

ΒΟΛΟΣ 2020

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΘΕΣΣΑΛΙΑΣ  
ΤΜΗΜΑ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

UNIVERSITY OF THESSALY  
DEPARTMENT OF CIVIL ENGINEERING



## Ανάλυση χωροχρονικών συσχετίσεων σε πεδίο ταχυτήτων τυρβώδους ροής



Χατζημπαλής Ηλίας

Επιβλέπων: Θ. Καρακασίδης, Καθηγητής

## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΡΟΛΟΓΟΣ .....	4
ΕΙΣΑΓΩΓΗ .....	5
<b>ΚΕΦΑΛΑΙΟ 1: ΒΑΣΙΚΟΙ ΣΤΑΤΙΣΤΙΚΟΙ ΔΕΙΚΤΕΣ</b>	
1.1 ΕΙΣΑΓΩΓΗ .....	7
1.2 ΘΕΩΡΗΤΙΚΗ ΑΝΑΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ .....	7
1.3 ΜΕΤΡΗΣΕΙΣ ΤΑΧΥΤΗΤΩΝ ΣΕ ΚΑΝΑΛΙ ΜΕ ΚΑΙ ΧΩΡΙΣ ΤΗ ΔΙΕΝΕΡΓΕΙΑ ΜΑΓΝΗΤΙΚΟΥ ΠΕΔΙΟΥ.....	9
1.4 ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΤΙΜΩΝ .....	9
<b>ΚΕΦΑΛΑΙΟ 2: Ο ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ PEARSON</b>	
2.1 ΓΕΝΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ .....	13
2.2 ΕΦΑΡΜΟΓΗ ΣΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΡΓΑΣΙΑΣ .....	16
2.3 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ-ΠΑΡΑΤΗΡΗΣΕΙΣ .....	20
<b>ΚΕΦΑΛΑΙΟ 3 ΕΥΡΕΣΗ MINIMUM SPANNING TREES</b>	
3.1 ΕΙΣΑΓΩΓΗ .....	22
3.2 ΕΦΑΡΜΟΓΕΣ ΤΩΝ SPANNING TREES .....	23
3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΩΝ SPANNING TREES .....	25
3.4 ΔΗΜΙΟΥΡΓΙΑ MINIMUM SPANNING TREE ΓΙΑ ΤΙΣ ΑΝΑΓΚΕΣ ΤΗΣ ΕΡΓΑΣΙΑΣ .....	29
3.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΕΦΑΛΑΙΟΥ .....	34
<b>ΚΕΦΑΛΑΙΟ 4: ΣΥΝΑΡΤΗΣΗ ΔΙΑΣΥΣΧΕΤΙΣΗΣ</b>	
4.1 ΕΙΣΑΓΩΓΗ .....	37
4.2 ΕΦΑΡΜΟΓΕΣ .....	38
4.3 ΕΦΑΡΜΟΓΗ ΣΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΡΓΑΣΙΑΣ .....	39

<b>4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΕΦΑΛΑΙΟΥ .....</b>	<b>48</b>
<b>ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ .....</b>	<b>50</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ .....</b>	<b>53</b>

Η φωτογραφία στο εξώφυλλο απεικονίζει τον ποταμό Νέστο και ανακτήθηκε από <https://www.xanthinea.gr/2020/02/nestos-o-entyposiakos-potamos-poy-latreytike-san-theos.html> στις 06/09/2020

## ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1: Στρωτή (πάνω) και τυρβώδης (κάτω) ροή .....	5
Εικόνα 2: Ροή καναλιού .....	8
Εικόνα 3: Πίνακες κανονικοποίησης .....	12
Εικόνα 4: Διάγραμμα μηνιαίας κατανάλωσης .....	15
Εικόνα 5: Πίνακες Pearson .....	16
Εικόνα 6: Διάγραμμα ταχύτητας U (με μαγνητικό πεδίο) .....	17
Εικόνα 7: Διάγραμμα ταχύτητας U (χωρίς μαγνητικό πεδίο) .....	17
Εικόνα 8: Διάγραμμα ταχύτητας V (με μαγνητικό πεδίο) .....	19
Εικόνα 9: Διάγραμμα ταχύτητας V (χωρίς μαγνητικό πεδίο) .....	19
Εικόνα 10: Διάγραμμα ταχύτητας W (με μαγνητικό πεδίο) .....	19
Εικόνα 11: Διάγραμμα ταχύτητας W (χωρίς μαγνητικό πεδίο) .....	19
Εικόνα 12: Προφίλ ταχυτήτων .....	20
Εικόνα 13: Spanning tree .....	22
Εικόνα 14: Spanning tree .....	23
Εικόνα 15: Συγκρότημα με οικίες .....	24
Εικόνα 16: Πρώτη προσπάθεια σύνδεσης .....	24
Εικόνα 17: Δεύτερη προσπάθεια σύνδεσης .....	25
Εικόνα 18: Πίνακας στοιχείων .....	27
Εικόνα 19: Πίνακας στοιχείων (επεξεργασία) .....	28
Εικόνα 20: Spanning tree Rum .....	30
Εικόνα 21: Διάγραμμα αποστάσεων U.....	33
Εικόνα 22: Διάγραμμα αποστάσεων V.....	33
Εικόνα 23: Διάγραμμα αποστάσεων W.....	34
Εικόνα 24: Ηλεκτρομαγνητική Αντλία.....	36
Εικόνα 25: Διάγραμμα cross U ( χωρίς μαγνητικό πεδίο) .....	39
Εικόνα 26: Διάγραμμα ti U ( χωρίς μαγνητικό πεδίο) .....	42
Εικόνα 27: Διάγραμμα cross V ( με μαγνητικό πεδίο) .....	42
Εικόνα 28: Διάγραμμα ti V ( με μαγνητικό πεδίο) .....	43
Εικόνα 29: Διάγραμμα cross V (χωρίς μαγνητικό πεδίο) .....	43
Εικόνα 30: Διάγραμμα ti V ( χωρίς μαγνητικό πεδίο) .....	44
Εικόνα 31 :Διάγραμμα cross W ( με μαγνητικό πεδίο) .....	44
Εικόνα 32: Διάγραμμα ti W ( με μαγνητικό πεδίο) .....	45
Εικόνα 33: Διάγραμμα cross W (χωρίς μαγνητικό πεδίο) .....	45
Εικόνα 34: Διάγραμμα cross W ( με μαγνητικό πεδίο).....	46
Εικόνα 35: Διάγραμμα ti W ( με μαγνητικό πεδίο) .....	46
Εικόνα 36: Διάγραμμα cross W (χωρίς μαγνητικό πεδίο) .....	47
Εικόνα 37: Διάγραμμα ti W ( χωρίς μαγνητικό πεδίο) .....	47

## ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1: Υπολογισμός αποστάσεων .....	31
Πίνακας 2: Υπολογισμός αποστάσεων .....	32

## ΠΡΟΛΟΓΟΣ

Η εργασία αυτή αποτελεί μέρος της ολοκλήρωσης των σπουδών μου στο Τμήμα Πολιτικών Μηχανικών της Πολυτεχνικής Σχολής του Πανεπιστημίου Θεσσαλίας. Έμπνευση για το θέμα της υπήρξε η τυρβώδης συμπεριφορά του νερού κάτω από ορισμένες συνθήκες που εμφανίζονται στη φύση αλλά και στο περιβάλλον του εργαστηρίου. Στη συνέχεια παρουσιάζονται τα στοιχεία που χρειάστηκαν για την εν λόγω μελέτη, η επεξεργασία τους, τα αποτελέσματα που προέκυψαν καθώς και πολλά σημαντικά συμπεράσματα για τη συμπεριφορά των ρευστών. Το παρόν κείμενο αποτελεί το βιβλιογραφικό κομμάτι της εργασίας. Το μεγαλύτερο μέρος της επεξεργασίας των δεδομένων έγινε σε περιβάλλον Matlab. Για αισθητικούς αλλά κυρίως και πρακτικούς λόγους οι εντολές και οι μακροσκελείς κώδικες του Matlab δεν έχουν συμπεριληφθεί στο κείμενο αυτό.

Η εκπόνησή της έγινε χάρη στην ουσιαστική συμβολή του κυρίου Αθανάσιου Φράγκου. Η βοήθειά του καθ' όλη την εκπόνηση της εργασίας ήταν καθοριστική για την τελική της έκβαση. Οι διαδρομές των ροών που διάλεξα είναι μακρές και μαζί τους έκανα μια καλή αρχή στην έρευνα. Γι' αυτό και του οφείλω ένα μεγάλο ευχαριστώ.

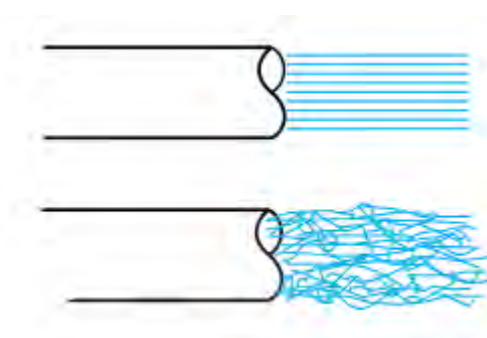
Πολλές ευχαριστίες, επίσης, ανήκουν στον κ. Καρακασίδη, Καθηγητή του Τμήματος Πολιτικών Μηχανικών, επιβλέποντα της εργασίας μου ο οποίος με εμπιστεύτηκε για την ανάθεση και την εκπόνησή της και χωρίς τον οποίο η παρούσα εργασία δεν θα είχε υλοποιηθεί.

## ΕΙΣΑΓΩΓΗ

Η παρακάτω εργασία επεξεργάζεται και σχολιάζει την τυρβώδη κίνηση ρευστού μέσα σε ένα κανάλι με στερεά όρια. Σκοπός της είναι η εξαγωγή συμπερασμάτων σχετικά με τις ταχύτητες των μορίων του ρευστού στις τρεις διευθύνσεις καθώς και η εξήγηση της επιρροής του μαγνητικού πεδίου, όταν αυτό υφίσταται στα τοιχώματα ενός καναλιού, από φυσική και μηχανική σκοπιά. Αρκετές πληροφορίες που παρουσιάζονται στη συνέχεια είναι ήδη γνωστές από τη μηχανική ρευστών ενώ χρησιμοποιήθηκαν και αρκετά στοιχεία και δείκτες από τη στατιστική.

Γενικότερα, η μηχανική ρευστών είναι η επιστήμη εκείνη που μελετά τη στατική αλλά και τη δυναμική κατάσταση ενός ρευστού το οποίο συνοδεύεται πάντοτε και από κάποια δικά του χαρακτηριστικά. Σε αντίθεση με τη μηχανική του στερεού που ασχολείται με σταθερά σχήματα, η μηχανική ρευστών ασχολείται με σώματα που λαμβάνουν το σχήμα του δοχείου τους και κατά συνέπεια έχουν τη δυνατότητα ροής (Παντζάλης, 2017). Τέτοια σώματα είναι τα υγρά και τα αέρια. Βασικά χαρακτηριστικά που ξεχωρίζουν τα ρευστά μεταξύ τους και μας δίνουν πληροφορίες για τα χαρακτηριστικά του καθενός, είναι η θερμοκρασία ( $T$ ), η πυκνότητα ( $\rho$ ) και η πίεση ( $p$ ) (Παντζάλης, 2017). Στη φύση εμφανίζονται πολλά διαφορετικά είδη ροών οι οποίες, μετά από μελέτη και πειραματισμούς, έχουν διαφορετικές παραμέτρους και περιγράφονται από διαφορετικές μαθηματικές εξισώσεις.

Οι ροές χωρίζονται σε δύο μεγάλες κατηγορίες (εικ. 1), ανάλογα με τις τροχιές που διαγράφουν κατά την κίνησή τους, στοιχειώδεις ποσότητες του ρευστού που ρέει. Στην περίπτωση που οι τροχιές είναι παράλληλες γραμμές και δεν διασταυρώνονται, η ροή ονομάζεται στρωτή ενώ εάν οι τροχιές συναντώνται η ροή ονομάζεται τυρβώδης (εικ. 1) (Παντζάλης, 2017).



Εικόνα 1: Στρωτή (πάνω) και τυρβώδης (κάτω) ροή

Έχουν δημιουργηθεί πολλά διαφορετικά μαθηματικά μοντέλα και εξισώσεις για την αναπαράσταση των περισσότερων ειδών ροών ( Navier-Stokes, Bernoulli, αριθμός

Reynolds κ.α.). Στην παρούσα περίπτωση εξετάζεται η τυρβώδης ροή ρευστού σε ένα κανάλι και σκοπός είναι η ενδελεχής μελέτη της συμπεριφοράς των ταχυτήτων με την επιρροή ή όχι μαγνητικού πεδίου σε αυτή. Τα δεδομένα αφορούν σε προσομοιώσεις DNS με παρουσία και απουσία μαγνητικού πεδίου. Η εφαρμογή του μαγνητικού πεδίου μπορεί να επηρεάσει τη ροή ρευστών παρουσία ιόντων. Αναλυτικότερα, έχουν γίνει μετρήσεις για τις συνιστώσες των διανυσμάτων των ταχυτήτων στις τρεις διευθύνσεις σε ένα κανάλι. Για την εύρεση συσχέτισης και την εξαγωγή όλων των στατιστικών δεικτών που απαιτούνται, έγινε αναζήτηση σε στατιστικές μελέτες και επιστημονικά άρθρα που μελετούν παρόμοια προβλήματα. Σε τέτοιου είδους μελέτες που ασχολούνται με ροές ρευστών με ύπαρξη μαγνητικού πεδίου, εντοπίζονται αποτελέσματα που επιβεβαιώνουν ότι πράγματι η επιρροή του μαγνητικού πεδίου στις τιμές και κατευθύνσεις των διανυσμάτων των ταχυτήτων είναι υπαρκτή (Fragkou et al, 2019) (Sarris et al. 2007). Επίσης, σε αρκετές περιπτώσεις η επιρροή αυτή είναι σημαντική και επηρεάζει κατά πολύ την κατεύθυνση της ροής ενώ σε άλλες, ανάλογα βέβαια και με την ένταση του μαγνητικού πεδίου, η επιρροή είναι μικρή έως αμελητέα. Σκοπός της παρακάτω ανάλυσης είναι η δημιουργία ενός προφίλ για τις ταχύτητες σε κανάλι με προδιαγραφές ίδιες με αυτό που χρησιμοποιείται καθώς και η εξαγωγή συμπερασμάτων που έχουν να κάνουν με τη μηχανική ρευστών. Όλοι οι υπολογισμοί και η επεξεργασία των δεδομένων έγινε στο υπολογιστικό λογισμικό του Matlab της εταιρίας Mathworks<sup>1</sup>.

---

<sup>1</sup> Το Matlab (<https://www.mathworks.com/products/matlab.html>), μια προγραμματιστική γλώσσα τέταρτης γενιάς, αποτελεί ένα χρήσιμο για επιστήμονες υπολογιστικό λογισμικό το οποίο προσφέρει στο χρήστη ταχύτητα και ακρίβεια στους υπολογισμούς και γενικότερα κάνει την ενασχόληση με προβλήματα όπως αυτό της εργασίας πολύ πιο εύκολη. Επιπλέον, καθιστά εφικτή την πολύ γρήγορη δημιουργία παραστατικών διαγραμμάτων που βοηθούν στην ευκολότερη κατανόηση των αποτελεσμάτων.

## **ΚΕΦΑΛΑΙΟ 1: ΒΑΣΙΚΟΙ ΣΤΑΤΙΣΤΙΚΟΙ ΔΕΙΚΤΕΣ**

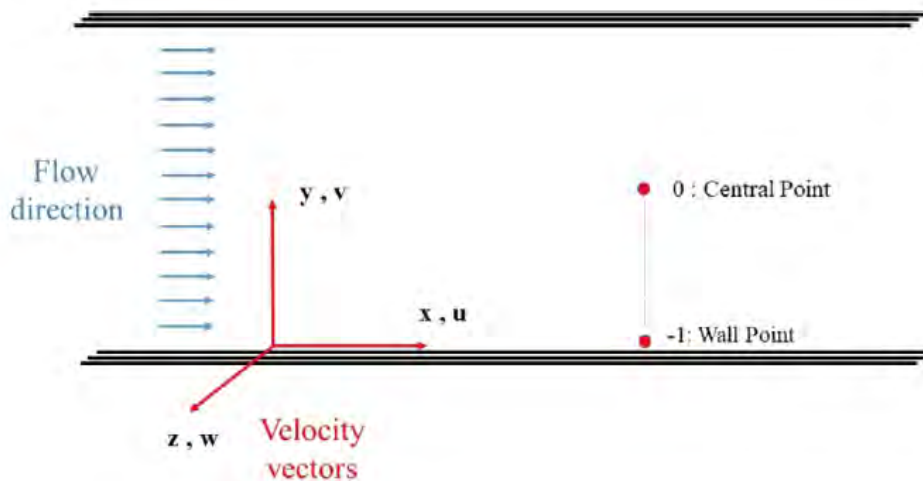
### **1.1 ΕΙΣΑΓΩΓΗ**

Για την επεξεργασία των δεδομένων του προβλήματος απαιτούνται κάποιοι στατιστικοί δείκτες όπως είναι οι συντελεστές συσχέτισης (Pearson coefficients), τα διαγράμματα “βάρους” των δεδομένων (spanning trees) και οι συντελεστές διασυσχέτισης (cross-correlation). Για τον υπολογισμό όλων των παραπάνω, χρειάζεται πρώτα ο υπολογισμός κάποιων βασικών στατιστικών δεικτών που χρησιμοποιούνται σχεδόν σε όλα τα προβλήματα που περιλαμβάνουν στατιστική υπόσταση. Στο κεφάλαιο αυτό γίνεται ο υπολογισμός όλων αυτών των βασικών στατιστικών δεικτών που απαιτούνται για την εξαγωγή των επόμενων πιο εξειδικευμένων δεικτών. Με αυτούς τους πρώτους δείκτες (μέση τιμή, διασπορά, τυπική απόκλιση κ.α.) είναι εφικτή και η δημιουργία της συνάρτησης cross correlation σε επόμενο στάδιο της εργασίας. Για την εύρεση της εν λόγω συνάρτησης απαιτούνται όλοι οι δείκτες που υπολογίζονται σε αυτό το κεφάλαιο. Επιπλέον, γίνεται κανονικοποίηση των τιμών ώστε να είναι ευκολότεροι οι επερχόμενοι υπολογισμοί και να υπάρχουν σε όλες τις διευθύνσεις τα ίδια σημεία αναφοράς ως προς τις τάξεις μεγέθους. Παρακάτω γίνεται μια αναλυτικότερη θεωρητική αναφορά του προβλήματος και παρουσιάζονται οι μέθοδοι υπολογισμού των πρώτων γενικών στατιστικών δεικτών.

### **1.2 ΘΕΩΡΗΤΙΚΗ ΑΝΑΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ**

Σκοπός της εργασίας, όπως αναφέρεται και παραπάνω, είναι η εξαγωγή συμπερασμάτων σχετικά με την επιρροή ή μη, του μαγνητικού πεδίου στη συμπεριφορά των συνιστωσών των ταχυτήτων στις 3 διευθύνσεις σε ένα κανάλι με ροή ρευστού και αναζήτηση και εντοπισμός χωροχρονικών συσχετίσεων στο πεδίο ταχυτήτων της ροής. Το μοντέλο που μελετάται φαίνεται στο παρακάτω διάγραμμα (εικ. 2) όπου απεικονίζεται το κανάλι, η κατεύθυνση της ροής καθώς και οι 3 συνιστώσες της ταχύτητας.





Εικόνα 2: Ροή καναλιού

Στην παραπάνω εικόνα (εικ. 2) δίνεται ουσιαστικά μια σχηματική απεικόνιση του προβλήματος που μελετάται. Με τα μπλε βέλη συμβολίζεται η κατεύθυνση ροής του καναλιού που ορίζουν οι άνω και κάτω μαύρες παράλληλες γραμμές. Επιπλέον, με κόκκινα βέλη φαίνονται οι τρεις διευθύνσεις καθώς και τα διανύσματα ταχυτήτων που αντιστοιχούν σε κάθε διεύθυνση. Κάθε διάνυσμα ταχύτητας εντός του καναλιού μπορεί να αναλυθεί σε αυτές τις συνιστώσες ( $u, v, w$ ) με σκοπό να υπάρχει μεγαλύτερο εύρος συμπερασμάτων σχετικά με την επιρροή του μαγνητικού πεδίου στα διανύσματα. Το σημείο 0 βρίσκεται στο κέντρο του καναλιού και το σημείο -1 αφορά το όριο του καναλιού. Το συγκεκριμένο σχήμα αντικατοπτρίζει πλήρως το μοντέλο των ταχυτήτων που μετρήθηκαν όπως φαίνεται παρακάτω.

Το πρόβλημα που αντιμετωπίζεται στο παρακάτω κεφάλαιο είναι η εύρεση χωροχρονικών συσχετίσεων μεταξύ ταχυτήτων σε ροή ρευστού μέσα σε κανάλι, μέσω μετρήσεων που έγιναν κατά την ύπαρξη ή μη μαγνητικού πεδίου στα τοιχώματα του καναλιού. Με έρευνες που έχουν γίνει μέχρι σήμερα έχει αποδειχθεί πως μπορεί να υπάρξει επηρεασμός μιας ροής ενός ηλεκτρικά αγώγιμου ρευστού όταν στην περιοχή της ροής υπάρχει επιρροή μαγνητικού πεδίου (Fragkou et al, 2019). Μάλιστα, η βάση της πρακτικής αυτής χρησιμοποιείται σε πολλές εφαρμογές όπως είναι διάφορες τεχνικές διόγκωσης κρυστάλλων (μέθοδος **Czochralski**, Lynn F. Schneemeyer, 2003), σε καλύμματα αντιδραστήρων σύντηξης (Molokov et al, 2007) καθώς και σε μαγνητοϋδροδυναμικές αντλίες (Al-Hababeh et al, 2016). Μέσω εργαστηριακής μελέτης που έγινε για όλες τις παραπάνω πρακτικές αλλά και με δοκιμές στο πεδίο, όπου υπάρχουν φυσικά ή και τεχνητά κανάλια (ποταμοί κ.α.), έγινε γνωστό πως το εξωτερικό μαγνητικό πεδίο επηρεάζει σημαντικά τη ροή ενός ηλεκτρικά αγώγιμου ρευστού (Fragkou et al, 2019). Στο συγκεκριμένο πρόβλημα και γενικότερα σε προβλήματα παρόμοιας φύσης που απλώνονται και σε άλλες επιστήμες (υδρολογία, βιολογία, οικονομική επιστήμη κ.α.), η ανάλυση των δεδομένων ακολουθεί τη θεωρία των χρονοσειρών καθώς τέτοια δυναμικά συστήματα προσομοιώνονται βέλτιστα και με τις λιγότερες δυνατές αποκλίσεις με

χρήση αυτής της θεωρίας. Η προσομοίωση των ταχυτήτων με τον παραπάνω τρόπο δίνει τελικά την ύπαρξη ή μη συσχέτισης μέσω της θεωρίας των χρονοσειρών. Επιπλέον, μέσω της συσχέτισης που φαίνεται στα διαγράμματα (correlation plots), βγαίνουν αρκετά συμπεράσματα που παρουσιάζονται παρακάτω.

### **1.3 ΜΕΤΡΗΣΕΙΣ ΤΑΧΥΤΗΤΩΝ ΣΕ ΚΑΝΑΛΙ ΜΕ ΚΑΙ ΧΩΡΙΣ ΤΗ ΔΙΕΝΕΡΓΕΙΑ ΜΑΓΝΗΤΙΚΟΥ ΠΕΔΙΟΥ**

Για το συγκεκριμένο τμήμα της εργασίας δόθηκαν μετρήσεις σε 45 σημεία εντός του καναλιού που αφορούν σε μετρήσεις ταχυτήτων με και χωρίς την ύπαρξη μαγνητικού πεδίου στα δύο παράλληλα με τη ροή τοιχώματα. Κάθε μέτρηση αποτελεί στην ουσία μια χρονοσειρά με τιμές που εξάγονται στις τρεις διευθύνσεις εντός του καναλιού. Επομένως, σε κάθε ένα από τα 45 σημεία έχουν πραγματοποιηθεί 6 μετρήσεις (3 για κάθε διεύθυνση με και χωρίς μαγνητικό πεδίο) ενώ για κάθε μία διεύθυνση υπάρχει μία αναλυτική χρονοσειρά 3601 τιμών.

Μετά τη συλλογή των μετρήσεων ζητήθηκε ο υπολογισμός διαφόρων στατιστικών δεικτών με σκοπό την εξαγωγή ενός προφίλ συμπεριφοράς των ταχυτήτων με και χωρίς μαγνητικό πεδίο. Συγκεκριμένα, έγινε κανονικοποίηση των μετρήσεων και υπολογίστηκαν οι συντελεστές Pearson για να βρεθεί αν και κατά πόσο υπάρχει συσχέτιση μεταξύ των ταχυτήτων. Οι υπολογισμοί έγιναν στο περιβάλλον του λογισμικού Matlab το οποίο προσφέρει ταχύτητα στους υπολογισμούς με αυτοματοποιημένες εντολές και παραστατικές λύσεις μέσω της εξαγωγής διαγραμμάτων.

Πιο συγκεκριμένα, το Matlab αποτελεί υπολογιστικό λογισμικό το οποίο εκτελεί αριθμητικές και λογικές πράξεις με χρήση των γλωσσών προγραμματισμού C και Fortran. Περιέχει πολλές βιβλιοθήκες με αυτοματοποιημένες εντολές που εκτελούν συγκεκριμένες εντολές και πολλούς τύπους διαγραμμάτων για κάθε περίπτωση. Αποτελεί ιδανικό υπολογιστικό περιβάλλον για ανάλυση δεδομένων και γραφικές λύσεις. Με τον τρόπο αυτό οι υπολογισμοί γίνονται ευκολότεροι για το χρήστη ενώ καταλαμβάνουν ακόμα πολύ μικρό χώρο ακόμα και σε ψηφιακό περιβάλλον. Στη συνέχεια, φαίνονται όλες οι εντολές που χρησιμοποιήθηκαν και γίνεται παρουσίαση της καθεμίας ώστε να γίνουν σαφείς οι σκοποί χρήσης τους.

### **1.4 ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΤΙΜΩΝ**

Μετά από την εισαγωγή των αρχικών μετρήσεων στο περιβάλλον **workspace** του Matlab, πρώτο βήμα για την επεξεργασία τους ήταν η κανονικοποίησή τους. Η διαδικασία της κανονικοποίησης γίνεται στα δεδομένα έτσι ώστε να μειωθεί ουσιαστικά το εύρος των τιμών και να μεταφερθούν οι αξίες τους σε αριθμούς που

είναι ευκολότεροι προς επεξεργασία. Κατά την κανονικοποίηση των δεδομένων μας η τυχαία μεταβλητή  $Y$  ακολουθεί την κανονική κατανομή με μέση τιμή μηδέν και διασπορά ένα. Σύμφωνα με τη θεωρία για την κανονικοποίηση ενός πλήθους τιμών είναι αναγκαίος ο υπολογισμός της μέσης τιμής αλλά και της τυπικής απόκλισής τους.

Η μέση τιμή ανήκει στα μέτρα θέσης της περιγραφικής στατιστικής, τα οποία προσδιορίζουν χαρακτηριστικές θέσεις μέσα στο εύρος των δεδομένων (Κουγιουμτζής, 2009). Συμβολίζεται με  $\bar{x}$  και η τιμή της δίνεται από τον παρακάτω τύπο:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

όπου:

- $x_i$  η τιμή της παρατήρησης
- $n$  το πλήθος των παρατηρήσεων

Η τυπική απόκλιση ενός πλήθους δεδομένων ανήκει στα μέτρα μεταβλητότητας της περιγραφικής στατιστικής, τα οποία δίνουν περιληπτικά τη διασκόρπιση και μεταβλητότητα των δεδομένων (Κουγιουμτζής, 2009) και υπολογίζεται μέσω της διακύμανσης ή διασποράς  $s^2$  που δίνεται από τον παρακάτω τύπο:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2)$$

όπου:

- $n$  το πλήθος των παρατηρήσεων
- $x_i$  η τιμή της παρατήρησης
- $\bar{x}$  η μέση τιμή των παρατηρήσεων

Η δειγματική τυπική απόκλιση  $s$  δίνεται από τη σχέση:

$$s = \sqrt{s^2} \quad (3)$$

Για την κανονικοποίηση των τιμών στο περιβάλλον του Matlab, χρησιμοποιήθηκε η παρακάτω ρουτίνα:

```
mesitimiw45nm=mean(w45nm);  
stdevw45nm=std(w45nm);  
kanonw45nm=zeros(3601,1);  
for i=1:3601  
    kanonw45nm(i)=(w45nm(i)-mesitimiw45nm)/stdevw45nm;  
end
```

όπου:

- η μεταβλητή `mesitimi` μας δίνει τη μέση τιμή των μετρήσεων σε κάθε σημείο μέσω της συνάρτησης του matlab `'mean'`,
- η μεταβλητή `stdev` περιέχει τιμές της τυπικής απόκλισης με την αυτοματοποιημένη συνάρτηση `'std'`,
- η μεταβλητή `kanon` μέσω της συνάρτησης `'zeros'` δίνει το εύρος της κανονικοποίησης (0,1) και
- η εντολή `'for... end'`, η οποία καταδεικνύει επαναληπτικό βρόγχο, μας δίνει την κανονικοποίηση κάθε τιμής μέσα στο ορισμένο εύρος.

Μέσω της παραπάνω ρουτίνας δημιουργήθηκαν συνολικά 270 κανονικοποιημένοι πίνακες με 3601 τιμές ο κάθε ένας. Ο αριθμός 270 προκύπτει από τη στιγμή που έχουμε 45 σημεία για κάθε διεύθυνση ταχύτητας ( $u,v,w$ ) για τις περιπτώσεις με και χωρίς μαγνητικό πεδίο. Οι παραπάνω τιμές αποθηκεύτηκαν στο πεδίο εργασίας του matlab (`workspace`) για περαιτέρω επεξεργασία.

Στη συνέχεια, οι κανονικοποιημένες τιμές που βρίσκονται σε μορφή διανυσμάτων, πίνακες μίας στήλης και 3601 γραμμών, ομαδοποιήθηκαν σε μεγαλύτερους πίνακες για υπολογισμό των συντελεστών Pearson. Συγκεκριμένα, δημιουργήθηκε ένας πίνακας  $A$  για κάθε διεύθυνση ταχύτητας με και χωρίς την ύπαρξη μαγνητικού πεδίου όπως φαίνεται παρακάτω (εικ. 3):

Aum	[kanonu1m kanonu2m kanonu3m kanonu4m kanonu5m kanonu6m kanonu7m kanonu8m... kanonu9m kanonu10m kanonu11m kanonu12m kanonu13m kanonu14m kanonu15m kanonu16m... kanonu17m kanonu18m kanonu19m kanonu20m kanonu21m kanonu22m kanonu23m kanonu24m... kanonu25m kanonu26m kanonu27m kanonu28m kanonu29m kanonu30m kanonu31m kanonu32m... kanonu33m kanonu34m kanonu35m kanonu36m kanonu37m kanonu38m kanonu39m kanonu40m... kanonu41m kanonu42m kanonu43m kanonu44m kanonu45m]
Avm	[kanonv1m kanonv2m kanonv3m kanonv4m kanonv5m kanonv6m kanonv7m kanonv8m... kanonv9m kanonv10m kanonv11m kanonv12m kanonv13m kanonv14m kanonv15m kanonv16m... kanonv17m kanonv18m kanonv19m kanonv20m kanonv21m kanonv22m kanonv23m kanonv24m... kanonv25m kanonv26m kanonv27m kanonv28m kanonv29m kanonv30m kanonv31m kanonv32m... kanonv33m kanonv34m kanonv35m kanonv36m kanonv37m kanonv38m kanonv39m kanonv40m... kanonv41m kanonv42m kanonv43m kanonv44m kanonv45m]
Awm	[kanonw1m kanonw2m kanonw3m kanonw4m kanonw5m kanonw6m kanonw7m kanonw8m... kanonw9m kanonw10m kanonw11m kanonw12m kanonw13m kanonw14m kanonw15m kanonw16m... kanonw17m kanonw18m kanonw19m kanonw20m kanonw21m kanonw22m kanonw23m kanonw24m... kanonw25m kanonw26m kanonw27m kanonw28m kanonw29m kanonw30m kanonw31m kanonw32m... kanonw33m kanonw34m kanonw35m kanonw36m kanonw37m kanonw38m kanonw39m kanonw40m... kanonw41m kanonw42m kanonw43m kanonw44m kanonw45m]
Aunm	[kanonul1nm kanonu2nm kanonu3nm kanonu4nm kanonu5nm kanonu6nm kanonu7nm kanonu8nm... kanonu9nm kanonu10nm kanonu11nm kanonu12nm kanonu13nm kanonu14nm kanonu15nm kanonu16nm... kanonu17nm kanonu18nm kanonu19nm kanonu20nm kanonu21nm kanonu22nm kanonu23nm kanonu24nm... kanonu25nm kanonu26nm kanonu27nm kanonu28nm kanonu29nm kanonu30nm kanonu31nm kanonu32nm... kanonu33nm kanonu34nm kanonu35nm kanonu36nm kanonu37nm kanonu38nm kanonu39nm kanonu40nm... kanonu41nm kanonu42nm kanonu43nm kanonu44nm kanonu45nm]
Avnm	[kanonv1nm kanonv2nm kanonv3nm kanonv4nm kanonv5nm kanonv6nm kanonv7nm kanonv8nm... kanonv9nm kanonv10nm kanonv11nm kanonv12nm kanonv13nm kanonv14nm kanonv15nm kanonv16nm... kanonv17nm kanonv18nm kanonv19nm kanonv20nm kanonv21nm kanonv22nm kanonv23nm kanonv24nm... kanonv25nm kanonv26nm kanonv27nm kanonv28nm kanonv29nm kanonv30nm kanonv31nm kanonv32nm... kanonv33nm kanonv34nm kanonv35nm kanonv36nm kanonv37nm kanonv38nm kanonv39nm kanonv40nm... kanonv41nm kanonv42nm kanonv43nm kanonv44nm kanonv45nm]
Awnm	[kanonw1nm kanonw2nm kanonw3nm kanonw4nm kanonw5nm kanonw6nm kanonw7nm kanonw8nm... kanonw9nm kanonw10nm kanonw11nm kanonw12nm kanonw13nm kanonw14nm kanonw15nm kanonw16nm... kanonw17nm kanonw18nm kanonw19nm kanonw20nm kanonw21nm kanonw22nm kanonw23nm kanonw24nm... kanonw25nm kanonw26nm kanonw27nm kanonw28nm kanonw29nm kanonw30nm kanonw31nm kanonw32nm... kanonw33nm kanonw34nm kanonw35nm kanonw36nm kanonw37nm kanonw38nm kanonw39nm kanonw40nm... kanonw41nm kanonw42nm kanonw43nm kanonw44nm kanonw45nm]

Εικόνα 3: Πίνακες κανονικοποίησης

## ΚΕΦΑΛΑΙΟ 2: Ο ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ PEARSON

### 2.1 ΓΕΝΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

Ο συντελεστής Pearson μας δίνει τη συσχέτιση μεταξύ δύο μεταβλητών ( $x, y$ ) και την εξάρτηση, εάν υπάρχει, που έχει η μια με την άλλη (Παπαδόπουλος, 2013). Ορίστηκε από τον Άγγλο μαθηματικό Karl Pearson ο οποίος του έδωσε το όνομά του και δίνεται από τον παρακάτω τύπο:

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{n s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (4)$$

όπου:

- $x_i, y_i$  οι τιμές των μεταβλητών,
- $n$  το πλήθος των μεταβλητών,
- $\bar{x}, \bar{y}$  οι δειγματικές μέσες τιμές των μεταβλητών και
- $s_x, s_y$  οι δειγματικές τυπικές αποκλίσεις των μεταβλητών

Ο συντελεστής Pearson αποτελεί ένα μέτρο της γραμμικής συσχέτισης μεταξύ δύο τυχαίων μεταβλητών. Επιπλέον, είναι το πρώτο ιστορικά μέτρο συσχέτισης και μέχρι σήμερα ίσως το πιο ευρέως χρησιμοποιούμενο (Burgess, 2002). Το εύρος των τιμών του συντελεστή είναι  $(-1, 1)$  όπου στην περίπτωση  $+1$  υπάρχει τέλεια, γραμμική συσχέτιση μεταξύ των μεταβλητών, ενώ η τιμή  $-1$  υποδηλώνει τέλεια αντισυσχέτιση. Για ενδιάμεσες τιμές προκύπτουν τα παρακάτω συμπεράσματα:

- $r = \pm 0.1$  : μικρή συσχέτιση
- $r = \pm 0.3$  : μέτρια συσχέτιση
- $r = \pm 0.5$  : ισχυρή συσχέτιση

Σε κάθε περίπτωση υποδηλώνεται ο βαθμός της γραμμικής συσχέτισης μεταξύ των μεταβλητών. Στην περίπτωση που ο συντελεστής είναι ίσος με το 0, δεν υφίσταται καμία συσχέτιση μεταξύ των μεταβλητών και αυτές ορίζονται ως ανεξάρτητες.

Σε μία δειγματοληπτική έρευνα εφόσον βρεθεί συσχέτιση μεταξύ δύο μεταβλητών  $X$  και  $Y$  αυτό σημαίνει πως οι δύο μεταβλητές συνδέονται απλώς με κάποια σχέση. Η ύπαρξη μη μηδενικού συντελεστή Pearson, συνεπάγεται μόνο την ύπαρξη σύνδεσης μεταξύ τους και σε καμία περίπτωση δεν συνδέεται η σύνδεση αυτή με αιτιότητα (Αγγελής, 2013). Αιτιότητα δύναται να υπάρχει μπορεί όμως και όχι. Υπάρχει πάντοτε και η πιθανότητα οι δύο μεταβλητές να συνδέονται λόγω της κοινής τους σύνδεσης με μια άλλη τρίτη μεταβλητή. Σε κάθε περίπτωση, αλληλοσυσχέτιση υπάρχει μόνο κατόπιν πειραματικής διαδικασίας ή εφόσον αυτή υπαγορεύεται από επιστημονική βάση.

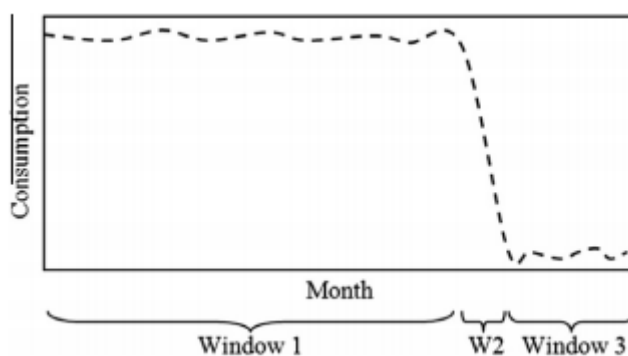
Ο συντελεστής Pearson και ο υπολογισμός του μπορούν να βοηθήσουν σε πολλά διαφορετικά πεδία που αναζητείται συσχέτιση μεταξύ μεταβλητών. Ενδεικτικά σύμφωνα με έρευνες και επιστημονικά άρθρα, ο συντελεστής Pearson μπορεί να χρησιμοποιηθεί στον εντοπισμό καρκινογόνων κληρονομικών κυττάρων (Hung-Chung Huang et al, 2010) στη δημιουργία φίλτρων που αποβάλλουν μοριακές ενώσεις (Zhou et al, 2016) καθώς και στον εντοπισμό απάτης ή δολιοφθοράς μέσα σε εταιρίες παραγωγής ηλεκτρικού ρεύματος (Monedero et al, 2012). Και στις τρεις παραπάνω περιπτώσεις λαμβάνονται δεδομένα που έχουν συγκεκριμένη διασπορά και στατιστικούς δείκτες και εντοπίζεται, εφόσον υπάρχει, συσχέτιση με άλλες μεταβλητές.

Στην περίπτωση των γενετικών κυττάρων που μπορούν να μεταφέρουν ασθένειες από έναν άνθρωπο στους απογόνους του, έγινε χρήση του συντελεστή Pearson καθώς και του δείκτη απόστασης Kolmogorov-Smirnov. Το μοντέλο βασίζεται στην αναγνώριση αυτού που οι επιστήμονες της ιατρικής ονομάζουν γονιδιακό προφίλ (Gene-Expression-Profile/GEP). Η έρευνα βασίστηκε στην υπόθεση ότι τα γονίδια του βιοδείκτη, τα οποία διαχωρίζουν τα μολυσμένα από τα φυσιολογικά δείγματα παρουσιάζουν ευδιάκριτα γονιδιακά προφίλ μεταξύ των δύο ομάδων (Huang et al, 2010). Με τη χρήση του συντελεστή Pearson προσδιορίζονται συγκεκριμένοι νοσούντες βιολογικοί δείκτες κάνοντας σύγκριση μεταξύ αυτών και των υγιών ομοειδών τους και παραβάλλοντας τη συγκεκριμένη μελέτη με άλλες σχετικές στον ιατρικό τομέα, η έρευνα επιδίωξε την ταυτοποίηση πιθανών υποψήφιων παθογενών προφίλ. Με αυτό τον τρόπο γίνεται εφικτή η πρόληψη αλλά κυρίως η άμεση αντιμετώπιση νοσημάτων για ανθρώπους των οποίων οι πρόγονοι πάσχουν από κάποια κληρονομική ασθένεια. Φαίνεται πως ακόμη και σε θέματα υγείας οι στατιστικοί δείκτες και ειδικότερα ο συντελεστής Pearson μπορούν να βοηθήσουν ως εργαλεία έρευνας.

Ο συντελεστής Pearson χρησιμοποιήθηκε και για τη βελτίωση των φίλτρων αποβολής μικροοργανισμών. Μέχρι σήμερα είναι αποδεδειγμένο πως τα μοριακά φίλτρα είναι πολύ αποδοτικά και αποτελεσματικά όταν χρησιμοποιούνται σε μη γραμμικά (non-Gaussian) συστήματα. Το μεγάλο μειονέκτημα της λειτουργίας τους εντοπίζεται στο γεγονός ότι προκαλούν εκφυλισμό στα σωματίδια. Μέσω της χρήσης του συντελεστή Pearson, έγινε προσπάθεια για τη μείωση του εκφυλισμού αυτού. Πιο συγκεκριμένα, γίνεται δειγματοληψία των σωματιδίων προτού αυτά περάσουν από το φίλτρο και δειγματοληψία των φιλτραρισμένων ομοειδών τους. Έπειτα, με υπολογισμό βασικών στατιστικών δεικτών (μέση τιμή, τυπική απόκλιση κ.α.) εντοπίζεται και ο συντελεστής Pearson που συνδέει τα φιλτραρισμένα με τα μη φιλτραρισμένα σωματίδια (Zhou et al, 2016). Σκοπός είναι οι τιμές να είναι κοντά στη μονάδα ώστε να υπάρχει έντονη σύνδεση μεταξύ των μορίων. Σίγουρα η συσχέτιση δεν θα είναι ποτέ απόλυτα γραμμική ( $r=+1$ ) διότι τα φίλτρα κατακρατούν συγκεκριμένες ενώσεις που θέλουν οι κατασκευαστές. Ωστόσο η ισχυρή συσχέτιση μετά το φιλτράρισμα, δείχνει πως ο εκφυλισμός των μορίων είναι ο ελάχιστος δυνατός. Μετά από πολλές διαφορετικές δοκιμές και συνεχή δειγματοληψία και σύγκριση, μπορεί να στοιχειοθετηθεί ένα μοριακό φίλτρο που να εξυπηρετεί το

σκοπό κατασκευής του με ελαχιστοποίηση των σφαλμάτων. Το παραπάνω παράδειγμα αποτελεί ένα από τα χιλιάδες στο οποίο η στατιστική μελέτη και πιο συγκεκριμένα η χρήση του συντελεστή Pearson, μπορεί να βοηθήσει κατά πολύ στην ανάπτυξη και βελτιστοποίηση ενός προϊόντος.

Ακόμη μία εφαρμογή του συντελεστή Pearson εμφανίζεται στον εντοπισμό και την αποκάλυψη απάτης και λανθασμένων χειρισμών σε επιχειρήσεις. Πιο συγκεκριμένα, σε επιχειρήσεις παραγωγής και διανομής ηλεκτρικού ρεύματος έγινε χρήση του συντελεστή (Monedero et al, 2012). Η εταιρεία Endesa, η οποία δραστηριοποιείται στη διανομή και παραγωγή ηλεκτρικού ρεύματος σε χώρες όπως η Ισπανία αλλά και σε χώρες της λατινικής Αμερικής, σε συνεργασία με το πανεπιστήμιο της Σεβίλλης κατάφερε με αλγορίθμους που εμπεριέχουν τον συντελεστή Pearson και άλλους στατιστικούς δείκτες να εντοπίζουν απώλειες εσόδων που οφείλονται σε μη τεχνικές βλάβες. Τέτοιου είδους απώλειες είναι αυτές που κοστίζουν περισσότερο από όλες σε τέτοιες εταιρείες καθώς δεν εντοπίζονται εγκαίρως και μπορούν να υφίστανται για μεγάλο χρονικό διάστημα. Πολλές φορές οι μειώσεις στην κατανάλωση ρεύματος μπορεί να οφείλονται σε αλλαγές συμβολαίων των καταναλωτών η και σε ποσοστά υπομέτρησης που οφείλονται σε παλιό ή προβληματικό μηχανισμό μέτρησης. Οι πτώσεις όμως ποτέ δεν μπορούν να είναι δραστικές καθώς πρέπει καθημερινά να καλύπτονται οι ανάγκες του ανθρώπου που απαιτούν τη χρήση ηλεκτρικού ρεύματος. Με τον συντελεστή Pearson εντοπίστηκε συσχέτιση στις ανθρώπινες ανάγκες και στην κατανάλωση ενέργειας. Με αυτό τον τρόπο δημιουργήθηκαν προφίλ ημερήσιων αναγκών ανά νοικοκυριό και σε περιπτώσεις που εντοπίστηκαν μεγάλες και με βάθος χρόνου αποκλίσεις έγινε έλεγχος για τυχόν απάτες από τους καταναλωτές. Σε αρκετές περιπτώσεις εντοπίστηκε παρέμβαση στους μετρητές από τους καταναλωτές, τους επιλήφθηκε πρόστιμο και έπειτα έγινε επισκευή των μετρητών. Στο παρακάτω σχήμα (εικ.4) φαίνεται η αλλαγή στην καμπύλη κατανάλωσης μέσα σε διάστημα ενός μήνα.



Εικόνα 4: Διάγραμμα μηνιαίας κατανάλωσης

Από όλες τις παραπάνω εφαρμογές του συντελεστή Pearson, καταλαβαίνει κανείς πως είναι πολύ μεγάλο το φάσμα που μπορεί να εξυπηρετήσει την ανθρώπινη



έρευνα. Όπως συμβαίνει και με άλλα επιστημονικά εργαλεία και πάντοτε σε συνεργασία με άλλους επιστημονικούς κλάδους, η χρήση τέτοιων δεικτών εξυπηρετεί την καθημερινότητα και την εξέλιξη της ανθρωπότητας.

Άλλα μέτρα που χρησιμοποιούνται για να εκφράσουν συσχέτιση είναι ο συντελεστής συσχέτισης του Spearman και το  $\tau$  του Kendall (Αγγελής, 2013). Ο συντελεστής του Spearman αποτελεί μη παραμετρικό στατιστικό μέτρο για δεδομένα όχι κατά ανάγκη κανονικά ή συνεχή. Βασίζεται σε διάταξη των δεδομένων και υπολογισμό του συντελεστή Pearson ενώ κρίνεται ιδανικός για μεταβλητές διάταξης. Το  $\tau$  του Kendall είναι όπως και ο συντελεστής Spearman μη παραμετρικός συντελεστής, ιδανικός για μικρά σύνολα δεδομένων. Επιπλέον, για έρευνες που αφορούν πληθυσμό θεωρείται ο καταλληλότερος και καλύτερος συντελεστής συσχέτισης.

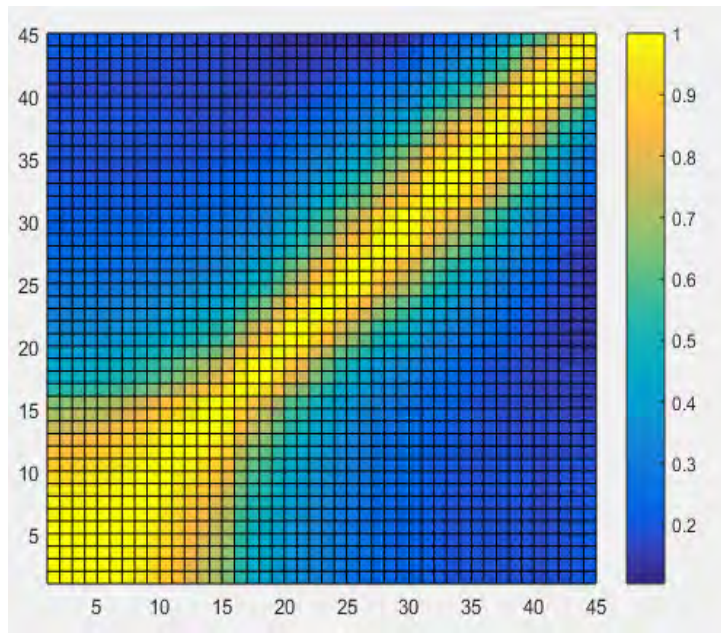
## 2.2 ΕΦΑΡΜΟΓΗ ΣΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΡΓΑΣΙΑΣ

Αφού έγινε η συλλογή των μετρήσεων, η εύρεση της μέσης τιμής, της τυπικής απόκλισης και η κανονικοποίηση των τιμών για τις ταχύτητες και στις τρεις διευθύνσεις, η εργασία συνεχίστηκε με την εύρεση του συντελεστή Pearson. Στην προκειμένη περίπτωση για την εξαγωγή των συντελεστών Pearson για εύρεση συσχέτισης μεταξύ των ταχυτήτων, χρησιμοποιήθηκε αυτοματοποιημένη συνάρτηση που βρίσκεται στο λογισμικό του Matlab (εικ. 5). Η συνάρτηση ονομάζεται "corrcoef" (correlation-coefficient) και οι πίνακες για τους οποίους υπολογίστηκαν οι συντελεστές περιέχουν τις κανονικοποιημένες τιμές που υπολογίστηκαν προηγουμένως.

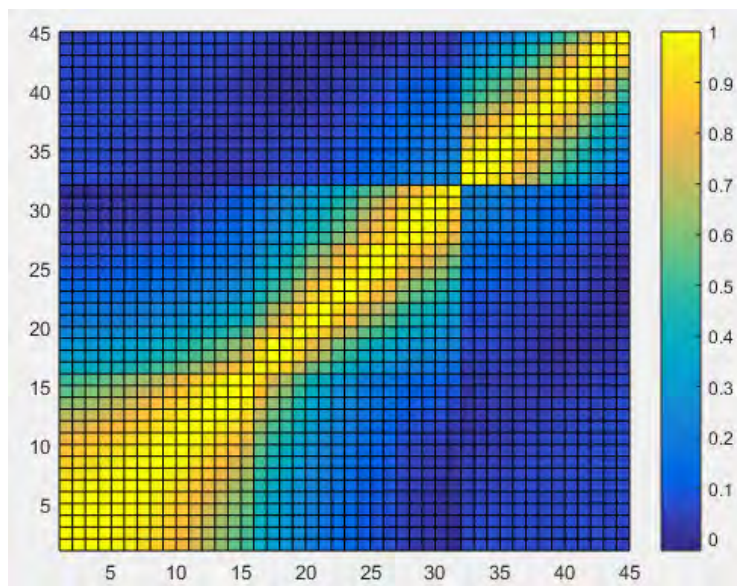
```
Rum=corrcoef (Aum)
Rvm=corrcoef (Avm)
Rwm=corrcoef (Awm)
Runm=corrcoef (Aunm)
Rvnm=corrcoef (Avnm)
Rwnm=corrcoef (Awnm)
Rumix=corrcoef (Aum, Aunm)
Rvmix=corrcoef (Avm, Avnm)
Rwmix=corrcoef (Awm, Awnm)
```

Εικόνα 5: Πίνακες Pearson

Όπως φαίνεται και στις παραπάνω εντολές, οι συντελεστές υπολογίστηκαν για τους πίνακες ταχυτήτων κάθε διεύθυνσης με και χωρίς την ύπαρξη μαγνητικού πεδίου. Οι συντελεστές εμφανίζονται στη μορφή διαγραμμάτων έτσι ώστε τα αποτελέσματα να είναι παραστατικά και κατανοητά προς όλους. Τα διαγράμματα που φαίνονται παρακάτω (εικ. 6,7) μπορούν να δώσουν πολλές πληροφορίες των οποίων η ανάλυση γίνεται στη συνέχεια.



Εικόνα 6: Διάγραμμα ταχύτητας  $U$  (με μαγνητικό πεδίο)



Εικόνα 7: Διάγραμμα ταχύτητας  $U$  (χωρίς μαγνητικό πεδίο)

Τα διαγράμματα δημιουργήθηκαν με τη συνάρτηση του Matlab “rcolor”. Έχουν ληφθεί ως παράδειγμα τα διαγράμματα για την ταχύτητα  $U$  (διεύθυνση  $X$ ) για μία πρώτη εξαγωγή συμπερασμάτων. Για την κατανόηση της διακύμανσης των τιμών τοποθετήθηκε στο διάγραμμα μέσω του λογισμικού η μπάρα αναφοράς (colorbar) κατά την οποία:

- με σκούρο μπλε χρώμα απεικονίζονται οι τιμές που είναι μηδενικές (ανεξάρτητες μεταβλητές),
- με γαλάζιο χρώμα οι τιμές που βρίσκονται κοντά στην τιμή 0,5 (μέτρια γραμμική συσχέτιση) και

- με κίτρινο χρώμα οι τιμές που είναι ίσες με 1 (τέλεια γραμμική συσχέτιση).

Οι ενδιάμεσες τιμές εμφανίζονται με συνδυασμούς των παραπάνω χρωματισμών.

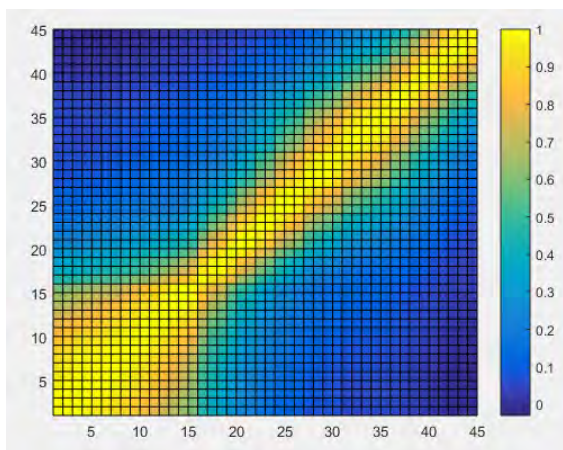
Όπως αναμενόταν, η κύρια διαγώνιος των πινάκων (εικ. 6, 7) είναι χρωματισμένη με κίτρινο χρώμα κάτι που σημαίνει πως οι τιμές της είναι ίσες με τη μονάδα. Αυτό συμβαίνει διότι σε κάθε θέση της κύριας διαγωνίου του πίνακα των συντελεστών συσχέτισης δίνεται ο συντελεστής συσχέτισης κάθε στήλης με τον εαυτό της και επομένως υπάρχει τέλεια συσχέτιση. Επιπλέον, στα διαγράμματα (εικ. 6, 7) φαίνεται πως στις περιοχές που περικλείουν την κύρια διαγώνιο, υπάρχει επίσης κίτρινο χρώμα (τιμές κοντά στη μονάδα) κάτι που σημαίνει πως υπάρχει και εκεί πολύ ισχυρή συσχέτιση. Κάτι τέτοιο δείχνει πως η συσχέτιση κάθε στήλης ταχυτήτων με τις κοντινές της είναι αρκετά έντονη. Επίσης με την απομάκρυνση από την κύρια διαγώνιο παρατηρείται και η εντονότερη εμφάνιση μπλε χρωματισμών, κάτι που σημαίνει ότι κάθε στήλη έχει μικρότερη συσχέτιση με τις πιο μακρινές ομοειδείς της. Ενδεικτικά, κοιτώντας τους πίνακες που δημιουργήθηκαν μέσω του Matlab και πιο συγκεκριμένα στις τιμές του συντελεστή συσχέτισης για την κατεύθυνση  $u$ , στα κελιά που επικοινωνούν άμεσα με την κύρια διαγώνιο οι τιμές των συντελεστών κυμαίνονται μεταξύ 0.999 και 0.9321. Με την απομάκρυνση από την κύρια διαγώνιο αυξάνεται η διασπορά των τιμών και μειώνεται αισθητά και η τιμή τους αφού στην ίδια διεύθυνση και μακριά από τη διαγώνιο συναντάμε τιμές της τάξης του 0.05.

Συγκριτικά φαίνεται να υπάρχουν αρκετές ομοιότητες μεταξύ του διαγράμματος με μαγνητικό (εικ. 6) πεδίο και εκείνου χωρίς μαγνητικό πεδίο (εικ. 7). Πιο συγκεκριμένα και στα δύο διαγράμματα φαίνεται πως η κύρια διαγώνιος είναι ίση με τη μονάδα. Επιπροσθέτως, φαίνεται ότι με την απομάκρυνση από την κεντρική διαγώνιο εμφανίζεται και μείωση των τιμών στους συντελεστές Pearson. Επομένως, με αύξηση της απόστασης υπάρχει μείωση της συσχέτισης. Τέλος, από τα διαγράμματα φαίνεται πως η κύρια διαγώνιος λειτουργεί και σαν άξονας συμμετρίας για τα χρώματα ανάντι και κατόντι αυτής. Με μια ματιά στους ακριβείς πίνακες των συντελεστών γίνεται αντιληπτό πως πρόκειται για συμμετρικούς πίνακες πράγμα που σημαίνει πως με αντιστροφή γραμμών και στηλών ο πίνακας παραμένει ίδιος.

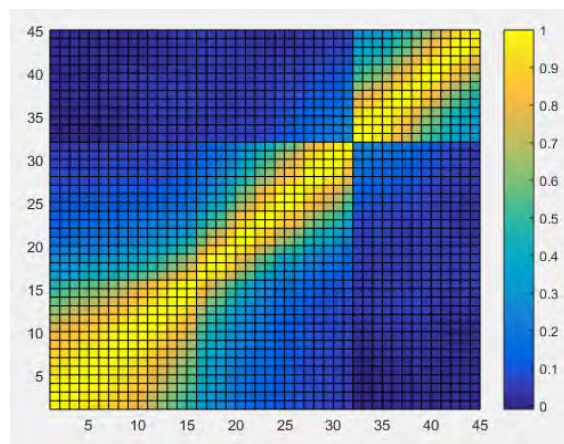
Με μεγαλύτερη προσοχή στη λεπτομέρεια, εντοπίζονται και διαφορές μεταξύ των δύο διαγραμμάτων. Αν και στις δύο περιπτώσεις φαίνεται να υπάρχει σχεδόν μία ομαλότητα στον τρόπο αλλαγής των συντελεστών Pearson με απομάκρυνση από την κύρια διαγώνιο, στο διάγραμμα που αφορά τη μη ύπαρξη μαγνητικού πεδίου αυτή η ομαλότητα φαίνεται να διαταράσσεται. Αυτό εντοπίζεται στο γεγονός πως υπάρχουν αποχρώσεις έντονου μπλε χρωματισμού ανάμεσα σε χρωματισμούς κίτρινου και γαλάζιου. Κάτι τέτοιο σημαίνει πως έχουμε ανώμαλη αυξομείωση των συντελεστών συσχέτισης ανάντη αλλά και κατόντη της κύριας διαγωνίου. Κάτι τέτοιο εντοπίζεται και στο διάγραμμα του μαγνητικού πεδίου αλλά με πολύ μικρότερη ένταση. Γενικότερα, οι χρωματισμοί συνάδουν στο μέγιστο των περιπτώσεων με την απομάκρυνση και με τη σύγκριση όλο και πιο μακρινών μεταξύ

τους στηλών. Στο σημείο αυτό θα μπορούσε να γίνει έλεγχος για το αν υπάρχει γραμμικότητα μεταξύ απομάκρυνσης και τιμών συντελεστών Pearson αλλά για να επιβεβαιωθεί κάτι τέτοιο θα πρέπει να ακολουθείται και να επιβεβαιώνεται από όλες τις στήλες. Σίγουρα γραμμικότητα δε συναντάται στην περίπτωση που απουσιάζει το μαγνητικό πεδίο αφού όπως φαίνεται και από το διάγραμμα (εικ. 7) η ομαλή αλλαγή στο χρωματισμό διαταράσσεται στη θέση της 32<sup>ης</sup> στήλης.

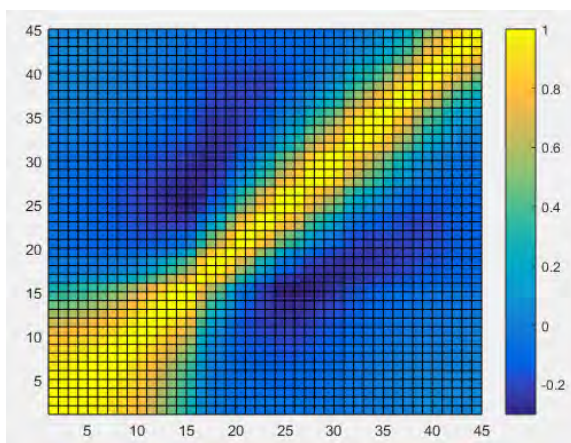
Αν γίνει παρατήρηση και των διαγραμμάτων συντελεστών στις υπόλοιπες διευθύνσεις (εικ.8-11), φαίνεται πως και εκεί εντοπίζονται ομοιότητες και διαφορές αλλά και διαφορετικές συμπεριφορές όσο αφορά στην συσχέτιση.



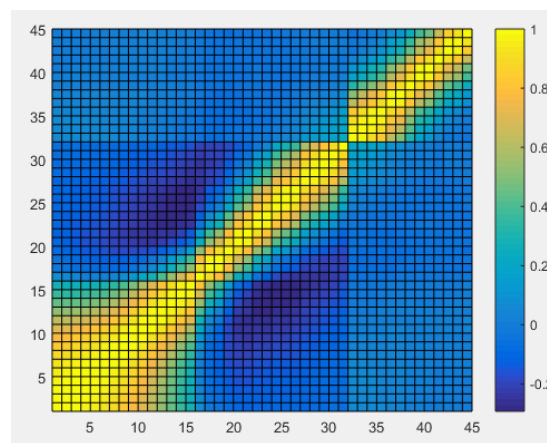
Εικόνα 8: Διάγραμμα ταχύτητας V (με μαγνητικό πεδίο)



Εικόνα 9: Διάγραμμα ταχύτητας V (χωρίς μαγνητικό πεδίο)



Εικόνα 10: Διάγραμμα ταχύτητας W (με μαγνητικό πεδίο)



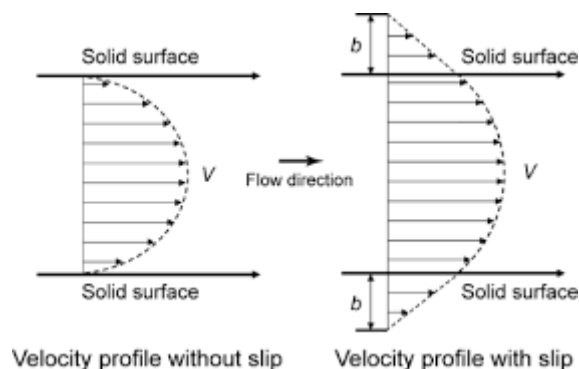
Εικόνα 11: Διάγραμμα ταχύτητας W (χωρίς μαγνητικό πεδίο)

Πιο συγκεκριμένα, στις περιπτώσεις των συνιστωσών V και W με την ύπαρξη μαγνητικού πεδίου, τα δύο διαγράμματα εμφανίζουν έντονη συσχέτιση και κοινούς χρωματισμούς γύρω από την περιοχή της κύριας διαγωνίου. Ωστόσο, στο διάγραμμα της ταχύτητας W κατά την απομάκρυνση από την κύρια διαγώνιο εμφανίζεται έντονο σκούρο μπλε χρώμα το οποίο στη συνέχεια γίνεται ξανά γαλάζιο. Αν ανατρέξει κανείς στους πίνακες με τις τιμές συντελεστών συσχέτισης για

την συγκεκριμένη περίπτωση θα παρατηρήσει πως πράγματι κατά την απομάκρυνση από την κύρια διαγώνιο προς τα άκρα εμφανίζεται απότομη πτώση της τιμής των συντελεστών και μετά ήπια αύξησή τους. Το ίδιο συμβαίνει και στην περίπτωση απουσίας μαγνητικού πεδίου για την ίδια συνιστώσα. Στα διαγράμματα της συνιστώσας  $V$  η αλλαγή των χρωματισμών και κατά συνέπεια των συντελεστών συσχέτισης είναι πιο ομαλή και δεν παρουσιάζει τις παραπάνω αυξομειώσεις. Αυτό μας οδηγεί στο συμπέρασμα πως οι ταχύτητες του άξονα  $z$  και οι συσχετίσεις μεταξύ τους επηρεάζονται από τα μορφολογικά στοιχεία του καναλιού περισσότερο από τις συνιστώσες των διευθύνσεων  $x$  και  $y$ .

### 2.3 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ-ΠΑΡΑΤΗΡΗΣΕΙΣ

Με ανάλυση όλων των διαγραμμάτων μπορεί να βγει το συμπέρασμα πως κάθε πίνακας με στοιχεία ταχυτήτων εμφανίζει τις δικές του διακυμάνσεις όσο αφορά τους συντελεστές Pearson. Ωστόσο, φαίνεται πως στα διαγράμματα που αφορούν την ύπαρξη μαγνητικού πεδίου εμφανίζεται εντονότερη συσχέτιση στις τιμές που βρίσκονται κοντά στην κύρια διαγώνιο, αφού η κίτρινη λωρίδα είναι μεγαλύτερου πάχους από την αντίστοιχη στα διαγράμματα χωρίς μαγνητικό πεδίο. Παρόλα αυτά η αυξομείωση των συντελεστών συσχέτισης μακριά από την κύρια διαγώνιο, φαίνεται να είναι ανεξάρτητη σε κάθε διεύθυνση όπως παρατηρήθηκε και προηγουμένως. Κομβικό ρόλο σε αυτό το φαινόμενο παίζει η μορφολογία του καναλιού αφού τα γεωμετρικά του χαρακτηριστικά επηρεάζουν κάποιες διευθύνσεις και τις συνιστώσες τους, περισσότερο από άλλες. Αυτό ισχύει για όλα τα διαγράμματα. Τέλος, τα διαγράμματα που παρουσιάζουν συντελεστές σε μετρήσεις χωρίς μαγνητικό πεδίο, εμφανίζουν μία κοινή συμπεριφορά στο κελί (32, 32). Φαίνεται εκεί μία απότομη μείωση του συντελεστή συσχέτισης γύρω από την κύρια διαγώνιο στα διαγράμματα και των τριών διευθύνσεων. Αυτό ίσως οφείλεται σε κάποια παράμετρο του πεδίου στο συγκεκριμένο σημείο από τη στιγμή που εμφανίζεται το φαινόμενο σε κάθε διεύθυνση. Από φυσικής άποψης, το προφίλ των συντελεστών συσχέτισης για όλες τις διευθύνσεις που μας δίνουν και οι παραπάνω εικόνες μπορεί να χρησιμοποιηθεί για εξαγωγή χρήσιμων συμπερασμάτων.



Εικόνα 12: Προφίλ ταχυτήτων

Όπως είναι ήδη γνωστό η μέγιστη ταχύτητα σε κανάλι με στερεά όρια εντοπίζεται στο κέντρο του καναλιού όπου οι τριβές από τα τοιχώματα είναι οι ελάχιστες δυνατές (εικ. 12). Οι συντελεστές συσχέτισης μας δείχνουν πως και η ελάχιστη επιρροή σε μία οποιαδήποτε ταχύτητα εντός των ορίων του καναλιού, μπορεί να προκαλέσει αντίστοιχα ελάχιστη ή και μεγαλύτερη επιρροή σε όλες τις υπόλοιπες τιμές. Μικρή αλλαγή στα τοιχώματα του καναλιού δημιουργεί αλλαγή και στις μέγιστες ταχύτητες που εντοπίζονται στο κέντρο του καναλιού. Τέλος από τα διαγράμματα φαίνεται πως στις περιπτώσεις που υπάρχει μαγνητικό πεδίο στα τοιχώματα του καναλιού, υπάρχει και πιο ομαλή μετάβαση από τους κίτρινους στους μπλε χρωματισμούς. Αυτό σημαίνει πως η μείωση στις τιμές των συντελεστών συσχέτισης ακολουθεί μία πιο ομαλή μεταβολή. Τα διαγράμματα στα οποία απουσιάζει το μαγνητικό πεδίο, εμφανίζουν εντονότερη χρωματική ασυνέχεια. Το φαινόμενο αυτό μπορεί να σημαίνει πως η επιρροή της έντασης του μαγνητικού πεδίου είναι υπαρκτή, και μάλιστα έντονη διότι τα διαγράμματα που αφορούν μετρήσεις με ύπαρξη μαγνητικού πεδίου δεν εμφανίζουν χρωματικές ασυνέχειες γύρω από τη θέση εκείνη ( κελί (32,32)). Στη θέση εκείνη υπάρχει σίγουρα μια διαφορετική δυναμική συμπεριφορά του μοντέλου από ότι σε άλλα σημεία που βρίσκονται γύρω από την κύρια διαγώνιο.

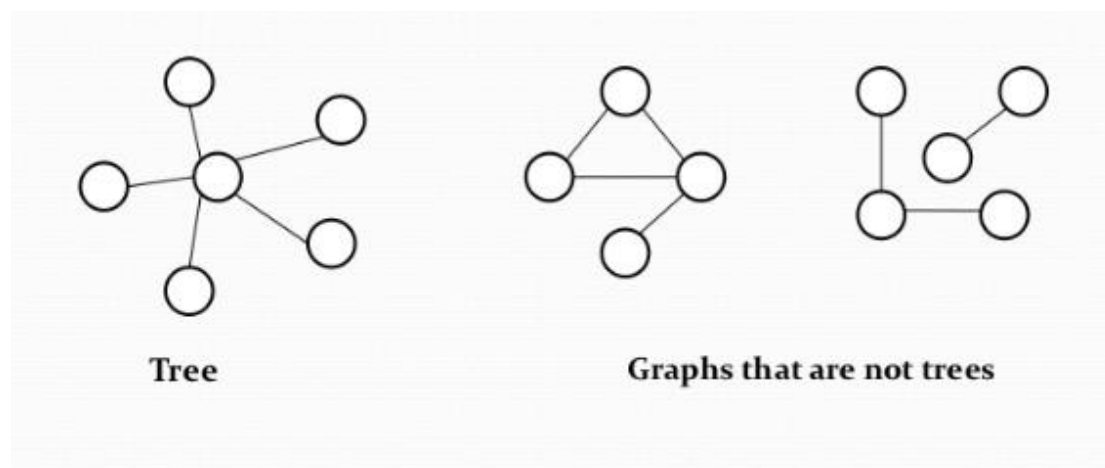
## ΚΕΦΑΛΑΙΟ 3 ΕΥΡΕΣΗ MINIMUM SPANNING TREES

### 3.1 ΕΙΣΑΓΩΓΗ

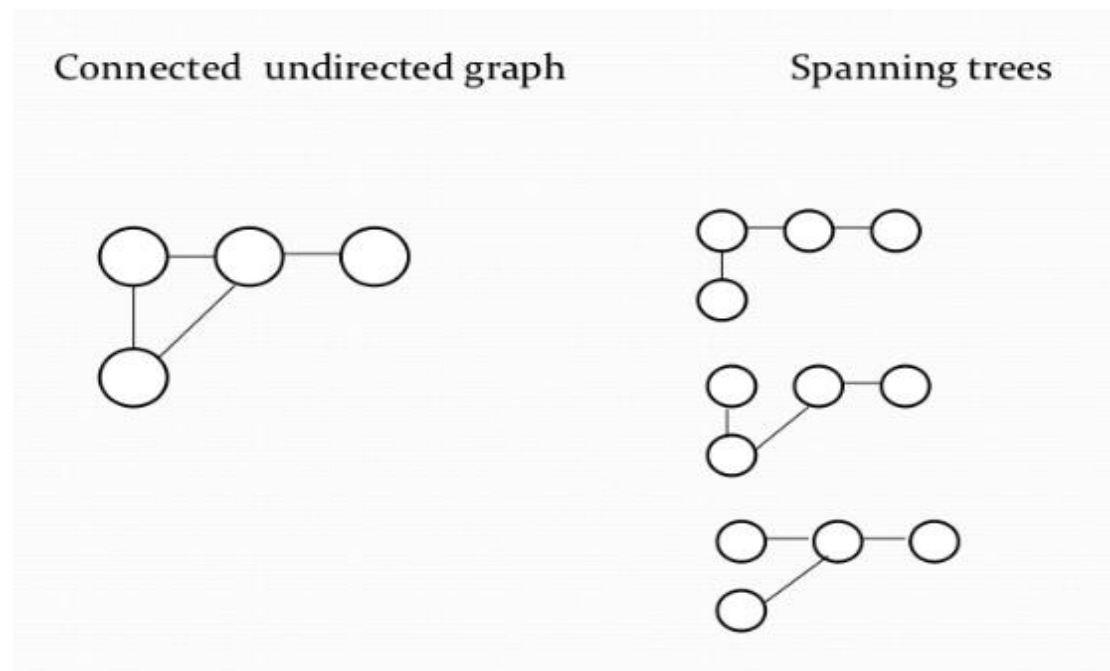
Μετά την εύρεση των συντελεστών Pearson και την εύρεση της συσχέτισης μεταξύ των κανονικοποιημένων τιμών των ταχυτήτων, έγινε η δημιουργία των γραφημάτων-δέντρων ή όπως είναι ευρέως γνωστά των γραφημάτων spanning trees. Αυτά τα διαγράμματα μας δείχνουν πως συνδέονται κάποια δεδομένα ή και κάποιες μεταβλητές μεταξύ τους αλλά και με τι αποστάσεις χωρίζονται αυτά τα δεδομένα. Ονομάζονται spanning trees διότι η μορφή τους παραπέμπει σε δέντρο με πολλά διαφορετικά κλαδιά (Attenborough et al, 2003). Για να θεωρηθεί ένα διάγραμμα διάγραμμα δέντρου θα πρέπει να πληρεί δύο πολύ βασικές προϋποθέσεις:

- δεν εντοπίζονται κενά στο διάγραμμα (είναι δυνατές όλες οι διαδρομές μεταξύ των δεδομένων)
- δεν υπάρχουν κλειστές κυκλικές διαδρομές (βρόγχοι)

Η διαφορά τους από απλά διαγράμματα σύνδεσης δεδομένων φαίνεται στις παρακάτω εικόνες:



Εικόνα 13: Spanning tree



Εικόνα 14: Spanning tree

Όπως φαίνεται και παραπάνω (εικ. 13,14) για το ίδιο πλήθος δεδομένων μπορούν να υπάρξουν διαφορετικές διατάξεις στο σχεδιασμό του spanning tree. Αυτό συνεπάγεται πως δεν υπάρχει ένα μοναδικό σωστό διάγραμμα-δέντρο για μία ομάδα δεδομένων απλώς η κάθε διάταξη έχει διαφορετικές προτεραιότητες και διαφορετικές αποστάσεις. Επομένως, για κάθε είδους πρόβλημα που δέχεται αντιμετώπισης, επιλέγεται το spanning tree που εξυπηρετεί καλύτερα τις παραμέτρους κάτω από τις οποίες πρέπει να εντοπιστεί η λύση. Αν υπάρχουν  $n$  μεταβλητές ή δεδομένα που πρέπει να δημιουργήσουν ένα spanning tree, τότε οι ακμές του διαγράμματος είναι ίσες με  $n-1$  (Attenborough et al, 2003).

### 3.2 ΕΦΑΡΜΟΓΕΣ ΤΩΝ SPANNING TREES

Τα spanning trees έχουν πολλές εφαρμογές και επεκτείνονται σε πολλές υποκατηγορίες του επιστημονικού κλάδου. Αυτό συμβαίνει διότι αποτελούν ένα πολύ απλό και ευκατανόητο διάγραμμα προς όλους. Επιπλέον, είναι μια μορφή οργανογράμματος που παρουσιάζει τη σειρά των εργασιών και τις προτεραιότητες που θέτονται σε ένα οποιοδήποτε project. Παρακάτω παρουσιάζεται ένα απλό πρόβλημα για το οποίο με τη χρήση της μεθοδολογίας των spanning trees δίνεται η βέλτιστη λύση.

Το πρόβλημα αφορά την εγκατάσταση ενός τηλεφωνικού δικτύου σε ένα συγκρότημα νοικοκυριών που βρίσκονται σε συγκεκριμένες αποστάσεις μεταξύ τους όπως φαίνεται στην παρακάτω εικόνα (εικ. 15):





Εικόνα 15: Συγκρότημα με οικίες

Σκοπός είναι το πέρασμα της τηλεφωνικής γραμμής να συνδέει τις οικίες με το κεντρικό γραφείο, που βρίσκεται στο κέντρο, το οποίο δίνει και το σήμα για τις τηλεπικοινωνίες. Παρακάτω φαίνεται μια πρώτη προσέγγιση:



Εικόνα 16: Πρώτη προσπάθεια σύνδεσης

Η παραπάνω (εικ. 16), αποτελεί λύση στο πρόβλημα αφού όπως φαίνεται η τηλεφωνική γραμμή συνδέει όλες τις κατοικίες με το κεντρικό γραφείο. Η τηλεφωνική γραμμή δεν συνδέει παρά μόνο κάθε οικία με το κεντρικό γραφείο και δεν υπάρχει άμεση σύνδεση μεταξύ των οικιών. Στη συνέχεια, φαίνεται και μία

λύση που ακολουθεί διαφορετική λογική, έτσι ώστε να μπορεί να γίνει σύγκριση και επιλογή της καλύτερης λύσης.



Εικόνα 17: Δεύτερη προσπάθεια σύνδεσης

Η ειδοποιός διαφορά που εντοπίζεται μεταξύ των δύο λύσεων είναι η σύνδεση μεταξύ οικιών. Στην πρώτη περίπτωση δεν υφίσταται τέτοια σύνδεση. Στη δεύτερη όμως (εικ. 17), βλέπει κανείς πως άμεση σύνδεση με το κεντρικό γραφείο έχουν μόνο τρεις οικίες. Οι υπόλοιπες συνδέονται ανά περιοχή με αυτές τις τρεις και λαμβάνουν με αυτό τον τρόπο τηλεφωνικό σήμα. Αυτή η λογική ακολουθείται μέχρι και σήμερα αφού η έμμεση σύνδεση και η σύνδεση ανά περιοχή μειώνει κατά πολύ το μήκος της συνολικής καλωδίωσης. Με τη σειρά της η λιγότερη καλωδίωση μειώνει κατά πολύ και το κόστος του έργου. Επομένως πρόκειται για βέλτιστη λύση. Με το ίδιο σκεπτικό που αναπτύχθηκε παραπάνω λειτουργούν όλες οι οικιακές συνδέσεις (ύδρευση, ηλεκτρική κ.α.).

Άλλες εφαρμογές των *spanning trees* εντοπίζονται σε διαδικασίες ελαχιστοποίησης του κόστους (Bergantiños, Navarro-Ramos, 2019) καθώς και στον προσδιορισμό κατανομής των τουριστών ανά περιοχή (Wei et al, 2018).

### 3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΩΝ SPANNING TREES

Μετά τους υπολογισμούς των συντελεστών συσχέτισης μεταξύ των ταχυτήτων κάθε διεύθυνσης με και χωρίς την ύπαρξη μαγνητικού πεδίου, οι παραπάνω συντελεστές χρησιμοποιήθηκαν για τη δημιουργία των *Minimum Spanning Trees*. Πρόκειται ουσιαστικά για διαγράμματα τα οποία μας δείχνουν τις ελάχιστες αποστάσεις μεταξύ των δεδομένων που περιέχονται, στην προκειμένη περίπτωση, στη μορφή πινάκων (Attenborough et al, 2003). Το μοντέλο των *Spanning Trees* χρησιμοποιεί τις αποστάσεις και τους συντελεστές συσχέτισης μεταξύ δεδομένων.

Ένα spanning tree αποτελεί τη σχηματική απεικόνιση ενός πλήθους δεδομένων με αριθμό ακμών κατά ένα μικρότερο του πλήθους των δεδομένων και το μικρότερο δυνατό άθροισμα μήκους των ακμών. Επιπλέον στα spanning trees απουσιάζουν οι κλειστές διαδρομές (loops) ενώ η απώλεια μίας ακμής οδηγεί και στο “σπάσιμο” του σχήματος σε δύο επιμέρους σχήματα. Στην πραγματικότητα οι αποστάσεις μπορούν να αναπαριστούν πραγματικές αποστάσεις, κυκλοφοριακή συμφόρηση ή γενικότερα φορτίο σε κάποιο σύστημα. Τα spanning trees και η εφαρμογή τους συναντώνται κυρίως σε σχεδιασμούς δομικών ή γενικότερα έργων του πολιτικού μηχανικού (δημιουργία χρονοδιαγράμματος έργου), σε προγραμματισμούς δικτύων από ηλεκτρολόγους μηχανικούς (επικοινωνία δικτυακών κόμβων) και σε αναλύσεις ομάδων που αποτελούνται από ομοειδή μέλη (πχ. συγκροτήματα κτιρίων). Η επεξεργασία των δεδομένων και η δημιουργία των διαγραμμάτων έγινε σε περιβάλλον Matlab.

Αρχικά για την διεκπεραίωση της παραπάνω διαδικασίας, χρησιμοποιήθηκαν τα αρχεία των πινάκων των συντελεστών συσχέτισης όπως αυτά είχαν υπολογιστεί από το προηγούμενο μέρος της εργασίας. Επομένως έγινε εισαγωγή έξι στοιχείων, με και χωρίς μαγνητικό πεδίο για τις τρεις διαστάσεις του χώρου, στην επιφάνεια **workspace** του Matlab. Στη συνέχεια έπρεπε να χρησιμοποιηθούν οι εξισώσεις που ορίζει η θεωρία για εύρεση των αποστάσεων μεταξύ των συντελεστών συσχέτισης για τη δημιουργία των spanning trees. Σύμφωνα με τη θεωρία, η ευκλείδεια απόσταση μεταξύ δύο στοιχείων ορίζεται από την παρακάτω εξίσωση:

$$d_{ij}^2 = \|\tilde{s}_i - \tilde{s}_j\|^2 = \sum_{k=1}^n (\tilde{s}_{ik} - \tilde{s}_{jk})^2 \quad (5)$$

Η παραπάνω εξίσωση (εξ. 5) αποτελεί ουσιαστικά μια γενίκευση του Πυθαγορείου θεωρήματος σε περισσότερες από δύο συντεταγμένες. Έπειτα από επεξεργασία η εξίσωση της απόστασης παίρνει την παρακάτω μορφή:

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \quad (6)$$

Τα στοιχεία  $\rho_{ij}$  (εξ.6) αφορούν τα στοιχεία των πινάκων με τους συντελεστές συσχέτισης. Μέσα από την εύρεση των αποστάσεων γίνεται πλέον εφικτή η δημιουργία των διαγραμμάτων spanning trees. Είναι ουσιαστικά διαγράμματα σε μορφή δέντρου όπου οι αποστάσεις μπορούν να ερμηνευθούν και ως το μήκος του κάθε κλαδιού αυτού του δέντρου. Με τη δημιουργία των παραπάνω διαγραμμάτων βρίσκουμε όχι μόνο τις αποστάσεις μεταξύ των στοιχείων, αλλά και τις διαδρομές και τις συνδέσεις που υπάρχουν μεταξύ τους. Για τα δεδομένα που συναντώνται στη μορφή πινάκων ακολουθείται ένας αλγόριθμος. Στον αλγόριθμο περιέχεται μια σειρά ενεργειών με την επανάληψη των οποίων, μπορεί να δημιουργηθεί το αντίστοιχο spanning tree για οποιοδήποτε μέγεθος πίνακα χρειάζεται. Η διαδικασία φαίνεται στη συνέχεια με το παρακάτω παράδειγμα:

Αν ζητείται το διάγραμμα spanning tree για τον πίνακα (εικ. 18):

	U	V	W	X	Y	Z
U	-	27	14	12	13	30
V	27	-	21	23	25	31
W	14	21	-	29	16	18
X	12	23	29	-	20	24
Y	13	25	16	20	-	17
Z	30	31	18	24	17	-

Εικόνα 18: Πίνακας στοιχείων

Ακολουθούνται με τη σειρά τα παρακάτω βήματα:

1. Επιλογή οποιασδήποτε στήλης και διαγραφή των στοιχείων της ίδιας σειράς εκτός του στοιχείου που ανήκει στην επιλεγμένη στήλη
2. Αναζήτηση ανάμεσα στα στοιχεία των επιλεγμένων στηλών για την εύρεση του μικρότερου στοιχείου και επιλογή του στοιχείου αυτού
3. Διαγραφή της σειράς στην οποία ανήκει το επιλεγμένο στοιχείο (εκτός του ίδιου) και επιλογή της στήλης στην οποία ανήκει
4. Επανάληψη των βημάτων 2 και 3 έως ότου γίνει διαγραφή όλων των σειρών του πίνακα
5. Σχεδιασμός του spanning tree

Μετά το πέρας της παραπάνω διαδικασίας θα έχει γίνει εύρεση των ελάχιστων δυνατών αποστάσεων μεταξύ των στοιχείων του πίνακα και ο πίνακας θα έχει την παρακάτω μορφή (εικ. 19):

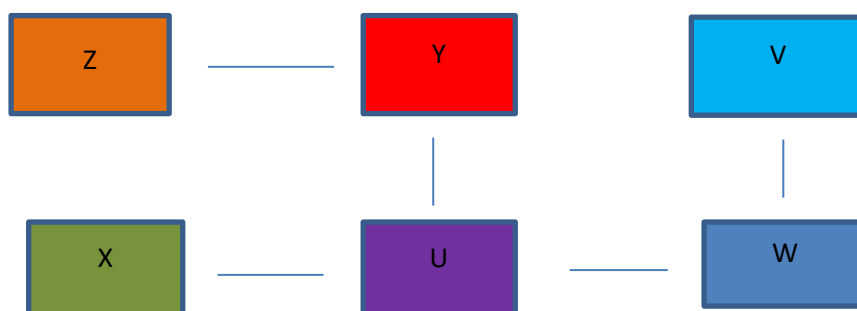
	U	V	W	X	Y	Z
U		<del>27</del>	<del>14</del>	12	<del>13</del>	<del>30</del>
V	<del>27</del>		21	<del>23</del>	<del>25</del>	<del>31</del>
W	14	<del>21</del>		<del>29</del>	<del>16</del>	<del>18</del>
X	<del>12</del>	<del>23</del>	<del>29</del>		<del>20</del>	<del>24</del>
Y	13	<del>25</del>	<del>16</del>	<del>20</del>		<del>17</del>
Z	<del>30</del>	<del>31</del>	<del>18</del>	<del>24</del>	17	

Εικόνα 19: Πίνακας στοιχείων (επεξεργασία)

Οι ελάχιστες αποστάσεις είναι τα στοιχεία που είναι κυκλωμένα και το άθροισμά τους μας δίνει την ελάχιστη δυνατή απόσταση μεταξύ των στοιχείων. Οι αποστάσεις είναι οι εξής:

- XU-12
- UY-13
- UW-14
- YZ-17
- VW-21

Συνολικά το ελάχιστο άθροισμα είναι ίσο με 78 και ο σχεδιασμός του spanning tree φαίνεται παρακάτω:



### 3.4 ΔΗΜΙΟΥΡΓΙΑ MINIMUM SPANNING TREE ΓΙΑ ΤΙΣ ΑΝΑΓΚΕΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Όπως γίνεται αντιληπτό στην περίπτωση των spanning trees είναι εφικτή η ύπαρξη δύο ή και περισσότερων ορθών spanning trees για τα στοιχεία προς επεξεργασία. Η ονομασία 'minimum' δίνεται σε εκείνο του οποίου το άθροισμα των αποστάσεων μεταξύ των στοιχείων είναι το ελάχιστο δυνατό. Επιπλέον, σε πλήθος στοιχείων που είναι σχετικά μικρό, όπως και στην περίπτωση του παραδείγματος, η δημιουργία του spanning tree είναι εφικτό και σχετικά εύκολο να γίνει με το χέρι. Παρόλα αυτά, σε περιπτώσεις μεγάλου αριθμού στοιχείων είναι πολύ ευκολότερο μια τέτοια διαδικασία να γίνεται σε κάποιο υπολογιστικό περιβάλλον. Ειδικότερα, όταν και τα δεδομένα χρειάζονται πριν από αυτό το στάδιο μία πρώιμη επεξεργασία για την κανονικοποίησή τους, διαδικασία που πολλές φορές περιέχει αρκετούς υπολογισμούς.

Στην περίπτωση που αφορά το θέμα της εργασίας μετά από την εξαγωγή των πινάκων συντελεστών συσχέτισης, συνεχίστηκε η επεξεργασία των στοιχείων στο περιβάλλον του Matlab, με τη δημιουργία ενός κώδικα για την εύρεση των αποστάσεων και των διαγραμμάτων. Στον κώδικα χρησιμοποιήθηκαν συναρτήσεις που υπάρχουν στο Matlab και βρίσκουν αυτοματοποιημένα τα στοιχεία που χρειάζονται για την δημιουργία των spanning trees.

```
Runmdist=sqrt(2*(1-Runm));  
Runmspars=sparse(Runmdist);  
UgRunm=tril(Runmspars);  
[RunmTree,pred]=graphminspantree(UgRunm);  
view(biograph(RunmTree,[],'ShowArrows','off','ShowWeights','on'));
```

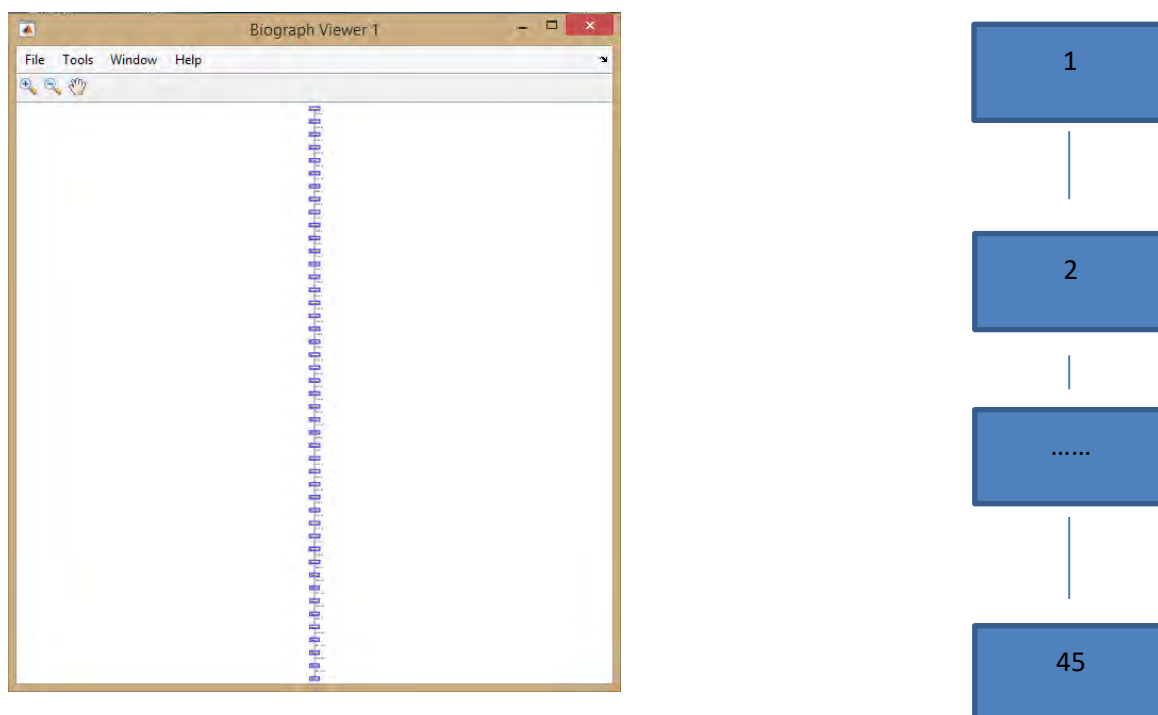
όπου:

- Η συνάρτηση **sqrt()** μας δίνει την τετραγωνική ρίζα
- Η συνάρτηση **sparse()** αφαιρεί από τον εισαγόμενο πίνακα όλα τα μηδενικά στοιχεία (χρησιμοποιείται συνήθως για μείωση απαιτούμενου αποθηκευτικού χώρου των δεδομένων)
- Η συνάρτηση **tril()** χρησιμοποιείται στον πίνακα sparse και επιστρέφει το μέρος του πίνακα που βρίσκεται κάτω από την κύρια διαγώνιο
- Η συνάρτηση **graphminspantree()** δημιουργεί το διάγραμμα spanning tree
- Η συνάρτηση **view()** επιστρέφει στην οθόνη το διάγραμμα που έχει δημιουργηθεί

Με την παραπάνω ακολουθία εντολών υπολογίζονται αρχικά οι αποστάσεις, στη συνέχεια ο πίνακας απαλλάσσεται από περιττά στοιχεία και στη συνέχεια δημιουργείται το spanning tree. Η διαδικασία επαναλήφθηκε έξι φορές με

αντικατάσταση των στοιχείων για τον πίνακα συντελεστών συσχέτισης σε κάθε περίπτωση.

Η απεικόνιση των spanning trees ακολούθησε θα έλεγε κανείς μια γραμμική συμπεριφορά αφού τα στοιχεία συνδέονται με τη σειρά και με ευθεία γραμμή με τον παρακάτω τρόπο:



Εικόνα 20: Spanning tree Rum

Η σύνδεση των κελιών των διαγραμμάτων μεταξύ τους εξαρτάται από το που βρίσκεται το μικρότερο στοιχείο κάθε φορά κατά τη διαδικασία δημιουργίας του spanning tree. Επομένως, το γεγονός πως τα spanning trees των δεδομένων της εργασίας έχουν μορφή ευθείας γραμμής σημαίνει πως κατά τη δημιουργία τους το μικρότερο στοιχείο βρισκόταν πάντοτε στη διπλανή ακριβώς στήλη. Το παραπάνω διάγραμμα αποτελεί το MST (Minimum Spanning Tree) για την διεύθυνση x με την παρουσία μαγνητικού πεδίου. Τα υπόλοιπα 5 διαγράμματα που δημιουργήθηκαν για τις υπόλοιπες περιπτώσεις είναι ίδια και για αυτό το λόγο δεν συμπεριλήφθηκαν στην εργασία για να αποφευχθεί η επανάληψη της ίδιας εικόνας. Μέσω των spanning trees υπολογίστηκαν και οι αποστάσεις η αλλιώς τα "βάρη" μεταξύ των δεδομένων τα οποία φαίνονται στους παρακάτω πίνακες (1,2):

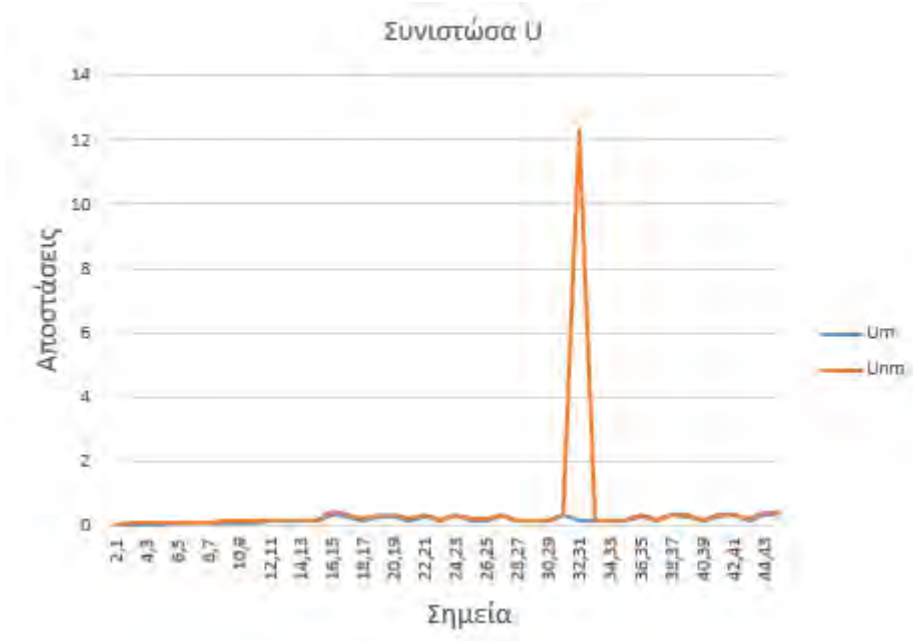
Rum		Rvm		Rwm	
(2,1)	0.0219	(2,1)	0.0649	(2,1)	0.0660
(3,2)	0.0409	(3,2)	0.0940	(3,2)	0.0925
(4,3)	0.0315	(4,3)	0.0554	(4,3)	0.0565
(5,4)	0.0410	(5,4)	0.0611	(5,4)	0.0674
(6,5)	0.0521	(6,5)	0.0681	(6,5)	0.0829
(7,6)	0.0646	(7,6)	0.0769	(7,6)	0.1020
(8,7)	0.0774	(8,7)	0.0874	(8,7)	0.1225
(9,8)	0.0895	(9,8)	0.0985	(9,8)	0.1419
(10,9)	0.0998	(10,9)	0.1085	(10,9)	0.1589
(11,10)	0.1081	(11,10)	0.1178	(11,10)	0.1731
(12,11)	0.1145	(12,11)	0.1264	(12,11)	0.1840
(13,12)	0.1187	(13,12)	0.1327	(13,12)	0.1908
(14,13)	0.1214	(14,13)	0.1370	(14,13)	0.1942
(15,14)	0.1229	(15,14)	0.1401	(15,14)	0.1961
(16,15)	0.3579	(16,15)	0.4139	(16,15)	0.5726
(17,16)	0.2508	(17,16)	0.2888	(17,16)	0.4007
(18,17)	0.1331	(18,17)	0.1495	(18,17)	0.2060
(19,18)	0.2730	(19,18)	0.2984	(19,18)	0.4148
(20,19)	0.2880	(20,19)	0.2955	(20,19)	0.4234
(21,20)	0.1519	(21,20)	0.1516	(21,20)	0.2135
(22,21)	0.2974	(22,21)	0.2963	(22,21)	0.4056
(23,22)	0.1562	(23,22)	0.1520	(23,22)	0.2051
(24,23)	0.3020	(24,23)	0.2963	(24,23)	0.4016
(25,24)	0.1612	(25,24)	0.1495	(25,24)	0.2013
(26,25)	0.1607	(26,25)	0.1488	(26,25)	0.2034
(27,26)	0.3056	(27,26)	0.2969	(27,26)	0.4135
(28,27)	0.1609	(28,27)	0.1511	(28,27)	0.2077
(29,28)	0.1650	(29,28)	0.1473	(29,28)	0.2036
(30,29)	0.1626	(30,29)	0.1448	(30,29)	0.2014
(31,30)	0.3125	(31,30)	0.2732	(31,30)	0.4066
(32,31)	0.1641	(32,31)	0.1366	(32,31)	0.2091
(33,32)	0.1631	(33,32)	0.1361	(33,32)	0.2062
(34,33)	0.1591	(34,33)	0.1373	(34,33)	0.2045
(35,34)	0.1534	(35,34)	0.1391	(35,34)	0.2003
(36,35)	0.2866	(36,35)	0.2710	(36,35)	0.3818
(37,36)	0.1502	(37,36)	0.1341	(37,36)	0.1913
(38,37)	0.3006	(38,37)	0.2621	(38,37)	0.3641
(39,38)	0.2994	(39,38)	0.2651	(39,38)	0.3606
(40,39)	0.1514	(40,39)	0.1334	(40,39)	0.1832
(41,40)	0.2858	(41,40)	0.2544	(41,40)	0.3351
(42,41)	0.3083	(42,41)	0.2445	(42,41)	0.3189
(43,42)	0.1654	(43,42)	0.1239	(43,42)	0.1590
(44,43)	0.3096	(44,43)	0.2460	(44,43)	0.3155
(45,44)	0.3126	(45,44)	0.2432	(45,44)	0.3304

Πίνακας 1: Υπολογισμός αποστάσεων

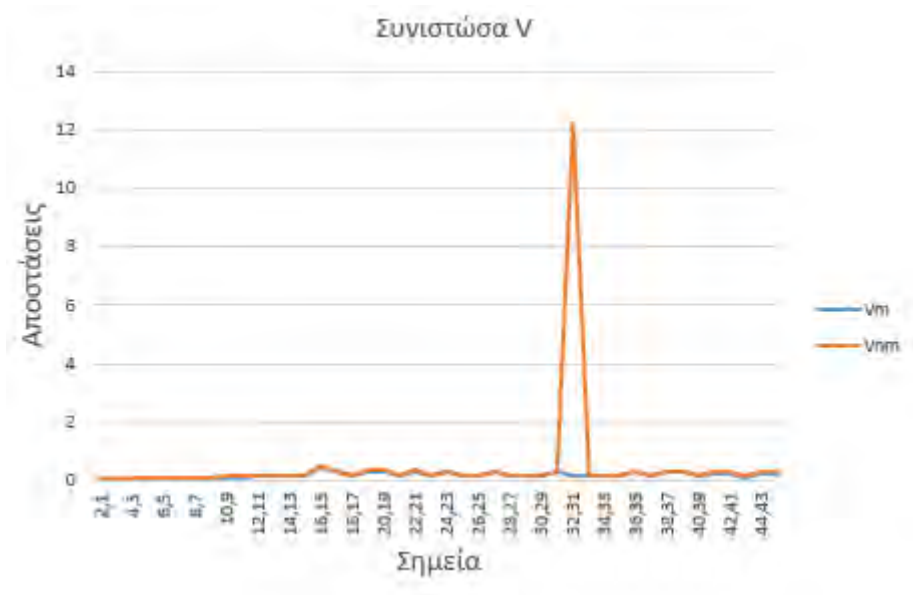


Runm		Rvnm		Rwnm	
(2,1)	0.0447	(2,1)	0.0783	(2,1)	0.0824
(3,2)	0.0754	(3,2)	0.1139	(3,2)	0.1121
(4,3)	0.0533	(4,3)	0.0689	(4,3)	0.0677
(5,4)	0.0656	(5,4)	0.0771	(5,4)	0.0809
(6,5)	0.0787	(6,5)	0.0859	(6,5)	0.0993
(7,6)	0.0917	(7,6)	0.0955	(7,6)	0.1200
(8,7)	0.1040	(8,7)	0.1069	(8,7)	0.1397
(9,8)	0.1155	(9,8)	0.1195	(9,8)	0.1578
(10,9)	0.1254	(10,9)	0.1313	(10,9)	0.1739
(11,10)	0.1336	(11,10)	0.1425	(11,10)	0.1878
(12,11)	0.1408	(12,11)	0.1531	(12,11)	0.1997
(13,12)	0.1470	(13,12)	0.1615	(13,12)	0.2092
(14,13)	0.1530	(14,13)	0.1681	(14,13)	0.2159
(15,14)	0.1582	(15,14)	0.1702	(15,14)	0.2200
(16,15)	0.4716	(16,15)	0.4721	(16,15)	0.6273
(17,16)	0.3404	(17,16)	0.3178	(17,16)	0.4418
(18,17)	0.1767	(18,17)	0.1630	(18,17)	0.2262
(19,18)	0.3521	(19,18)	0.3327	(19,18)	0.4433
(20,19)	0.3525	(20,19)	0.3352	(20,19)	0.4437
(21,20)	0.1788	(21,20)	0.1689	(21,20)	0.2278
(22,21)	0.3369	(22,21)	0.3316	(22,21)	0.4404
(23,22)	0.1702	(23,22)	0.1671	(23,22)	0.2260
(24,23)	0.3360	(24,23)	0.3136	(24,23)	0.4233
(25,24)	0.1766	(25,24)	0.1571	(25,24)	0.2117
(26,25)	0.1752	(26,25)	0.1588	(26,25)	0.2119
(27,26)	0.3391	(27,26)	0.3171	(27,26)	0.4120
(28,27)	0.1662	(28,27)	0.1632	(28,27)	0.2182
(29,28)	0.1653	(29,28)	0.1611	(29,28)	0.2234
(30,29)	0.1645	(30,29)	0.1581	(30,29)	0.2183
(31,30)	0.3113	(31,30)	0.2999	(31,30)	0.4279
(32,31)	1.2280	(32,31)	1.2234	(32,31)	1.2517
(33,32)	0.1745	(33,32)	0.1493	(33,32)	0.2237
(34,33)	0.1735	(34,33)	0.1474	(34,33)	0.2146
(35,34)	0.1699	(35,34)	0.1474	(35,34)	0.2045
(36,35)	0.3170	(36,35)	0.2895	(36,35)	0.4059
(37,36)	0.1623	(37,36)	0.1471	(37,36)	0.2173
(38,37)	0.3222	(38,37)	0.2899	(38,37)	0.4177
(39,38)	0.3225	(39,38)	0.2857	(39,38)	0.4075
(40,39)	0.1645	(40,39)	0.1429	(40,39)	0.2053
(41,40)	0.3278	(41,40)	0.2745	(41,40)	0.4071
(42,41)	0.3430	(42,41)	0.2750	(42,41)	0.4032
(43,42)	0.1851	(43,42)	0.1376	(43,42)	0.2010
(44,43)	0.3709	(44,43)	0.2633	(44,43)	0.3832
(45,44)	0.3684	(45,44)	0.2701	(45,44)	0.3877

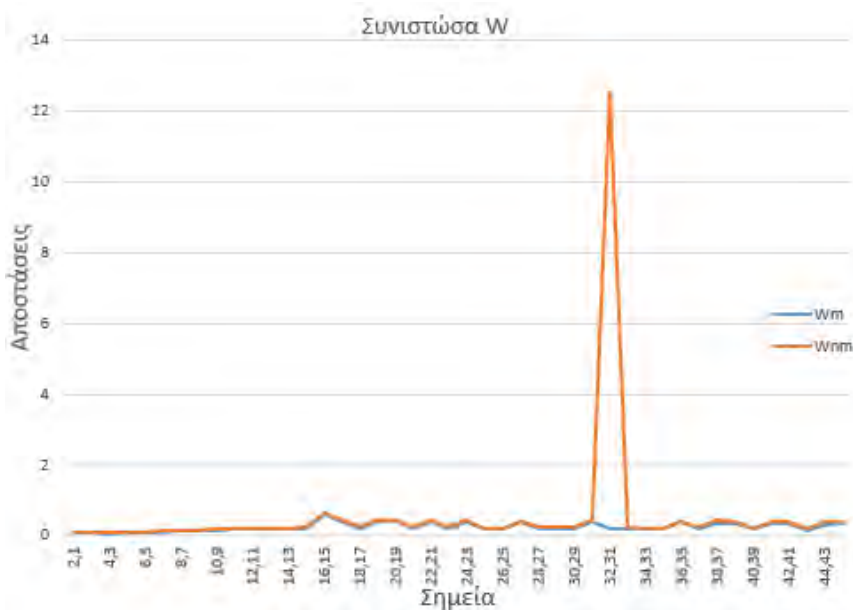
Πίνακας 2: Υπολογισμός αποστάσεων



Εικόνα 21: Διάγραμμα αποστάσεων U



Εικόνα 22: Διάγραμμα αποστάσεων V



Εικόνα 23: Διάγραμμα αποστάσεων W

### 3.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΕΦΑΛΑΙΟΥ- ΦΥΣΙΚΗ ΕΡΜΗΝΕΙΑ

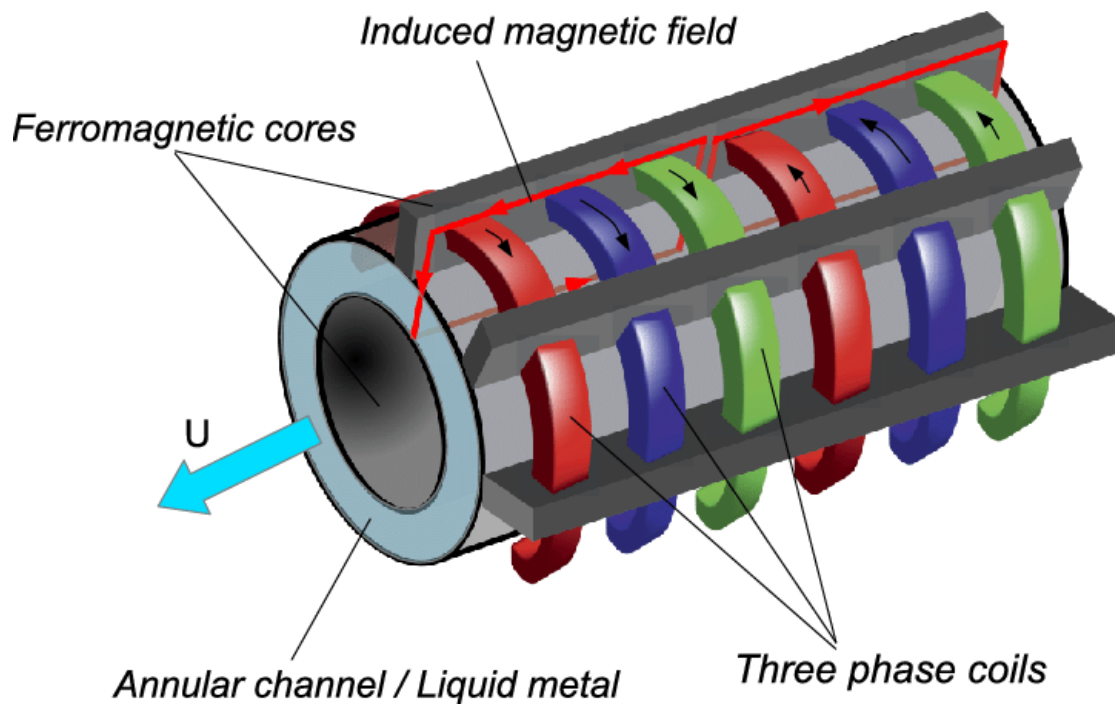
Τα spanning trees που δημιουργήθηκαν έχουν όλα γραμμική μορφή. Αυτό σημαίνει ότι τα σημεία των μετρήσεων βρίσκονται και αυτά σε μία γραμμή. Αυτό το καταλαβαίνει κανείς από το γεγονός ότι οι αποστάσεις που οφείλουν να είναι ελάχιστες, αφού δημιουργείται Minimum spanning tree, ενώνουν πάντοτε συνεχόμενα αριθμητικά σημεία. Αυτό το συμπέρασμα θα μπορούσε να βγάλει κανείς αναλύοντας και τα διαγράμματα συσχέτισης του κεφαλαίου για την εύρεση του συντελεστή Pearson. Οι κοντινές στήλες έχουν αρκετά μεγάλους συντελεστές. Αυτό μετά και τη δημιουργία των spanning trees, επιβεβαιώνει ότι οι μετρήσεις έγιναν σε ευθεία γραμμή. Όπως φαίνεται στις αποστάσεις οι τιμές στις περιπτώσεις που απουσιάζει το μαγνητικό πεδίο είναι μεγαλύτερες από εκείνες που εμφανίζονται με το μαγνητικό πεδίο. Αυτό οδηγεί στο συμπέρασμα πως το μαγνητικό πεδίο που βρίσκεται στα τοιχώματα του καναλιού φέρνει τις κατευθύνσεις των ταχυτήτων πιο κοντά.

Από τα διαγράμματα που έπονται των πινάκων (εικ. 21-23) μπορούν να βγουν σημαντικά συμπεράσματα. Δημιουργήθηκαν τρία διαγράμματα κάθε ένα για την συνιστώσα κάθε διεύθυνσης. Με πορτοκαλί χρώμα φαίνεται η γραμμή της περίπτωσης με απουσία μαγνητικού πεδίου ενώ με την μπλε γραμμή φαίνεται η

γραμμή των αποστάσεων στις περιπτώσεις που υπάρχει το μαγνητικό πεδίο. Η ασυνέχεια που είχε παρατηρηθεί προηγουμένως στα διαγράμματα Pearson στο κελί (32,32) εμφανίζεται και σε αυτά τα διαγράμματα. Πράγματι εάν ανατρέξει κανείς στην τιμή της απόστασης μεταξύ των σημείων 31 και 32 θα δει πως αυτή είναι πολύ μεγάλη και μάλιστα είναι σε άλλη τάξη μεγέθους σε σύγκριση με τις υπόλοιπες τιμές των αποστάσεων. Σε όλες τις υπόλοιπες αποστάσεις μεταξύ των σημείων οι τιμές είναι πανομοιότυπες και αυτός είναι και ο λόγος που τα διαγράμματα αλληλοκαλύπτουν το ένα το άλλο και δεν είναι τόσο εμφανής η μπλε γραμμή. Η απόσταση μεταξύ των σημείων 31 και 32 είναι μεγάλη και πλέον γίνεται ξεκάθαρη και η μειωτική συμπεριφορά που εμφάνιζαν οι συντελεστές συσχέτισης προηγουμένως σε εκείνη την περιοχή.

Από φυσικής άποψης μπορεί να πει κάποιος πως το μαγνητικό πεδίο αναγκάζει τις τροχιές που διαγράφει το ρευστό να ακολουθούν συγκεκριμένες κατευθύνσεις. Για το λόγο αυτό μειώνονται και οι αποστάσεις μεταξύ των ταχυτήτων στις τρεις διευθύνσεις. Οι τρεις συνιστώσες των ταχυτήτων αποκτούν ένα νέο σύστημα αναφοράς που ορίζεται από το ίδιο το μαγνητικό πεδίο. Η φυσική αυτή ιδιότητα των ηλεκτρικά αγώγιμων ρευστών να επηρεάζονται από μαγνητικά πεδία χρησιμοποιείται κυρίως στις ηλεκτρομαγνητικές αντλίες (electromagnetic pumps).

Η αλήθεια είναι πως εντοπίζονται αρκετές διαφορές μεταξύ του προβλήματος που επεξεργάζεται η παρούσα εργασία και της λειτουργίας μιας τυπικής ηλεκτρομαγνητικής αντλίας. Ωστόσο η λογική επηρεασμού του ρευστού από το μαγνητικό πεδίο μπορεί να γίνει εύκολα κατανοητή από την εφαρμογή της ηλεκτρομαγνητικής αντλίας. Στη συνέχεια, (εικ. 21) εμφανίζεται η απεικόνιση μίας τυπικής ηλεκτρομαγνητικής αντλίας και γίνεται επεξήγηση της λειτουργίας της.



Εικόνα 24: Ηλεκτρομαγνητική αντλία

Όπως φαίνεται στην εικόνα, (εικ. 24) ηλεκτρομαγνητική ονομάζεται μία αντλία η οποία γύρω από την κυλινδρική διατομή του αγωγού περιέχει σύρμα. Με τη σειρά του το σύρμα μέσω χρήσης ηλεκτρικού ρεύματος, δημιουργεί μαγνητικό πεδίο συγκεκριμένης έντασης έτσι ώστε να κατευθύνει τη ροή του ηλεκτρικά αγωγίμου ρευστού προς συγκεκριμένη κατεύθυνση. Το σύρμα φαίνεται στην εικόνα 24 με τα χρώματα κόκκινο, πράσινο και μπλε γύρω από τον αγωγό και η ταχύτητα ροής φαίνεται με το γαλάζιο βέλος στα αριστερά της εικόνας.

Σίγουρα η κατανομή της έντασης του μαγνητικού πεδίου στην περίπτωση της αντλίας εμφανίζει πολλές διαφορές με την κατανομή που εμφανίζει το μαγνητικό πεδίο στα τοιχώματα ενός καναλιού. Η βασικότερη διαφορά εντοπίζεται στο γεγονός πως στην περίπτωση της αντλίας το μαγνητικό πεδίο βρίσκεται περιμετρικά και καλύπτει εξ' ολοκλήρου τον αγωγό ενώ στο κανάλι αυτό υπάρχει μόνο στα τοιχώματά του και έχει σημεία ασυνέχειας πάνω και κάτω από τη ροή του καναλιού. Με τον ίδιο όμως τρόπο που στην αντλία οι ταχύτητες αναγκάζονται να ακολουθήσουν συγκεκριμένη ροή, έτσι και στο πρόβλημα της εργασίας το μαγνητικό πεδίο ορίζει κάποιες συγκεκριμένες κατευθύνσεις για κάθε διεύθυνση και "πιέζει" τα μόρια του ρευστού που στην προκειμένη περίπτωση πρόκειται για νερό να τις ακολουθήσουν.

## ΚΕΦΑΛΑΙΟ 4: ΣΥΝΑΡΤΗΣΗ ΔΙΑΣΥΣΧΕΤΙΣΗΣ

### 4.1 ΕΙΣΑΓΩΓΗ

Όπως και στα παραπάνω κεφάλαια έτσι και σε αυτό αναλύεται ακόμη μία διαδικασία στατιστικού περιεχομένου η οποία βοηθά στην κατεύθυνση επίλυσης του προβλήματος που αντιμετωπίζεται στην εργασία. Η διαδικασία αυτή ονομάζεται διασυσχέτιση (cross-correlation) και ενώ παρουσιάζει κοινά χαρακτηριστικά με τους συντελεστές συσχέτισης που εμφανίστηκαν προηγουμένως, παρουσιάζει και αρκετές διαφορές.

Διασυσχέτιση ονομάζεται ένα μέτρο συσχέτισης μεταξύ δύο χρονοσειρών δεδομένων, όταν σε μία εκ των δύο έχει επιβληθεί συγκεκριμένη χρονοκαθυστέρηση (time-lag) (Koenig et al, 1999). Είναι η μέθοδος σύμφωνα με την οποία επιλέγεται η ανάλυση ενός γνωστού στοιχείου που βρίσκεται μέσα σε ένα άγνωστο μίγμα στοιχείων. Σε περιπτώσεις όπου έχουμε συνεχείς συναρτήσεις η συνάρτηση της διασυσχέτισης ορίζεται με τον παρακάτω τύπο:

$$(f * g) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(\tau)g(\tau + t)d\tau \quad (7)$$

Σε θεωρητικό πλαίσιο, η συσχέτιση ή αλλιώς διασυσχέτιση, μπορεί να υπάρχει μεταξύ παρατηρήσεων μιας χρονοσειράς, ή να αποτελεί δείκτη μεταξύ δύο ή και περισσότερων χρονοσειρών. Με τη διαδικασία της διασυσχέτισης (cross-correlation) εκφράζεται ένα μέτρο συσχέτισης μεταξύ δύο χρονοσειρών ( $\rho_{i,j}$ ). Και οι δύο παραπάνω διαδικασίες είναι αδιάστατες και μας δίνουν τιμές συντελεστών από -1 έως +1 ( $\rho \in (-1,1)$ ) όπου με τον αριθμό +1 εκφράζεται η απόλυτη/άμεση συσχέτιση και με τον αριθμό -1 η αντίστροφη συσχέτιση. Οι τιμές προκύπτουν από μια αναλογία μεταβλητών που μετρούνται με την ίδια μονάδα μέτρησης και είναι επίσης αδιάστατες. Λόγω του ότι οι αρχικές χρονοσειρές αποτελούν ουσιαστικά δείγματα, οι τιμές που προκύπτουν από διαδικασίες συσχέτισης και διασυσχέτισης σε μη πεπερασμένες σειρές αποτελούν εκτιμήσεις και δεν είναι ποτέ ακριβείς.

Εφόσον υπογραμμίστηκαν οι διαφορές μεταξύ της απλής συσχέτισης και της διασυσχέτισης παρακάτω φαίνονται οι τύποι που ακολουθούνται για την εξαγωγή τιμών διασυσχέτισης (cross correlation):

$$\rho_{i,j} = \frac{\gamma_{i,j}}{\sqrt{\sigma_i^2 \sigma_j^2}} \quad (8)$$

όπου:

- $\gamma_{i,j}$  η τιμή της συνδιακύμανσης

- $\sigma_i^2, \sigma_j^2$  οι τιμές διασποράς των δύο μεταβλητών

Η διασυσχέτιση μεταξύ δύο χρονοσειρών  $x_i$  και  $x_j$  καθορίζεται από τη σχέση μεταξύ της συνδιακύμανσης και των ξεχωριστών διακυμάνσεων της κάθε χρονοσειράς όπως φαίνεται στον παραπάνω τύπο. Όπου  $\gamma_{i,j}$  η τιμή της συνδιακύμανσης και δίνεται από την παρακάτω εξίσωση:

$$\hat{\gamma}_{i,j} = \frac{1}{N} \sum_{t=1}^N [(x_i^t - \bar{x}_i)(x_j^t - \bar{x}_j)] \quad (9)$$

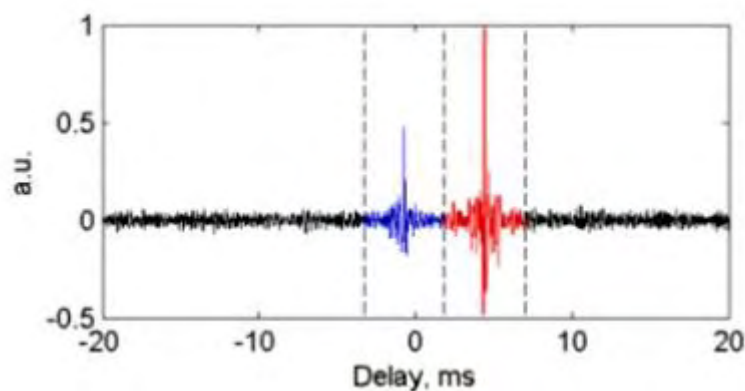
Η εξίσωση είναι ουσιαστικά ίδια με εκείνη της απλής διακύμανσης απλώς αφορά δύο μεταβλητές. Με ανάλυση των παραπάνω εξισώσεων, η αναλυτική μορφή της εξίσωσης της διασυσχέτισης είναι η εξής:

$$\hat{\rho}_{i,j} = \frac{\sum_{t=1}^N [(x_i^t - \bar{x}_i)(x_j^t - \bar{x}_j)]}{\sqrt{\sum_{t=1}^N (x_i^t - \bar{x}_i)^2 \sum_{t=1}^N (x_j^t - \bar{x}_j)^2}} \quad (10)$$

## 4.2 ΕΦΑΡΜΟΓΕΣ

Η διαδικασία της διασυσχέτισης γνωρίζει πολλές εφαρμογές σε πραγματικά προβλήματα της καθημερινότητας του ανθρώπου. Η εφαρμογή της εντοπίζεται σε ανιχνευτές σήματος, στα προγνωστικά δελτία καιρού, στο σχεδιασμό συστημάτων ασφαλείας κ.α. Στη συνέχεια, αναλύεται η εφαρμογή της διασυσχέτισης (cross-correlation) στον έλεγχο και την “αποτύπωση” της συμφόρησης που δημιουργείται από την κίνηση σκαφών στο νερό. Κάθε πλωτό σκάφος όπως και γενικότερα κάθε συσκευή έχει κάποιο ακουστικό αποτύπωμα που το διαχωρίζει από τα υπόλοιπα. Τα παραγόμενα ηχητικά κύματα μπορούν να μετρηθούν εξ’ αποστάσεως από αισθητήρες και οι μετρήσεις να χρησιμοποιηθούν για την ιχνηλάτιση του κάθε σκάφους. Το ηχητικό αποτύπωμα που ταξιδεύει υποθαλάσσια δημιουργείται λόγω της προπέλας του σκάφους, τα μηχανικά του στοιχεία και του σχήματος της καμπίνας (Fillinger et al, 2009). Η μέτρηση γίνεται υποθαλάσσια καθώς το νερό είναι πυκνότερο μέσο από τον αέρα και μπορούν να καταγραφούν περισσότερες συχνότητες σε αυτό. Τα ηχητικά κύματα και η χρονοκαθυστέρηση που υπάρχει κατά τη διάδοσή τους χρησιμοποιούνται ως δεδομένα για εύρεση διασυσχέτισης. Η χρονοκαθυστέρηση ανάλογα και με το που βρίσκεται ο δέκτης των σημάτων μπορεί

να δώσει πληροφορίες για την κατεύθυνση του σκάφους. Στην παρακάτω εικόνα (εικ.11) φαίνεται ένα διάγραμμα που αναπαριστά τη διαφορά φάσης μεταξύ δύο σημάτων που ελήφθησαν από δύο διαφορετικά σκάφη.



Εικόνα 25: Διάγραμμα διαφοράς φάσης

Ο οριζόντιος άξονας αφορά χρόνο και ο κάθετος καθυστέρηση μεταξύ των δύο σημάτων.

Η χωρική πυκνότητα των σημάτων σχετίζεται άμεσα με την χωρική πυκνότητα των διασυσχετίσεών τους. Ενώ πολλές φορές τα σήματα από διαφορετικά σκάφη φτάνουν την ίδια χρονική στιγμή στους αισθητήρες, μέσω της διασυσχέτισης και της χρονικής καθυστέρησης που αυτή επιβάλλει γίνεται εφικτός και ο διαχωρισμός τους. Έπειτα από το διαχωρισμό γίνεται ταυτοποίηση του σήματος και ανάλογα με αυτή μπορεί να γίνει εφικτός ο υπολογισμός της απόστασης μεταξύ των σκαφών που βρίσκονται σε κίνηση. Η μέθοδος αυτή αν και αποτελεσματική δεν είναι αποδοτική σε πολυσύχναστους κόμβους με πολλά σκάφη λόγω του μεγάλου πλήθους των σημάτων. Σε αυτές τις περιοχές χρησιμοποιείται η πιο εξελιγμένη μορφή διασυσχέτισης.

#### 4.3 ΕΦΑΡΜΟΓΗ ΣΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΡΓΑΣΙΑΣ

Σε αυτό το στάδιο της εργασίας και εφόσον έχουν υπολογιστεί οι συντελεστές Pearson για τις μετρήσεις των ταχυτήτων και έχουν δημιουργηθεί τα spanning trees για κάθε τύπο διαφορετικής μέτρησης, υπολογίστηκαν οι συντελεστές διασυσχέτισης (cross correlation). Έπειτα δημιουργήθηκαν εκ νέου διαγράμματα, παρόμοια με εκείνα που είχαν αναπτυχθεί για τους συντελεστές Pearson, και όπως θα φανεί και στη συνέχεια εμφανίστηκαν αρκετές ομοιότητες ως προς τα διαγράμματα και τον κώδικα, εντοπίστηκαν όμως και αρκετές διαφορές. Γίνεται κατανοητό πως η συσχέτιση μεταξύ των μετρήσεων, που έχουν πραγματοποιηθεί για τις ταχύτητες, είναι πολύ σημαντική διότι με αυτή θα μπορέσει να δημιουργηθεί ένα προφίλ συμπεριφοράς της ταχύτητας σε κάθε διεύθυνση. Επομένως, θα γίνει εφικτή και η εξαγωγή συμπερασμάτων για πιθανούς τρόπους με τους οποίους θα



μπορούν οι ταχύτητες να επηρεαστούν και να εξυπηρετήσουν τυχόν μηχανικούς σκοπούς για τον άνθρωπο. Με τον ίδιο τρόπο που επηρεάζει το μαγνητικό πεδίο τη συσχέτιση μεταξύ των ταχυτήτων όπως φάνηκε στους συντελεστές Pearson, έτσι με πειραματισμούς όσο αφορά την ένταση και την σωστή τοποθέτηση του μαγνητικού πεδίου σε κανάλι, μπορεί ο άνθρωπος να επηρεάσει σημειακά και συνολικά τη συμπεριφορά των ταχυτήτων για οποιονδήποτε σκοπό. Οι υπολογισμοί και η επεξεργασία έγινε και πάλι στο περιβάλλον του Matlab.

Για την επεξεργασία των χρονοσειρών των ταχυτήτων και την εύρεση των δεικτών διασυσχέτισης δημιουργήθηκε κώδικας στο λογισμικό του Matlab. Αρχικά, εμφανίστηκε η ανάγκη κανονικοποίησης των δεδομένων για ακόμα μία φορά. Για λόγους ευκολίας και ταχύτητας, έγινε αντιγραφή των κανονικοποιημένων τιμών των ταχυτήτων από το πεδίο εργασίας (**workspace**) των απλών συντελεστών συσχέτισης του δεύτερου κεφαλαίου. Για τη δυνατότητα χρήσης του ίδιου κώδικα σε όλες τις περιπτώσεις, έπρεπε να γίνει εισαγωγή των τιμών των μεταβλητών των κανονικοποιημένων ταχυτήτων σε μία νέα μεταβλητή ( $s_1, s_2, \dots, s_n$ ). Παρακάτω φαίνονται ενδεικτικά οι εντολές εισαγωγής τιμής μεταβλητών:

```
s1=kanonwlmCopy;
```

Στην εντολή που προηγείται, φαίνεται ενδεικτικά το παράδειγμα για τις ταχύτητες της διεύθυνσης  $w$  στην περίπτωση ύπαρξης μαγνητικού πεδίου στη ροή. Οι ίδιες εντολές έγιναν και για τις 45 χρονοσειρές της συγκεκριμένης κατηγορίας και εδώ εμφανίζονται οι πρώτες τιμές για λόγους συντομίας. Η ένδειξη "Copy" μετά το τέλος κάθε τιμής δημιουργήθηκε αυτόματα από το λογισμικό του Matlab για να ειδοποιήσει τον χρήστη πως η τιμή αυτή αποτελεί αντίγραφο από άλλο φύλλο εργασίας (workspace του συντελεστή συσχέτισης). Η κανονικοποίηση είναι ακριβώς η ίδια που έγινε και στο στάδιο των συντελεστών Pearson με τη συνάρτηση "zeros" του Matlab.

Στη συνέχεια, για κάθε μια από τις 45 χρονοσειρές χρησιμοποιήθηκε η εντολή του Matlab "xcorr" η οποία αφορά συγκεκριμένα τη διασυσχέτιση μεταξύ μεταβλητών και ακολουθήθηκε η παρακάτω σύνταξη:

```
[C1_1,lag1_1] = xcorr(s1,s1,'coeff');
```

Στο βήμα αυτό γίνεται ουσιαστικά η σύγκριση μεταξύ των χρονοσειρών και η εύρεση της "καθυστέρησης" (lag). Στις τιμές τύπου "C1\_1, C1\_2,..." εισάγονται οι τιμές των συντελεστών διασυσχέτισης και στις τιμές "lag" εισάγονται οι καθυστερήσεις ή αλλιώς μετακινήσεις που γίνονται σε κάθε περίπτωση. Η συνάρτηση "xcorr" σε αυτή την περίπτωση μετράει τη συσχέτιση που υπάρχει

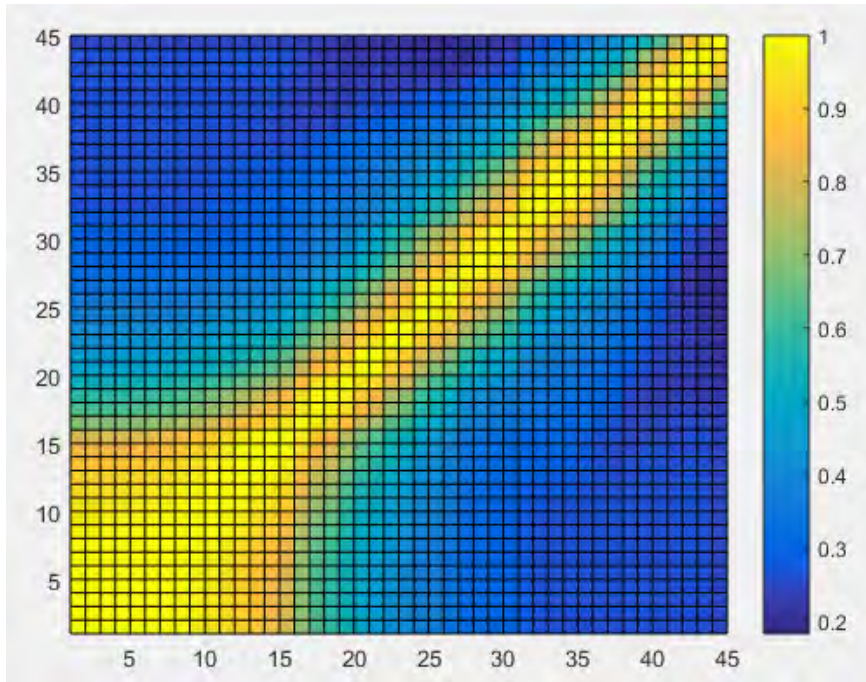
μεταξύ της πρώτης χρονοσειράς με τον εαυτό της καθώς και με τις υπόλοιπες 44 όταν αυτές έχουν υποστεί μία μεταφορά θέσεων στις τιμές τους ή αλλιώς καθυστέρηση (lag). Στην περίπτωση που οι δύο συγκρίσιμες τιμές έχουν διαφορετικό μήκος τότε η συνάρτηση τοποθετεί αυτόματα την τιμή 0 στις θέσεις που υπολείπονται στη μικρότερη χρονοσειρά.

Έπειτα από την εύρεση των τιμών συσχέτισης μεταξύ των χρονοσειρών αλλά και των αντίστοιχων lags, δημιουργήθηκε κώδικας για την εύρεση των θέσεων όπου υπάρχουν οι μέγιστες τιμές των συντελεστών συσχέτισης.

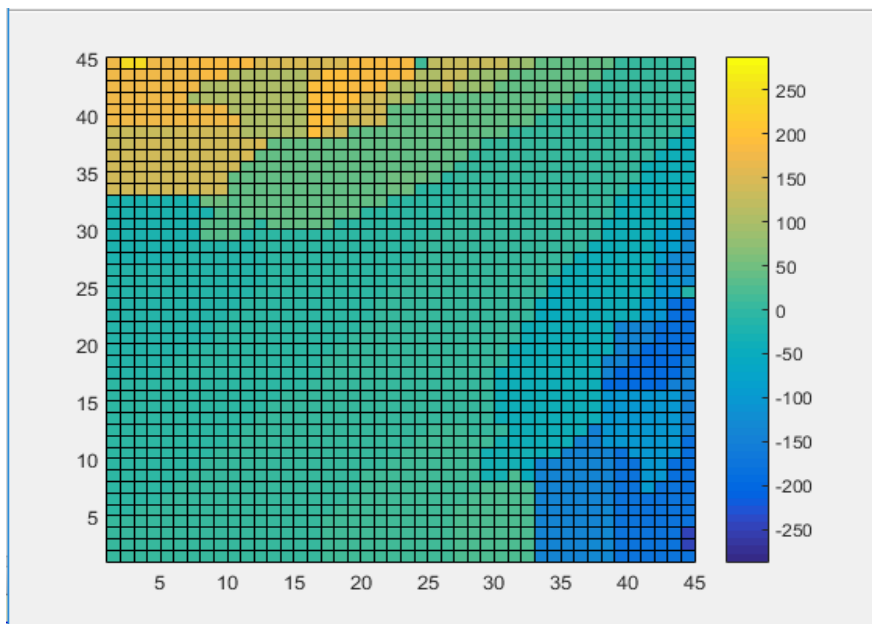
```
[M1_1,I1_1] = max(C1_1); t1_1 = lag1_1(I1_1);
```

Η συνάρτηση “[M1\_1,I1\_1]=max();” εντοπίζει τις θέσεις των μέγιστων τιμών των συντελεστών συσχέτισης και κάνει την εισαγωγή τους σε ένα διάνυσμα I1\_1. Εάν η μέγιστη τιμή εμφανίζεται περισσότερες από μία φορές τότε διατηρείται η θέση στην οποία η τιμή αυτή εμφανίστηκε πρώτη. Στη συνέχεια, οι χρονικές διαφορές κρατούνται σε μεταβλητές με την ονομασία t1\_1,t1\_2 κλπ. Η διαδικασία όπως και παραπάνω έγινε για όλες τις 45 χρονοσειρές. Οι τιμές των χρονικών διαφορών κρατήθηκαν σε πίνακα (ti=[t1\_1, t1\_2, t1\_3, t1\_4, t1\_5,...).

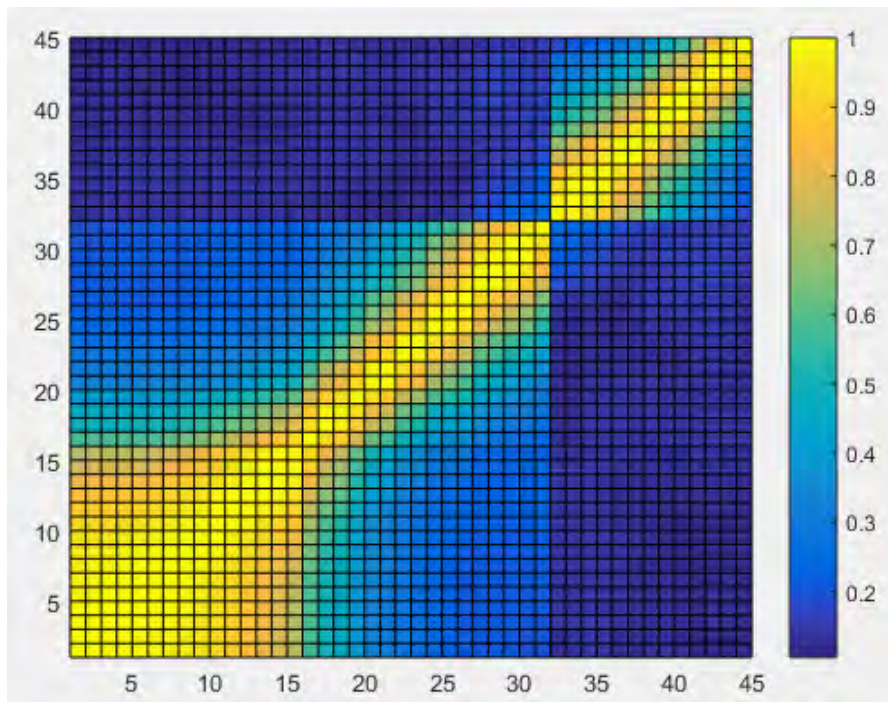
Η παραπάνω διαδικασία ακολουθείται και πραγματοποιεί τη σύγκριση μίας χρονοσειράς με τις υπόλοιπες και δίνει τις αντίστοιχες τιμές συντελεστών συσχέτισης. Συγκεκριμένα, οι παραπάνω εντολές αφορούν τη σύγκριση της χρονοσειράς νούμερο 1 με τον εαυτό της και τις υπόλοιπες 44. Οι ίδιες εντολές εκτελούνται για όλες τις υπόλοιπες μεταβλητές για την εξαγωγή όλων των τιμών συσχέτισης όλων των πιθανών συνδυασμών χρονοσειρών. Επομένως, μετά από επανάληψη των εντολών για όλες τις 45 χρονοσειρές, οι τιμές των συντελεστών συσχέτισης τοποθετούνται σε έναν πίνακα που, όπως είναι αναμενόμενο, έχει διαστάσεις 45x45 και του οποίου τα στοιχεία της κύριας διαγωνίου είναι ίσα με τη μονάδα. Ο πίνακας αυτός παρουσιάζει αρκετές ομοιότητες με τους πίνακες των συντελεστών Pearson που είχαν δημιουργηθεί σε προηγούμενη διαδικασία κάτι που φαίνεται και από τα διαγράμματα που δημιουργήθηκαν με τη συνάρτηση **pcolor()**. Τα διαγράμματα παρουσιάζονται παρακάτω (εικ 26-37). Επιπλέον, παρακάτω φαίνονται και τα διαγράμματα του πίνακα ti με τους χρονικούς συντελεστές κάθε σύγκρισης χρονοσειρών και για κάθε περίπτωση.



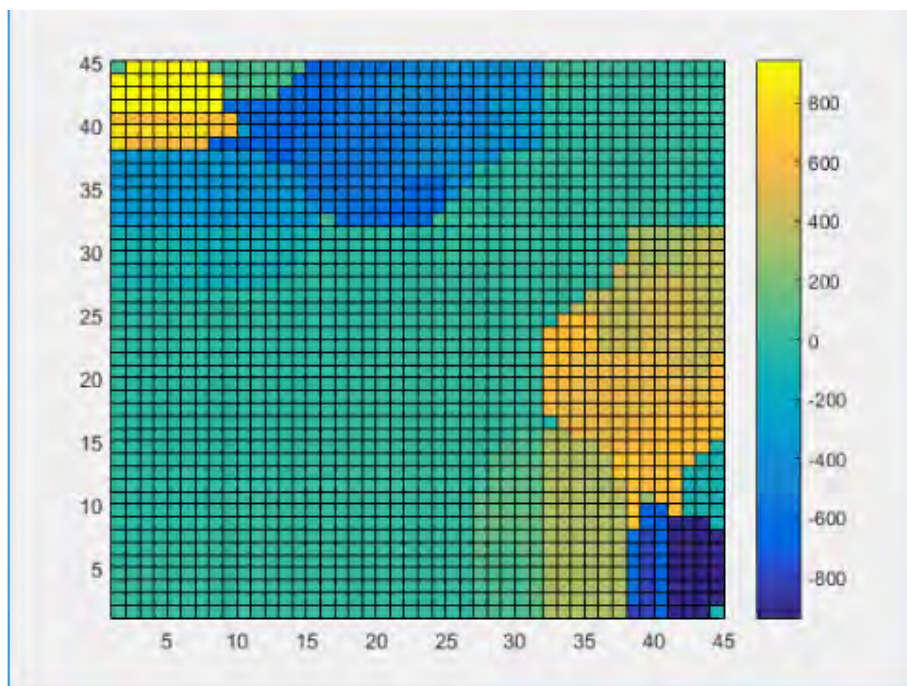
Εικόνα 26: Διάγραμμα cross U (με μαγνητικό πεδίο)



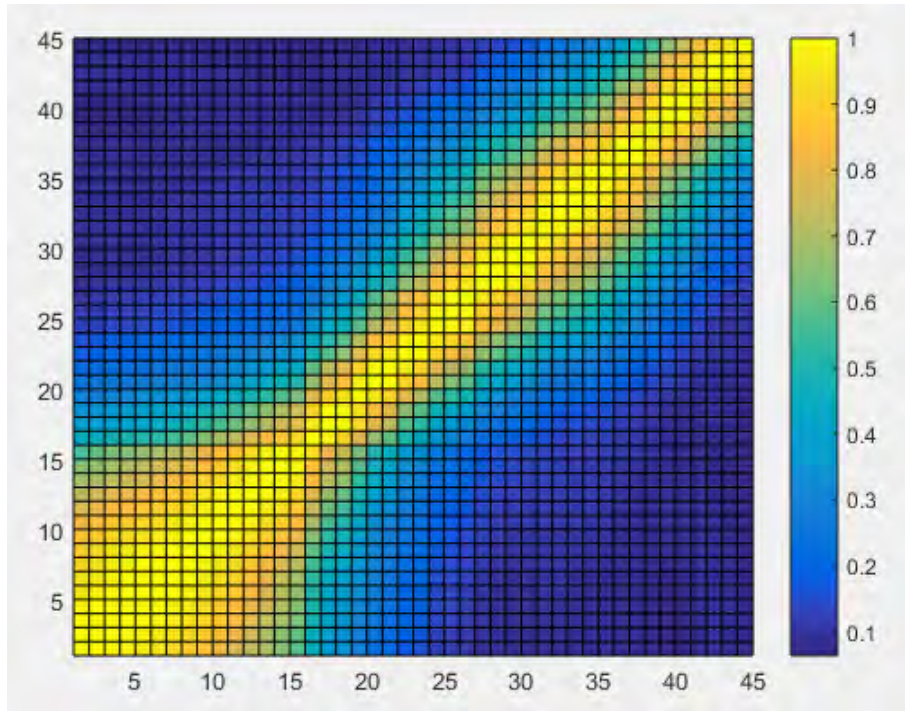
Εικόνα 27: Διάγραμμα  $t_i U$  ( με μαγνητικό πεδίο)



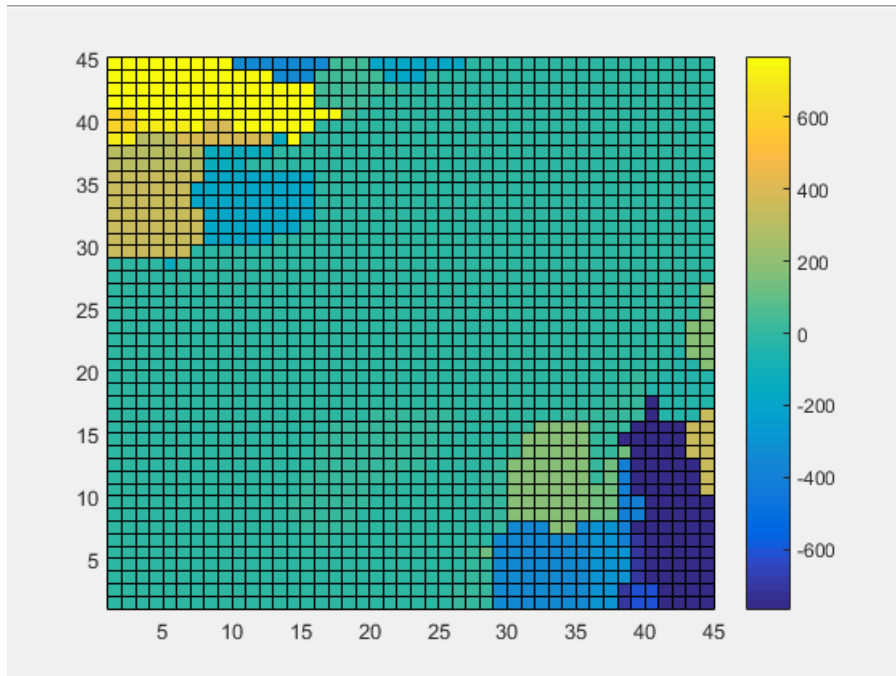
Εικόνα 28: Διάγραμμα cross U (χωρίς μαγνητικό πεδίο)



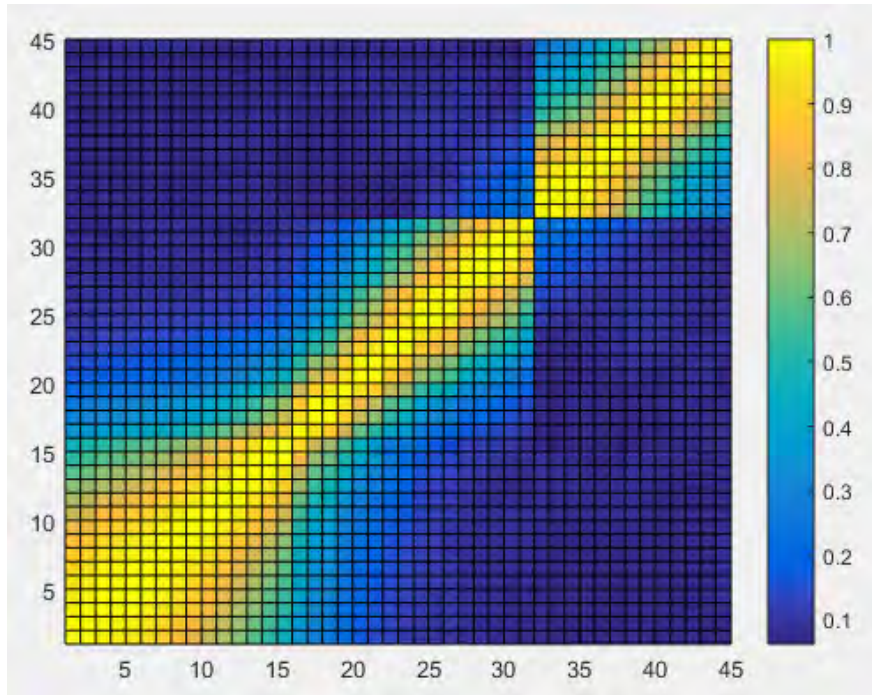
Εικόνα 29: Διάγραμμα ti U (χωρίς μαγνητικό πεδίο)



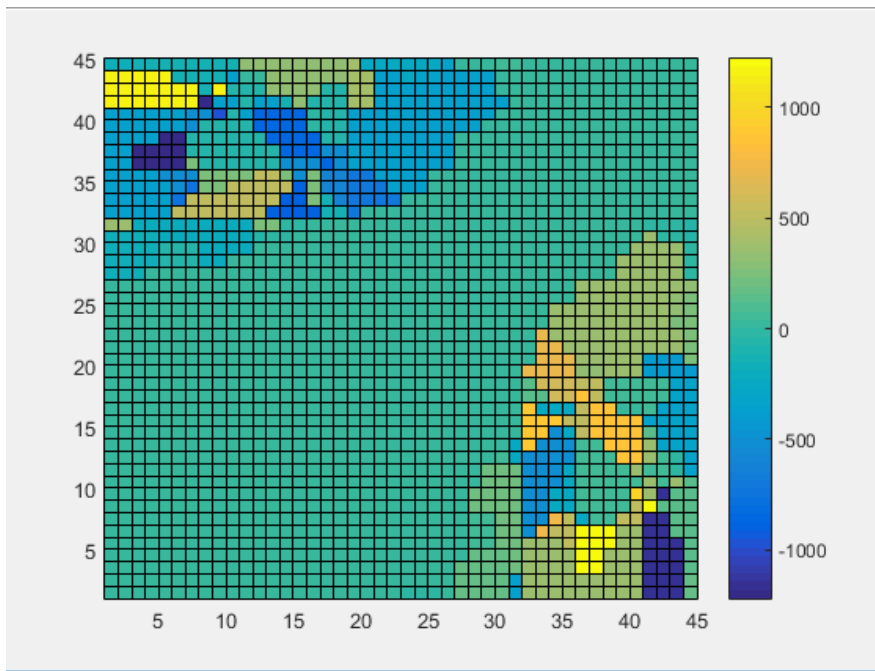
Εικόνα 30: Διάγραμμα cross V ( με μαγνητικό πεδίο)



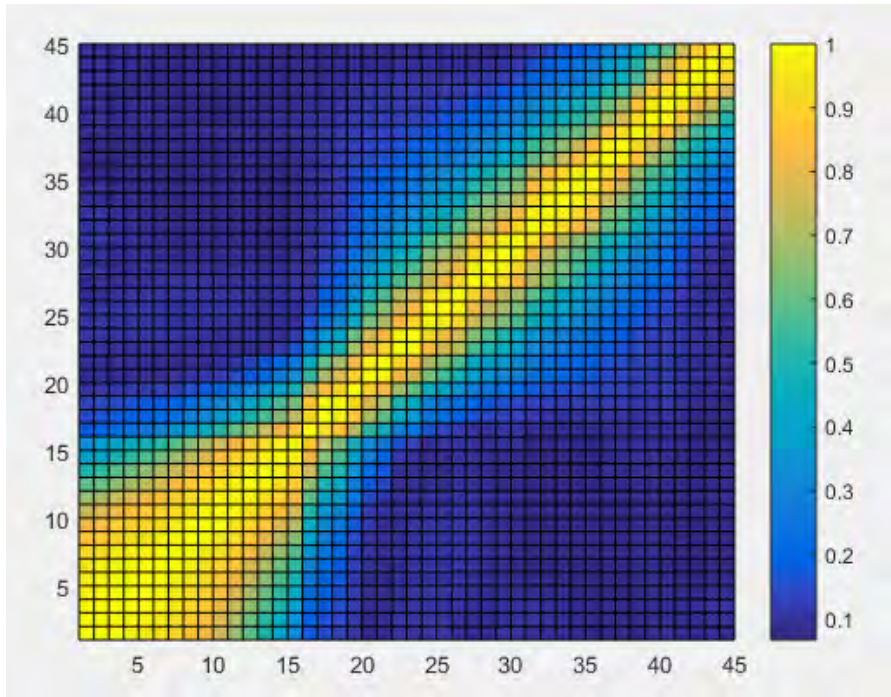
Εικόνα 31: Διάγραμμα ti V ( με μαγνητικό πεδίο)



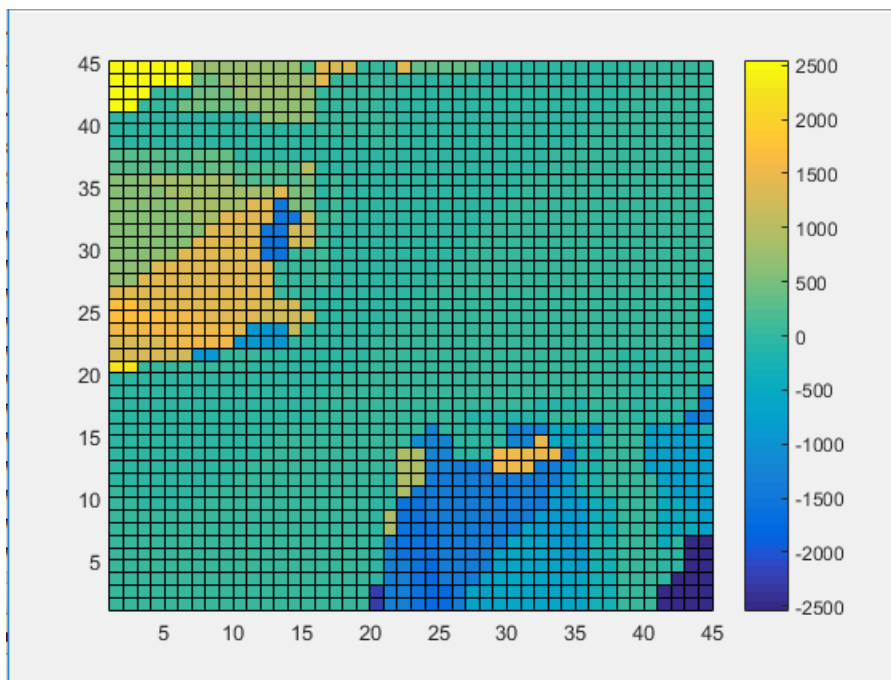
Εικόνα 32: Διάγραμμα cross V (χωρίς μαγνητικό πεδίο)



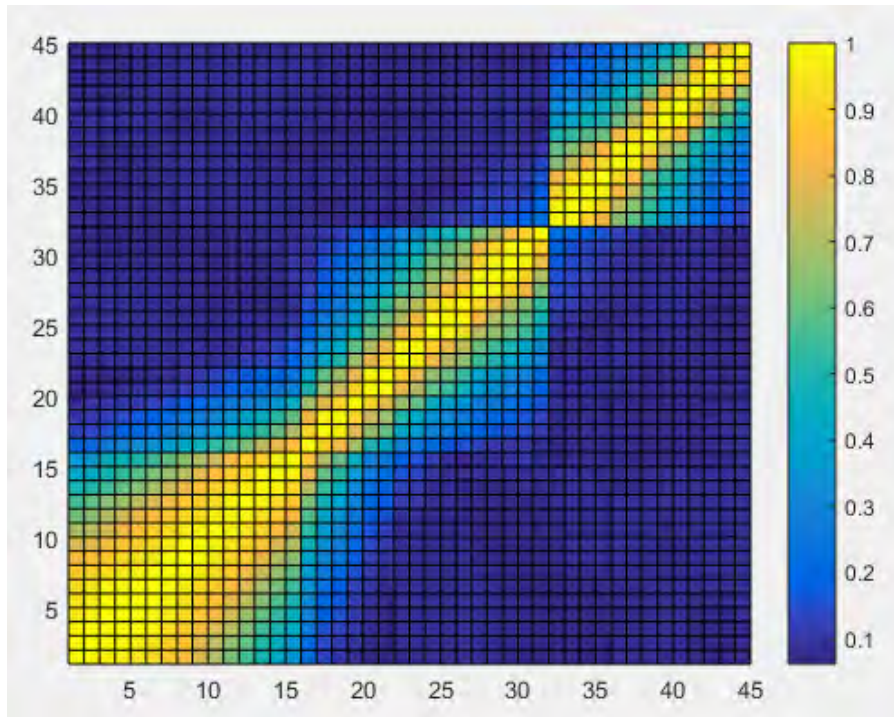
Εικόνα 33: Διάγραμμα ti V (χωρίς μαγνητικό πεδίο)



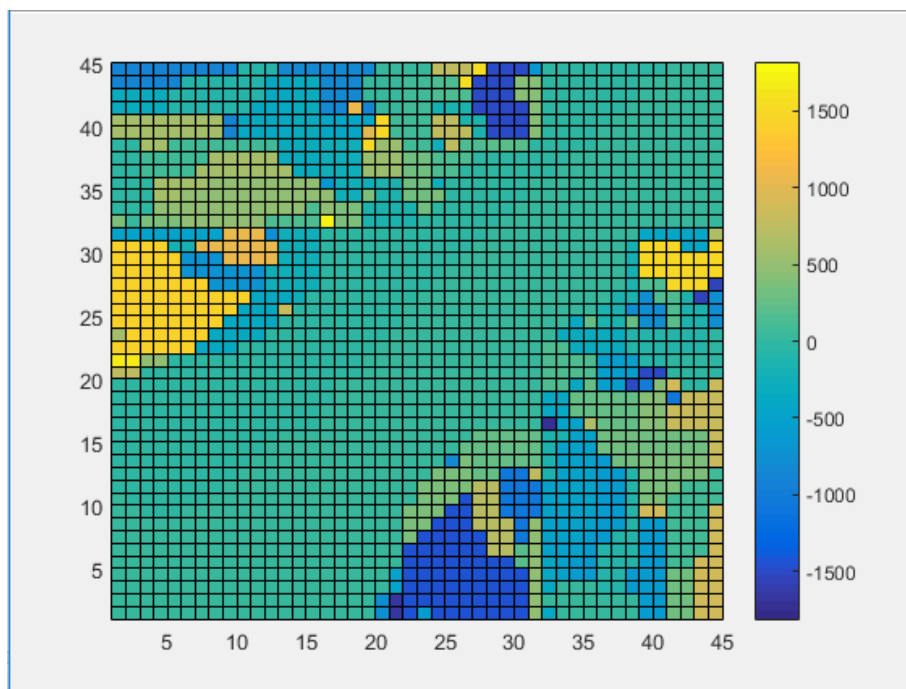
Εικόνα 34: Διάγραμμα cross W ( με μαγνητικό πεδίο)



Εικόνα 35: Διάγραμμα ti W ( με μαγνητικό πεδίο)



Εικόνα 36: Διάγραμμα cross  $W$  (χωρίς μαγνητικό πεδίο)



Εικόνα 37: Διάγραμμα  $ti W$  (χωρίς μαγνητικό πεδίο)



Μέσω παρατήρησης των διαγραμμάτων εντοπίζονται με σχετική ευκολία αρκετές ομοιότητες με τα ομοειδή διαγράμματα των συντελεστών Pearson. Οι πίνακες για ακόμη μια φορά όπως και στην προηγούμενη περίπτωση είναι συμμετρικοί και εμφανίζουν την τιμή της μονάδας στην κεντρική διαγώνιο για αυτό και εκεί υπάρχει έντονη εμφάνιση του κίτρινου χρώματος. Επιπλέον, εμφανίζεται σχετικά μία ομαλή χρωματική συμπεριφορά και μεταβολή πράγμα που σημαίνει πως όσο υπάρχει απομάκρυνση από την κεντρική διαγώνιο, οι χρωματισμοί μεταβάλλονται από κίτρινο σε γαλάζιο χρώμα και έπειτα σε πιο σκούρο μπλε. Αυτό σημαίνει πως η συσχέτιση μεταξύ χρονοσειρών είναι κατά κάποιο τρόπο αντιστρόφως ανάλογη της απόστασης που υπάρχει μεταξύ τους. Επομένως, όσο αυξάνεται η απόσταση, μειώνεται η συσχέτιση. Λαμβάνοντας υπόψη και τα διαγράμματα των χρονικών καθυστερήσεων, γίνεται αντιληπτό πως υπάρχει εμφάνιση μεγάλων τιμών διασυσχέτισης μεταξύ των δεδομένων για μικρές χρονικές καθυστερήσεις. Η ερμηνεία αυτού του φαινομένου μας δίνει για ακόμα μία φορά την απάντηση πως όπως και η συσχέτιση έτσι και η διασυσχέτιση είναι εντονότερη μεταξύ μεγεθών των οποίων οι αποστάσεις είναι μικρότερες. Στις ομοιότητες των διαγραμμάτων συντελεστών διασυσχέτισης με τα αντίστοιχα Pearson, έρχεται να προστεθεί και η απότομη αλλαγή χρωματισμού στις θέσεις γύρω από το κελί (32,32) στις περιπτώσεις όπου υπάρχει απουσία μαγνητικού πεδίου. Σε εκείνη την περιοχή δεν εμφανίζεται ομαλή μετάβαση από το κίτρινο στο μπλε χρώμα όπως και στις περιπτώσεις των διαγραμμάτων Pearson. Τέλος, μεταξύ διαγραμμάτων συντελεστών Pearson και διασυσχέτισης παρατηρείται διαφορά στις αποχρώσεις εκτός της κυρίας διαγώνιου. Στα διαγράμματα Pearson παρατηρούνται γενικότερα πιο ανοιχτές αποχρώσεις του μπλε κάτι που σημαίνει πως οι τιμές σε όλο το φάσμα του πίνακα είναι πιο κοντά στη μονάδα και εκφράζουν ισχυρότερη συσχέτιση από αυτή που εμφανίζεται στα παραπάνω διαγράμματα.

#### **4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΕΦΑΛΑΙΟΥ**

Στη σύγκριση των διαγραμμάτων διασυσχέτισης με αυτά των χρονικών συντελεστών παρατηρούνται επίσης ομοιότητες. Αρχικά το μοτίβο που ακολουθείται στα διαγράμματα συντελεστών Pearson και συντελεστών διασυσχέτισης στην κύρια διαγώνιο εμφανίζεται και στα διαγράμματα των χρονικών συντελεστών  $t_i$ . Όπως στις δύο πρώτες περιπτώσεις έχουμε εμφάνιση της μονάδας στην κύρια διαγώνιο, εδώ φαίνεται πως στην κύρια διαγώνιο και γύρω από αυτή υπάρχουν τιμές που είναι κοντά στο μηδέν. Επιπλέον κατά την απομάκρυνση από την κύρια διαγώνιο παρατηρείται αύξηση των χρονικών συντελεστών, όπως συμβαίνει και στα διαγράμματα συντελεστών συσχέτισης. Οι χρωματικές ανομοιομορφίες, όπου αυτές εμφανίζονται, ακολουθούνται και από τα διαγράμματα  $t_i$ . Χαρακτηριστικά παραδείγματα αποτελούν τα διαγράμματα στα οποία απουσιάζει το μαγνητικό πεδίο. Πιο συγκεκριμένα στην περιοχή του κελιού (32, 32) εμφανίζεται και στους

χρονικούς συντελεστές απότομη αύξηση των τιμών τους κάτι που φαίνεται από την αλλαγή των χρωμάτων. Γενικότερα επιβεβαιώνεται και σε αυτά τα διαγράμματα το “προφίλ” που έχει αναφερθεί και νωρίτερα και έχει να κάνει με τη συμπεριφορά γύρω από την κύρια διαγώνιο αλλά και με την απομάκρυνση από αυτή. Αυτό σημαίνει πως η διασυσχέτιση μεταξύ κοντινών χρονοσειρών συνοδεύεται από μικρές “χρονοδιαφορές” σχεδόν μηδενικές, ενώ στις περιπτώσεις που συγκρίνονται πιο μακρινές χρονοσειρές υπάρχει και μεγαλύτερος χρονικός συντελεστής που τις συνδέει.

Οι μικρές χρονοδιαφορές που εμφανίζονται γύρω από την κύρια διαγώνιο μας οδηγούν στα ίδια συμπεράσματα που μας οδήγησαν προηγουμένως και τα διαγράμματα των spanning trees. Τα αποτελέσματα και των δύο μεθόδων είναι ίδια όμως εκφράζονται με διαφορετικό τρόπο. Στα spanning trees φαίνεται άμεσα η απόσταση μεταξύ των ταχυτήτων μέσω των διαγραμμάτων. Στα διαγράμματα της διασυσχέτισης η μεγάλη ή μικρή απόσταση μεταξύ των ταχυτήτων φαίνεται από την αντίστοιχα μεγάλη η μικρή χρονοκαθυστέρηση που εμφανίζεται μεταξύ των ταχυτήτων. Όπως αναφέρθηκε και προηγουμένως, οι τιμές των χρονοκαθυστερήσεων είναι μικρές στην κύρια διαγώνιο διότι σε εκείνες τις περιοχές συγκρίνονται μετρήσεις ταχυτήτων που βρίσκονται χωρικά κοντά μεταξύ τους. Επομένως τα παραπάνω διαγράμματα αποτελούν ακόμη μία επιβεβαίωση πως τα σημεία των μετρήσεων αποτελούν γεωμετρικά μια ευθεία γραμμή.

Το συμπέρασμα πως τα σημεία συνθέτουν γεωμετρικά μια ευθεία γραμμή θα μπορούσε βεβαίως να βγει και από τους απλούς συντελεστές συσχέτισης του δεύτερου κεφαλαίου. Δεν θα ήταν όμως τίποτα περισσότερο από μία υποψία του παρατηρητή και δεν θα αποτελούσε ασφαλές συμπέρασμα χωρίς περαιτέρω επεξεργασία. Τη δυνατότητα της σιγουριάς για το συμπέρασμα αυτό δίνει η παράμετρος του χρόνου στη διαδικασία της διασυσχέτισης. Από τη στιγμή που οι χρονικές καθυστερήσεις παίρνουν τις παραπάνω τιμές και εμφανίζουν αυτό το προφίλ είναι εύκολα κατανοητό πως τα σημεία είναι σε σειρά. Οι τιμές δείχνουν, όπως είναι λογικό, ότι απαιτείται λιγότερος χρόνος για να πάει κανείς από ένα σημείο σε ένα διπλανό του και πως όσο πιο πολύ απομακρύνεται κανείς τόσο περισσότερο γίνεται και ο χρόνος μεταφοράς.

## ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Αναλύοντας όλα τα παραπάνω και παρατηρώντας τα αποτελέσματα των διαδικασιών που ακολουθήθηκαν μπορούν να βγουν αρκετά συμπεράσματα. Αρχικά, σε όλες τις περιπτώσεις είναι φανερό πως εντοπίζεται διαφορά, στις τιμές των ταχυτήτων και των συντελεστών συσχέτισης, μεταξύ των περιπτώσεων στις οποίες υπάρχει μαγνητικό πεδίο και σε εκείνες στις οποίες αυτό απουσιάζει. Επομένως η επιρροή του πεδίου είναι υπαρκτή και μάλιστα σε αρκετές περιπτώσεις έντονη. Στα στοιχεία που ανήκουν στην κύρια διαγώνιο η παραπάνω επιρροή δεν εντοπίζεται, μόλις όμως υπάρξει απομάκρυνση από αυτή, αυτή είναι εύκολα παρατηρήσιμη. Στα στοιχεία αυτά υπάρχει τέλεια συσχέτιση κατά μήκος της κύριας διαγωνίου ενώ με την απομάκρυνση μειώνεται η συσχέτιση αυτή. Αυτό μπορεί να φανεί και από τα διαγράμματα των μέγιστων χρονικών συντελεστών που δημιουργήθηκαν κατά τη διαδικασία της διασυσχέτισης καθώς στην κύρια διαγώνιο και στις γύρω περιοχές της οι χρονικοί συντελεστές καθυστέρησης τείνουν στο 0. Ως γενικότερο συμπέρασμα μπορεί να επισημανθεί πως ανάλογα με τη διεύθυνση, οι ταχύτητες ακολουθούν ένα προφίλ συμπεριφοράς και πιο συγκεκριμένα στις περιπτώσεις που απουσιάζει το μαγνητικό πεδίο η συσχέτιση αλλάζει ομαλά μαζί με την απόσταση.

Οι μέθοδοι που χρησιμοποιήθηκαν για την εξαγωγή συμπερασμάτων στα προηγούμενα κεφάλαια παρουσιάζουν ομοιότητες αλλά και διαφορές μεταξύ τους. Αρχικά όπως είναι και εύκολα αντιληπτό και από τα διαγράμματα, έντονη ομοιότητα εμφανίζει η διαδικασία των συντελεστών Pearson με αυτή της διασυσχέτισης. Και οι δύο μέθοδοι χρησιμοποιούνται για την εύρεση συσχέτισης μεταξύ δεδομένων μεταβλητών ή συνόλων. Η διαφορά εντοπίζεται στο γεγονός πως οι συντελεστές Pearson δίνουν τη γραμμική συσχέτιση μεταξύ των μεταβλητών ενώ η διασυσχέτιση μπορεί να δώσει τη χρονική καθυστέρηση που υφίσταται μεταξύ τους. Για το παράδειγμα που αναπτύχθηκε στο κεφάλαιο της διασυσχέτισης με τα πλωτά σκάφη και την εύρεση των αποστάσεων μεταξύ τους, η διαφορά στη χρονική φάση είναι πολύ σημαντική και επομένως μία αντίστοιχη λύση με χρήση συντελεστών Pearson δεν είναι εφικτή. Σε προβλήματα που δεν εμπλέκεται ο χρόνος και ενδιαφέρει μόνο η συσχέτιση μεταξύ δεδομένων προτείνεται η χρήση του συντελεστή Pearson στην πληθώρα των περιπτώσεων.

Αντίστοιχα τα *spanning trees* δίνουν πληροφορίες για τις αποστάσεις μεταξύ των δεδομένων. Αποτελούν ίσως την ευκολότερη μέθοδο εκ των τριών που αναλύθηκαν προηγουμένως αφού είναι και η πιο εύκολα κατανοητή για το ευρύ κοινό. Η παρατήρηση των διαγραμμάτων συσχέτισης και διασυσχέτισης πρέπει να συνοδεύεται και από μεγαλύτερη επεξήγηση σε έναν απλό άνθρωπο. Σίγουρα περιέχουν ίσως και περισσότερη πληροφορία από τα *spanning trees* όμως η πληροφορία αυτή είναι “κρυμμένη”. Τα *spanning trees* δίνουν αμεσότερα την πληροφορία στον παρατηρητή αφού περιέχουν στοιχεία και τις μεταξύ τους αποστάσεις. Δεν απαιτούν τον ίδιο όγκο επεξήγησης σε κάποιον που δεν είναι εξοικειωμένος με το πρόβλημα. Στο παράδειγμα του προβλήματος που

πραγματεύεται η εργασία φάνηκε πως τα δεδομένα που εμφάνιζαν τους μεγαλύτερους συντελεστές συσχέτισης μεταξύ τους, εμφάνιζαν και τις μικρότερες αποστάσεις στα *spanning trees*. Αυτό δεν είναι πάντοτε έτσι. Η μικρότερη απόσταση δεν εγγυάται σε όλα τα προβλήματα και την ισχυρότερη συσχέτιση μεταξύ δεδομένων. Τα *spanning trees* είναι ιδανικά για σχεδιασμούς δικτύων είτε όταν πρόκειται για δίκτυα ύδρευσης είτε για υπολογιστικά ηλεκτρονικά δίκτυα.

Για το πρόβλημα που επεξεργάζεται η συγκεκριμένη εργασία περισσότερη πληροφορία δίνουν οι μέθοδοι των *spanning trees* και της διασυσχέτισης (*cross correlation*). Ο συντελεστής Pearson δίνει και εκείνος σημαντικές πληροφορίες όχι όμως στον βαθμό που τις παρέχουν οι άλλες δύο μέθοδοι. Πιο παραστατική από όλες τις μεθόδους, τουλάχιστον στο συγκεκριμένο πρόβλημα είναι αυτή των *spanning trees*. Είναι η μέθοδος εκείνη που μας έδωσε γραφικά την ακολουθία των σημείων μέτρησης και με πολύ απλό τρόπο. Σίγουρα και η διαδικασία της διασυσχέτισης μπορεί να βοηθήσει στη εξαγωγή των ίδιων συμπερασμάτων όμως για να γίνει αυτό τα δεδομένα χρειάζονται περαιτέρω επεξεργασία. Από την άλλη, η διαδικασία της διασυσχέτισης παρέχει πληροφορία για τη συσχέτιση όλων των δεδομένων μεταξύ τους και όχι μόνο εκείνων που διαδέχονται το ένα το άλλο. Για παράδειγμα, για εύρεση συσχέτισης των ταχυτήτων της χρονοσειράς 1 με τη χρονοσειρά 45 αυτό που αρκεί στην περίπτωση της διασυσχέτισης είναι η εύρεση της τιμής του κελιού (1,45). Η πληροφορία δίνεται άμεσα δίχως να ληφθούν υπόψη οι ενδιάμεσες 43 χρονοσειρές. Στο συγκεκριμένο θέμα τα *spanning trees* υστερούν διότι παρέχουν μεν τη βέλτιστη διαδρομή μεταξύ των σημείων, εντοπίζοντας όμως μόνο τις αποστάσεις συνεχόμενων σημείων. Για το ίδιο παράδειγμα που αναλύθηκε στην περίπτωση των *spanning trees* θα ήταν αναγκαίο το άθροισμα όλων των αποστάσεων των σημείων που μεσολαβούν μεταξύ του 1 και του 45. Ίσως και το γεγονός πως τα σημεία βρίσκονται σε ευθεία γραμμή να μην ευνοεί τη λύση του προβλήματος μέσω της διαδικασίας διαγραμμάτων “δέντρων” (*spanning trees*).

Λαμβάνοντας υπόψη όλα τα παραπάνω, φαίνεται πως ιδανικότερη λύση για το παρόν πρόβλημα αποτελεί η διαδικασία της διασυσχέτισης (*cross correlation*). Η διαδικασία αυτή με παραστατικούς πίνακες που απεικονίζουν το βαθμό συσχέτισης μεταξύ των δεδομένων αλλά και την χρονική καθυστέρηση που μεσολαβεί μεταξύ τους φαίνεται ως βέλτιστη ανάμεσα σε αυτές που εξετάστηκαν. Για ένα δυναμικό σύστημα όπως είναι και αυτό που εξετάζεται, φαίνεται πως η μέθοδος αυτή μας δίνει το μεγαλύτερο όγκο αποτελεσμάτων προς επεξεργασία επομένως δίνει και το μεγαλύτερο όγκο πιθανών συμπερασμάτων. Επιπλέον είναι εύκολα κατανοητά για το χρήστη τα γραφήματα που δημιουργήθηκαν μέσω της μεθόδου και παρέχεται και πληροφορία στη χρονική διάσταση μέσω της οποίας μπορούν να βγουν συμπεράσματα σχετικά με τη χωρική κατανομή των σημείων που έγιναν οι μετρήσεις. Επομένως η διαδικασία της διασυσχέτισης είναι και αυτή που προτείνεται για το συγκεκριμένο πρόβλημα με τα συγκεκριμένα χαρακτηριστικά.

Το γενικότερο συμπέρασμα που μπορεί να βγει από όλη την παραπάνω ανάλυση είναι το γεγονός πως κάθε μέθοδος είναι ιδανική για διαφορετικούς τύπους

προβλημάτων και ως μηχανικοί, οφείλουμε να επιλέγουμε κάθε φορά εκείνη που μας δίνει τα βέλτιστα αποτελέσματα. Για προβλήματα που έχουν να κάνουν με απλή εύρεση συσχέτισης επιλέγεται συνήθως ο συντελεστής Pearson ενώ για προβλήματα που απαιτούν χωρική (spanning trees) είτε χρονική (cross correlation) ανάλυση προτιμούνται οι άλλοι μέθοδοι που αναλύθηκαν. Είναι πραγματικά χρήσιμο και σημαντικό να γνωρίζουμε την καταλληλότητα κάθε μεθόδου όταν εμφανίζεται ένα πρόβλημα προς επίλυση έτσι ώστε να κερδηθεί χρόνος, κόπος ή και πολλές φορές οικονομικοί πόροι.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

### Ξενόγλωσση

- Al-Hababbeh, O., Al-Saqqa, M., Safi, M., & Abo Khater, T. (2016). Review of magnetohydrodynamic pump applications. *Alexandria Engineering Journal*, 55(2), 1347-1358. <https://doi.org/10.1016/j.aej.2016.03.001>
- Applications of minimum spanning trees University of Dallas [Course notes]. Retrieved from <https://personal.utdallas.edu/~besp/teaching/mst-applications.pdf>
- Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Oxford: Newnes
- Bergantiños, G., Navarro-Ramos, A. (2019). A characterization of the folk rule for multi-source minimal cost spanning tree problems. *Operations Research Letters*, Volume 47 (547), 366-370. <https://doi.org/10.1016/j.orl.2019.07.002>
- Burgess, T. F. (2002). Systems analysis and modeling: A macro-to-Micro approach with multidisciplinary applications. *Journal of the Operational Research Society*, 53 (12), 1398-1398. <https://doi.org/10.1057/palgrave.jors.2601465>
- Fillinger, L., Sutin, A., & Sedunov, A. (2009). Cross-correlation of ship noise for water traffic monitoring. *The Journal of the Acoustical Society of America*, 126(4), 2251. <https://doi.org/10.1121/1.3249259>
- Fragkou, A., Karakasidis, T., & Sarris, I. (2019). Recurrence quantification analysis of MHD turbulent channel flow. *Physica A: Statistical Mechanics and its Applications*, 531, 121741. <https://doi.org/10.1016/j.physa.2019.121741>
- Gao, W., Zhang, Q., Lu, Z., Wu, D., Du, X. (2018). Modelling and application of fuzzy adaptive minimum spanning tree in tourism agglomeration area division. *Knowledge-Based Systems*, 143, 317-326. <https://doi.org/10.1016/j.knosys.2017.06.007>
- Huang, H., Zheng, S., & Zhao, Z. (2010). Application of Pearson correlation coefficient (PCC) and Kolmogorov-Smirnov distance (KSD) metrics to identify disease-specific biomarker genes. *BMC Bioinformatics*, 11(S4). <https://doi.org/10.1186/1471-2105-11-s4-p23>
- Koenig, J. L. (1999). *Spectroscopy of Polymers*. Elsevier Science. <https://www.sciencedirect.com/book/9780444100313/spectroscopy-of-polymers?via=ihub=>
- Molokov, S., Moreau, R., Moffatt, K. (2007). Magnetohydrodynamics Historical Evolution and Trends, <https://link.springer.com/book/10.1007/978-1-4020-4833-3>

- Monedero, I., Biscarri, F., León, C., Guerrero, J. I., Biscarri, J., & Millán, R. (2012). Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees. *International Journal of Electrical Power & Energy Systems*, 34(1), 90-98. <https://doi.org/10.1016/j.ijepes.2011.09.009>
- Sarris, I. E., Kassinos, S. C., & Carati, D. (2007). Large-eddy simulations of the turbulent Hartmann flow close to the transitional regime. *Physics of fluids*, 19(8), 085109. <https://doi.org/10.1063/1.2757710>
- Schneemeyer, L. F. (2003). Crystal growth. *Encyclopedia of Physical Science and Technology*, 79-89. <https://doi.org/10.1016/b0-12-227410-5/00158-7>
- Wei, Gao, Zhang, Qiuluan, Lu, Zi, Wu, Dianshuang, Du, Xiaohui (2018). Modelling and application of fuzzy adaptive minimum spanning tree in tourism agglomeration area division. <https://doi.org/10.1016/j.knosys.2017.06.007>
- Zhou, H., Deng, Z., Xia, Y., Fu, M. (2016). A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing*, 216, 208-215. <https://doi.org/10.1016/j.neucom.2016.07.036>

## Ελληνόγλωσση

- Αγγελής, Ε. (2013). *Ανάλυση Δεδομένων* [Πανεπιστημιακές παραδόσεις]. Ανακτήθηκε από <https://thalis.math.upatras.gr/~adk/lectures/ida/lab6/slides6.pdf>
- Κουγιουμτζής, Δ. (2009). *Μέρος Β του μαθήματος* [Σημειώσεις]. Ανακτήθηκε από <http://users.auth.gr/dkugiu/Teach/CivilEngineer/descriptive.pdf>
- Παντζάλης, Ν. (2017). *Μηχανική των ρευστών*. Αθήνα: Ίδρυμα Ευγενίδου
- Παπαδόπουλος, Γ. (2013). *Συσχέτιση δύο μεταβλητών -εργαστήριο μαθηματικών και στατιστικής* [Σημειώσεις]. Γεωπονικό Πανεπιστήμιο Αθηνών, <https://www.aua.gr/gpapadopoulos/files/sisxetisi091.pdf>