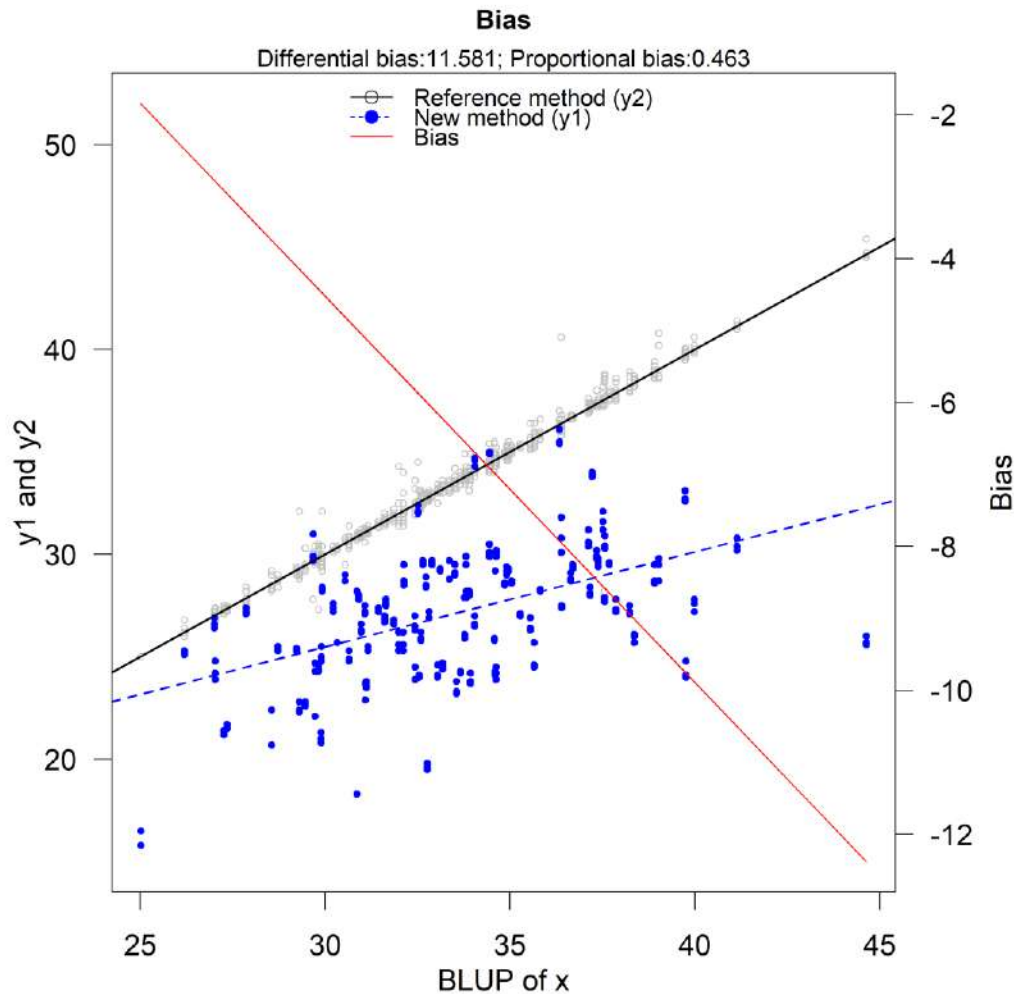


Figure 2.4. Bias plot. The differential bias is 11.581 while the proportional bias is 0.463 (data from Chapter 4).

The modification of the classical measurement error model Taffé applied considers that the heteroscedasticity depends on the latent trait and not on the observed average. The model is assumed to be linear even though non-linear functions of x_i can be used and easily interpreted. To visually assess the plausibility of the straight-line model a graphical representation of $|\hat{\varepsilon}_{2ij}^*|$ versus \hat{x}_i provides a good start. The term $|\hat{\varepsilon}_{2ij}^*|$ is the regression of the absolute values of the residuals $\hat{\varepsilon}_{2ij}^*$ from the linear regression model $y_{2ij} = \alpha_2^* + \beta_2^* \hat{x}_i + \varepsilon_{2ij}^*$ on \hat{x}_i (the estimate of the latent trait) by ordinary least squares. This is similar to Bland and Altman's method to compute a smooth estimate of the heterogeneous variance of the measurement errors.

The estimation method is not described since it is beyond the scope of this thesis. Taffé published a STATA package (Taffé et al., 2017) and an R package (Taffé et al., 2019) which is available in the R repository. The R function *bias_plot()* (Figure 2.4) outputs a plot where the fixed and proportional bias are reported along with their confidence intervals and a plot. Specifically, it is a scatter plot of y_{1ij} and y_{2ij} versus the best linear unbiased predictor (BLUP) \hat{x}_i along with two regression lines. Another scale is added (y-axis on the right) showing the relationship between the estimated amount of bias and \hat{x}_i . The *bland_altman_plot()* function outputs the extended Bland Altman plot when there are repeated measurements with possibly heteroscedastic variance of measurement errors. The *precision_plot()* (Figure 2.5) allows the visual comparison of the precision of the new measurement method after recalibration with that of the reference standard, by creating a scatter plot of the estimated standard deviations against the best linear prediction of the true latent trait \hat{x} . The recalibration procedure is performed by computing the following:

$$y_{1ij}^* = \frac{y_{1ij} - \hat{\alpha}_1^*}{\hat{\beta}_1^*}, \quad (2.28)$$

where $\hat{\alpha}_1^*$ is the estimate of the proportional bias and $\hat{\beta}_1^*$ is the estimate of the fixed bias and y_{1ij}^* is the recalibrated value. The method performs well, according to simulations, with a sample size of 100 subjects and 10 to 15 repeated measurements per individual from the reference method and only 1 from the new. It is possible that after the recalibration procedure the novel method turns out to be more precise than the reference. *Compare_plot()* function displays the agreement before and after recalibration (Figure 2.6).

The authors assumed that the latent trait is constant for every subject when taking repeated values. This assumption may be relaxed in case there is a trend in the repeated measurements. Thus, the data are considered longitudinal, and the author suggests a method to account for time dependency between the clustered measurements.

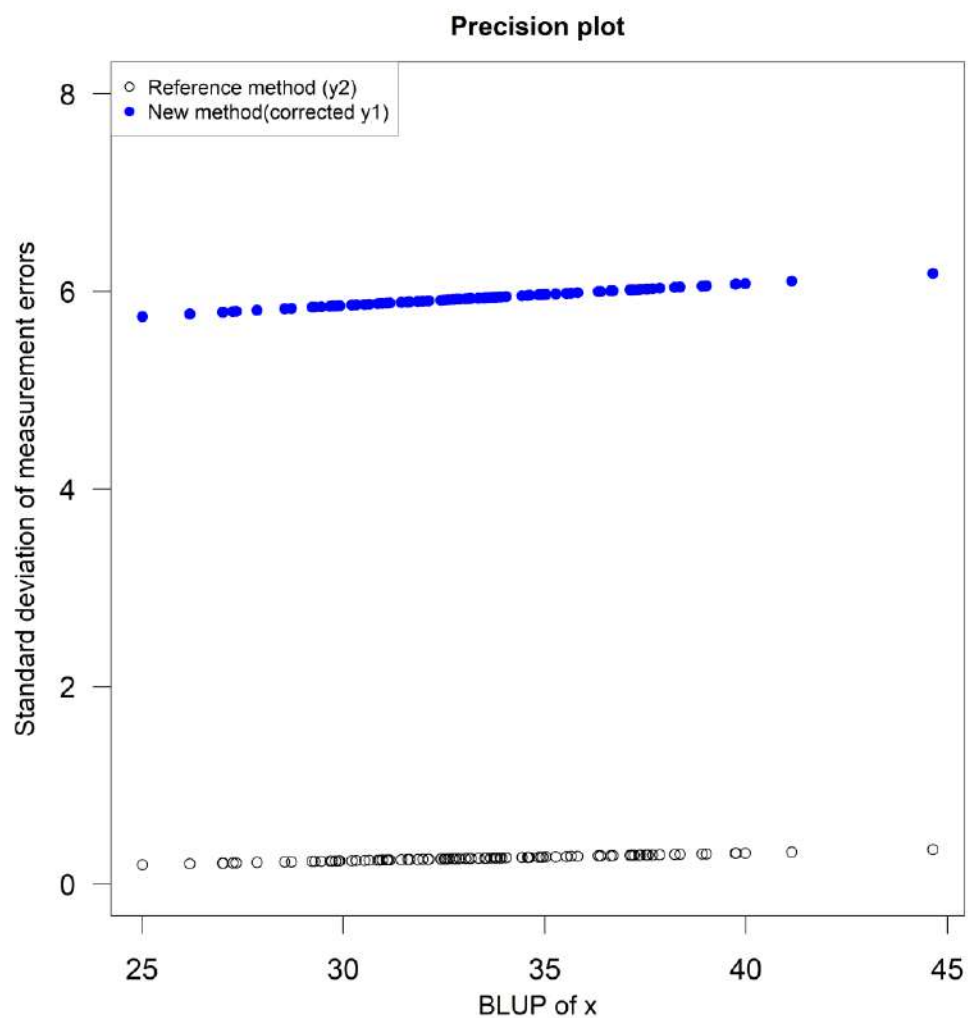


Figure 2.5. Precision plot. The reference method is more precise compared to the new method (data from Chapter 4).

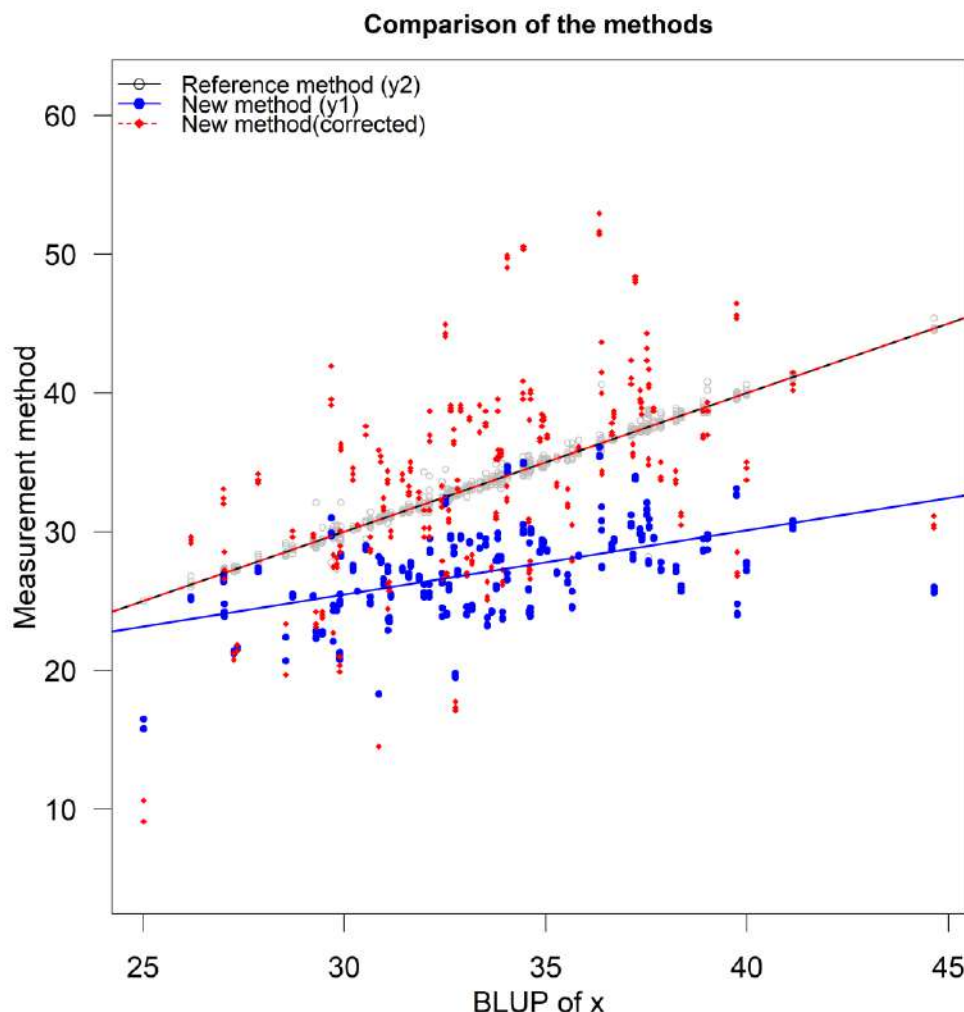


Figure 2.6. Compare plot. The methods agreement improves after recalibration (data from Chapter 4).

Passing – Bablok, Deming and Thiel – Sen regression

Most of the method comparison procedures evaluate agreement between two devices/methods using distributional assumptions for the experimental data which are rarely met. Deming regression (Cornbleet & Gochman, 1979) (Figure 2.7) fits a line through paired measurements and assumes that the dependent and independent variables are measured with error. However, it requires that the error variance ratio is known. Thiel – Sen regression (Theil, 1950) is considered non-parametric but assumes no measurement errors in the independent variable. Passing – Bablok is a non-parametric rank-based alternative to Deming regression. A plot can be generated to display the regression line (Figure 2.8). A test of zero intercept and unit slope is provided by Passing and Bablok (1983, 1984) and Bablok et al. (1988). An assumption for Passing – Bablok regression is that the square of the slope in model (2.3) equals to the error variances ratio of the new method over the reference. Both Thiel – Sen and Passing – Bablok approaches are based on the slopes of the connecting lines between pairwise measurements. Linnet (1993) and Stockl et al. (1998) provide a comparison between Deming and Passing – Bablok regressions using both simulated and real data. Last, Passing – Bablok regression is popular mainly in Clinical Biochemistry, Pharmacology and

Laboratory Medicine when it comes to comparing two alternative measurement methods. Moreover, it is considered as an approach for method comparison studies in a guideline of the Clinical and Laboratory Standards Institute and suggested in a protocol (Jensen & Kjelgaard-Hansen, 2006) for clinical laboratories.

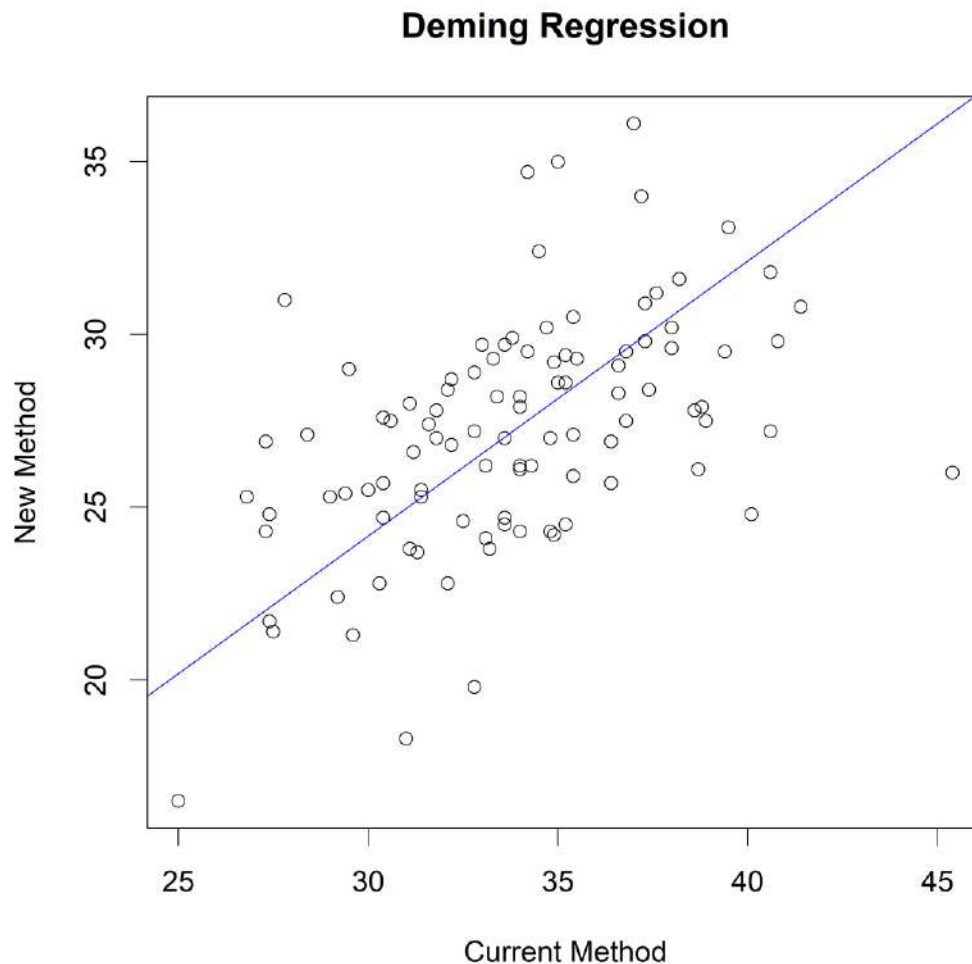
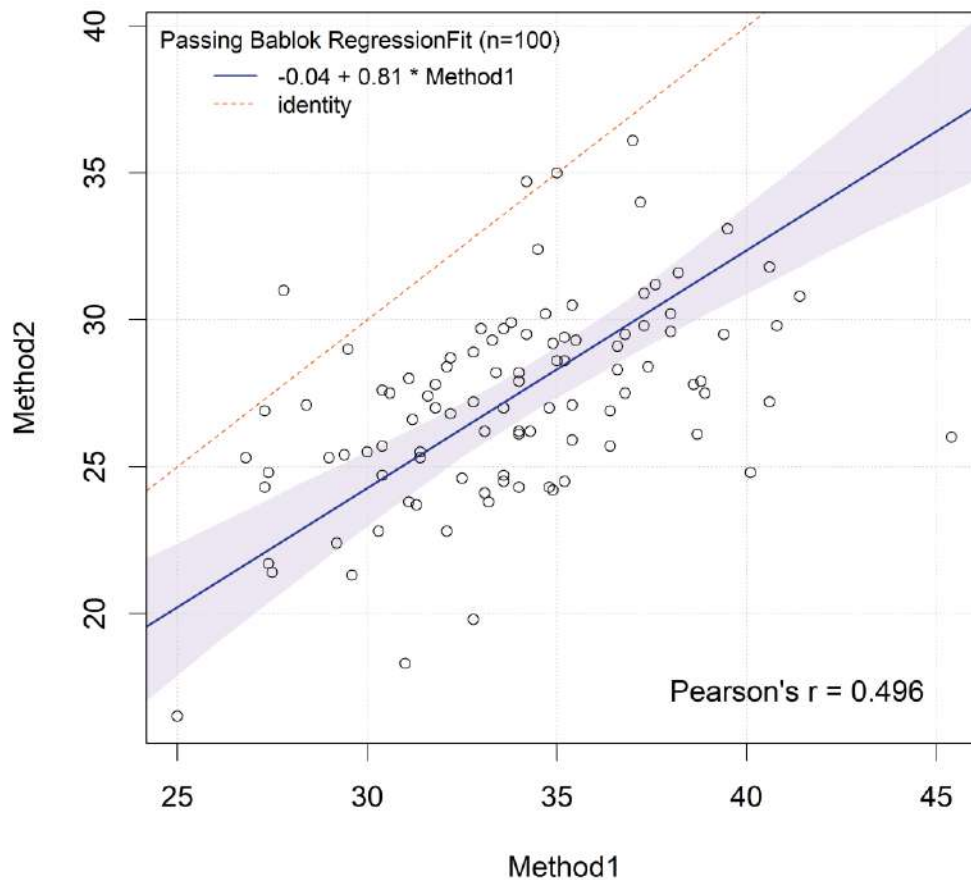


Figure 2.7. Deming Regression for unreplicated data using “mcr” R package (data from Chapter 4).

Passing Bablok Regression Fit



The 0.95-confidence bounds are calculated with the `bootstrap(quantile)` method

Figure 2.8. Passing-Bablok for unreplicated using "mcr" R package (data from Chapter 4).

2.1.5. Indices for assessing Agreement and Similarity

Many agreement indices can be found in the bibliography. Each index is used depending on the experimental design applied to the research and the conceptual intent of the study. One must cautiously choose the one appropriate to their setting. Indices can quantify the agreement and similarity between two or more devices. Their estimates can be retrieved either via their sample counterparts or by modeling the data and then using model components. When it comes to modeling the data, models like the measurement-error or mixed-effects are used to add covariates to the analysis, include repeated measurements and fix problems when assumptions are violated. Furthermore, the measurement-error model is an alternative to the simple linear model, where repeated measurements are included and problems with assumptions can be fixed. Researchers might choose it for its simplicity compared to mixed-effects models, depending on the nature of the agreement study. There are two categories of indices. The absolute (or unscaled) and the relative (or scaled) indices (Lin et al., 2012).

Absolute indices

The absolute indices report measures according to the magnitudes of the actual data. They are unscaled and independent of between-sample variation. For example, Pearson's correlation coefficient is not absolute since its range is always between -1 and 1. Usually these indices are defined as the expectation of a function of the difference or features of the distribution of the absolute difference. The absolute indices comprise of the mean squared deviation, the repeatability coefficient (described in Section 2.1.4), the reproducibility variance, the limits of agreement (described in Section 2.1.4), the coverage probability (CP) and the total deviation index (TDI). A detailed review about the indices can be found in Barnhart et al. (2007). Three of the absolute indices that were used for the case-studies in this thesis are described below.

i. Mean Squared Deviation (MSD)

In the case target values are random, the joint distribution of Y and X which are the paired measurements per device is assumed to have a bivariate distribution with means μ_Y and μ_X , variances σ_Y^2 and σ_X^2 respectively and covariance σ_{XY} . Mean squared deviation evaluates an aggregated deviation from the identity line.

$$MSD = E(Y - X)^2$$

$$\text{or } \varepsilon^2 = (\mu_Y - \mu_X)^2 + \sigma_Y^2 + \sigma_X^2 - 2\sigma_{YX}$$

Estimated by sample statistics with a log transformation, $W = \ln \varepsilon^2$ has an asymptotic normal

distribution with mean $\ln \varepsilon^2$ and variance $\sigma_w^2 = \frac{2}{n-2} \left[1 - \frac{(\mu_Y - \mu_X)^4}{\varepsilon^4} \right]$.

Since the MSD is not easily interpreted, two indices based on MSD put a meaningful interpretation on this basic index.

ii. Total Deviation Index (TDI)

TDI is an index that captures a predefined proportion (π_0) of data within a boundary (δ_0) from target values (Lin, 2000). Specifically, two measurement methods may be considered to have sufficient agreement if a large proportion of their differences is small. Thus, we define π_0 as the proportion of their differences and δ_0 as the sufficient difference. For example, if the researcher wants to investigate where do the differences lie in 90% of the population then TDI(0.9) is the appropriate index for this study. If the researcher sets a threshold (or clinically accepted difference of 2 then π_0 is 0.9 and TDI should be less than or equal to 2 ($\delta_0 = 2$) for sufficient agreement. TDI is similar in concept to a tolerance limit, where it bounds the range of values which is likely to contain a certain proportion of a population. The interpretation is focused on the deviations from the individual target values and not their own mean.

Several methods for estimation and inference about the TDI estimate have been proposed by (Lin, 2000; Lin et al., 2002; Lin et al., 2012) using the following method:

Assuming that $D=Y-X$ has a normal distribution with mean $\mu_D = \mu_Y - \mu_X$ and variance $\sigma_D^2 = \sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}$.

$TDI_{\pi_0} = \sqrt{(\chi^2)^{-1}\left(\pi_0, 1, \frac{\mu_D^2}{\sigma_D^2}\right)}$, where $\chi^2(\cdot)$ is the chi square distribution with 1 degree of freedom and non-centrality parameter $\frac{\mu_D^2}{\sigma_D^2}$. According to the authors, the estimate of this index has intractable asymptotic properties, thus the following approximation is suggested:

$$\delta_{\pi_0(\text{approx})} \doteq \Phi^{-1}\left(1 - \frac{1 - \pi_0}{2}\right) |\varepsilon|$$

Where $\varepsilon^2 = E(D^2) = (\mu_Y - \mu_X)^2 + \sigma_D^2$, is the mean square deviation. Furthermore, the approximation concludes satisfactory agreement when mean differences are small but can be conservative when the relative bias squared ($\Delta = \frac{\mu_D^2}{\sigma_D^2}$) is unreasonably large (Lin, 2000). Below, the combinations of π_0 and Δ for satisfactory agreement of the approximation are summarized.

1. $\pi_0 = 0.75$ and $\Delta \leq \frac{1}{2}$
2. $\pi_0 = 0.8$ and $\Delta \leq 8$
3. $\pi_0 = 0.85$ and $\Delta \leq 2$
4. $\pi_0 = 0.9$ and $\Delta \leq 1$
5. $\pi_0 = 0.95$ and $\Delta \leq \frac{1}{2}$

For inference on the approximation $\delta_{\pi_0(\text{approx})}$ or the natural logarithm of TDI_{π_0} , the author proposes the following hypothesis:

$$H_0 : \delta_{\pi_0(\text{approx})} \geq \delta_0, H_a : \delta_{\pi_0(\text{approx})} < \delta_0$$

For the approximation and similarly to the natural logarithm of TDI_{π_0} . Where δ_0 is the threshold (acceptable difference) set by the researcher.

Choudhary and Nagaraja (2007) show that Lin's (2000) approach is not consistent and may not retain its asymptotic nominal level. Thus, they propose an upper bound for the estimate of Lin's TDI for no repeated measurements from an exact test. Further, Choudhary (2008) and Choudhary and Nagaraja (2017a) extend the method based on the asymptotic distribution of the logarithm of Lin's TDI including repeated measurements by modelling the observed data via mixed-effects models and construct a relevant asymptotic tolerance interval for the distribution of appropriately defined differences. The method incorporates covariates, heteroscedasticity, and serial correlation in the errors. Rathnayake and

Choudhary (2017) propose a method for the case where the data are longitudinal. Further, Quiroz and Burdick (2009) proposed a method for inference about the TDI estimate when repeated measures are present, and they are paired over time. First, they fit the data using an ANOVA model. Then, they construct generalized confidence intervals about the TDI estimate based on replacing Lin's parameters (2000) with generalized pivotal quantities.

Escaramís et al. (2010a) propose to use the following model to estimate TDI:

$$y_{ijl} = x_{ijl}\delta + \beta_j + \alpha_i + \gamma_{ij} + e_{ijl} \quad (2.29)$$

- y_{ijl} is the l th measurement from subject i by device j .
- $i = 1, \dots, n$ indexes the subjects.
- $j = 1, 2$ denotes the methods.
- $l = 1, \dots, m$ indexes the repeated measurements.
- δ is the vector of fixed effects parameters common to both devices.
- x_{ijl} is the corresponding row of the design matrix for covariates.
- β_j is the device's fixed effect.
- α_i is the individual random effect assuming that $\alpha_i \sim N(0, \sigma_\alpha^2)$.
- γ_{ij} is the individual-method interaction random effect with $\gamma_{ij} \sim N(0, \sigma_\gamma^2)$.
- e_{ijl} is a random error assuming that $e_{ijl} \sim N(0, \sigma_e^2)$.

The author uses restricted maximum likelihood (REML) to estimate the model parameters. Lin (2000) defined TDI as a boundary k_p , which captures a large proportion p , of paired measurement differences from two methods within the boundary i.e., the value of k_p , that yields $P(|D| < k_p) = p$, where D is the paired differences variate. Under the assumptions of model (2.29), D is the paired differences based on any one of the replicates, $D = y_{ijl} - y_{ij'l'}$. Based on D , k_p is known as the total-TDI for evaluating total agreement according to (Lin et al., 2007). The distribution of D is $D \sim N(\mu_D, \sigma_D^2)$. The mean (μ_D) and standard deviation (σ_D^2) are estimated via the model parameters:

$$\mu_D = \beta_j - \beta_{j'}$$

$$\sigma_D^2 = 2(\sigma_Y^2 + \sigma_e^2) \text{ when the error variances per method are the same.}$$

$$\sigma_D^2 = 2(\sigma_Y^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2) \text{ when the error variances per method are different.}$$

TDI is estimated by finding the appropriate probability interval for the distribution of the paired-measurement difference variable, therefore, deriving its tolerance interval. TDI sets a boundary such that the proportion p percent of the differences in paired observations are within the boundary: $P(|D| < k_p) = p$ or $P(-k_p < D < k_p) = p$. Therefore, $[-k_p; k_p]$ determines the probability interval of D centered at 0, regardless of the mean value of D . Assuming that D follows a normal distribution with mean μ_D and standard deviation σ_D , k_p can be derived as follows:

$$k_p = \mu_D + z_{p_1} \sigma_D$$

$$-k_p = \mu_D + z_{p_2} \sigma_D,$$

where z_{p_1} and z_{p_2} are the p_s -th percentile ($s = 1, 2$) of the standard normal distribution such that $z_{p_2} = \frac{-2\mu_D}{\sigma_D - z_{p_1}}$ and $p_1 - p_2 = p$. One can find p_1 by using its link with p :

$$\Phi(z_{p_1}) - \Phi\left(\frac{-2\mu_D}{\sigma_D - z_{p_1}}\right) = p$$

Where $\Phi(\cdot)$ is the cumulative standard normal distribution. The parameter p_1 cannot be found in closed form, hence, the authors propose a recursive formula based on the binary search selection.

For unlinked and linked data, using models (2.10, 2.11) and following the approach described by Choudhary and Nagaraja (2017a) TDI can be calculated using the following formula:

$$TDI(p) = \sqrt{\left\{ \left(2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2 \right) \chi_{1,p}^2 \left(\frac{\beta_0^2}{\left\{ 2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2 \right\}} \right) \right\}}.$$

Last, for longitudinal data, Rathnayake and Choudhary (2017) proposed the following formula, using model (2.12) for methods 1 and 2 to estimate TDI:

$TDI_{12}(p, t) = 100p$ th percentile of the absolute value of the measurement differences $|Y_1(t) - Y_2(t)|$ is the following:

$$TDI_{12}(t, p) = \sqrt{\left\{ \left(2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2 \right) \chi_{1,p}^2 \left(\frac{(\mu_1(t) - \mu_2(t))^2}{\left\{ 2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2 \right\}} \right) \right\}}$$

Where p_0 is a specified large probability and χ_{1,p_0}^2 represents the $100p_0$ th percentile of a noncentral χ^2 distribution with one degree of freedom and noncentral parameter $\frac{(\mu_1(t) - \mu_2(t))^2}{2\psi^2 + \sigma_{e_1}^2 + \sigma_{e_2}^2}$.

This approach can be generalized for multiple methods. Denoting $TDI_{jr}(p_o, t)$ for method j and r . For inference, the authors used an upper band for the two methods. For the general case of more than two methods inference for TDI is applied using an upper band which is estimated for each pair of methods of interest. The methodology for the upper band is described in by Rathnayake and Choudhary (2017). Following this approach TDI is assumed to be estimated when the errors are homoscedastic. For heteroscedastic data we replace $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$ with $\sigma_{e_1}^2 g_1^2(u_i, \delta_1)$ and $\sigma_{e_2}^2 g_2^2(u_i, \delta_2)$.

Nonparametric methods for TDI estimation

Assumptions for parametric approaches for TDI estimation are the normality of the data and the linearity between response and effects. To deal with violations of these assumptions Perez-Jaume and Carrasco (2015) introduce a nonparametric methodology for estimation and inference for TDI. The authors compare their new method with already established. Lin et al. (2016) discuss nonparametric estimation of TDI using quantile regression. This method works satisfactorily for moderate samples ($n \geq 40$).

Inference for TDI

Escaramis et al. (2010a) proposed tolerance intervals. The value k_p is obtained by replacing μ_D and σ_D by their REML estimate counterparts derived from model (2.29) in expression $\hat{k}_p = \hat{\mu}_D + z_{p_1} \hat{\sigma}_D$. For inference, a one-sided tolerance interval is computed that covers the p_1 -percent of the population from D with a stated confidence.

Let T be the studentized variable of $\hat{\mu}_D + z_{p_1} \hat{\sigma}_D$. T follows a non-central Student-t distribution with non-centrality parameter $z_{p_1} \sqrt{N}$, $T \sim t_\nu(z_{p_1} \sqrt{N})$, where $N = 2 \cdot n \cdot m$ is the total possible paired-measurement differences between the two method/devices and the degrees of freedom ν , are derived from the residual degrees of freedom. For the case of using individual-device interaction or discarding it, Escaramis et al. (2010a) described different cases for obtaining ν . An upper bound for TDI estimate can be constructed by using the following $(1-a) \cdot 100\%$ one-sided tolerance interval, where a is the type I error rate:

$$UB(1-a) \cdot 100\% (k_p) = \mu_D + t_\nu(1-a, z_{p_1} \sqrt{N}) \frac{\sigma_D}{\sqrt{N}}$$

It corresponds to the exact one-sided tolerance interval for at least p_1 proportion of the population of defined by Hahn (1970) and Hahn and Meeker (1991).

To perform a hypothesis test if the interest is to ensure that at least p -percent of the absolute differences between paired measurements are less than a predefined constant k_0 , Lin's form of hypothesis can be followed.

$$H_0 : k_p \geq k_0, H_1 : k_p < k_0$$

H_0 would be rejected at level α if:

$$UB(1-a) \cdot 100\% (k_p) = \mu_D + t_\nu(1-a, z_{p_1} \sqrt{N}) \frac{\sigma_D}{\sqrt{N}} < k_0$$

Escaramis et al. (2010b) provided the R script to calculate the TDI estimate and its corresponding upper bound.

Choudhary and Nagaraja (2017a) use the large-sample theory of ML estimators to compute standard errors, confidence bounds and tolerance intervals (pp. 17, 18, 76, 77). When the sample is not large enough bootstrap confidence intervals are produced. The estimators of the model-based counterparts are obtained from models (2.10 – 2.12), depending on the nature of the data (unlinked, linked or longitudinal). To compute simultaneous confidence intervals and bounds, the percentiles of appropriate functions of multivariate normally distributed pivots are needed. The authors use the method proposed by Hothorn et al. (2008) using the R-package “multcomp”. Furthermore, they proposed a Bootstrap-t UCB and a modified nearly unbiased estimator (MNUT approximation) for computing the critical value, the p-value and the upper confidence bound (UCB) (Choudhary & Nagaraja, 2007).

iii. Coverage Probability (CP)

The coverage probability provides an equivalent approach to the Total Deviation Index. It is considered to give the inverse results. For a specified difference δ , the CP index is defined as

$$CP(\delta) = P(|D| \leq \delta) \text{ or } \pi_{\delta_0} = \chi^2\left(\frac{\delta_0^2}{\varepsilon^2}, 1\right)$$

CP is interpreted as the proportion of the population of differences (D) contained within $|\delta_0|$. In other words, it asks for more than 100p% of the differences distribution to be contained within δ_0 without the researcher to predefine π . Thus, to find CP, a δ needs to be predefined.

The choice whether to use the TDI or CP indices should be based on whether the researcher is interested in the boundary or the percentage of the differences for the population.

Relative Indices

Relative indices are scaled values for a predefined range, and usually lie between 0 and 1. In early studies, the intraclass correlation coefficient (ICC) was used to evaluate the agreement between quantitative measurements. There are numerous versions of the ICC and a review can be found in Barnhart et al. (2007). The concordance correlation coefficient (CCC) is considered the most popular index for assessing agreement between quantitative measurements. ICC and CCC are identical under ANOVA assumptions. However, ICC considers the methods/devices fixed and random, but the CCC treats them as only fixed. In this thesis CCC will be used instead of ICC as the latter can be evaluated under ANOVA assumptions while there is no need for this restriction for the former. Moreover, the statistical literature for CCC is rich and covers many case studies. Chen and Barnhart (2008) compare ICC and CCC for assessing agreement for data without and with replications. Last, the coefficient of individual agreement (CIA) is similar to CCC but is scaled relative to the minimum or acceptable agreement. Before considering the two methods/devices under investigation, their replication errors must be evaluated to check if they are acceptable. This procedure must be standard in every method comparison study because if the replication errors are significantly different then there is no point assessing agreement as the methods/devices will disagree regardless of any fixed biases. Then, interchangeability can be achieved if the difference between the measurements is relatively small and close to the difference of replicated measurements within the method/device.

i. Concordance Correlation Coefficient (CCC)

The concordance correlation coefficient was introduced by Lin (1989) and is defined as the standardized version of the Mean Square Deviation. It is an index that evaluates the agreement between two methods/devices and their corresponding paired measurements are Y_1 for the first method and Y_2 for the second. Both methods'/devices' measurements are attained with error.

Let Y_1 and Y_2 be two random variables with a joint normal distribution:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right)$$

The value of the squared difference of Y_1 and Y_2 can be used as an agreement value. However, this value ranges between 0 and infinity. This is difficult to interpret, hence, Lin (1989) suggested its standardization which ranges between 0 and 1.

$CCC = 1 - \frac{MSD}{MSD_0}$, where MSD_0 is the MSD assuming that Y_1 and Y_2 are independent. MSD is defined in section 2.1.5.

$$\text{Thus } CCC = \frac{2\rho\sigma_1\sigma_2}{(\mu_1 - \mu_2)^2 + \sigma_1^2 + \sigma_2^2}$$

For $CCC = 1$, there is perfect agreement,

For $CCC = 0$, there is no agreement,

For $CCC = -1$, there is perfect negative agreement.

Another useful form of the CCC is the following:

$$CCC = \chi_a \cdot \rho$$

where χ_a is the accuracy coefficient and ρ is the Pearson coefficient.

CCC - Accuracy Coefficient

Lin (1989) defined the accuracy coefficient which measures the closeness of the marginal distributions of Y and X . In other words, how far the line of best-fit deviated from the 45° identity line through the origin.

$$\chi_a = \frac{2}{\bar{\omega} + \frac{1}{\bar{\omega}} + v^2}, \text{ where } v = \frac{\mu_Y - \mu_X}{\sqrt{\sigma_Y \sigma_X}} \text{ and } \bar{\omega} = \frac{\sigma_Y}{\sigma_X} \text{ or } \frac{\sigma_X}{\sigma_Y}$$

v is defined as the location shift and $\bar{\omega}$ is defined as the scale shift.

$\chi_a = 1$ signifies equal means and variances.

$\chi_a = 0$ signifies that the absolute difference in means and/or variance approach infinity.

CCC - Precision Coefficient

This is the typical Pearson correlation coefficient:

$$\rho = \frac{\sigma_{YX}}{\sigma_Y \sigma_X}$$

It measures how far each observation deviates from the line of best fit.

To derive CCC, Lin uses the squared perpendicular distance of any paired observation to the identity line. This is defined as accuracy. When this distance is fixed, the precision, which measures how close the observation is from the regression line, is not fixed. Thus, regardless of the precision the accuracy is always fixed. Even if the precision is very low the accuracy will not be affected. This is a consequence of using one point. The authors instead use two points to minimize the area defined by the regression segment and the line of identity.

Although Lin's approach does not use mixed-effects models and is easy to estimate CCC, Carrasco and Jover (2003) reported, using simulations, that Lin's approach has a minimal bias in point estimation, as well as in standard error estimation. Bias grows in certain conditions compared to the estimation using variance components of a mixed-effects model. An improved Concordance Correlation Coefficient was proposed by Liao (2003). Moreover, Carrasco et al. (2013) calculate CCC using both U-statistics and variance components. U-statistics has been shown to be more robust than variance components in case the data are non-normal and skewed with low sample size. However, the U-statistic method presented in their paper cannot handle missing data or unbalanced designs, or multiple methods compared to the variance components method.

CCC for repeated values

In a method comparison study, taking multiple measurements for the same subject for many time points is quite common. Repeated measures, corresponding to observations on the same subject or experimental unit over time can be used to evaluate CCC. Chinchilli et al. (1996) used generalized multivariate analysis of variance to compute a weighted version of the CCC for repeated measurements. Carrasco et al. (2003) demonstrated the equivalence between CCC and ICC (intraclass correlation coefficient) by evaluating the CCC using a mixed-effects model accounting for repeated measurements which improves the estimation precision of the variance components. King et al. (2007a) introduced CCC for repeated measurements for two methods for longitudinal and non-longitudinal data. Carrasco et al. (2009) developed a CCC for longitudinal repeated measurements through the appropriate specification of ICC from a variance components linear mixed-effects model.

For unlinked data, under model (2.10) Choudhary and Nagaraja (2017a) proposed the following formula to calculate CCC:

$$CCC = \frac{2\sigma_b^2}{\beta_0^2 + 2(\sigma_b^2 + \psi^2) + \sigma_{e_1}^2 + \sigma_{e_2}^2} \quad (2.30)$$

While for linked data under model (2.11) :

$$CCC = \frac{2(\sigma_b^2 + \sigma_{b^*}^2)}{\beta_0^2 + 2(\sigma_b^2 + \psi^2 + \sigma_{b^*}^2) + \sigma_{e_1}^2 + \sigma_{e_2}^2} \quad (2.31)$$

Rathnayake and Choudhary (2017) proposed a method to estimate CCC for longitudinal data and multiple methods. The variance components model that accounts for many possible sources of variation (2.12) is used to estimate CCC between method 1 and 2 and is expressed using the following formula:

$$CCC_{12}(t) = \frac{2(\sigma_b^2 + \sigma_{b^*}^2)}{\{\mu_1(t) - \mu_2(t)\}^2 + 2(\sigma_b^2 + \psi^2 + \sigma_{b^*}^2) + \sigma_{e_1}^2 + \sigma_{e_2}^2} \quad (2.32)$$

The method can be generalized for multiple methods. Denoting $CCC_{jr}(t)$ for method j and r . For inference, the authors used a lower band for the two methods. For the general case of more than two methods inference for CCC is applied using an upper band which is estimated for each pair of methods of interest. The methodology for the lower band is described in (Rathnayake & Choudhary, 2017). CCC (2.30-2.32) is estimated when the errors are homoscedastic. For heteroscedastic errors $\sigma_{e_1}^2$ and $\sigma_{e_2}^2$ are replaced with $\sigma_{e_1}^2 g_1^2(u_i, \delta_1)$ and $\sigma_{e_2}^2 g_2^2(u_i, \delta_2)$. Model (2.12) can be extended for additional covariates. Choudhary and Nagaraja (2017a) use the same formula (2.56) but model the means using polynomials when considering time continuous unlike Rathnayake and Choudhary (2017) that use penalized-splines regression (discussed in Section 2.1.2.).

Parametric, semiparametric and nonparametric estimation and inference for CCC

For CCC estimation and inference, parametric, semiparametric and nonparametric approaches have been proposed. For parametric approaches, asymptotic distribution of the estimated CCCs can be used for inference if the data are modeled (Carrasco & Jover, 2003) for large sample sized. For estimating CCC using sample-based moments asymptotic distribution of the estimated CCC may be used for inference (Lin, 1989). For the semiparametric approach Barnhart and Williamson (2001) proposed modeling CCC via generalized estimating equations. For nonparametric approaches Choudhary (2010) proposed nonparametric confidence intervals treating agreement indices as a functional of the joint cumulative distribution function of the measurements and King et al. (2007a; 2007b) proposed inference for CCC using U-statistics for longitudinal paired data. Williamson et al. (2007) compared and proposed permutation and bootstrap test for testing equalities of CCCs in various conditions. For longitudinal data Oliveira et al. (2018) estimated CCC's components (LPC and LA) by replacing the model's coefficients and variance components by their REML estimated. A confidence interval can be constructed using a nonparametric bootstrap based on M (for example 5000) bootstrap samples with either the percentile method, or otherwise a normal approximation confidence interval. Choudhary and Nagaraja (2017a) use asymptotic distribution of the estimated CCCs to produce an upper confidence bound when the sample is large and bootstrap methods when the sample is small.

Rothery (1979) proposed a nonparametric measure of intraclass correlation which is based on the probability of specific types of concordances among the observations. This measure can be estimated from ranked data. Since the concordance correlation coefficient is related to the intraclass correlation coefficient inference methods for ICC can be used for CCC (Carrasco & Jover, 2003).

Hutson and Yu (2021) developed a robust permutation test for CCC. The test is based on appropriately studentized statistics. The asymptotical validation of the test was demonstrated across a range of distributional assumptions and sample sizes via simulations.

ii. Coefficient of Individual Agreement (CIA)

CIA assesses agreement between methods/devices where each method/device evaluates each subject at once. The interpretation of this index lies in the concept of deciding whether a single measurement made by method/device 1 can be replaced by a single measurement of method/device 2, when each subject is measured only once by each method/device. Barnhart et al. (2007) defined CIA in two cases. First, one of the two methods/devices is a reference. Second, there is no reference. This index is defined under ANOVA assumptions.

Comparison of the indices

Even though CCC is very informative and straightforward, Atkinson and Nevill (1997) pointed out, an increase in the between subject variability results in a larger value of CCC even if individual difference between measurements by the two methods remain the same. Furthermore, Barnhart et al. (2007) reported the dependency of CCC from between-subject variability based on its scaled nature which is relative to the maximum disagreement defined as the expected MSD under independence. There are two advantages of TDI compared to scaled indices. First, it does not depend on the measurement range. Thus, not affected by the between-subject variance. Second, it has a straightforward interpretation since it is reported at the same scale as the device's measurements under investigation. However, TDI depends on covariates explaining within-subject variation. In assessing agreement, the null and alternative hypothesis should be reversed. The conventional rejection region is the region of declaring one sided agreement. Thus, asymptotic power and sample size calculation should follow the above guideline. The powers of CCC, TDI and CP were compared in (Lin et al., 2002) and the results were similar for TDI and CP. However, TDI and CP estimates are superior to CCC but are only valid under the normality assumption. Last, Lin and Chinchilli (1997) suggest that measurements range should be compared first before comparing CCC estimates based on them. Last, Barnhart et al. (2007) compared CIA and CCC when there is no reference method/device. Both indices are affected by between- and within-subject variability. However, CIA is less

dependent than the CCC on the relative magnitude, $\frac{\sigma_B^2}{\sigma_W^2}$, of the between- and within subjects' variability.

Comparison of the methods producing confidence intervals, and confidence bounds

Choudhary and Nagaraja (2007) proposed an alternative test to Lin's TDI test and Wang and Hwang (2001) the nearly unbiased test (NUT). They demonstrate that TDI test may not be consistent as $n \rightarrow \infty$. Furthermore, TDI test and NUT are not consistent and may reach different conclusions for a common level of type I error. They suggest a modification for NUT called MNUT, which produces a simple closed form approximation numerically equivalent to the exact test for sample size $n \leq 30$. MNUT appears to be slightly more powerful than the equivalent NUT. Last, the authors provide an asymptotic level α parametric bootstrap-t test which has better properties than the asymptotic TDI test for moderate size sample sizes. An exact confidence interval is valid even for small sample sizes. Compared to the exact test, the bootstrap-t test's type-I error rates are closer to the nominal level than the rate of the exact test. After simulations performed by Choudhary and Nagaraja (2007), their proposed bootstrap-t test displays slightly increased power than the exact, when the sample size is moderately large. However, the

calculation of sample size for their proposal is complicated and involves a great amount of computational complexity. For CCC, if joint distributions of the observations under various conditions are valid (Williamson et al., 2007) a permutation test can be used, but this is an assumption that can be difficult to confirm. Thus, bootstrap-based tests may be preferable.

2.1.6. Assessing Similarity

The earliest paper assessing similarity measures is by (Grubbs, 1948) and focused on paired data. Mandel and Stiehler (1954) discuss precision and accuracy, arguing that a measurement method evaluation cannot be based only on these two. Fixed and proportional bias (β_0 and β_1), precision (2.15) and sensitivity (2.16) are considered measures of similarity. Following Choudhary and Nagaraja (2017a), to evaluate similarity the marginal distributions of Y_1 and Y_2 are examined via estimates and two-sided confidence intervals. Their distributions are given by equations (2.10 - 2.12), for unlinked, linked and longitudinal data. The fixed bias and the precision ratio are the two measures of similarity that will be evaluated using mixed-effects model. Last, proportional bias and precisions are evaluated under measurement-error models.

The fixed bias will be estimated via the model's counterparts. According to models (2.10 – 2.12) the fixed bias is estimated using: $\mu_1 - \mu_2$ for unlinked and linked data and $\mu_1(t) - \mu_2(t)$ for longitudinal data.

The precision ratio is evaluated in two different cases.

First, for models that ignore subject x method interactions:

$$\lambda = \frac{\sigma_{e1}^2}{\sigma_{e2}^2}$$

Second, for models that include subject x method interactions:

$$\lambda = \frac{\sigma_{e1}^2 + \psi^2}{\sigma_{e2}^2 + \psi^2}$$

For longitudinal data the precision ratio does not depend on t.

The precision ratios above are assumed to be estimated when the errors are homoscedastic. For heteroscedastic data we replace σ_{e1}^2 and σ_{e2}^2 with $\sigma_{e1}^2 g_1^2(u_i, \delta_1)$ and $\sigma_{e2}^2 g_2^2(u_i, \delta_2)$ thus fixed bias remains the same, but the precision ratio is given by:

$$\lambda = \frac{\sigma_{e1}^2 g_1^2(\tilde{u}_1, \delta_1)}{\sigma_{e2}^2 g_2^2(\tilde{u}_2, \delta_2)}$$

For inference, the method described by Nawaratha and Choudhary (2015) is used for heteroscedastic data. Specifically, if θ is a vector of model's counterparts then the measure of similarity is a function of θ . Denoting the measure of similarity as φ , and b^* a value in the measurement range then $\varphi(b^*)$ is any measure of similarity in a specific value (the measure is assumed to be scalar). Substituting θ with its corresponding ML estimate, $\hat{\theta}$, in its expression gives its ML estimator $\hat{\varphi}(b^*)$. From delta method

(Lehmann, 1998), when the sample size is large $\hat{\varphi}(b^*) \sim N_1(\varphi(b^*), G'(b^*)I^{-1}G(b^*))$, where $G(b^*) = (\frac{\partial}{\partial \theta})\varphi(b^*)|_{\theta=\hat{\theta}}$ can be computed numerically. Thus, approximate $100(1-a)\%$ two-sided pointwise confidence interval for $\varphi(b^*)$ on a grid of values of the measurement range can be computed as $\hat{\varphi}(b^*) \pm z_{1-a/2} \{G'(b^*)I^{-1}G(b^*)\}^{1/2}$.

2.1.7. Assessing Repeatability

Repeatability evaluation is essential and can be used to identify possible sources of disagreement. It is considered as intra-method agreement and assists the agreement study. This is possible in cases where a method/device has low intra-method agreement. Thus, if it does not agree with itself, it will certainly not agree with another method/device. For unlinked data, repeated measurements are replications of the same underlying measurement. Instead of using the bivariate distributions (Y_1, Y_2) for measurements of the two methods on a randomly selected subject from a population, Y_j^* is defined as a replication of Y_j , where $j=1,2$ denote the two methods/devices. By definition Y_j and Y_j^* have the same distribution. CCC and TDI are modified and are calculated. By dropping the subscripts for model (2.10) for unlinked data:

$$Y_1^* = b + b_1 + e_1^*, Y_2^* = b + b_2 + e_2^*$$

is induced, similar to (2.10) by dropping the subscripts.

Then for method 1:

$$\begin{pmatrix} Y_1 \\ Y_1^* \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{e_1}^2 & \sigma_b^2 + \psi^2 \\ \sigma_b^2 + \psi^2 & \sigma_b^2 + \psi^2 + \sigma_{e_1}^2 \end{pmatrix} \right) \quad (2.33)$$

For method 2:

$$\begin{pmatrix} Y_2 \\ Y_2^* \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \beta_0 + \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{e_2}^2 & \sigma_b^2 + \psi^2 \\ \sigma_b^2 + \psi^2 & \sigma_b^2 + \psi^2 + \sigma_{e_2}^2 \end{pmatrix} \right), \quad (2.34)$$

where e_1^* and e_2^* are independent copies of e_1 and e_2 as defined in (2.10).

Defining $D_j = Y_j - Y_j^*$ as the difference in two replications of method j. From (2.33) and (2.34):

$$D_j \sim N_1(0, 2\sigma_{e_j}^2), \quad j=1,2. \quad (2.35)$$

Thus,

$$CCC_j = \frac{\sigma_b^2 + \psi^2}{\sigma_b^2 + \psi^2 + \sigma_{e_j}^2}$$

$$TDI_j = \left\{ 2\sigma_{e_j}^2 \chi_{1,p}^2(0) \right\}^{\frac{1}{2}} \text{ for } j=1,2.$$

By dropping the subscripts for model (2.11) for linked data:

$$Y_1^* = b + b_1 + b^{**} + e_1^*, Y_2^* = b + b_2 + b^{**} + e_2^*$$

For method 1:

$$\begin{pmatrix} Y_1 \\ Y_1^* \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_b \\ \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 & \sigma_b^2 + \psi^2 \\ \sigma_b^2 + \psi^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_1}^2 \end{pmatrix} \right) \quad (2.36)$$

And for method 2:

$$\begin{pmatrix} Y_2 \\ Y_2^* \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \beta_0 + \mu_b \\ \beta_0 + \mu_b \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 & \sigma_b^2 + \psi^2 \\ \sigma_b^2 + \psi^2 & \sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_2}^2 \end{pmatrix} \right) \quad (2.37)$$

Where b^{**} , e_1^* and e_2^* are independent copies of b^* , e_1 and e_2 as defined in (2.11).

Defining $D_j = Y_j - Y_j^*$ as the difference in two replications of method j. From (2.36) and (2.37):

$$D_j \sim N_1(0, 2(\sigma_{b^*}^2 + \sigma_{e_j}^2)), \quad j=1,2.$$

Thus,

$$CCC_j = \frac{\sigma_b^2 + \psi^2}{\sigma_b^2 + \psi^2 + \sigma_{b^*}^2 + \sigma_{e_j}^2}$$

$$TDI_j = \left\{ 2(\sigma_{e_j}^2 + \sigma_{b^*}^2) \chi_{1,p}^2(0) \right\}^{\frac{1}{2}} \text{ for } j=1,2.$$

2.1.8. R code to assess Agreement and Similarity

There are many R packages that evaluate indices like TDI, CCC and limits of agreement. The “SimplyAgree” package (Caldwell, 2022) offers calculation of LOA, CCC and others. Agreement tests can be performed constituting a hypothesis if the agreement between two measurements is within a certain limit using Shieh exact test for agreement (Shieh, 2020). Then the limits of agreement, with confidence limits, are included. For repeated measurements, where they expected to vary or not with balanced or unbalanced number of repeated measurements per subject. A function to assess the assumptions for normality, heteroscedasticity and proportional bias is provided. The estimates to calculate the LOA can be retrieved either via sample statistics or by retrieving mixed-effects models counterparts. Last, power analysis can be performed based on Lu et al. (2016). For LOA the package offers a nonparametric method for estimation based on Bland and Altman (1999), including a hypothesis test based on binomial proportions within the clinical allowable difference.

The package “cccrm” (Carrasco & Martinez, 2022) offers the calculation of CCC with repeated measurements and CCC for longitudinal data, including their confidence intervals, for data using two different methods. The U-statistic and the variance component methods. The calculations are based on Carrasco et al. (2009, 2013), Hoef (2012), King et al. (2007a), King and Chinchilli (2001) and Pinheiro et al. (2022).

The package “lcc” (Oliveira et al., 2020) provides the calculation of the Longitudinal concordance correlation estimated by fixed effects and variance components using polynomial mixed effects regression model. It also calculates the longitudinal Pearson correlation and longitudinal accuracy. These statistics can be estimated using different structures for the variance-covariance matrix for random effects and variance functions to model the within-group errors using or not the time as a covariate and is based on Oliveira et al. (2018).

R-code scripts are available by Choudhary and Nagaraja (2017b) that apply the methodology described in their book. Furthermore, R-code is available for computing the critical value and the UCB (exact approach) for TDI, the MNUT approximation for computing the critical value, the p-value, the UCB and LCB for TDI and the Bootstrap-t UCB for TDI.

Taffé et al. (2019) developed the “MethodCompare” R-package that evaluates fixed and proportional bias, precision and uses recalibration methods to correct the agreement between two methods. Moreover, the package can create a bias, precision, compare plot and the extended Bland-Altman plot with limits of agreement. The model used to estimate the parameters is the measurement-error model (2.25 – 2.26). The code of the package was adopted for the case-studies and was used to create some of the plots for Chapters 4, 5 and 6.

Kuiper et al. (2022) developed the R-package called “nopacko” for a non-parametric test for multi-observer concordance and differences between concordances in balanced and unbalanced data.

For the Bland-Altman plot with limits of agreement and their corresponding confidence intervals “blandr” (Datta, 2018) package can be used.

The package “mcr” (Potapov et al., 2023) can be used to produce graphs for Passing-Bablok and Deming Regression. Generally, the “mcr” can be used to perform regression methods and to quantify the relation between two measurement methods with or without repeated measurements.

2.1.9. Examples using Agreement and Similarity techniques.

Bongers et al. (2018) evaluated the validity, reliability, and inertia of four different temperature capsule systems using a custom-made water bath. The sample size was 40 (10 capsules per system) and the agreement between the capsules and the water bath measurements was reported using LOA and a modified Bland and Altman plot. The plot summarized each system readings using boxplots with mean difference between capsule and water bath temperature for the capsule systems \pm the LOA.

Furthermore, Laerhoven et al. (2022) validate an open-source ambulatory assessment system using Bland-Altman plot and reports the experimental design of the experiment in detail.

Igic et al. (2010) compared a micrometer-and a scanning electron microscope-based measurements of avian eggshell thickness. The authors assessed the instrument measurements agreement using the TDI index. They first modeled the data using a linear mixed effects model and then constructed the corresponding asymptotic tolerance interval for the distribution of appropriately defined differences.

Perez-Jaume and Carrasco (2015) compared two different questionnaires for alcohol intake assessment. The authors used a non-parametric method to calculate TDI. They use the same method to evaluate the agreement between the measurements taken from the left and the right side of the jaw. The sample size

was small (10 individuals). Since the measurements of interest were not normally distributed nor even continuous (count data), the authors used a non-parametric method to evaluate agreement.

2.2. Experimental Design and Statistical Analysis

The importance of proper designs in experiments is of utmost importance for valid results and adequate reproducibility. Many papers provide scanty information to assess whether the experimental design and statistical analysis are appropriate. Repeated and replicate measurements are both multiple response measurements taken at the same combination of factor settings. However, repeated measurements are taken during the same experimental run or consecutive runs, while replicate measurements are taken during identical but different experimental runs, which are often randomized. Their differences affect the structure of the dataset and the statistical analysis applied to process the data. Experiments may have both repeated and replicate measurements depending on the sources of variability the researcher wishes to investigate. Usually, replicates are more expensive than repeated measurements since they are taken from different subjects. Since resources are limited, it is vital to get the most information from an experiment. Significantly more information can be retrieved from well-designed experiments and often require fewer runs than poorly designed experiments. In many situations, researchers mistakenly take for granted the sample's independence even though they sample from the same subject. This occurs when the experimental unit is not defined properly and instead of replicates the researchers provide repeated measurements. Moreover, measurements might be taken from the same subject repeatedly through time while the characteristic measured changes. These measurements, coming from the same subject, are clustered and proper statistical modeling should be applied to consider the measurement's dependency. These data are defined as longitudinal and clustered data. The primary goal of a longitudinal study is to distinguish the change in response over time and the factors that affect change (Fitzmaurice et al., 2011). This theory can be applied in a method or measurement comparison study since many experiments require to capture the within-individual changes. Chapter 3 applies this design since repeated measurements are taken in different occasions through time.

A well-designed experiment must include seven important characteristics (Lazic, 2016).

- Effects can be estimated unambiguously and without bias.
- Estimates are precise.
- Estimates are protected from possible one-off events that might compromise the results.
- The experiment is easy to conduct.
- The data is easy to analyze and interpret.
- Maximum information is obtained for fixed time, resources, and samples.
- The findings are applicable to a wide variety of subjects, conditions, and situations.

The key steps to perform a method comparison study are discussed by Hanneman (2008) and Choudhary and Nagaraja (2017a):

Before the experiment:

- The two instruments under investigation must measure the exact same quantity.
- The sample must be taken simultaneously by both methods/devices unless the subject's characteristic being measured by the devices will not change instantly. For example, the soil moisture is possibly stable for an hour without water been added. Thus, the researcher can obtain separate samples from the two devices within the allowable time interval. This is a study with

sequential measurements. Randomizing the order of measurement so that any real time differences would be spread across the two methods of measurement is considered a good practice. In case of wind speed measurements, where we consider wind speed changing immediately, simultaneous measurements must be taken.

- The sample size is important. Sufficient sample size of paired measurements can reduce chance findings. It is important to perform a power analysis before any experiment. However, in many studies this is not the case. For agreement studies, (Lu et al., 2016) propose a method to calculate sample size for Bland-Altman method. To estimate the power:

$$power = 1 - \beta = 1 - (\beta_1 + \beta_2) = probt(t_{1-\frac{\alpha}{2}, n-1}, n-1, \tau_1) + probt(t_{1-\frac{\alpha}{2}, n-1}, n-1, \tau_2)$$

- $probt[\bullet, n-1, \tau_1]$ is the cumulative distribution function of a Student's non central t-distribution with n-1 degrees of freedom and non-centrality parameter τ_1 .
- β_1 is the type II error for the upper limit of LOA.
- β_2 is the type II error for the lower limit of LOA.

$$\tau_1 = \frac{\delta - \bar{D} - z_{1-\frac{\gamma}{2}} s_D}{s_D \sqrt{\frac{1}{n} + \frac{z_{1-\frac{\gamma}{2}}^2}{2(n-1)}}} \text{ and } \tau_2 = \frac{\delta + \bar{D} - z_{1-\frac{\gamma}{2}} s_D}{s_D \sqrt{\frac{1}{n} + \frac{z_{1-\frac{\gamma}{2}}^2}{2(n-1)}}} \text{ are the non-centrality parameters.}$$

When $\mu = 0$ the sample size formula can be written as follows:

$$n = \frac{\left(2 + z_{1-\frac{\gamma}{2}}^2\right) \left[\text{tinv}\left(1 - \frac{\beta_1 + \beta_2}{2}, n-1, t_{1-\frac{\alpha}{2}, n-1}\right) \right]^2 s_D^2}{2 \left(z_{1-\frac{\gamma}{2}} s_D - \delta \right)^2}$$

and $(-\delta, \delta)$ are the CAD. A R-package called "blandPower" (Wisniewski, 2023) illustrates the method.

- The measurement range should be as wide as possible. Thus, the researcher can evaluate the possible intervals where the methods show acceptable agreement. The methods/devices then can be used interchangeably only for intervals that meet the predefined CAD.

After the Experiment (Analysis):

- The available data should be explored by using exploratory analysis either graphically or using tables with measures of central tendencies and dispersion. A Bland-Altman plot should be used to assess the data for heteroscedasticity, dependency of the difference from the measurement range, and outliers. Moreover, a scatterplot should be used as a supplementary plot to investigate the relationship between the two methods. Last, a trellis plot is useful to visualize the spread of the repeated values and possible biases regarding the two methods. A trellis plot (Choudhary & Nagaraja, 2017a) is constructed by using the x-axis as the measurement range and the y axis as

the subjects' id. The two methods are differentiated by using two different symbols for each measurement per subject. Chapters 4, 5 and 6 include Trellis plots to describe the data.

- Instead of using statistics of the sample, the data should be modeled. Modeling the data provides more flexibility in assessing various indices and sources of disagreement by using model-specific parameters. Diagnostics should be performed, and the researcher should use goodness of fit techniques.
- Confidence intervals should be used to assess measures of similarity.
- Confidence bounds should be used to assess agreement indices.
- In case of significant disagreement, identify the sources of disagreement and investigate corrective actions.

2.2.1. Inappropriate methods to assess Agreement.

The most prevalent methods for comparing two or more instruments, methods, operators, assays or formulas to measure a quantity from a unit (a person, an animal, a plant etc) is the correlation coefficient, the Mean Square Error, ANOVA, paired t-test and others. Numerous articles based on this topic (Byra-Cook et al., 1990; Earp & Finlayson, 1991; Lunney et al., 2018) use these methods. However, these methods are not testing agreement. Jacobsen et al. (1991) reports that the use of only correlation coefficients and paired t-tests to assess agreement comes with the risk of making a type III error since the researchers are finding the right answer to the wrong question.

Using the paired t-test to assess Agreement.

Traditionally, researchers tested the agreement between two measurements of two devices using a paired t-test for the paired measurements (Y_1, Y_2) . It is equivalent to the one-sample t-test of zero mean applied to the paired differences D_i . According to Choudhary and Nagaraja (2017a) when computing the

T statistic $T = \sqrt{n} \frac{\bar{D}}{s_D}$, where \bar{D} is the sample mean of the differences of the pairs of observations and

s_D is the sample standard deviation of the differences, and set up the null and its alternative hypothesis as:

$$H_0 : \mu = 0 \quad \text{and} \quad H_1 : \mu \neq 0$$

T follows the t distribution with $n - 1$ degrees of freedom under null hypothesis. In addition to bivariate normality, if the data also follow the mixed-effects model, in case the null hypothesis is not rejected, it may be because $\frac{s_D^2}{n}$ is large, either due to large variability in the differences or due to small sample size.

This implies that the test may have low power.

Furthermore, under the mixed-effects model assumptions, by testing $\mu = 0$, is equivalent to testing for $\beta_0 = 0$. However, in the presence of unequal proportional bias, this equivalency fails. The same applies when a measurement error model is used to fit the data. The mean is in the following form:

$\mu = b_0 + (b_1 - 1)\mu_b$. Testing for $\mu = 0$ is not equivalent to testing for the fixed bias equality to zero.

Last, according to Lin et al. (2012) the rejection region, alternative hypothesis, is the one based on strong evidence in the data. Failing to reject will not prove agreement. Reversing the alternative and the null hypothesis will grant the researcher with the region that provides strong data driven evidence for declaration of agreement.

Using the Linear least squares estimates (Least Squares Regression Analysis) to assess Agreement.

Calculating the estimates of the slope and intercept for the ordinary linear regression is also one of the common practices to measure agreement. According to Ludbrook (1997) “Investigators sometimes use least squares (Model I) regression analysis to calibrate one method of measurement against another. In this technique, the sum of squares of the vertical deviations of y values from the line is minimized.” This is not a valid approach. Even though both y and x values are attended by random error the linear least square regression assumes that only the dependent variable is measured with error (Cornbleet & Gochman, 1979).

Using the Correlation Coefficient to assess Agreement.

It is very common, for many reports that were published in biomedical and other journals (Kirkeby, 1995; Tschakovsky, 1995), to consider the correlation coefficient as a method of assessing agreement between two methods that measure the same characteristic. The correlation coefficient is falsely used for agreement studies, as Lin (1989) suggests. It is defined as the quantification of the linear association between two continuous variables. It is inadequate to assess agreement in various instances, because it evaluates only the association of two sets of observations (Altman & Bland, 1983). Moreover, it is susceptible to the range of the data measured. Studies that compare assays that measure the same variable using a wide range will always report a higher correlation coefficient compared to studies that compare the exact same assays but for a narrower range of measurements. The coefficient can easily be inflated by adding higher and lower values of measurements. Researchers should not compare the correlation coefficients in different studies, before assessing the range of observation values. Even if the new methods measurements deviate fairly from the reference measurements, but are related via a specific transformation, the correlation coefficient will report a high value.

Specifically, using the measurement-error model as an example, Choudhary and Nagaraja (2017a) report that the correlation depends on the between-subject variation in the population being measured (p. 10). Using the measurement-error model's-based moments (2.4), (2.5) the correlation coefficient is defined as:

$$\rho = \frac{\beta_1 \sigma_b^2}{(\sigma_b^2 + \sigma_{e_1}^2)^{\frac{1}{2}} (\beta_1 \sigma_b^2 + \sigma_{e_2}^2)^{\frac{1}{2}}}$$

The correlation can be interpreted as the square root of the product of the reliabilities of method 1 and 2, given by

$$\frac{\sigma_b^2}{\sigma_b^2 + \sigma_{e_1}^2} \cdot \frac{\sigma_b^2}{\beta_1 \sigma_b^2 + \sigma_{e_2}^2}$$

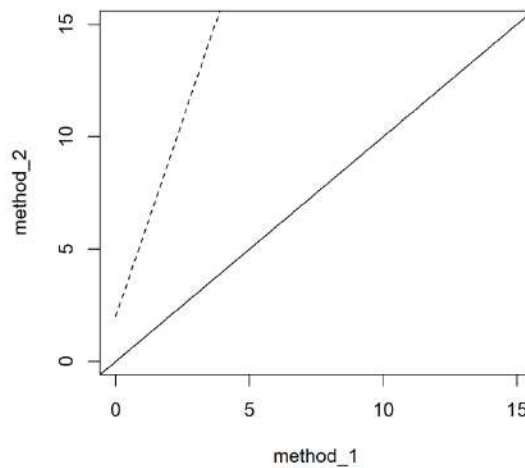


Figure 2.9. The scatterplot of two sets of measurements with a fixed bias of 2 and a proportional bias of 2.5. The dashed line shows the set of measurements, and the solid line shows the line of equity.

Thus, the correlation also depends on the between-subject variation in the population being measured. If the methods are used in two populations, one with greater heterogeneity than the other, the methods would exhibit higher correlation in the population with greater heterogeneity.

Ludbrook (1997) states that “Investigators often use the Pearson product-moment correlation coefficient to compare methods of measurement. This cannot detect systematic biases, only random error.” As Janse et al. (2021) and Liehr et al. (1995) report, the correlation coefficient will assess how well the observations will fit on any straight line, regardless of the deviation from the line of equity. For this reason, a scatterplot should always be used to assess the relation of the two variables in case no linear association is formed. An example of the correlation coefficient inadequacy of measuring agreement is a comparison of two methods when there is a constant deviation of n units (fixed bias). Even though the two methods will never report perfect agreement the correlation coefficient might be close to 1. Figure 2.9 displays the scatterplot of two sets of measurements with a fixed bias of 2 and a proportional bias of 2.5. The dashed line shows the set of measurements, and the solid line shows the line of equity. The correlation coefficient value is 1. Finally, simple correlation cannot cope with replicated data as Altman and Bland (1983) report.

Examples using misleading methods to assess agreement.

Most of the open-source devices developed for agricultural sciences are validated against an industrial. Unfortunately, a vast number of articles follow inappropriate methods or limit their research into descriptive measures that certainly don't cover the goals of method comparison studies. First, statistical inference tests are used, regardless of their assumptions been violated. As an example, Mesas-Carrascosa et al. (2015) are using ANOVA and t-test assuming independent measurements. This is not valid since the measurements were taken from the same device in 5-minute intervals. Another violation of the same assumption can be found in Barnard et al. (2014). The authors are performing linear regression on the instantaneous measurements from the two systems. The assumption of independent subjects is violated since the measurements are taken from the same devices over time. Second, Mesas-

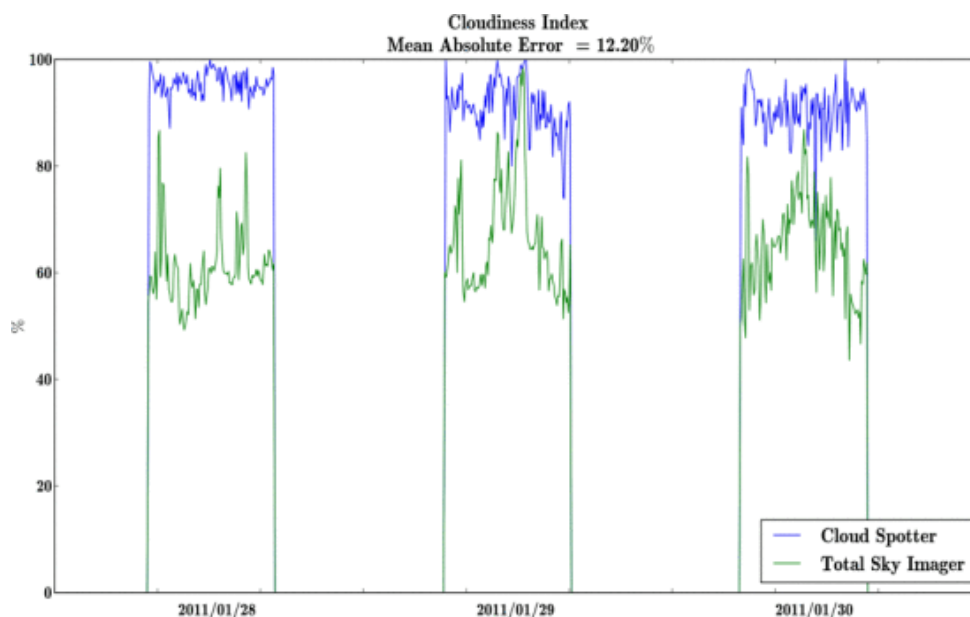


Figure 2.10. Lineplots. The x-axis presents the time, while on the y-axis a percentage of cloudiness index is presented. To differentiate between the two devices the authors use green color for the reference device and blue for their prototype. The disagreement intervals are not obvious neither the sources of bias (fixed or proportional). (source: Laerhoven et al. (2022))

Carrascosa et al. (2015), are not clearly reporting the experimental design since no experimental unit is defined.

Third, in many articles, indices like RMSE (Root Mean Square Error) or MAE (Mean Absolute Error) are used to assess the agreement between two devices. RMSE is just a measure of deviation between paired data sets and is considered only a descriptive measurement of the data. Thus, it does not report possible sources of disagreement, or ranges of disagreement. RMSE is not a formal test and the results provided by this index will not allow any researcher to make any inference. Last, many authors use inadequate graphs to report agreement. Laerhoven et al. (2022) use a lineplot (Figure 2.10). The x-axis presents time, while on the y-axis a percentage of cloudiness index is presented. To differentiate between the two devices the authors use green color for the reference device and blue for their prototype. The disagreement intervals are not obvious neither the sources of bias (fixed or proportional). The same method can be found in Mesas-Carrascosa et al. (2015). The authors compare the proposed and a reference device, using ten 5-minute intervals of paired measurements. The results cannot be generalized for all the devices that use the specific sensors since the analysis lacks statistical inference. In this case there is a need for more experimental units to diminish any lurking variables that affect the agreement. No sources of disagreement can be reported using this method.

To overcome all these challenges, this thesis focuses on the experimental design of every experiment presented, correctly defining the experimental unit, fixed or random factors that must be modeled correctly and possible covariates and lurking variables that might affect the analysis.

2.3. Designing and Using the Hardware

For the past 10 years, the dependency on smart devices for many tasks is evident. The flourishing of development boards, especially the ones that follow the open-source mentality, ignited a chain reaction

on the production of multiple microcontrollers and microprocessors. Intel and Texas Instruments were the pioneers in the field. The emergence of the microprocessor and the microcontroller created a new era.

A microprocessor, also called an integrated circuit or chip, is a system that has three different functions. First, an input, then a process to run and last, an output. These inputs and outputs are a series of voltages that are used to command external devices. The microprocessor needs input, then according to prior instructions, outputs are provided.

When the microprocessor is inserted into a circuit then we have a microprocessor-based system. The microprocessor will do nothing by itself unless it is embedded in a circuit.

A microcontroller is a complete microprocessor-based control system built into a single chip. It contains all the vital parts into a single device. The difference between the two is the general use. The microprocessor is a system for general use while the microcontroller is designed for a single task or purpose. For example, microcontrollers can be found inside a pocket calculator, a camera, a washing machine, or a car's electronic systems and are built for executing specific tasks. On the other hand, microprocessors can be found inside computers and execute general purpose tasks like word processing, video imaging, computer games etc. The microcontroller lacks processing power while the microprocessor's processing power has increased exponentially over the past decades. However, great processing power means more power consumption. Thus, microcontrollers are efficient for applications that are battery based but not limited to them. Washing machines do not need a 4-core microprocessor since they only execute a few simple tasks. For this reason, engineers built them with microcontrollers making the design simple and cost-effective.

The descriptions of the parts used to develop the devices of this thesis are presented from a maker's perspective. The technical details useful for the reader to understand this thesis's case-studies are mentioned and analyzed without discussing advanced technical information since it is beyond the scope of this thesis. However, all the references are provided in case the reader wishes to delve deeper into more advanced topics.

2.3.1. Microcontrollers

Microcontroller (MCU) designers focus on supporting peripherals used to carry out various tasks such as climate control, DC motors, position or motion control, environmental monitoring, mobile phones, cameras, appliances, security systems etc. Due to the integration of all functions on a single chip, their sizes and power consumption is significantly smaller compared to microprocessors. Maintenance and troubleshooting are another advantage of using MCUs instead of other more complex systems. However, depending on the application the researcher/individual must choose from a vast variety of microcontrollers. The introduction of EEPROM granted MCUs the ability to be electrically erased. Thus, allowing for swift prototyping and In-System-Programming (ISP). In 1993 flash memory was introduced by Atmel. There are many types of MCUs such as PIC, ARM, 8051, AVR, Intel, MSP etc. Each of them has their own characteristics. AVR is the most common MCU used by companies like Arduino, Adafruit and others. The AVR MCUs consist of a processor core, a flash memory, and RAM. It also includes additional parts, like input and output ports (I/O), timers, serial communication ports, analog to digital converters (ADC) and a USB port. The ATMEGA32u4 (Atmel, 2016a) MCU (figure 2.11a) is used in most of the case-studies of this thesis since it comes with rich documentation. It is a low power CMOS 8-bit MCU with many interfaces

and communication protocols. The ATmega 32u4 power supply ranges from 2.7 to 5.5 Volts. The MCU's speed can be either 8 or 16 MHz. Figure 2.11b. displays the 44 pins of the MCU. The pin mapping table is available at (Arduino, 2023d). ATmega 32u4 is equipped with a 10-bit analog to digital converter (ADC), which means that the MCU can read voltage signal in increments of 4.8mV if the reference voltage is 5 Volts $\left(\frac{5}{2^{10}}\right)$ or 3.2mV if the reference voltage is 3.3 Volts $\left(\frac{3.3}{2^{10}}\right)$. The input voltage increments for the 5 Volts reference are displayed in Figure 2.12. The ATmega 32u4 MCU provides 14 digital and 6 analog pins. These pins work as inputs or outputs. ATmega 32u4 supports four important communication protocols. The 1-Wire protocol, the Serial Peripheral Interface (SPI), the Universal Asynchronous Receiver-Transmitter (UART) and the Inter-Integrated Circuit (I2C) protocol. These protocols are used to communicate with peripherals such as SD cards, Real Time Clocks, or sensors. The protocols are described in Section 2.3.4.

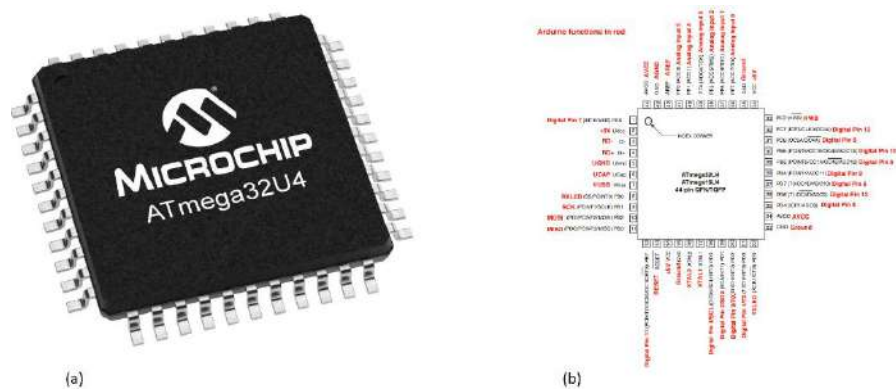


Figure 2.11. (a) Atmega 32u4 on the left (source: Microchip Technology Inc). (b) Pinout ATmega 32u4 on the right (source: Arduino documentation).

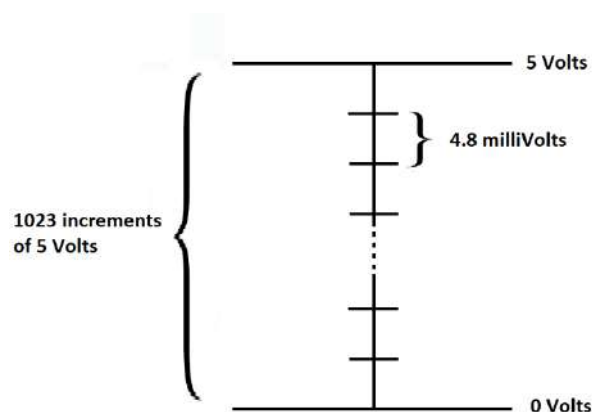


Figure 2.12. The 10-bit resolution of the Analog to Digital converter. The 5 Volts are divided into 1023 increments of 4.8 milli Volts.

The ESP 32 (ESPRESSIF, 2023) (Figure 2.13) series is a MCU with embedded Wi-Fi and Bluetooth. It is designed to achieve the best RF performance and low power consumption. It is designed for mobile, wearable electronics and Internet-of-Things (IOT) applications. It integrates an internal antenna switch, and a variation of the chip has a surface-mount uFL connector for external antennas. The MCU's characteristics can be found in Appendix A (Table A2). ESP 32 offers 10 capacitive touch GPIO, 16 Pulse Width Modulation Signals (PWM) output channels, 18 ADC input channels with 12-bit resolution, 2 digital to analog converter (DAC) pins, RTC GPIOs (General Pin Input Output) routed to the RTC low-power subsystem, which can be used when the ESP 32 is in deep sleep, 3 SPI interfaces, 2 I2C interfaces and 3 UART interfaces for communication and other features which are describe in the datasheet.

The ESP 32 MCU is implemented in Chapter 3.

2.3.2. Development Boards

Microcontrollers are used to build development boards like Arduino, Adafruit (2023), ARM9 etc. The most famous Arduino's development board is the Arduino Uno (Figure 2.14a). It is based on the ATmega328P microcontroller, and it is a user-friendly board.

Adafruit developed the feather proto 32u4 (Fried, 2023a) which is based on the ATmega32u4 microcontroller. It is very popular since it has a small footprint and other "shields" can be stacked on top of it and perform more complex tasks. Shields are additional boards that usually will not include a microcontroller rather than some electronics that perform very specific tasks. For example, the adalogger feather wing (Fried, 2023c) is a real time clock (RTC) combined with an SD card reader. When it is stacked on top of the feather proto 32u4 (Figure 2.14b), the maker can construct a powerful logger without the need of soldering and wiring. These are open-source electronic platforms. They are easy to work with, since they are equipped with pins and buttons that allow the user to connect electronics, sensors and other modules that execute different tasks. Arduino and Adafruit, host countless tutorials on their sites and other independent documentations and videos can be found in the web (Instructables, n.d.; Fried, 2023d; Huth, 2023; Mechatronics, 2023).



Figure 2.13. (a) ESP-32 WROOM-32E, with Internal antenna, on the left (source: espressif, esp32 documentation) . (b) ESP32-WROOM-32UE, with surface-mount uFL connector for external antennas, on the right (source: espressif, esp32 documentation)

Specifically, modestmaker (Huth, 2023) focuses on environmental monitoring and describes step-by-step the DIY process to create loggers that store air temperature and humidity, water level and other environmental characteristics remotely, in open-source servers. The author of the channel is a hydrologist and describes the usefulness of open-source devices for his work.

Arduino's popularity increased at high rates and for the past 10 years, the company created cross-curricular content for almost all the educational levels (Arduino, 2023c). These tools are essential for STEAM (Yakman, 2008) classes that empower and support students to connect the topics they study using real world applications based on Science Technology Engineering Art and Mathematics. The company extended their range of consumers by creating Arduino Pro (Arduino, 2023d) which consists of industrial development boards and countless documentation which remains loyal to the Arduino's user-friendly application.

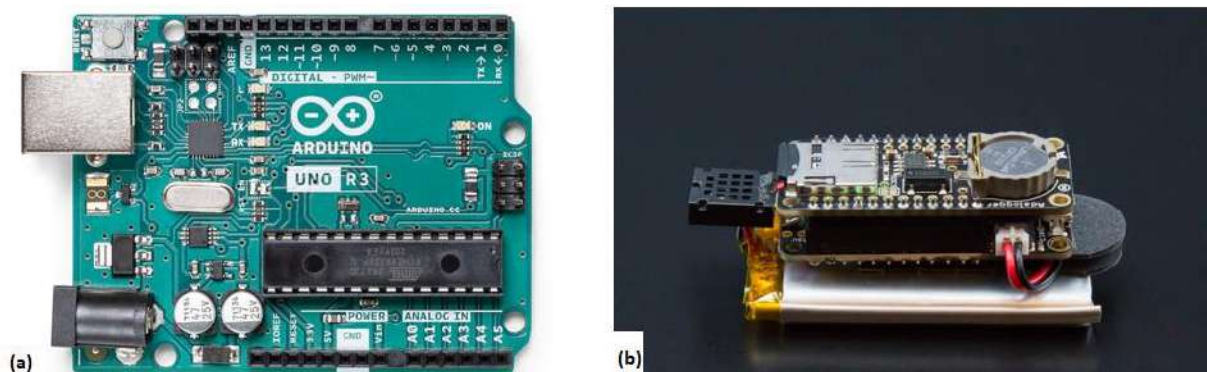


Figure 2.14. (a) Arduino Uno development board on the left (source: Arduino official web page) . (b) Adafruit feather 32u4 proto stacked with an adalogger on the right (source: Adafruit official web page).



Figure 2.15. Raspberry pi 3, model B stacked with a LoraWan Gateway with RAK2245 Pi Hat. This is the gateway where all the LoRa nodes send their data. The Raspberry Pi connects via WiFi or 4G (dongle) to upload the data to an IOT server in case no internet connection is available.

Another development board which is very popular and surpasses in processing power the Arduino-like products based is the Raspberry pi. It uses a microprocessor and is built in a way that can easily interface with hardware. Unlike the typical Arduino and Arduino clone microcontrollers, Raspberry pi can run an Operating System. The most common operating system used in Raspberry pi is Raspbian (Raspberry pi, 2023) and is a free operating system based on Debian, optimized for Raspberry pi. It can also be equipped with “hats” which are similar to the shields discussed previously. Raspberry pi is used in Chapter 4. A LoraWan Gateway with RAK2245 Pi HAT (RAK, 2023b) is used along with the Raspberry Pi 3, model B (Figure 2.15). The board is equipped with Wi-Fi or 4G (using an external dongle) to upload data to an IOT server (discussed in Section 2.3.8). The specifications of the board are summarized in Appendix A (Table A3).

2.3.3. Arduino IDE

Apart from hardware, Arduino developed an interface called Arduino Integrated Development Environment (IDE) which uses a variant of C/C++ programming language with an addition of special methods and functions developed for electronics. Another alternative to Arduino IDE is MicroPython (2013) which was developed by Damien George and CircuitPython (2023) created by Adafruit specifically for microcontrollers. Both programming languages are a variant of python and are supported by specific microcontrollers like ATSAMD21 (Microchip, 2023) or ESP32. For CircuitPython, any text editor allows the user to develop code while Arduino code and MicroPython use specific interfaces.

In this thesis Arduino IDE was used to upload “sketches” on the development boards. Sketches refer to code used for projects. When the Arduino project was born, Arduino IDE was used to upload code to microcontrollers via a programmer. The programmer can be either internal or external and aids the “burning” procedure. Burning refers to uploading the code to the board. The code is stored in the flash memory and is not erased even if the power is disrupted. Using the IDE, a user can upload another code replacing the old. Arduino IDE 2 (Figure 2.16) was released on September 12, 2022, adding more tools compared to the older version. Specifically, autocompletion, a built-in debugger and syncing sketches with Arduino Cloud (Arduino, 2023f) are three of the most important additions.

2.3.4. Electrical Connections and Communication Protocols

A microcontroller is a complete microprocessor-bases control system built into a single chip. Thus, one of its three main functions is the inputs, processing the data and outputs. A microcontroller has digital and analog inputs. Digital inputs can read two different states while the analog inputs can sense a range of voltages. There is always a need to standardize communication. For example, when it comes to two people interacting, person A wants to initiate conversation with person B the first action is saying “hi”. If person B replies, then person A will start the conversation. If person B is not replying, then person A will either repeat his first attempt by saying hi or stop the action. This can be considered as an example of a communication protocol to start a conversation. The procedure is described in the diagram (Figure 2.17) below.



Figure 2.16. Arduino IDE 2. The interface where users can write, compile and upload sketches to the Arduino and Arduino compatible development boards.

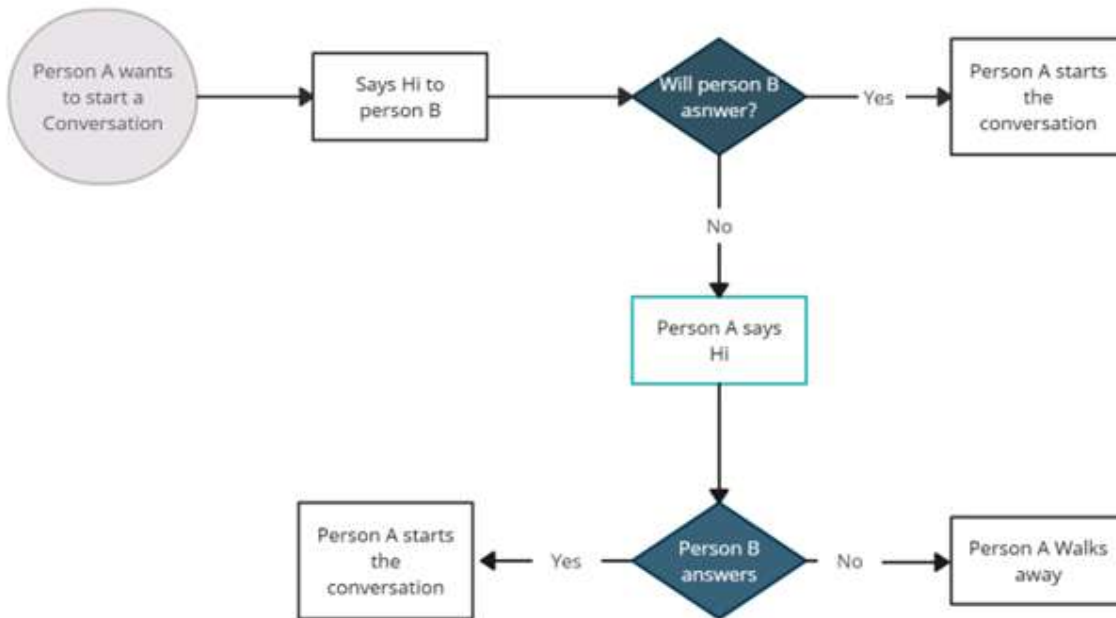


Figure 2.17. A simplified analog as an example of communication protocols. Person A wants to initiate conversation with person B. The first action is to say "hi". If person B replies, then person A will start the conversation. If person B is not replying, then person A will either repeat his first attempt by saying hi or stop the action.

The same idea using different variations applies to the way electronic parts communicate with the microcontroller. There are four communication protocols for the development boards described previously. The W1 (1-Wire), the I2C (Inter-Integrated-Circuit, 2 Wire), the SPI (Serial Peripheral Interface, 4 Wire) and the UART/RS232. Depending on the sensors or modules that perform a procedure, the

manufacturers can choose a protocol to communicate with the microcontroller or simply output voltage or current.

The 1-Wire protocol (for AVR microcontrollers)

The W1 protocol is a half-duplex bidirectional protocol. This means that information is sent from one device to the other via one wire. Its advantage lies in the fact that only one pin per MCU is used plus a ground reference. Most of the devices that use this protocol are powered by the protocol's pin (parasitic power supply). Having two contacts, 1-Wire devices are preferred in some cases since they are considered as the most economical way to add electronic functionality to objects that are not electronic for identification, authentication, delivery of calibration information and manufacturing data (Linke, 2008). For every 1-Wire slave, there is a unique, unalterable factory-programmed 64-bit identification number (ID), which corresponds to the devices address on the specific protocol.

Popular devices that use 1-Wire is the DS18B20 (Maxim Integrated Products, 2019) temperature sensor, the iButton (Hindman, 2014) and others.

Summing up, the advantages and disadvantages of using 1-Wire are the following (Giovino, 2021):

- It uses only one wire for communication and a common ground compared to other protocols.
- Carries data for up to 100 meters.
- Sensors can be powered by the 1-Wire (parasitic power).
- Can be used combined to other communication protocols.

The disadvantages of using 1-Wire are the following (RF Wireless World, 2012):

- Low speed compared to other communication protocols.
- Since it is applied in both software and hardware the data synchronization of data at the receiver must be implemented in software. This is considered a complex task.
- Even if it is designed for long distance communication compared to other protocols, noise can be inserted due to cable capacitance.
- Slave devices are manufactured only by one manufacturer (Dallas semiconductors).

UART (Universal Asynchronous Receiver/Transmitter) protocol

UART is one the most prevalent communication protocols in microcontrollers. This protocol uses asynchronous means, which means that there is no clock that synchronizes the output of every device transmitting. UART works with different types of serial protocols, when properly configured. It is considered a hardware communication protocol and uses two wires (signals). The two signals are the Transmitter (Tx) and the Receiver (Rx). The data are transmitted bit-to-bit serially using a baud rate. The baud rate is the rate (bits per second) at which the data are sent and since there is no clock synchronizing the devices, communication synchronization is achieved via common baud rate. Figure 2.18 displays the protocol. Serial communications need less circuitry and wires, thus, reducing the cost of the application (Pena & Legaspi, 2020). One of UART's advantages is based on the designer side. UART based devices can be robust and can assure quality. Properly designed UART devices can guarantee data transfer with no errors. As Pena and Legaspi (2020) discuss, UART can be used in debugging, manufacturing function-level tracing, customer of client updates and testing/verification of products.

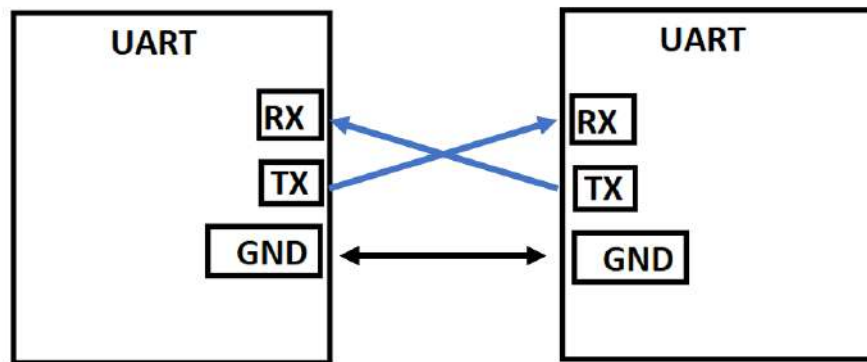


Figure 2.18. UART communication protocol. The two signals are the Transmitter (Tx) and the Receiver (Rx). The devices share a common ground.

Summing up, the advantages and disadvantages of using UART are the following (Yida, 2019):

- It is a very simple and well documented protocol.
- There is no need for a clock to achieve synchronization.
- It uses a parity bit which allows for error checking.

The disadvantages of using UART are the following:

- The data frame size is restricted to 9 bits.
- Multiple Master systems and Slaves cannot be used.
- Baud rates should be the same for the two devices communicating or within 10% of each device to avert data loss.
- The transmission speeds are low.

The I2C (Inter-Integrated Circuit) protocol

The I2C communication protocol allows multiple peripheral devices to communicate sharing the same bus. It only requires two signals, the serial clock (SDA) and the serial data (SCL). Those wires can support up to 1008 peripheral devices. It supports a multi-controller system, allowing more than one MCU to communicate with peripheral devices using the same bus. Most I2C devices can communicate at 100-400 KHz. Figure 2.19 displays the I2C communication protocol. In hardware level, usually two 4.7 K Ω pull-up resistors are needed. The value of the resistors depends on the amount of capacitance on the I2C lines are needed to guarantee the smooth communication of the devices. The controller or master communicates with the slaves (peripherals or other MCUs). The slave will only transmit data when the master requires it. Every device has a unique address in an isolated system. Every slave device must be configured before startup to define each operation.

The I2C communication is summarized as follows (Valdez & Becker, 2015):

1. When the master sends data to the slave:
 - a. Master (transmitter) sends a START statement and addresses the slave (receiver).
 - b. Master (transmitter) sends data to slave (receiver).
 - c. Master (transmitter) ends the transfer with a STOP statement.

2. If the master wants to receive (read) data from the slave:
 - a. Master (receiver) sends a Start statement and addresses the slave (transmitter).
 - b. Master (receiver) sends the requested register to read to slave (transmitter).
 - c. Master (receiver) receives data from the slave (transmitter).
 - d. Master (receiver) ends the transfer with a STOP statement.

Summing up, the advantages and disadvantages of using I2C are the following (Yida, 2019):

- It supports multi-master and slave communication.
- It is simple to establish amidst multiple devices since it only uses two bidirectional wires.
- Adopts to user's requirements.
- Multiple master support.
- Contains an acknowledgement mechanism.

The disadvantages of using UART are the following:

- Requires pull-up resistors thus has slower speed than SPI.
- As the number of devices increases it becomes more complex.

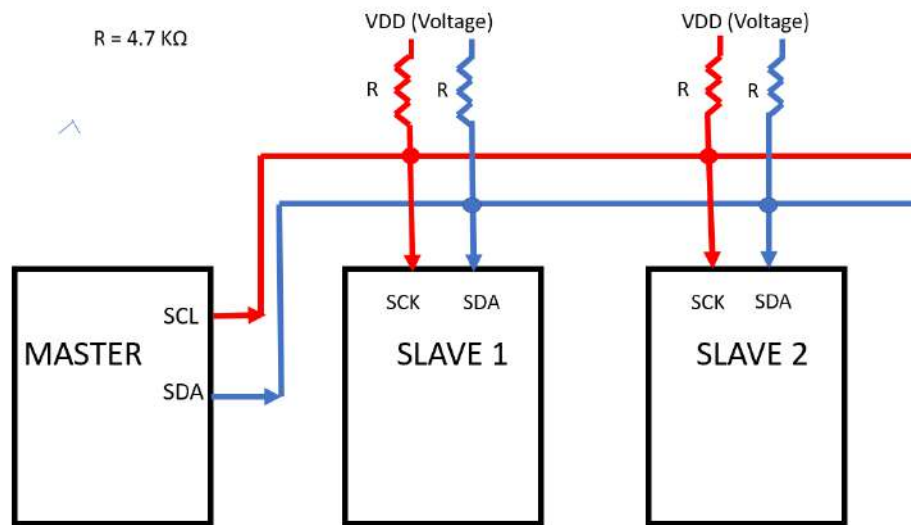


Figure 2.19. The I2C communication protocol with one master and two slaves. A 4.7 KΩ resistor is placed between VDD and the two wires (SCL, SDA).

SPI (Serial Peripheral Interface) Protocol

The SPI protocol is a synchronous, full duplex main-subnode-based interface mostly used for communication with peripherals for short distances. The protocol has one master and many slaves. The master is the microcontroller, and the slaves are the peripherals. It uses three or four wires. The following description is based on the four-wire variation. Three of them are common for all the devices. The clock signals are denoted as SCLK, CLK or SPI. The main out subnode MOSI (for slave). The main in subnode MISO (for master). Last, the fourth wire (SS) is the slave wire where each device has its own in the master's side (Figure 2.20). According to (Usach, 2015) this interface provides relatively high transmission rates with adaptable configurations. SPI is not officially specified. Even though this is an advantage since the designer can maximize its configurability it entangles the interconnection between various parts. Figure 2.21 displays the four pins needed for SPI communication along with their connection. The slaves, usually the peripherals, are connected in different designated pins on the master's side.

Summing up, the advantages and disadvantages of using SPI are the following (Yida, 2019):

- No complicated address system like I2C.
- The fastest protocol compared to UART and I2C.
- Data can be continuously transmitted compared to UART.
- Data can be received and transmitted at the same time since it is equipped with two different wires for this task (MOSI, MISO)

The disadvantages of using UART are the following:

- Four pins are required compared to the two pins of UART and I2C.
- No acknowledgement mechanism compared to I2C.
- No error check procedure compared to UART.
- It can support only one master.

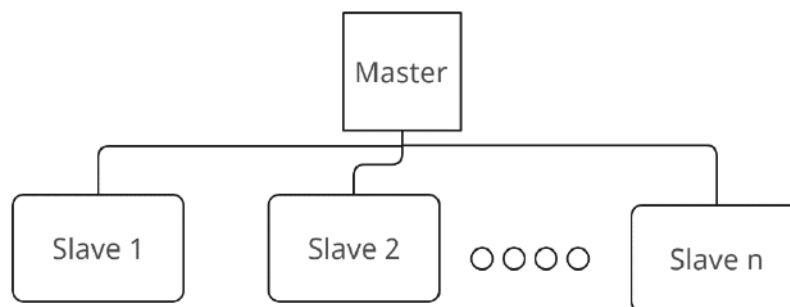


Figure 2.20. SPI communication with one master and many slaves.

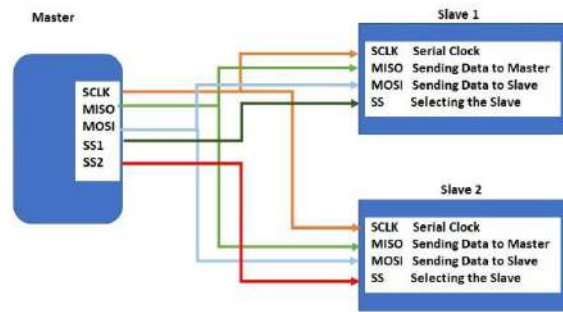


Figure 2.21. Pin diagram for SPI communication.

2.3.5. Wireless Communication

Many wireless communication protocols have been designed in the past decades. Wi-fi, Bluetooth Low-power, LoRa, GSM, 4G-LTE, Zigbee and NB-IOT are dominant in the IOT community. Depending on the application, each protocol has its own benefits and drawbacks. LoRa is more popular on power consumption and distance range, although it is new compared to the other protocols. Wi-Fi, Zigbee and Bluetooth protocols are the most prevalent but will not surpass LoRa when it comes to distance range and power consumption. However, LoRa protocol offers less baud rates. Thus, it is up to the researcher/individual to decide which protocol to use. This thesis describes only Wi-Fi, LoRa, GSM and 4G protocols since they were the ones used in the case-studies.

Wi-Fi (802.11)

Wi-Fi is a well-established wireless protocol. In the microcontroller world, many chips like Espressif's ESP 32 or ESP 8266 (ESPRESSIF, 2023), Microchip's ATSAMW25 (Atmel, 2016b) and others embed it on the MCU. Shields and individual modules that include 802.11 technology can be attached to the development board using communication protocols as discussed in Section 2.3.4. There are many Wi-Fi protocol standards created so far. ESP32 supports 802.11b/g/n/e/I (Banerji & Chowdhury, 2013). Figure 2.22a displays the Adafruit Huzzah32 -ESP32 (Fried, 2023b) development board and the Adafruit WINC1500 (Fried, 2022a), Wi-Fi shield for Arduino Uno is displayed on Figure 2.22b.

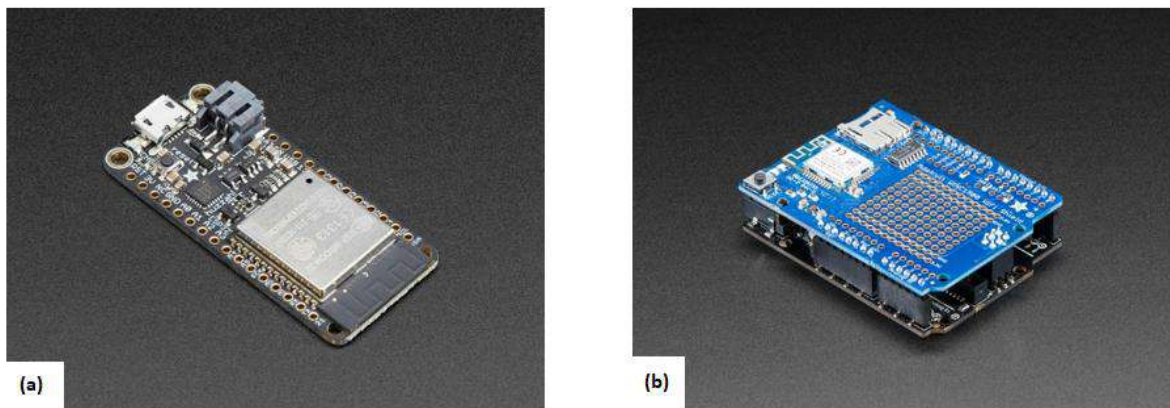


Figure 2.22. (a) Adafruit HUZZAH32 – ESP32 feather board (source: Adafruit official web site). (b) Adafruit WING1500 Wi-Fi shield for Arduino Uno (source: Arduino official web site).

ESP 32 MCU has two different modes. First, the Station mode (STA) needs an access point to connect to the internet and send the data to a server. It acts like a normal computer, getting an IP, behaving like a client by sending requests or as a server, when other devices connected to the network send requests to the ESP 32. In this case, Chapter 3 uses this mode to send the weather station's data via Wi-Fi to a public server. Second, the Access Point mode, where ESP 32 is simulating a router and other devices can connect to it.

Wi-Fi protocol transmission frequency bands are 2.4GHz or 5GHz. The range covered by Wi-Fi is up to 70 meters with no extender. Moreover, in case the applications require a high bitrate then Wi-Fi is an ideal choice. This protocol is ideal for areas where there is Wi-Fi coverage. Unfortunately, for precision agriculture this is not the case most of the time. In the field there is rarely Wi-Fi coverage. Moreover, greenhouses are usually equipped with Wi-Fi access points, but this is not the general rule. Furthermore, the typical Wi-Fi module power consumption is around 95-240mA, which depends on the standard and the operations of the Wi-Fi module. Table 2.1 presents the RF Power-Consumption Specifications for the ESP32 series. The power consumption measurements are taken with a 3.3 V supply at 25°C of ambient temperature at the RF port. All transmitters' measurements are based on a 50% duty cycle.

There are many libraries developed for the ESP 32. A list of the libraries can be found in Humfrey and per1234 (2022). Although ESP 32 is a very popular module, other modules, and wireless technologies like LoRa are more power efficient and cover a better distance range. However, Wi-Fi provides higher data rate compared to LoRa. Moreover, Wi-Fi's documentation is rich since it is a rather old protocol. Vinod et al. (2023) developed a smart IOT-based drip irrigation system for precision farming using the ESP-32 module.

Even though 802.11 is not considered as power efficient as other wireless technologies and the range is quite limited, in 2016 the IEEE 802.11ah Wi-Fi standard was released (Tian et al., 2021) promising lower power consumption and better coverage. It enables Wi-Fi to be used as an IOT solution with a range coverage of 1000 meters allowing for high data rate (from 150 Kbps – 78 Mbps).

Summing up, the advantages of using Wi-Fi communication protocol are the following:

- High data rates.

- Easy to install.

Disadvantages:

- High power consumption.

- Short-distance communication.

LoRa (Long Range)

LoRa, which stands for long range, is a wireless protocol for low power wide area wireless network (LPWAN) used by devices that implement IOT applications. The devices or end-nodes, are usually battery powered thus, making the low power feature attractive for the network designers. LoRaWAN aims to provide key IOT specifications such as bi-directional communication, end-to-end security infrastructure, and services that focus on mobility and localization. It is a rather new wireless communication protocol that is used for large scale wireless sensor networks (WSNs). The effective cost, long distance range and energy efficiency makes this protocol appropriate for various applications. LoRa was created for low-bitrate applications, like smart city, agricultural, biomedical sensors, and other (Costa & Duran-Faundez, 2018; Ji et al., 2019; Dragulinescu et al., 2020; Andrade & Yoo, 2019).

The architecture of LoRaWan is based on a star-of-stars topology, where gateways act as the relays to forward messages between end-devices and a centralized network server (Lora Alliance, 2023b). The gateways act as a transparent bridge and are connected to the network server (public or private) via standard IP connections by converting RF packets to IP packets. Converse, the procedure is applied for IP packets to RF packets. The communication is bi-directional, meaning that both gateway and end-node can talk and listen to each other. This is a practical application since useful features like Firmware Over-the Air (FOTA) (Giovino , 2019), which is upgrading the end-nodes firmware wirelessly, can be applied with no physical contact. This is efficient, especially for agricultural applications, where sometimes hundreds of end-nodes (depending on the gateway) are spread around many acres of land.

Table 2.1. RF Power-Consumption specifications for 802.11 (WiFi). The power consumption measurements are taken with a 3.3 V supply at 25°C of ambient temperature at the RF port. All transmitters' measurements are based on a 50% duty cycle. (source: ESPRESSIF, 2023)

Mode	Min	Typ	Max	Unit
Transmit 802.11b, DSSS 1 Mbps, POUT = +19.5 dBm	-	240	-	mA
Transmit 802.11g, OFDM 54 Mbps, POUT = +16 dBm	-	190	-	mA
Transmit 802.11n, OFDM MCS7, POUT = +14 dBm	-	180	-	mA
Receive 802.11b/g/n	-	95~100	-	mA

LoRa's physical layer is based on Chirp Spread Spectrum (CSS) modulation (Reynders & Pollin, 2016) combined with Frequency Shift Keying (FSK) (Bensky, 2019). In other words, to send information LoRa needs to convert those bits of information to electromagnetic waves, to travel via air. CSS and FSK are two methods that make this conversion possible. The frequency the protocol uses is 169 and 433 MHz (for Asia), 869 MHz (for Europe) and 915 MHz (for North America). These frequency bands are considered license-free, thus LoRa Alliance recommends a duty cycle limitation of 1% in the European band. In other words, the IOT device can transmit less than or equal to 1% of the time. This limitation is performed to avoid interference between the devices. Moreover, LoRaWAN uses a range of 24 – 80 channels with either 125 kHz or 250 kHz bandwidth (The Things Network, 2023).

Examples of LoRa end-nodes include OEM brands (Figure 2.23a) which might include embedded OLED screens along with wireless LoRa functionality. Figure 2.23b displays the Adafruit LoRa radio FeatherWing (Fried, 2023f) which acts as a shield for the Adafruit feather development boards (described in Section 2.3.2). Figure 2.24 displays LoRaWAN applied in agriculture and summarizes all the applications and benefits. Up to 1000 end-nodes can establish connection with a gateway. A LoRaWAN gateway can receive and transmit signals over 15 kilometers in rural areas and up to 5 kilometers in dense urban environments depending on the location of the end-node (indoors or outdoors) and the configuration of the data rate and spreading factor (Semtech Corporation, 2020). The spreading factor is a parameter of LoRa's physical layer and is part of a mechanism that adapts the emission power and transmission data rate to the network conditions around the device. Table 2.2 displays the spreading factor, the data rate, the bit rate (bit/s) for Europe's default channels. The highest the spreading factor, the longer the range covered but the data rate is lowered. Moreover, the higher the spreading factor the longer it takes to send a message with a fixed size thus, power consumption increases. Reduction of packet loss can be achieved by sending the same message on different radio channels. The tradeoff is increased power consumption. The data rates (DR) of LoRa protocol ranges from 0.3 to 11 Kbps. The DR adaptation allows a trade-off between communication range and message duration. Moreover, due to the spread spectrum of the protocol, different DRs are not interfering with each other and thus establish virtual channels improving the capacity of the gateway. For power consumption efficiency the network server optimizes the spreading factor and data rate individually for each end-node by means of Adaptive Data Rate (ADR) strategy.

Table 2.2. Spreading factors, Bandwidth, Data Rate and Bit Rate combinations for LoRa's transmission in Europe.

Spreading Factor + Bandwidth	Data Rate	Bit Rate (bit/s)
SF12 / 125 KHz	0	250
SF11 / 125 KHz	1	440
SF10 / 125 KHz	2	980
SF9 / 125 KHz	3	1760
SF8 / 125 KHz	4	3125
SF7 / 125 KHz	5	5470
SF7 / 250 KHz	6	11000

Figure 2.25 displays the LoRaWan network architecture. The end-nodes send packets of data to the gateway. The gateway sends the data to the network server. Through the backend layers and the link layer the data reach the application layer where they are gathered, displayed, and processed by the application (described in Section 2.3.8).

There are three classes of end-point devices that serve the needs of the users according to the application as described by Lora Alliance (2023b). First, Class A uses the lowest power and supports bi-directional communication. This is the default class. It is launched by the end-device and is fully asynchronous. The protocol used for this type of communication is the ALOHA (Baccelli et al., 2006). The end-device can enter low-power mode, for a period of time set by the user. Second, Class B, which is bi-directional and synchronize the end-node with network by using periodic beacons and schedule open downlink ping slots. The tradeoff for this configuration is additional power usage since the end-nodes have to “wake up” from power down periods. Third, Class C, keeps the end-node active all the time thus, reducing the latency. However, the power drain of the receiver increases which makes this class choice efficient for continuous power supply only. This thesis uses class A end-nodes, since the case-studies are based on renewable sources of energy (solar-panels) or limited battery applications (Chapter 4, 6). LoRa’s power consumption of different transceiver manufacturers are displayed in Table 2.3.

Summing up the advantages of the protocol:

- Long distance range.
- Low cost.
- Low power.
- A single gateway for 500 to 1000 end-nodes, depending on the gateway.
- Resists multipath fading (Mahender et al., 2018) and Doppler effect.
- Firmware over-the-air update.

The disadvantages of the protocol:

- Low baud rate.
- Open frequencies are different for each country.

Table 2.3. Main current consumption details on LoRa/LoRaWan common transceivers from different manufacturers. (source: Ibanez et al., 2017)

Transceiver	Sleep	Transmit	Receive
Semtech SX1272	0.1 mA (max 1 μ A)	Min: 18 mA (7 dBm) Max: 125 mA (20 dBm)	10.5 or 11.2 mA
Semtech SX1276	0.2 μ A (max 1 μ A)	Min: 20 mA (7 dBm) Max: 120 mA (20 dBm)	10.8, 11.5 or 12 mA
HopeRF HM-TRLF-LF/HFS	2 μ A (min 1,2 μ A, max 3 μ A)	Min: 35 mA (13 dBm) Max: 120 mA (20 dBm)	16 ma (min. 15 mA, max. 18 mA)
Micochip RN2483	Up to 100-150 μ A	Min: 17.3 mA (-4.0 dBm) Max: 38.9 mA (14.1 dBm)	14.2 mA



Figure 2.23. (a) OEM LoRa development board embedded with an OLED screen (source: Wirelesslan.gr) (b) Adafruit LoRa radio FeatherWing. A shield for the Adafruit feather development board series (source: Adafruit official web site).

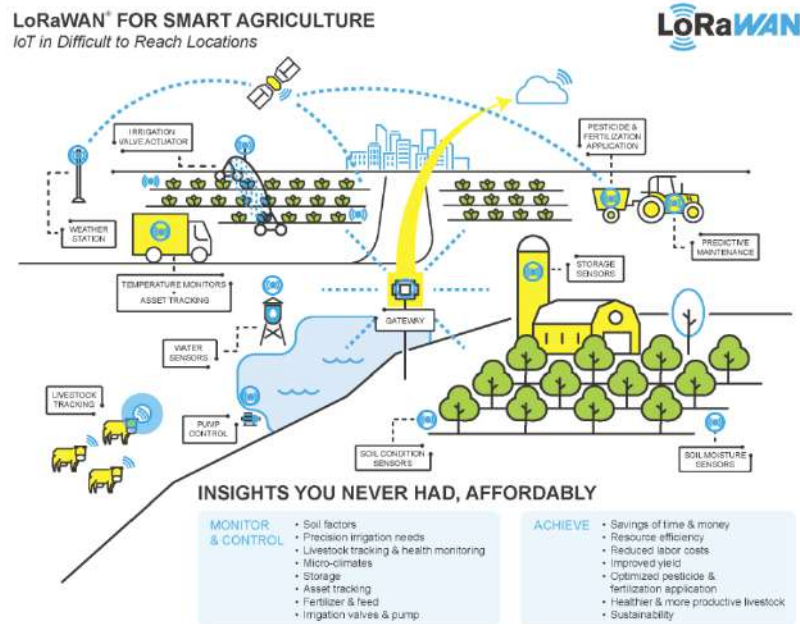


Figure 2.24. Benefits of using LoRaWAN in agriculture. Up to 1000 end-nodes can establish connection to the gateway. (source: LoRa alliance web site)

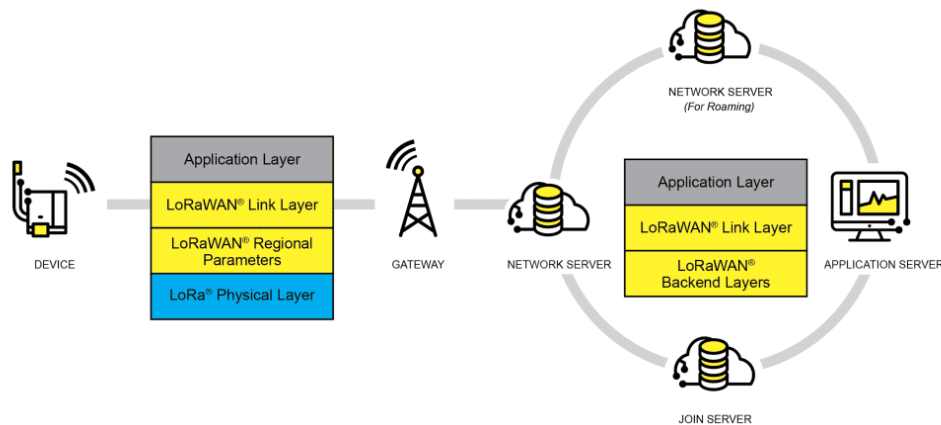


Figure 2.25. LoRaWAN infrastructure. End-node device route to the application server (source: LoRa alliance web site).

GSM (Global System for Mobile Telecommunications)

GSM protocol (Rahnema, 1993) provides mobility for IOT devices thanks to its high coverage. Even though many countries deprecated the protocol, Greece and other European countries are still using it. Its bit rate is low, compared to 3G, 4G/LTE and 5G technologies. However, lower bitrate means lower power consumption making this technology useful for transmitting sensor data from battery-powered devices. The over-the-air bit rate is 270 Kbps. The uplink frequency range is 933-960 MHz.

GSM networks are generally compatible with a broader variety of devices. Thus, using GSM as the bridge between end-nodes and IOT servers is an ideal combination for power efficiency and coverage. Moreover, GSM modules are less expensive than 4G modules.

Adafruit Feather Fona (Fried, 2023e) is a development board that supports GSM technology. It uses a 32u4 MCU and the community offers abundant documentation. The power consumption of the Adafruit module for the enabling GSM is 70mA on average. For a TCP/IP connection it requires 150 mA on average and disabling GPRS requires 113mA on average. The module can be put to sleep mode and will drop the current draw, but the cellular connection will still be on. To disable the cellular connection, the module is equipped with a dedicated pin that switches the GSM module off.

Last, the module is equipped with a uFL connector for external antennas.

A significant disadvantage to the other wireless communication protocols is the need for an internet service provider. Usually, the fees are charged monthly, and a SIM card is needed. There are many alternatives to local Greek providers such as Monogoto (2023). Monogoto is a card Sim supplier for global connectivity. Gartner Peer Insights (2023) provides a IOT service provider review with ratings. Users can rate their experience with the listed providers.

Chapter 6 case-study uses GSM cellular connection to upload the data to an IOT server.

4G LTE networks

4G/LTE is a fourth-generation mobile network technology which offers high-speed data transfer services. Compared to GSM, 4G/LTE provides higher data rates and can reach up to 50 Mbit/s for the upload

channel. However, its power consumption is higher than GSM's. Arduino MKR1500 (Arduino, 2023g) is an open-source development board that provides 4G services. Moreover, a 4G dongle can be used to add 4G network connectivity to a Raspberry pi. Chapter 4 uses this configuration combining 4G and LoRa wireless connectivity to create a gateway using a Raspberry pi 3 B+. More information about 4G/LTE networks performance can be found on Krasniqi et al. (2018).

2.3.6. Designing the Hardware

There are many open-source programs for designing hardware prototypes. Fritzing (Knorig et al., 2009) and Kicad (Evans, 2021) are very popular and easy to use.

Fritzing is an open-source software that makes electronics hardware design attainable for any individual interested. It is used by engineers, artists, researchers, and makers for sharing their electronic devices' documentation. Fritzing produces layouts for printed circuit boards (PCB) for mass production and can also be used for educational purposes. It has a vast variety of libraries, allowing the user to input any kind of electronic part, microcontroller, development board, shields etc. in their design. Figure 2.26a displays Fritzing's interface. Three layouts are available. A breadboard, a schematic, and a PCB layout. Depending on the application the user chooses the appropriate output. When the user designs his project using one of the three layouts, Fritzing converts the design in the other layouts simultaneously. Thus, making it easier to output the design for different uses. A printed circuit board can be made at home using DIY etching processes (monusuthar2016, 2017) or by sending the schematic Fritzing outputs in a PCB development company. All the designs for the case-studies in this thesis were developed using Fritzing.

Kicad (Figure 2.26b) is an open-source Electronics Design Automation (EDA) suite. Similar to Fritzing, one of its features is schematic entry, which is the step where the electronic diagram of the electronic circuit is developed. However, Kicad adds more features and offers 3D rendering and a circuit simulator. Circuit simulation is an important attribute. Usually, makers use the solderless breadboard to create a prototype. Circuit simulation allows makers to create their prototype virtually and check for the circuit output for troubleshooting in case the circuit is not working as expected during the breadboard prototyping procedure.

2.3.7. Loggers for Environmental Monitoring and Sensor Working Principles.

There are many sensors developed for various applications. Air and liquid temperature sensors, air humidity sensors, soil moisture sensors, PH sensors, Electrical Conductivity (EC) sensors, pressure sensors etc. Each sensor has its own working principle. For example, thermistors are temperature sensitive resistances. Resistance is affected by changes in temperature. Using Ohms law, a simple circuit called voltage divider converts the resistance output to voltage. Figure 2.27a displays a voltage divider with a 5 Volt reference voltage. The output is connected to an analog input of Arduino Uno development board. This is essential since an MCU can only read voltage. The resistance value can be calculated using the following formula:

$$V_{out} = V_{in} \frac{R_1}{R_1 + R_2} \quad \text{thus, } R_1 = V_{out} \frac{R_2}{V_{in} - V_{out}} .$$

R_1 is the thermistors resistance value. R_2 is the reference resistor (in this example 10 kOhms) and is usually chosen to have the same value as the thermistors when temperature is 25°C. V_{in} is 5 Volts in this example and V_{out} is the voltage read by the microcontroller. Further, R_1 is mapped to a specific

temperature using functions provided by the manufacturer. The procedure for using thermistors including the wiring and equations for converting resistance to temperature values along with methods to maximize the accuracy of the thermistor's measurement are described in Chapter 6.

Similarly, Fisher et al. (2018) developed an open-source cloud-connected sensor-monitoring platform for measuring soil humidity using Watermark 200SS granular matrix sensors. The sensor's measurement

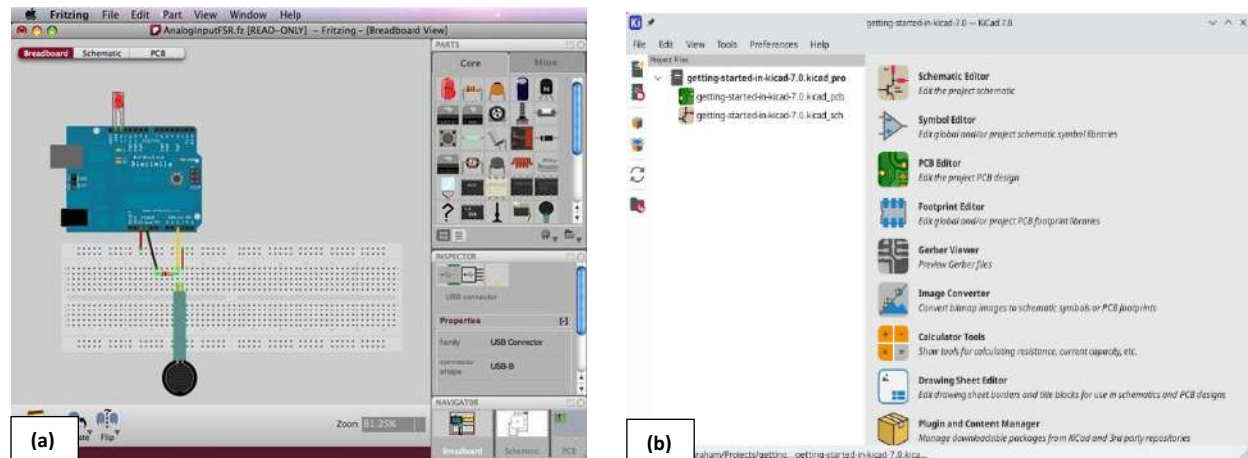


Figure 2.26. (a) Fritzing interface. A tool for designing electronics and producing documentation, on the left (source: Fritzing official web site). (b) Kicad user interface, on the right. Kicad is an open-source Electronics Design Automation Suite (source: Kicad official website).

circuitry consists of voltage dividers (the system is equipped with four sensors) and uses the same working principle as the thermistors. The resistance of the sensor changes according to the soil moisture.

Every sensor uses its own communication protocol to connect with a logger. Sensor manufacturers usually develop loggers that are compatible with their sensors allowing for plug and play features. This is because the logger is already programmed for a specific sensor. Development boards allow a maker to construct a high-quality logger and adapt its code to any sensor with a known communication protocol (discussed in Section 2.3.4).

Arduino compatible sensors are built with one or more of the standard microcontroller communication protocols. Unfortunately, these sensors are mostly used for educational purposes since the quality of the readings is not always high and sometimes the measurements are unreliable. On the other hand, industrial sensors might be compatible with the typical MCU communication protocols but there are cases where their output is current, RS485 (Wang & Liu, 2021) and Serial Digital Interface at 1200 baud (SDI-12) (SDI-12 Support Group, 2021). These protocols can be simulated using additional hardware or software with an Arduino or Arduino compatible development board. Chapter 4 interfaces a soil moisture sensor with RS485 output, with an Arduino compatible logger. Current output can be read by an Arduino or Arduino compatible logger using additional hardware with an operational amplifier for simple linear signal processing and a resistor for dissipating current. DFRobot is an electronics manufacturer that developed an analog current to voltage converter (DFRobot, 2020) with a range of 4 to 20 mA (Figure 2.27b).

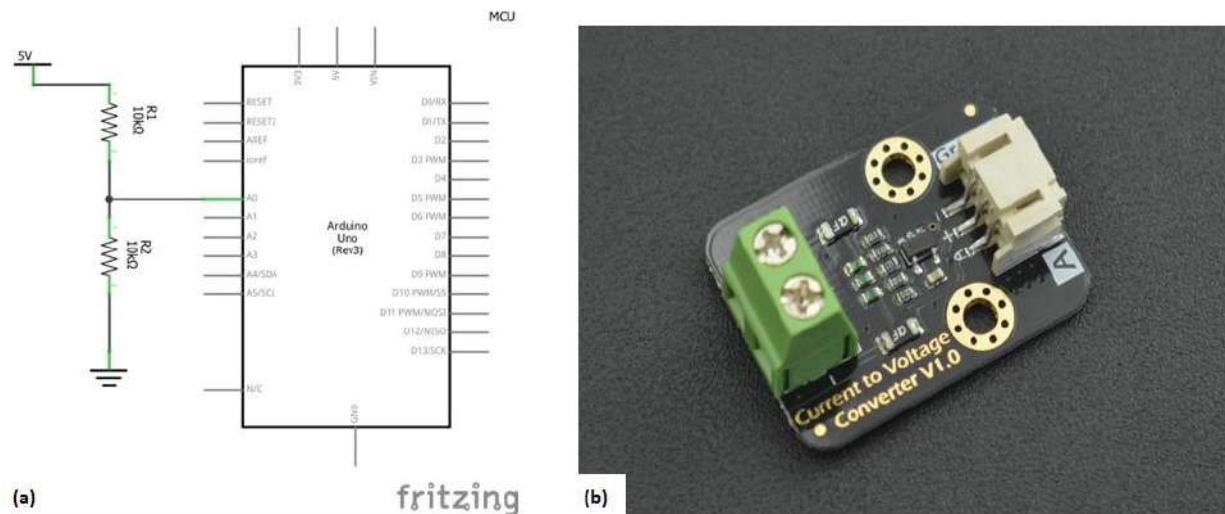


Figure 2.27. (a) A voltage divider with input voltage 5 Volts. The output is connected to an analog input of the Arduino Uno development board. (source: created using Fritzing) (b). SEN0262 Current-to-voltage module. The module converts current that ranges from 4 to 20 mA to voltage (source: DFrobot official web site).

2.3.8. Internet of Things platforms.

IOT devices upload sensor or other forms of data in an IOT platform. The IOT platforms form a bridge between devices and users by providing Application Programming Interface (API) (Red Hat, 2022), to ease their control and management. Thus, administrators, developers, researchers and IOT users are facilitated and relaxed from challenging configurations the API layer requires. IOT platforms offer space for connecting, managing, monitoring, and controlling the data flow send by devices. Further, the device can be controlled from the platform using a downlink. For example, an IOT device responsible for controlling the lights in a warehouse can be coordinated by the user via the IOT platform, simply by sending a signal to command switching on the lights for the user to enter the warehouse. In other words, the user can issue any command that is not preprogrammed.

Babun et al. (2021) provide a survey for IOT platforms in a communication, security and privacy perspective. The authors consider seven technical comparison criteria. The topology design, programming languages, third-party support, extended protocol support, event handling, security and privacy. Moreover, they present viable solutions for the platforms, to strengthen security and privacy. Last, advantages and drawbacks of every IOT platform are presented to aid IOT administrators, researchers, and users to make a choice based on academic judgement.

In this Section, three IOT platforms are presented since they were used in Chapter 3, 4 and 6. For more information on IOT cloud providers, Zdravkovic et al. (2016) provide a survey in a functional and design perspective to current IOT platforms and related research. For small or large-scale projects, there are many free IOT platforms like ThingSpeak (2023) or Adafruit IO (2023). Prototype designs are established as small-scale projects and are usually free of charge. Later, if the user is pleased with the performance of the system, the designer introduces the corresponding fees the platform requires.

ThingSpeak

ThingSpeak is an IOT analytics platform service that allows the user to store, visualize and analyze their data in real time. The toolbox this platform provides includes alert commands using web services like Twitter, Twilio or email. Moreover, with MATLAB analytics the user can write and execute MATLAB code (MathWorks Inc, 2022) to carry out preprocessing, visualizations, and statistical analyses. ThingSpeak provides a vast number of tutorials for various IOT devices to connect.

First, the user signs up and creates a new channel (Figure 2.28). Then she/he sets up the fields where the data will be stored. For example, if the IOT device sends air temperature and humidity data, Field 5 is for humidity and Field 6 is for temperature (Figure 2.29).

ThingSpeak allows for various visualizations such as a filled area 2-D plot, a multiple series chart or a time series chart. Other graphs can be created using MATLAB visualization tab, interacting directly with data of any Field. Moreover, statistical analysis or any kind of calculation can be performed using the MATLAB Analysis Tab. ThingSpeak, offers Widgets for every Channel. For example, Gauge meters, numeric displays, Lamp indicators or image displays. The data gathered from the Fields can be exported in JSON, XML or CSV format. Last, various Apps like Plugins, TimeControl, React, Talkback and ThingHTTP are provided. Plugins is an App that allows the user to create Google Gauges, charts, or custom displays using HTML, CSS, and JavaScript. Timecontrol works with other ThingSpeak Apps to perform an action at a specific time or on a regular schedule. React works with other ThingSpeak Apps just like Timecontrol to perform actions when a channel data meets a certain condition. For example, if the battery value of the IOT device runs low, maybe below 3.7 Volts in case the device is powered by a LiPo battery, then the platform can email the user. TalkBack enables any device to act on queued commands. ThingHTTP enables communication among devices and web services without having to implement the protocol on the device level. For example, React can work with ThingHTTP and send a request to a custom made HTTP server. All the available documentation can be found in ThingSpeak (2023) or other resources online.

Adafruit IO

Similar to ThingSpeak, Adafruit IO is a server provided by Adafruit, a company founded in 2005 by an MIT engineer, Limor Fried. Her goal was to create an online shop/forum for learning and supplying electronics. Her team provides numerous documentations based on their products but not limited to them. The procedure to use the Adafruit IO services is similar to ThingSpeak. The user has to sign up and choose a data plan. The user can choose the Adafruit IO Basic plan, which is free and provides 2 devices, 5 groups and 10 feeds. The server stores the data for 30 days and the user can retrieve them in an excel format. Adafruit IO+ provides unlimited devices, groups and feeds plus other privileges for a monthly fee. Graphs and apps are available for data processing and can be scheduled to notify the user if a condition is met via email. Chapter 3 uses Adafruit IO.

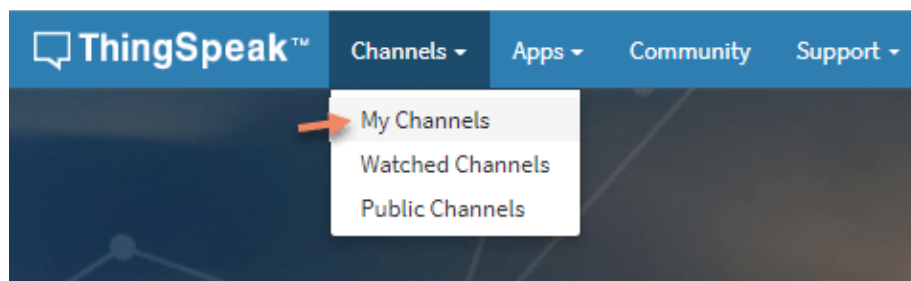


Figure 2.28. ThingSpeak menu. Creating a channel.

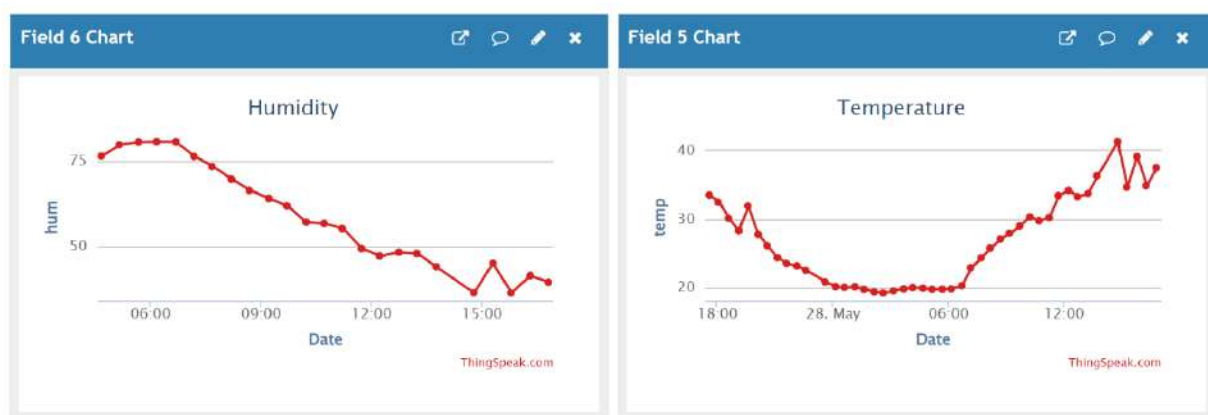



Figure 2.29. ThingSpeak data visualization from Field 5 and 6. The x-axis is the timeline while the y-axis displays the value of the characteristic the sensor measures.

The Things Network (TTN)




The Things Network is an IOT ecosystem that establishes networks, devices and other solutions using LoRaWAN. It uses the Things Stack Community Edition, which is crowdsourced and decentralized LoRaWAN network supporting the open-source community. The services provided by TTN are free for limited use and there are many options for other plans depending on the user's needs. It is used as an intermediate server and forwards the data in other IOT servers that offer data storage. The user can establish a gateway or can use a gateway (owned by another user) which is open. By setting the open setting on, the users support the community by offering LoRaWAN access to other users. Figure 2.30 presents The Things Network settings for the LoRa gateway used in the present research (Chapter 4). The frequency plan is chosen as EU_863_870_TTN for Europe.



biometry_uth
 ID: biomet-uth

Disconnected

General information

Gateway ID	biomet-uth	
Gateway EUI	BS 27 EB FF FE 02 3F 50	<> 
Gateway description	None	
Created at	Feb 4, 2022 13:00:26	
Last updated at	Feb 5, 2022 23:39:37	
Gateway Server address	eu1.cloud.thethings.network	

LoRaWAN information


Frequency plan	EU_863_870_TTN
Global configuration	 Download global_conf.json

Figure 2.30. The Things Network settings for the LoRa gateway used in the present research. The frequency plan is chosen as EU_863_870_TTN for Europe.

3. A method comparison study between open-source and industrial weather stations.

3.1. Goal of the Project

Meteorological parameter monitoring is essential in the agricultural field. Precision agriculture is based on the characteristics measured and stored by individual sensors or a collection of sensors embedded in one system. The latter is an example of weather stations. Many researchers publish their work on inexpensive meteorological parameter monitoring systems deployed by open-source development boards. Specifically, Kim et al. (2022) developed a Weather Station (KOSEN) and a server for provision of weather information to farmers. Gunawardena et al. (2018) developed a low cost open-source distributed sensor network and used an artificial neural network to connect the radiation errors that caused disagreement between the novel device and the corresponding research grade sensors. The sensor network development is based on an Arduino Mega. Khattab et al. (2019) developed an IoT-based cognitive monitoring system for early plant disease forecast based on Arduino Mega. Sagheer et al. (2020) developed a cloud-based IoT platform for precision control of soiless greenhouse cultivation using an Arduino microcontroller.

An open-source weather station was developed and constructed for meteorological parameter monitoring inside or outside greenhouses. The station can measure only air temperature and humidity, but due to its modular nature other sensors can be added effortlessly since documentation for open-source systems is rich and accessible to everyone. Two types of system evaluation were performed to the novel device. First, the evaluation of the device's operation. Second, the evaluation of the agreement between the open-source and its corresponding industrial weather station, regarding their temperature measurements. The experimental design and the statistical analysis are described in detail in Section 3.6 and 3.7.

3.2. Software

The device was programmed using Arduino IDE. The libraries used are listed in Table 3.1. The code was simple since the station's readings regarded only air temperature and humidity. In Figure 3.1, the flowchart presents the sequence of the weather station's functions. Arduino IDE consists of two functions. The setup function, which runs the code only once from the moment the module is powered. The loop function, which runs repeatedly until the module is powered down.

Table 3.1. Libraries used in the code and their functionality.

Library	Functionality
FS	Data Storage
SD	Data Storage
RTCLib	Real Time Clock
SHTC3	Sensor
AdafruitIO_WiFi	Connecting to Adafruit IO

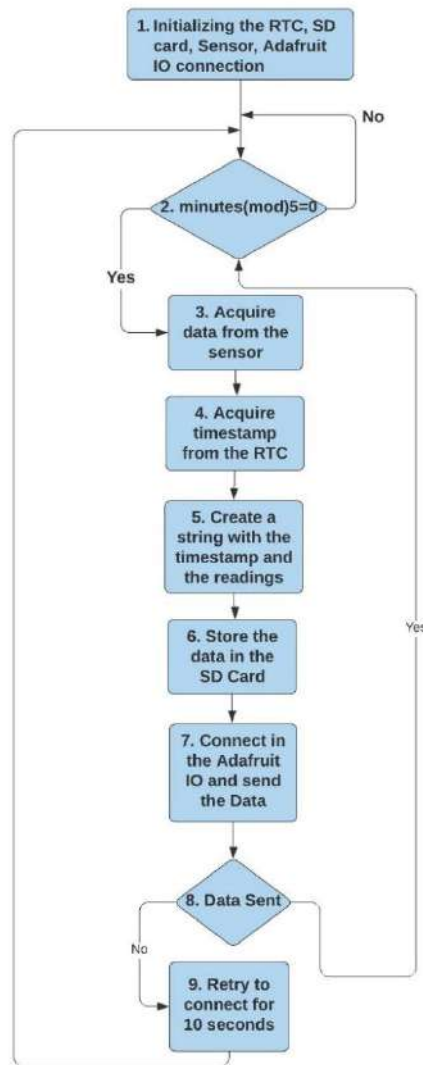


Figure 3.1. Flowchart that summarizes the sequence of the weather station's functions.

First, in the setup function of the Arduino IDE, we initialize the RTC, SD card, temperature/humidity sensor functionalities and then the module connects to the Adafruit IO server. If it fails to connect, the flow continues, with the system not taking any action. Second, in the loop function of the Arduino IDE the system is repeatedly acquiring the minutes of the timestamp from the RTC and is compared to the operation minutes modulo 5 which is the remainder of the present minutes divided by 5. If the result is 0 then the temperature is acquired from the SHTC3 temperature/humidity sensor and the timestamp which includes the present year, month, day, hour, minute, second. Third, a string is created and is stored as a text file in the SD card. Finally, it connects to the Adafruit IO and sends the data. If the connection is not established, the system retries for 10 seconds. If it fails to connect the system returns to the point before the first decision-making procedure, acquiring repeatedly the minutes from the timestamp. The same applies also in case it succeeds, taking no further action.

3.3. Cost of the device

The total cost of the open-source weather station was about 55 Euros in 2021 market prices. Table 3.2 displays the cost of every component in detail.

3.4. Industrial Weather Station

The industrial weather station is from Symmetron company, equipped with a Thygro SDI-12 air temperature and humidity sensor (Figure 3.2), a solar radiation sensor, solar panels, battery, logger, and a solar radiation shield. The air temperature accuracy is $\pm 0.2^{\circ}\text{C}$ and the resolution is $\pm 0.015^{\circ}\text{C}$. The total cost including the logger, the sensors, the battery and all the software was around 3000 euros. The service costs for remote monitoring are additional.

3.5. Hardware

The open-source weather station constructed for this experiment consists of four major components. First, the coordinator of the station's functions is the ESP-WROOM-32 (ESPRESSIF, 2023) microcontroller embedded in a wemos lolin evaluation board (Wemos Electronics, n.d.) (Figure 3.3b), which targets a wide variety of applications, ranging from low-power sensor networks to the most demanding tasks, according to the datasheet. Second, the logging function is performed by two components, a microSD card module (Figure 3.3d) and a real time clock (RTC) DS3231 (Maxim Integrated Products, 2015) (Figure 3.3c). Third, the temperature/humidity sensor is an Adafruit Sensirion SHTC3 Temperature and Humidity sensor (Siepert, 2022) (Figure 3.3a), which has an accuracy of $\pm 0.2^{\circ}\text{C}$, response time 8 seconds and 0.01°C resolution. Last, the device can store data in a server provided by the Adafruit company, Adafruit IO (2023). The medium used for remote data logging is Wi-Fi, available in the greenhouse.



Figure 3.2. Temperature and humidity sensor enclosed in a radiation shield of Thygro SDI-12.

Table 3.2. Detailed cost of every component for the construction of the open-source weather station.

Components	Cost (In Euros)
LOLIN ESP-WROOM-32	10
SHTC3	3.5
SD-Card module	1
Real Time Clock	1
Box Enclosure	5
PCB	2
Tripod	20
Material for the radiation shield	5
Angle Brackets	2
Screws and Nuts	1
Wires	2
Screw terminals	2
Resistors	0.6
Switch	0.5
Total	55.6

3.5.1. Real Time Clock

The real time clock module is a DS3231 low-cost and highly accurate I2C module, with integrated temperature compensated crystal oscillator. A CR2032 coin cell battery can be inserted into the battery dedicated slot to maintain the time regardless of the MCU's power state. It is a backup solution in case of power failures or power saving modes. The information that can be retrieved from the module regarding time is years, months, day, hours, minutes, and seconds. Due to its embedded sensor, temperature is also available for compensation. It features a 32Kb memory for data storage. Furthermore, two programmable time-of-day alarms and a programmable square-wave output are available. The last one is very useful especially for low-power applications that need to wake the MCU up in predefined intervals. A temperature-compensated voltage reference and a comparator circuit constantly monitors the voltage external supply and switches to battery power if necessary. There are many libraries available for this module. The one used for this project is RTCLib from evaherrada (2023).

3.5.2. Temperature and Humidity Sensor

The sensor used for this project is an Adafruit SHTC3 (Siepert, 2022) temperature and humidity sensor by sensirion. The breakout board was made by Adafruit which makes the connection with the MCU easy and straightforward. All the necessary electronic components and circuitry to make the communication between the sensor and the MCU are available in the breakout board. The communication protocol is I2C. There is no need for soldering wires as the breakout board is equipped with a SparkFun Qwiic (Sparkfun, 2023) which is compatible with STEMMA QT (Fried, 2022b) connectors. These connectors are 3 and 4 pin JST PH and are plug-n-play directly from the sensors to the STEMMA QT compatible development boards. Last, there is a power regulator which allows the user to power the sensor either by 3.3 or 5 volts. The library used for this module is the Adafruit SHTC3 (PaintYourDragon, 2022). The documentation with the technical specifications, the guide to install the library and the wiring is available in Adafruit's website. Table 3.3 summarizes the technical specifications of the sensor.

The device stores data in the sd-card and can transmit it to the IOT server every five minutes, simultaneously. The sd-card module works as a back-up system in case the WiFi connectivity is lost.

The sensor's solar radiation shield (Jakub, 2017) is custom made, using inexpensive and widely available materials to the public. Specifically, four flowerpot plates were used which were painted with white spray and 3 cm nylon stand-offs were placed between them to be separated so air can flow freely. The temperature sensor is placed inside to avoid direct sunlight exposure and prevent biased readings.

The box that houses the system is an IP66 plastic box (Kafkas, n.d.) and the body of the station is a camera tripod (Hama, n.d.) embedded with aluminum angle brackets that contain mounting holes, which are useful for placing the box on the desirable height. The circuit is displayed on Figure 3.4. The complete finalized product is displayed in Figure 3.5. Headers were soldered on the Printed Circuit Board for the vital and more expensive parts. This implementation makes it easier for developers to replace broken parts and there is no need to construct the PCB from scratch or desolder components. The device was tested in the laboratory for one week to make sure everything will work as planned during the field experiment.

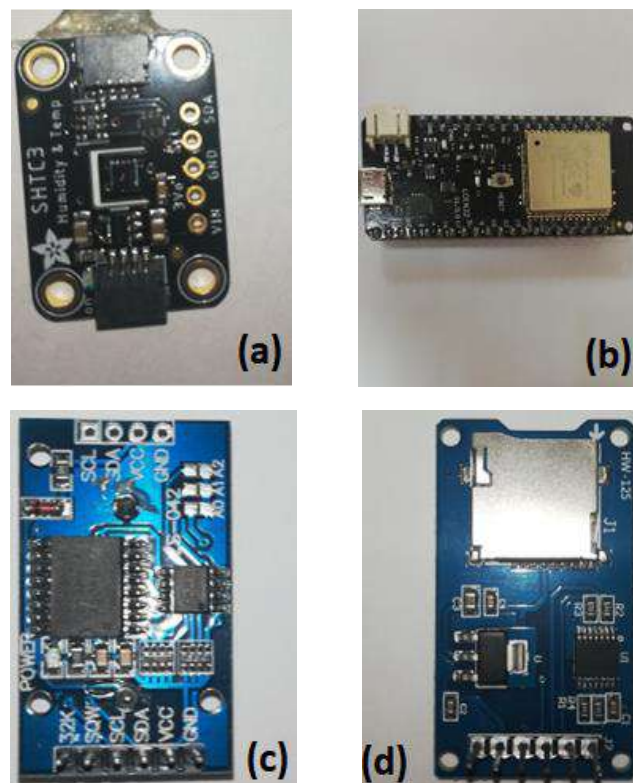


Figure 3.3. (a) Adafruit Sensirion SHTC3 temperature/humidity sensor. (b) ESP-WROOM-32 module, for functionalities coordination and wireless connection. (c) Real Time Clock SD3231 to acquire the timestamp. (d) SD card module to store the data

Table 3.3. Technical specifications for the Adafruit SHTC3 temperature and humidity sensor.

Technical Specifications	
Temperature	
Accuracy	$\pm 0.2^{\circ}C$ from 0 to $60^{\circ}C$
Resolution	$0.01^{\circ}C$
Measurement Range	-40 to $125^{\circ}C$
Humidity	
Accuracy	$\pm 2\%$ from 20 to 80%
Resolution	0.01%
Measurement Range	0 to 100%
Power Consumption (typical scenarios)	Idle: $45\ \mu A$ Sleep Mode: $0.6\ \mu A$ Measurement: $430\ \mu A$
Response Time	8 seconds
Communication Protocol	I2C



Figure 3.4. Finalized PCB, which does not include the sensor. All the parts were soldered, and the device was tested for 1 week in the lab to make sure everything will work as planned during the field experiment



Figure 3.5. The final product for a single system. The box that houses the system is an IP66 plastic box and the body of the station is a camera tripod embedded with aluminum corners that contain mounting holes, which are useful for placing the box in the desirable height.

3.6. Experimental Design

The experiment took place in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. The topology consists of 3 industrial weather stations and 30 open-source, surrounding each industrial as shown in Figure 3.6. This configuration reduces the cost of the experiment substantially since getting the same amount of industrials as open source increases the cost. The logging interval is 5 minutes, and the measurement is instant and not averaged. Due to Adafruit IO server's (discussed in Chapter 2.3.8.) limitation, since the basic no fee plan was used, only one of the 10 devices per cluster (Figure 3.6) was transmitting to the IOT server. The rest stored the data in the SD card. This configuration was made to evaluate the wireless performance of the station.

Three tripods for the open-source weather station were placed next to each industrial. Each tripod contained three radiation shields and each radiation shield contain 3 to 4 sensors and two boxes. Each box contained five PCBs, one for each system. This design was followed since the space around the industrial weather station was limited. Each sensor was placed 4 cm away from each other to avoid interference.

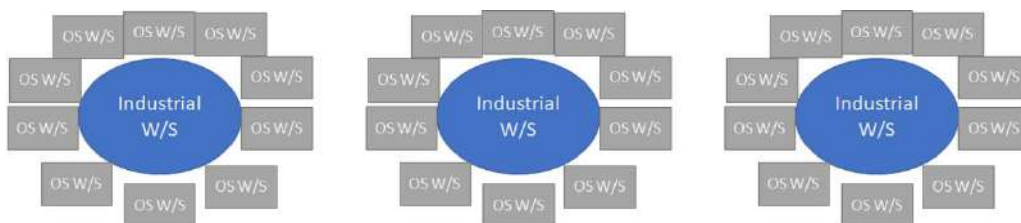


Figure 3.6. The topology of the experiment. Every industrial weather station creates a cluster of ten open-source weather stations.

Each PCB containing the open-source device, including the sensor, is defined as the experimental unit in this experiment. The tripods can be considered as a random factor which creates a cluster of sensors. The same applies for each box that contains the PCBs and each radiation shield that contains the sensors. However, these factors will not be included in the model since it will make it more complex and the variation coming from them will be included in the error term. The response of the experiment is the air temperature. Since each device was capturing temperature measurements in a 5-minute interval for 7 days, the data are considered longitudinal, and the appropriate analysis will be followed. Theoretically the devices would store 288 measurements (occasions) but due to network failure for specific times during the day not all occasions were transmitted from the industrial weather stations. Furthermore, power failures would cause the same issue for the open-source weather station since it is not equipped with a battery. Table 3.4 summarizes the total occasions recorded per day that both devices transmitted successfully. In case one of the weather stations missed a measurement, the corresponding occasion for the other devices is discarded.

The cultivation was tomato and the industrial weather stations 1 and 2 were in the same chamber close to the vegetation but the industrial weather station 3 was further away with no vegetation present. The duration of the experiment was 7 days.

Table 3.4. Common occasions for both devices recorded per day.

Days	Occasions
1	221
2	222
3	244
4	145
5	124
6	261
7	221

Table 3.5. The experimental design. The experimental unit is defined as the sensor and logger system. The data are considered longitudinal. The design is balanced.

Experimental Design	
Experimental Unit	Sensor and logger system
Repeated Measurements	Yes, averaged per time interval
Number of repeated measurements	51
Data category	Paired, longitudinal data, 7 days
Sample size	30
Balanced/Unbalanced measurements	Balanced

3.7. Method Comparison Study Techniques

Rathnayake and Choudhary (2017) proposed semiparametric modelling and analysis for longitudinal method comparison data. However, Choudhary and Nagaraja (2017a) (C & N) provided similar method where the mean functions are modeled as polynomials (2.13) instead of penalized splines. Choudhary's and Nagaraja's approach was followed, modelling the data using mixed-effects models assuming the proportional bias is 1, to produce three assessment indices. Moreover, the methodology proposed by Carasco et al. (2009, 2013) was also used to calculate CCC for longitudinal data. Last, Longitudinal Concordance Correlation (Oliveira et al., 2018) index was evaluated by implementing the corresponding R-package (Oliveira, 2020). The shared advantage for the first and the third methodology is that CCC is calculated for all the time occasions and not summarized in one value like the second method. The decision to follow these methodologies was based on the nature of the data. The present data are longitudinal, and an attempt to calculate indices and intervals was made to assess agreement (CCC, TDI and LOA) and one index to assess similarity (Fixed Bias) between the two systems. Choudhary's and Nagaraja's R code (2017b) is available to fit a mixed-effects model (discussed in Section 2.1.5), considering the occasions as repeated values. However, due to heavy computational burden, the model is unable to converge and estimate its counterparts due to the large number of occasions per day. Thus, using National Observatory of Athens (2023) summary method to report daily temperature, the days were divided into 8 time zones and temperature recordings were averaged per time-zone (Table 3.6). Then, the indices were calculated for 51 occasions. Moreover, CCC was calculated using the "cccrm" (Carrasco & Martinez, 2022) and "lcc" package (Oliveira, 2020). The difference between the three analyses was based on three main parts. First, the analysis using C & N methodology provided CCC, TDI and limits of agreement for all time occasions. However, for Carrasco's methodology CCC was summarized as one value. Oliveira's methodology provided only CCC for all the time occasions plus the longitudinal Pearson correlation and the longitudinal accuracy statistics. Second, C & N methodology fit the data using a different formulation for the mixed-effects model compared to Carrasco and Oliveira. Table 3.7 summarizes the model formulation per methodology for CCC calculation. Third, C & N methodology includes CCC, TDI and limits of agreement while Carrasco's and Oliveira's only CCC.

Table 3.6. Summary of time zones used to report daily temperature from the National Observatory of Athens.

Time Intervals	Time-Zone
00:00 - 02:59	1
03:00 - 05.59	2
06:00 - 08.59	3
09:00 - 11:59	4
12:00 - 14:59	5
15:00 - 17:59	6
18:00 - 20:59	7
21:00 - 23:59	8

Table 3.7. Model formulation for the mixed-effects model per methodology for CCC evaluation.

	C & N	Carrasco	Oliviera
Fixed Effects	method	method	method
Random Effects	subject, subject x occasions	subject, subject x method	subject
Correlation Structure	AR(1), ARMA, etc	Compound Symmetry/ AR(1)	NA
Variance-Covariate	method	NA	method
Estimation Method	ML	REML	REML

3.8. Results

3.8.1. Exploratory Analysis

The data consist of air temperature measurements taken by two kinds of devices – open-source (new method) and the industrial (reference) on a method comparison study of 30 open-source weather stations and 3 industrials. There are 30 subjects, 2 methods, 8 time zones and 7 days. Each day is divided into 8 time-zones as described in Section 3.6. Thus, 3360 measurements are expected to be counted. However, there were 3060 of measurements in total, since not all days include all the occasions. The loss of data is based on power failures or unsuccessful transmission from the industrial. Table 3.8 displays a summary for the number of time-zones per day.

Trajectories of the differences per subject are displayed in Figure 3.7. The x-axis presents the occasions of the experiment. The first digit defines the day and the second the time-zone. For example, occasion 11 represents the first day and first time zone. Occasion 24 represents the fourth time-zone of the second day. There are 51 occasions in total. A trend throughout the days is obvious. The differences start with around -3 units, decrease until they reach zero then increase, and reach a peak of approximately 4 units. The open-source device underestimates then overestimates the temperature. The positive peaks are noticed in the 5th to 6th time zone which corresponds to 12:00 am until 18:00 pm where the highest temperatures are measured. Figure 3.8 presents the scatterplot of the industrial vs open-source device per time-zone and the identity line is included in every plot. The open-source device underestimates the temperature for time-zones 1, 2, 8 and overestimates for time-zones 5 and 6. For time-zones 3,4 and 7 both effects are noticed. Figure 3.9 presents the Bland-Altman plot without the limits of agreement per time-zone. A trend is visible for time-zones 3,7 and 8. Figure 3.10 and 3.11 display boxplots of the temperature differences per time zone and per time-occasion respectively.

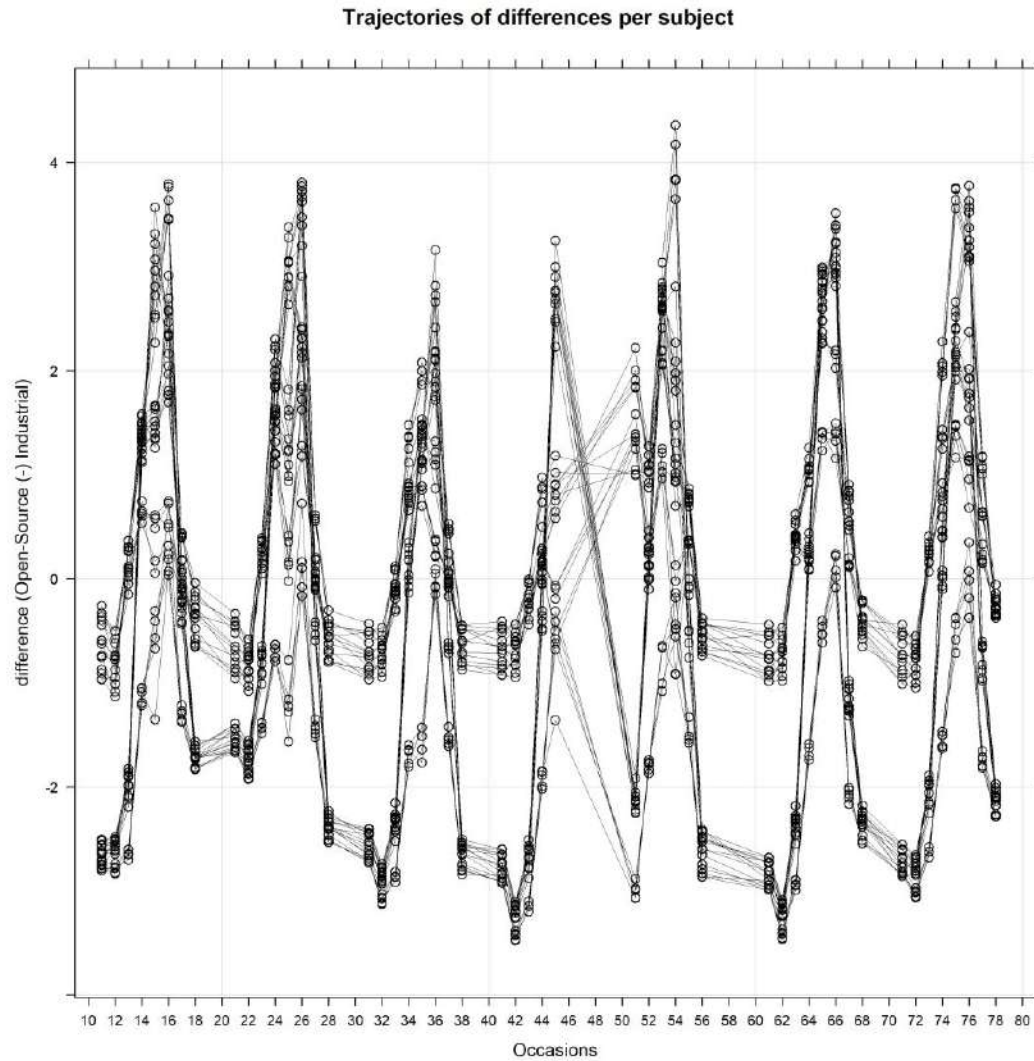


Figure 3.7. Trajectories of the differences per subject. Occasion 11 represents the first day and first time zone. Occasion 24 represents the fourth time zone of the second day. There are 51 occasions in total. A trend throughout the days is obvious. The differences start with around -3 units, increase, and reach a peak of approximately 4 units. Thus, the open-source device underestimates then overestimates the temperature.

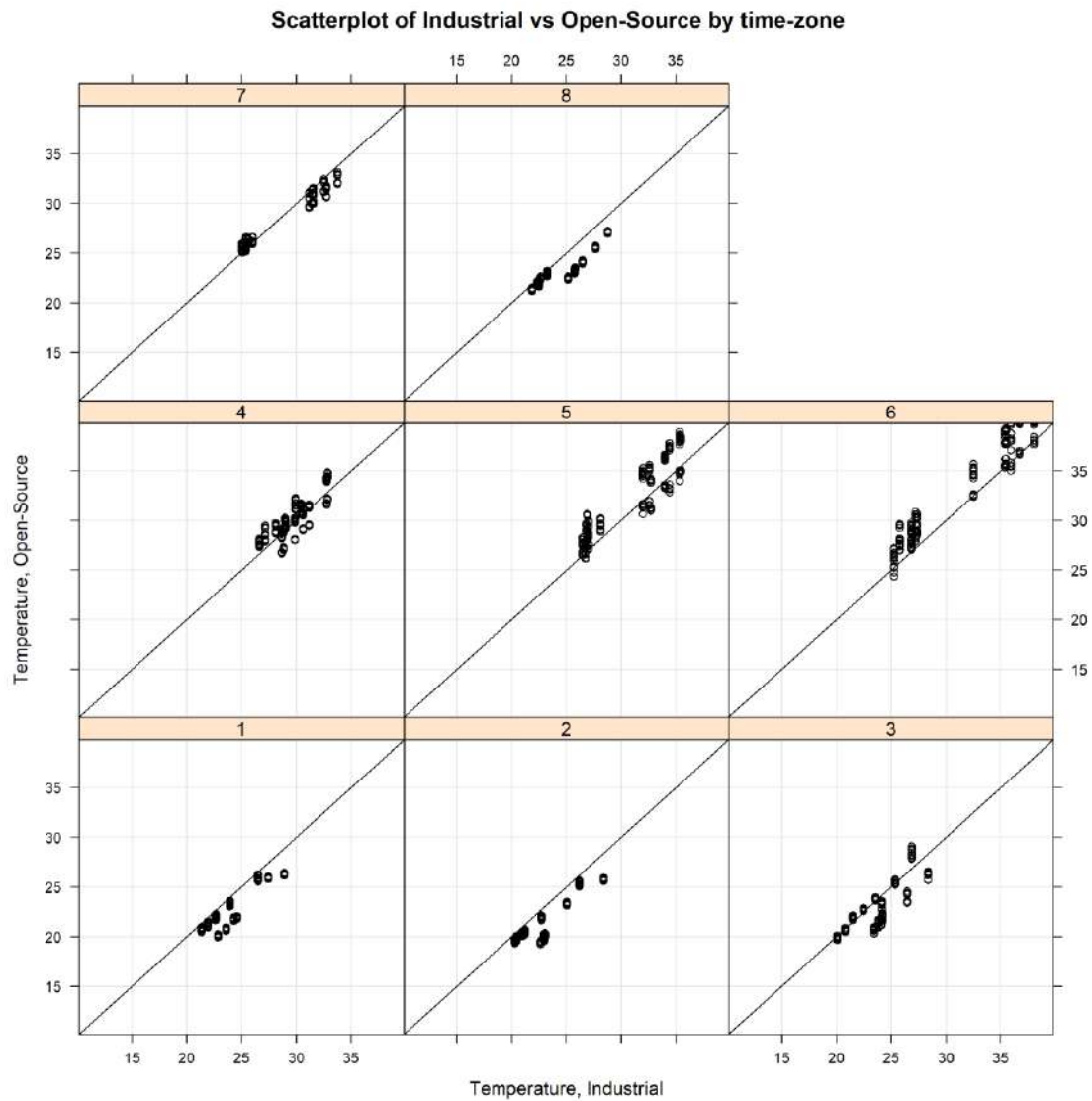


Figure 3.8. Scatterplot of Industrial vs open-source by time-zone. The open-source device underestimates the temperature for time-zones 1, 2, 8 and overestimates for time-zones 5 and 6. For time-zones 3,4 and 7 both effects are noticed.

Table 3.8. Time-zones per day. Days 1,2,3,6,7 include measurements for all the time-zones. However, day 4 and 5 include 5 and 6 time-zones respectively.

Days	Time-Zones
1	8
2	8
3	8
4	5
5	6
6	8
7	8

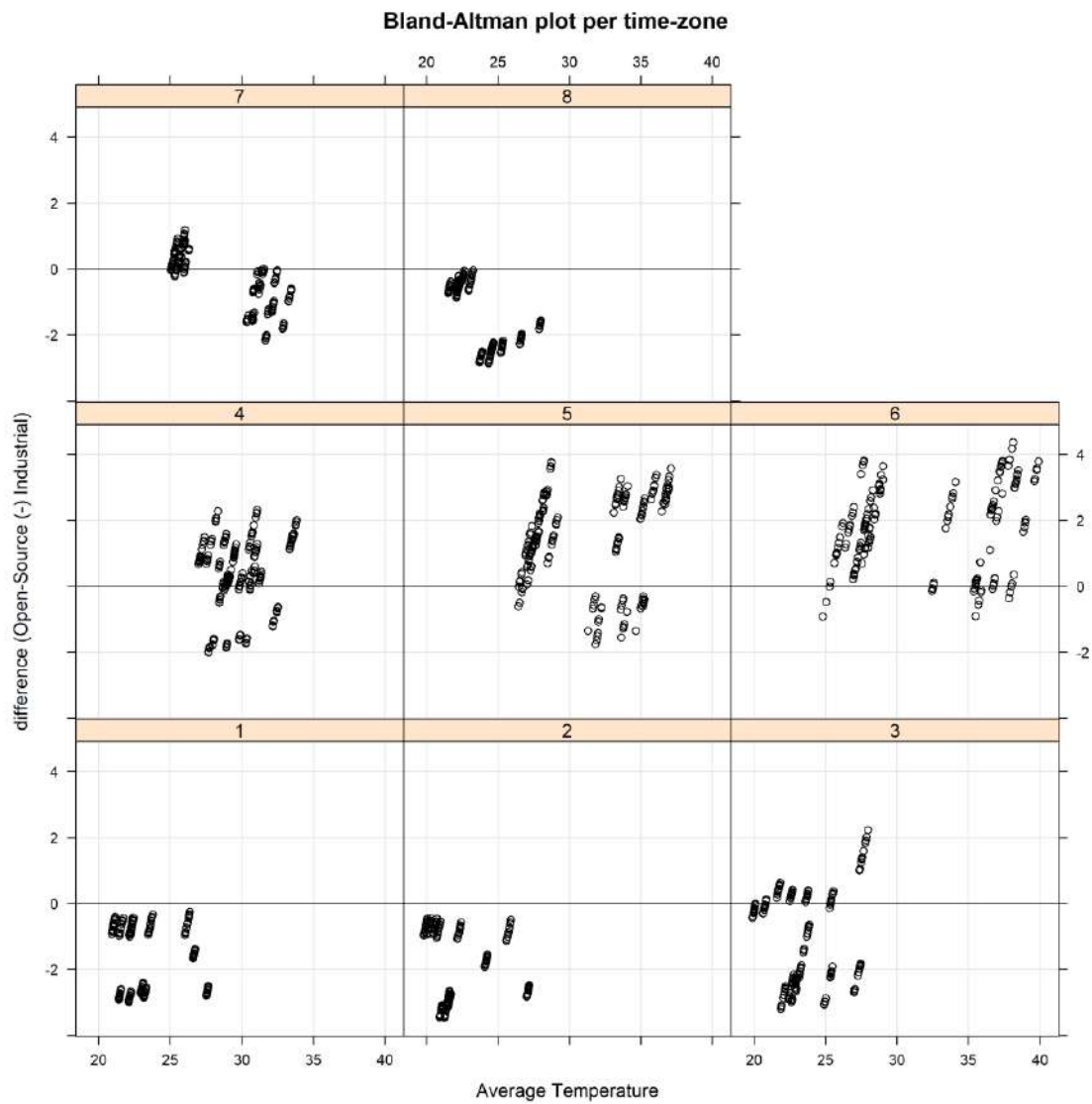


Figure 3.9. Bland-Altman plot without the limits of agreement per time-zone. A trend is visible for time-zones 3,7 and 8.

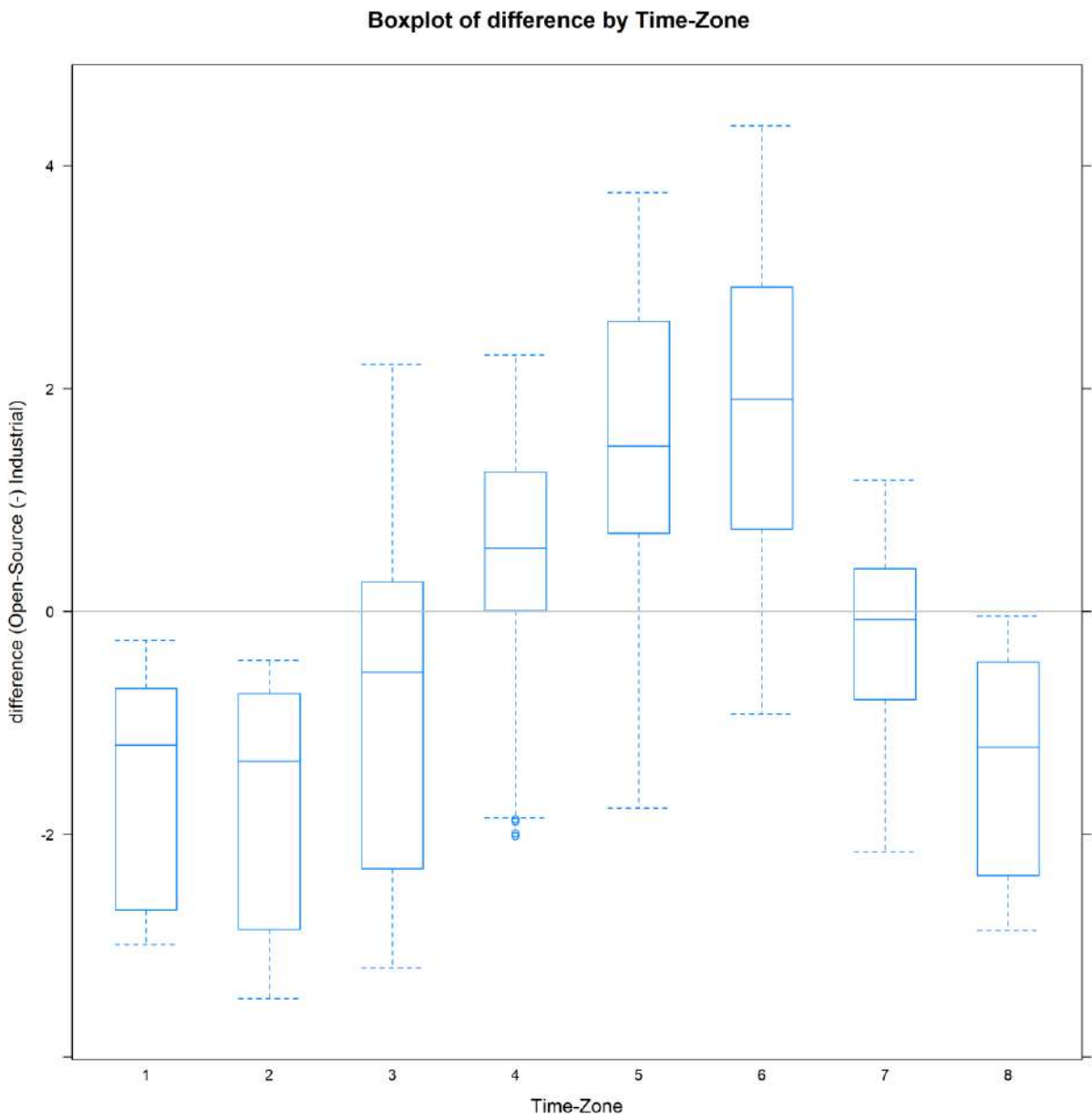


Figure 3.10. Boxplot of the temperature differences per time zone.

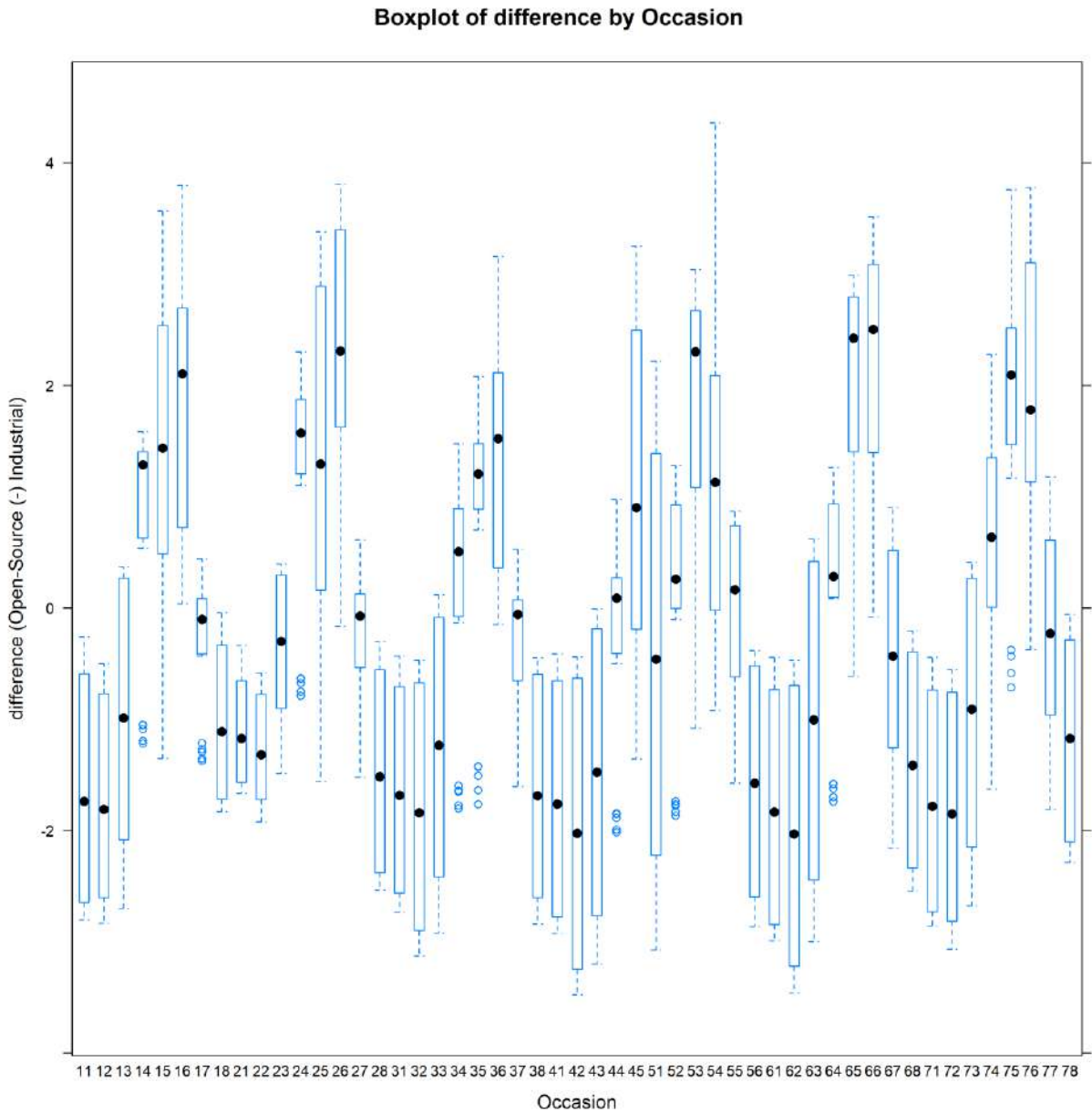


Figure 3.11. Boxplot of difference by occasion. A cubic polynomial provides a reasonable fit.

3.8.2. Agreement Analysis

A cubic polynomial provides a reasonable fit in Figure 3.11. Thus, a degree of $q = 3$ is chosen for the mean functions (2.13). Model (2.12) is fit to the temperature data. However, regardless of the choice for the correlation structure, the model either did not converge, or the standard deviation of the errors was extremely high. This is a sign of inappropriate model fitting. The analysis continued using Carrasco's methodology. The model formulation considered AR(1) as the correlation structure, with correlation parameter estimate $\varphi_1 = 0.5295$. The random effects formulation was the same as the previous model. The model was fit using restricted ML instead of ML as for the previous model. Table 3.9 summarizes CCC estimate and the corresponding confidence intervals. The agreement is considered low since CCC estimate is 0.3346 and its corresponding confidence interval is [0.2642, 0.4016]. Investigating model diagnostics, (Figure C1, Appendix C), there is evidence of great deviation from normality thus the estimates are not considered robust.

The analysis continues with the implementation of the *lcc* function from the “lcc” R-package by Oliveira et al. (2020). Using the formulation described in Table 3.7, LCC is evaluated for all time occasions. Apart from LCC, the function estimates the longitudinal accuracy and longitudinal Pearson correlation. These indices are considered as the accuracy (LA) and precision (LPC) coefficient respectively (discussed in Section 2.1.5.). Table A.10 (Appendix A) summarizes LCC, LPC and LA per occasion. Moreover, Figure 3.12 presents the estimates and 95% confidence intervals for the LCC, LPC and LA throughout the time occasions. For LCC, the agreement is considered low. LCC ranges from 0.1146 to 0.2236. LPC ranges from 0.1234 to 0.2243 which reveals high deviation from the line of best fit. LA ranges from 0.9286 to 0.9969 thus the marginal distributions of the two measurement methods are close. The points on the plots represent the concordance correlation coefficient considering independent measurements taken over time. There is a significantly high difference between CCC and LCC.

Low values on the LPC and high values on LA are evidence of accurate but not precise measurements. Thus, low precision is a possible source of disagreement between the two methods.

Table 3.9. CCC estimate and its corresponding confidence intervals. The agreement is considered low.

CCC estimated by variance components			
CCC	LL CI 95%	UL CI 95%	SE CCC
0.3346	0.2642	0.4016	0.0351

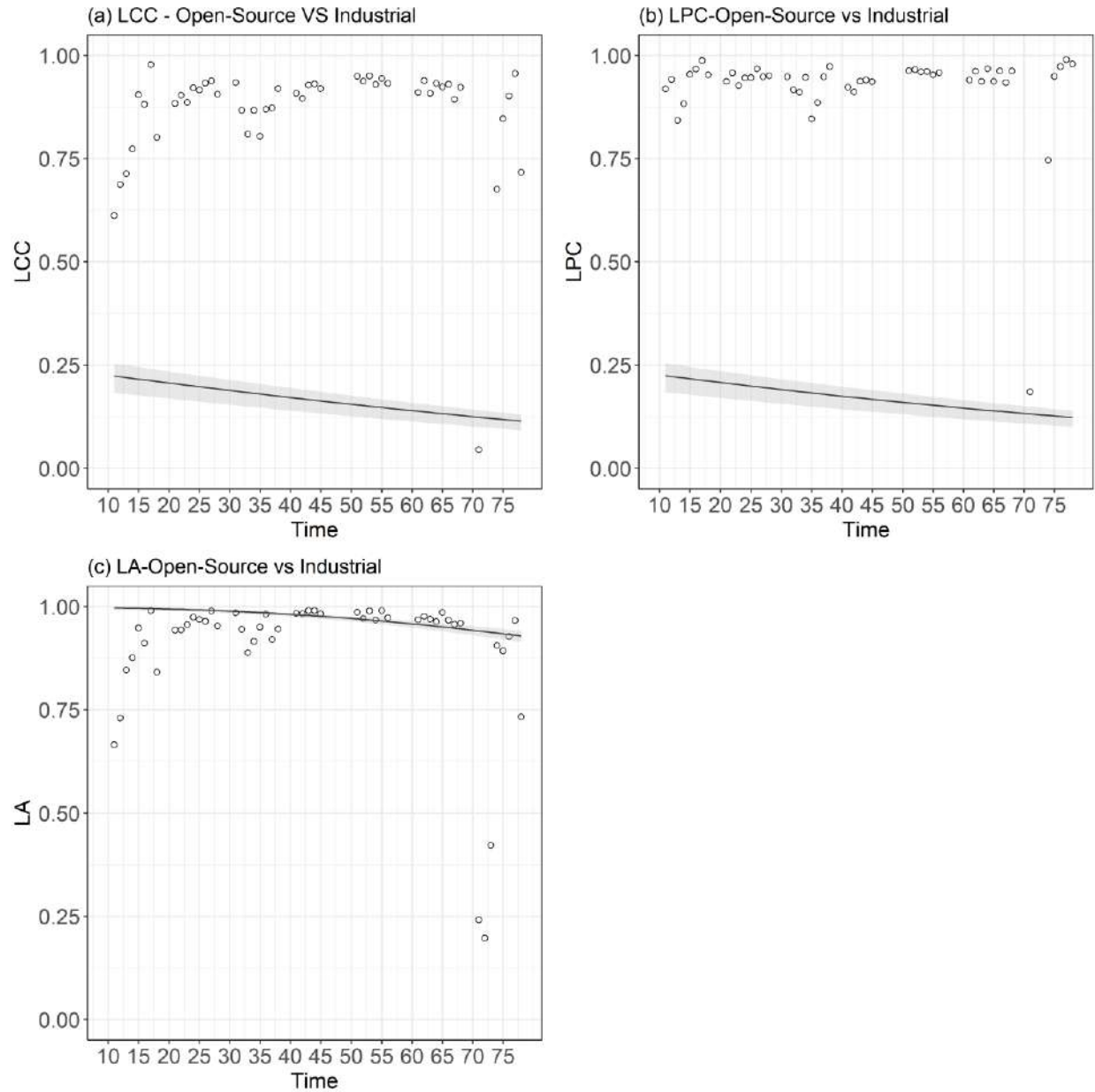


Figure 3.12. (a) Estimates and 95% confidence intervals for the longitudinal concordance correlation. LCC starts from 0.2236 for occasion 11 and reaches 0.1146. As the occasions increase LCC reduces. (b) Estimates and 95% confidence intervals for the longitudinal Pearson correlation. (c) Estimates and 95% confidence intervals for the longitudinal accuracy. The points on the plots represent the concordance correlation coefficient considering independent measurements taken over time.

3.9. Discussion

The construction of open-source devices for environmental remote monitoring forms a new trend for affordable and fully configurable industrial and experimental equipment. Solely, the construction and functionality validation are not enough to evaluate the devices' performance and reliability. Thus, a formal methodology that covers the experimental design and the statistical analysis was implemented to investigate the areas of agreement and disagreement between a reference and a new device. This is the first attempt to evaluate such devices under specific conditions (longitudinal data) using statistical methods to assess agreement and there is certainly space for improving the methodology or extending to other conditions.

Methods for assessing agreement for longitudinal data with a great number of repeated measurements using time series can extend the current approach. The first attempt to fit the current data to mixed-effects models using Choudhary's and Nagaraja's methodology failed to produce a valid model, regardless of the different formulations of the correlation structure. Moreover, covariates such as humidity and cluster effects (sensors hosted in the same radiation shield) can be considered as possible sources of disagreement. Considering this information further actions can be considered to improve the agreement and identify possible sources of disagreement between the reference and the new device.

Carasco's methodology fits a proper model for the data. The results indicated low agreement since CCC estimate is 0.3346 and its corresponding confidence interval is [0.2642, 0.4016]. However, possible covariates like humidity or location of the weather stations should be taken into consideration in future research.

Olivier's methodology produced a valid model and LCC, LPC and LA were estimated. However, there was no specific correlation structure defined for the repeated measurements. The author's methodology revealed low agreement (LCC ranged from 0.1146 to 0.2236). LPC and LA provided an explanation about the inadequate agreement. Since LPC is low [0.1234, 0.2243] and LA is high [0.9286, 0.9969], lack of precision is one of the possible sources of disagreement.

For the hardware part, upgrades and additions can be considered for the open-source device, such as power efficiency, adding solar panels and sleep modes to save energy, expand the environmental characteristics measured such as solar radiation sensors (PAR sensors, pyranometers), soil moisture and conductivity sensors etc. This task is supported by the huge amount of documentation and guides available online.

Last, the wireless technology can be exchanged between Wi-Fi and LoRa, depending on the conditions surrounding the area of coverage. If the infrastructure (field or greenhouse) the logger is established is not equipped with Wi-Fi, LoRa wireless technology can be used (discussed in Section 2.3.5), equipped with LTE or GSM cellular modules. Such applications are discussed in Chapter 4 and 6 where power outlets are either available, thus using more power consuming cellular applications (LTE) or not available, using less power consuming cellular technology (GSM).

4. A method comparison study between an open-source and an industrial system for measuring water content from perlite slabs.

4.1. Goal of the study

Perlite is a generic name for an amorphous volcanic rock. Markoska et al. (2018) describe Perlite's features and applications. When water is entering its structure vaporizes and escapes. This procedure causes its expansion to 7-16 times its original volume. Perlite's properties are useful for commercial application since it has low bulk density, low thermal conductivity, high heat resistance etc. Due to its non-toxic nature, it is applied in hydroponics as a substrate for various cultivations.

Perlite provides excellent water retention and suitable aeration for root growth in plants. Perlite's particle size can influence the water retention capacity of the expanded perlite. Finer particles are packed together and provide an increased surface area thus retaining more moisture. On the other hand, coarser particles provide less surface area per volume and retain less moisture but release it more rapidly. Thus, according to the needs of the user, regarding moisture, particle size can vary (Lee et al., 2022).

According to Burés et al. (1997a; 1997b) the available volumetric water content in perlite with 0 – 4 mm diameter is 13.6% and unavailable, 36.5 %.

Since perlite's water content is non-uniform throughout the substrate, models are produced to predict it (Lee et al., 2022). Thus, it is important to place multiple water content sensors simultaneously to different location-points of the perlite slab, to collect data and estimate the model's coefficients. However, one perlite slab is just one sample. Multiple samples are needed to collect data under the same conditions. Furthermore, soil sensors are expensive and usually a logger is needed which might not support wireless communication or is substantially expensive to add one. A solution to this issue is the development of open-source loggers that interface with industrial soil sensors. Moreover, the addition of any wireless technology such as Bluetooth, Wi-Fi, LoRa, GSM or LTE to open-source loggers is considered a trivial task when it comes to open-source devices. This is possible since open-source communities provides a vast majority of documentation and straightforward tutorials.

Measuring more than one soil characteristic at the same time has proven to be more insightful than measuring only soil water content. Measuring other soil characteristics, such as electrical conductivity, temperature, water content and pH simultaneously, is a useful tactic since prediction models might include covariates that could improve their specificity and sensitivity.

Industrial and reliable sensors usually need their corresponding logger. However, the industrial loggers are limited to specific functions and are mostly compatible with the manufacturer's sensors. Further, industrial loggers are expensive with restricted configurability.

The present research focuses on two main goals. First, the design and construction of an open-source logger for soil moisture sensors but not limited to them. Second, the evaluation of the agreement between the open-source system (including the sensor) and its corresponding industrial. Other sensors can be plugged into the logger using one or more of the communication protocols discussed in Section 2.3.4. If the sensor's communication protocol is not supported by the development board, then it can be simulated using electronic parts that are usually provided with a breakout board making it easy to communicate with the MCU. A Modbus RS-485 soil, temperature, conductivity and pH sensor was used to take measurements from perlite slabs during an aquaponics experiment in the greenhouse facilities of the

Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. Tables B1 and B2 (Appendix B) summarize the chemical composition and the physical properties of the perlite slab during the experiment. The product's name is Geoflor Hydro and the distributor is Perlite Hellas S.A. According to the company, the product is the traditional coarse perlite for all agricultural uses. It provides suitable aeration, water absorption and is chemically stable, clear of any pest seeds and diseases. The cultivation during the experiment was cucumber.

4.2. Design of the Device

The device was designed as a portable logger that can be used for in situ measurements and at the same time as laboratory equipment. Thus, it is battery operated or can be powered via an electrical outlet using micro-USB to USB-A cable. When the battery is charged the device will be powered directly via the electrical outlet. Furthermore, it was purposely fitted in a rather small enclosure box to increase portability for field measurements. Wireless connectivity via LoRa was embedded via the development board and a gateway is needed for cloud storage. In case there is no available internet connection a portable 4G/LTE module can be plugged into the gateway. Alternatively, the gateway can connect to the IOT server using Wi-Fi or ethernet). Figure 4.1 displays the final product.

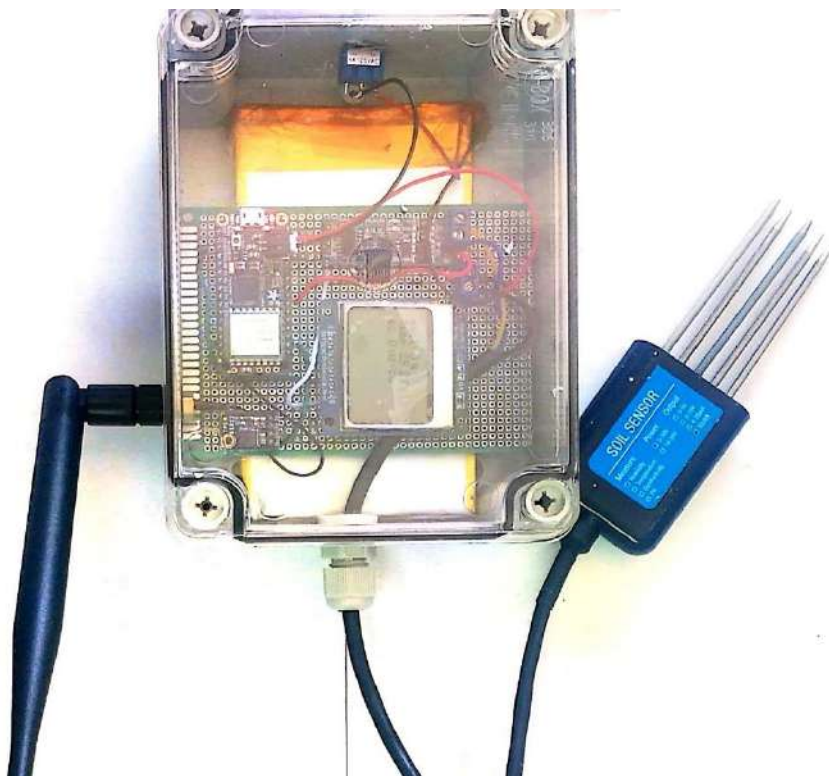


Figure 4.1. The open-source logger with the water content, temperature, electrical conductivity and pH sensor.

4.2.1. Hardware

Soil Moisture Sensor

The sensor (Figure 4.2) is a soil moisture, temperature, electrical conductivity and pH sensor probe constructed by ComWinTop (ComWinTop, 2023). The trait of the perlite slab under investigation is water content. The sensor's principle is based on the measurement of the dielectric constant of the substrate. When the moisture in the substrate changes, then a considerable amount of the dielectric changes. Then, it can measure the substrate's volumetric water content. The volumetric water content (%) is calculated by multiplying the volumetric water content by bulk density. The sensor measures a cube of 7^3 cm^3 since the probe is 7 cm long. It is suitable for soil, water, nutrient solution, and other substrates immersion for long-term monitoring. The sensor is equipped with automatic temperature compensation. It is durable and anticorrosive since it is made from flame-retardant epoxy resin. The communication protocol with the MCU is RS-485. Since the MCU is not equipped with an RS-485 driver, an external module is used. The sensor's manufacturer offers an executable file where the user can interact with the sensor (Figure 4.3). A USB to RS-485 Waveshare industrial isolated converter (Waveshare, 2018) was used to connect the sensor with the computer for configuration and debugging purposes. The sensor's technical specifications are summarized in Table 4.1.



Figure 4.2. Soil moisture, temperature, electrical conductivity and pH Sensor by ComWinTop. (source: ComWinTop official webpage).

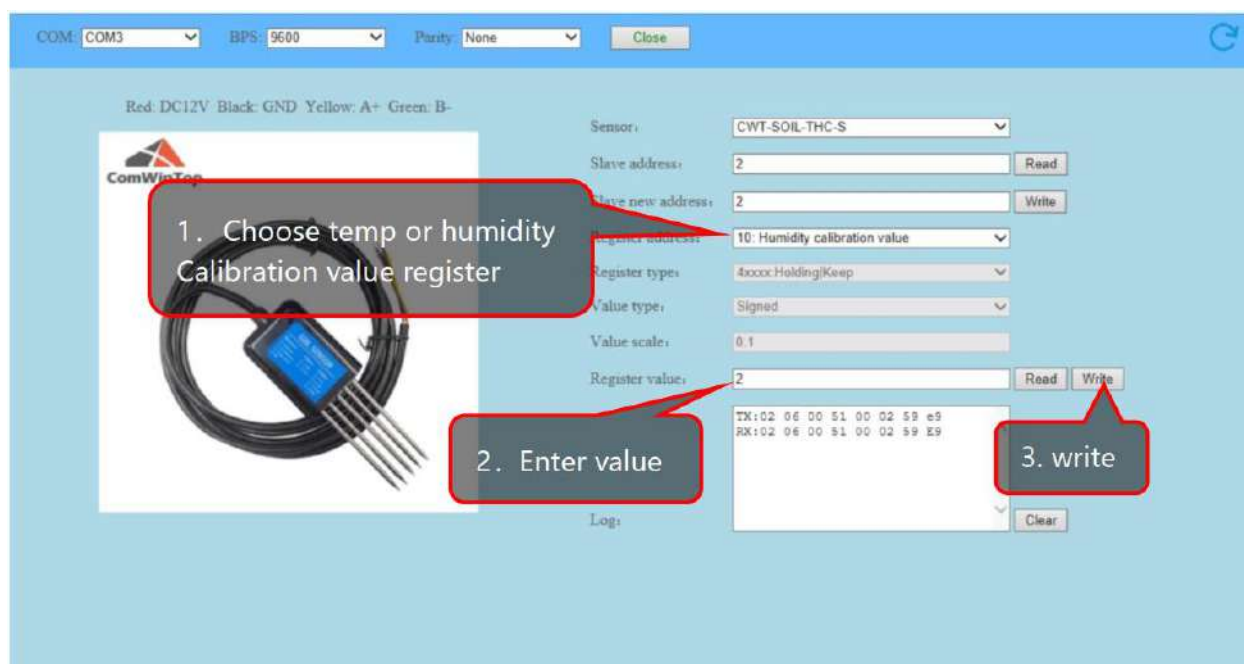


Figure 4.3. Sensor's interface. The program is available by ComWinTop company (source: ComWinTop official webpage).

Table 4.1. Technical Specifications for the Industrial Soil Moisture, Temperature and Electrical Conductivity sensor.

Technical Specifications	
Power Supply	5~30 VDC
Operating Temperature	-40°C ~ 80°C
Interface	RS-485
Current Consumption	6mA @ 24V DC
IP Rating	IP68
Soil Moisture	
Range	From completely dry to fully saturated (From 0 % to 100 % of saturation)
Accuracy	±2 % (0~50 %) ±3 % (50~100 %)
Long term stability	≤ 1 % RH / year
Response Time	≤ 4 s
Soil Temperature	
Range	-40°C ~ 80°C
Accuracy	±5°C (25°C)
Response time	≤ 15 s
Long-term stability	≤ 0.1 % °C / year
Electrical Conductivity	Built-in temperature compensation sensor

Range	0~20,000 $\mu\text{S} / \text{cm}$
Accuracy	$\pm 3 \%$ for 0 – 10,000 $\mu\text{S} / \text{cm}$
	$\pm 5 \%$ for 10,000 – 20,000 $\mu\text{S} / \text{cm}$
Long Term Stability	$\leq 1 \%$ $\mu\text{S} / \text{year}$
Response time	$\leq 1 \text{ s}$
pH	
Range	3 – 9 pH
Accuracy	$\pm 0.3 \text{ pH}$
Long Term Stability	$\leq 5 \%$ / year
Response Time	$\leq 10 \text{ s}$

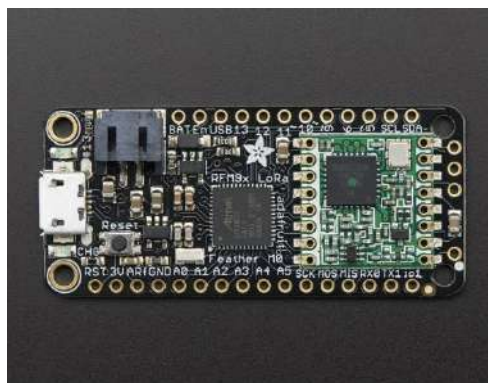


Figure 4.4. The development board, Adafruit feather M0 RFM95 LoRa Radio (900 MHz). (source: Adafruit official website)

The development board

The development board is an Adafruit feather M0 RFM95 LoRa radio module (Figure 4.4) (Fried, 2023f) for European bands (900 MHz). The microcontroller is an ATSAMD21G18 ARM Cortex M0 processor, clocked at 48 MHz at 3.3 Volts logic. Since it is a feather product, it has a 3.7 Volt Lithium polymer connector and built-in battery charging. Technical Specifications can be found in Appendix Table A7. More details about the pinout can be found in Appendix Figure A1. A uFL connector can be soldered on the board and the user can choose from a wide variety of LoRa antennas. According to Adafruit's official website the range can be between 2 km line of sight by soldering an 8.2cm wire antenna, or 20 km by setting tweaking and directional antennas. The manufacturer provides a detailed guide to run a simple LoRa sketch using the board. The code can be adapted to the user's needs.

MAX485 TTL to RS485 module

MAX485 (Maxim Integrated Products, 2014) is the intermediate between the RS-485 Sensor and the MCU. It is a low-power transceiver for RS-485 communication protocol. It minimizes Electro Magnetic Interference (EMI) and reduces reflections caused by improperly terminated cables, thus allowing for error-free transmission up to 250 Kbps. The module allows half-duplex communication. In stand-by mode

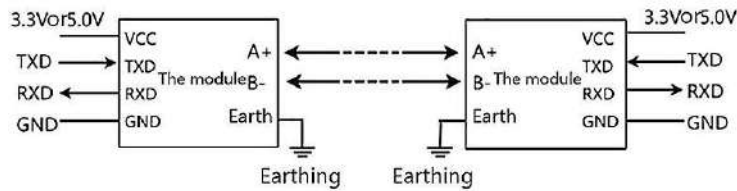
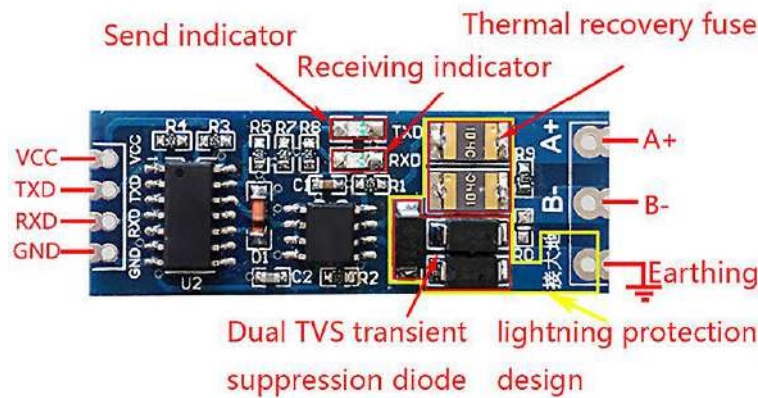


Figure 4.5. Uart serial port to RS485 Converter module and its connection diagram with the MCU (source: fruugo (2023)).

the module draws 300 μA and it allows for 32 receivers on the same bus. Its supply voltage is 5 Volts. Figure 4.5 presents the breakout board for the module and the connection with the MCU. The communication protocol between the MCU and the RS-485 module is UART (discussed in Section 2.3.4.).

Adafruit Miniboost 5V (TPS61023)

The sensor's operating voltage ranges from 5 to 30 Volts DC. The development board's output while it operates on battery is 3.3 Volts. Thus, a boost converter is needed to increase the voltage to at least 5 Volts. The TPS61023 (Texas Instruments, 2019) is built on a breakout board from Adafruit. Its input is 2 to 5 Volts and the maximum current drawn from a LiPoly battery while on nominal voltage is 1100 mA. The module provides a stable 5.2 Volts.

Nokia 5110 LCD Monitor and the PCD854 controller

Nokia 5110 LCD monitor was first manufactured for the NOKIA cell phone company for the respective cell phone. Its controller, PCD8544 is a low power CMOS LCD driver designed to drive a graphic display of 84 columns and 48 rows. Its supply voltage ranges from 2.7 to 3.3 Volts making it ideal to work with a feather development board. To connect to the MCU, 5 pins will be dedicated. A backlight is available, but it is not recommended for low power applications. The choice of the lcd monitor was based on its low power consumption. It is easy to integrate as it requires an SPI-like serial bus interface to communicate with the MCU and displays text or pictures using 84x48 pixels.

LoRa Gateway

The LoRa gateway used in this research was developed using a Raspberry pi 3 B+ with a RAK2245 Pi Hat (Figure 2.15). The greenhouse is equipped with internet connectivity via Wi-Fi but is unreliable. Thus, a 4G/LTE USB dongle is connected to the Raspberry pi to provide continuous connectivity with the server. A detailed guide to establish the RAK2245 Pi Hat is provided by the manufacturer of the RAK2245 Pi Hat (RAK ,2023a). Figure 4.6 displays the LoRa Gateway including the 4G/LTE dongle.

4.2.2. Software

The sketch has three main functions. First, the measurement function, which is responsible for measuring the soil's water content, temperature, electrical conductivity and pH. Second, the communication with TTN (The Things Network) function to upload the data to ThingSpeak (discussed in Section 2.3.8). Third, the activate sleep mode for power saving function. The MCU communicates with the sensor via UART.

The measurement function uses the microcontroller's UART communication protocol. The microcontroller sends an eight-byte code (in hexadecimal form) to request data from the sensor. The sensor responds and sends eleven-byte code (in hexadecimal form). The manufacturer gives an example of how to convert from hexadecimal to decimal. Figure 4.7 displays a response frame from an example provided by the manufacturer. The fourth and fifth byte corresponds to the water content. The sixth and seventh to the temperature, the eighth and ninth to the electroconductivity. To convert the hexadecimal to decimal in the sketch, the following lines of code are used.

```
soil_int = int(values_code[3]<<8|values_code[4]);  
soil_temp = int(values_code[5]<<8|values_code[6]);  
soil_ec = int(values_code[7]<<8|values_code[8]);  
soil_pH = int(values_code[9]<<8|values_code[10]);
```

The values are stored in global variables which are defined as real numbers.

The communication function is based on the "LMIC" framework (IBM et al., 2015) provided by IBM. It supports LoRaWAN Class A devices which are compatible with the development board's radio (HopeRF). For the LMIC to run on top of Arduino environment the library "hal.h" is used. To connect to TTN two keys are needed. The API key, which is provided by the TTN when the user adds a new application and the DevEUI (Device Extended Unique Identifier) which is assigned by the manufacturer of the LoRa module. The payload sent to the gateway is received by TTN. A complete guide to send data to TTN using Adafruit Feather Lora M0 is available by LoRa Vsb Cz (2023). TTN receives the payload, decrypts it using a payload decoder and forwards it to Thingspeak using an integration. Integrations are used to process and act on data by triggering events. To allow TTN to communicate with Thingspeak, a webhook is created. A detailed guide can be found on The Things Stack (2023). The Arduino sketch for this research and the additional code for the TTN payload formatter is available on GitHub (Bataka, 2023).

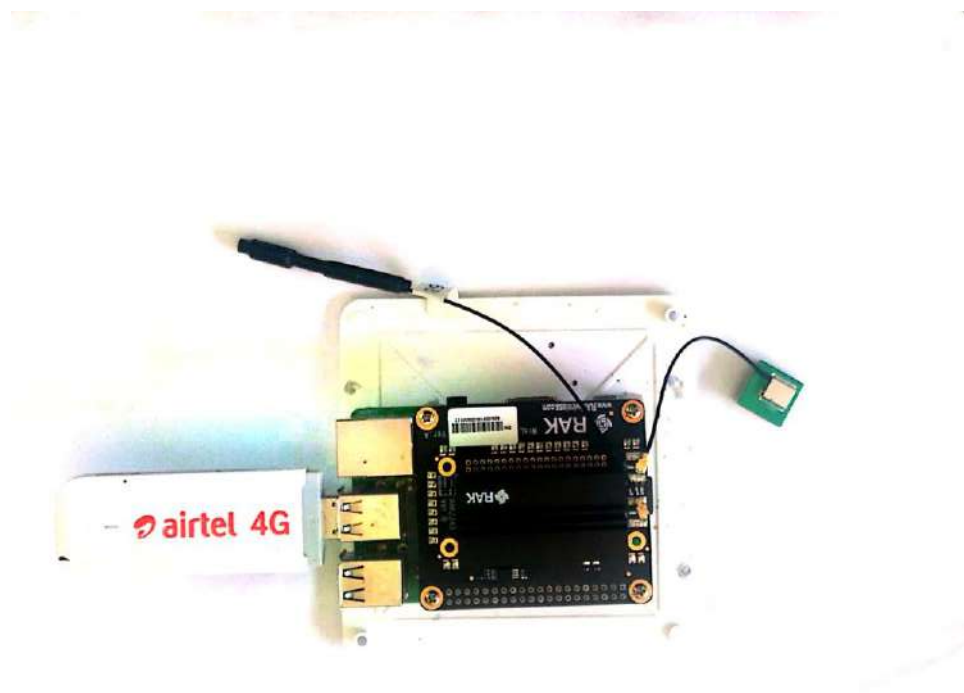


Figure 4.6. Raspberry pi 3 B+ combined with a RAK2245 Pi Hat for LoRa gateway deployment. To connect to the server a 4G/LTE dongle is added with a SIM card.

Sensor responds:

Address	Function Code	Number of byte	humidity	temperature	conductivity	PH	Error Check (Lo)	Error Check (Hi)
0x01	0x03	0x08	0x02 0x92	0xFF 0x9B	0x03 0xE8	0x00 0x38	0x38	0x75

Temperature: FF9B H= -101 => temperature= -10.1°C

Humidity: 292 H= 658 => humidity= 65.8%

Conductivity: 3E8 H= 1000 => Conductivity = 1000 us/cm

PH: 3E8 H= 56 => PH= 5.6

Figure 4.7. The example of a response frame provided by the manufacturer of the sensor. The fourth and fifth byte corresponds to the water content. The sixth and seventh to the temperature, the eighth and ninth to the electrical conductivity and the tenth and eleventh corresponds to pH.

4.2.3. Cost of the Device

The cost of the device is summarized in Table 4.2 and for the LoRa gateway in Table 4.3. The list is based on end of summer 2023 prices for local Greek vendors, mouser (Europe) and ComWinTop (Chinese vendor).

Table 4.2. Summary of the cost for the construction of the open-source device.

Cost per part and total (in Euro, no VAT included)	
Adafruit Feather M0 LoRa	32.85
MAX485 TTL to RS-485 module	0.94
Soil,EC,Temperature Sensor	34
Adafruit Adalogger FeatherWing	8.41
Enclosure Box	6.37
Switch	0.65
LiPoBattery 5000 mAh	16.13
Cable Gland PG-07	0.2
Antenna	15.59
Nokia 5110 module	4.68
uFl connector	0.67
Adafruit MiniBoost 5V @ 1A – TPS61023	4.8
Wires	2
PCB	2.5
Total	129.39

Table 4.3. Summary of the cost for the construction of the LoRa gateway.

Gateway Cost (in Euros, Vat included)	
Raspberry Pi 3 B+	50
RAK2245 pi Hat	137.09
Huawei E3372H-320 4G Mobile Router	62.99
Total	250.08



Figure 4.8. HH150 Handheld Meter and SM150 Soil Moisture Sensor used as the industrial device for the method comparison study.

4.3. The reference device

The industrial device (Figure 4.8) used for the method comparison study is a Handheld Meter (HH150) and SM150 Soil Moisture Sensor. The device's technical characteristics are listed in Table A8 (Appendix A). The cost of the device is around 760 Euro including Vat, excluding shipping costs (end of August 2023 German market prices).

4.4. Experimental Design

The perlite slabs are displayed in Figure 4.9, 4.10 and 4.11. A circle of 5 cm diameter was drawn on every slab (Figure 4.9). Each circle was 10 cm away from the planting area. All the measurements were taken inside the circle to minimize the water content variability. However, location bias could be still present regardless of the sensor's precision. This will be included in the error term. All the slab's rows are assumed homogeneous. The selection of the sample was performed using R and its *sample* function. The R script

for this experiment is available in GitHub (Bataka, 2023). All the slabs were numbered, and 100 numbers were selected using *sample's* output. Then, a coin was flipped to choose the device that will take the first measurement. Repeated measurements were taken using a slightly different procedure for each device. For SM150, after insertion, the measurement button was pressed, a measurement was taken and then for the next repeated measurement the button was pressed again. For the open-source device, after insertion, a measurement was taken, then the device was switched off and on and another measurement was taken. Another way to take the repeated measurements was to insert the sensor, then re-insert it in the same spot. However, once the perlite particles are disturbed, there might be bias if one attempts to remeasure the same spot.

A preliminary experiment was performed to investigate the data and the experimental conditions. The analysis revealed proportional bias significantly different from one. Thus, the mixed-effects model formulation was inappropriate to fit the data. The approach followed after the examination of the preliminary data was Taffé's (2018), since the author is using measurement error models. Taffé's approach requires a sample size of 100, 10 to 15 repeated measurements from the reference and one or more from the new device. Thus, 10 repeated measurements were taken for the SM150 and 3 for the open-source device. Figures 4.10 and 4.11 display the SM150 and the open-source device respectively. Table 4.4 summarizes the experimental design.

Table 4.4. The experimental design details summarized. The experimental unit is the perlite slab. There were 3 repeated measurements for the open-source device and 10 for SM150 per slab. The data are considered unlinked since the repeated measurements are not paired. The sample size is 100 and the data are unbalanced.

Experimental Design	
Experimental Unit	Perlite slab
Repeated Measurements	Yes
Number of repeated measurements	3 for OS, 10 for SM150
Data category	Unlinked
Sample size	100
Balanced/Unbalanced repeated measurements	Unbalanced repeated measurements



Figure 4.9. A circle of 5 cm diameter was drawn on every slab. Each circle was 10 cm away from the planting area. All the measurements were taken inside the circle to minimize the water content variability.



Figure 4.10. HH150 handheld meter and SM150 soil water content sensor. All the measurements are taken with the sensor inserted inside the circle.



Figure 4.11. Water content sensor connected to the open-source device. The measurement is taken inside the circle, next to the one taken from SM150. Its measurement marks are obvious on the left side of the circle.

4.5. Agreement and Similarity Techniques

The analysis to evaluate the agreement and similarity between SM150 consists of two steps. First, exploratory data analysis was performed to present the data and investigate the model's assumptions. The preliminary experiment revealed a proportional bias of 0.335 which violates the mixed-effects model's assumption of similar proportional bias between the two devices. Thus, the measurement error model approach was followed. Specifically, Taffé's approach (2018) was followed fitting (2.25 – 2.27). For the agreement and similarity analysis, "methodCompare" package was implemented. Bland-Altman plot with confidence intervals for the LOAs was also produced using "Blandr" R-package. Moreover, for the exploratory analysis parts of Choudhary's and Nagaraja's R-script (2017b) was used.

4.6. Results

4.6.1. Exploratory data analysis

Figure 4.12 displays a boxplot for the water content (%) measurements per device. SM150's measurements are significantly higher than the open-source's. Moreover, the plot reveals outliers for both devices.

The trellis plot (Figure 4.13) for the soil water content (%) measurements reveals several important findings. The vertical axis is divided into rows and each row displays all the repeated measurements for

both devices using method-specific colors. Pink color represents the measurements for the open-source device while blue represents the measurements for SM150. The open-source device underestimates the soil water content for most of the subjects by a non-constant quantity. The between-subject variation is higher compared to the within-subject variation.

Figure 4.14 displays a scatterplot for the water content (%) measurements of the SM150 versus the open-source device. Each subject is represented by a unique id number and the repeated values share the same id subject number symbol. Since the SMM150 has 10 repeated values but the open-source 3, the values that are not available for both devices are discarded. The open-source device underestimates the water content measurements for all subjects apart from subject 6, 28 and 47. Moreover, the scatterplot reveals low correlation between the methods. Table 4.5 presents a summary of the measurement range for the water content (%) for both devices.

Table 4.5. A summary for the water content (%) measurements. The table presents the minimum, 1st, 2nd and 3rd quartile, the mean and the maximum value of the measurements.

Summary for the water content (%) measurements					
Minimum	Q1	Median	Mean	Q3	Max
15.8	29.2	32.2	31.98	34.8	45.4

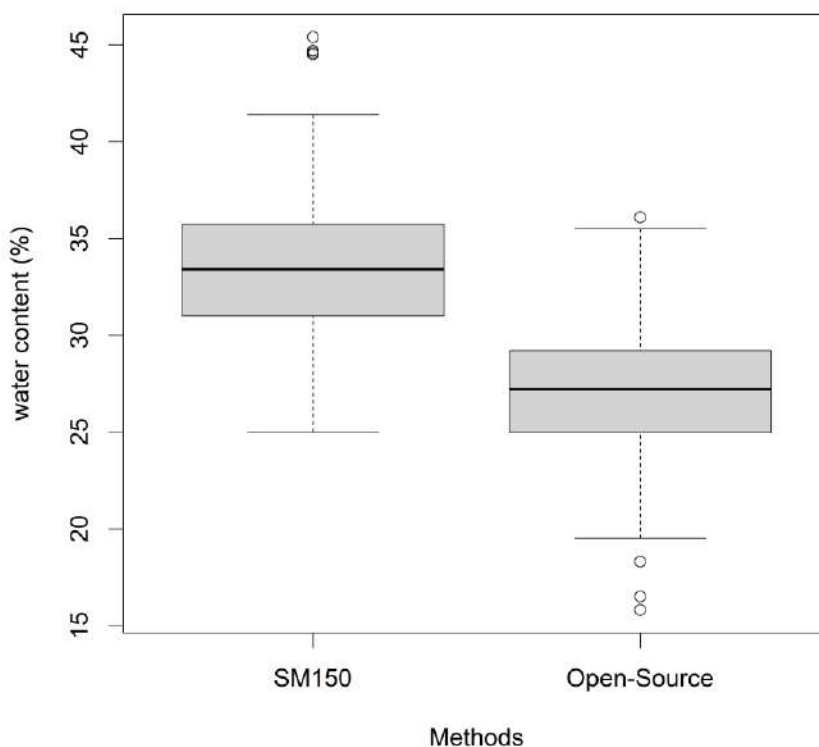


Figure 4.12. Boxplot for the water content (%) measurements per device. SM150's measurements are significantly higher than the open-source's

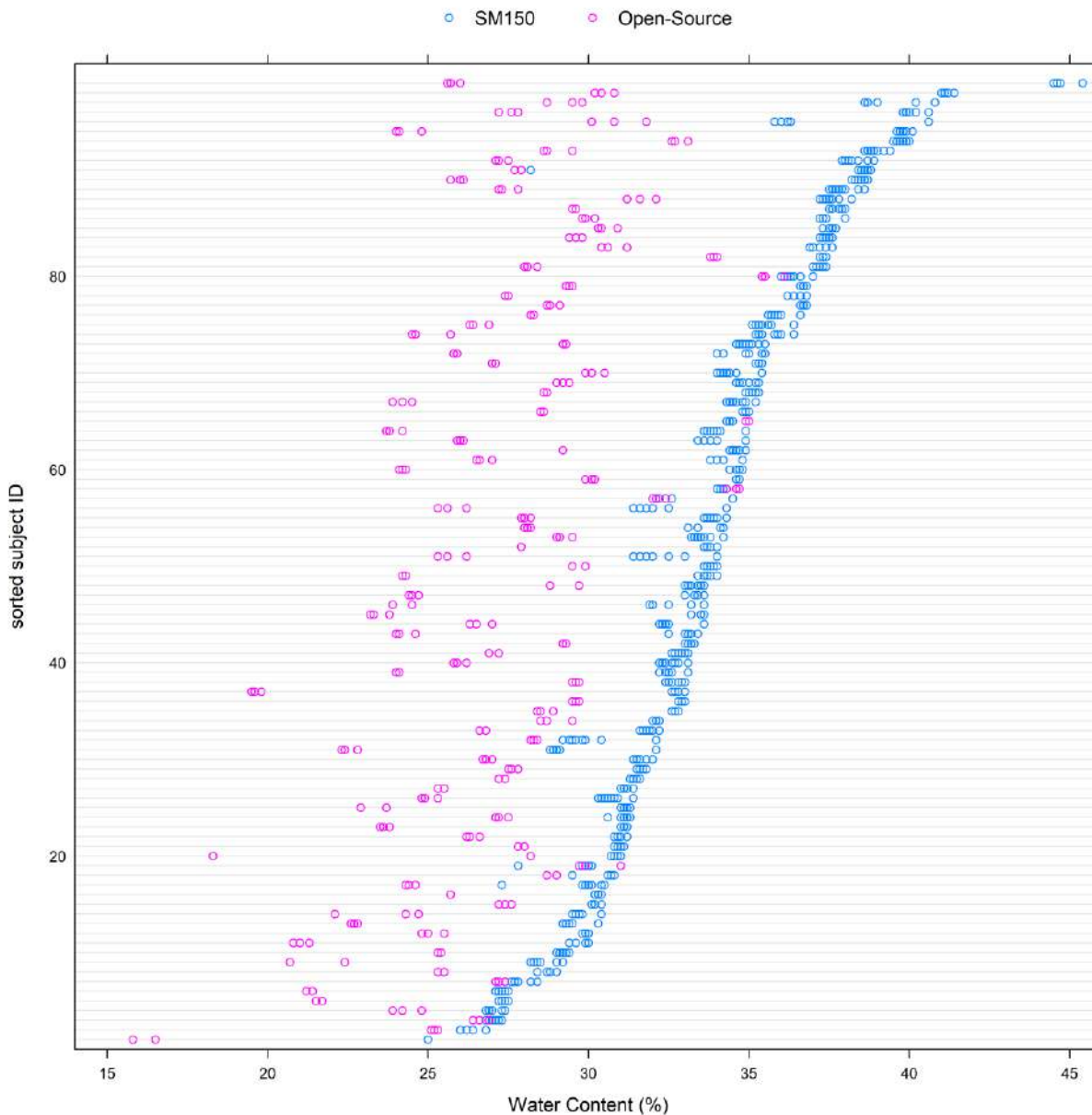


Figure 4.13. Trellis plot for the soil water content measurements. The vertical axis is divided into rows and each row displays all the repeated measurements for both devices using method-specific colors. Pink color represents the measurements for the open-source device while blue represents the measurements for the reference device. The open-source device underestimates the soil water content for most of the subjects by a non-constant quantity. As the measurement range increases the difference decreases revealing dependance of the difference on the magnitude of the measurements.

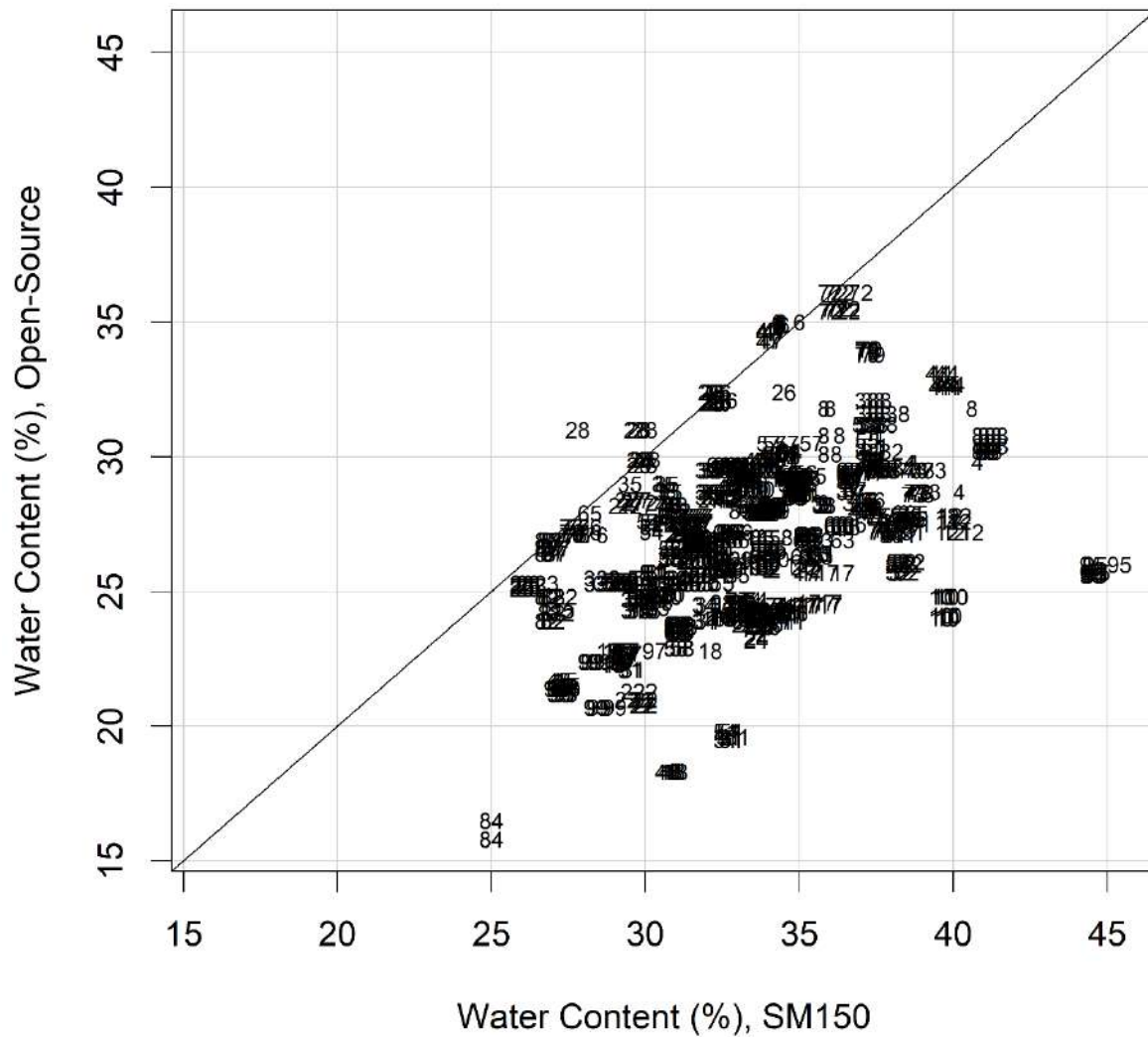


Figure 4.14. Scatterplot for the water content (%) measurements of the reference device versus the open-source. Each subject is represented by a unique id number and the repeated values share the same id subject number symbol. The open-source device underestimates the water content for all subjects apart from subject 6, 28 and 47. The scatterplot reveals low correlation between the methods.

4.6.2. Agreement Evaluation

The *measure_compare* function of the “methodCompare” package fits a measurement-error model (2.25 - 2.27) by assuming that $a_2 = 0$ and $b_2 = 1$. Figure 4.15 presents the extended Bland-Altman plot. The points in the plot are not centered at zero, suggesting a difference in the fixed bias (as reported previously in the exploratory data analysis). There is an obvious non-constant, fan-shaped spread thus, revealing heteroscedastic differences. Regression based limits of agreement for difference in water content (%) determined by the open-source and SM1500 devices are superimposed in the plot. It is expected that the 95% limits of agreement include 95% of differences between the two devices.

Figure 4.16 presents the bias plot. The *bias_plot* function visually assesses the fixed (differential) and proportional bias relative to the reference method using the methodology described in Section 2.1.4. The plot is produced by graphing a scatter plot of the open-source device and of SM150 versus the best linear unbiased predictor (BLUP) of SM150 (reference device) with their two corresponding regression lines. A second scale is added on the right axis, displaying the relationship between the estimated amount of bias and the BLUP of SM150. Moreover, the analysis performed to estimate the bias considers the heteroscedasticity effect. The fixed and proportional bias estimates are 11.669 (95% CI [8.6113, 14.7276])) and 0.46 (95% CI [0.3696, 0.5513]) respectively. Table 4.6 presents the bias estimates and their corresponding (95%) confidence intervals. The equal scale assumption for the mixed-effects models is not valid according to the proportional bias estimate and 95% confidence intervals. The interpretation of the bias equation is simple and straightforward. For example, if the true unobserved value (approximated by the BLUP) is 25%, then the open-source device will measure on average 23.169 % ($0.46 \cdot 25 + 11.669$). The blue dashed line refers to the open-source device regression line while the black solid line refers to the SM150 regression line. The red solid line corresponds to the total bias produced by the open-source device when measuring the true unobserved value (approximated by its BLUP). As the magnitude of the water content (%) increases the bias increases.

Figure 4.17 presents the compare plot, which is produced using the *compare_plot* function. This plot allows the visualization of the bias-corrected values of the open-source device. The blue solid line presents the open-source device regression line before recalibration and its corresponding measurements are presented with a blue solid circle. The red dashed line presents the open-source device regression line after recalibration and its corresponding measurements are presented by red solid rhombus symbols. The black solid line presents the SM150 regression line, and the corresponding measurements are presented with a black transparent circle. After recalibration, the open-source device’s measurements are centered around the SM150 measurements. However, the spread around the regression line seems to be higher than before for lower and higher values. This can be justified from the proportional bias value which is lower than one. The recalibration formula was produced by substituting the differential and proportional bias in (2.28):

$$y_{ij}^* = \frac{y_{ij} - 11.669}{0.46}$$

Figure 4.18 presents the precision plot. This plot allows the visual comparison for the precision of the open-source device with SM150. A scatter plot is created for the estimated standard deviations against the BLUP of the true unobserved value. The standard deviations are estimated after the open-source device is recalibrated. The open-source device’s precision decreases as the measurement magnitude

increases. The SM150 has higher precision compared to the open-source device. Specifically, the SM150 device is almost 6 times more precise than the open-source.

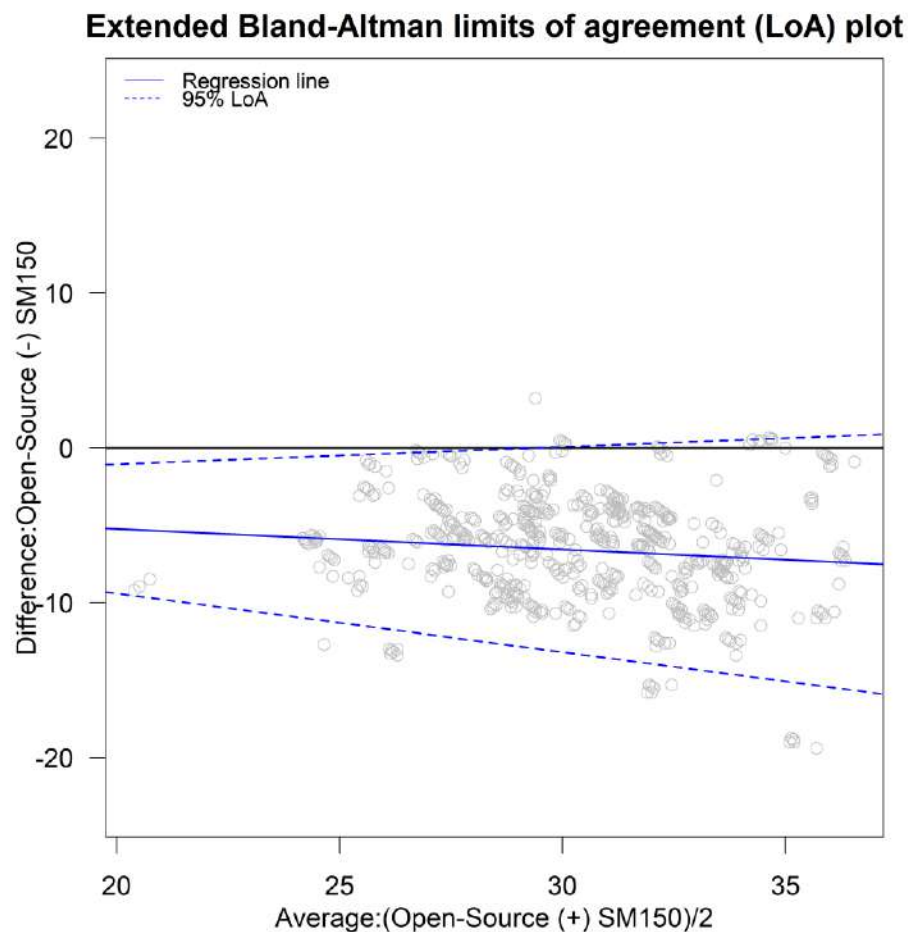


Figure 4.15. Extended 95% Bland-Altman limits of agreement. The differences depend on the magnitude of the water content since they increase when it increases. Moreover, the variability is not constant across the measurement range since it increases when the measurement range increases. The points in the plot are not centered at zero, suggesting a difference in the fixed bias.

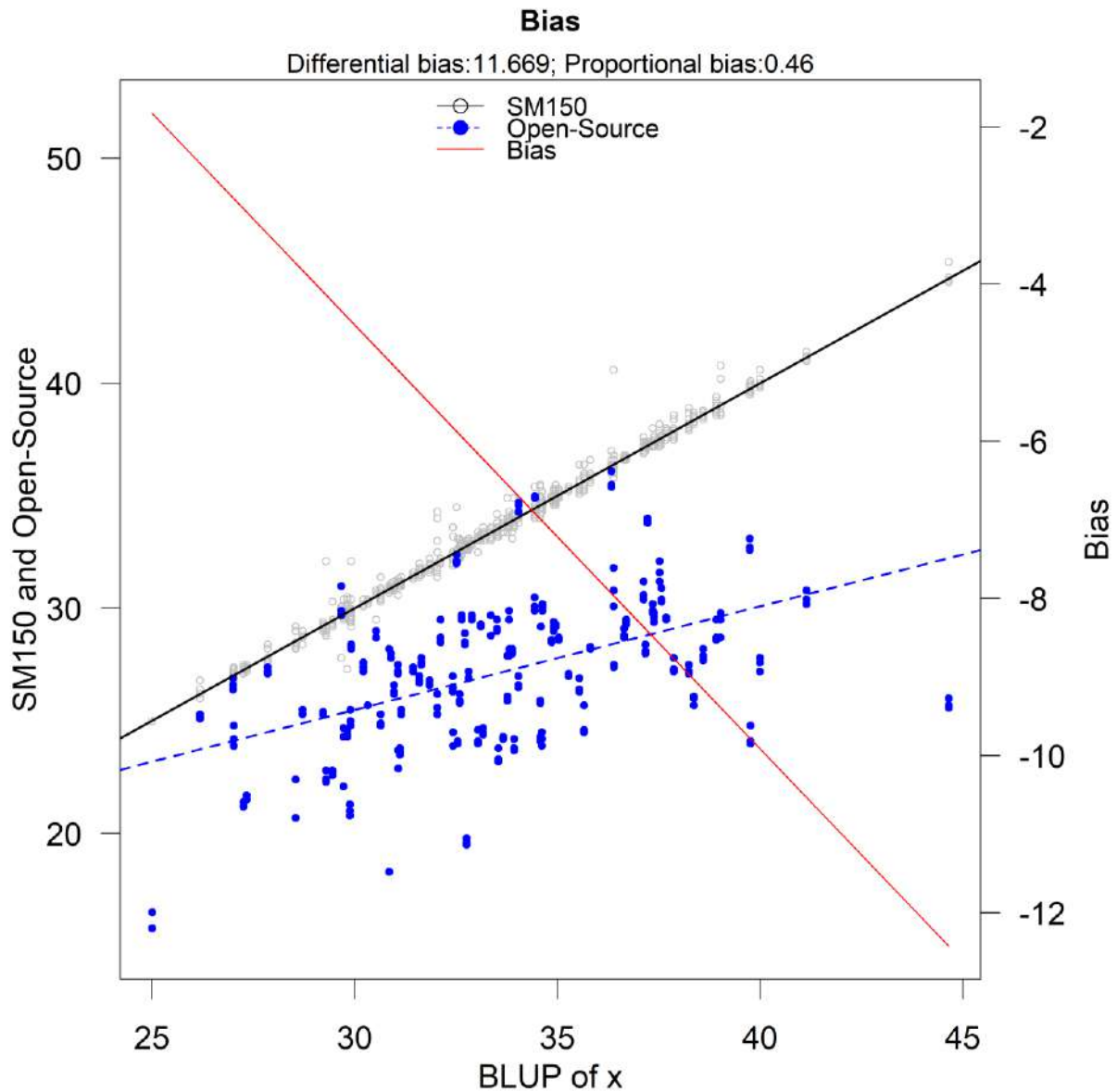


Figure 4.16. Bias-plot. The blue dashed line refers to the open-source's device regression line while the black solid line refers to the SM150 regression line. The red solid line corresponds to the total bias produced by the open-source device when measuring the true unobserved value (approximated by its BLUP). As the magnitude of the water content (%) increases the bias increases.

Table 4.6. The differential and proportional bias of the open-source device compared to SM1500 (reference).

Bias	Estimate	95 % Confidence Intervals	
		2.50 %	97.50 %
Differential	11.6694	8.6113	14.7276
Proportional	0.4605	0.3696	0.5514

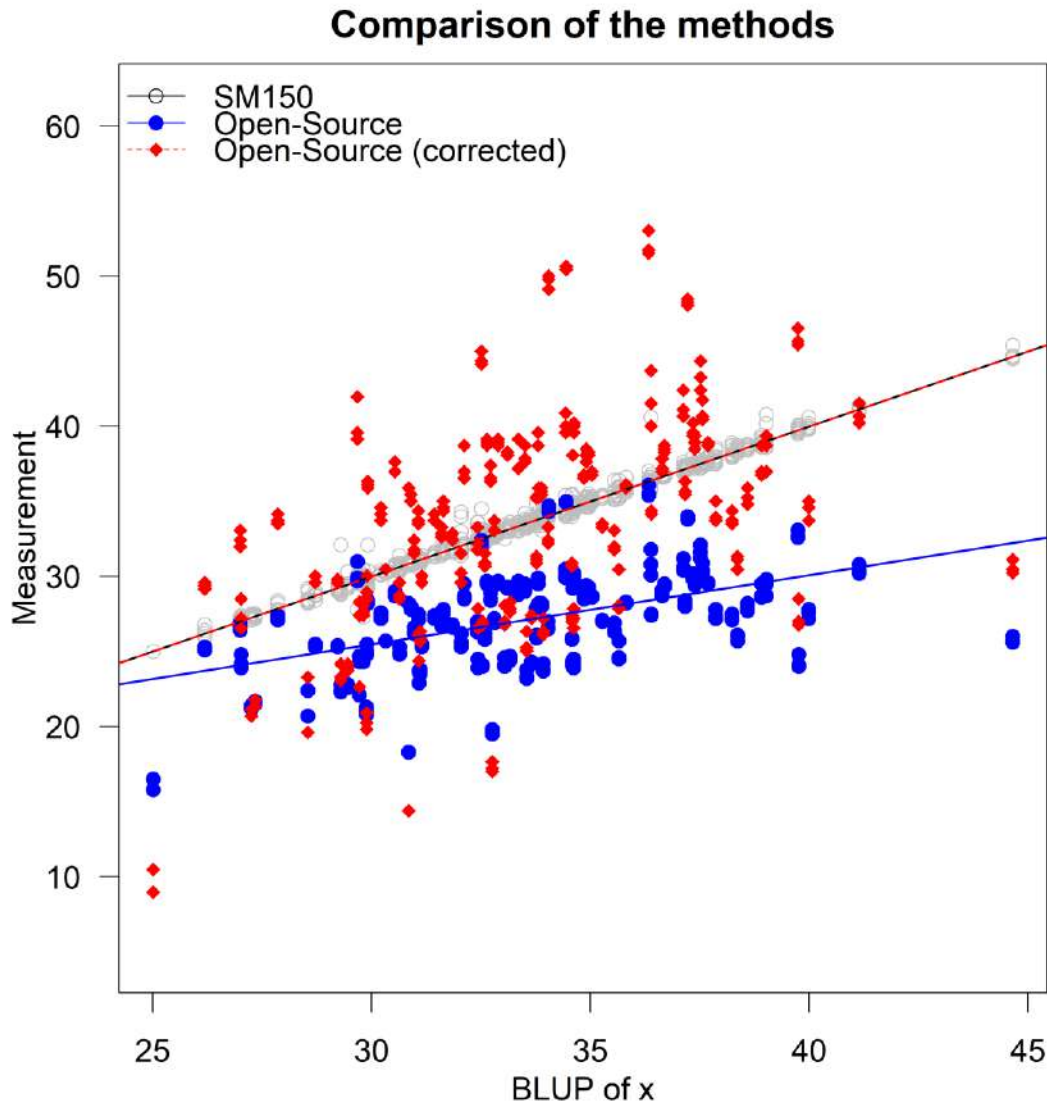


Figure 4.16. Compare plot. This plot allows the visualization of the bias-corrected values of the open-source device. The blue solid line presents the open-source device regression line before recalibration and its corresponding measurements are presented with a blue solid circle. The red dashed line presents the open-source device regression line after recalibration and its corresponding measurements are presented by red solid rhombus symbols. The black solid line presents the SM150 regression line, and the corresponding measurements are presented with a circle.

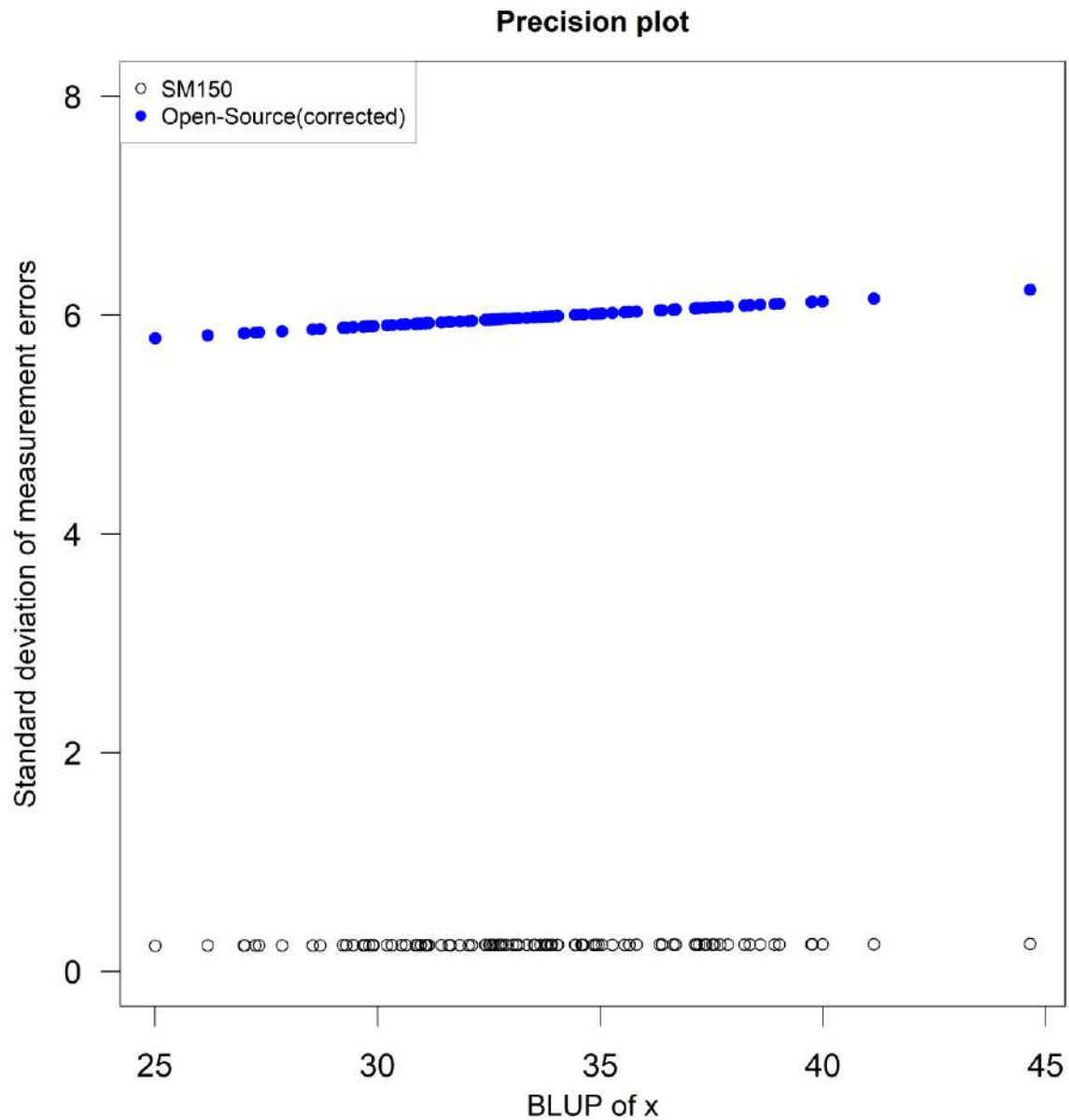


Figure 4.17. Precision plot. This plot allows the visual comparison of the precision of the open-source device with that of SM150. A scatter plot is created for the estimated standard deviations against the BLUP of the true unobserved value. The standard deviations are estimated after the open-source device is recalibrated. The open-source device's precision decreases as the measurement magnitude increases. The SM150 has higher precision compared to the open-source device. Specifically, the SM150 device is almost 6 times more precise than the open-source.

4.7. Discussion

An open-source device for measuring perlite slab's water content was designed and constructed to assist aquaponics experiments in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. To investigate the validity of the device's measurements a method comparison study was designed. An SM150 water content (%) sensor with its corresponding Handheld Meter (HH150) was used as the reference device. A sample of 100 perlite slabs were chosen randomly and measured. The experimental design included repeated measurements. The samples were balanced but the repeated measurements were unbalanced. The between-subject variation of water content (%) ranged from 15.8 to 45.4.

Investigating the data from a preliminary experiment a measurement-error model was fit to assess the proportional bias of the open-source device compared to the SM150. The findings discouraged the use of mixed-effects models to assess the agreement and similarity of the two methods since the proportional bias was significantly different to 1 which is a violation of the model's assumptions. The extended Bland-Altman plot revealed heteroscedastic differences and an underestimation of the open-source device compared to the SM150 for the water content (%) measurements. The methodology used for the analysis was Taffé's (discussed in Section 2.1.2) since the author fits the data using a measurement-error model. The fixed bias estimate is 11.669 (95% CI [8.6113, 14.7276]) units of water content (%) and the proportional bias 0.46 (95 % CI [0.3696, 0.5514]). According to the bias plot, lower bias seems to be found in lower values of the measurement range and increases as the magnitude of the measurements increase.

If the accepted difference between the two methods is 5%, then a researcher could use the device for the measurement range of [25, 35]% of water content. Overall, the bias plot can be a guide for using (or not) the two devices interchangeably. The measurement range used by a researcher is based on the quality characteristics of the perlite slab according to Burés et al. (1997a; 1997b).

After recalibration, the agreement seemed to improve. However, the spread around the regression line seems to be higher than before for lower and higher values.

During the experiment, two more characteristics of the perlite slabs were measured. The electrical conductivity and the temperature. Taffé's approach does not include covariates thus, further research should be carried out using the data from the current experiment that could reveal possible sources of disagreement. Last, further research is vital for evaluating indices like CCC and TDI produced by measurement-error model counterparts.

The open-source device can be modified for measuring more environmental characteristics regarding aquaponics applications. The data can be fed to a prediction model for crop damage prevention or growth estimation. Open-source devices are essential since their low cost allow researchers to increase replications and improve the prediction models' precision.

5. A method comparison study between open-source and industrial pH meter logger and sensors.

5.1. Goal of the Study

The acidity of fleshy fruit, as measured by titratable acidity and/or pH, is an important component of fruit organoleptic quality. Fruit acidity is associated with the presence of organic acids, with malic and citric acids to be the most abundant in most ripe fruits (Etienne et al., 2013). There is an interrelated relationship between pH and titratable acidity. Titratable acidity is determined by neutralizing the acid present in a known quantity of food sample using a standard base, while the endpoint for titration is usually a target pH (or the color change of a pH-sensitive dye). The titratable acidity of fruits is used, along with sugar content (sweetness), as an indicator of maturity (Tyl & Sadler, 2017).

Citrus is one of the most important commercial fruit crops in the world, and fruit weight, size, acidity and maturity index, harvest time, chemical and nutritional composition are important quality traits for fresh citrus consumption and acceptance by the citrus industry. Organic acids and sugars expressed as total soluble solids (TSS) vary according to species, varieties, and environmental and horticultural conditions such as climate, rootstock, and irrigation (Lado et al., 2014). In citrus fruits, the perceived flavor depends on the combination of taste and aroma, in which the sweet and sour taste attributes are principally result from the presence of sugars and acids in the juice, and the aroma depends on many volatile organic compounds (Benjamin et al., 2013; Zheng et al., 2016). The TSS/acidity ratio has been used worldwide as the main commercial maturity indicator of citrus fruit internal quality. In general, a TSS/acidity ratio of at least 6 or higher is acceptable for commercial marketability, however important differences may exist depending on the citrus species. Ratios acceptable for commercialization usually range from 7-9:1 for oranges and mandarins to 5-7:1 for grapefruits. For lemon this index is not applied even though in India, the ripening period, juice content and a total acidity of 9% are considered as good indicators of a mature fruit (Lado et al., 2014).

An open-source device for measuring pH was designed and constructed for the needs of this research. The benefits of using an open-source device interchangeably with a corresponding industrial are mostly based on the lower cost and configurability of the former. Thus, a method comparison study between the open-source device and an industrial was designed to evaluate their agreement and similarity. This study will reveal measurement ranges where the difference between the two devices might be acceptable or not. Using recalibration methods the agreement might increase and improve.

5.2. Design of the Device

The device was designed as a portable logger that can be used for in situ measurements and at the same time as laboratory equipment. Thus, it is battery operated or can be powered via an electrical outlet using micro-USB to USB-A cable. When the battery is charged the device will be powered directly via the electrical outlet. Furthermore, it was purposely fitted in a rather small enclosure box to increase portability for field measurements. The power consumption was measured using an AXIOMET AX-178 polimeter (AXIOMET, n.d.). A measurement is taken every one minute and stored in the micro-SD card. The consumption is 15.3761 mAH with no sleep mode activated.

5.2.1. Hardware

The open-source logger is equipped with two sensors. A pH sensor (DFRobot, 2022) from Seeed studio and a temperature sensor DS18B20 (Maxim Integrated Products, 2019). A 16-bit analog to digital

converter is added to the design to improve the precision of the voltage reading since the output of the sensor is analog. The development board for this device is the Adafruit feather proto 32u4 (discussed in Section 2.3.1). An Adafruit Featherwing logger was added for instant capture of the measurement in a microSD card embedded with a timestamp by pressing a push button. A Nokia 5110 LCD monitor (Philips, 1999) was added to display the values of voltage, pH, temperature, and battery. A 1200 mAH LiPo battery is the main power source of the device which can be charged via micro-USB to USB-A while the device is operating.

The pH Sensor

SEN0169 (Figure 5.1) is an analog pH meter, specifically designed for Arduino and Arduino-compatible microcontrollers. The electrode is considered industrial. The sensor has long life (>0.5 years in long-term monitoring mode), is highly accurate (0.1 pH in room temperature), it has fast response (≤ 1 minute),



Figure 5.1. DFRobot PH meter (SEN0169) (source: DFRobot official website).



Figure 5.2. Waterproof DS18B20 digital temperature sensor for Arduino and Arduino compatible microcontrollers. (source: DFRobot official website).

has a measuring range from 0 to 14 pH and includes a gain adjustment potentiometer. The output voltage of the electrode is linear and is capable of long-term monitoring. The sensor has an industrial built and uses a BNC connector and PH2.0 sensor interface. Table 5.1 summarizes the technical specifications of the probe. The communication between the sensor and the MCU is one-way since the sensor transmits data using an analog MCU pin. Since the 32u4 MCU uses a 10-bit ADC, an ADC1115 16bit ADC and gain amplifier is added to increase the sensor's precision.

The temperature sensor

DS18B20 (Figure 5.2) is a waterproof digital temperature sensor designed for Arduino or Arduino compatible microcontrollers. According to the manufacturer, since the sensor's signal is digital no signal degradation is present even if the distances between the MCU and the sensor are very long. The sensor provides 9-to-12-bit resolution temperature readings (configurable via software). The communication protocol between the MCU and the sensor is 1-Wire (discussed in Section 2.3.4). Multiple DS18B20 sensors can connect on the same 1-Wire bus since they are produced with a unique silicon serial number. Table 5.2 summarizes the technical specifications of the sensor.

Table 5.1. Technical Specifications of pH sensor SEN0169.

Specifications	
Module Power	3.3 or 5 Volts
Measuring Range	0–14 pH
Measuring Temperature	0 ~ 60 °C
Accuracy	±0.1 pH (25 °C)
Response Time	≤ 1 minute
Interface	Analog Output

Table 5.2. Waterproof DS18B20 digital temperature sensor specifications.

Specifications	
Module Power	3 to 5.5 Volts
Measuring Range*	-55 °C to 125 °C
Accuracy	±0.5 °C from -10 °C to 85 °C
Resolution	9 to 12 bits ADC
Interface	1-Wire
Steel tube dimensions	6mm diameter by 35 mm long
Notes	

- *The temperature range is between -55 °C to 125 °C. However, since the junction between the cylinder and the cable is a heat tube, the manufacturer recommends -55 °C to 100 °C.
- Multiple sensors share the same pin.
- Includes a temperature-limit alarm system.
- Query time is less than 750ms .

ADS1115 16bit ADC with gain amplifier

This module is a precision ADC with 16 bits of resolution. The first 15 bits are used for the value and the last one for the sign of the value. It is equipped with a voltage reference and an oscillator. It uses I2C (discussed in Section 2.3.4) communication protocol to interact with the MCU. Four different slave addresses can be selected allowing four different ADS1115 modules to be connected in the same bus. Its operating voltage ranges from 2 to 5.5 Volts. Furthermore, it can converge signals at rates up to 860 samples per second. Its second functionality includes programmable gain amplifier that provides input ranges from inputs to as low as ± 256 mV with increments of 0.0078125 mV, thus measuring both small and large signals with high resolution. Moreover, it offers an input multiplexer which provides two differential or four single-ended inputs. Last, the module operates in continuous conversion mode or a single-shot mode. This means that it automatically powers down in single-shot mode, reducing the power consumption during the measuring periods. To avoid damaging the module, the gain should be set more than or equal to the input voltage of the channel. Table 5.3 summarizes the technical specifications of the module.

5.2.2. Software

DFRobot provides a library for the SEN0169 via GitHub (Arya11111, 2022). The library includes a calibration mode. However, the calibration was performed manually and is described in Section 5.2.5 due to the MCU's incompatibility. Furthermore, the code was developed without using the library.

The code functionality is described as follows. First, the MCU reads the signal of the pH sensor via the ADS1115, in continuous mode using a single input channel. Second, in case an instantaneous measurement needs to be taken and stored the user will press the push button and the measurement embedded with a time stamp will be stored in the microSD card. The function *button()* and *store()* provide these functionalities. After the calibration procedure, the equation is stored in the sketch and the *measure()* function returns the pH measurement after inserting the input voltage. The function *measure()* returns the proper calibration line, depending on the temperature of the liquid. The sketch is available in GitHub (Bataka, 2023).

Table 5.3. ADS1115 16bit ADC technical specifications.

Specifications	
Module Power	2 to 5.5 Volts
Current Consumption	150 μ A continuous mode
Programmable Data Rate	8 to 860 SPS
Interface	I2C
Input Channels	4 single / 2 differential
Notes	
<ul style="list-style-type: none">• Internal low-drift voltage reference• Internal Oscillator• Internal PGA	

5.2.3. Calibration Method

The calibration procedure was performed using two pH buffers. Eight measurements were taken. The first two were taken from 4.01 and 7.01 pH buffers when the liquid's temperature was 7.5°C. The same procedure followed for temperatures of 12.5°C, 17.5°C and 22.5°C. The probe was removed from the solution 1 minute after its insertion to reach the response time according to the sensor's datasheet. The temperature of the buffer solution was measured using the DS18B20 temperature sensor. Table 5.4 summarizes the voltage and their corresponding pH values. Each temperature interval is using the calibration equation of the corresponding midpoint temperatures. For example, equation (5.1) will be used for the range between 5°C and 10°C. Figure 5.3 presents the four calibration lines per temperature range (5.1 – 5.4).

$$y_{7.5^{\circ}\text{C}} = -6.27615x + 16.4456 \quad (5.1)$$

$$y_{12.5^{\circ}\text{C}} = -6.1349695x + 16.25767 \quad (5.2)$$

$$y_{17.5^{\circ}\text{C}} = -6.0241x + 16.09036 \quad (5.3)$$

$$y_{22.5^{\circ}\text{C}} = -6.04351x + 16.1196615 \quad (5.4)$$

Table 5.4. Temperature of the pH buffer during the calibration procedure. Eight measurements were taken. Four for 4.01 pH buffer and four for 7.01 pH buffer.

Temperature	Temperature Range (°C)	Voltage	pH
7.5°C	[5, 10]	1.983	4.01
13.5°C	(10, 15]	1.505	7.01
17.5°C	(15, 20]	1.998	4.01
22.4°C	(20, 25]	1.509	7.01
7.5°C	[5, 10]	2.007	4.01
13.5°C	(10, 15]	1.509	7.01
17.5°C	(15, 20]	2.010	4.01
22.4°C	(20, 25]	1.514	7.01

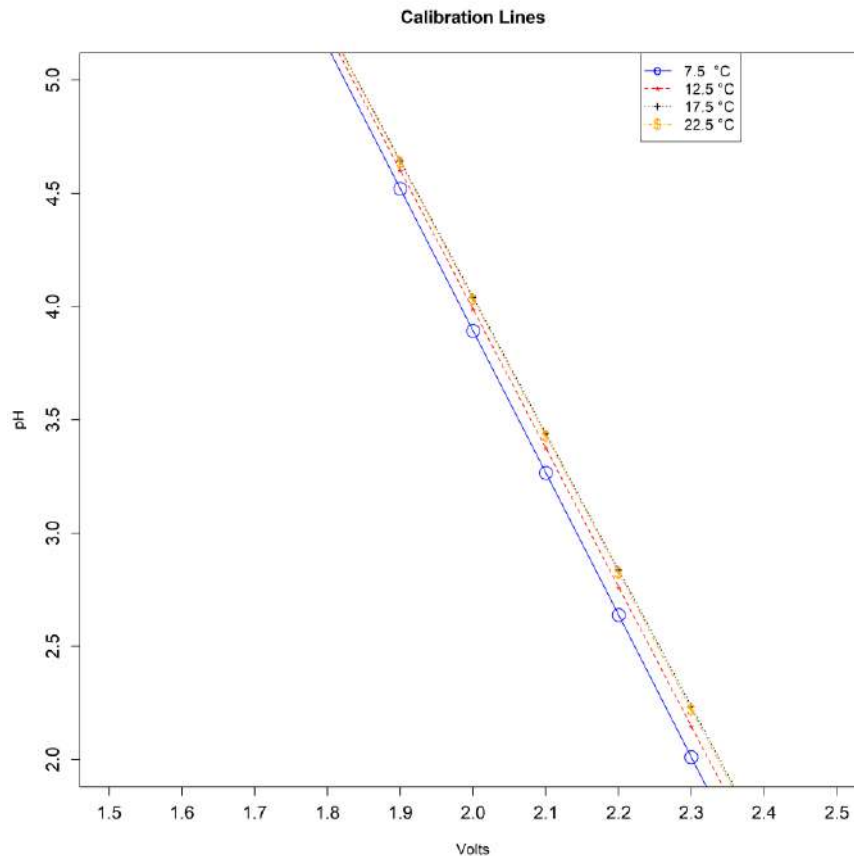


Figure 5.3. Calibration curves for 7.5°C, 13.5°C, 17.5°C and 22.4°C.

The equations were added to the Arduino sketch in the *measure()* function. PH was automatically calculated depending on the temperature measurement.

5.2.4. Cost of the Device

Table 5.5 summarizes the cost per component and the total cost. The cost can be reduced if parts like the development board and the ADS1115 can be replaced by cheaper equivalents from other brands. Furthermore, the final product is usually not using development boards, removing the inessential parts. Thus, the cost and the device's footprint reduce, especially when the PCB is designed and printed with Surface Mounted Discrete (SMD) electronic parts replacing the through hole equivalents.

Table 5.5. Device cost in 2023 local vendor prices and international vendor Mouser.

Item	Cost
DFRobot PH meter Pro Kit (SEN0619)	70
Temperature Sensor DS18B20	3
Adafruit Feather proto 32u4	18.75
Adafruit Adalogger FeatherWing	8.41
Adafruit ADS1115 ADC 16bit	18.4
microSD Card	3
Coin Cell Battery	1
IP66 Enclosure Box	5
PCB	1.5
Nokia 5110 module	4.68
Others (Wires, Solder)	2
Total	127.63

5.3. The reference device

The reference device is a Hanna Instruments HI9024 Waterproof pHMeter (Hanna Instruments, 1996)(Figure 5.4b). It is a heavy-duty pH meter designed for laboratory use and its accuracy is sustained even under harsh industrial conditions. It can easily be calibrated and has three memorized buffer values (4.01, 7.01 and 10.01). The device has automatic buffer recognition thus avoiding errors during the calibration procedure. Moreover, it is equipped with a temperature compensation function. The temperature can be measured using a temperature sensor probe or can be entered manually. Since there was no temperature sensor available, the temperature was set manually using the DS18B20 sensor which was embedded in the open-source logger. Thanks to its waterproof cylindrical case the temperature sensor was inserted in the solution that was intended to be measured during the experiment. The technical specifications of the reference device are summarized in Table A9 (Appendix A). The specific pH meter model is not available in the market since it is considered obsolete. An equivalent but contemporary model is HANNA HI 99171. Its cost in Greek local vendors is around 585 Euro including the shipping cost.

5.4. Experimental Design

The target solution was retrieved from two varieties of citrus fruits. Each fruit is considered an experimental unit. In total, 15 grapefruits and 15 juice oranges (Valencia variety). Each unit was hand squeezed (Figure 5.4a), and its juice was measured by the open-source device and by a Hanna HI 9024 pH meter. The order of measurement was randomized using R's *sample* function, and 10 repeated measurements were collected. The procedure to retrieve the repeated measurements was sequential. First, the *sample* function was used to define the instrument that will measure first. Then, the other instrument was used. In this way 9 more measurements of the same juice were taken by first cleaning each instrument using deionized water and then taking the measurement. The data are considered linked since they are paired over the measurement times. Figure 5.4c displays the open-source pH sensor and the measurement procedure. Table 5.6 summarizes the experimental design information.



Figure 5.4. (a) The hand squeezing procedure. (b) Hanna pH meter. (c) DFRobot pH meter probe.

Table 5.6. Summary of the experimental design information. The sample size is 30, since each fruit is considered one experimental unit. There were ten repeated measurements, and the design was balanced. The data are considered linked.

Experimental Design	
Experimental Unit	Fruit
Repeated Measurements	Yes, sequentially
Number of repeated measurements	10
Data category	linked data
Sample size	30
Balanced/Unbalanced measurements	Balanced
Possible Covariates	Temperature, juice quantity, instrument turn

5.5. Method Comparison Study Techniques

To evaluate the agreement and similarity between Hanna HI9024 and the open-source device CCC was calculated using model (2.11) and the methodology of Choudhary and Nagaraja (2017a) and Carrasco et al. (2013). TDI was calculated using the methodologies by Choudhary and Nagaraja (2017a) and Escaramis et al. (2010a). The measures of similarity and repeatability were calculated using the methodology from Choudhary and Nagaraja (2017a).

R-code, R-packages and datasets are available online (Choudhary & Nagaraja, 2017b; Carrasco & Martinez, 2022; Escaramis et al., 2010b). Specifically, the R-package is “cccrm” from Carrasco and Martinez (2022) and the function `cccUst()` which uses the U-statistic approach (King et al., 2007a).

Moreover, using model (2.25 – 2.27) and “MethodCompare” R-package, the proportional and fixed bias are evaluated. Last, the extended Bland-Altman plot, the bias plot and the plots after recalibration were produced using the same package.

Bland-Altman plot with confidence intervals for the LOAs was produced using “Blandr” R-package.

The code was modified and adjusted to the experiment's needs and is available on GitHub (Bataka, 2023).

5.6. Results

The "Chapter5.R" R-script (Bataka, 2023) was used to perform the analysis. In this experiment the type of the fruit (grapefruit or juice orange), temperature, quantity of the juice and the turn which the instruments measured first are considered covariates.

5.6.1. Data Exploration

Data exploration follows using three different graphs. Figure 5.5a displays a scatterplot for the pH measurements of Hanna pH meter versus the open-source. To avoid using the same plotting symbol per subject and visualize the repeated measurements, each subject is represented by a unique id number and the repeated values share the same id subject symbol. Using this method, a dependence structure is developed, and more information is available in one plot. A systematic underestimate of the open-source device for pH measurements is clear in the graph. There are two clusters formed in the data. The lower left corresponds to the grapefruit pH while the upper right to the juice oranges. Figure 5.5b displays a Bland-Altman plot (averages vs differences) without the limits of agreement. For higher values of pH, the differences seem to have slightly higher spread compared to lower values of pH. This is a sign of possible heteroscedastic errors. There is no obvious trend in the Bland-Altman plot suggesting a common scale for the assays, verifying the common scale assumption for the mixed-effects model. This is also obvious in the extended Bland-Altman plot (Figure 5.8).

Figure 5.6 displays a trellis plot. The vertical axis is divided into rows and each row displays all the repeated measurements for one subject and both devices using method-specific colors. Blue color represents the measurements for Hanna pH meter while pink represents the measurements for the open-source device. Since the repeated measurements are plotted in one row, within-subject variability is visible and easy to compare with the between-subject variability. The open-source device shows slightly less within-subject variation compared to Hanna. The between-subject variation ranges between 2.78 to 3.7 and a summary is presented in Table 5.7. A consistent bias is also visible in the graph, suggesting a constant fixed bias throughout the measurement range. The open-source device underestimates the pH measurements by approximately 0.22 units.

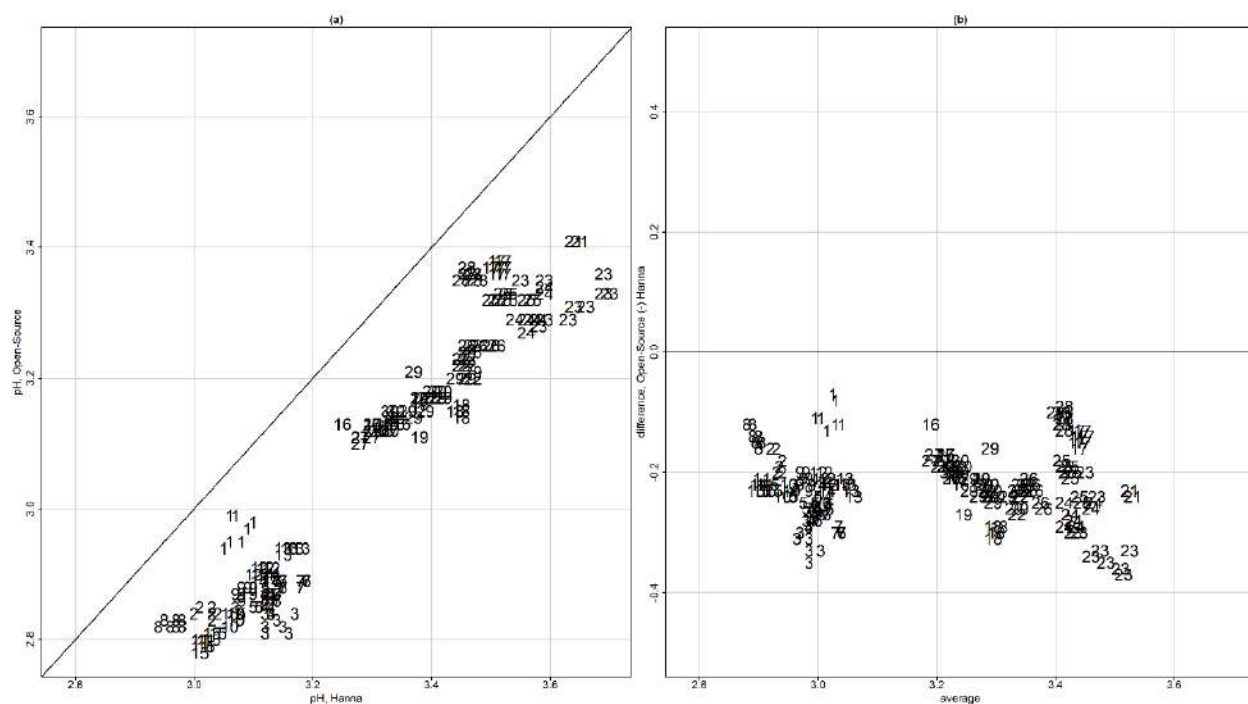


Figure 5.5. (a) Scatterplot for the pH measurements of Hanna pH meter versus the open-source. Each subject is represented by a unique id number and the repeated values share the same id subject number symbol. A systematic underestimate of the open-source device for pH measurements is clear in the graph. There are two clusters formed in the data. The lower left corresponds to the grapefruit pH while the upper right corresponds to the juice oranges. (b) Bland-Altman plot (averages vs differences) without the limits of agreement. For higher values of pH the differences seem to have slightly higher spread compared to lower values of pH. This is a sign of possible heteroscedastic errors. There is no obvious trend suggesting a common scale for the assays, verifying the common scale assumption for the mixed effects model.

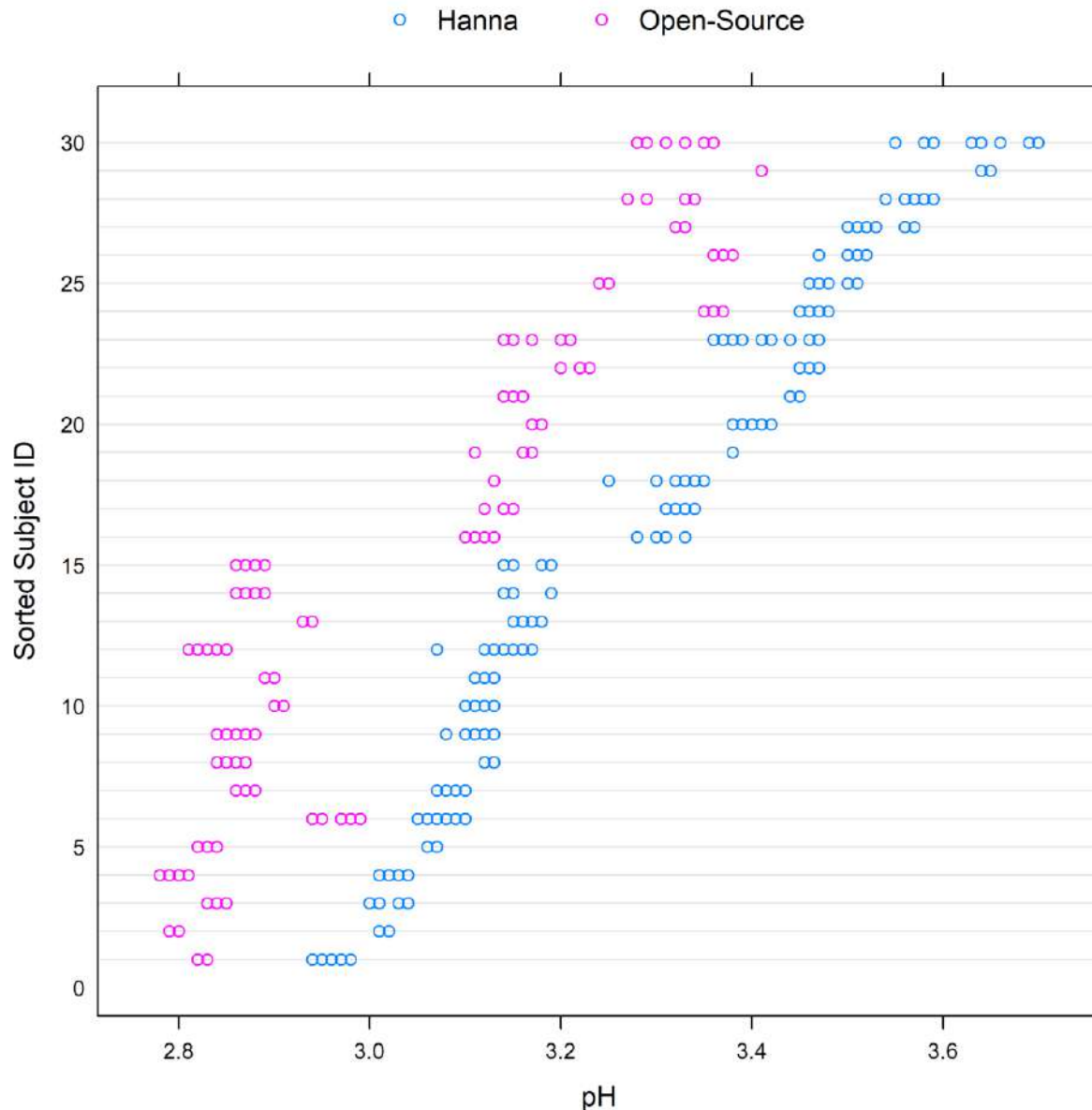


Figure 5.6. Trellis plot for the pH measurements. The vertical axis is divided into rows and each row displays all the repeated measurements for one subject using method-specific colors. Blue color represents the measurements for the Hanna pH meter while pink represents the measurements for the open-source device. The open-source device underestimates pH measurements for most of the subjects by approximately 0.22 units. Moreover, the within subject-variation for the open-source device is less than Hanna's. The between-subject variability ranges between 2.78 and 3.7.

Figure 5.7a and 5.7b display the interaction plots for subject x method and subject x time interaction respectively. For the subject x method interaction plot the average per subject for every measurement is plotted on the vertical axis and each method on the horizontal (Figure 5.7a). There is evidence of subject x method interaction since the lines intersect. Figure 5.7b displays the subject x time interaction. The repeated measurements are averaged per method for each subject (vertical axis) and the time points are

displayed on the horizontal axis. A few of the lines intersect, providing evidence of possible interaction between subjects and time.

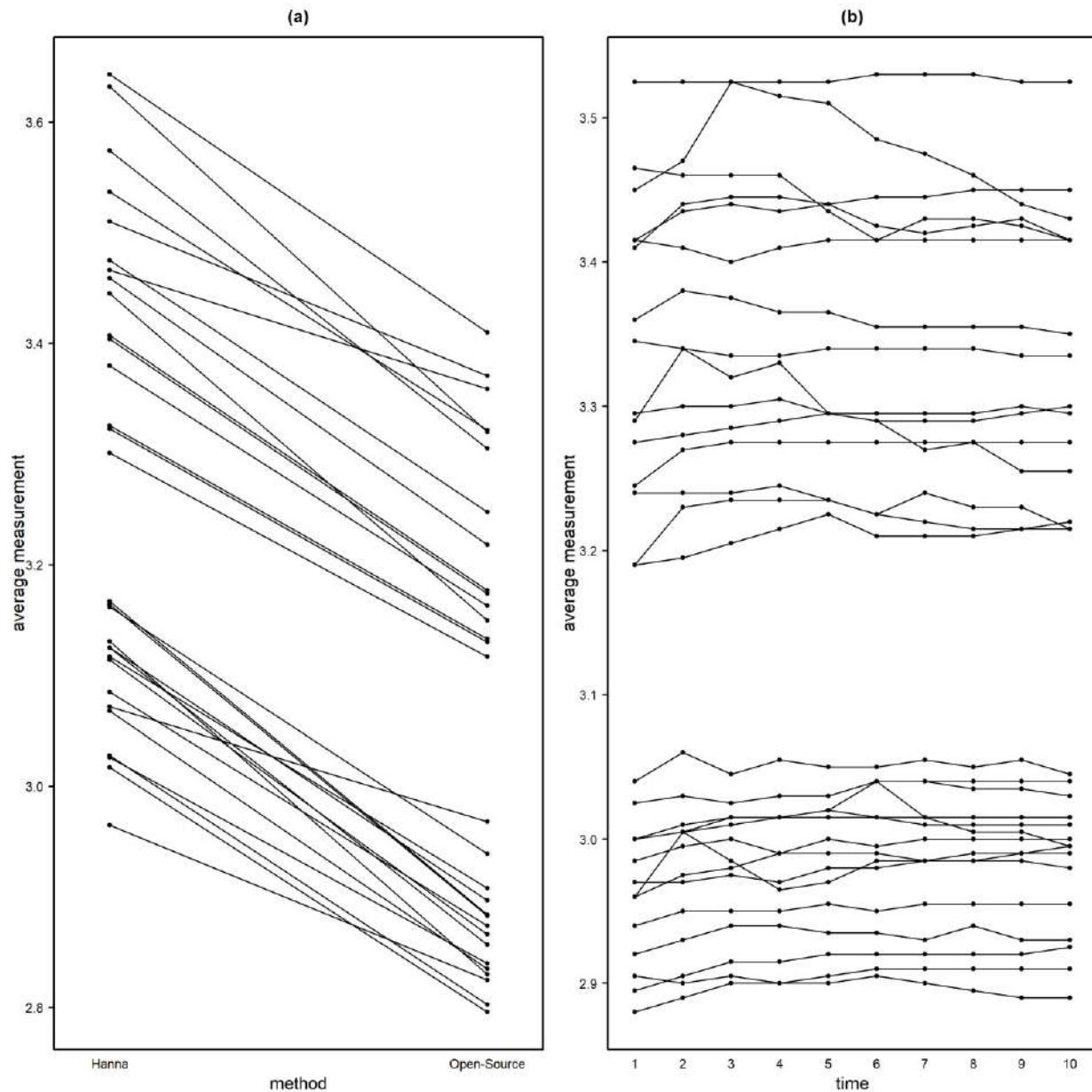


Figure 5.7. (a) Interaction plot between method and subjects. There is evidence of subject x method interaction since the lines intersect. (b) Interaction plot between subjects and time. A few of the lines intersect, providing evidence of possible, but not strong interaction between subjects and time.

Extended Bland-Altman limits of agreement (LoA) plot

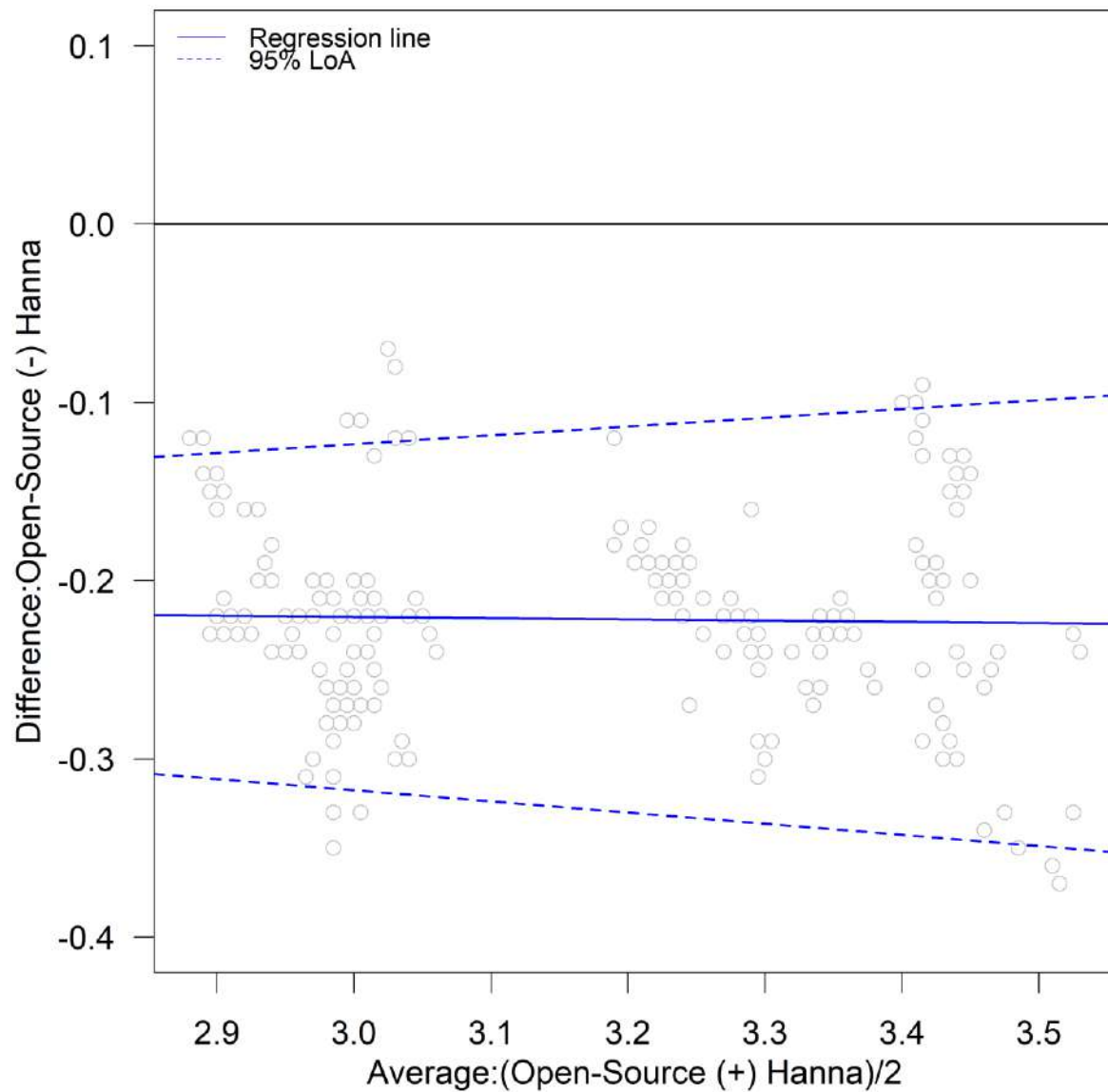


Figure 5.8. Extended Bland-Altman plot and LOA. There is slight evidence of heteroskedastic errors. No trend is apparent; thus, a common scale is assumed for the assays.

Table 5.7. The minimum, 1st quartile, median, mean 3rd quartile, and maximum pH values per device. The measurement range lies between 2.78 and 3.7.

	Minimum	Q1	Median	Mean	Q3	Maximum
Open-Source	2.780	2.868	3.045	3.053	3.220	3.410
Hanna	2.940	3.110	3.220	3.275	3.460	3.700

5.6.2. Choosing the proper model

Initially, the data were fit to the homoscedastic model with no covariates for linked data (2.11), then the corresponding heteroscedastic. However, the additional computational burden provided by the subject to occasion interaction for the linked data hindered the procedure to calculate the confidence bounds for the indices. Thus, the unlinked homoscedastic and heteroscedastic were chosen to proceed with the analysis. Figure C2a, in Appendix C, displays the standardized residuals versus fitted values of the homoscedastic model for Hanna pH meter and the open-source device and Figure C2b displays the absolute values of standardized residuals versus fitted values for the same model. There is no obvious sign of a fan shape. A formal criterion will be used to compare the homoscedastic and the heteroscedastic model. AIC and BIC were calculated, and the heteroscedastic model was chosen (Table 5.8). The heteroscedastic model is slightly better than the homoscedastic model. Either can be chosen. In this case the heteroscedastic is chosen for illustrative purposes. The analysis for the homoscedastic model can be found in Appendix C.

Later, model (2.14) was used to fit the data, which includes the covariates (discussed in Section 5.4) without interactions. According to AIC and BIC criteria, the model without covariates was chosen. Table 5.9 displays AIC, BIC, Log-Likelihood and degrees of freedom for the heteroscedastic models with and without covariates.

Table 5.8. AIC, BIC, Log-Likelihood and degrees of freedom for the homoscedastic and the heteroscedastic model. The heteroscedastic model is slightly better than the homoscedastic model. Either can be chosen. In this case the heteroscedastic is chosen for illustrative purposes. The analysis for the homoscedastic model can be found in Appendix C.

	AIC	BIC	Log-Likelihood	Degrees of Freedom
Homoscedastic	−2945.889	−2919.507	1478.945	6
Heteroscedastic	−2999.128	−2963.952	1507.564	8

Table 5.9. AIC, BIC, log-likelihood and degrees of freedom for model selection. The model which does not include covariates is slightly better than the one which includes the covariates. Thus, the model with no covariates is selected for further analysis and indices calculation.

Covariates	AIC	BIC	Log-Likelihood	Degrees of Freedom
No	−2999.128	−2963.952	1507.564	8
Yes	−2641.498	−2597.528	1330.749	10

To account for heteroscedasticity, a sequence of 20 values starting from 2.78, which is the minimum value for the average values of the two methods, and 3.7 which is the maximum value for the average of the two methods was created. Then, the variance function (Section 2.1.2) was defined as $g(u_i, \delta) = |u_i|^\delta$, where $\tilde{u}_i = h(\bar{y}_{i1}, \bar{y}_{i2}) = \frac{\bar{y}_{i1} + \bar{y}_{i2}}{2}$. The variance function parameter \tilde{u}_i can also be chosen as the average values per subject of the reference. No significant changes are reported regardless of the choice of \tilde{u}_i . The parameter $\delta_1 = 4.07$ for Hanna and $\delta_2 = 3.28$ for the open-source device. The model's counterparts are displayed in Table 5.10.

Diagnostics for the optimal model (Figure C3, Appendix C) display the standardized residuals on the horizontal axis vs the quantiles of the standard normal distribution. The plot reveals a slight deviation from the normal distribution. The standard errors for the estimates are reasonable thus the agreement and similarity indices evaluation proceeds using this model.

Substituting the ML estimates from Table 5.10 in (2.21, 2.23) to obtain the fitted distribution (Y_1, Y_2) given the pH level \tilde{u} is

$$N_2 \left(\begin{pmatrix} -0.22 \\ 3.05 \end{pmatrix}, \begin{pmatrix} 0.04062325 + 2.615606 \times 10^{-8} \tilde{u}^{4.07} & 0.03934918 \\ 0.03934918 & 0.04062325 + 7.374885 \times 10^{-8} \tilde{u}^{3.28} \end{pmatrix} \right)$$

And of D given \tilde{u} is

$$D|\tilde{u} \sim N_1(-0.22, 0.002548149 + 2.615606 \times 10^{-8} \tilde{u}^{4.07} + 7.374885 \times 10^{-8} \tilde{u}^{3.28})$$

The intra-method differences distribution given \tilde{u} are produced by substituting the parameters from Table 5.10 in (2.35):

$$D_1|\tilde{u} \sim N_1(0, 5.231212 \times 10^{-8} \tilde{u}^{4.07})$$

$$D_2|\tilde{u} \sim N_1(0, 1.474977 \times 10^{-8} \tilde{u}^{3.28})$$

D_1 denotes the differences for Hanna pH meter and D_2 denotes the differences for the open-source device .

The range of \tilde{u} is taken to be $U = (2.78, 3.70)$, the observed measurement range. The findings from the exploratory analysis are confirmed from the model's ML estimates. Specifically, Hanna's estimated error standard deviation is larger than the open-source's. The error standard deviation for Hanna increases from 0.0104 to 0.0334. For the open-source the standard error deviation is smaller than Hanna's and increases from 0.0078 to 0.0199. Both error standard deviations are dominated by the between-subject standard deviation which is evaluated as $\sqrt{\exp(\sigma_b^2) + \exp(\sigma_\psi^2)} = 0.2011$.

Table 5.10. Model counterparts. The fixed bias is -0.13 . Recalibration actions can be taken to improve the agreement by adding the constant. The standard errors seem reasonable for the model's validity.

Parameter	Estimate	SE	95% Interval
β_0	-0.22	0.01	[-0.23, -0.20]
μ_b	3.27	0.04	[3.20, 3.35]
$\log(\sigma_b^2)$	-3.24	0.26	[-3.75, -2.72]
$\log(\psi^2)$	-6.67	0.26	[-7.18, -6.15]
$\log(\sigma_{\varepsilon_1}^2)$	-17.46	1.61	[-20.61, -14.31]
$\log(\sigma_{\varepsilon_2}^2)$	-16.42	1.61	[-19.58, -13.27]
δ_1	4.07	0.70	[2.71, 5.44]
δ_2	3.28	0.70	[1.91, 4.65]

5.6.3. Similarity Assessment

For the similarity assessment, two indices are evaluated. First, fixed bias which represents the difference in means of the two devices under the equal scale assumption (discussed in Section 2.1.3). Second, the precision ratio. Since the errors are heteroscedastic and the precision is defined as the error variance of Hanna over the error variance of the open-source, Figure 5.9 displays the precision as a function of magnitude of measurement. Table 5.11 summarizes the two indices. The open-source device is twice to three times more precise than Hanna. The fixed bias is -0.22 units for the open-source device compared to Hanna. The open-source device underestimates the pH measurement 0.22 units since the entire interval is below zero. Overall, the two devices cannot be considered similar. The open-source is considered superior because of its higher precision. These findings are consistent with the exploratory analysis.

Table 5.11. Precision ratio estimates as a function of magnitude of measurement. The open-source device is twice to three times more precise than Hanna. The fixed bias is -0.22 units for the open-source device compared to Hanna. The open-source device underestimates the pH measurement 0.22 units since the entire interval is below zero.

Similarity Assessment			
Grid		Lambda Estimate	Confidence Interval
	2.78	1.7987	[1.0397, 2.7374]
	3.70	2.8322	[1.7702, 5.4713]
		Estimate	Confidence Interval
	Fixed Bias	−0.2215	[−0.2398, −0.2033]

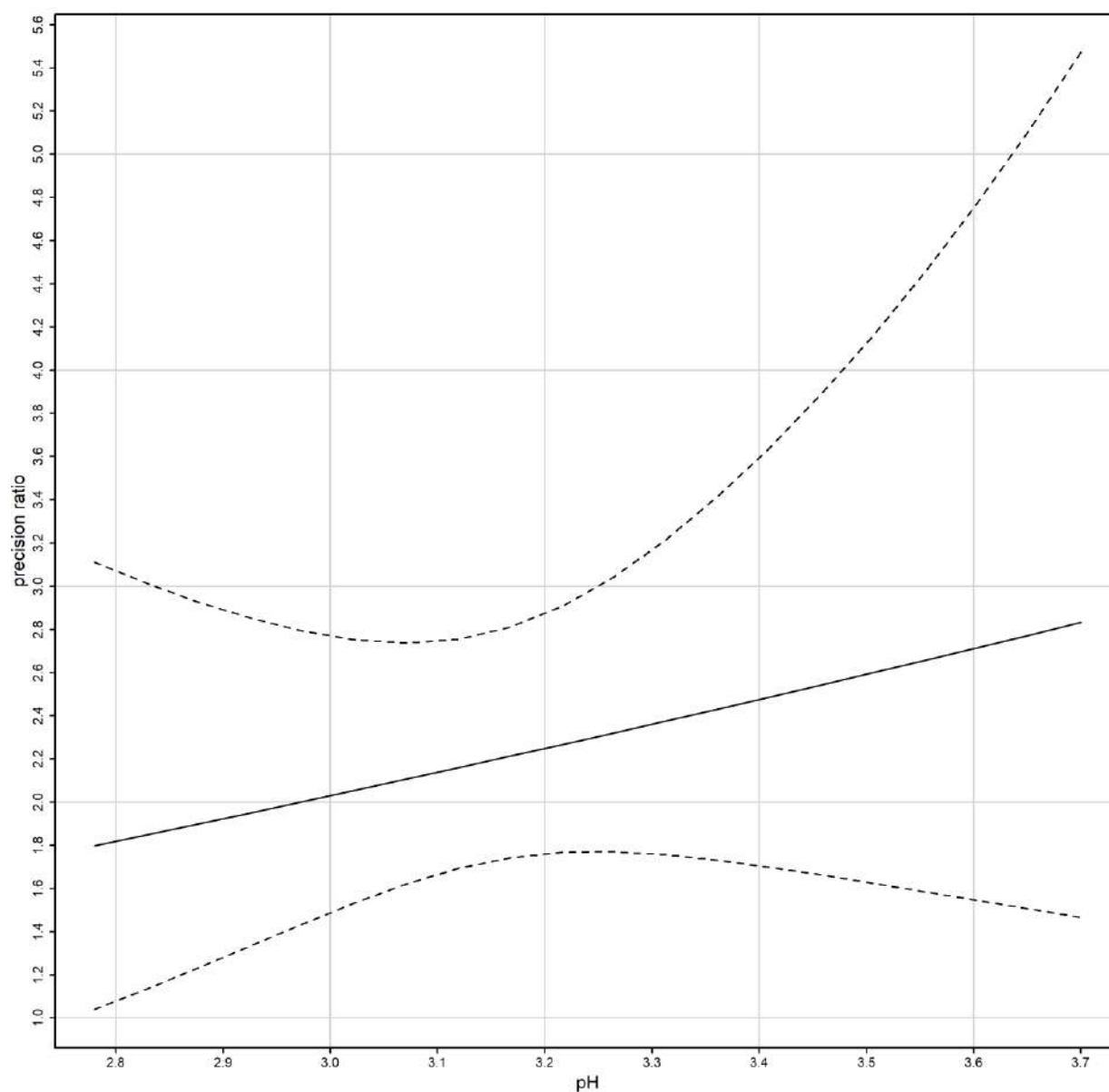


Figure 5.9. Precision ratio with the corresponding confidence intervals. The open-source device is more precise than Hanna, especially while the magnitude of the measurements increases. The open-source device is twice to three times more precise than Hanna.

5.6.4. Evaluation of Repeatability

CCC, TDI and the limits of agreement are calculated for the intra-agreement of each device separately. Figure 5.10 displays the limits of agreement as a function of magnitude of measurement. The limits of agreement for the open-source device are included in Hanna's LOA. Table 5.13 summarizes the ranges for the 95% limits of Agreement for pH data as a function of magnitude of measurement. The open-source device LOA are narrower comparing to the Hanna pH meter. Thus, the open-source device has better repeatability compared to Hanna. Based on the distributions of D_1 and D_2 , the intra-method limits are

centered at zero. In Figure 5.14a, CCC index is presented for inter- and intra- agreement. The dashed and dotted line represent the intra-method agreement for Hanna and open-source device respectively. For Hanna the lower bound ranges between 0.9955 and 0.9534 and for the open-source between 0.9975 and 0.9830. The intra-method agreement for both devices is considered excellent. However, the open-source device has higher intra- method agreement compared to Hanna. This conclusion is expected since the similarity assessment reported smaller error variation for the open-source device. Figure 5.14b, presents TDI (0.9). For the open-source device, which is represented by the dotted line, TDI (0.9) upper bound ranges between 0.0181 and 0.0462 while for Hanna, dashed line, between 0.0286 and 0.0945. The interpretation for TDI (0.9) is as follows : the bound of 0.0181 implies that 90% of the time the difference between two replications of the open-source device on the same subject falls within ± 0.0181 when the true pH value is 2.78. The TDI bounds for both devices are only 0.67 – 1.25 % of the magnitude of measurement, indicating a high degree on intra-method agreement. Table 5.12 displays CCC and TDI(0.9) along with their corresponding bounds for the minimum and maximum range of the measurements.

5.6.5. Agreement Assessment

Using formula (2.20,2.21) to calculate the limits of agreement using the model's counterparts in Table 5.10 and the variance function, the inter and intra agreement are displayed in Figure 5.10. For the inter-agreement of the devices, Table 5.12 summarizes the ranges for the 95% limits of Agreement for pH data as a function of magnitude of measurement. The inter-method limits, based on the distribution of D, are centered at -0.2215 . For lower pH values the LOA are narrower compared to the higher pH values and range from $[-0.3464, -0.3237]$ for lower LOA and $[-0.1193, -0.0966]$ for upper LOA. The intervals reveal a systematic underestimation of the pH measurements from the open-source device. Figure 5.11 presents Bland-Altman plot and limits of agreement and their corresponding confidence intervals superimposed. The two assumptions for the LOA are considered valid (discussed in Section 2.1.4) thus no regression is used. "Blandr" R package was used to produce the plot.

Table 5.12. Summary for CCC and TDI(0.9) for the intra- method agreement.

Grid	CCC for repeatability: Hanna	Lower Confidence Bound
2.78	0.9973	0.9955
3.7	0.9732	0.9534
Grid	CCC for repeatability: Open-Source	Lower Confidence Bound
2.78	0.9985	0.9975
3.7	0.9904	0.9830
Grid	TDI for repeatability: Hanna	Upper Confidence Bound
2.78	0.0243	0.0286
3.7	0.0777	0.0945
Grid	TDI for repeatability: Open-Source	Upper Confidence Bound
2.78	0.0181	0.0213
3.7	0.0462	0.0562

Table 5.13. 95% limits of agreement for inter-method Agreement for pH data as a function of magnitude of measurement. The inter-method limits, based on the distribution of D , are centered at -0.22. For lower pH values the LOA are narrower comparing to the higher pH values and range from [-0.3464, -0.3237] for lower LOA and [-0.1193, -0.966] for upper LOA.

Limits of Agreement						
Inter-Method						
	Minimum	Q1	Median	Mean	Q3	Max
Lower	-0.3464	-0.3369	-0.3305	-0.3321	-0.3263	-0.3237
Upper	-0.1193	-0.1167	-0.1125	-0.1109	-0.1060	-0.0966

Table 5.14. 95% limits of agreement for intra-method Agreement for pH data as a function of magnitude of measurement. The open-source device has narrower LOA compared to Hanna pH meter. Thus, the open-source device has better repeatability compared to Hanna. Based on the distributions of D_1 and D_2 , the intra-method limits are centered at zero.

Limits of Agreement						
Intra-Method						
Hanna	Minimum	Q1	Median	Mean	Q3	Max
Lower	-0.0926	-0.0713	-0.0540	-0.0564	-0.0400	-0.0289
Upper	0.0289	0.0400	0.0540	0.0564	0.0713	0.0926
Open-Source	Minimum	Q1	Median	Mean	Q3	Max
Lower	-0.0550	-0.0446	-0.0356	-0.0366	-0.0280	-0.0216
Upper	0.0216	0.0280	0.0356	0.0366	0.0446	0.0550

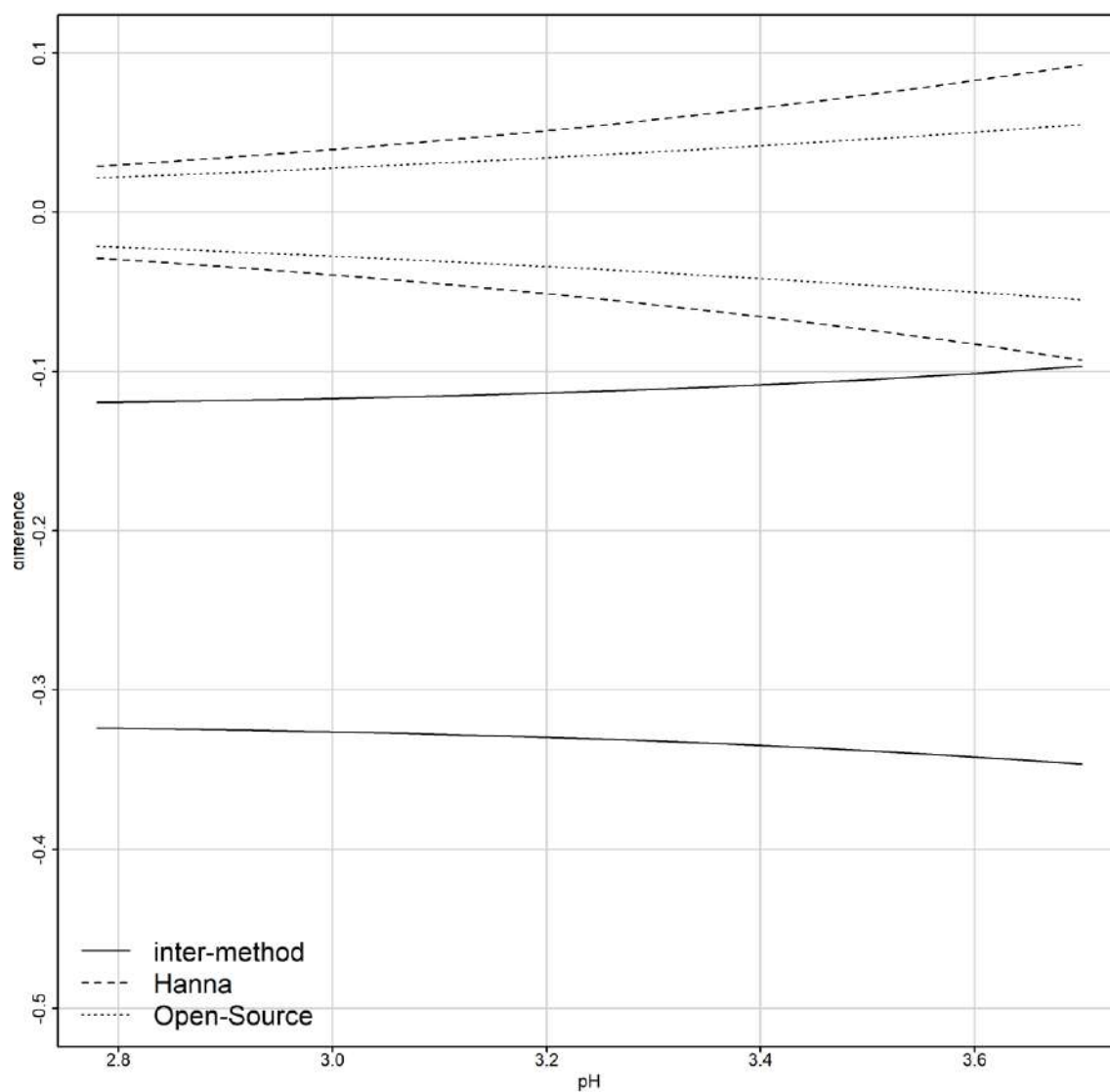


Figure 5.10. 95% Limits of inter- and intra- method agreement. The intervals reveal a systematic underestimation of the pH measurements from the open-source device. For the 95% limits of the intra-method agreement the open-source device has narrower LOA compared to Hanna pH meter; thus, the open-source device has better repeatability compared to Hanna. For lower pH values the LOA are narrower compared to the higher pH values.

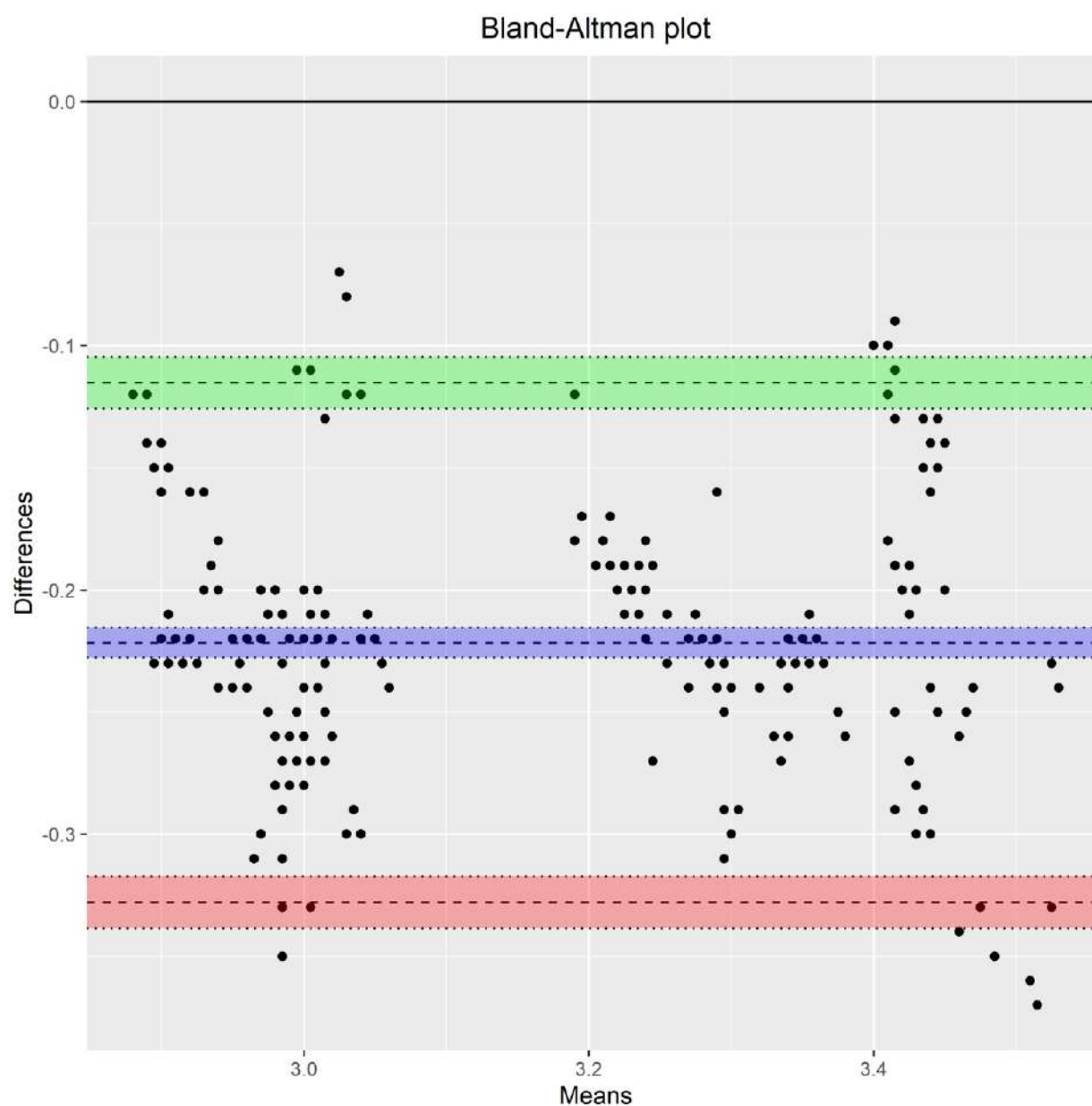


Figure 5.11. Bland - Altman plot using "blandr" package. Apart from the limits of agreement and the mean difference, their corresponding confidence intervals are plotted.

Table 5.15 presents CCC and TDI estimates, and lower and upper confidence bounds respectively before recalibration. The estimates for CCC range between 0.5970 and 0.6032 and the corresponding lower confidence bounds range between 04776 and 04839 throughout the pH measurement range. TDI (0.9)

estimates range between 0.2883 and 0.3031 and their corresponding upper confidence bounds range between 0.3095 and 0.3232 throughout the pH measurement range.

Figure 5.12a presents one-sided 95% pointwise confidence bands for CCC as a function of the magnitude of the measurements. The solid line represents CCC lower confidence bound for the inter-method agreement and ranges between 0.4776 and 0.4839. CCC lower band decreases as the pH level increases. Thus, the agreement becomes progressively worse but only by a small amount. The inter-method agreement is not considered satisfactory. Figure 5.12b presents the one-sided 95% pointwise upper confidence bands for inter- and intra- method versions of TDI (0.9) and their reflections over the horizontal line at zero. For the inter- method agreement TDI (0.9), which is represented by the solid line, upper confidence bound ranges between 0.3095 and 0.3232. As the pH level increases from 2.78 to 3.7 TDI increases. The bound of 0.3232 shows that 90% of differences in measurements from the devices fall within ± 0.3232 when the true value is 3.7. Such a difference is unacceptably large for many applications. The bounds of 0.3095 and 0.3232 are, respectively, 8.36 and 8.74 % of the true value. No big difference seems to exist for the inter-method agreement throughout the pH measurement range. The similarity evaluation reveals that a difference in the means of the devices is a contributor to disagreement. TDI and CCC will improve after recalibration.

Overall, as the magnitude increases TDI increases and CCC decreases. This means that the intra- method agreement becomes worse as the magnitude increases. High intra- method agreement CCC values reflect that the within-subject variations for both assays are very small compared to the between-subject variation.

Table 5.15. CCC and TDI estimates with their corresponding lower and upper confidence bounds throughout the pH measurement range.

Before Recalibration		
Grid	CCC	Lower Confidence Bound
2.78	0.6032	0.4839
3.7	0.5970	0.4776
Grid	TDI	Upper Confidence Bound
2.78	0.2883	0.3095
3.7	0.3031	0.3233

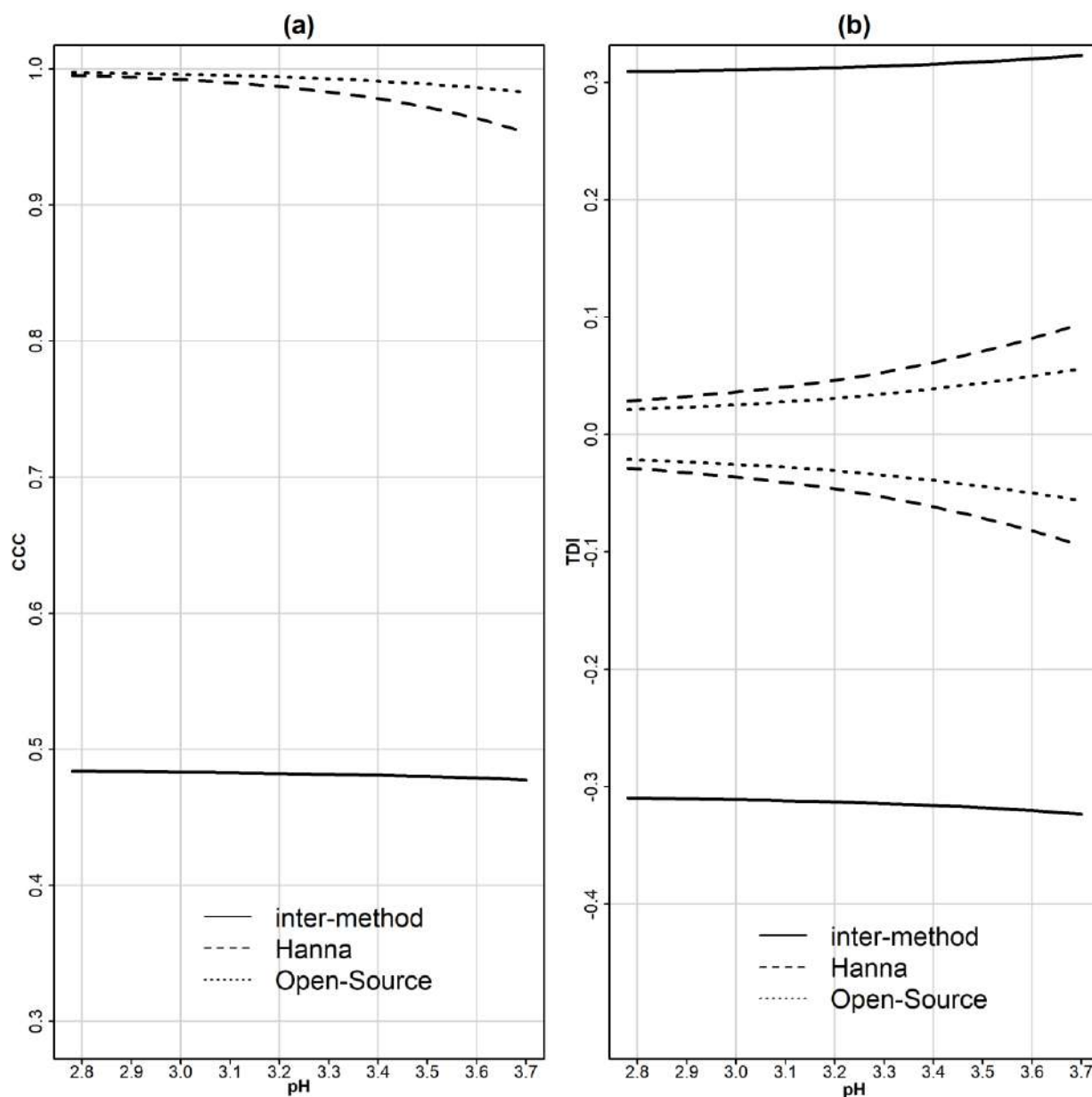


Figure 5.12. (a) One-sided 95% pointwise confidence bands for CCC as a function of the magnitude of the measurements. The solid line represents CCC lower confidence bound for the inter-method agreement and ranges between 0.4776 and 0.4839. The agreement is considered insufficient. The dashed and dotted line represents the intra-method one-sided 95% pointwise confidence band for Hanna and open-source respectively. For Hanna the band ranges between 0.9979 and 0.9987 and for the open-source between 0.9986 and 0.9997. The intra-method agreement for both devices is considered excellent. (b) One-sided 95% pointwise upper confidence bands for intra-method versions of TDI (0.9) and their reflections over the horizontal line at zero. The open-source device has higher intra- method agreement compared to Hanna.

Using Escaramis et al. (2010a) methodology to calculate TDI (0.9), the estimate is 0.2767 and the upper bound is 0.2880. Table 5.16 summarizes the results. Compared to Choudhary's and Nagaraja's

methodology, Escaramis TDI (0.9) is a little lower. This is probably because the devices are assumed to have equal variances compared to (2.11) since the mixed-effects model formulation is not the same.

Using Carrasco's methodology to evaluate CCC using U statistic, CCC estimate is 0.6024 and the corresponding 95% confidence interval is [0.5328 0.6639]. Carrasco's method provides similar results compared to Choudhary's and Nagaraja's methodology estimate.

Last, Figure 5.13 displays the Bias-plot. The proportional bias is 0.965 (95% CI [0.9352, 0.9938]) and the fixed bias -0.1052 (95% CI $[-0.2013, -0.091]$). The fixed biases estimate is different compared to Choudhary's and Nagaraja's method probably because the parameter estimation method is different. However, the red solid line which corresponds to the total bias confirms the findings in Table 5.11 since the total bias ranges from $[-0.235, -0.21]$.

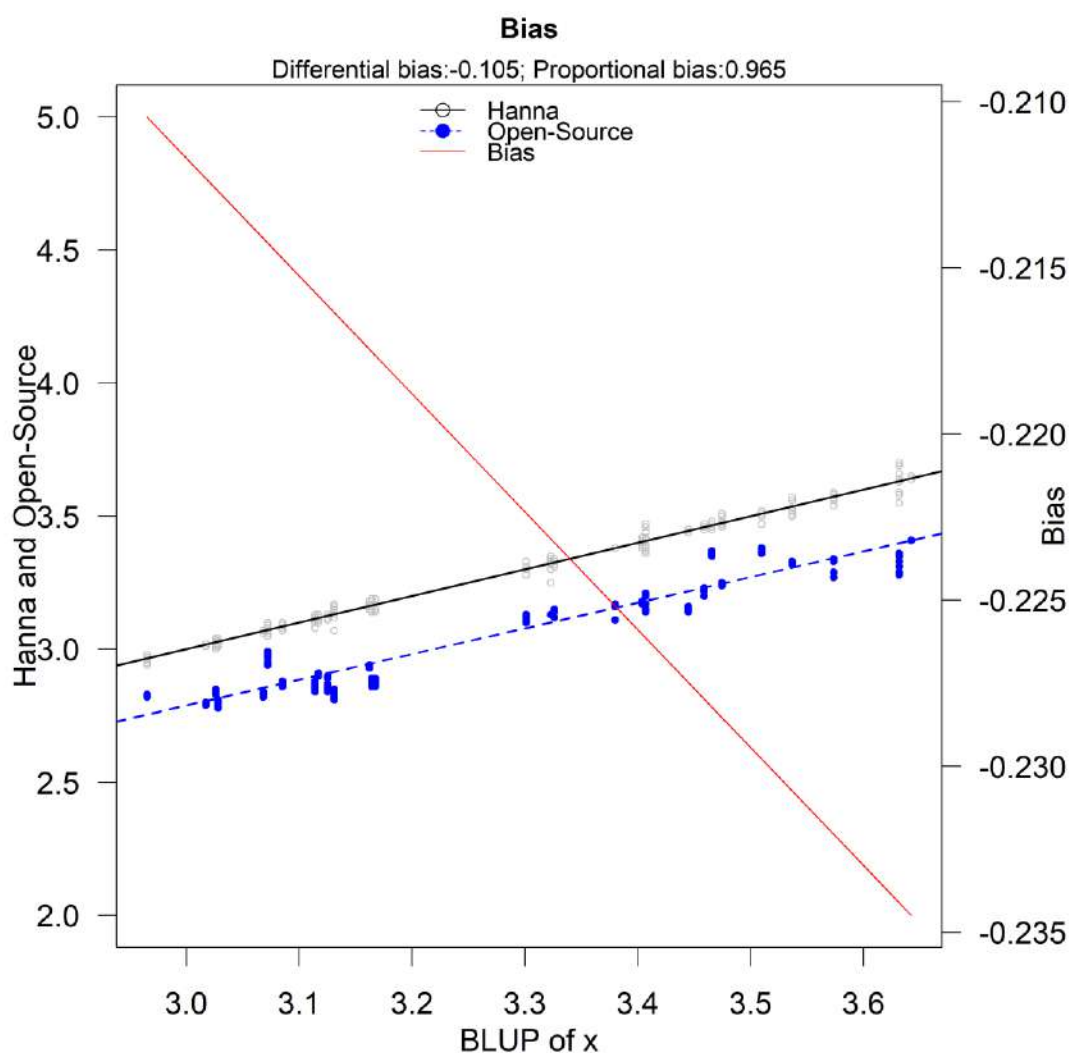


Figure 5.13. Bias-plot. The proportional bias is bias is 0.965 (95% CI [0.9352, 0.9938]) and the fixed bias -0.1052 (95% CI $[-0.2013, -0.091]$). The fixed biases estimate is different compared to Choudhary's and Nagaraja's method. However, the red solid line which corresponds to the total bias confirms the findings in Table 5.11.

5.6.6. Recalibrating the open-source device

The similarity evaluation reveals that the fixed bias (difference in the means) contributes to the disagreement between the two devices. Recalibration of the open-source devices by subtracting -0.22 from its measurements makes the mean difference zero and improves the extend of agreement substantially. Table 5.17 reports CCC and TDI estimates and confidence bounds after recalibration. CCC improves significantly. The lower confidence bands range from 0.9194 to 0.9407 revealing excellent agreement throughout the measurement range. TDI also improves and ranges from 0.1052 to 0.1215 throughout the measurement range. The agreement for this research's case study is considered acceptable. TDI (0.9) and CCC were also calculated after recalibration using methodology developed by Escaramis and Carrasco respectively (Table 5.18). They are both close to Table 5.17 values with TDI (0.9) and CCC lower than Table 5.17 values.

Table 5.16. Summary of TDI (0.9) and CCC with their corresponding bounds using Escaramis and Carrasco's methodology respectively.

Before Recalibration	
TDI (0.9) - Escaramis	
Estimate	0.2767
Upper Bound	0.2880
CCC - Carrasco	
Estimate	0.6024
Confidence Interval	[0.5328, 0.6639]

Table 5.17. CCC and TDI estimates after recalibration. Both indices improve significantly and report excellent agreement between the devices.

After Recalibration		
Grid	CCC	Lower Confidence Bound
2.78	0.9666	0.9407
3.7	0.9509	0.9194
Grid	TDI	Upper Confidence Bound
2.78	0.0857	0.1052
3.7	0.1048	0.1215

Table 5.18. TDI (0.9) from Escaramis and CCC from Carrasco after recalibration. The agreement improved significantly and is now considered excellent.

After Recalibration	
TDI (0.9) - Escaramis	
Estimate	0.0707
Upper Bound	0.0734
CCC - Carrasco	
Estimate	0.9641
Confidence Interval	[0.9510,0.9740]

5.7. Discussion

Soil analysis is very important before the establishment of cultivation. Soil pH is one of the main measurements of the soil analysis. Soil pH measurement indicates whether the soil is acidic or basic but also shows the availability of essential nutrients and toxicity of other elements if their relationship with pH is known (Thomas, 2018). The knowledge of soil pH reveals the possible necessity of its adjustment especially in extreme soil pH values (acidic or alkaline). Moreover, soil pH determines the type of fertilizers that can be used to ensure nutrient availability and a balanced plant nutrition. By monitoring the pH values regularly (every 2 to 3 years) in a field, it might be essential to raise the pH of the soil from acidic to near neutral pH values by liming (Jensen, 2010). Apart from the effect of soil pH to nutrient absorption by plant, it is mentioned that substrate pH may also influence the rhizosphere microbiome (Ferrarezi et al., 2022). In hydroponics, the pH of the nutrient solution is very important to reach a balanced nutrition for the cultivation. In-situ analysis methods based on pH sensors can improve efficiency and reduce costs. Compared to soil, soilless culture substrate typically has a relatively simple, stable, and artificial-controlled composition thus, mitigating the disadvantages of direct measurement method. Therefore, direct measurement appears to be suitable for frequent, rapid, and accurate pH measurement in soilless culture substrate. However, the large error of this technique still needs to be solved (Chen et al., 2021).

A portable open-source device that measures the pH of the juice of grapefruits and oranges was designed and constructed for laboratory experiments and in situ measurements. To evaluate its functionality, a method comparison study between the open-source device and a corresponding industrial was designed. The statistical analysis to assess their agreement was based on indices and graphical methods using mixed-effects models. The agreement indices evaluated were the Concordance Correlation Coefficient (CCC) and the Total Deviation Index (TDI). TDI estimates and confidence bounds were evaluated using (2.11) and methodology described in Section 2.1.5 and by methodology proposed by Escaramis et al. (2010a) described in Section 2.1.5. There were small differences between the two methods probably due to the mixed-effects models' different formulation.

Overall, agreement between the two devices is not satisfactory but improves to excellent levels after recalibration since the main source of disagreement is the fixed bias (0.22 pH units).

Further experiments can be conducted to investigate the agreement for an extended range of measurements and a greater variety of fruits. Moreover, the open-source device can be used in other applications that include soil pH or substrate pH in soilless cultivations and future method comparison

studies can be designed to assess the agreement and similarity between the two devices for various applications.

6. FruiTemp: Design, Implementation and Analysis for an Open-Source Temperature Logger Applied to Fruit Fly Host Experimentation

6.1. Goal of the Study

The goal of this study was to create an inexpensive custom-made device that records core, surface temperature and external air temperature of a fruit and test it in laboratory conditions and further in field experiments of interest. The essential specifications for the device are portability, power efficiency, user-friendly interface, and a weatherproof enclosure. The sensor specifications include sensor tips as less invasive as possible to reduce the insertion hole diameter on the fruit, accurate and precise readings. Additionally, an important consideration was the reduction of its development cost that would make feasible to bulk produce in-house for increased sample size in experimental trials.

The Mediterranean fruit fly (medfly), *Ceratitidis capitata* (Wiedemann) (*Diptera Tephritidae*) is one of the most notorious agricultural pests worldwide. Medfly originated from the Afrotropical region and exist in wide range of climates in many parts of the world (Liquido et al., 1991; White & Elson-Harris, 1992; De Meyer et al., 2002; Papadopoulos et al., 2013). It is an extremely polyphagous species that completes many generations per year according to environmental conditions. It can infest approximately 353 hosts from 67 different families (Liquido et al., 1991; Papadopoulos et al., 1996; Benjamin et al., 2015). It is extremely damaging for many fruit species and cultivars since the infestation rates may reach 100% (White & Elson-Harris, 1992; Fimiani, 1989; Tzanakakis & Katsoyannos, 2003). In Greece, it completes 3–7 generations per year. It overwinters at the larval stage, in locations where the temperature during the cold months is low, inside infested fruits, mostly in apples and bitter oranges (Papadopoulos et al., 1996). Those fruits, such as apples remaining on the trees in autumn, often after harvest, provide the last breeding resources for female medflies and they receive the last ovipositions (Tzanakakis & Katsoyannos, 2003). Prevailing climatic conditions determine both developmental process and survival of these last immatures and whether the life cycle can be completed yielding adults, next spring. The rate of successful overwintering and hence the fate of these individuals largely determine the population growth and infestation rates of the next season (Papadopoulos et al., 2001).

Being ectotherm, as all arthropods, egg and larvae development and survival of medflies depend on the temperature conditions within the last infested fruit. Living at the edge of environmental stress (i.e., winter conditions of temperate areas) small fluctuations may become detrimental for immatures survival within fruit. On the other hand, collectively small fluctuation may greatly affect the developmental process that is extremely low under such conditions (Papadopoulos et al., 2003). Hence, acquiring precise data on temperature fluctuation is important for studying overwintering dynamics for medfly and other similar insect species. Thus far, most population growth models rely on air or soil temperature data and often place little emphasis on the overwintering generation and its particulates (Gutierrez et al., 2009; Gutierrez & Ponti, 2011; Ponti et al., 2009; Vera et al., 2002). As mentioned above, medfly's larvae overwintering success is determined by the temperature and the type of the host (Papadopoulos et al., 1996, 2001, 2003). Thus, measuring the core temperature of the fruit and comparing it with its surface temperature and the temperature of the environment, would greatly enhance our understanding of the overwintering dynamics of medfly in temperate areas and would pave the way for furthermore detailed approaches on insects overwintering. Agreement between core temperature and fruit surface temperature needs to be established to use the latter as a proxy for the former in fruit fly related experiments allowing for the least possible damage and intervention regarding the host.

An open-source logger integrated with precise and slightly intrusive point medical temperature sensors was developed to conduct a series of experiments regarding the FF-IPM project (FF-IPM). There are many devices available in the market to measure the inner fruit temperature (Hengko, n.d.) but most of them are either expensive, highly intrusive or incapable of logging temperature in remote places or outdoors in general, such as a tree, for many days. FruiTtemp is an open-source system designed to measure accurately and precisely the temperature inside the core and the surface of a fruit and the environment surrounding it. Thermistor probes were used to precise the estimation of the temperature inside the apples which are considered ideal winter hosts of medfly.

The workflow of this project is summarized in Figure 6.1.

6.2. Design of the device

The device (Figure 6.2, 6.3a,b) is a portable, waterproof temperature data logger that is equipped with two fine precision medical temperature sensors and a harsh environment temperature probe. It can log data according to the interval prespecified by the user and has low power consumption needs. It is equipped with a real time clock powered by a backup battery to include a time stamp for every reading and a micro-SD card module to store the data. Further, it has an on/off waterproof switch and is enclosed in an IP54 flanged box.

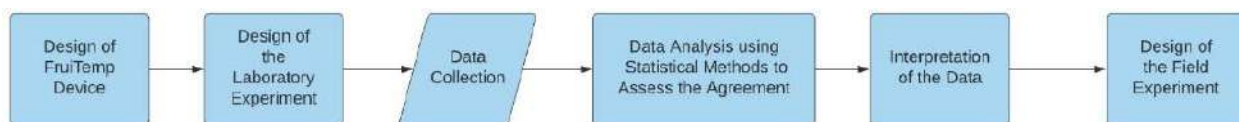


Figure 6.1. Initially the device was designed according to specifications based on the nature of the experiment. Then, the laboratory experiment was designed, and data collection followed. Data analysis involved statistical methods to assess the Agreement. The experiment's findings confirmed the initial hypothesis thus designing the field experiment.

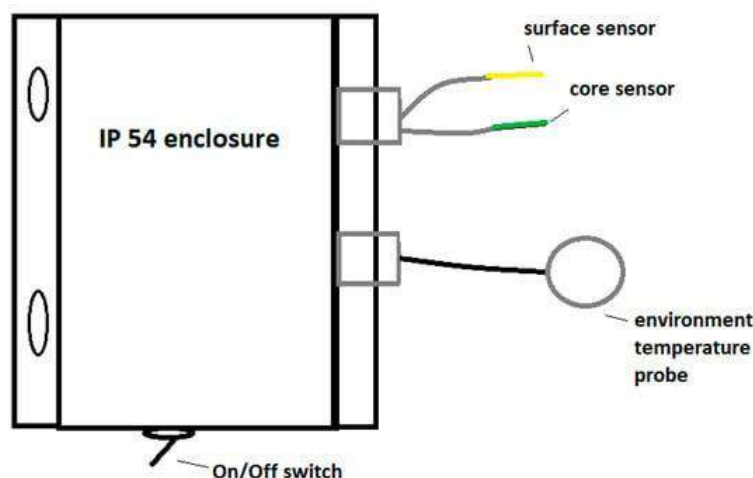


Figure 6.2. The device consists of two precise medical sensors inserted in the fruit and an environmental temperature probe. There is also an On/Off waterproof switch. The box is waterproof (IP 54) and flanged.

6.2.1. Hardware

The main parts of the system are the microcontroller, Adafruit feather 32u4 proto board, (Fried, 2023a) which is the processing module that coordinates all the others. The second most important parts are the Amphenol NTC thermistor MC65 series (Amphenol, 2014) and the Amphenol Industrial Temperature Sensor JS8746A (Amphenol, 2019). The former sensor is a very fine point medical temperature sensor (1.65 mm diameter maximum), and the latter is a harsh-environment sensor for air temperature with a cylindrical enclosure. The size of the MC65 series makes it the best available solution for the experiment, which requires the least intrusive way to measure the temperature of the fruit's core. Furthermore, a smaller diameter thermistor for the sensing part was not chosen although available, because, as the thermistor is inserted in the core of the fruit it might be damaged due to the friction between the flesh of the fruit and the thermistor during its insertion. The vital parts that contribute to the precision and accuracy of the system are the ADS1115 16-bit ADC—4 channel with Programmable Gain Amplifier (Texas Instruments, 2009)—which increases the ADC resolution from 10 to 15 bits and amplifies the analog signal of the sensor and the LM4040 Voltage Reference Breakout (Texas Instruments, 2013), which stabilizes the reference voltage to achieve more accurate readings. Table A4 (Appendix A) displays the specification of the thermistors used.

The battery used for the system to be portable is a 1200 Mah Li-po battery, which can be recharged by means of a micro-USB connector while in operation. The use of headers in a Printed Circuit Board (PCB) allows researchers to easily replace parts that might be damaged or not working properly. A detailed list of the components can be found in Appendix B. Moreover, since the Adafruit feather series has the same footprint for every module, the microcontroller board can be switched with a wireless development board without affecting the design of the hardware.

The design is simple but efficient and is based on open-source technology. The MCU used is a 32u4, 8-bit microcontroller, clocked in 8 Mhz. This choice was based on low consumption of the microcontroller and the vast number of guides and tutorials available that can provide guidance even for amateur users. Since the 32u4 is equipped with a 10-bit analog to digital converter and the input range to the ADC is 0 to 2.048 Volts, a 16-bit analog to digital converter (ADS1115) equipped with an amplifier was added to the device. The readings from the thermistor are sent directly to the ADS1115, which uses an I2C connection with the microcontroller and the signal is amplified and digitized using 15 bits, instead of the microcontroller's 10-bit ADC. The amplification gain according to the datasheet and the input voltage from the thermistor is 2, which means that the Vcc used to calculate the value of the thermistor in volts (Table A6, Appendix A, ADC value formula) is 2.048 Volts. Moreover, since the microcontroller reference voltage is not stable, an LM4040 voltage reference breakout is used (0.1% output voltage tolerance) by defining the reference as 2.048 volts. Last, all the reference resistors used to create a voltage divider to read the resistance of each thermistor are of high precision and are described in detail in Table A4 (Appendix A). A logger shield was used to write all the data to a micro-SD card integrated with a timestamp. The current time was acquired by the real time clock PCF8523. The logger shield uses a 3 Volt CR1220 battery to keep track of the time even if the power is cut from the microcontroller to the logger shield. The schematic is available in Figure A1 (Appendix A).

Physical Features of the Device

The dimensions of the device (Figure 6.3b) are 136 mm × 83 mm × 44 mm. The box is equipped with mount holes useful in securing the device on the tree with nylon plastic cable ties. The MC65 (Figure 6.3c) and JS8746A (Figure 6.3d) sensors' length is 400 mm and 1000 mm, respectively. Heat-Tubes are used to protect the MC65 sensors and silicone is applied in the junction between the sensor and the heat-tube to make it stiff and protect it from damage while we insert it in the fruit.

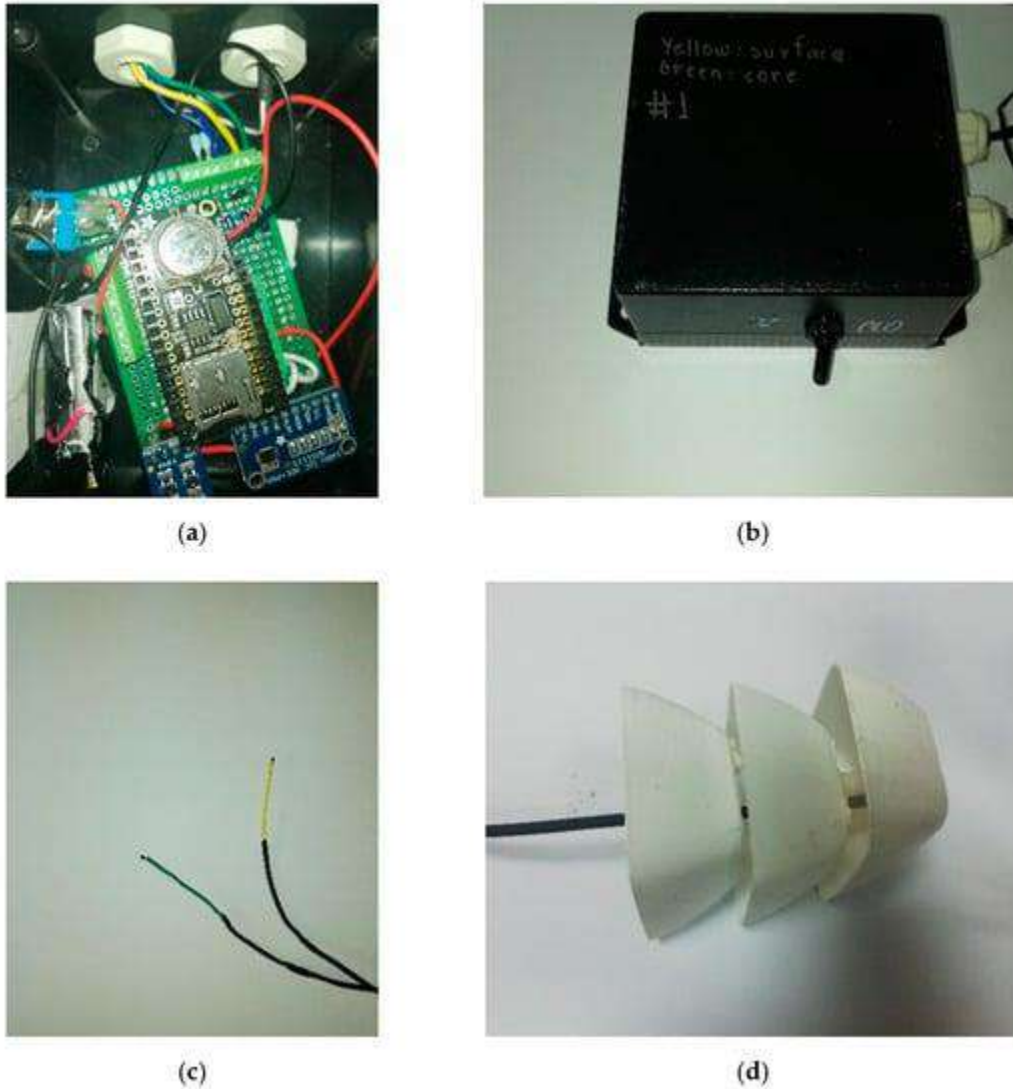


Figure 6.3. (a) The data logger interior consists of an Adafruit 32u4 proto, an Adafruit FeatherWing logger, an ADS1115 ADC with a gain amplifier and an LM4040 Voltage reference breakout. (b) The case of the system is an IP 54 flanged box. (c) The MC65 precise medical sensors. (d) The JS8746A environment temperature probe with its enclosure.

6.2.2. Software

The datasheet of the ADS1115 includes the choice of samples per second that has a range of 8 to 860. The Adafruit library for the ADS1115 (Herrada, 2023) was used and modified in the main code, which by default uses 128 samples per second.

The logging interval is adjusted according to the users' needs. For the FF-IPM experiments the device was configured for 1 and 15-minute intervals.

The code was based on GitHub (OSBSS, 2018) and the instructions to provide it are in GitHub (Bataka, 2023). The GitHub example is based on another thermistor of the manufacturer Vishay, the ADS1115 usage, the Hoge-2 equation, the resistance reading and the manual sample average. Logger functionality was implemented by adding the Real Time Clock and the SD-card code. Sleep functionality for the MCU was used to reduce power consumption. Reduction of the self-heating effect functionality was added by powering the thermistor only when a measurement is taken. The coefficients of the thermistors were configured accordingly using the method described in Section 6.2.3. The code is also available in GitHub (Bataka, 2023).

Code Description

First the libraries used in the sketch are included.

```
#include <Wire.h>
#include <SD.h>
#include <Adafruit_ADS1015.h>
#include "RTCLib.h"
#include <stdio.h>
#include <Adafruit_SleepyDog.h>
```

Libraries	Use
<Wire.h>	library for the I2C
<SD.h>	library for the SD module
<Adafruit_ADS1015.h>	library for the ADS1115 ADC
"RTCLib.h"	library for the real time clock
<stdio.h>	library for additional c++ functions
<Adafruit_SleepyDog.h>	library for the sleep mode

```
#define VBATPIN A9
```

```
RTC_PCF8523 rtc;
Adafruit_ADS1115 ads;
File myFile;
```

Construct objects
from the
corresponding
constructors.

A voltage divider is connected between the BAT pin (pin where the battery is connected) and pin 9. Thus, this is the pin where we will retrieve the battery voltage using the battery() function.

```
void setup() {
```

```
  pinMode(6, OUTPUT);
  digitalWrite(6, LOW);
```

```
  Wire.begin();
  Serial.begin(115200);
  ads.begin();
  ads.setGain(GAIN_ONE);
```

```
  if (!SD.begin(10)) {
    while (1);
  }
```

```
  //rtc.adjust(DateTime(F(__DATE__), F(__TIME__)));
```

```
}
```

```
void loop() {
```

Initializing
constructors for
Wire, ADS1115,
SD card, RTC.

```

//Entering sleep Mode
for(int index=0;index<7;index++){
Watchdog.sleep(8000);
}
//Retrieving Date and Time
DateTime now = rtc.now();
//powering the thermistors
digitalWrite(6,HIGH);
delay(1000);
int16_t adc0_average = samples(0);
int16_t adc1_average = samples(1);
int16_t adc2_average = samples(2);
digitalWrite(6,LOW);
// powering down the thermistor
// get the resistance
float R0_avg = resistance(adc0_average);
float R1_avg = resistance(adc1_average);
float R2_avg = resistance(adc2_average); // get the temperature using the
Hoge-2 equations
float temperature0_avg = hoge_ext(R0_avg); //for the external
float temperature1_avg = hoge(R1_avg); //for the MC65
float temperature2_avg = hoge(R2_avg); //for the MC65
writesd(temperature1_avg,temperature2_avg,temperature0_avg,battery());
}

```

Perform multiple iterations to get higher accuracy ADC values

```

//Calculating battery capacity
float battery(){
float measuredvbat = analogRead(VBATPIN);
measuredvbat *= 2; // we divided by 2, so multiply back
measuredvbat *= 3.3; // Multiply by 3.3V, our reference voltage
measuredvbat /= 1024; // convert to voltage
return(measuredvbat);
}

```

Function to read battery voltage.

```

// Get resistance -----//
float resistance(int16_t adc)
{
float ADCvalue = adc*(4.096/2.03); // Vcc = 4.096 on GAIN_ONE setting, Arduino Vcc = 3.3V in this case
float R = 10000/(32767/ADCvalue-1); // 32767 refers to 15-bit number
return R;
}

```

Function to read the resistance.

```

// Get temperature from Steinhart equation -----//
float steinhart(float R)
{
float Rref = 10000.0;
//float A = 0.003368620;
float A = 0.0033540154;
//float B = 0.00024057263;
float B = 0.00025627725;
//float C = -0.0000026687093;
float C = 0.000002082921;
// float D = -0.00000040719355;
float D = 0.000000073003206;
float E = log(R/Rref);

```

Function that uses the resistance as input and outputs temperature using Steinhart's equation.


```

    float T = 1/(A + (B*E) + (C*(E*E)) + (D*(E*E*E)));
    return T-273.15;
}

```

```

void writesd(float t1, float t2, float ext, float measuredvbat)
{
    myFile = SD.open("fruit1.txt", FILE_WRITE);
    // if the file opened okay, write to it:
    if (myFile) {

        DateTime now = rtc.now();
        myFile.print(now.year(), DEC);
        myFile.print("/");
        myFile.print(now.month(), DEC);
        myFile.print("/");
        myFile.print(now.day(), DEC);
        myFile.print(" ");
        myFile.print(now.hour(), DEC);
        myFile.print(":");
        myFile.print(now.minute(), DEC);
        myFile.print(":");
        myFile.print(now.second(), DEC);
        myFile.print(" , ");
        myFile.print(ext,3);
        myFile.print(",");
        myFile.print(t1,3);
        myFile.print(",");
        myFile.print(t2,3);
        myFile.print(",");
        myFile.print(measuredvbat);
        myFile.println();
        Serial.println("DONE writing on sd");
        // close the file:
        myFile.close();
    } else {
        // if the file didn't open, print an error:
        Serial.println("error opening test.txt");
    }
}

```

Function to
write the
data in the
microSD
card.

```

// Perform multiple iterations to get higher accuracy ADC values (reduce
noise) -----//
int16_t samples(int pin)
{
    int n=10; // number of iterations to perform
    int32_t sum=0; //store sum as a 32-bit number
    for(int i=0;i<n;i++)
    {
        int16_t value = ads.readADC_SingleEnded(pin);
        sum = sum + value;
        delay(10);
    }
    int32_t average = sum/n; //store average as a 32-bit number
    return average;
}

```

Function
that
performs
multiple
iterations to
get higher
accuracy
ADC values
to reduce
noise.

6.2.3. Calibration Method

The thermistors were calibrated using conventional non-expensive methods by accessing the sensor's temperature resistance curves (Thermometrics, 2014). The procedure was based on Ohm's Law and the Hoge-2 equation (Hoge, 1988). Only the MC65 thermistor is described since the JS8746A was calibrated the same way.

First, resistance ratio of the thermistor was recorded in a range of 0 to 44°C to choose the category of coefficients according to Table A5 (Appendix A). This method is implemented by using a voltage divider and the formula R_t in Table A6 (Appendix A). The reference resistance is 10 KOhms with a tolerance of 0.05%. The ADCvalue in the formula corresponds to the readings of the ADS1115 analog to digital converter (15 bits of precision and 1 bit for the sign) and amplifier which was used to increase the precision and resolution of the reading since the ADC of the microcontroller has only 10 bits. The ADC value was calculated using the formula in Table A6. Vcc on gain two is 2.048 as the dataset of the ADS1115 suggests. The signal was amplified by only a gain of two since the output of the thermistor can only range between 0 and less than 2.048 volts. LM4040 was used to set a reference of 2.048 with a 0.1% tolerance. This addition improved the accuracy and precision of the system since it keeps the voltage reference stable, and the values achieve less fluctuation than using the MCU's reference voltage. The first formula on Table A6 calculates the resistance of the thermistor. The value 10,000 is the reference resistance and $2^{15} - 1$ which is the ADC number of bits.

The material of the thermistor is type F. For that reason, Table A5 was used to choose which coefficients were appropriate to form the Hoge-2 equation which is the best calibration equation comparing to nine more for the MF501 NTC thermistor according to Liu (2018). For the present application the values of the resistor ranged between 0.5 to 3.223 so the second row was chosen (0.36035 to 3.274), this ratio is close to optimal linearity, according to Rudtsch and von Rohden (2015) reporting that for a microK-type of instrument, optimum linearity is achieved in the resistance ratio range between 0.2 and 1.2. Figure 6.4 summarizes the calibration procedure.

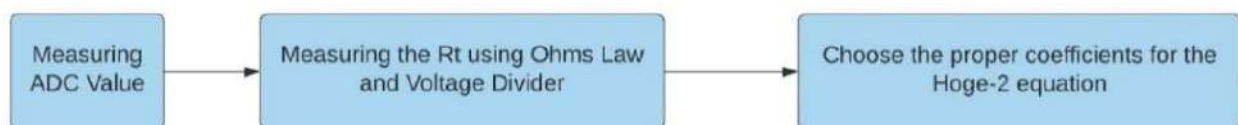


Figure 6.4. The calibration procedure. First, the ADC value is measured. Second, R_t is retrieved using Ohm's Law and a voltage divider. Last, proper coefficients from the manufacturer are chosen as an input to Hoge-2 equation.

6.2.4. Power Consumption and Power Saving Techniques

FruiTempt runs on a 1200 mAh Li-po battery which can be recharged by means of a micro-USB connector while in operation. Power consumption reduction was achieved using software and hardware techniques. The microcontroller periodically goes to sleep, and the watchdog timer (Matalon, 2016) awakes it for temperature sampling. The microcontroller is clocked at a low frequency of 8 Mhz. When the device is measuring and recording, which means that it is awake from the sleep function, it takes 6 s to measure and log the data and has an averaged power consumption of approximately 9.735 mA. When the device enters sleep mode the power consumption drops to an average of 1.944 mA approximately. The average consumption of the device including both modes is 3.148 mA, for a logging interval of 1 sample per minute. Unfortunately, the watchdog library will enable sleep mode in increments of 8 seconds. For example, if the user wishes to wake the board every 5 minutes, a loop needs to be created that will run 38 times to reach 5.06667 seconds. Thus, the MCU will not be able to wake up exactly after five minutes. Moreover, every MCU will act differently after the watchdog function is enabled. In Section 6.7 a solution is proposed.

6.2.5. Cost of the Device

The trade-off between cost and linearity (Amphenol, 2021) affected the design. The thermistor manufacturer suggests a circuit including a differential amplifier that improves the linearity of the thermistor. Since the ADS1115 includes 4 single channels or 2 differentials, the single channels were used, since the sensors were 3, leaving one channel unused. By using 2 differential channels, only 2 thermistors can be supported.

Other devices, like HOBO MX2303 which is weatherproof, power efficient, has high accuracy (± 0.2 °C) and has an additional Bluetooth connection, could be a good fit to the experiments demand but it is only equipped with two probe sensors instead of three, the diameter of the sensors (0.53 cm) is wider than the proposed system's internal sensors (0.165 cm) and the housing of the sensor is a cylindrical stainless steel compared to the proposed system's internal sensor, which is a point epoxy material which senses a significantly smaller area than the HOBO MX2303. Overall, the HOBO MX2303 is inappropriate because it provides only two sensors, compared to our device, which provides three, and the sensors of the HOBO MX2303 are much more intrusive compared to the proposed systems sensors. Last, the Hengko temperature and relative humidity data recorder sensor is specifically made for fruits and vegetables, but it only provides one sensor, has an a non-acceptable for the experiment accuracy (± 0.5 °C) and its probe is highly intrusive compared to the proposed system.

The total cost was around EUR 110 (end 2020 market costs) which can be reduced substantially by not using breakout and development boards (for example Adafruit 32u4 proto), instead create a PCB that include only the vital parts for the system that will also decrease the size of the device. The costs of the individual parts are displayed in Table 6.1.

Table 6.1. FruiTTemp individual parts' cost.

Part	Cost (In Euro)	Quantity
Adafruit ADS1115 16-bit ADC-4 channel PGA breakout board	18.4	1
Adafruit LM4040 breakout board	6.35	1
Adafruit Feather 32u4 Proto	16.9	1
Adafruit Adalogger Feather Wing	7.58	1
Amphenol NTC Thermistor MC65 series	6.59	2
Amphenol Industrial Temperature sensor JS8746A	18.91	1
Metal Film Resistor Vishay/Dale 10Kohms, 0.05% tolerance 5ppm	2.18	2
Metal Film Resistor Vishay/Dale 10Kohms, 0.1% tolerance 5ppm	1.13	1
Lithium Polymer single cell 3.7 Volt/1200 mah	8.9	1
Switch and waterproof cap	0.6	1
Cable glands PG7	0.6	1
IP54 Hammond flanged box	6.78	1
Wires/PCB/Heat Tubes	4	1
Total	107.69	

6.2.6. Bias Removing Techniques, Accuracy and Measurement Range

According to Ebrahimi-Darkhaneh (2019), a thermistor dissipates power in the form of heat when current flows through it due to its nature since thermistors are resistors. For that reason, the input voltage of the LM4040 was connected to pin 6 of the MCU and adjusted the code to power the thermistor only when a reading is going to take place.

Manual averaging was also added to the code. The function was getting 50 readings using a 10-millisecond interval, which according to Goumopoulos (2018), for the thermistor the author used, 50 readings for every cycle delivered an optimal accuracy after proper calibration.

The MC65 thermistor reports an accuracy of 0.05°C and the JS8746A reports an accuracy of 0.15°C according to the manufacturer's datasheet. The external ADC and the bias removing techniques used are not expected to add extra bias. The measuring range of the MC65 as reported in the datasheet is -40°C to 105°C and for JS8746A is -40°C to 120°C, but since the coefficients used to form the Hoge-2 equations were retrieved for a range of 0°C to 44°C, the present configuration suggests using this measurement range and is not limiting the device since it can be programmed to switch the parameters when the temperature is below 0°C or above 44°C.

6.3. Design of the Method Comparison Study: Measuring the Agreement Between the Core and the Surface Thermistor.

The difference between the core and the surface temperature of apples was tested by inserting the green MC65 sensor (Figure 6.3c) in the core, 3.3 cm below the surface and the yellow MC65 sensor, 0.5 cm below the surface of 80 apples of the cultivar "Golden Delicious". The apples had similar diameter (mean

= 6.94 cm, st.dev = 0.355) and firmness (mean = 7.49 and st.dev = 0.535). To simulate temperature fluctuations 8 fixed temperature conditions of 2 °C, 5 °C, 15 °C, 20 °C, 27 °C, 30 °C, 34 °C and 43 °C were used.

The external temperature sensor was measuring the fixed temperature during the experiment since all devices like fridges or furnaces fluctuate within an interval around the fixed value. The values of the readings are provided in Table 6.2. For the first fixed temperature condition 10 apples were placed in a commercial fridge (PITSOS P1KCL3606D) set at a fixed temperature of 2°C. For the 5°C condition 10 apples were placed in a commercial fridge with a fixed temperature of 5°C. For the 15°C condition 10 apples were placed inside an Elvem CLP 600 chamber (Figure 6.5b) fixed at 15°C. For the 20°C 10 apples were placed in a controlled temperature room fixed at 20 °C. For the 25°C 10 apples were placed in a controlled temperature room fixed at 25 °C.

For the rest of the temperature conditions, a WTC binder 78,532 furnace was used and set the temperature for 27°C, 34°C and 43°C.

To reach the desired temperature the apples were placed for 15 hours in the commercial fridges, chamber and the controlled environment room and 4 h in the furnace per desired temperature. Then, an apple was picked and pierced with a sharp needle of 33 mm length and 3 mm diameter, inserted the green sensor and fixed it with sticky tape, pierced the apple 0.5 cm below the surface, inserted the yellow sensor, fixed it with duct tape and placed the apple in one of the aforementioned places with the logger. After 15 minutes, the logger was switched on to make the temperature stable in case it increased/decreased during the insertion routine and 15 measurements were taken every 3 seconds. The procedure was repeated for 10 apples per fixed temperature. The experimental design of the laboratory experiment is summarized in Table 6.3.

6.4. Agreement Techniques

The study of agreement comprises of two important parts. The first part is to quantify the extent of agreement between two measurement methods/instruments/or procedures and to determine whether this is sufficient so that they can be used ‘interchangeably’. The second part is to compare important characteristics of the measurement methods such as bias and precision (Choudhary & Nagaraja, 2017a). The first goal of the study is the quantification of the difference between the core and the surface temperature. The second goal is the comparison of important characteristics such as bias and precision of the measurements since the bias reveals the trend of the difference along various ranges of temperature between the core and surface temperature of the fruit. The precision might reveal possible variations of the temperature readings probably due to difference in temperature fluctuations in the two points of interest.

Table 6.2. The means and standard deviations of the reference temperatures measured by the external temperature sensor per fixed condition.

	2 °C	5 °C	15 °C	20 °C	25 °C	27 °C	34 °C	43 °C
Mean	3.663	5.897	15.439	18.8	26.948	33.931	31.414	43.289
Standard Deviation	1.88	0.954	0.049	0.237	0.376	0.554	0.554	1.156

The approach followed in this case-study is based on measurement error models that predict the true latent trait using an empirical Bayes approach (Taffé, 2018) as described in Section 2.1.4. An R package is available for the implementation of the method (Taffé et al., 2019). A graphical illustration of the estimated bias between measurement approaches and the precision of each is produced as well. Since the package focuses on recalibration of the two methods the source code of the package is modified to remove the recalibration step, which is not relevant in our case. The adapted code is presented in detail in GitHub (Bataka, 2023) following model (2.25), (2.26) and equation (2.27).

The temperature of the thermistor placed in the core of the fruit is considered the reference. Thus, y_{2ij} is the j^{th} repeated measurement of the i^{th} subject for the thermistor placed in the core. Then y_{1ij} is the measurement for the thermistor placed on the surface of the fruit, where $i=1,...,80$ the number of subjects and $j=1,...,15$ the number of repeated measurements. The data are considered linked since they were paired.

This configuration was chosen since the goal of the field experiment targets the difference between the core and the surface temperature of the fruit. The measurement-error model was chosen since the goal of this case-study was to quantify the differential and proportional bias compared to mixed-effects models which assume that the proportional bias is the same for both methods.

6.5. Design of the Field Experiment

The experiment took place in apple orchards, in the mountainous village of Drakeia, south Pelion, at 500m altitude, Volos, Greece. The climatic conditions of the area during the summer months exceed 30°C and during the winter can drop below zero especially during nighttime. The sensors were installed in apple varieties Granny Smith and Red Delicious (Starking) after the harvest period. First, a tree was randomly selected and installed the first FruiTtemp logger on the peak of a tree's crown, the second to the east side of the tree, and the third on a low full shaded point (Figure 6.5a). The selection of fruits/subjects varied according to their daily sun exposure. The apple located on the peak of the tree was exposed to the sun continuously during the day, the apple located to the south-east side of the tree was exposed to the sun for less time than the one on the peak and the last one was full shaded during the whole day. One sensor (MC65) was placed inside the core of the apple just before the side of stem and the other (MC65) 1.5 – 2 mm away from the surface of the fruit. The third sensor (JS8746A) was also placed on the tree next to the fruit to measure the temperature of the environment (Figure 6.3d.). The system logging interval was adjusted to 15 min. After 24 h the next apple was measured for a total of 15 fruits. After the 24-h procedure each fruit was evaluated for its qualitative characteristics.

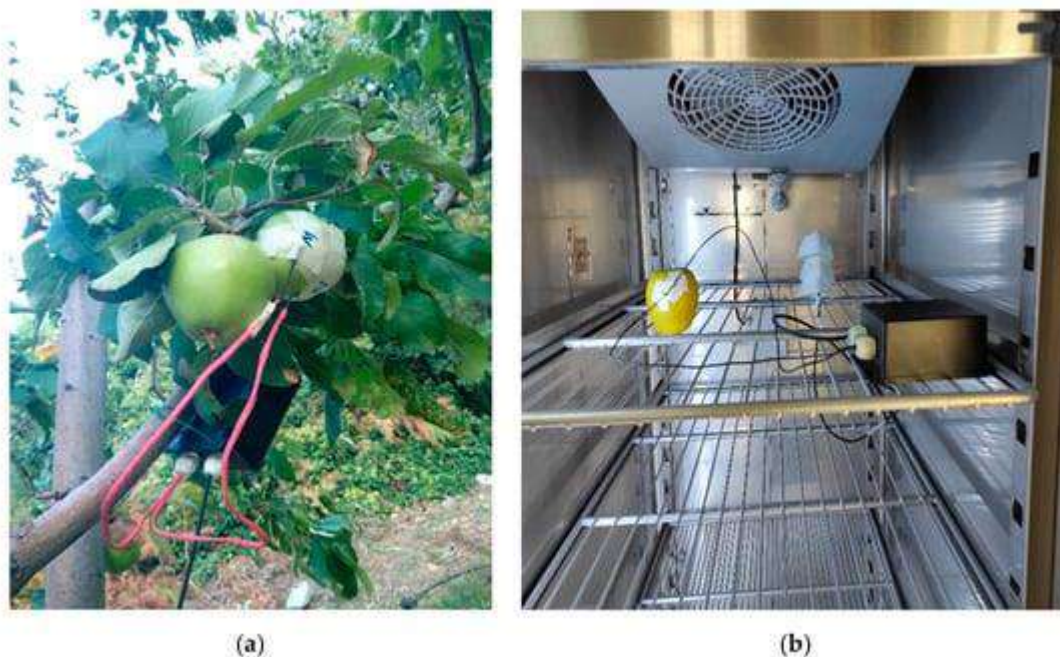


Figure 6.5. (a) The proposed device set during the field experiment. (b) The proposed device set in the Elvem CLP 600 chamber during the laboratory experiment.

Table 6.3. Summary of the experimental design for the laboratory experiment.

Experimental Design	
Experimental Unit	Fruit
Repeated Measurements	Yes, simultaneously
Number of repeated measurements	15
Data category	Linked data
Sample size	80
Balanced/Unbalanced measurements	Balanced

6.6. Results

Table 6.2 displays the temperature for every fixed condition (fixed temperature inside the fridges, room and furnaces described in Section 6.3) during the laboratory experiment, measured by the external temperature sensor, including the mean and standard deviations. The external sensor was placed 20 cm away from the fruit in all cases.

According to the scatterplot in Figure 6.6 the pairs of the measurements do not seem to deviate more than 0.5 units from the identity line. This is interpreted as minor differences among the core and surface temperatures especially for temperatures ranging from 9°C to 19°C. In Figure 6.7 the Bland–Altman and limits of agreement (LOA) plot is displayed, which was generated using the “MethodCompare” package in R version 4.0.2 (The R Foundation for Statistical Computing, Vienna, Austria). The limits of agreement in the plot (blue dashed lines and a solid line for the regression) indicate that there is a positive bias of the measurement method for low temperature values and a negative bias for high values. The linear trend of the Bland–Altman plot indicates possible correlation between differences and averages, and difference in

precision of the methods which is confirmed in the precision plot (Figure 6.8). There were no outliers detected.

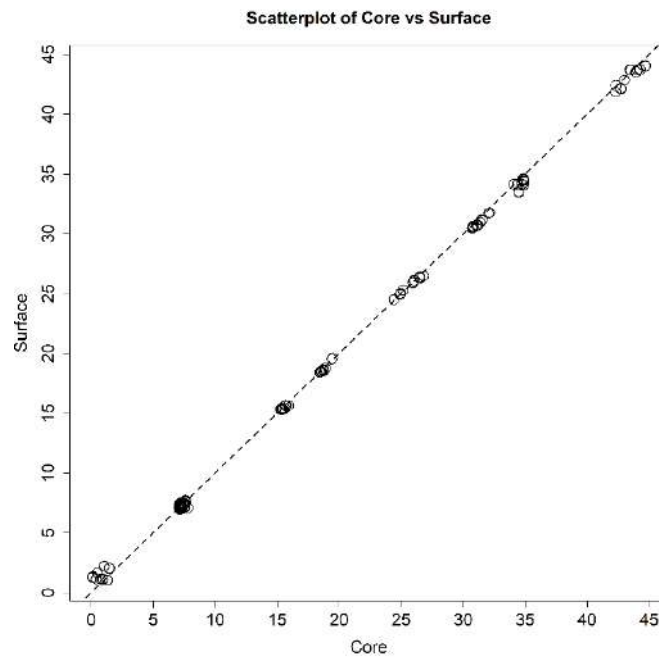


Figure 6.6. Scatterplot of measurement pairs including the identity line.

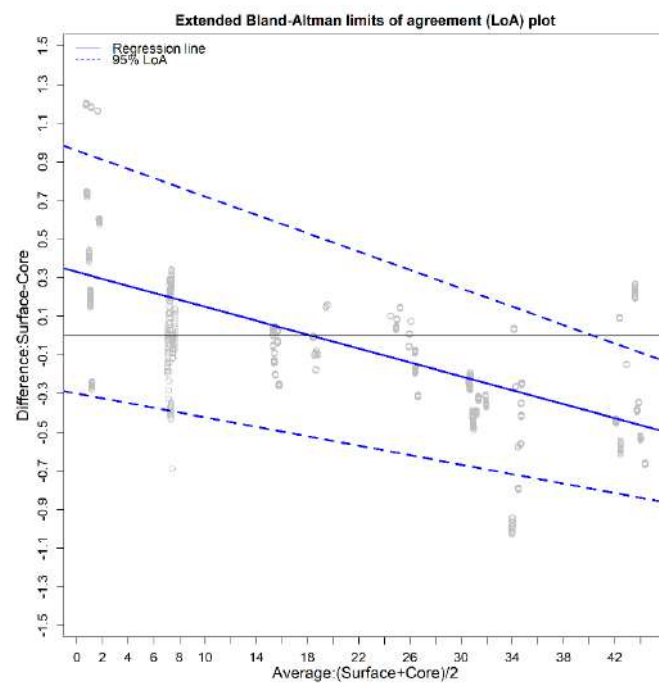


Figure 6.7. Bland–Altman LOA plot. The blue dashed lines are the limits of agreement, and the solid blue line is the regression line.

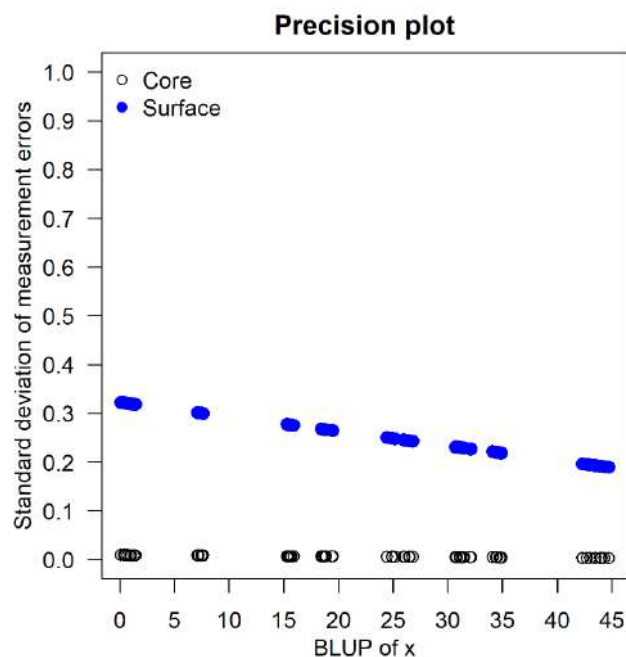


Figure 6.8. Precision plot showing the precision of each measurement method. The blue dots correspond to the precision of the surface readings along the temperature and the black dots to the precision of the core.

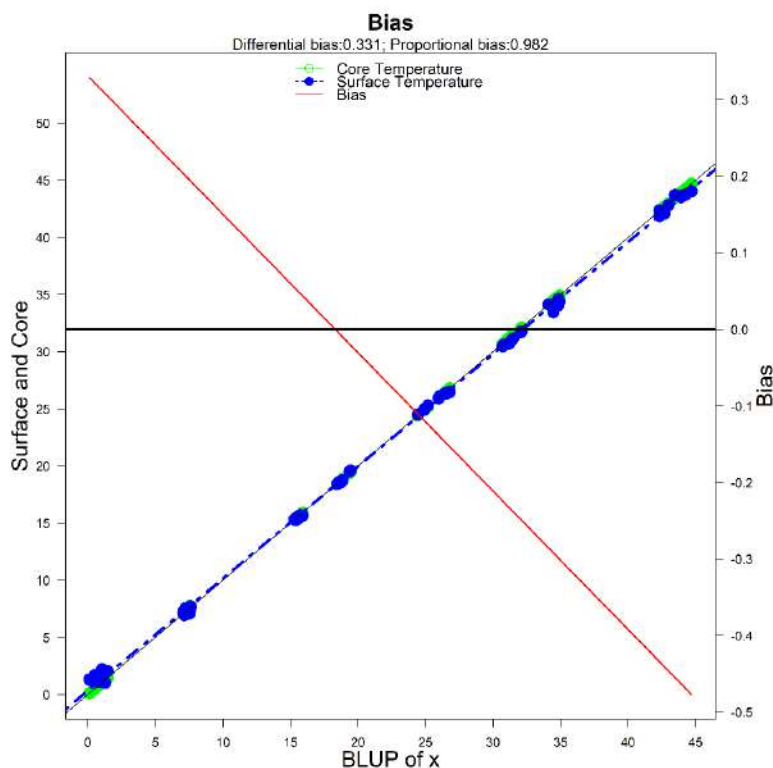


Figure 6.9. The Bias Plot. The bias is displayed as the red line and reveals the bias throughout the range of measurements. For the value of 30°C the bias is approximately -0.2°C . The black horizontal line is the zero reference for bias, the dashed blue and green line is regression of the Surface and Core temperature readings, respectively. The solid black line is the identity line.

After the analysis, the results revealed a differential bias of 0.331 (95% CI: [0.299,0.363]) and a proportional bias of 0.982 (95% CI: [0.981, 0.983]). The bias is positive for low values and negative for high values. Table 6.4 summarizes the differential and proportional bias for the measurement at the surface relative to the measurement at the core of the fruit and associated 95% confidence intervals. Bland–Altman LOA plot on the other hand estimates a differential bias of 0.958 (95% CI: [0.658, 1.28]) and a proportional bias of 0.976 (95% CI: [0.958, 0.989]) revealing the inadequacy of the specific method since it only assesses the variability of the differences but not each method separately.

The bias is displayed in Figure 6.9 and decreases as the true values increase until it reaches 0 at around 19°C then increases at the opposite direction. Specifically, almost a zero bias was found around 19°C, -0.2°C at 30°C and almost -0.4°C at 40°C. According to our results for a temperature of 40°C in the core, the surface sensor will measure 39.612°C ($0.331 + 0.982 * 40$), which gives a total bias of around -0.4°C compared to the core. The precision of each sensor is displayed in Figure 6.8 indicating that the core measurements are more precise than the surface temperatures revealing fluctuations of maximum 0.1°C on absolute value when the surface temperature is measured and a constant difference of precision between the core and the surface of 0.2°C with the core being more precise. This indicates that the core retains the temperature for more time than the surface much more in lower temperatures rather than in higher.

6.7. Improving Fruitemp – Future Work

The development board used for Fruitemp is an Adafruit feather proto 32u4. Feather series (Fried, 2023d) includes other microcontrollers with embedded wireless modules which have the same footprint for every solution. This feature makes Fruitemp a modular system and allows the user to add different development boards of the feather series, to include other functionalities such as wireless communication like LoRa, GSM, 4G, WiFi, Bluetooth LE etc. No additional wiring or soldering is needed. Moreover, FeatherWings are shields placed on top of any Feather development board extending the system according to the individual's needs. For example, DS3231 Precision RTC FeatherWing offers various wake up modes. When the MCU is in sleep mode the DS3231 can send a signal and wake the MCU at a prespecified interval. This is not possible using PCF8523 as discussed in Section 6.2.4. To add wireless functionality the Adafruit Feather M0 with RFM95 LoRa Radio can be added instead of Feather proto 32u4. For the LoRa gateway (Figure 6.10), another Feather M0 with RFM95 LoRa Radio is used connected with an Adafruit Feather Fona (Fried, 2023e) via I2C (Section 2.3.4). The LoRa development board acts as an intermediate between the end-nodes and the Feather Fona which is equipped with a GSM module and sends data to ThingSpeak (Section 2.3.8) via GPRS. Figure 6.11 displays ThingSpeak's interface. The 3 visualizations are for the core, surface and air temperature. Since the core and surface differences are small a change in the scale of the graph can make the discrepancies more obvious. Red dots represent the measurements. The gateway is powered using a 12 Volt lead acid battery. The battery is charged using a 30 Watt solar panel. The new design is under development.

Table 6.4. Differential and Proportional bias for the measurement at the surface relative to the measurement at the core of the fruit and associated 95% confidence intervals.

Bias	Estimate	2.5%	97.5%
Differential	0.331	0.299	0.363
Proportional	0.982	0.981	0.983

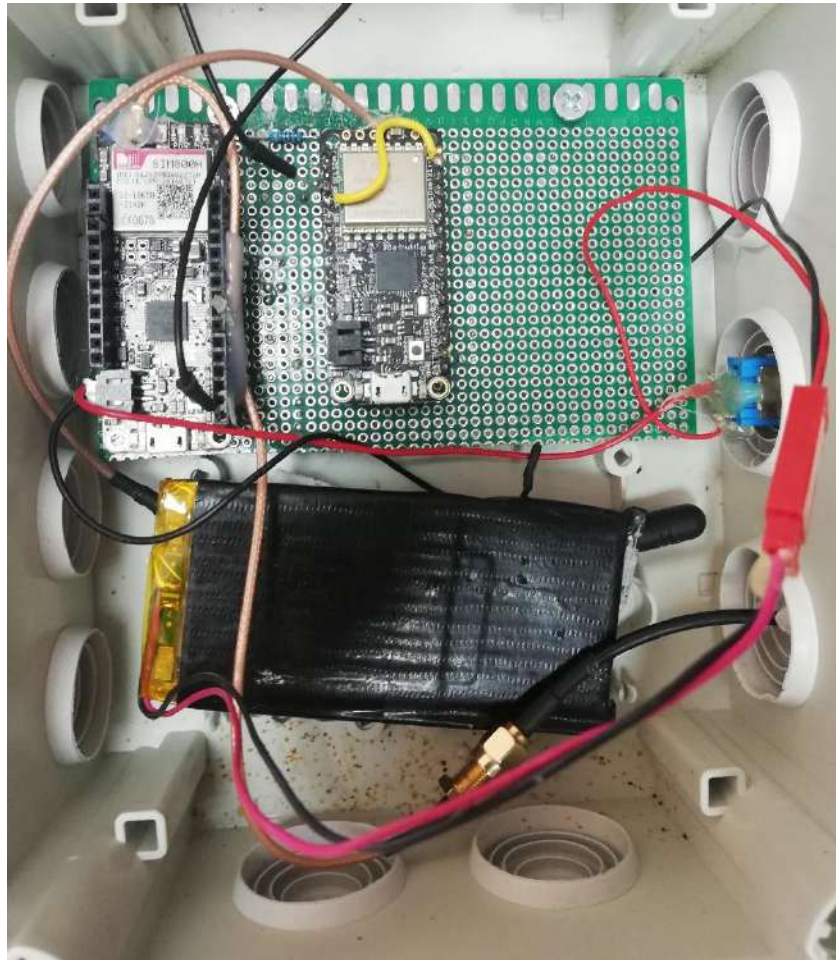


Figure 6.10. LoRa gateway consists of an Adafruit feather M0 with RFM95 LoRa module connected via I2C with an Adafruit Feather Fona which sends the data to ThingSpeak via GPRS.

FruiTTemp

Channel ID: 1667263
Author: mwa00002443449
Access: Private

Private View Public View Channel Settings Sharing API Keys Data Import / Export

Add Visualizations Add Widgets Export recent data

MATLAB Analysis

MATLAB Visualization

Channel 4 of 4

Channel Stats

Created: [about a year ago](#)
Last entry: [39 minutes ago](#)
Entries: 3612



Figure 6.11. ThingSpeak interface. The core, surface and air temperature are visualized using linegraphs. Red dots represent the measurements.

6.8. Discussion

The FruiTTemp prototype was built to aid the needs of gathering data for the overwintering dynamics of medfly in marginal for its existence areas in the frameworks of the Horizon 2020 funded project FF-IPM. The system promotes the open-source community and inexpensive methods that can be adopted in other studies in entomology and agriculture. This project can be a guide for non-specialized researchers to build systems that can aid them on a wide range of similar experiments.

The proposed device has low cost, fine sensors, portability, energy efficiency and accuracy as assessed in laboratory conditions. A field experiment was further designed and implemented successfully. The proposed device combines all the characteristics that can be found in widely used commercial devices such as the Hobo Mx2303 (Hobo, 2019) and the Hengko which are much more expensive and more intrusive to the fruit samples.

Agreement methods utilized were straightforward to implement and graphical tools make it easier for an individual with basic statistical background to interpret the results.

Other upgrades and optimizations can be applied to increase the reliability, precision, and accuracy of the device. The method described by Liu (2018) can be used to retrieve the coefficients for the Hoge-2 equation, to improve the calibration part. Since the thermistor curves provided by the manufacturer delivers a typical scenario of the sensors, an oil bath can be used to retrieve the Hoge-2 coefficients. In that case, calibration can be applied for each individual thermistor.

The FruitTemp prototype can be easily adapted to serve additional experimental needs in entomology and in ecology. Understanding insect phenology (seasonal occurrence), population growth and dynamics, as well as patterns of aging in the wild is still a major challenge in both applied ecology and entomology. Precise temperature recording in micro-habitats, such as a host fruit for medfly larvae, is of paramount importance for gaining insights regarding field biology of pikilotherms. FruitTemp provides an inexpensive open-source system that can be used in a wide range of studies besides overwintering. The data generated from these kinds of experiments will feed population modeling with reliable data that increase the precision of the projection and assist the management activities often implemented against insect pests.

Future research involves a detailed assessment of temperature fluctuations in the core of the fruit given standard temperature external conditions and the study of different hosts. Last, more experiments will be conducted to assess the wireless version of FruitTemp using LoRa and GSM wireless technology.

7. General Discussion

This thesis consists of two main topics. First, the design and construction of open-source devices to assess agricultural and industrial experiments. Second, the development of method comparison studies to assess the agreement between open-source devices and their corresponding industrials.

The method comparison study provides exploratory and inferential analysis. Depending on the experiment, indices are produced to assess the agreement and the similarity between the two devices. Four experiments were conducted and applied the methodology described in Chapter 2.

First, an open-source wireless weather station was designed and developed to assist experiments conducted in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. The advantages for such devices include two parts. First, the inexpensive design and second the configurability. The device was equipped with only one sensor that measures air temperature and humidity. However, the great number of tutorials and documentations allows users to integrate the device with multiple sensors and wireless technologies. Since the data points (repeated measurements) were 1,438 (every 5 minutes for 7 days with some data loss) a summary method was applied based on the National Observatory of Athens, the measurements were averaged using 8 daily time zones. For the analysis three different methodologies were applied to evaluate the concordance correlation coefficient, based on mixed-effects models, for longitudinal data. First, Choudhary and Nagaraja (2017a) proposed a method to evaluate CCC but their approach did not produce a valid model. Second, Carrasco's (2009) approach produced a single index for CCC and its estimate is 0.3346 [0.2642,0.4916]. Oliviera's approach estimated a Concordance Correlation Coefficient based on longitudinal data called Longitudinal Concordance Coefficient (LCC) which is the product of two components, the longitudinal Pearson correlation (LPC) and the Longitudinal Accuracy (LA). LPC evaluates the deviation of the measurements from the line of best fit while LA evaluates the deviation from the identity line. LCC estimate ranged from 0.1146 to 0.2236 since the index is produced for every time occasion of the data. LPC estimates were close to LCC and LA estimates ranged from 0.9286 to 0.9969. Overall, the analysis revealed low agreement between the two devices. A possible source of disagreement is the precision of the devices. Moreover, several factors such as humidity, location of the stations and cluster effects could have affected the agreement. Last, since the averaging in time-zones leads to loss of possible vital information, time series theory should be further researched and implemented to method comparison studies.

Second, a water content logger system was designed and developed to assist experiments conducted in the greenhouse facilities of the Laboratory of Agricultural Constructions and Environmental Control, in the area of Velestino (latitude 39°44', longitude 22°79', altitude 85 m), Greece. The system consists of a wireless LoRa portable logger equipped with RS-485 communication protocol and a water content, temperature and electroconductivity sensor for soil and soilless substrates. The substrate for the current experiment was perlite. The data were unlinked, and the sample size was balanced unlike the repeated measurements which were unbalanced. A preliminary experiment revealed proportional bias significantly different from 1 thus, Taffé's approach was followed since there was no assumption for proportional bias unlike the mixed-effects models. The analysis reported 11.669 units of water content (%) fixed bias and 0.46 proportional bias. Taffé also provided precision assessment and a recalibration method. Precision analysis reported lower precision for the open-source device which is almost 6 times less precise than the corresponding industrial. After recalibrating the open-source device the agreement was improved for

most of the values, but not the lower and the higher values in the measurement range. The agreement was adequate, assuming the acceptable difference is 5%, for the measurement range of [15,35]% of water content. Future work includes experiments for more substrates and measurement ranges.

Third, a pH logger was designed and constructed to assist experiments for acidity evaluation in citrus fruits. The device is almost 5 times less expensive than its corresponding industrial. Apart from pH, a temperature sensor was added for temperature compensation since pH readings are affected by temperature. The open-source device was calibrated using pH buffers for four different temperature ranges. A method comparison study was designed between the open-source system and a Hanna HI9024 Waterproof pH meter. For the agreement and similarity assessment the data were fit to mixed-effects models and Choudhary's and Nagaraja's approach was used to produce two indices (CCC and TDI) and an interval (limits of agreement) for assessing agreement and repeatability. For similarity assessment, fixed bias and precision ratio were estimated along with their 95% confidence intervals. Since the errors were heteroscedastic, the indices were estimated throughout the measurement range. Agreement is considered mediocre since CCC estimate is approximately 0.597 to 0.6032 with a lower bound of approximately 0.4776 to 0.4839 throughout all the measurement range. TDI(0.9) estimate is 0.2883 to 0.3031 with an upper bound of 0.3095 to 0.3232 throughout all the measurement range. After recalibration the agreement was excellent since CCC estimate is approximately 0.9667 to 0.9509 with a lower bound of approximately 0.9407 to 0.9194. For TDI(0.9) the estimate is 0.0857 to 0.1048 with an upper bound of 0.1052 to 0.1215. CCC and TDI(0.9) estimated by Carrasco et al. (2009) was 0.9781 and Escaramis et al. (2010a) 0.0670. The fixed bias is -0.2215 pH units and the main source of disagreement. The precision ratio reported that the open-source device is twice to three times more precise than Hanna. Repeatability is considered excellent since CCC estimate is close to 1 for both methods and TDI(0.9) is approximately 0.0243 to 0.0777 with an upper bound estimate of 0.0286 to 0.0949 for Hanna pH meter and 0.0181 to 0.0462 with an upper bound estimate of 0.0213 to 0.0563 for the open-source device throughout the measurement range. Overall, after recalibration the agreement is excellent, and the two devices can be used interchangeably.

Fourth, an open-source logger integrated with precise and slightly intrusive point medical temperature sensors was developed to conduct a series of experiments regarding the FF-IPM project. The logger's gain is based on its portability, low cost and the sensors' design. Specifically, the fine medical sensors are inserted in fruit's core and surface with the lowest possible hole diameter compared to other commercial and industrial sensors. The cost is almost half compared to other devices and can be reduced even more, by designing it with only parts essential for its functionality. Moreover, wireless technology can be easily implemented to the open-source device, reducing substantially the labor cost for data gathering. Apart from the device, a method comparison study was designed, and statistical methodologies were adopted to evaluate the agreement between temperature measurements taken from the core and the surface of the fruit. For the agreement analysis Taffé's approach was followed since the needs of this research regarded evaluation of the fixed and proportional bias of the surface compared to the core temperature. The results reported a fixed bias of 0.331°C and proportional bias 0.982. The FruitTemp prototype could be easily adapted to serve additional experimental needs in entomology and in ecology. Understanding insect phenology (seasonal occurrence), population growth and dynamics, as well as patterns of aging in the wild is still a major challenge in both applied ecology and entomology. Precise temperature recording in micro-habitats, such as a host fruit for medfly larvae, is of paramount importance for gaining insights regarding field biology of pikilotherms.

This thesis focuses on the promotion of the open-source philosophy for assisting research and industry in the agricultural field. Apart from the proposal of four devices applied in different agricultural fields, a formal method was investigated to assess the agreement between open-source and commercial or industrial devices that are proven to work efficiently. This work goal is to encourage the development of a formal protocol to validate open-source device using proper experimental design and statistical tools. The four case-studies can be used as a format and adopted to other future experiments. Moreover, the methodology discussed in this thesis can be extended and improved, staying faithful to the open-source philosophy which is evolving during the past decade. Business models and licenses for open-source devices are discussed in Appendix D1. This framework sets the ground for shared knowledge and grants the opportunity to any person, regardless of her/his educational level or financial status, to build their own business or community.

8. Conclusions

This thesis highlights the need to develop a protocol that assesses open-source devices, regarding their functionality and the validity of their measurements. The most effective way to validate the measurements of a novel device is to compare them to established commercial/industrial ones. The official and reliable process to accomplish this task is to design and apply a method comparison study which includes proper experimental design and statistical tools to assess the agreement and similarity between the two devices. This methodology is applied mostly in medical research (Bland & Altman, 1999; Khan et al. 2020) but not limited to it. This thesis adapts the statistical tools to assess agreement of medical devices/methods, to open-source devices developed for agricultural applications. The four case-studies can be used as a detailed guide to design and perform method comparison studies for open-source devices in the specific field.

Further, all of the established and widely used methodologies for assessing agreement, graphical and index-based, are presented and compared depending on the nature of the method comparison study in four case-studies unlike other research (Essack et al., 2009; Davis et al., 2007; Khan et al. 2021; Lunney et al., 2018; Igic et al., 2010; Perez-Jaume & Carrasco, 2015) where only one or a few of them are presented. Moreover, compared to already published papers (Essack et al., 2009; Laerhoven et al., 2022), the current research proposes a step-by-step guide to validate open-source devices, including the experimental design, descriptive statistics and a variety of formal statistical assessment and encourages the development of a protocol applied to this highly blooming field. Bongers et al. (2018) describes parts of the proposed guide but only reports Bland-Altman plot and ICC.

Last, novel devices for commercial and laboratory use are presented, and validated using the approach proposed by this thesis. The procedure to develop and evaluate them using the concept of agreement is available online (Bataka, 2023) thus, respecting the spirit of the open-source philosophy.

References

- Adafruit (2023). *Adafruit*. Retrieved April 22, 2023, from Adafruit: <https://adafruit.com/>
- Adafruit IO. (2023). *Adafruit Internet of Things Platform*. Retrieved April 22, 2023, from Adafruit IO: <https://io.adafruit.com/>
- Aliagas, C., Pérez-Foguet, A., Meseguer, R., Millán, P., & Molina, C. (2022). A Low-Cost and Do-It-Yourself Device for Pumping Monitoring in Deep Aquifers. *Electronics*, 11(22), 3788. <https://doi.org/10.3390/electronics11223788>
- Altman, D. G. (1993). Construction of age-related reference centiles using absolute residuals. *Statistics in Medicine*, 12(10), 917–924. <https://doi.org/10.1002/sim.4780121003>
- Altman, D. G., & Bland, J. (1983). Measurement in Medicine : the Analysis of Method Comparison Studies. *The Statistician*, pp. 307 - 317.
- Amphenol. (2014). *NTC Interchangeable Type 65 Series*. Retrieved April, 20, 2020, from <https://gr.mouser.com/datasheet/2/18/AAS-920-306C-NTC-Type-65-Series-031314-web-1315871.pdf>
- Amphenol. (2019). *JS8746 Harsh Environment Temperature Sensor*. Retrieved April, 20, 2020, from https://gr.mouser.com/datasheet/2/18/Amphenol_07152019_AAS-920-747B-Thermometrics_JS874-1622552.pdf
- Amphenol. (2021). *Amphenol-Sensors*. Retrieved April, 20, 2020, from <https://www.amphenol-sensors.com/hubfs/Documents/AAS-930-272B-Thermometrics-ColdChain-TempSensors%20021621-web.pdf>
- Analytika. (2023). *Analytika*. Retrieved May, 10, 2023, from https://www.analytika.gr/product-categories-46075/metrites-edafikis-ygrasias-soil-moisture-meters-46131/3680sm12-irrigation-station-2-sensor_43104/
- Andrade, R. O., & Yoo, S. G. (2019). A Comprehensive Study of the Use of LoRa in the Development of Smart Cities. *Applied Sciences*, 9(22), 4753. <https://doi.org/10.3390/app9224753>
- Arduino. (2015). *Arduino - Introduction*. Retrieved May 6, 2023, from <http://arduino.cc/en/guide/introduction>
- Arduino. (2023a). *Arduino Hardware*. Retrieved May 6, 2023, from <https://www.arduino.cc/en/hardware>
- Arduino. (2023b). *Arduino Software*. Retrieved May 6, 2023, from <https://www.arduino.cc/en/software>
- Arduino. (2023c). *Arduino Education*. Retrieved May 6, 2023, from <https://www.arduino.cc/education>
- Arduino. (2023d). *Arduino Pro*. Retrieved May 6, 2023, from <https://www.arduino.cc/pro/>
- Arduino. (2023e). *Arduino Docs*. Retrieved January 20, 2021, from <https://docs.arduino.cc/hacking/hardware/PinMapping32u4>
- Arduino. (2023f). *Arduino Cloud*. Retrieved August 14, 2023, from <https://cloud.arduino.cc/>

Arduino. (2023g). *Arduino MKR NB 1500*. Retrieve January 20, 2023, from <https://docs.arduino.cc/hardware/mkr-nb-1500>

Astrua, M., Ichim, D., Penneccchi, F., & Pisani, M. (2007). Statistical techniques for assessing the agreement between two instruments. *Metrologia*, 44(385). doi:10.1088/0026-1394/44/5/015

Atkinson, G., & Nevill, A. (1997). Comment on the Use of Concordance Correlation to Assess the Agreement between Two Variables. *Biometrics*, 53(2), 775–777. <http://www.jstor.org/stable/2533978>

Atkinson, R. K., Sabo, K., & Conley, Q. (2012). The participatory web. In *Handbook of Technology in Psychology, Psychiatry and Neurology: Theory, Research, and Practice* (pp. 91–120). Nova Science Publishers, Inc. <https://doi.org/10.1080/1369118x.2012.665935>

Atmel. (2016a). *Atmel ATmega16U4, ATmega32U4 Datasheet Summary*. Retrieved, May, 4, 2023, from <https://datasheet.chipsmall.com/ATMEGA32U4-AUR/586138.pdf>

Atmel. (2016b). *ATSAMW25-MR210PB Datasheet*. Retrieved May, 12, 2023, from https://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-42618-SmartConnect-ATSAMW25-MR210PB_Datasheet.pdf

AXIOMET. (n.d.). *AXIOMET AX-178*. Retrieved May, 20, 2023, from <https://www.tme.eu/Document/93bf63070827429f8d0f447adbc603bd/INSTR-AX-178%20-%20EN.pdf>

Arya111111. (2022). *DFRobot_PH library*. https://Github.Com/DFRobot/DFRobot_PH/Graphs/Contributors.

Babun, L., Denney, K., Celik, Z. B., McDaniel, P., & Uluagac, A. S. (2021). A survey on IoT platforms: Communication, security, and privacy perspectives. *Computer Networks*, 192, 108040. <https://doi.org/https://doi.org/10.1016/j.comnet.2021.108040>

Baccelli, F., Blaszczyszyn, B., & Muhlethaler, P. (2006). An Aloha Protocol for Multihop. *Transactions On Information Theory*, 52(2), 421-436. doi:10.1109/TIT.2005.862098

Banerji, S., & Chowdhury, R. (2013). On IEEE 802.11: Wireless Lan Technology. *International Journal of Mobile Network Communications & Telematics*, 3. <https://doi.org/10.5121/ijmnc.2013.3405>

Barnard, H. R., Findley, M. C., & Csavina, J. (2014). PARduino: a simple and inexpensive device for logging photosynthetically active radiation. *Tree Physiology*, 34(6), p. <https://doi.org/10.1093/treephys/tpu044>. doi:<https://doi.org/10.1093/treephys/tpu044>

Barnhart, H. X., & Williamson, J. M. (2001). Modeling Concordance Correlation via GEE to Evaluate Reproducibility. *Biometrics*, 57(3), 931–940. <https://doi.org/https://doi.org/10.1111/j.0006-341X.2001.00931.x>

Barnhart, H. X., Lokhnygina, Y., Kosinski, A. S., & Haber, M. (2007). Comparison of concordance correlation coefficient and coefficient of individual agreement in assessing agreement. *Journal of biopharmaceutical statistics*, 17(4), 721–738. <https://doi.org/10.1080/10543400701329497>

Barnhart, H. X., Song, J., & Haber, M. J. (2005). Assessing intra, inter and total agreement with replicated readings. *Statistics in Medicine*, 24(9), 1371–1384. <https://doi.org/10.1002/sim.2006>

Bataka, E. P., Miliokas, G., Katsoulas, N., & Nakas, C. T. (2021a). A Method Comparison Study between Open Source and Industrial Weather Stations. *The 13th EFITA International Conference*.
<https://doi.org/10.3390/engproc2021009008>

Bataka, E. P., Rodovitis, V. G., Zarpas, K. D., Papadopoulos, N. T., & Nakas, C. T. (2021b). FruiTTemp: Design, Implementation and Analysis for an Open-Source Temperature Logger Applied to Fruit Fly Host Experimentation. *Applied Sciences*, 11(13), 6003. <https://doi.org/10.3390/app11136003>

Bataka, E.P. (2023). *Statistical methods and experimental design for agreement and similarity studies between open-source software and hardware devices and their corresponding commercials for applications in agriculture*. Retrieved September 1, 2023, from
<https://github.com/kersee112358?tab=repositories>

Benjamin, B., Billah, M., Afreh-Nuamah, K., Obeng-Ofori, D., & Nyarko, G. (2015). Review of the pest status, economic impact and management of fruit-infesting flies (Diptera: Tephritidae) in Africa. *African Journal of Agricultural Research*, 10, 1488–1498. <https://doi.org/10.5897/AJAR2014.9278>

Benjamin, G., Tietel, Z., & Porat, R. (2013). Effects of rootstock/scion combinations on the flavor of citrus fruit. *Journal of agricultural and food chemistry*, 61(47), 11286–11294.
<https://doi.org/10.1021/jf402892p>

Bensky, A. (2019). Radio system design. In A. Bensky, *Short-range Wireless Communication*. (pp. 163-198). Elsevier.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet (London, England)*, 1(8476), 307–310.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, 8(2), 135–160. <https://doi.org/10.1177/096228029900800204>

Bland, J. M., & Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of biopharmaceutical statistics*, 17(4), 571–582.
<https://doi.org/10.1080/10543400701329422>

Bongers, C. C. W. G., Daanen, H. A. M., Bogerd, C. P., Hopman, M. T. E., & Eijsvogels, T. M. H. (2018). Validity, Reliability, and Inertia of Four Different Temperature Capsule Systems. *Medicine and Science in Sports and Exercise*, 50(1), 169–175. <https://doi.org/10.1249/MSS.0000000000001403>

Bouten, W., Baaij, E. W., Shamoun-Baranes, J., & Camphuysen, K. C. J. (2013). A flexible GPS tracking system for studying bird behaviour at multiple scales. *Journal of Ornithology*, 154(2), 571–580.
<https://doi.org/10.1007/s10336-012-0908-1>

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-29854-2>

Buonaccorsi, J. P. (2010). *Measurement Error*. Chapman and Hall/CRC.
<https://doi.org/10.1201/9781420066586>

- Burés, S., Carmen Gago, M., Morales, O., Marfà, O., & Xavier Martinez, F. (1997a). WATER CHARACTERIZATION IN GRANULAR MATERIALS (REFEREED). *Acta Horticulturae*, 450, 389–396. <https://doi.org/10.17660/ActaHortic.1997.450.48>
- Burés, S., Marfà, O., Pérez, T., Tébar, J., & Lloret, A. (1997b). MEASURE OF SUBSTRATES UNSATURATED HYDRAULIC CONDUCTIVITY (REFEREED). *Acta Horticulturae*, 450, 297–304. <https://doi.org/10.17660/ActaHortic.1997.450.35>
- Byra-Cook, C. J., Dracup, K. A., & Lazik, A. J. (1990). Direct and indirect blood pressure in critical care patients. *Nursing research*, 39(5), 285–288.
- Caldwell, A. R. (2022). SimplyAgree: An R package and jamovi Module for Simplifying Agreement and Reliability Analyses. *Journal of Open Source Software*, 7(71), 4148. <https://doi.org/10.21105/joss.04148>
- Carkeet A. (2015). Exact parametric confidence intervals for Bland-Altman limits of agreement. *Optometry and vision science : official publication of the American Academy of Optometry*, 92(3), e71–e80. <https://doi.org/10.1097/OPX.0000000000000513>
- Carrasco, J. L., King, T. S., & Chinchilli, V. M. (2009). The concordance correlation coefficient for repeated measures estimated by variance components. *Journal of biopharmaceutical statistics*, 19(1), 90–105. <https://doi.org/10.1080/10543400802527890>
- Carrasco, J. L., Phillips, B. R., Puig-Martinez, J., King, T. S., & Chinchilli, V. M. (2013). Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Computer methods and programs in biomedicine*, 109(3), 293–304. <https://doi.org/10.1016/j.cmpb.2012.09.002>
- Carrasco, J.L. and Jover, L. (2003), Estimating the Generalized Concordance Correlation Coefficient through Variance Components. *Biometrics*, 59: 849-858. <https://doi.org/10.1111/j.0006-341X.2003.00099.x>
- Carrasco, JL, & Martinez JP (2022). _cccrm: Cooncordance Correlation Coefficient for Repeated (and Non-Repeated) Measures_. R package version 2.1.0, <<https://CRAN.R-project.org/package=cccrm>>.
- Carroll, R. J., & Ruppert, D. (1988). *Transformation and Weighting in Regression*. Springer US. <https://doi.org/10.1007/978-1-4899-2873-3>
- Carstensen, B., Simpson, J., & Gurrin, L. C. (2008). Statistical models for assessing agreement in method comparison studies with replicate measurements. *The international journal of biostatistics*, 4(1), . <https://doi.org/10.2202/1557-4679.1107>
- Casals, L., Mir, B., Vidal, R., & Gomez, C. (2017). Modeling the Energy Performance of LoRaWAN. *Sensors*, 17(10), 2364. <https://doi.org/10.3390/s17102364>
- Chen, C.C., & Barnhart, H.X. (2008). Comparison of ICC and CCC for assessing agreement for data without and with replications. *Computational Statistics and Data Analysis*, 53(2), 554–564. <https://doi.org/10.1016/j.csda.2008.09.026>
- Chen, C., Zhang, X., Xu, K., Su, X., & Sun, J. (2021). Compensation Methods for pH Direct Measurement in Soilless Culture Substrates Using the All-Solid-Stated pH Sensor. *IEEE Sensors Journal*, 21(23), 26856–26867. <https://doi.org/10.1109/JSEN.2021.3123658>

Chen, L., Chapman, J. L., Yee, B. J., Wong, K. K. H., Grunstein, R. R., Marshall, N. S., & Miller, C. B. (2018). Agreement between electronic and paper Epworth Sleepiness Scale responses in obstructive sleep apnoea: Secondary analysis of a randomised controlled trial undertaken in a specialised tertiary care clinic. *BMJ Open*, 8(3). <https://doi.org/10.1136/bmjopen-2017-019255>

Cheng, C., Van Ness, J.W. (1999). *Statistical Regression with Measurement Error*. John Wiley, Chichester, UK.

Chinchilli, V. M., Martel, J. K., Kumanyika, S., Lloyd, T., Lloyd, T. (1996). *A Weighted Concordance Correlation Coefficient for Repeated Measurement Designs Published by : International Biometric Society* Stable URL : <https://www.jstor.org/stable/2533172> International Biometric Society is collaborating with JSTOR to digitize , pres. 52(1), 341–353.

Choudhary, P. K. (2008). A tolerance interval approach for assessment of agreement in method comparison studies with repeated measurements. *Journal of Statistical Planning and Inference*, 138(4), 1102–1115. <https://doi.org/10.1016/j.jspi.2007.03.056>

Choudhary, P. K., & Nagaraja, H. N. (2007). Tests for assessment of agreement using probability criteria. *Journal of Statistical Planning and Inference*, 137(1), 279–290. <https://doi.org/10.1016/j.jspi.2005.11.009>

Choudhary, P. K., & Nagaraja, H. N. (2017a). *Measuring Agreement*. Wiley. <https://doi.org/10.1002/9781118553282>

Choudhary, P. K., & Nagaraja, H. N. (2017b). *Measuring Agreement: Models, Methods, and Applications*. Retrieved, May, 14, 2023, from https://personal.utdallas.edu/~pankaj/agreement_book/

Choudhary, P. K., & Tony Ng, H. K. (2006). Assessment of agreement under nonstandard conditions using regression models for mean and variance. *Biometrics*, 62(1), 288–296. <https://doi.org/10.1111/j.1541-0420.2005.00422.x>

Circuit Python. (2023). *Circuit Python*. Retrieved, May, 21, 2023, from <https://circuitpython.org/>: <https://circuitpython.org/>.

Circuit Tree (2016). *Circuit Tree*. Retrieved, May, 29, 2023, from <https://circuit-tree.com/>.

Coleman, G., Salter, W., & Walsh, M. (2022). OpenWeedLocator (OWL): an open-source, low-cost device for fallow weed detection. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-021-03858-9>

ComWinTop. (2023). *ComWinTop*. Retrieved, May, 22, 2023, from <https://store.comwintop.com/>

Cornbleet, P. J., & Gochman, N. (1979). Incorrect least-squares regression coefficients in method-comparison analysis. *Clinical chemistry*, 25(3), 432–438.

Costa, D., & Duran-Faundez, C. (2018). Open-Source Electronics Platforms as Enabling Technologies for Smart Cities: Recent Developments and Perspectives. *Electronics*, 7(12), 404. <https://doi.org/10.3390/electronics7120404>

Cox, N. J. (2006). Assessing agreement of measurements and predictions in geomorphology. *Geomorphology*, 76(3–4), 332–346. <https://doi.org/10.1016/j.geomorph.2005.12.001>

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley & Sons, Inc.
<https://doi.org/10.1002/9781119115151>

Datta, D. (2018). blandr: Bland-Altman Method Comparison. 0.5.1. Retrieved, 4, 2023, from
<https://cran.r-project.org/web/packages/blandr/index.html>

Davis, J. A., Rozenek, R., DeCicco, D. M., Carizzi, M. T., & Pham, P. H. (2007). Comparison of three methods for detection of the lactate threshold. *Clinical Physiology and Functional Imaging*, 27(6), 381–384. <https://doi.org/10.1111/j.1475-097X.2007.00762.x>

Dawoud, D. S., & Dawoud, P. (2020). Serial Communication Protocols and Standards RS232/485, UART/USART, SPI, USB, IN-STEON, Wi-Fi and WiMAX. In *Serial Communication Protocols and Standards RS232/485, UART/USART, SPI, USB, INSTEON, Wi-Fi and WiMAX* (pp. i–xl).

Vet (de), H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59(10), 1033–1039.
<https://doi.org/10.1016/j.jclinepi.2005.10.015>

DFRobot. (2020). *SEN0262 datasheet*. Retrieved April, 2, 2023, from DFRobot:
<https://www.farnell.com/datasheets/3161991.pdf>

DFRobot. (2022). *SEN0161-V2*. Retrieved April, 2, 2023 from
https://wiki.dfrobot.com/Gravity__Analog_pH_Sensor_Meter_Kit_V2_SKU_SEN0161-V2

Dhraief, M. Z., Bedhiaf, S., Dhehibi, B., Oueslati-Zlaoui, M., Jebali, O., & Ben-Youssef, S. (2019). FACTORS AFFECTING INNOVATIVE TECHNOLOGIES ADOPTION BY LIVESTOCK HOLDERS IN ARID AREA OF TUNISIA. *New Medit*, 18(4). <https://doi.org/10.30682/nm1904a>

Drăgulinescu, A. M. C., Manea, A. F., Fratu, O., & Drăgulinescu, A. (2020). LoRa-Based Medical IoT System Architecture and Testbed. *Wireless Personal Communications*, 126(1), 25–47.
<https://doi.org/10.1007/s11277-020-07235-z>

Dunn, G. (2004). *Statistical Evaluaton of Measurement Errors: Design and Analysis of Reliability Studies*. Oxford University Press.

Earp, J., & Finlayson, D. (1991). Relationship between urinary bladder and pulmonary artery temperatures: a preliminary study. *Heart Lung*, 20(3), pp. 265-270.

Ebrahimi-Darkhaneh, H. (2019). *Measurement error caused by self-heating in NTC and PTC thermistors*. 2019. Retrieved April, 20, 2020, from
<https://www.ti.com/lit/an/slyt774/slyt774.pdf?ts=1624707191874>

Electro Boy. (2022). *hackster.io*. Retrieved May 5, 2023, from
<https://www.hackster.io/electroboy001/nrf24l01-testing-circuit-using-arduino-b6707e#toc-pcb-for-nrf-shield-4>

Embedded, E. (n.d.). *AVR Architecture*. Retrieved May 20, 2023, from Explore Embedeed:
http://www.exploreembedded.com/wiki/1._AVR_Architecture

- Escaramís, G., Ascaso, C., & Carrasco, J. L. (2010a). The total deviation index estimated by tolerance intervals to evaluate the concordance of measurement devices. *BMC Medical Research Methodology*, 10. <https://doi.org/10.1186/1471-2288-10-31>
- Escaramís, G., Ascaso, C., & Carrasco, J. L. (2010b). *BMC Medical Research Methodology*. Retrieved April 20, 2023, from https://static-content.springer.com/esm/art%3A10.1186%2F1471-2288-10-31/MediaObjects/12874_2009_434_MOESM1_ESM.PDF
- ESPRESSIF. (2023). *ESP32 Series*. Retrieved May 25, 2023, from https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf
- Essack, Y., Hoffman, M., Rensburg, M., van Wyk, J., Meyer, C., & Erasmus, R. (2009). A comparison of five glucometers in South Africa. *Journal of Endocrinology, Metabolism and Diabetes of South Africa*, 14(2), 102–105. <https://doi.org/10.1080/22201009.2009.10872202>
- Etienne, A., Génard, M., Lobit, P., Mbéguié-A-Mbéguié, D., & Bugaud, C. (2013). What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. *Journal of Experimental Botany*, 64(6), 1451–1469. <https://doi.org/10.1093/jxb/ert035>
- evaherrada. (2023). *Github/adafruit/RTCLib*. Retrieved June, 5, 2023 from <https://github.com/adafruit/RTCLib>
- Evans, J. (2021). *KiCad Documentation*. Retrieved April 22, 2023, from <https://docs.kicad.org/7.0/en/introduction/introduction.html>
- Fahad, E. (2023). *ChatGPT Arduino Programming Ultimate Test, ChatGPT for Arduino Projects*. Retrieved August 14, 2023, from <https://www.electronicclinic.com/chatgpt-arduino-programming-ultimate-test-chatgpt-for-arduino-projects/>
- FDA. (1999). *Accuracy*. Retrieved May, 10, 2023 from <https://www.govinfo.gov/content/pkg/FR-1999-06-08/html/99-14401.htm>.
- Ferrández-Pastor, F. J., Alcañiz-Lucas, S., García-Chamizo, J. M., & Platero-Horcajadas, M. (2019). Smart Environments Design on Industrial Automated Greenhouses. *13th International Conference on Ubiquitous Computing and Ambient Intelligence UCAmI 2019*, 36. <https://doi.org/10.3390/proceedings2019031036>
- Ferrarezi, R. S., Lin, X., Gonzalez Neira, A. C., Tabay Zambon, F., Hu, H., Wang, X., Huang, J.-H., & Fan, G. (2022). Substrate pH Influences the Nutrient Absorption and Rhizosphere Microbiome of Huanglongbing-Affected Grapefruit Plants. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.856937>
- Ferreira, M. V., Moraes, A. C., Rendo-Urteaga, T., Bel-Serrat, S., Torres-Leal, F. L., Moreno, L. A., & Carvalho, H. B. (2018). Ordinary least products regression is a simple and powerful statistical tool to identify systematic disagreement between two measures: fixed and proportional bias assessment. *2018 Stata Conference*, 34.
- FF-IPM. (n.d.). FF-IPM. Retrieved April 20, 2021, from <https://fruitflies-ipm.eu/>

- Fimiani, P. (1989). Mediterranean Region. In A. H. Robinson, *Fruit Flies: Their Biology, Natural Enemies, and Control*. (pp. 39-55). Amsterdam.: Elsevier.
- Fisher, D. K., Fletcher, R. S., Anapalli, S. S., & III, H. C. P. (2018). Development of an Open-Source Cloud-Connected Sensor-Monitoring Platform. *Advances in Internet of Things*, 08(01), 1–11. <https://doi.org/10.4236/ait.2018.81001>
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied Longitudinal Analysis*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119513469>
- Fleiss, J. L. (1999). *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118032923>
- Frey, M. E., Petersen, H. C., & Gerke, O. (2020). Nonparametric Limits of Agreement for Small to Moderate Sample Sizes: A Simulation Study. *Stats*, 3(3), 343–355. <https://doi.org/10.3390/stats3030022>
- Fried, L. (2022a). *Adafruit WINC1500 WiFi Shield for Arduino*. Retrieved May 23, 2023, from <https://cdn-learn.adafruit.com/downloads/pdf/adafruit-winc1500-wifi-shield-for-arduino.pdf>
- Fried, L. (2022b). *Adafruit Stemma & Stemma QT*. Retrieved May 29, 2023, from <https://cdn-learn.adafruit.com/downloads/pdf/introducing-adafruit-stemma-qt.pdf>
- Fried, L. (2023a). *Adafruit Feather 32u4*. Retrieved May 29, 2023, from <https://learn.adafruit.com/adafruit-feather-32u4-fona/overview>
- Fried, L. (2023b). *Adafruit HUZZAH32 - ESP32 Feather*. Retrieved May 23, 2023, from <https://cdn-learn.adafruit.com/downloads/pdf/adafruit-huzzah32-esp32-feather.pdf>
- Fried, L. (2023c). *Adafruit Adalogger FeatherWing*. Retrieved April 22, 2023, from <https://cdn-learn.adafruit.com/downloads/pdf/adafruit-adalogger-featherwing.pdf>
- Fried, L. (2023d). *Adafruit Learn*. Retrieved May 30, 2023, from <https://cdn-learn.adafruit.com/downloads/pdf/adafruit-feather-32u4-fona.pdf>
- Fried, L. (2023e). *Adafruit Feather 32u4 Fona*. Retrieved May 29, 2023, from <https://www.adafruit.com/product/2771>
- Fried, L. (2023f). *Adafruit Feather m0 LoRa*. Retrieved May 30, 2023, from [cdn-learn.adafruit.com: https://cdn-learn.adafruit.com/downloads/pdf/adafruit-feather-m0-radio-with-lora-radio-module.pdf](https://cdn-learn.adafruit.com/downloads/pdf/adafruit-feather-m0-radio-with-lora-radio-module.pdf)
- Fruugo. (2023). Stable Uart Serial Port To Rs485 Converter Function Module Rs485 To Ttl Module Hfmqv. Retrieved July 11, 2023, from https://www.fruugo.gr/stable-uart-serial-port-to-rs485-converter-function-module-rs485-to-ttl-module-hfmqv/p-123137211-258844746?language=en&ac=KelkooCSS&asc=pmax&gclid=EAlaIqobChMI9ZCM8K2GgAMVqI6DBx1E7AjCEAYYAIBEGKnud_BwE
- Gartner Peer Insights. (2023). *Managed IOT Connectivity Services Reviews and Ratings*. Retrieved April 10, 2023, from <https://www.gartner.com/reviews/market/managed-iot-connectivity-services>
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>

Gerke, O. (2020). Nonparametric Limits of Agreement in Method Comparison Studies: A Simulation Study on Extreme Quantile Estimation. *International Journal of Environmental Research and Public Health*, 17(22), 8330. <https://doi.org/10.3390/ijerph17228330>

Gershenfeld, N. A. (2005). *Fab: The Coming Revolution on Your Desktop--from Personal Computers to Personal Fabrication*. Basic Books.

Giovino, B. (2019). *How to Implement LoRa Firmware Over the Air (FOTA) with Minimal Power Consumption*. Retrieved April 10, 2023, from digikey: <https://www.digikey.gr/en/articles/how-to-implement-lora-fota-minimal-power-consumption>

Giovino, B. (2021). *How to Efficiently Connect to Sensors in IoT Endpoints Using 1-Wire Communication*. Retrieved from DigiKey: <https://www.digikey.gr/en/articles/how-to-efficiently-connect-to-sensors-in-iot-endpoints-using-1-wire-communication>

Github. (2018). *M0 Lorawan TTN*. Retrieved May 20, 2023, from <https://github.com/werktag/m0-lorawan-ttn>

Github. (2022). *DFRobot pH sensor*. Retrieved May, 12, 2023, from https://github.com/DFRobot/DFRobot_PH

GmbH, E., & Temmen., N. (2021). *ECOS*. Retrieved April 10, 2023, from <https://www.ecos.eu/en/>: https://www.ecos.eu/files/content/downloads/publikationen/REPORT_Smart_Farming.pdf

Google Open Source (2023). *Github: Open Source Casebook*. Retrieved May 31, 2023, from <https://google.github.io/opencasebook/patents/#notes>

Goumopoulos, C. (2018). A High Precision, Wireless Temperature Measurement System for Pervasive Computing Applications. *Sensors*, 18(10), 3445. <https://doi.org/10.3390/s18103445>

Grubbs, F. E. (1948). On Estimating Precision of Measuring Instruments and Product Variability. *Journal of the American Statistical Association*, 43(242), 243. <https://doi.org/10.2307/2280371>

Gunawardena, N., Pardyjak, E. R., Stoll, R., & Khadka, A. (2018). Development and evaluation of an open-source, low-cost distributed sensor network for environmental monitoring applications. *Measurement Science and Technology*, 29(2), 024008. <https://doi.org/10.1088/1361-6501/aa97fb>

Gutierrez, A. P., & Ponti, L. (2011). Assessing the invasive potential of the Mediterranean fruit fly in California and Italy. *Biological Invasions*, 13(12), 2661–2676. <https://doi.org/10.1007/s10530-011-9937-6>

Gutierrez, A. P., Ponti, L., & Cossu, Q. A. (2009). Effects of climate warming on Olive and olive fly (*Bactrocera oleae* (Gmelin)) in California and Italy. *Climatic Change*, 95(1–2), 195–217. <https://doi.org/10.1007/s10584-008-9528-4>

Hahn, G. J. (1970). Statistical Intervals for a Normal Population Part I. Tables, Examples and Applications. *Journal of Quality Technology*, 2(3), 115–125. <https://doi.org/10.1080/00224065.1970.11980426>

Hahn, G. J., & Meeker, W. Q. (1991). *Statistical Intervals*. Wiley. <https://doi.org/10.1002/9780470316771>

- Hall Ellis Solicitors. (2023). *Hal Ellis Solicitors*. Retrieved March 8, 2023, from <https://hallellis.co.uk/software-solicitors/open-source-lawyers/>
- Hama. (n.d.). *Hama*. Retrieved May 20, 2023 from <https://www.hama.com/00004469/hama-star-black-tripod-153-3d>
- Hamilton, C., & Stamey, J. (2007). Using Bland–Altman to Assess Agreement between Two Medical Devices – Don’t Forget the Confidence Intervals! *Journal of Clinical Monitoring and Computing*, 21(6), 331–333. <https://doi.org/10.1007/s10877-007-9092-x>
- Hammond Manufacturing. (2022). *ABS Plastic Multi-Purpose Enclosures w/ PC Board Standoffs*. Retrieved May 29, 2023, from <https://gr.mouser.com/datasheet/2/177/1591xx-1224884.pdf>
- Hanna Instruments. (1996). *Transcat*. Retrieved May 10, 2023, from <https://www.transcat.com/media/pdf/HI9024.pdf>
- Hanneman, S. K. (2008). Design, Analysis, and Interpretation of Method-Comparison Studies. *AACN Advanced Critical Care*, 19(2), 223–234. <https://doi.org/10.1097/01.AACN.0000318125.41512.a3>
- Hengko. (n.d.). Handheld Temperature and Relative Humidity Data Recorder Sensor in Fruit and Vegetable Warehouses Glasshouses. Retrieved April 2, 2021, from <https://www.hengko.com/control-the-atmosphere-relative-humidity-and-temperature-sensor-in-fruit-and-vegetable-warehouses-glasshouses-products/>
- Heradio, R., Chacon, J., Vargas, H., Galan, D., Saenz, J., de La Torre, L., & Dormido, S. (2018). Open-Source Hardware in Education: A Systematic Mapping Study. *IEEE Access*, 6, 72094–72103. <https://doi.org/10.1109/ACCESS.2018.2881929>
- Herrada, E. (2023). Adafruit ADS1X15 library. Retrieved May 29, 2023, from https://github.com/adafruit/Adafruit_ADS1X15
- Hindman, B. (2014). *Analog Devices - What is an iButton device?* Retrieved May 23, 2023, from <https://www.analog.com/media/en/technical-documentation/tech-articles/what-is-an-ibutton-device.pdf>
- Hobo. (2019). *bmc*. Retrieved January 10, 2021, from <https://www.bmc.de/mediafiles/datenblaetter/HOBO%20MX2303%20Data%20Logger%20Datasheet.pdf>
- Hoef, J. V. (2012). Who Invented the Delta Method? *The American Statistician*, 66(2), pp. 124-127. doi:<https://www.jstor.org/stable/23339471>
- Hoge, H. J. (1988). Useful procedure in least squares, and tests of some equations for thermistors. *Review of Scientific Instruments*, 59(6), 975–979. <https://doi.org/10.1063/1.1139762>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, 50(3), 346–363. <https://doi.org/10.1002/bimj.200810425>
- Huang, X., Lu, Q., Wang, L., Cui, M., & Yang, F. (2020). Does aging and off-farm employment hinder farmers’ adoption behavior of soil and water conservation technology in the Loess Plateau?

International Journal of Climate Change Strategies and Management, 12(1), 92–107.

<https://doi.org/10.1108/IJCCSM-04-2019-0021>

Humfrey, N., & per1234. (2022). *Arduino Library List*. Retrieved March 10, 2023, from

<https://www.arduinolibraries.info/>

Huth, H. (2023). *Modest Maker*. Retrieved April 10, 2023, from <https://modestmaker.com/>

Hutson, A. D., & Yu, H. (2021). A robust permutation test for the concordance correlation coefficient.

Pharmaceutical Statistics, 20(4), 696–709. <https://doi.org/10.1002/pst.2101>

IBM, Kooijman, M., Moore, T., Won, C., & Rose, F. (2015). IBM LMIC framework. Retrieved March 20,

2023, from <https://www.arduinolibraries.info/libraries/ibm-lmic-framework>

Ibrahim, M., Elgamri, A., Babiker, S., & Mohamed, A. (2015). Internet of things based smart environmental monitoring using the Raspberry-Pi computer. *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, 159–164.

<https://doi.org/10.1109/ICDIPC.2015.7323023>

Igic, B., Hauber, M. E., Galbraith, J. A., Grim, T., Dearborn, D. C., Brennan, P. L. R., Moskát, C., Choudhary, P. K., & Cassey, P. (2010). Comparison of micrometer- and scanning electron microscope-based measurements of avian eggshell thickness. *Journal of Field Ornithology*, 81(4), 402–410.

<https://doi.org/10.1111/j.1557-9263.2010.00296.x>

InstaDeep. (2019). *DeepPCB*. Retrieved October 14, 2023, from <https://deeppcb.ai/>

Instructables. (n.d.). *Autodesk Instructables*. Retrieved October 10, 2019, from

<https://www.instructables.com/>

ISO. (2023). *Bias*. Retrieved May 10, 2023, from <https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:ed-2:v1:en>.

ISO. (2017). *ISO Online Browsing Platform*. Retrieved March 10, 2023, from

<https://www.iso.org/obp/ui/#iso:std:iso:21748:ed-2:v1:en>

Jacobsen, B. S., Tulman, L., & Lowery, B. J. (1991). Three sides of the same coin: the analysis of paired data from dyads. *Nursing Research*, 40(6), 359–363.

Jakub, N. (2017). *Instructables*. Retrieved May 14, 2020, from <https://www.instructables.com/DS18B20-Radiation-Shield/>

Jan, S.-L., & Shieh, G. (2018). The Bland-Altman range of agreement: Exact interval procedure and sample size determination. *Computers in Biology and Medicine*, 100, 247–252.

<https://doi.org/10.1016/j.combiomed.2018.06.020>

Janse, R. J., Hoekstra, T., Jager, K. J., Zoccali, C., Tripepi, G., Dekker, F. W., & van Diepen, M. (2021).

Conducting correlation analysis: important limitations and pitfalls. *Clinical Kidney Journal*, 14(11), 2332–2337. <https://doi.org/10.1093/ckj/sfab085>

Jensen, T. (2010). Soil pH and the Availability of Plant Nutrients. *IPNI Plant Nutrition TODAY*, 2.

- Jensen, A. L., & Kjelgaard-Hansen, M. (2006). Method comparison in the clinical laboratory. *Veterinary Clinical Pathology*, 35(3), 276–286. <https://doi.org/10.1111/j.1939-165X.2006.tb00131.x>
- Ji, M., Yoon, J., Choo, J., Jang, M., & Smith, A. (2019). LoRa-based Visual Monitoring Scheme for Agriculture IoT. *2019 IEEE Sensors Applications Symposium (SAS)*, 1–6. <https://doi.org/10.1109/SAS.2019.8706100>
- Jolles, J. W. (2021). Broad-scale applications of the Raspberry Pi: A review and guide for biologists. *Methods in Ecology and Evolution*, 12(9), 1562–1579. <https://doi.org/10.1111/2041-210X.13652>
- Kafkas. (n.d.). *Kafkas*. Retrieved May 20, 2021, from https://static.kafkas.gr/Uploads/Pdfs/197385/2007077_656007077_01_Z05.PDF
- Khan, S., Saultry, B., Adams, S., Kouzani, A. Z., Decker, K., Digby, R., & Bucknall, T. (2021). Comparative accuracy testing of non-contact infrared thermometers and temporal artery thermometers in an adult hospital setting. *American Journal of Infection Control*, 49(5), 597–602. <https://doi.org/10.1016/j.ajic.2020.09.012>
- Khattab, A., Habib, S. E. D., Ismail, H., Zayan, S., Fahmy, Y., & Khairy, M. M. (2019). An IoT-based cognitive monitoring system for early plant disease forecast. *Computers and Electronics in Agriculture*, 166, 105028. <https://doi.org/10.1016/j.compag.2019.105028>
- Kim, J., Minagawa, D., Saito, D., Hoshina, S., & Kanda, K. (2022). Development of KOSEN Weather Station and Provision of Weather Information to Farmers. *Sensors*, 22(6), 2108. <https://doi.org/10.3390/s22062108>
- King, T. S., & Chinchilli, V. M. (2001). A generalized concordance correlation coefficient for continuous and categorical data. *Statistics in Medicine*, 20(14), 2131–2147. <https://doi.org/10.1002/sim.845>
- King, T. S., Chinchilli, V. M., & Carrasco, J. L. (2007a). A repeated measures concordance correlation coefficient. *Statistics in Medicine*, 26(16), 3095–3113. <https://doi.org/10.1002/sim.2778>
- King, T. S., Chinchilli, V. M., Wang, K., Carrasco, J. L., King, T. S., Chinchilli, V. M., Wang, K., Carrasco, J. L., King, T. S., Chinchilli, V. M., Wang, K., & Carrasco, J. L. (2007b). *A Class of Repeated Measures Concordance Correlation Coefficients*. 3406. <https://doi.org/10.1080/10543400701329455>
- Kirkeby, O. J., Rise, I. R., Nordsletten, L., Skjeldal, S., Hall, C., & Risoe, C. (1995). Cerebral blood flow measured with intracerebral laser-Dopplerflow probes and radioactive microspheres. *Journal of Applied Physiology*, 79(5), 1479–1486. <https://doi.org/10.1152/jappl.1995.79.5.1479>
- Knörig, A., Wettach, R., & Cohen, J. (2009). Fritzting. *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, 351–358. <https://doi.org/10.1145/1517664.1517735>
- Kota, V. D., Sharma, H., Albert, M. v., Mahbub, I., Mehta, G., & Namuduri, K. (2023). A Low-Power Wireless System for Predicting Early Signs of Sudden Cardiac Arrest Incorporating an Optimized CNN Model Implemented on NVIDIA Jetson. *Sensors*, 23(4), 2270. <https://doi.org/10.3390/s23042270>
- Krasniqi, F., Maraj, A. & Blaka, E. (2018). "Performance analysis of mobile 4G/LTE networks". *South-Eastern European Design Automation, Computer Engineering, Computer Networks and Society Media*

Conference (SEEDA_CECNSM), Kastoria, Greece, 2018, pp. 1-5, doi: 10.23919/SEEDA-CECNSM.2018.8544937.

Kuiper, R., Hoogenboezem, R., Huisman, S., Sonnevelf, P., & van Duin, M. (2022). *R-package nopacko*. Retrieved March 20, 2023, from <https://cran.r-project.org/web/packages/nopaco/nopaco.pdf>

Lado, J., Rodrigo, M., & Zacarías, L. (2014). Maturity indicators and citrus fruit quality. *Stewart Postharvest Review*.

Laerhoven, K. V., Hoelzemann, A., Pahmeier, I., Teti, A., & Gabrys, L. (2022). Validation of an open-source ambulatory assessment system in support of replicable activity studies. *German Journal of Exercise and Sport Research*, 52, pp. 262–272. doi: <https://doi.org/10.1007/s12662-022-00813-2>

Lai, D., & Shiao, S. (2005). Comparing two clinical measurements: A linear mixed model approach. *Journal of applied statistics*, 32(8), pp. 855–860. doi:10.1080/02664760500080157

Last Minute Engineers. (2023). Last Minute Engineers. Retrieved from <https://lastminuteengineers.com/arduino-micro-sd-card-module-tutorial/>

Lazic, S. E. (2016). *Experimental design for laboratory biologists: Maximizing information and improving reproducibility*. Cambridge University Press. <https://doi.org/10.1017/9781139696647>

Lee, K.-S., Lee, D.-S., Lim, C.-S., Lee, S.-P., Yang, J.-E., & Chung, D.-Y. (2022). Water Retention Characteristics of Various Sizes of Expanded Perlite Produced from Two Different Types of Rocks. *Horticulturae*, 8(9), 805. <https://doi.org/10.3390/horticulturae8090805>

Lehmann, E. L., & Casella, G. (1998). *Theory of Point Estimation*. Springer-Verlag. <https://doi.org/10.1007/b98854>

Li, X., Chen, Z., Wang, J., & Jin, J. (2023). LeafSpec-Dicot: An Accurate and Portable Hyperspectral Imaging Device for Dicot Leaves. *Sensors*, 23(7), 3687. <https://doi.org/10.3390/s23073687>

Liao, J. J. Z. (2003). An improved concordance correlation coefficient. *Pharmaceutical Statistics*, 2(4), 253–261. <https://doi.org/10.1002/pst.52>

Liehr, P., Dedo, Y. L., Torres, S., & Meininger, J. C. (1995). Assessing agreement between clinical measurement methods. *Heart & Lung*, 24(3), 240–245. [https://doi.org/10.1016/S0147-9563\(05\)80043-3](https://doi.org/10.1016/S0147-9563(05)80043-3)

Lin, L. I. K. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, 45(1), 255. <https://doi.org/10.2307/2532051>

Lin, L. I. K. (2000). Total deviation index for measuring individual agreement with applications in laboratory performance and bioequivalence. *Statistics in Medicine*, 19(2), 255–270. [https://doi.org/10.1002/\(SICI\)1097-0258\(20000130\)19:2<255::AID-SIM293>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1097-0258(20000130)19:2<255::AID-SIM293>3.0.CO;2-8)

Lin, L. I., & Chinchilli, M. V. (1997). Rejoinder to the letter to the editor from Atkinson and Nevill. *Biometrics*, 53, pp. 777–778.

Lin, L., Hedayat, A. S., & Wu, W. (2007). A Unified Approach for Assessing Agreement for Continuous and Categorical Data. *Journal of Biopharmaceutical Statistics*, 17(4), 629–652. <https://doi.org/10.1080/10543400701376498>

- Lin, L., Hedayat, A. S., & Wu, W. (2012). *Statistical Tools for Measuring Agreement*. Springer New York. <https://doi.org/10.1007/978-1-4614-0562-7>
- Lin, L., Hedayat, A. S., Sinha, B., & Yang, M. (2002). Statistical Methods in Assessing Agreement. *Journal of the American Statistical Association*, 97(457), 257–270. <https://doi.org/10.1198/016214502753479392>
- Lin, L., Pan, Y., Hedayat, A. S., Barnhart, H. X., & Haber, M. (2016). A simulation study of nonparametric total deviation index as a measure of agreement based on quantile regression. *Journal of Biopharmaceutical Statistics*, 26(5), 937–950. <https://doi.org/10.1080/10543406.2015.1094812>
- Linke, B. (2008). *Overview of 1-Wire Technology and Its Use*. Retrieved May 23, 2023, from <https://www.analog.com/media/en/technical-documentation/tech-articles/guide-to-1wire-communication--maxim-integrated.pdf>
- Linnet, K. (1993). Evaluation of regression procedures for methods comparison studies. *Clinical Chemistry*, 39(3), 424–432. <https://doi.org/10.1093/clinchem/39.3.424>
- Liquido, N., Shinoda, L., & Cunningham, R. (1991). Host Plants of the Mediterranean Fruit Fly (Diptera: Tephritidae): An Annotated World Review. *Miscellaneous publications of the Entomological Society of America*, p. 77.
- Liu, G., Guo, L., Liu, C., & Wu, Q. (2018). Evaluation of different calibration equations for NTC thermistor applied to high-precision temperature measurement. *Measurement*, 120, 21–27. <https://doi.org/10.1016/j.measurement.2018.02.007>
- LoRa Alliance. (2023a). *lora-developers.semtech.com*. Retrieved April 7, 2023, from <https://lora-developers.semtech.com/documentation/tech-papers-and-guides/lora-and-lorawan/>
- Lora Alliance. (2023b). *what is LoRaWAN specifications*. Retrieved April 7, 2023, from LoRa Alliance: <https://lora-alliance.org/about-lorawan/>
- LoRa Vsb Cz. (2023). Complete guide to connect fethar m0 to TTN. Retrieved April 7, 2023, from <https://lora.vsb.cz/index.php/adafruit-feather-m0/>
- Louis, L. (2016). Working Principle of Arduino and Using it as a Tool for Study and Research. *International Journal of Control, Automation, Communication and Systems*, 1(2), 21–29. <https://doi.org/10.5121/ijcacs.2016.1203>
- Lova Raju, K., & Vijayaraghavan, V. (2022). A Self-Powered, Real-Time, NRF24L01 IoT-Based Cloud-Enabled Service for Smart Agriculture Decision-Making System. *Wireless Personal Communications*, 124(1), 207–236. <https://doi.org/10.1007/s11277-021-09462-4>
- Lu, M.-J., Zhong, W.-H., Liu, Y.-X., Miao, H.-Z., Li, Y.-C., & Ji, M.-H. (2016). Sample Size for Assessing Agreement between Two Methods of Measurement by Bland–Altman Method. *The International Journal of Biostatistics*, 12(2). <https://doi.org/10.1515/ijb-2015-0039>
- Lucas, A., Hudson, G. J., Simpson, P., Cole, T. J., & Baker, B. A. (1987). An automated enzymic micromethod for the measurement of fat in human milk. *Journal of Dairy Research*, 54(4), 487–492. <https://doi.org/10.1017/S0022029900025693>

- Ludbrook, J. (1997). SPECIAL ARTICLE COMPARING METHODS OF MEASUREMENT. *Clinical and Experimental Pharmacology and Physiology*, 24(2), 193–203. <https://doi.org/10.1111/j.1440-1681.1997.tb01807.x>
- Luiz, R.R., Costa, A.J.L., Kale, P.L., Werneck, G.L. (2003). Assessment of agreement of a quantitative variable: a new graphical approach. *Journal of Clinical Epidemiology*, 56(10), 963–967. [https://doi.org/10.1016/S0895-4356\(03\)00164-1](https://doi.org/10.1016/S0895-4356(03)00164-1)
- Lunney, M., Tonelli, B., Lewis, R., Wiebe, N., Thomas, C., MacRae, J., & Tonelli, M. (2018). A comparison of temporal artery thermometers with internal blood monitors to measure body temperature during hemodialysis. *BMC Nephrology*, 19(1), 137. <https://doi.org/10.1186/s12882-018-0938-x>
- Mahender, K., Kumar, T. A., & Ramesh, K. S. (2018). *Analysis of multipath channel fading techniques in wireless communication systems*. 020050. <https://doi.org/10.1063/1.5032012>
- Mandel, J. (1978). Accuracy and Precision: Evaluation and interpretation of analytical results. In I. Kolthoff, & P. Elving, *Treatise on Analytical Chemistry: Theory and practise*. (2 ed., Vol. 1, pp. 243-298). New York: Wiley.
- Mandel, J., & Stiehler, R. D. (1954). Sensitivity--a criterion for the comparison of methods of test. *Journal of Research of the National Bureau of Standards*, 53(3), 155. <https://doi.org/10.6028/jres.053.018>
- MARKOSKA, V., SPALEVIC, V., & GULABOSKI, R. (2018). A RESEARCH ON THE INFLUENCE OF POROSITY ON PERLITE SUBSTRATE AND ITS INTERACTION ON POROSITY OF TWO TYPES OF SOIL AND PEAT SUBSTRATE. *The Journal "Agriculture and Forestry,"* 64(3). <https://doi.org/10.17707/AgricultForest.64.3.02>
- Matalon, N. (2016). *Arduino Libraries: Watchdog*. Retrieved September 17, 2023, from <https://github.com/nadavmatalon/WatchDog>
- Mathew, B., Gawad, S.G., & Mathai, T. (2017). Design of Wireless Notice Boards for Long Distance Data Transmission by using Bluetooth and nRF24l01.
- MathWorks Inc. (2022). *Statistics and machine learning toolbox documentation*. Retrieved September 18, 2023, from <https://www.mathworks.com/help/stats/index.html>
- MathWorks. (2023). *Things speak*. Retrieved May 10, 2023, from <https://thingspeak.com/>
- Maxim Integrated Products. (2014). *MAX481/MAX483/MAX485/ MAX487–MAX491/MAX1487*. Retrieved June 12, 2023, from <https://www.analog.com/media/en/technical-documentation/data-sheets/MAX1487-MAX491.pdf>
- Maxim Integrated Products. (2015). *Extremely Accurate I2C-Integrated RTC/TCXO/Crystal*. Retrieved May 15, 2020, from <https://www.analog.com/media/en/technical-documentation/data-sheets/ds3231.pdf>
- Maxim Integrated Products. (2019). *DS18B20*. Retrieved March 20, 2023, from <https://www.analog.com/media/en/technical-documentation/data-sheets/ds18b20.pdf>
- Mechatronics (2023). *How to Mechatronics*. Retrieved April 24, 2023, from <https://howtomechatronics.com/>

Mesas-Carrascosa, F. J., Verdú Santano, D., Meroño, J. E., Sánchez de la Orden, M., & García-Ferrer, A. (2015). Open source hardware to monitor environmental parameters in precision agriculture. *Biosystems Engineering*, 137, 73–83. <https://doi.org/10.1016/j.biosystemseng.2015.07.005>

Meteoshop. (2023). *Meteoshop*. Retrieved March 10, 2023, from <https://www.meteoshop.gr/product/569/meteorologikoi-stathmoi-davis-6263-wireless-vantage-pro2-plus/>

Meyer, M., Copeland, R., Wharton, R., McPheron, B., & Barnes, B. (2002). Proceedings of the 6th International Fruit Fly Symposium. *On the Geographic Origin of the Medfly Ceratitis Capitata (Wiedemann) (Diptera: Tephritidae)*. Stellenbosch: Isteg Scientific Publications.

Microchip, I. (2023, 4 24). *SAM D21/DA1 Family*. Retrieved May 10, 2023, from <https://ww1.microchip.com/downloads/aemDocuments/documents/MCU32/ProductDocuments/DataSheets/SAM-D21-DA1-Family-Data-Sheet-DS40001882H.pdf>

MicroPython (2013). MicroPython. Retrieved May 10, 2023, from <https://www.kickstarter.com/projects/214379695/micro-python-python-for-microcontrollers>

Mies, R., Bonvoisin, J., & Stark, R. (2020). DEVELOPMENT OF OPEN SOURCE HARDWARE IN ONLINE COMMUNITIES: INVESTIGATING REQUIREMENTS FOR GROUPWARE. *Proceedings of the Design Society: DESIGN Conference*, 1, 997–1006. <https://doi.org/10.1017/dsd.2020.38>

Minitab. (2023). *Support minitab*. Retrieved May 18, 2023, from <https://support.minitab.com:https://support.minitab.com/en-us/minitab/21/help-and-how-to/statistical-modeling/doe/supporting-topics/basics/replicates-and-repeats-in-designed-experiments/#:~:text=Repeat%20and%20replicate%20measurements%20are,experimental%20runs%2C%20which%20are%20oft>

Monogoto. (2023). *Monogoto*. Retrieved May 18, 2023, from https://monogoto.io/iot-m2m-sim-card/?utm_source=google&utm_medium=cpc&utm_campaign=brand&utm_term=monogoto%20ltd&utm_campaign=Brand&utm_source=adwords&utm_medium=ppc&hsa_acc=2593602194&hsa_cam=9635664446&hsa_grp=101081544602&hsa_ad=642785941917&hsa_src=g

monusuthar2016. (2017). *PCB Etching*. Retrieved May 18, 2023, from <https://www.instructables.com/PCB-Etching/>

National Instruments Corp. (2023). *NI Engineer Ambitiously*. Retrieved April 24, 2023, from <https://www.ni.com:https://www.ni.com/en-us/shop/data-acquisition/sensor-fundamentals/sensor-terminology.html>

National Observatory of Athens. (2023). *Meteo*. Retrieved May 20, 2023 from <https://www.meteo.gr/index.cfm>

- Nawarathna, L. S., & Choudhary, P. K. (2015). A heteroscedastic measurement error model for method comparison data with replicate measurements. *Statistics in Medicine*, 34(7), 1242–1258. <https://doi.org/10.1002/sim.6424>
- Ni, J., Zhang, J., Wu, R., Pang, F., & Zhu, Y. (2018). Development of an Apparatus for Crop-Growth Monitoring and Diagnosis. *Sensors*, 18(9), 3129. <https://doi.org/10.3390/s18093129>
- Niezen, G., Eslambolchilar, P., & Thimbleby, H. (2016). Open-source hardware for medical devices. *BMJ Innovations*, 2(2), 78–83. <https://doi.org/10.1136/bmjinnov-2015-000080>
- Nordic., S. (2008). *Nrf24L01+ Single Chip 2.4Ghz Transceiver*. Retrieved April 18, 2023, from https://www.sparkfun.com/datasheets/Components/SMD/nRF24L01Plus_Preliminary_Product_Specification_v1_0.pdf
- Oellermann, M., Jolles, J. W., Ortiz, D., Seabra, R., Wenzel, T., Wilson, H., & Tanner, R. L. (2022). Open Hardware in Science: The Benefits of Open Electronics. *Integrative and Comparative Biology*, 62(4), 1061–1075. <https://doi.org/10.1093/icb/icac043>
- Oliveira, T. de P., Hinde, J., & Zocchi, S. S. (2018). Longitudinal Concordance Correlation Function Based on Variance Components: An Application in Fruit Color Analysis. *Journal of Agricultural, Biological and Environmental Statistics*, 23(2), 233–254. <https://doi.org/10.1007/s13253-018-0321-1>
- Oliveira, T. P., Moral, R. A., Zocchi, S. S., Demetrio, C. G. B., & Hinde, J. (2020). lcc: an R package to estimate the concordance correlation, Pearson correlation and accuracy over time. *PeerJ*, 8, e9850. <https://doi.org/10.7717/peerj.9850>
- OpenAI. (2023). *ChatGPT*. Retrieved August 14, 2023, from <https://chat.openai.com/chat>
- Open-Source Community. (2023). *Open-source*. Retrieved February 2, 2023 from <https://opensource.com/resources/what-open-hardware>
- Open-Source Hardware Association. (2023). *Open Source Hardware Association*. Retrieved from [oshwa.org: https://www.oshwa.org/definition/](https://www.oshwa.org/definition/)
- Open-Source Initiative. (2023). *Open Source Initiative*. Retrieved February 2, 2023 from <https://opensource.org/>
- OSBSS. (2018). *Thermistor_v2*. Retrieved May 29, 2019, from Github: https://github.com/OSBSS/Thermistor_v2
- OSHW. (2023). *Open Source Hardware Association*. Retrieved February 2, 2023, from <https://www.oshwa.org/>: <https://www.oshwa.org/>
- PaintYourDragon. (2022). *Github/Adafruit/Adafruit_SHTC3*. Retrieved February 2, 2021, from Github: https://github.com/adafruit/Adafruit_SHTC3
- Papadopoulos, N. T., Carey, J. R., Katsoyannos, B. I., & Kouloussis, N. A. (1996). Overwintering of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Northern Greece. *Annals of the Entomological Society of America*, 89(4), 526–534. <https://doi.org/10.1093/aesa/89.4.526>

- Papadopoulos, N. T., Katsoyannos, B. I., Kouloussis, N. A., Economopoulos, A. P., & Carrey, J. R. (1998). Effect of adult age, food, and time of day on sexual calling incidence of wild and mass-reared *Ceratitis capitata* males. *Entomologia Experimentalis et Applicata*, 89(2), 175–182. <https://doi.org/10.1046/j.1570-7458.1998.00397.x>
- Papadopoulos, N. T., Plant, R. E., & Carey, J. R. (2013). From trickle to flood: the large-scale, cryptic invasion of California by tropical fruit flies. *Proceedings of the Royal Society B: Biological Sciences*, 280(1768), 20131466. <https://doi.org/10.1098/rspb.2013.1466>
- Papadopoulos, N., Katsoyannos, B., Carey, J., & Kouloussis, N. (2001). Seasonal and Annual Occurrence of the Mediterranean Fruit Fly (Diptera: Tephritidae) in Northern Greece. *Ann. Entomol. Soc. Am.*, 94(1), pp. 41-50. doi:[https://doi.org/10.1603/0013-8746\(2001\)094\[0041:SAAOOT\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0041:SAAOOT]2.0.CO;2)
- Papavaslopoulou, S., Giannakos, M. N., & Jaccheri, L. (2017). Empirical studies on the Maker Movement, a promising approach to learning: A literature review. *Entertainment Computing*, 18, 57–78. <https://doi.org/10.1016/j.entcom.2016.09.002>
- Passing, H., & Bablok, W. (1983). A New Biometrical Procedure for Testing the Equality of Measurements from Two Different Analytical Methods. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part I. *Cclm*, 21(11), 709–720. <https://doi.org/10.1515/cclm.1983.21.11.709>
- Patle, K. S., Saini, R., Kumar, A., & Palaparthi, V. S. (2022). Field Evaluation of Smart Sensor System for Plant Disease Prediction Using LSTM Network. *IEEE Sensors Journal*, 22(4), 3715–3725. <https://doi.org/10.1109/JSEN.2021.3139988>
- Pearce, J. M. (2015). Quantifying the Value of Open Source Hard-ware Development. *Modern Economy*, 06(01), 1–11. <https://doi.org/10.4236/me.2015.61001>
- Pedersen, S. M., & Lind, K. M. (2017). *Precision Agriculture – From Mapping to Site-Specific Application* (pp. 1–20). https://doi.org/10.1007/978-3-319-68715-5_1
- Pena, E., & Legaspi, M. G. (2020). *UART: A Hardware Communication Protocol Understanding Universal Asynchronous Receiver/Transmitter*. Retrieved March 25, 2023, from <https://www.analog.com/media/en/analog-dialogue/volume-54/number-4/uart-a-hardware-communication-protocol.pdf>
- Perez-Jaume, S., & Carrasco, J. L. (2015). A non-parametric approach to estimate the total deviation index for non-normal data. *Statistics in Medicine*, 34(25), 3318–3335. <https://doi.org/10.1002/sim.6544>
- Philips. (1999). *Sparkfun*. Retrieved April 15, 2023, from <https://www.sparkfun.com/datasheets/LCD/Monochrome/Nokia5110.pdf>
- Pierce, F. J., & Nowak, P. (1999). *Aspects of Precision Agriculture* (pp. 1–85). [https://doi.org/10.1016/S0065-2113\(08\)60513-1](https://doi.org/10.1016/S0065-2113(08)60513-1)
- Pinheiro J, Bates D, R Core Team (2022). *_nlme: Linear and Nonlinear Mixed Effects Models_*. R package version 3.1-162, <https://CRAN.R-project.org/package=nlme>.
- Pinheiro, J. C., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Ney Work.: Springer.

PONTI, L., COSSU, Q. A., & GUTIERREZ, A. P. (2009). Climate warming effects on the *Olea europaea* - *Bactrocera oleae* system in Mediterranean islands: Sardinia as an example. *Global Change Biology*, 15(12), 2874–2884. <https://doi.org/10.1111/j.1365-2486.2009.01938.x>

Potapov S, Model F, Schuetzenmeister A Schuetzenmeister A, Dufey F, Raymaekers J (2023). *_mcr: Method Comparison Regression_*. R package version 1.3.2 <<https://CRAN.R-project.org/package=mcr>>.

Quiroz, J., & Burdick, R. K. (2009). Assessment of Individual Agreements with Repeated Measurements Based on Generalized Confidence Intervals. *Journal of Biopharmaceutical Statistics*, 19(2), 345–359. <https://doi.org/10.1080/10543400802622576>

R Core, T. (2023). R: A language and environment for statistical. Vienna, Austria. Retrieved May 28, 2023, from <https://www.R-project.org/>

Rahnema, M. (1993). Overview of the GSM system and protocol architecture. *IEEE Communications Magazine*, 31(4), 92–100. <https://doi.org/10.1109/35.210402>

RAK. (2023a). *RAK2245 Pi HAT WisLink LPWAN Concentrator Datasheet*. Retrieved June 21, 2023, from <https://docs.rakwireless.com/Product-Categories/WisLink/RAK2245-Pi-HAT/Datasheet/#overview>: <https://docs.rakwireless.com/Product-Categories/WisLink/RAK2245-Pi-HAT/Datasheet/#overview>

RAK. (2023b). *RAK2245 Pi HAT Quick Start Guide*. Retrieved June 21, 2023, from <https://docs.rakwireless.com/api/render/?emulateScreenMedia=false&pdf.format=legal&url=https%3A%2F%2Fdocs.rakwireless.com%2FProduct-Categories%2FWisLink%2FRAK2245-Pi-HAT%2FQuickstart%2F>

Raspberry Pi Foundation (2012). *Get started with Raspberry pi*. Retrieved May 15, 2023, from <https://www.raspberrypi.org/help/>

Raspberry pi. (2023). *Raspbian*. Retrieved June 18, 2023, from <http://www.raspbian.org>: <http://www.raspbian.org>

Rathnayake, L. N., & Choudhary, P. K. (2017). Semiparametric modeling and analysis of longitudinal method comparison data. *Statistics in Medicine*, 36(13), 2003–2015. <https://doi.org/10.1002/sim.7261>

Red Hat. (2022). *What is an API?* Retrieved February 23, 2023, from Red Hat: <https://www.redhat.com/en/topics/api/what-are-application-programming-interfaces>

REKHA, C., POORNIMA, G., MANASA, M., ABHIPSA, V., DEVI, J. P., KUMAR, H. T. V., & KEKUDA, T. R. P. (2012). Ascorbic Acid, Total Phenol Content and Antioxidant Activity of Fresh Juices of Four Ripe and Unripe Citrus Fruits. *Chemical Science Transactions*, 1(2), 303–310. <https://doi.org/10.7598/cst2012.182>

Rencher, A. C., & Schaalje, G. B. (2007). *Linear Models in Statistics*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470192610>

Reynders, B., & Pollin, S. (2016). Chirp spread spectrum as a modulation technique for long range communication. *2016 Symposium on Communications and Vehicular Technologies (SCVT)*, 1–5. <https://doi.org/10.1109/SCVT.2016.7797659>

Rf Wireless World (2012). *1 Wire protocol basics*. Retrieved May 20, 2023, from <https://www.rfwireless-world.com/Tutorials/1-wire-protocol-basics.html>.

- ROTHERY, P. (1979). A nonparametric measure of intraclass correlation. *Biometrika*, 66(3), 629–639. <https://doi.org/10.1093/biomet/66.3.629>
- Rovný, P. (2016). The Analysis of Farm Population with Respect to Young Farmers in the European Union. *Procedia - Social and Behavioral Sciences*, 220, 391–398. <https://doi.org/10.1016/j.sbspro.2016.05.513>
- Rudtsch, S., & von Rohden, C. (2015). Calibration and self-validation of thermistors for high-precision temperature measurements. *Measurement*, 76, 1–6. <https://doi.org/10.1016/j.measurement.2015.07.028>
- Sagheer, A., Mohammed, M., Riad, K., & Alhajhoj, M. (2020). A Cloud-Based IoT Platform for Precision Control of Soilless Greenhouse Cultivation. *Sensors*, 21(1), 223. <https://doi.org/10.3390/s21010223>
- SDI-12 Support Group. (2021). *SDI-12 A Serial-Digital Interface Standard for Microprocessor-Based Sensors*. Retrieved May 5, 2023, from https://www.sdi-12.org/current_specification/SDI-12_version-1_4-Jan-30-2021.pdf
- Seeed Studio. (2020). *Seeed Studio*. Retrieved June 12, 2023, from <https://files.seeedstudio.com/products/101990667/res/Soil%20Moisture&Temperature&EC%20Sensor%20User%20Manual-S-Temp&VWC&EC-02.pdf>
- Semtech Corporation. (2020). LoRa and LoRaWAN : A technical Overview. Retrieved May 29, 2023, from https://lora-developers.semtech.com/uploads/documents/files/LoRa_and_LoRaWAN-A_Tech_Overview-Downloadable.pdf
- Shieh, G. (2020). Assessing Agreement Between Two Methods of Quantitative Measurements: Exact Test Procedure and Sample Size Calculation. *Statistics in Biopharmaceutical Research*, 12(3), 352–359. <https://doi.org/10.1080/19466315.2019.1677495>
- Siepert, B. (2022). *Adafruit Sensirion SHTC3 – Temperature & Humidity Sensor Breakout*. Retrieved May 31, 2023, from <https://learn.adafruit.com/adafruit-sensirion-shtc3-temperature-humidity-sensor/overview>
- Sparkfun. (2023). *Sparkfun/qwiic*. Retrieved May 31, 2023, from <https://www.sparkfun.com/qwiic>
- Stevens, N. T., Steiner, S. H., & MacKay, R. J. (2017). Assessing agreement between two measurement systems: An alternative to the limits of agreement approach. *Statistical Methods in Medical Research*, 26(6), 2487–2504. <https://doi.org/10.1177/0962280215601133>
- Stöckl, H. (1998). Multimediale Diskurswelten zwischen Text und Bild. in *Mediendiskurse. verbal-Workshop Graz 1996* (S. 73-92). (Sprache im Kontext; Band (Hg.) Bernhard Kettemann & Martin Stegu & Hartmut Stöckl). Lang.
- Symmetron. (n.d.). *Symmetron*. Retrieved May 24, 2021, from https://symmetron.gr/uploads/ThygroSDI-12Brochure_EN.pdf
- Taffé, P., Peng, M., Stagg, V., & Williamson, T. (2017). Biasplot: A Package to Effective Plots to Assess Bias and Precision in Method Comparison Studies. *The Stata Journal: Promoting Communications on Statistics and Stata*, 17(1), 208–221. <https://doi.org/10.1177/1536867X1701700111>

Taffé, P. (2018). Effective plots to assess bias and precision in method comparison studies. *Statistical Methods in Medical Research*, 27(6), 1650–1660. <https://doi.org/10.1177/0962280216666667>

Taffé, P., Halfon, P., & Halfon, M. (2020). A new statistical methodology overcame the defects of the Bland–Altman method. *Journal of Clinical Epidemiology*, 124, 1–7. <https://doi.org/10.1016/j.jclinepi.2020.03.018>

Taffé, P., Peng, M., Stagg, V., & Williamson, T. (2019). MethodCompare: An R package to assess bias and precision in method comparison studies. *Statistical Methods in Medical Research*, 28(8), 2557–2565. <https://doi.org/10.1177/0962280218759693>

Texas Instruments. (2009). *ADS1115 Ultra-Small, Low-Power, 16-Bit, Analog-to-Digital Converter with Internal Reference*. Retrieved September 20, 2021, from.: <https://cdn-shop.adafruit.com/datasheets/ads1115.pdf>

Texas Instruments. (2013). *LM4040-N/LM4040Q-N Precision Micropower Shunt Voltage Reference*. Retrieved September 20, 2021, from <https://cdn-shop.adafruit.com/datasheets/lm4040-n.pdf>

Texas Instruments. (2019). *TPS61023 3.7-A Boost Converter with 0.5-V Ultra-low Input Voltage*. Retrieved May 10, 2023, from https://www.ti.com/lit/ds/symlink/tps61023.pdf?HQS=dis-mous-null-mouser-mode-dsf-pf-null-ww&ts=1691346237065&ref_url=https%253A%252F%252Fgr.mouser.com%252F

The Things Network. (2023). Regional Limitations of RF Use in LoRaWAN. Retrieved May 29, 2023, from [https://www.thethingsnetwork.org/docs/lorawan/regional-limitations-of-rf-use/#:~:text=The%20LoRa%20Alliance%C2%AE%20recommends,Power\)%%20of%20%2B16%20dBm.](https://www.thethingsnetwork.org/docs/lorawan/regional-limitations-of-rf-use/#:~:text=The%20LoRa%20Alliance%C2%AE%20recommends,Power)%%20of%20%2B16%20dBm.)

The Things Stack. (2023). Integrations – ThingSpeak. Retrieved August 20, 2023, from <https://www.thethingsindustries.com/docs/integrations/cloud-integrations/thingspeak/>

Theil, H. (1950) A Rank Invariant Method of Linear and Polynomial Regression Analysis, i, ii, iii. *Proceedings of the Koninklijke Nederlandse Akademie Wetenschappen, Series A Mathematical Sciences*, 53, 386-392, 521-525, 1397-1412.

Thermometrics. (2014). *Sensor Temperature Resistance Curves*. Retrieved September 18, 2019, from <https://www.amphenol-sensors.com/hubfs/Documents/AAS-913-318C-Temperature-resistance-curves-071816-web.pdf?Itemid=8466+%27>

ThingSpeak. (2023). *ThingSpeak documentaion*. Retrieved April 15, 2023, from <https://www.mathworks.com/help/thingspeak/>

Thomas, G. W. (2018). *Soil pH and Soil Acidity* (pp. 475–490). <https://doi.org/10.2136/sssabookser5.3.c16>

Thu Ha, V., Thanh, N. T., & Ha, V. T. (2023). Design and fabrication of a moving robotic glove system. *International Journal of Electrical and Computer Engineering (IJECE)*, 13(3), 2704. <https://doi.org/10.11591/ijece.v13i3.pp2704-2710>

- Tian, L., Santi, S., Seferagić, A., Lan, J., & Famaey, J. (2021). Wi-Fi HaLow for the Internet of Things: An up-to-date survey on IEEE 802.11ah research. *Journal of Network and Computer Applications*, 182, 103036. <https://doi.org/10.1016/j.jnca.2021.103036>
- Trevennor, A. (2012). *Practical AVR Microcontrollers*. Apress. <https://doi.org/10.1007/978-1-4302-4447-9>
- Tschakovsky, M. E., Shoemaker, J. K., & Hughson, R. L. (1995). Beat-by-beat forearm blood flow with Doppler ultrasound and strain-gauge plethysmography. *Journal of Applied Physiology*, 79(3), 713–719. <https://doi.org/10.1152/jappl.1995.79.3.713>
- Tyl Catrin and Sadler, G. D. (2017). pH and Titratable Acidity. In S. S. Nielsen (Ed.), *Food Analysis* (pp. 389–406). Springer International Publishing. https://doi.org/10.1007/978-3-319-45776-5_22
- Tzanakakis, M., & Katsoyannos, B. (2003). Insect Pests of Citrus. In M. Tzanakakis, & B. Katsoyannos, *Insects of Fruit Trees and Grapevine*. (pp. 213-219). Athens: Agrotipos.
- Usach, M. (2015). *AN-1248 Application Note : SPI Interface*. Retrieved April 18, 2023, from: <https://www.analog.com/media/en/technical-documentation/application-notes/AN-1248.pdf>
- Valdez, J., & Becker, J. (2015). *Understanding the I2C Bus*. Retrieved from April 18, 2023, <https://www.ti.com/lit/an/slva704/slva704.pdf>
- Vera, M. T., Rodriguez, R., Segura, D. F., Cladera, J. L., & Sutherst, R. W. (2002). Potential Geographical Distribution of the Mediterranean Fruit Fly, *Ceratitis capitata* (Diptera: Tephritidae), with Emphasis on Argentina and Australia. *Environmental Entomology*, 31(6), 1009–1022. <https://doi.org/10.1603/0046-225X-31.6.1009>
- Vinod Kumar, S., Singh, C. D., Rao, K. V. R., Kumar, M., & Rajwade, Y. A. (2023). Development of a smart IoT-based drip irrigation system for precision farming. *Irrigation and Drainage*, 72(1), 21–37. <https://doi.org/10.1002/ird.2757>
- Vishay. (2016). *Metal Film Resistors, Axial, High Precision, High Stability*. Retrieved September 18, 2023, from <https://gr.mouser.com/datasheet/2/427/ptf-1763051.pdf>
- Voigt, C., Montero, C. S., & Menichinelli, M. (2016). An empirically informed taxonomy for the maker movement. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9934 LNCS, 189–204. https://doi.org/10.1007/978-3-319-45982-0_17
- Wang, K., & Liu, H. (2021). *RS-485 Basics Series - Technical White paper*. Retrieved April 19, 2023, from <https://www.ti.com/lit/wp/slla545/slla545.pdf?ts=1685251929193>
- Wang, W., & Gene Hwang, J. T. (2001). A nearly unbiased test for individual bioequivalence problems using probability criteria. *Journal of Statistical Planning and Inference*, 99(1), 41–58. [https://doi.org/10.1016/S0378-3758\(01\)00074-X](https://doi.org/10.1016/S0378-3758(01)00074-X)
- Wavehsare. (2018). *Waveshare*. Retrieved March 29, 2023, from https://www.waveshare.com/w/upload/e/e4/USB_TO_RS232_485_TTL_-user-manual-en.pdf

Weinberg, M. (2015). Licensing Open Source Hardware. In S. A. Alicia Gibb, *Building Open Source Hardware*. Pearson Education.

Weinberg, M. (2023). *Open-Source Hardware Association*. Retrieved February 15, 2023, from <https://www.oshwa.org/faq/#what-is-a-license>: <https://www.oshwa.org/faq/#what-is-a-license>

Wemos Electronics. (n.d.). *Wemos Electronics*. Github. <https://github.com/wemos>

White, I., & Elson-Harris, M. (1992). Fruit Flies of Economic Significance: Their Identification and Bionomics. In I. White, & M. Elson-Harris, *Fruit Pest Species*. (1 ed., pp. 291-297). Wallingford: CAB.

Williamson, J. M., Crawford, S. B., & Lin, H. (2007). Resampling Dependent Concordance Correlation Coefficients. *Journal of Biopharmaceutical Statistics*, 17(4), 685–696.
<https://doi.org/10.1080/10543400701329471>

Wisniewski N. (2023). remote: Tools for Bland-Altman Analysis. R package version 0.1.0.
<https://github.com/nwisn/blandPower>

World, R. W. (2012). *Advantages of 1-wire interface | disadvantages of 1-wire interface*. Retrieved May 11, 2023, from <https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-1-wire-interface.html#:~:text=Drawbacks%20or%20disadvantages%20of%201%2Dwire%20interface&text=The%20synchronization%20of%20data%20at,supports%20slower%20speed%20of%20communicatio>

Yakman, Georgette. (2008). STEAM Education: an overview of creating a model of integrative education.

Yida. (2019). *UART vs I2C vs SPI – Communication Protocols and Uses*. Retrieved May 11, 2023, from <https://www.seeedstudio.com/blog/2019/09/25/uart-vs-i2c-vs-spi-communication-protocols-and-uses/>

Zdravković, M., Trajanovic, M., Sarraipa, J., Jardim-Goncalves, R., Lezoche, M., Aubry, A., & Panetto, H. (2016). 6th International Conference on Information Society and Technology. *Survey of Internet-of-Things platforms*.

Zheng, H., Zhang, Q., Quan, J., Zheng, Q., & Xi, W. (2016). Determination of sugars, organic acids, aroma components, and carotenoids in grapefruit pulps. *Food Chemistry*, 205, 112–121.
<https://doi.org/10.1016/j.foodchem.2016.03.007>

Appendix A

Table A1. ATmega32u4 pin mapping (source: Arduino (2013))

Pin Number	Pin Name	Mapped Pin Name
1	PE6 (INT.6/AIN0)	Digital pin 7
2	UVcc	+5V
3	D-	RD-
4	D+	RD+
5	UGnd	UGND
6	UCap	UCAP
7	VUSB	VBus
8	(SS/PCINT0) PB0	RXLED
9	(PCINT1/SCLK) PB1	SCK
10	(PDI/PCINT2/MOSI) PB2	MOSI
11	(PDO/PCINT3/MISO) PB3	MISO
12	(PCINT7/OC0A/OC1C/#RTS) PB7	Digital pin 11 (PWM)
13	RESET	RESET
14	Vcc	+5V
15	GND	GND
16	XTAL2	XTAL2
17	XTAL1	XTAL1
18	(OC0B/SCL/INT0) PD0	Digital pin 3 (SCL)(PWM)
19	(SDA/INT1) PD1	Digital pin 2 (SDA)
20	(RX D1/AIN1/INT2) PD2	Digital pin 0 (RX)
21	(TXD1/INT3) PD3	Digital pin 1 (TX)
22	(XCK1/#CTS) PD5	TXLED
23	GND1	GND
24	AVCC	AVCC

Pin Number	Pin Name	Mapped Pin Name
25	(ICP1/ADC8) PD4	Digital pin 4
26	(T1/#OC4D/ADC9) PD6	Digital pin 12
27	(T0/OC4D/ADC10) PD7	Digital Pin 6 (PWM)
28	(ADC11/PCINT4) PB4	Digital pin 8
29	(PCINT5/OC1A/#OC4B/ADC12) PB5	Digital Pin 9 (PWM)
30	(PCINT6/OC1B/OC4B/ADC13) PB6	Digital Pin 10 (PWM)
31	(OC3A/#OC4A) PC6	Digital Pin 5 (PWM)
32	(ICP3/CLK0/)OC4A) PC7	Digital Pin 13 (PWM)
33	(#HWB) PE2	HWB
34	Vcc1	+5V
35	GND2	GND
36	(ADC7/TDI) PF7	Analog In 0
37	(ADC6/TDO) PF6	Analog In 1
38	(ADC5/TMS) PF5	Analog In 2
39	(ADC4/TCK) PF4	Analog In 3
40	(ADC1) PF1	Analog In 4
41	(ADC0) PF0	Analog In 5
42	AREF	AEF
43	GND3	GND
44	AVCC1	AVCC

Table A2. ESP 32 variations characteristics.

	ESP32	ESP32-S2	ESP32-S3	ESP32-C3	ESP32-C6
Announcement Date	2016, September	2019, September	2020, December	2020, November	2021, April
Main processor	Tensilica Xtensa 32-bit LX6 (up to 240MHz) (optionally dual core)	Tensilica Xtensa 32-bit LX7 (up to 240MHz)	Tensilica Xtensa 32-bit LX7 dual core (up to 240MHz)	RISC-V 32-bit (up to 160MHz)	RISC-V 32-bit (up to 160MHz)
SRAM	520KB	320KB	512KB	400KB	400KB
ROM	448KB	128KB	384KB	384KB	384KB
JTAG	✓	✓	?	✓	✓
Cache	64KB	8/16KB (configurable)	?	16KB	?
WiFi	Wi-Fi 4	Wi-Fi 4	Wi-Fi 4	Wi-Fi 4	Wi-Fi 6
Bluetooth	BLE 4.2 (upgrade to 5.0, with limitations)	X	BLE 5.0	BLE 5.0	BLE 5.0
Ethernet	✓	X	?	X	?
RTC memory	16KB	16KB	16KB	8KB	?
PMU	✓	✓	?	✓	?
ULP coprocessor	✓	ULP-RISC-V	?	X	?
Cryptographic Accelerator	SHA, RSA, AES, RNG	SHA, RSA, AES, RNG, HMAC, Digital Signature	SHA, RSA, AES, RNG, HMAC, Digital Signature	SHA, RSA, AES, RNG, HMAC, Digital Signature	SHA, RSA, AES, RNG, HMAC, Digital Signature
Secure boot	✓	✓	✓	✓	✓
Flash encryption	✓	XTS-AES-128/256	✓	XTS-AES-128	XTS-AES-128
SPI	4	4	?	3	?
I2C	2	2	?	1	?
I2S	2	1	?	1	?
UART	3	2	?	2	?
SDIO Host	1	0	2	0	0
SDIO Slave	1	0	0	0	0
GPIO	34	43	44	22	22
LED PWM	16	8	?	6	?
MCPWM	6	0	2	0	0
Pulse counter	8	4	?	0	X
GDMA*	0	0	?	6	?
USB	X	USB OTG 1.1	?	Serial/JTAG	?
TWAI**	1	1	?	1	?
ADC	2x 12-bit SAR, up to 18 channels	2x 13-bit SAR, up to 20 channels	?	2x 12-bit SAR, up to 6 channels	?
DAC	2x 8-bit	2x 8-bit	?	X	X
RMT	8x transmission/reception	4x transmission/reception	?	2x transmission + 2x reception	?
Timer	4x 64-bit	4x 64-bit	?	2x 54-bit + 1x 52-bit	?
Temperature Sensor	✓	✓	?	✓	?

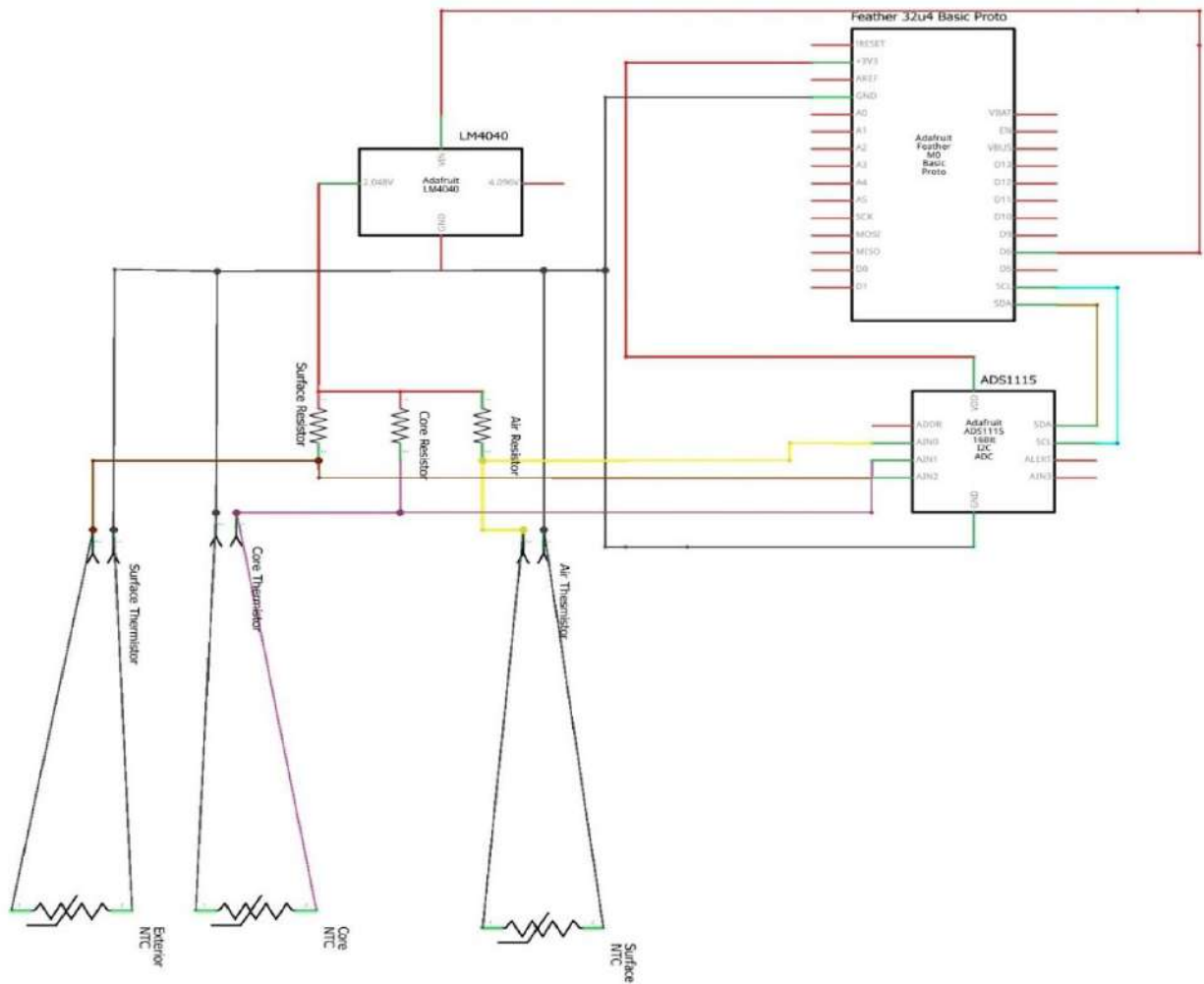
Hall Sensor	✓	X	?	X	?
Touch Sensor	10	14	?	X	?

Table A3. Raspberry Pi 3, model B Specifications

<i>Processor</i>	<i>Quad Core 1.2 GHz Broadcom BCM2837 CPU</i>
<i>RAM</i>	<i>1 GB</i>
<i>Wireless Connectivity 1</i>	<i>BCM43438 wireless LAN (WiFi)</i>
<i>Wireless Connectivity 2</i>	<i>Bluetooth Low Energy (BLE)</i>
<i>Wired Connectivity 1</i>	<i>100 Base Ethernet</i>
<i>Wired Connectivity 2</i>	<i>40-pin extended GPIO</i>
<i>Wired Connectivity 3</i>	<i>4 USB 2 ports</i>
<i>Media Output 1</i>	<i>4 pole stereo output</i>
<i>Media Output 2</i>	<i>1 composite video port</i>
<i>Media Output 3</i>	<i>1 full size HDMI port</i>
<i>Media Input 1</i>	<i>1 CSI camera port for connecting Raspberry Pi camera</i>
<i>Media Input/Output 1</i>	<i>1 DSI display port for connecting a Raspberry Pi touchscreen display</i>
<i>Media Input/Output 1</i>	<i>1 Micro SD port for loading an operating system and storing data</i>
<i>Power Source</i>	<i>1 Micro-USB power source up to 2.5 A</i>

Table A4. Specifications of the thermistors.

Thermistor	Tolerance	Resistance at 25 °C Nominal	Thermistor Diameter
MC65F103A	0.05 °C	10 KOhms	1.65 mm maximum
JS8746A-0.15	0.15 °C	10 KOhms	5.8 mm



fritzing

Figure A1. The circuit sketch describes the connections among the electronic parts. The sketch was created using Fritzing.

Table A5. The coefficients of the Hoge-2 equation based on the values of the fraction R_t/R_{25} , which is the value of the resistance measured using Ohm's Law over the value of the reference resistance (10 KOhm in our case).

Rt/R25 range	a	b	c	d
68.600 to 3.274	3.3538646E-03	2.5654090E-04	1.9243889E-06	1.0969244E-07
3.274 to 0.36036	3.3540154E-03	2.5654090E-04	2.0829210E-06	7.3003206E-06
0.36036 to 0.06831	3.3539264E-03	2.5609446E-04	1.9621987E-06	4.6045930E-08
0.06831 to 0.01872	3.3368620E-03	2.4057263E-04	-2.6687093E-06	-4.0719355E-07

Table A6. Formulas used to calculate the resistance of the thermistor (R_t), the temperature reading (Hoge-2) and the ADC value.

RT	Temperature Reading (Hoge-2)	ADC Value
$\frac{10000}{\frac{32767}{ADCvalue} - 1}$	$\frac{1}{A + B \log\left(\frac{R_t}{R_{ref}}\right) + C \left(\log\left(\frac{R_t}{R_{ref}}\right)\right)^2 + D \left(\log\left(\frac{R_t}{R_{ref}}\right)\right)^3}$	$adc \frac{V_{cc} \text{ on gain two}}{V_{cc} \text{ reference}}$

Table A7. Adafruit feather M0 RFM95 Lora Radio technical specifications.

Technical Specifications	
Flash Memory	256 K
RAM	32K
Clock	48 MHz
Microcontroller	ATSAMD21G18 ARM Cortex M0
Lipo connector	Yes
Footprint	51mm x 23mm x 8mm
Logic/Power	3.3 Volts
GPIO pins	20
Analog Outputs	1
Analog Inputs	10
Communication Protocols	Hardware Serial, I2C, SPI
Battery Readings	Yes / Pin A7
ADC Resolution	12 bit
DAC Resolution	12 bit
LoRa module	SX1276 with SPI interface (Hope RF)

Adafruit Feather M0 RFM69

<https://www.adafruit.com/products/3176>

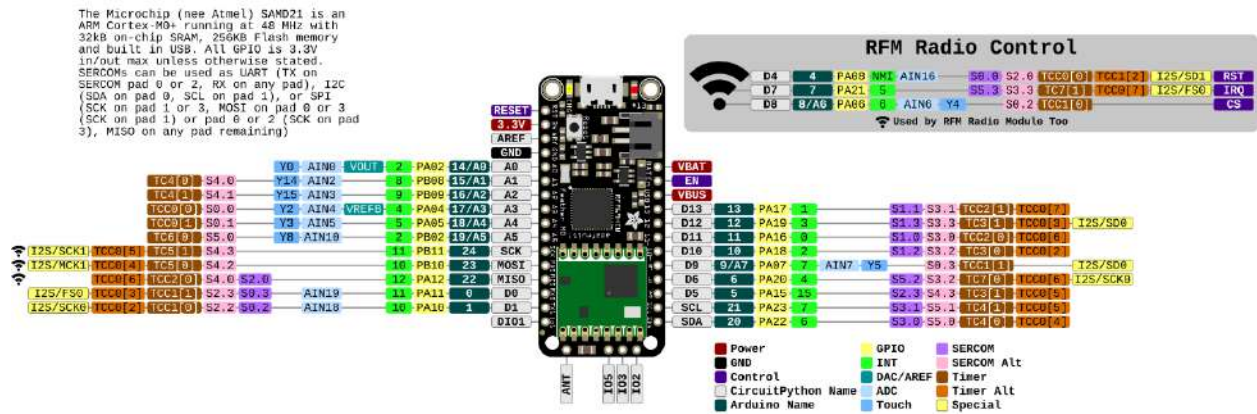


Figure A2. Adafruit Feather M0 RFM69 pinout (source: Adafruit official website).

Table A8. SM150 soil moisture sensor and HH150 meter technical specifications.

SM150 soil moisture sensor	
Accuracy	±3% vol over 0 to 70% vol and 0~60°C
Measurement Range	0~100% vol but less accurate above 70% vol
Salinity Error	±5% vol over 100 to 1000 mS.m ⁻¹ and 0~60% vol
Output Signal	0~1V differential 0 to 60% nominal
Output Compatible with	HH150, HH2, GP2, GP1, DL6, DL2e
Maximum Cable Length	1m
Power Requirement	5 – 14 VDC, 18mA for 1 s.
Operating Range	–20 to 60°C
Environment	IP68
Sample Volume	55 × 70mm diameter
Dimensions/Weight	143 × 40mm diameter/0.1 kg
HH150 Meter	
Accuracy	±7.5 mV (negligible effect on SM150 accuracy)
Resolution	0.1% of volumetric reading or 1mV
Battery/life/stand by life	2 ×GP alkaline AAA/10000 readings/1 year
Environmental	Noncondensing
Compliance	CE, FCC & ROHS

Table A9. Technical specifications for Hanna Instruments HI9024 Waterproof pH meter.

Technical Specifications	
Range	0.00 to 14.00 pH ±399.9 mV (ISE); ±1999 mV (ORP) 0.0 to 100°C ; 0.0 to 70°C
Resolution	0.01 pH 0.1 mV (ISE); 1mV (ORP) 0.1°C
Accuracy (at 20°C)	±0.1 pH ±0.2 mV (ISE); ±1 mV ±0.2 °C; ±0.5°C
Calibration	Automatic with 3 mentioned standard buffers (4.01, 7.01, 10.01)
Temperature Compensation	Automatic or manual 0 to 70°C
Environment	0 to 50°C; 100% RH
Dimension	196 x 80 x 60 mm (meter) 340 x 230 x 80 mm (kit)

Table A10. Summary of the LCC, LPC and LA for the assessment of agreement between the industrial and open-source device.

	Time	LCC	LPC	LA
1	11	0.2236	0.2243	0.9969
2	12	0.2217	0.2225	0.9967
3	13	0.2199	0.2206	0.9964
4	14	0.218	0.2188	0.9962
5	15	0.2161	0.217	0.9959
6	16	0.2142	0.2152	0.9956
7	17	0.2124	0.2134	0.9953
8	18	0.2105	0.2116	0.9949
9	21	0.205	0.2063	0.9937
10	22	0.2032	0.2046	0.9933
11	23	0.2014	0.2028	0.9928
12	24	0.1996	0.2011	0.9924
13	25	0.1978	0.1994	0.9918
14	26	0.196	0.1977	0.9913
15	27	0.1942	0.196	0.9908
16	28	0.1924	0.1943	0.9902
17	31	0.1871	0.1893	0.9883
18	32	0.1853	0.1877	0.9876
19	33	0.1836	0.186	0.9869

20	34	0.1819	0.1844	0.9862
21	35	0.1801	0.1828	0.9854
22	36	0.1784	0.1812	0.9846
23	37	0.1767	0.1796	0.9838
24	38	0.175	0.178	0.983
25	41	0.1699	0.1734	0.9803
26	42	0.1683	0.1718	0.9794
27	43	0.1666	0.1703	0.9784
28	44	0.165	0.1688	0.9774
29	45	0.1633	0.1673	0.9764
30	51	0.1537	0.1585	0.9698
31	52	0.1521	0.157	0.9686
32	53	0.1505	0.1556	0.9674
33	54	0.149	0.1542	0.9661
34	55	0.1474	0.1528	0.9649
35	56	0.1459	0.1514	0.9636
36	61	0.1384	0.1446	0.9567
37	62	0.1369	0.1433	0.9552
38	63	0.1354	0.142	0.9538
39	64	0.1339	0.1407	0.9522
40	65	0.1325	0.1394	0.9507
41	66	0.1311	0.1381	0.9492
42	67	0.1296	0.1368	0.9476
43	68	0.1282	0.1355	0.946
44	71	0.124	0.1318	0.941
45	72	0.1226	0.1306	0.9393
46	73	0.1213	0.1293	0.9376
47	74	0.1199	0.1281	0.9359
48	75	0.1186	0.1269	0.9341
49	76	0.1172	0.1257	0.9323
50	77	0.1159	0.1246	0.9305
51	78	0.1146	0.1234	0.9286

Appendix B

Detailed list of the components of FruiTemp (Chapter 6)

1. Amphenol NTC thermistor MC65 series, 10K 25 °C is a precision solid state temperature sensor. The tolerance and interchangeability of this sensor is ± 0.05 °C. It is resin coated for good mechanical strength and resistance to solvents. The sensor itself is a 1.65 mm diameter “point”-sensor with 0.1 mm diameter heavy isomid insulated bifilar nickel lead wires. Its thermal time constant is 0.5 s in stirred oil and 8 s in still air. The datasheet is available in the reference (Amphenol, 2014).
2. Metal Film Resistor Vishay/Dale (Vishay, 2016). Through Hole 1/4-watt 10 Kohms with 0.05% tolerance and 5 ppm.
3. Metal Film Resistor Vishay/Dale (Vishay, 2016). Through Hole 1/4-watt 10 Kohms 0.1% tolerance 5 ppm.
4. Amphenol Industrial Temperature Sensor JS8746A (Amphenol, 2019). The tolerance and interchangeability of this sensor is ± 0.15 °C. According to the manufacture notes the sensor has been designed to address all aspects of temperature measurement for HVAC (Heating, Ventilation and air conditioning) control systems. Employing materials and build standards that enhance the sensor’s ability to withstand water ingress and degradation, the sensor meets statutory requirements for temperature measurement and performance expectations. It is equipped with environmental protection IP68, it is resistant to salt solutions, ozone, UV and a variety of marine environment cleaning detergents. It is equipped with stainless steel hard shell, with class corrosion resistance A2 or better according to ISO3506.
5. Precision LM4040 Voltage Reference Breakout 2048 V and 4096 V (Texas Instruments, 2013). This breakout was used due to the noisy voltage reference of the adafruit feather 32u4. To achieve high precision, we used the breakout to have a reference of 2048 V for the thermistors. Using Ohm’s Law, the reference had to be stable as possible to calculate the temperature with high precision and accuracy.
6. ADS1115 16-bit ADC-4 channels with Programmable Gain Amplifier (Texas Instruments, 2009). This part was used due to the lack of the feather’s 32u4 precision in the conversion of analog to digital signal. The feather has a 10-bit precision ADC, so we used this part to reach a 16-bit ADC for the system to be more sensitive to small changes of the voltage. For that reason, the system can be more precise. Moreover, we used the amplifier of the part to amplify the signal. We kept the amplification factor low to make the noise as low as possible.
7. Adafruit Feather 32u4 Basic Proto (Fried L, Adafruit Feather 32u4 Basic Proto, 2023). This development board was used because it has proven to be robust and tested for many years since it was released. The MCU is clocked at 8 Mhz and at 3.3 V logic. It has a built-in USB-to-serial program and debug capability built in. Additionally, it is a low consumption microcontroller, and the board is equipped with a connector for any 3.7 Lithium Polymer battery and built in charging.
8. Adafruit Adalogger FeatherWing (Fried L, Adafruit Adalogger FeatherWing, 2023). This shield is equipped with an SD-card module and a battery-packed Real Time Clock. It can be plugged on top of the microcontroller board and can be easily used.
9. Lithium Polymer single cell 3.7 Volt and a capacity of 1200 mah.
10. A switch.
11. A switch waterproof cap.

12. Cable Glands PG7 .
13. An IP 54 box by HAMMOND (Hammond Manufacturing, 2022).
14. Heat-Shrinkable tubes, 3 mm and 0.8 mm.
15. Prototyping hook-up single wires 22 AWG.
16. Double-Side Printed Circuit Board.
17. Radiation shield. A custom-made radiation shield made of plastic was designed and constructed using CNC machinery.

Table B1. Chemical composition of the substrate during the Chapter 4 experiment.

Chemical Composition	
SiO ₂	74.0% - 78.0%
Al ₂ O ₃	10.0% - 16.5%
Fe ₂ O ₂	0.5% - 2.0%
MgO	0.0% - 1.0%
CaO	0.0% - 2.0%
Na ₂ O	1.0% -5.0%
K ₂ O	1.0% - 4.0%
L.O.I. (SO ₂ ,CO ₂ ,H ₂ O)	2.0% - 6.0%
Traces	0.2% - 0.8%

Table B2. Physical properties of the perlite substrate during the Chapter 4 experiment.

Physical Properties	
Color	White
Odor	Odorless
Bulk Density	65 kg/m ³ (±10%)
Specific Gravity	2.1- 2.3
Moisture	Max. 1%
pH	6.5 – 8.5
Thermal Conductivity	0.035 – 0.040 W/mK
Softening Point	890 – 1100 ⁰ C
Size of perlite	0.5 - 3 mm
Origin	Greek

Appendix C

Chapter 3

Diagnostics for the model

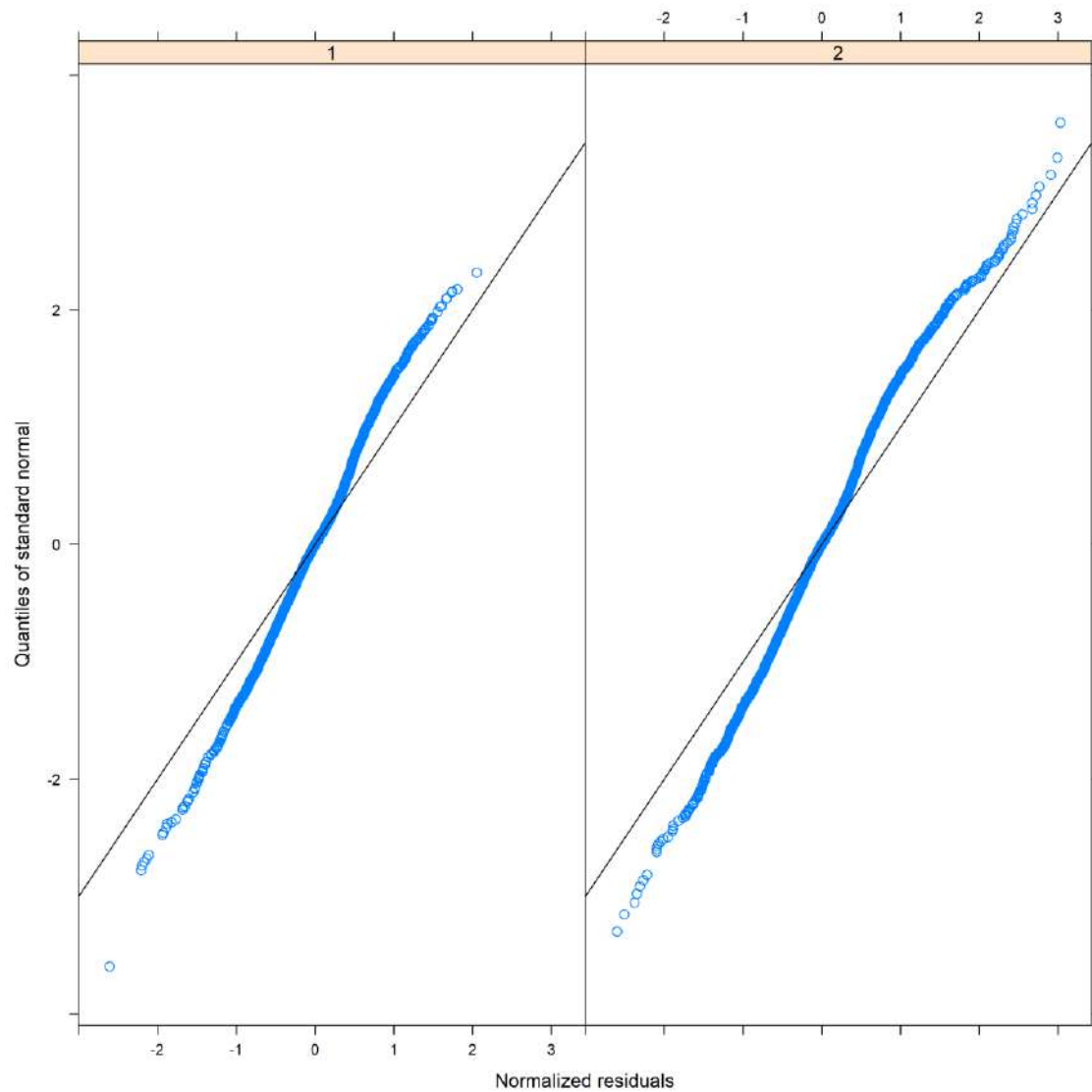


Figure C1. Diagnostics for the model. There is evidence of significant deviation from normality.

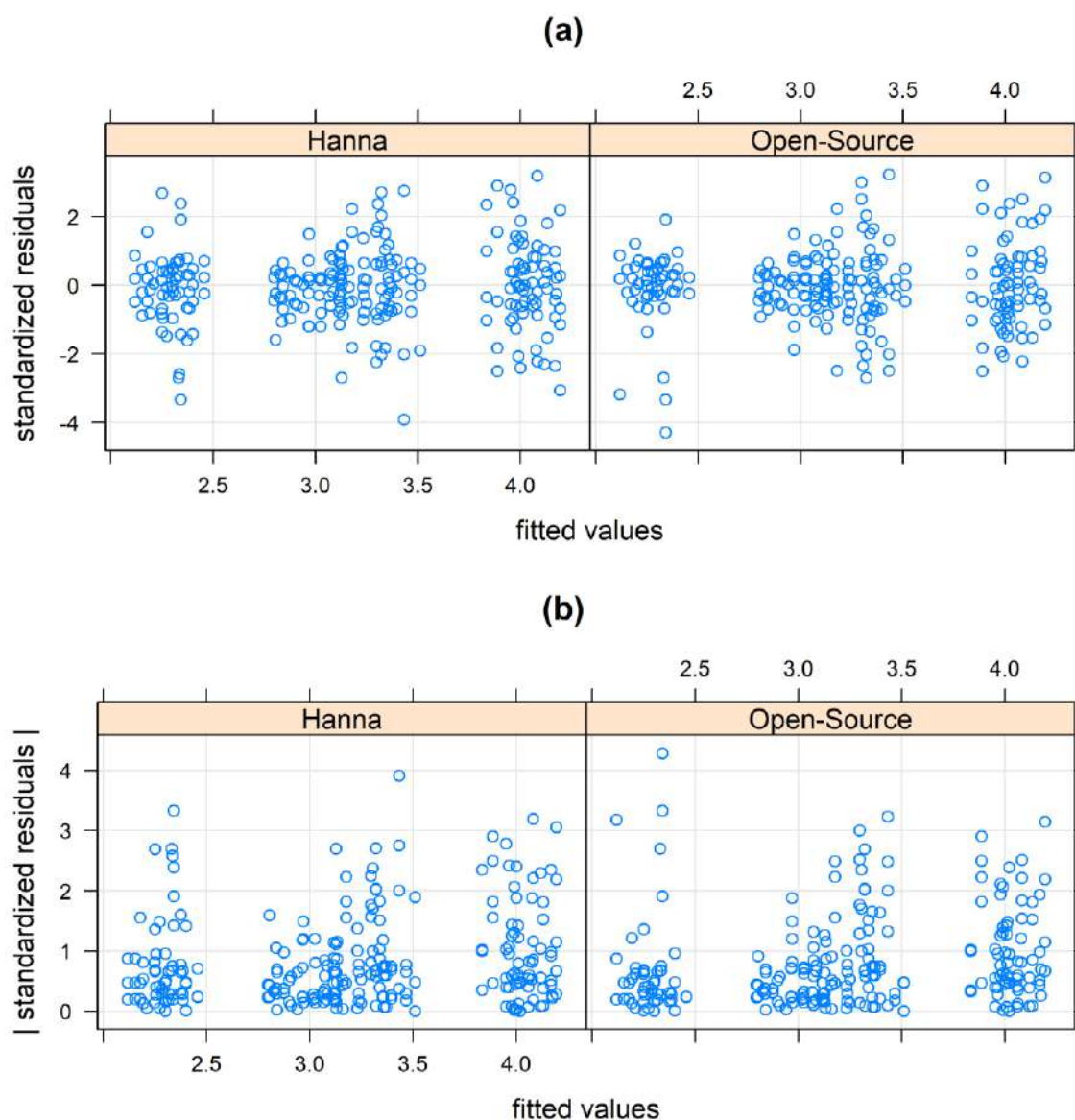


Figure C2. (a) Standardized residuals versus fitted values of the homoscedastic model for Hanna pH meter and the open-source device. (b) Absolute values of standardized residuals versus fitted values of the homoscedastic model for Hanna pH meter and the open-source device. There is no obvious sign of a fan shape thus a formal criterion (AIC, BIC and log-likelihood) will be used to compare the homoscedastic and the heteroscedastic model.

Analysis for the homoscedastic model

CCC and TDI for assessing agreement.

Before Recalibration		
	Estimate	Lower Confidence Bound
CCC	0.6015361	0.4821898
	Estimate	Upper Confidence Bound
TDI	0.29239970	0.31309797
After Recalibration		
	Estimate	Lower Confidence Bound
CCC	0.9626004	0.9355935
	Estimate	Upper Confidence Bound
TDI	0.09096271	0.10916581

Reproducibility Assessment		
Reproducibility	Estimate	Lower Confidence Bound
CCC (Hanna)	0.9915	0.9868
CCC (Open-Source)	0.9962	0.9942
	Estimate	Upper Confidence Bound
TDI (Hanna)	0.0435	0.0467
TDI (Open-Source)	0.0286	0.0307

Similarity Assessment		
	Estimate	Confidence Interval
Lambda	2.3061	[1.8167, 2.9274]
	Estimate	Confidence Interval
Fixed Bias	-0.2215	[-0.2398, -0.2033]

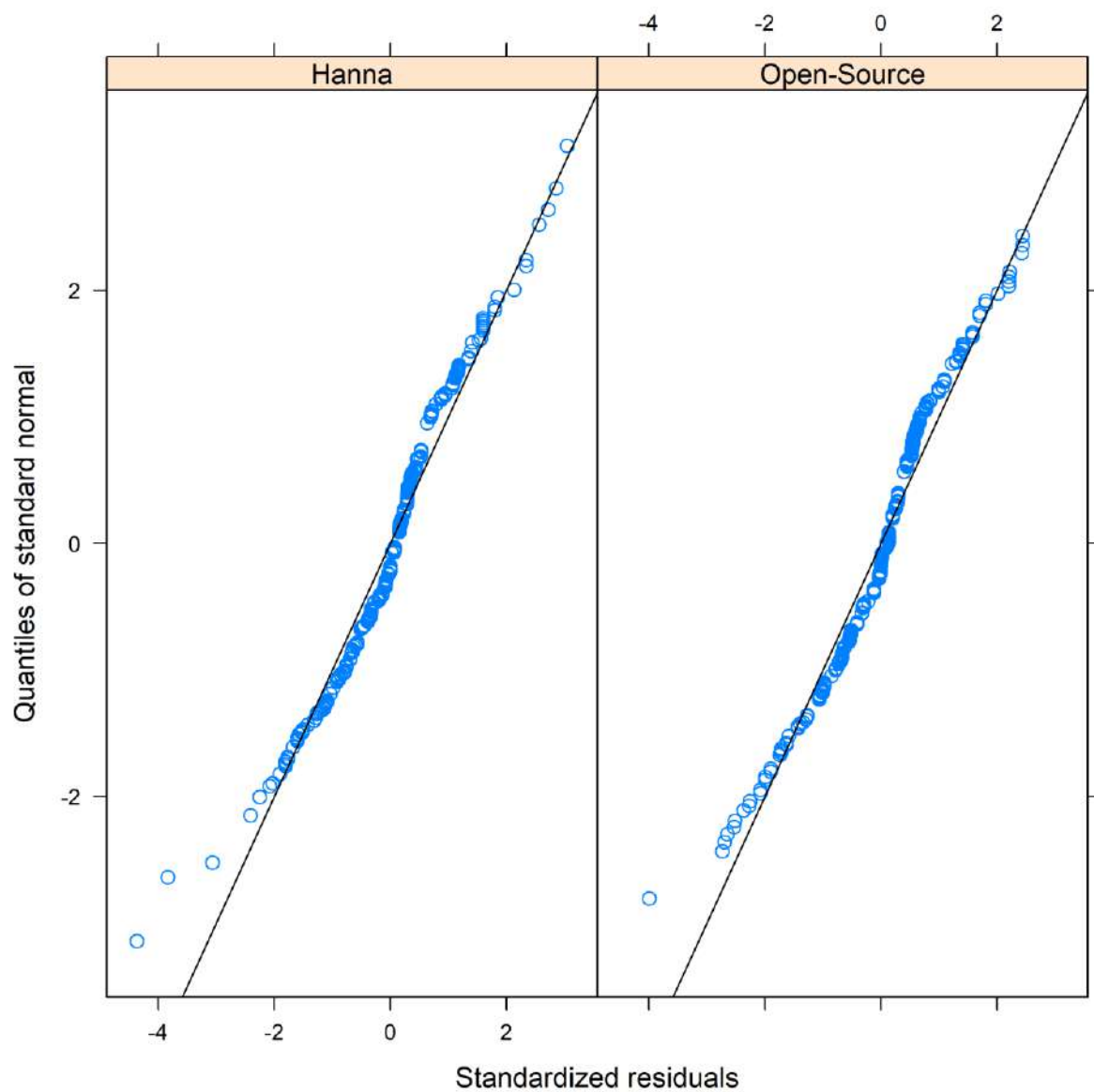


Figure C3. Standardized residuals on the horizontal axis vs the quantiles of the standard normal distribution. The plot reveals a slight deviation from the normal distribution.

Appendix D

D1. Business Models

A qualitative description of open-source development practices (OSPD) is delivered by Mies (2020) from an engineering and management aspect. According to the setting of Mies (2020) “OSPD is defined as the development of open-source hardware products in a collaborative process permitting the participation of any person interested”. Specifically, the authors focus on four standpoints. First, the business models that apply for the open-source practices. Second, the organization and structure of the surrounding community. Third, the design process. Fourth, the online support tools used.

According to Pearce (2015) three methods to quantify the value of open-source hardware development are reviewed. This is essential because it is rather difficult to use the prevalent funding models to support the vital resources for open-source devices development. The first method is the downloaded substitution valuation. This method is based on the number of times that an open hardware device is accessed on the internet. The second method is based on the hours and wage of the workers needed to replicate the open-source device. The last method is based on significant market share of the product. For all the methods a formula is provided for quantification of the value of the device.

D2. Licenses, copyrights, patents and trademarks

As far as the legal part of patenting open-source devices, things might be complicated comparing to the traditional open-source creations. This is a consequence of the hardware addition. According to Weinberg (2015) which is based in the United States laws “in the context of open-source hardware (OSHW), most of the licenses involved will relate to intellectual property rights. By and large, these licenses will grant permission to do things like copy, incorporate, and build upon existing projects.” The most common licenses in the open-source community are the GPL (GNU Public License) and CC (Creative Commons). GPL licenses whoever receives software to use it as they wish, including copying, modify, or distribute. This license grants the right to create derivative software, only if the new source code is also released. Furthermore, if someone does not include an open license for her/his files in any place they are stored, they automatically receive copyright. This is not considered as a limitation in the sharing process. On the contrary, open licenses are viewed as an opportunity to forward and embolden sharing. “Using an open license is an affirmative declaration in support of building a commons” (Weinberg, 2015).

When hardware enters the open-source community the open-source licensing scheme is not very straightforward. As (Google Open Source, 2023) states “A patent gives its owner the right to exclude others from making, using, and selling the claimed invention. In contrast, open-source licenses grant broad rights to modify compile, distribute, and use the software”. When it comes to different licensing schemes Weinberg (2015) covers many aspects and quotes many examples. The author of Licensing Open Source Hardware (Weinberg, 2015) differentiates between copyright, patent and trademark and considers them as “complementary sets of rights”. For the open-source devices copyright is used mostly for creative works, meaning parts of the software that the device uses or any design file, a note or a sketch. The copyright duration is lifetime plus 70 years after death of the author.

Patents cover the parts that are functional. In other words, things that do something. For example, a logger uses an amplifier that makes it feasible to read the signal of a sensor, without it there is no way for the logger to read the signal and the device cannot work as a logger for the specified sensor. On the other hand, the choice of the color for the blinking LEDs to define the ready state of the sensor or not is a

creative part. One could choose another color or another indication. This covers better the hardware part of the open-source device. To apply for a patent, a laborious procedure must follow including a huge amount of paperwork and a significant amount of money. The claimer must also prove that the device is unique and novel. If the claimer is granted with the patent, then its duration is 20 years.

As Weinberg (2023) states, "Patent law and copyright are mutually exclusive. In other words, something either fits within the scope of patent law or it fits within the scope of copyright law. In cases where an object seems to combine both creative and functional part, the law does its best to separate the two elements out."

When it comes to trademark, the law covers the name and the way to identify authentic goods that claim to be a product of a specific company or individual. It is considered as an easier alternative to getting a patent since it is less expensive and time consuming. Since open hardware consists of functional parts apart from the intellectual, copyright will not be able to protect each aspect of its entity. The nonfunctional parts of the device will be certainly licensable from open licenses. Another easier way to indirectly protect a makers' device is protecting his name and reputation. By licensing a trademark for an open-source device, the maker registers her/his name on it. If a maker creates a new project, the public will know it comes from her/him. Other makers might modify it or reproduce it, but they cannot build it using the original makers' name. The device will remain open source by staying loyal to the idea of free access, but the original maker will have the right to use the name of the device only. Thus, the individuals interested in buying the device will have a reference on where to look since there might be numerous similar devices using the same architecture of functions. Trademarks are easier and cheaper to issue unlike patents.

To sum up, by using parts of open-source documentation and resources, the user can give away or sell his/her creation legally according to the open-source hardware licenses. In other words, if a maker wants to sell their device, where its construction is based on open-source licenses, then she/he is free to do so, but comply with the original license's conditions. A person can apply for a patent even if they used chunks of code or pieces of hardware design that is based on open-source license. However, the device will not receive the standard protection that a typical patent is granted since the one who applies for a patent must respect the licensing and terms of the software or hardware, he used to create her/his device. The need to sort issues of patenting devices that are based on open-source material is essential more than ever. Thus, many lawyers are specialized in these matters (Hall Ellis Solicitors, 2023). It is wise to consult with such professionals since this topic is still new, sometimes unclear and can legally change in a short period of time.

Appendix E: Extended Greek Abstract

Η παρούσα διατριβή έχει ως στόχο να συνεισφέρει στις μεθοδολογικές προσεγγίσεις που αφορούν την αξιολόγηση της συμφωνίας και ομοιότητας μεταξύ συσκευών ανοιχτού κώδικα και υλικού σε σχέση με αντίστοιχες εμπορικές με τη χρήση στατιστικών μεθόδων και πειραματικού σχεδιασμού. Οι συσκευές ανοιχτού κώδικα και υλικού χρησιμοποιούνται σε μεγάλο αριθμό εφαρμογών σε τομείς υγείας, εκπαίδευσης, γεωπονίας, αθλημάτων, βιολογίας, έρευνας και άλλων. Η παρούσα διατριβή εστιάζει στους τομείς της γεωπονίας και έρευνας.

Το Κεφάλαιο 1 παρουσιάζει μία σύντομη εισαγωγή για τους παράγοντες που έχουν συνεισφέρει στην διάδοση των συσκευών ανοιχτού κώδικα και υλικού. Το κίνημα “Maker”, (Papavlasopoulou et al., 2017), το συμμετοχικό Web 2.0 (Atkinson et al., 2012), η κοινότητα Open-Source, το μειωμένο κόστος των ηλεκτρονικών εξαρτημάτων και οι κοινωνικές επιρροές, είναι κάποιοι από τους παράγοντες που έχουν συνεισφέρει σημαντικά στην εξάπλωση των συσκευών ανοιχτού κώδικα και υλικού. Επιπλέον, πλακέτες ανάπτυξης όπως η Arduino (Louis, 2018; Arduino, 2015) και η Raspberry-pi (Jolles, 2021), ενίσχυσε αυτό το φαινόμενο με την απλοποίηση εξειδικευμένων ηλεκτρονικών συνδεσμολογιών χρησιμοποιώντας πληθώρα πακέτων λογισμικού και γλώσσες προγραμματισμού. Τέλος, σύντομες παρουσιάσεις (tutorials) παρέχονται από πολλές εταιρείες που κατασκευάζουν πλακέτες ανάπτυξης, για χρήστες με μερική ή καθόλου εμπειρία. Στη συνέχεια, παραθέτονται ορισμοί για τον ανοιχτό κώδικα και το υλικό (Open-Source Community 2023, Open-Source Hardware Association, 2023; Open-Source Initiative, 2023) όπως και την περιγραφή των εφαρμογών τους στη γεωργία ακριβείας, έξυπνη γεωργία και έρευνα. Συγκεκριμένα, αναφέρει τη συνεισφορά και παραδείγματα εφαρμογής της φιλοσοφίας ανοιχτού υλικού και κώδικα στην έξυπνη γεωργία και γεωργία ακριβείας (Pedersen & Lind, 2017; Patle et al. 2022; Li et al. 2023). Ο Oellerman et al. (2022) επισημαίνει τρία σημεία που συνεισφέρουν οι συσκευές ανοιχτού κώδικα και υλικού στην επιστημονική κοινότητα. Ακόμα επισημαίνεται η φιλοσοφία του ανοιχτού κώδικα και υλικού όπως και η ελεύθερη διάδοση της γνώσης για κάθε άνθρωπο που επιθυμεί να αντιγράψει, τροποποιήσει ή δημιουργήσει συσκευές της συγκεκριμένης ιδέας, ενώ ταυτόχρονα μοιράζεται τη διαδικασία και όλα τα υλικά και μεθόδους που χρησιμοποίησε. Παρουσιάζονται παραδείγματα δημοσιεύσεων που αφορούν τη συγκεκριμένη φιλοσοφία (Coleman et al., 2022; Carrascosa et al., 2015; Aliagas et al., 2022; Ji et al., 2019) όπως και τα πλεονεκτήματα χρήσης τους ενώ αναγνωρίζεται η ανάγκη ανάπτυξης πλαισίου για την αξιολόγηση των λειτουργιών και της αξιοπιστίας των μετρήσεων των συγκεκριμένων συσκευών. Συγκεκριμένα, ορίζεται η έννοια των ερευνών συμφωνίας και ομοιότητας (Dunn, 2004; Choudhary & Nagaraja 2017a; Altman & Bland, 1983) και η συνεισφορά τους στην αξιολόγηση και στον θετικό ή αρνητικό προσδιορισμό της χρήσης των δύο συσκευών εναλλακτικά χωρίς να επηρεάζει κάποιο αποτέλεσμα. Ο Altman και Bland (1983) αναφέρουν δύο στόχους των ερευνών συμφωνίας και ομοιότητας. Πρώτον, η αξιολόγηση της αναπαραγωγής των μετρήσεων των μεθόδων/συσκευών σε πανομοιότυπες συνθήκες και κατά πόσο συμφωνούν. Δεύτερον, κατά πόσο οι δύο μέθοδοι/συσκευές μετράνε την ίδια ποσότητα κατά μέσο όρο και πόση παραλλακτικότητα εξηγείται από τους προκαθορισμένους παράγοντες. Σε περίπτωση μη-συμφωνίας, ερευνώνται οι πηγές ασυμφωνίας. Τέλος, αναφέρονται λόγοι για την ανάπτυξη ερευνών για την αξιολόγηση της συμφωνίας και αξιοπιστίας συσκευών ανοιχτού υλικού και λογισμικού σε σχέση με αντίστοιχες συσκευές (αναφοράς) όπως το υψηλό κόστος των συσκευών αναφοράς (Luiz et al., 2003), η ευκολία χρήσης συσκευών ανοιχτού κώδικα και υλικού αντί της αντίστοιχης βιομηχανικής (Bland & Altman, 1986), η βελτιστοποίηση στον τρόπο μέτρησης και η χρονοβόρες εφαρμογές των συσκευών αναφοράς.

Το Κεφάλαιο 2 περιγράφει χρήσιμους ορισμούς και μεθοδολογίες που χρησιμοποιήθηκαν για την αξιολόγηση της συμφωνίας και ομοιότητας μεταξύ συσκευών ανοιχτού κώδικα και των αντίστοιχων εμπορικών. Συγκεκριμένα, ορίζει το συστηματικό σφάλμα (bias), που αποτελείται από το σταθερό και το

αναλογικό σφάλμα (International Organization of Standardization, 2023). Την εγκυρότητα (accuracy) των μετρήσεων μίας μεθόδου σύμφωνα με τον φορέα US Food Drug Administration (FDA) (1999) που διαφέρει από αυτόν του φορέα ISO όπως και την εγκυρότητα συνάρτηση των μετρήσεων αισθητήρων. Επιπλέον, ορίζεται η ακρίβεια (precision) των μετρήσεων μίας μεθόδου, σύμφωνα με τον φορέα FDA (1999), την ευαισθησία (sensitivity) μίας μεθόδου (National Instruments Corp., 2023), την αξιοπιστία (reliability) μετρήσεων μίας μεθόδου (Choudhary & Nagaraja, 2017a), την ανάλυση (resolution) του αισθητήρα και την επαναληψιμότητα των μετρήσεων μίας μεθόδου (ISO, 2023). Επιπλέον, παρέχει μία εκτεταμένη συζήτηση για τη χρήση διαφορετικών μοντέλων ανάλογα με τα δεδομένα που παράγονται από συσκευές υπό διερεύνηση και την αξιολόγηση δεικτών και γραφημάτων που ερμηνεύονται εύκολα ακόμα και από άτομα χωρίς ιδιαίτερη εμπειρία στη στατιστική. Συγκεκριμένα, παρουσιάζεται το γραμμικό μοντέλο (Fleiss, 1999; Dunn, 2004), το μοντέλο μετρήσεων-σφαλμάτων (measurement-error model) (Buonacorsi, 2010; Choudhary & Nagaraja, 2017a) και το μοντέλο μεικτών-επιδράσεων (Gelman & Hill, 2006; Choudhary & Nagaraja, 2017a). Επιπλέον, παρουσιάζεται η έννοια των επαναλαμβανόμενων μετρήσεων οι οποίες διακρίνονται σε τρεις κατηγορίες: τις μη-συνδεδεμένες, τις συνδεδεμένες και τις χρονικές (Cartensen et al., 2008; Fitzmaurice et al., 2011). Ανάλογα με τον πειραματικό σχεδιασμό της έρευνας συμφωνίας και ομοιότητας χρησιμοποιούνται μία από τις τρεις κατηγορίες επαναλαμβανόμενων μετρήσεων ή σε περίπτωση μη-επαναλαμβανόμενων μετρήσεων χρησιμοποιούνται μετρήσεις κατά ζεύγη (paired). Ανάλογα με την επιλογή των μετρήσεων γίνεται χρήση συγκεκριμένου μοντέλου. Σε περίπτωση που το αναλογικό σφάλμα διαφέρει σημαντικά από τη μονάδα, τα μοντέλα μεικτών-επιδράσεων δε μπορούν να χρησιμοποιηθούν. Επιπλέον, μπορούν να εισαχθούν συμπαράγοντες στο μοντέλο, έτσι ώστε να ανιχνευτούν πιθανές πηγές ασυμφωνίας. Σε περίπτωση που παρατηρείται ετεροσκεδαστικότητα των σφαλμάτων, χρησιμοποιούμε κατάλληλα εργαλεία για την αντιμετώπιση της χρησιμοποιώντας μοντελοποίηση.

Μετά τη μοντελοποίηση, μπορούν να παραχθούν γραφήματα που αξιολογούν τη συμφωνία των δύο μεθόδων/συσκευών. Συγκεκριμένα, το πρώτο γράφημα που περιγράφεται είναι το Bland-Altman και τα 95% όρια συμφωνίας (Bland & Altman, 1983). Παρουσιάζονται οι προϋποθέσεις για την εφαρμογή του και μέθοδοι αντιμετώπισης περιπτώσεων όπου δε πληρούνται (Bland & Altman, 1999). Επιπλέον, περιγράφονται παραμετρικές και μη-παραμετρικές μέθοδοι παραγωγής των ορίων συμφωνίας και μέθοδοι συμπερασματολογίας. Τέλος, αναφέρονται μειονεκτήματα του γραφήματος Bland-Altman και των 95% ορίων συμφωνίας όπως και εναλλακτικές μεθοδολογίες. Η δεύτερη γραφική μέθοδος αφορά το γράφημα πιθανότητας συμφωνίας (Probability of Agreement plot) (Stevens et al., 2017) που προτείνεται ως εναλλακτικό γράφημα αντί του Bland-Altman. Η τρίτη γραφική μέθοδος αφορά το γράφημα σφάλματος και ακρίβειας (Bias and Precision Plot), το διάγραμμα ακρίβειας (Precision plot) και το γράφημα σύγκρισης (Compare plot) που προτείνονται από τον Taffé (2018). Οι τελευταίες γραφικές προσεγγίσεις αφορούν τη μέθοδο Passing – Bablok (Passing & Bablok, 1983, 1984; Bablok et al., 1988) και παλινδρόμηση Demming (Cornbleet & Gochman, 1979).

Το επόμενο βήμα μετά τις γραφικές μεθόδους αξιολόγησης της συμφωνίας είναι η παραγωγή δεικτών συμφωνίας με τη χρήση των παραμέτρων του μοντέλου. Οι δείκτες χωρίζονται σε δύο κατηγορίες, στους απόλυτους (Absolute) και τους σχετικούς δείκτες (Relative). Οι απόλυτοι δείκτες που αναφέρονται στην παρούσα διατριβή είναι της Μέσης Τετραγωνικής Απόκλισης (Mean Squared Deviation, MSD) και της Ολικής Απόκλισης (Total Deviation Index, TDI) (Lin, 2000). Οι σχετικοί δείκτες που αναφέρονται στην παρούσα διατριβή είναι ο Συντελεστής Συσχέτισης Συμφωνίας (Concordance Correlation Coefficient, CCC) (Lin, 1989) και ο Συντελεστής Ατομικής Συμφωνίας (Coefficient of Individual Agreement, CIA) που περιγράφεται από τον Barnhart et al. (2007). Για τον TDI και CCC περιγράφονται παραμετρικοί και μη-παραμετρικοί μέθοδοι εκτίμησης τους, όπως και μέθοδοι συμπερασματολογίας. Μετά τον υπολογισμό των δεικτών συμφωνίας, αξιολογείται και η ομοιότητα των μετρήσεων των δύο μεθόδων με την παραγωγή δεικτών ομοιότητας όπως το σταθερό σφάλμα και το κλάσμα ακρίβειας (Precision Ratio). Τέλος, αξιολογείται η επαναληψιμότητα κάθε συσκευής ξεχωριστά με την παραγωγή δεικτών TDI και CCC

όπως και 95% ορίων συμφωνίας. Έτσι αξιολογείται κατά πόσο συμφωνούν οι μετρήσεις μεταξύ τους όταν εκτελούνται σε πανομοιότυπες συνθήκες. Σε περίπτωση που το συμπέρασμα της έρευνας καταλήγει σε μη-αποδεκτή συμφωνία, πιθανές πηγές ασυμφωνίας μπορούν να εντοπιστούν και μέθοδοι διόρθωσης της συμφωνίας μπορούν να εφαρμοστούν εξετάζοντας την ομοιότητα των δύο συσκευών καθώς και την επαναληψιμότητα τους. Ακόμα, αναφέρονται τα πλεονεκτήματα και μειονεκτήματα κάθε προσέγγισης για την αξιολόγηση της συμφωνίας και ομοιότητας συγκρίνοντας κάθε δείκτη και γράφημα.

Περαιτέρω παρουσιάζεται ο πειραματικός σχεδιασμός και η στατιστική ανάλυση που ακολουθείται ανάλογα με τη φύση της έρευνας καθώς και τα βήματα για τον έγκυρο σχεδιασμό της (Lazic, 2016; Choudhary & Nagaraja, 2017a).

Επιπλέον παρουσιάζονται λανθασμένες ή ελλείψεις μέθοδοι αξιολόγησης της συμφωνίας και ομοιότητας μεταξύ δύο συσκευών. Συγκεκριμένα, το t-test κατά ζεύγη, η μέθοδος ελαχίστων τετραγώνων και ο συντελεστής συσχέτισης Pearson καθώς και παραδείγματα δημοσιεύσεων που τους εφαρμόζουν.

Επιπλέον, περιγράφονται συνοπτικά το υλικό και οι εφαρμογές που χρησιμοποιήθηκαν για την υλοποίηση των πειραμάτων. Αρχικά περιγράφονται οι λειτουργίες των μικροελεγκτών και μικροεπεξεργαστών. Μετά αναφέρονται οι βασικές λειτουργίες των πλακετών ανάπτυξης που χρησιμοποιήθηκαν στα πειράματα της διατριβής. Αργότερα αναφέρεται συνοπτικά η διεπαφή που χρησιμοποιήθηκε για τον προγραμματισμό των πλακετών (Arduino IDE). Επιπλέον, αναλύονται τα πρωτόκολλα επικοινωνίας των μικροελεγκτών, όπως I2C, 1-Wire, UART, SPI όπως και τα πλεονεκτήματα και μειονεκτήματα τους. Ακόμα περιγράφονται τα ασύρματα πρωτόκολλα επικοινωνίας όπως το Wi-Fi, LoRa, GSM, 4g-LTE και πλατφόρμες IOT όπως το Thingspeak (MathWorks Inc, 2022), Adafruit IO και The Things Network (TTN). Τέλος αναφέρονται εφαρμογές σχεδιασμού πλακετών ανοιχτού κώδικα, όπως το Fritzing (Knorig et al., 2009) και Kicad (Evans, 2021) και υλικά που χρησιμοποιήθηκαν για την κατασκευή καταγραφών.

Στο Κεφάλαιο 3 παρουσιάζεται ο σχεδιασμός και η ανάπτυξη ενός ασύρματου μετεωρολογικού σταθμού ανοιχτού κώδικα και υλικού που μετρά τη θερμοκρασία και υγρασία του περιβάλλοντος. Περιγράφεται το κομμάτι του λογισμικού αναφέροντας τις βιβλιοθήκες που χρησιμοποιήθηκαν για τον προγραμματισμό της πλακέτας ανάπτυξης όπως και το διάγραμμα λειτουργιών του κώδικα και το κομμάτι των υλικών για την ανάπτυξη της συσκευής μαζί με το αναλυτικό κόστος. Αρχικά, αξιολογείται η λειτουργία του σταθμού παρακολουθώντας τις μετρήσεις μέσω ασύρματης τεχνολογίας Wi-Fi για επτά μέρες. Έπειτα, σχεδιάζεται πείραμα αξιολόγησης συμφωνίας και ομοιότητας του σταθμού με αντίστοιχο εμπορικό με αισθητήρα Thygro και τα αντίστοιχα περιφερειακά από την εταιρεία Symmetron για τη θερμοκρασία αέρα. Για το πείραμα κατασκευάστηκαν τριάντα συσκευές ανοιχτού κώδικα και υλικού και χρησιμοποιήθηκαν τρεις αντίστοιχοι εμπορικοί σταθμοί. Για τον λόγο αυτό δημιουργήθηκαν τρεις ομάδες. Κάθε ομάδα περιέχει 10 σταθμούς ανοιχτού κώδικα και υλικού και μία εμπορική. Κάθε πειραματική μονάδα παράγει ένα σύμπλεγμα μετρήσεων σε συνάρτηση του χρόνου. Το πείραμα τερματίστηκε μετά από επτά μέρες συνεχόμενων μετρήσεων. Καθώς οι μετρήσεις ανά πειραματική μονάδα ήταν πάνω από 220 για κάποιες μέρες (αφού υπήρχαν απώλειες δεδομένων λόγω κακής ποιότητας του δικτύου) οι μετρήσεις συγχωνεύτηκαν σε χρονικές ζώνες, με τη μέθοδο που χρησιμοποιεί το Εθνικό Αστεροσκοπείο Αθénas. Αναφέρεται το πλαίσιο του πειραματικού σχεδιασμού όπου τα δεδομένα θεωρούνται χρονικά, ο αριθμός δείγματος είναι 30 (ισορροπημένος), ως πειραματική μονάδα θεωρείτε ο μετεωρολογικός σταθμός, ο αριθμός των επαναλαμβανόμενων τιμών είναι 51 και θεωρούνται ισορροπημένες. Έπειτα υλοποιείται περιγραφική στατιστική όπως και συμπερασματολογία με τη χρήση τριών μεθοδολογιών. Η πρώτη αφορά τη μεθοδολογία που αναπτύχθηκε από τον Choudhary και Nagaraja (2017a) όπου παράγονται μοντέλα μεικτών-επιδράσεων για την παραγωγή δεικτών όπως ο TDI, CCC και 95% διαστημάτων συμφωνίας για την αξιολόγηση της συμφωνίας των δύο συσκευών, το συστηματικό σφάλμα και το κλάσμα ακρίβειας για την αξιολόγηση της ομοιότητας τους όπως και γραφήματα Bland-Altman. Η δεύτερη μεθοδολογία (Carrasco et al., 2009, 2013) αφορά την παραγωγή

μοντέλου μεικτών-επιδράσεων για την εκτίμηση του δείκτη CCC. Η τρίτη μεθοδολογία (Oliviera et al., 2018) αφορά την παραγωγή μοντέλου μεικτών-επιδράσεων για την εκτίμηση χρονικού δείκτη συσχέτισης συμφωνίας (Longitudinal Concordance Correlation) και των παραγόντων του. Τα σφάλματα των παραμέτρων του μοντέλου της πρώτης μεθόδου ήταν σημαντικά μεγάλα και η σύγκλιση του αλγορίθμου εκτίμησης των δεικτών απέτυχε. Για το λόγο αυτό εφαρμόστηκαν οι άλλες δύο μέθοδοι. Συγκεκριμένα η τιμή του δείκτη της δεύτερης μεθοδολογίας CCC είναι 0,3346 με διαστήματα εμπιστοσύνης [0,2642 , 0,4016]. Η τρίτη μεθοδολογία εκτιμά το χρονικό δείκτη συσχέτισης για κάθε χρονική στιγμή με διάστημα [0,1146 , 0,2236] ξεκινώντας με την υψηλότερη τιμή και καταλήγοντας με τη μικρότερη. Ο δείκτης χρονικής ακρίβειας (Longitudinal Precision) παρουσιάζει παρόμοια συμπεριφορά. Αντίθετα με τους προηγούμενους δείκτες, ο δείκτης χρονικής εγκυρότητας (Longitudinal Accuracy) παρουσιάζει υψηλές τιμές κοντά στο 0,9. Η συμφωνία μεταξύ των συσκευών θεωρήθηκε ανεπαρκής και ως πιθανή πηγή ασυμφωνίας θεωρήθηκε η ακρίβεια των αισθητήρων λόγω της διαφοράς του δείκτη χρονικής ακρίβειας και εγκυρότητας. Περεταίρω έρευνα μπορεί να υλοποιηθεί για να συμπεριληφθούν όλα τα σημεία των μετρήσεων όπως και επιπλέον παράγοντες που πιθανόν να επηρεάζουν τη συμφωνία, όπως η υγρασία του αέρα ή η τοποθεσία κάθε σταθμού. Τέλος, εναλλακτικές μορφές ασύρματης επικοινωνίας όπως η LoRa μπορούν να προστεθούν στο σταθμό όπως και επιπλέον αισθητήρες. Τα πλεονεκτήματα χρήσης της νέας συσκευής στηρίζονται στο χαμηλό κόστος παραγωγής του όπως και στην εύκολη μορφοποίηση και προσθήκη επιπλέον αισθητήρων.

Στο Κεφάλαιο 4 παρουσιάζεται ο σχεδιασμός και η κατασκευή ενός ασύρματου καταγραφέα LoRa ανοιχτού κώδικα και υλικού για τη χρήση σε πειράματα ενυδριοπονίας σε σάκους υποστρώματος περλίτη. Περιγράφεται ο σχεδιασμός της συσκευής, το υλικό όπως και το λογισμικό που χρησιμοποιήθηκε. Συγκεκριμένα, η νέα συσκευή είναι συμβατή με αισθητήρες περιεκτικότητας νερού, με πρωτόκολλο επικοινωνίας RS-485. Ο αισθητήρας που χρησιμοποιήθηκε για το συγκεκριμένο πείραμα μετρά επιπλέον τη θερμοκρασία του σάκου υποστρώματος, την ηλεκτρική αγωγιμότητα καθώς και το pH. Η συσκευή μπορεί να αποστείλει τα δεδομένα σε διακομιστή IOT με τη χρήση ασύρματης τεχνολογίας LoRa. Μετά τον έλεγχο καλής λειτουργίας του καταγραφέα, σχεδιάστηκε το πείραμα αξιολόγησης συμφωνίας και ομοιότητας της νέας συσκευής με αντίστοιχη εμπορική. Η εμπορική συσκευή που χρησιμοποιήθηκε για το πείραμα είναι η Handheld Meter (HH150) και SM150 Soil Moisture Sensor. Επιπλέον, αναφέρεται το κόστος των δύο συσκευών. Συγκεκριμένα η συσκευή ανοιχτού υλικού και κώδικα κόστισε 130 ευρώ συμπεριλαμβάνοντας τον αισθητήρα και το κόστος του gateway στα 250 ευρώ περίπου. Το κόστος του εμπορικού αισθητήρα και καταγραφέα ανέρχεται στα 760 ευρώ. Έπειτα, παρουσιάζεται ο πειραματικός σχεδιασμός. Συγκεκριμένα, ως πειραματική μονάδα ορίζεται ο σάκος υποστρώματος, οι επαναλαμβανόμενες μετρήσεις θεωρούνται μη-ισορροπημένες (δέκα για τη συσκευή αναφοράς και τρεις για τη συσκευή ανοιχτού κώδικα και υλικού) με ισορροπημένο μέγεθος δείγματος, 100. Ακολούθησε περιγραφική στατιστική με την κατασκευή θηκογράμματος, Trellis-plot, και τροποποιημένου διαγράμματος διασποράς. Μετά την εκτέλεση πιλοτικού πειράματος, το αναλογικό σφάλμα εκτιμήθηκε με τιμή σημαντικά διαφορετική από το 1. Έτσι, για την ανάλυση χρησιμοποιήθηκε μοντέλο μέτρησης-σφαλμάτων με τη χρήση της μεθοδολογίας του Taffé (2018) για την εκτίμηση της διαφοράς στις μετρήσεις σε σχέση με τον καλύτερο γραμμικό αμερόληπτο παράγοντα πρόβλεψης της πραγματικής τιμής της μέτρησης που παράχθηκε από την εμπορική συσκευή. Τα αποτελέσματα ανέφεραν σημαντική απόκλιση από την εκτίμηση της πραγματικής τιμής των μετρήσεων παραθέτοντας το σταθερό (11,669) και αναλογικό σφάλμα (0,46). Η συμφωνία μεταξύ των δύο οργάνων θεωρήθηκε ικανοποιητική για συγκεκριμένα διαστήματα μετρήσεων όπως από [25, 35]% όπου το μέγιστο σφάλμα φτάνει στο ποσοστό 5%. Πιθανές πηγές ασυμφωνίας είναι η ακρίβεια του συστήματος ανοιχτού κώδικα και υλικού που είναι από 5 έως 6 περίπου φορές χειρότερη συγκριτικά με αυτή του εμπορικού μετά την επαναβαθμονόμηση όπως και το σταθερό σφάλμα. Τα πλεονεκτήματα χρήσης της νέας συσκευής συγκριτικά με της εμπορικής είναι το χαμηλό κόστος παραγωγής και η ευκολία μορφοποίησης της

ανάλογα με τις ανάγκες των πειραμάτων. Ο σάκος υποστρώματος περλίτη έχει αρκετή παραλλακτικότητα ως προς την περιεκτικότητα του σε νερό. Επομένως η χρήση πολλαπλών αισθητήρων σε διαφορετικά σημεία του σάκου είναι θεμελιώδης και αποτελεί βασική απαίτηση ως εισροή πληροφορίας σε αλγόριθμους πρόβλεψης περιεκτικότητας νερού στο σάκο.

Στο Κεφάλαιο 5 παρουσιάζεται ο σχεδιασμός και η κατασκευή καταγραφέα ανοιχτού κώδικα και υλικού που μετρά pH διαλυμάτων για τη χρήση σε πειράματα Δενδροκομίας. Η συσκευή ανοιχτού κώδικα και υλικού αποτελείται από δύο αισθητήρες. Πρώτον, τον αισθητήρα μέτρησης pH DFRobot SEN0169 ο οποίος είναι συμβατός με πλακέτες ανάπτυξης και ταυτόχρονα έχει κατασκευαστεί με βιομηχανικά πρότυπα. Η έξοδος του αισθητήρα είναι αναλογική (τάση). Για το λόγο αυτό χρησιμοποιήθηκε ADC με 15 bit ανάλυσης έναντι των 10 bit της πλακέτας ανάπτυξης. Έτσι, οι μετρήσεις πραγματοποιήθηκαν με βελτιστοποιημένη ανάλυση. Δεύτερον, στη συσκευή προστέθηκε αδιάβροχος αισθητήρας θερμοκρασίας για την εξισορρόπηση της τιμής pH. Η συσκευή έχει λειτουργίες καταγραφής των αποτελεσμάτων σε κάρτα microSD όπως και χρονοσφραγίδα. Για το λογισμικό, δε χρησιμοποιήθηκε έτοιμη βιβλιοθήκη, καθώς δεν ήταν συμβατή με την πλακέτα ανάπτυξης. Το κόστος της συσκευής ανοιχτού κώδικα και υλικού ανέρχεται στα 130 ευρώ περίπου. Για τη βαθμονόμηση των δύο συσκευών χρησιμοποιήθηκαν ρυθμιστικά διαλύματα pH τιμών 4,01 και 7,01. Έπειτα αξιολογήθηκε η καλή λειτουργία της συσκευής μετρώντας το pH καρπών εσπεριδοειδών. Επιπλέον, σχεδιάστηκε μελέτη αξιολόγησης της συμφωνίας και ομοιότητας της νέας συσκευής με αντίστοιχη εμπορική. Συγκεκριμένα, η εμπορική συσκευή που επιλέχθηκε είναι η Hanna HI9024 Waterproof pHMeter. Η εμπορική συσκευή έχει εξαντληθεί και η τιμή της αντίστοιχης σύγχρονης συσκευής ανέρχεται στα 585 ευρώ. Περιγράφεται ο πειραματικός σχεδιασμός, όπου ως πειραματική μονάδα θεωρούμε το χυμό κάθε εσπεριδοειδούς, το μέγεθος δείγματος (σύνολο 30) θεωρείται ισορροπημένο, οι επαναλαμβανόμενες μετρήσεις (σύνολο 10) θεωρούνται ισορροπημένες και θεωρήθηκαν πιθανοί συμπαράγοντες όπως η θερμοκρασία του υγρού, η ποσότητα του δείγματος και η σειρά μέτρησης του οργάνου. Τέλος, ο τύπος των επαναλαμβανόμενων μετρήσεων θεωρείται συνδεδεμένος. Για την ανάλυση χρησιμοποιήθηκαν τρεις μεθοδολογίες με την παραγωγή μοντέλων μεικτών-επιδράσεων. Αρχικά εφαρμόστηκε η μεθοδολογία του Choudhary και Nagaraja (2017a) για την υλοποίηση της περιγραφικής στατιστικής, την παραγωγή δεικτών CCC και TDI, των 95% ορίων συμφωνίας, της αξιολόγησης της ομοιότητας αλλά και επαναληψιμότητας κάθε συσκευής. Η δεύτερη μεθοδολογία βασίζεται στον Carrasco et al. (2013) για την εκτίμηση του δείκτη CCC. Η τρίτη μεθοδολογία βασίζεται στην Escaramis et al. (2010a) για την εκτίμηση του δείκτη TDI. Αρχικά για την υλοποίηση της περιγραφικής στατιστικής παράχθηκαν πέντε γραφήματα. Πρώτον, ένα τροποποιημένο διάγραμμα διασποράς, όπου κάθε υποκείμενο παρουσιάζεται με τον αύξοντα αριθμό του αντί για σημείο. Δεύτερον, το διάγραμμα Bland-Altman χωρίς τα όρια συμφωνίας. Τρίτον, το Trellis-plot. Τέταρτων, το διάγραμμα αλληλεπιδράσεων για τους παράγοντες συσκευή και υποκείμενο. Πέμπτων, το διάγραμμα αλληλεπιδράσεων για τους παράγοντες χρονική στιγμή και υποκείμενο. Στα τρία πρώτα γραφήματα φαίνεται ότι το αναλογικό σφάλμα δε διαφέρει σημαντικά από τη μονάδα ενώ παρατηρείται απόκλιση σταθερού μεγέθους για το μεγαλύτερο αριθμό των μετρήσεων. Δε παρατηρείται σημαντική ετεροσκεδαστικότητα. Για το τέταρτο και πέμπτο διάγραμμα παρατηρούνται σημαντικές αλληλεπιδράσεις μεταξύ των παραγόντων συσκευή και υποκείμενο αλλά όχι μεταξύ των παραγόντων χρονική στιγμή και υποκείμενο. Για την αξιολόγηση της συμφωνίας παράγεται το εκτεταμένο διάγραμμα Bland-Altman με τα αντίστοιχα 95% όρια συμφωνίας όπου είναι εμφανές το σταθερό σφάλμα, ενώ το αναλογικό σφάλμα δε θεωρείται σημαντικά διαφορετικό από τη μονάδα. Οι γραμμές παλινδρόμησης για τα όρια συμφωνίας παρουσιάζουν πιθανή ετεροσκεδαστικότητα των σφαλμάτων. Αρχικά, για την παραγωγή των δεικτών CCC και TDI δεδομένου της φύσης των δεδομένων, κατασκευάστηκε μοντέλο μεικτών-επιδράσεων για συνδεδεμένες επαναλαμβανόμενες τιμές. Δεδομένου της μη σημαντικής αλληλεπίδρασης των χρονικών στιγμών με τα υποκείμενα (γράφημα αλληλεπιδράσεων) και της αποτυχίας σύγκλισης του αλγορίθμου για την εκτίμηση των παραμέτρων, επιλέχθηκε το μοντέλο για μη-

συνδεδεμένες επαναλαμβανόμενες τιμές. Έπειτα, κατασκευάστηκαν δύο μοντέλα. Το πρώτο μοντέλο συμπεριλαμβάνει την μεθοδολογία για την επίδραση της ετεροσκεδαστικότητας σε αντίθεση με το δεύτερο. Για να επιλεχτεί το κατάλληλο μοντέλο, χρησιμοποιήθηκαν τα επίσημα κριτήρια AIC, BIC και log-likelihood. Έτσι, επιλέχθηκε το μοντέλο που συμπεριλαμβάνει τη μεθοδολογία για τη αντιμετώπιση της επίδρασης της ετεροσκεδαστικότητας. Επιπλέον, κατασκευάστηκε μοντέλο που συμπεριλαμβάνει τους συμπαράγοντες θερμοκρασία του υγρού, ποσότητα του δείγματος και σειρά μέτρησης του οργάνου. Τα μοντέλα με συμπαράγοντες και χωρίς συμπαράγοντες συγκρίθηκαν με τη χρήση επίσημων κριτηρίων AIC, BIC και log-likelihood και επιλέχθηκε το μοντέλο χωρίς συμπαράγοντες. Αργότερα εκτιμήθηκαν οι δείκτες CCC, TDI, 95% όρια συμφωνίας, σταθερό σφάλμα και κλάσμα ακρίβειας, με τη χρήση των παραμέτρων του τελικού μοντέλου και χρήσιμα γραφήματα υλοποιώντας παραπάνω από μία μεθοδολογίες. Επιπλέον, αξιολογήθηκε η επαναληψιμότητα κάθε οργάνου ξεχωριστά. Για τη συμφωνία, ο δείκτης CCC εκτιμήθηκε σε όλο το εύρος των μετρήσεων [0.5970, 0.6032] με κατώτερο όριο εμπιστοσύνης [0.4776, 0.4839]. Ο δείκτης TDI (0.9) εκτιμήθηκε σε όλο το εύρος των μετρήσεων [0.2883, 0.3031] με ανώτερο όριο εμπιστοσύνης [0.3095, 0.3232]. Τα όρια εμπιστοσύνης εκτιμήθηκαν σε όλο το εύρος των μετρήσεων. Το διάστημα για το κάτω όριο εμπιστοσύνης βρίσκεται μεταξύ των τιμών [-0.3464, -0.3237] και το άνω όριο εμπιστοσύνης βρίσκεται μεταξύ των τιμών [-0.1193, -0.0966]. Η συμφωνία με βάση τους δείκτες δε θεωρείται αποδεκτή. Για την αξιολόγηση της ομοιότητας των δύο συσκευών εκτιμήθηκε το κλάσμα ακρίβειας και το σταθερό σφάλμα σε όλο το εύρος των μετρήσεων. Οι τιμές του κλάσματος ακρίβειας βρίσκονται μεταξύ των τιμών [1.7987, 2.8322]. Έτσι το σύστημα ανοιχτού κώδικα θεωρείται 2 με τρεις περίπου φορές πιο ακριβές συγκριτικά με τη συσκευή αναφοράς. Η εκτίμηση του σταθερού σφάλματος είναι -0.2215. Το σταθερό σφάλμα και η μικρότερη ακρίβεια της συσκευής αναφοράς συγκριτικά με τη συσκευή ανοιχτού κώδικα θεωρούνται πιθανές πηγές ασυμφωνίας. Αργότερα, αξιολογήθηκε η επαναληψιμότητα εκτιμώντας τους δείκτες CCC, TDI και όρια εμπιστοσύνης για κάθε συσκευή ξεχωριστά. Για τη συσκευή ανοιχτού κώδικα η εκτίμηση της τιμής του CCC βρίσκεται μεταξύ των τιμών [0.9903, 0.9985] με κατώτερο όριο εμπιστοσύνης [0.9973, 0.9982]. Η εκτίμηση της τιμής του TDI βρίσκεται μεταξύ των τιμών [0.0181, 0.0462] με ανώτερο όριο εμπιστοσύνης [0.0213, 0.0563]. Η τιμή της εκτίμησης των ορίων εμπιστοσύνης βρίσκονται μεταξύ των τιμών [-0.0550, -0.0216] για το κάτω όριο και [0.0216, 0.0550] για το άνω όριο. Για τη συσκευή αναφοράς η εκτίμηση της τιμής του δείκτη CCC βρίσκεται μεταξύ των τιμών [0.9732, 0.9973] με κατώτερο όριο εμπιστοσύνης [0.9500, 0.9951]. Η εκτίμηση της τιμής του TDI βρίσκεται μεταξύ των τιμών [0.0243, 0.0777] με ανώτερο όριο εμπιστοσύνης [0.0286, 0.0949]. Η τιμή της εκτίμησης των ορίων εμπιστοσύνης βρίσκονται μεταξύ των τιμών [-0.0926, -0.0289] για το κάτω όριο και [0.0289, 0.0926] για το άνω όριο. Η επαναληψιμότητα των δύο συσκευών θεωρείται άριστη. Παράχθηκαν γραφήματα που παρουσιάζουν το εύρος των δεικτών για τη συμφωνία των δύο συσκευών και για τη συμφωνία κάθε συσκευής ξεχωριστά. Επιπλέον, κατασκευάστηκε το εκτεταμένο Bland-Altman διάγραμμα και τα 95% όρια εμπιστοσύνης, το διάγραμμα Bland-Altman και τα 95% όρια συμφωνίας με τα αντίστοιχα 95% διαστήματα εμπιστοσύνης και το Bias Plot. Όλα τα διαγράμματα παρουσιάζουν το σταθερό σφάλμα ως πηγή ασυμφωνίας. Το σταθερό και αναλογικό σφάλμα που παράχθηκαν από το μοντέλο μετρήσεων-σφαλμάτων συμφωνούν με τα ευρήματα των δεικτών ομοιότητας που εκτιμήθηκαν από το μοντέλο μεικτών-επιδράσεων. Μετά τη χρήση μεθόδων αναβαθμονόμησης η συμφωνία θεωρείται άριστη. Η συσκευή είναι φορητή και μπορεί να χρησιμοποιηθεί για “επι τόπου” μετρήσεις στον αγρό ή στο θερμοκήπιο. Τα πλεονεκτήματα χρήσης της συσκευής είναι το χαμηλό κόστος παραγωγής, περίπου 4 φορές μικρότερο συγκριτικά με το Hanna HI9024. Περαιτέρω έρευνα μπορεί να υλοποιηθεί για τη χρήση της συσκευής για τη μέτρηση pH εδάφους ή θρεπτικών υποστρωμάτων. Τέλος, επιπλέον αισθητήρες μπορούν να τοποθετηθούν εύκολα στη συσκευή για την ταυτόχρονη άντληση περισσότερων πληροφοριών για τα χαρακτηριστικά του δείγματος.

Στο Κεφάλαιο 6 περιγράφεται η κατασκευή συσκευής ανοιχτού κώδικα και υλικού για τις ανάγκες πειραμάτων που αφορούν τη μύγα της Μεσογείου (*Ceratitis Capitata*). Η συσκευή καταγράφει τη θερμοκρασία στην επιφάνεια και τον πυρήνα καρπών όπως και την εξωτερική θερμοκρασία αέρα. Οι αισθητήρες που χρησιμοποιήθηκαν για την κατασκευή της συσκευής είναι δύο ιατρικοί σημειακοί θερμίστορες ακριβείας με πολύ μικρή διάμετρο (1,65 mm μέγιστη διάμετρος) και ένας θερμίστορας με αλουμινένιο κυλινδρικό περίβλημα. Επιπλέον, χρησιμοποιήθηκε η ηλεκτρονική πλακέτα ADS1115 για τη βελτίωση της ανάλυσης του σήματος εισόδου, από 10 σε 15 bit όπως και η πλακέτα LM4040 για την εξισορρόπηση της τάσης αναφοράς των θερμιστόρων. Τέλος, η συσκευή διαθέτει σύστημα καταγραφείας που αποθηκεύει τα δεδομένα σε κάρτα microSD και ρολόι για την εισαγωγή χρονοσφραγίδας για κάθε μέτρηση. Η μπαταρία που χρησιμοποιήθηκε είναι τύπου LiPo με χωρητικότητα 1200 mAh. Για τη βαθμονόμηση των θερμιστόρων χρησιμοποιήθηκε ο τύπος Hoge-2 και οι καμπύλες σχέσης αντίστασης-θερμοκρασίας του κατασκευαστή. Η κατανάλωση της συσκευής είναι στα 3.148 mA κατά μέσο όρο. Το κόστος της συσκευής ανέρχεται στα 110 ευρώ περίπου (τιμές αγοράς τέλη 2020). Αρχικά, αξιολογήθηκε η λειτουργία της στο εργαστήριο αλλά και στον αγρό. Για τα πειράματα στον αγρό η συσκευή τοποθετήθηκε σε συγκεκριμένα σημεία δέντρων. Συγκεκριμένα, οι συσκευές τοποθετήθηκαν στην κορυφή δέντρου, στην ανατολική πλευρά του δέντρου και σε χαμηλό σκιασμένο σημείο του δέντρου. Η συχνότητα καταγραφής των μετρήσεων ήταν 15 λεπτά για 24 ώρες και αργότερα επιλέγονταν διαφορετικά δέντρα για την τοποθέτησή τους. Μετά τις 24 ώρες καταγράφηκαν τα ποιοτικά χαρακτηριστικά του καρπού. Το σύνολο των δέντρων όπου εκτελέστηκε η καταγραφή ήταν 15. Στο εργαστήριο σχεδιάστηκε έρευνα συμφωνίας και αξιοπιστίας για τη σύγκριση των μετρήσεων στον πυρήνα και στην επιφάνεια μήλων σε οχτώ διαφορετικές συνθήκες θερμοκρασίας. Συγκεκριμένα 10 μήλα τοποθετήθηκαν σε εμπορικό ψυγείο (PITSOS P1KCL3606D) σε θερμοκρασία 2°C και 10 μήλα σε θερμοκρασία 5°C. Αργότερα 10 μήλα τοποθετήθηκαν μέσα σε θάλαμο (Elvem CLP 600) σε σταθερή θερμοκρασία 15°C, 10 μήλα σε δωμάτιο ελεγχόμενης θερμοκρασίας 20°C και 10 μήλα σε δωμάτιο ελεγχόμενης θερμοκρασίας 25°C. Τέλος, για τις θερμοκρασίες 27°C, 34°C και 43°C ίδιος αριθμός μήλων τοποθετήθηκε σε φούρνο (WTC binder 78,532). Ο αριθμός του δείγματος είναι 80 και το δείγμα είναι ισορροπημένο (10 μήλα ανά συνθήκη θερμοκρασίας), ως πειραματική μονάδα θεωρήθηκε το μήλο και οι επαναλαμβανόμενες τιμές είναι συνδεδεμένες και ισορροπημένες με αριθμό 15. Για την αξιολόγηση της συμφωνίας των μετρήσεων θερμοκρασίας των δύο σημείων χρησιμοποιήθηκαν μοντέλα μέτρησης-σφαλμάτων. Τα αποτελέσματα αναφέρουν σταθερή απόκλιση 0,331°C καθώς και αναλογική απόκλιση 0,982. Επιπλέον, η σύγκριση ακρίβειας των μετρήσεων στα δύο σημεία υποδηλώνει ότι ο πυρήνας διατηρεί τη θερμοκρασία περισσότερο συγκριτικά με την επιφάνεια σε διαφορετικές συνθήκες. Η συσκευή είναι φορητή και λειτουργεί με μπαταρία. Η αυτονομία της μπορεί να φτάσει στις 7 έως 30 μέρες ανάλογα με το χρονικό διάστημα καταγραφής. Το σχήμα αλλά και το μέγεθος της συσκευής καθιστά ιδανική την τοποθέτησή της σε οποιοδήποτε σημείο του δέντρου επιθυμεί ο ερευνητής. Τέλος, σε μελλοντικές μελέτες, ασύρματη τεχνολογία LoRa μπορεί να προστεθεί στη συσκευή, ηλιακοί συλλέκτες για περισσότερη αυτονομία καθώς και ρολόι ακριβείας για την καλύτερη λειτουργία των κύκλων αδρανοποίησης της συσκευής. Με τις συγκεκριμένες τροποποιήσεις η συσκευή μπορεί να λειτουργεί για μήνες λόγω των λειτουργιών χαμηλής κατανάλωσης.

Το κεφάλαιο επτά αποτελεί μία γενική συζήτηση για τα προηγούμενα κεφάλαια καθώς και τα ευρήματα κάθε μελέτης. Τέλος αναφέρονται μελλοντικές έρευνες για την τροποποίηση των συσκευών καθώς και για τη βελτίωση των μεθόδων συμφωνίας και ομοιότητας.

Το κεφάλαιο οκτώ συμπεριλαμβάνει τα συμπεράσματα αλλά και τη συνεισφορά και καινοτομία της διατριβής. Συγκεκριμένα αναφέρει την υιοθέτηση της μεθοδολογίας συμφωνίας και ομοιότητας από τον τομέα της ιατρικής στον γεωπονικό τομέα με την πρόταση πρωτοκόλλου που αξιολογεί συσκευές ανοιχτού κώδικα συγκριτικά με αντίστοιχες εμπορικές για εφαρμογές στη γεωπονία. Τέλος, προτείνονται καινοτόμες συσκευές και περιγράφεται αναλυτικά η μεθοδολογία αξιολόγησης της συμφωνίας και ομοιότητας τους με αντίστοιχες εμπορικές σε τέσσερις διαφορετικές εφαρμογές.