UNIVERSITY OF THESSALY
SCHOOL OF ENGINEERING
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Platforms for (Self) Evolving Smart Energy Systems

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Rafik Fainti

December 2022

UNIVERSITY OF THESSALY
SCHOOL OF ENGINEERING
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Platforms for (Self) Evolving Smart Energy Systems

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Rafik Fainti

December 2022

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Πλατφόρμες (αυτό) Εξελισσόμενων Ευφυών Δικτύων Ενεργειας

Διατριβή η οποία υποβλήθηκε για τη μερική εκπλήρωση
των υποχρεώσεων απόκτησης του Διδακτορικού Διπλώματος

Ραφίκ Φάιντι

Δεκέμβριος 2022

UNIVERSITY OF THESSALY
SCHOOL OF ENGINEERING
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Platforms for (Self) Evolving Smart Energy Systems

PhD Dissertation

Rafik Fainti

Advisory committee

Lefteri H. Tsoukalas, Professor, University of Thessaly (Supervisor)

Dimitrios Bargiotas, Professor, University of Thessaly

Michail Vasilakopoulos, Professor, University of Thessaly

Examination committee

Lefteri H. Tsoukalas, Professor, University of Thessaly (Supervisor)

Dimitrios Bargiotas, Professor, University of Thessaly

Michail Vasilakopoulos, Professor, University of Thessaly

Miltiadis Alamaniotis, Assistant Professor, University of Texas at San Antonio

Nikolaos Andritsos, Professor, University of Thessaly

Alexandros Chronaios, Professor, University of Thessaly

Aspassia Daskalopulu, Associate Professor, University of Thessaly

December 2022

DISCLAIMER ON ACADEMIC ETHICS AND INTELLECTUAL PROPERTY RIGHTS

Being fully aware of the implications of copyright laws, I expressly state that this PH.D. dissertation, as well as the electronic files and source codes developed or modified in the course of this thesis, are solely the product of my personal work and do not infringe any rights of intellectual property, personality and personal data of third parties, do not contain work / contributions of third parties for which the permission of the authors / beneficiaries is required and are not a product of partial or complete plagiarism, while the sources used are limited to the bibliographic references only and meet the rules of scientific citing. The points where I have used ideas, text, files and / or sources of other authors are clearly mentioned in the text with the appropriate citation and the relevant complete reference is included in the bibliographic references section. I also declare that the results of the work have not been used to obtain another degree. I fully, individually and personally undertake all legal and administrative consequences that may arise in the event that it is proven, in the course of time, that this thesis or part of it does not belong to me because it is a product of plagiarism.

The declarant

Rafik Fainti

Acknowledge

The current dissertation represents the research that was conducted towards my PhD degree from the Department of Electrical and Computer Engineering, University of Thessaly, Greece.

I would like to thank Prof. Lefteri H. Tsoukalas for being a truly mentor and teacher. He gave me the opportunity and the honor to work with him, supporting me through this great adventure not only in my academic quests but also in my life. Working and learning from him was a true honor and advantage. His mentoring and his priceless advice made me a better person and scientist.

I would like to thank Prof. Manolis Vavalis for his guidance, encouragement and advice. He was the first who trusted me as an undergraduate student and introduced me to research. His mentoring, advice and support was priceless. I have been extremely lucky working with him.

I would like to show my appreciation and thank Prof. Miltiadis Alamaniotis who was there for me and helped me upon my arrival at the USA. His mentoring, support and guidance was also priceless.

I would also like to thank Prof. Dimitrios Bargiotas for his advice, comments and encouragement while I was writing my dissertation.

I would like to express my appreciation to the members of the dissertation committee; Prof. Lefteri H. Tsoukalas, Prof. Dimitrios Bargiotas, Prof. Michail Vasilakopoulos, Assistant Prof. Miltiadis Alamaniotis, Prof. Nikolaos Andritsos, Associate Prof. Aspassia Daskalopulu, Prof. Alexandros Chronaios.

At last but not least, I would like to thank my parents and my brother for their love, patience and support. Special thanks to my wife for her great support and love and to my daughter for not crying too much while I was writing my dissertation. I wouldn't be the same person without them.

To my beloved wife and daughter

PhD Dissertation

Platforms for (Self) Evolving Smart Energy Systems

Rafik Fainti

Abstract

The modern society is strongly dependent on electricity which means that the constant supply of such a commodity is essential. On the contrary, while the energy needs of the end-users keep increasing very fast, the expansion of the electricity infrastructure, either talking about the power grid and the transmission system or the power plants themselves does not follow with the same pace that means that the power grid is constantly congested.

The advent of Renewable Energy sources (RES) mitigates the supply issues. In contrast, these sources add unpredictability and stochasticity on the grid since nobody can guarantee their uninterrupted contribution in power supply. Moreover, conventional power units have an inherent inertia. They cannot reduce or increase their production and supply instantaneously that means that these units must supply enough amounts that can cover a temporary loss of RES units. Obviously, the extra amounts of power produced by the conventional power plants are wasted if RES operate normally. Electricity cannot be stored. At least not efficiently.

We should also take into serious consideration the introduction of Electric Vehicles (EVs) which adds more stress on a grid that is already operating at its limits. EVs add an extra load on the system and moreover this load is not diversified too much during the day. More or less, EVs charging cycle follows the same pattern. People leave their homes at morning, go to their works at the same time and charge their cars in the parking lot. They leave their work and return back to their homes where they charge their cars during the afternoon or night.

The stability of the power grid is responsibility of the system operators. They are responsible of the load management in order to guarantee the uninterrupted operation of the system. So far, this task is met empirically meaning that system operators have a knowledge about the days or the times that the grid is most congested or stressed and make decisions about how many and which units to dispatch in order to meet the demand and guarantee the safety of the system at the same time.

The electricity market is undergoing a transformation from its vertically integrated nature to a competitive one based on a security-constrained, bid-based auctiong system for selling and buying electricity. In this concept that introduced by William W. Hogan in [1], the afformentioned practices become obsolete. Moreover, since producers and consumers compete each other in order to sell/buy energy in the best possible price, Artificial Intelligence (AI) algorithms and methods can offer a solution that maximizes their portfolio on the one hand and on the other hand guarantees the stability of the system in real time.

The subject of the current dissertation is on the identification, exploitation and development of novel methods and algorithms focused on the transformation and restructuring of the vertically integrated electricity market into a competitive one where power producers, wholesalers, retailers and end-users compete to optimize their portfolio. A transformation which is of course unavoidable, considering the fossil fuel depletion, and already in progress in the most of the developed countries and especially in the US. To that end, through extensive simulation and experimentation I tried to investigate efficient, non discriminatory ways to exploit electricity auctioning by developing models for a C++ based simulation platform called GridLAB-D [2]. The examination of the feasibility, accuracy and stability of such systems without neglecting or jeopardizing security issues was also under serious consideration. Moreover, the restructuring of the power market and the upcoming ”smartness” of the electricity infrastructure has brought up new challenges regarding security and stability issues as well as the handling of the vast amount of data produced in real time. Consequently, the use Big Data analytics along with the use of Artificial Intelligence techniques such as Neural Networks (NNs) and Fuzzy Logic (FL) in order to accurately anticipate, prevent or fix risky conditions was of high priority.

Διδακτορική Διατριβή

Πλατφόρμες (αυτό) Εξελισσόμενων Ευφυών Δικτύων Ενεργειας

Ραφίκ Φάιντι

Περίληψη

Η σύγχρονη κοινωνία εξαρτάται σε μεγάλο βαθμό από την χρήση ηλεκτρικής ενέργειας κάτι που σημαίνει ότι η συνεχής παροχή του συγκεκριμένου αγαθού είναι απαραίτητη. Αντίθετα, ενώ η ζήτηση των καταναλωτών αυξάνεται πολύ γρήγορα, η επέκταση της υποδομής του ηλεκτρικού δικτύου, είτε μιλάμε για το ίδιο το δίκτυο είτε για τα εργοστάσια παραγωγής ηλεκτρικής ενέργειας, δεν ακολουθεί με τον ίδιο ρυθμό κάτι που οδηγεί σε ένα μόνιμα συμφορημένο δίκτυο.

Η έλευση των Ανανεώσιμων Πηγών Ενέργειας (ΑΠΕ) βελτιστοποιεί τα ζητήματα παροχής ηλεκτρικής ενέργειας. Σε αντίθεση, οι συγκεκριμένες πηγές ενέργειας προσθέτουν στο δίκτυο το στοιχείο του απρόβλεπτου καθώς και στοχαστικότητα μιας και κανείς δεν μπορεί να εγγυηθεί την αδιάλειπτη συνεισφορά τους στην παραγωγή ηλεκτρικής ενέργειας. Επιπλέον οι συμβατικές μονάδες παραγωγής έχουν μια εγγενή αδράνεια. Δεν μπορούν να αυξήσουν ή να μειώσουν την παραγωγή τους στιγμιαία κάτι που σημαίνει ότι θα πρέπει να παράγουν ποσότητες ενέργειας ικανές να αντισταθμίσουν μια προσωρινή απώλεια των μονάδων ΑΠΕ. Προφανώς τα επιπλέον αυτά ποσά ενέργειας απλά χάνονται στην περίπτωση που οι μονάδες ΑΠΕ λειτουργούν κανονικά. Η ηλεκτρική ενέργεια δεν μπορεί να αποθηκευτεί ή τουλάχιστον όχι με αποδοτικό τρόπο.

Κάτι που επίσης πρέπει να ληφθεί πολύ σοβαρά υπόψιν είναι η εισαγωγή των Ηλεκτρικών Αυτοκινήτων (ΗΑ) και ο επιπλέον φόρτος που προσθέτουν σε ένα ήδη συμφορημένο δίκτυο. Τα ΗΑ προσθέτουν ένα επιπλέον φορτίο στο δίκτυο και επιπλέον το φορτίο αυτό δεν διαφοροποιείται πολύ κατά την διάρκεια της μέρας. Λίγο ή πολύ ο κύκλος φόρτισης των ΗΑ ακολουθεί το ίδιο μοτίβο. Οι άνθρωποι αφήνουν το σπίτι τους το πρωί και πηγαίνουν στις δουλειές τους όπου φορτίζουν τα αυτοκίνητά τους στο πάρκινγκ. Παρόμοια, αφήνουν τις δουλειές τους και επιστρέφουν στο σπίτι τους όπου φορτίζουν ξανά τα αυτοκίνητά τους κατά την διάρκεια του απογεύματος ή της νύχτας.

Παραδοσιακά, υπεύθυνος για την ευστάθεια του δικτύου ηλεκτρικής ενέργειας είναι ο διαχειριστής συστήματος. Είναι υπεύθυνος για την διαχείριση του φορτίου με τρόπο που

εγγυάται την απρόσκοπτη λειτουργία του δικτύου. Μέχρι τώρα αυτό ήταν μια διαδικασία που γίνονταν εμπειρικά υπό την έννοια ότι ο διαχειριστής του συστήματος έχοντας γνώση των ημερών ή των ωρών που το σύστημα λειτουργεί με μεγάλο φόρτο, έπαιρνε αποφάσεις για το ποια εργοστάσια παραγωγής ενέργειας θα έπρεπε να λειτουργήσουν ούτως ώστε και να καλυφθεί η ζήτηση και να διασφαλιστεί η ασφάλεια του δικτύου.

Η αγορά ηλεκτρικής ενέργεια υφίσταται αυτή τη στιγμή μία αναδιαμόρφωση από την καθετοποιημένη φύση της προς μία ανταγωνιστική η οποία στηρίζεται σε ένα σύστημα δημοπρασιών τόσο για την παραγωγή και πώληση όσο και για την αγορά ηλεκτρικής ισχύος. Μέσα σ' αυτό το πλαίσιο, το οποίο εκφράστηκε πρώτη φορά από τον William W. Hogan στο [1], οι παραπάνω μέθοδοι καθίστανται παρωχημένες. Επιπλέον, αφού τόσο οι παραγωγοί όσο και οι καταναλωτές ανταγωνίζονται ο ένας τον άλλο ούτως ώστε να πουλήσουν/αγοράσουν ηλεκτρική ενέργεια στην καλύτερη δυνατή τιμή, μέθοδοι και αλγόριθμοι Τεχνητής Νοημοσύνης (ΤΝ) μπορούν να προσφέρουν λύσεις οι οποίες από την μια μεγιστοποιούν το χαρτοφυλάκιο τους και από την άλλη εγγυώνται την ασφάλεια του δικτύου σε πραγματικό χρόνο.

Το αντικείμενο της παρούσας διατριβής είναι η ταυτοποίηση, η παραγωγή και η χρήση καινούριων μεθόδων και αλγορίθμων οι οποίοι θα επικεντρώνονται στον μετασχηματισμό της καθετοποιημένης αγοράς ηλεκτρικής ενέργειας σε μία ανταγωνιστική όπου παραγωγοί ηλεκτρικής ενέργειας, χονδρέμποροι, λιανοπωλητές και τελικοί χρήστες ανταγωνίζονται μεταξύ τους σε μία προσπάθεια να μεγιστοποιήσουν τα κέρδη τους. Ο μετασχηματισμός αυτός ο οποίος είναι προφανώς αναπόφευκτος, ιδίως αν αναλογιστούμε την συνεχή μείωση των ορυκτών καυσίμων, και ήδη σε εξέλιξη στις περισσότερες από της αναπτυγμένες χώρες και κυρίως τις Ηνωμένες Πολιτείες της Αμερικής. Σ' αυτό το πλαίσιο και μέσα από εκτενή πειραματισμό και προσομοιώσεις προσπάθησα να ερευνήσω αποτελεσματικούς και χωρίς διακρίσεις τρόπους εκμετάλλευσης του προαναφερθέντος συστήματος δημοπρασιών αναπτύσσοντας μοντέλα σε μία πλατφόρμα ανοιχτού λογισμικό βασισμένη σε C++, το GridLAB-D [2]. Η μελέτη του κατά πόσο τέτοια συστήματα είναι εφικτά, της ακρίβειας αλλά και της ευστάθειας τους χωρίς να υποβαθμίσουμε ή να αδιαφορήσουμε για ζητήματα ασφάλειας ήταν επίσης μεγάλου ενδιαφέροντος. Επιπλέον, η αναδιάρθρωση της αγοράς ηλεκτρικής ενέργειας και η επερχόμενη «ευφυΐα» της υποδομής του ηλεκτρικού δικτύου έφερε νέες προκλήσεις σχετικά με θέματα ασφάλειας και σταθερότητας, καθώς και τον χειρισμό του τεράστιου όγκου δεδομένων που παράγονται σε πραγματικό χρόνο. Συνεπώς, η χρήση και επεξεργασία μεγάλου όγκου δεδομένων μαζί με τη χρήση τεχνικών τεχνητής νοημοσύνης όπως τα Νευρωνικά

Δίκτυα (ΝΔ) και η Ασαφής Λογική (ΑΛ) για την ακριβή πρόβλεψη, πρόληψη ή διόρθωση επικίνδυνων συνθηκών ήταν υψηλής προτεραιότητας.

Table of contents

List of figures

List of tables

Abbreviations

AI	Artificial Intelligence
AP	Ambient Pressure
ANN	Artificial Neural Network
BREP	Bagging REP Tree
BR	Bayesian Regularization
CFNN	Cascade-Forward Neural Network
CCPP	Combined Cycle PowerPlant
CHP	Combined Heat Power
DR	Demand-Response
DOE	Department of Energy
DER	Distributed Energy Resources
DOH	Degree of Hybridization
DSO	Distribution System Operator
EV	Electric Vehicle
ENN	Elman Neural Network
ESS	Energy Storage System
FFNN	Feed-Forward Neural Network
FL	Fuzzy Logic
HVAC	Heating, Ventilation, and Air Conditioning
HV	Hybrid Vehicle
IEEE	Institute of Electrical and Electronics Engineers
kW	kilowatt
kWh	Kilowatt-Hour
k-NN	K-Nearest Neighbour
MLP	Multi-Layer Perceptron

MFFNN	Maltilayer Feed-Forward Neural Network
NR	Newton-Raphson
NN	Neural Network
NPV	Negative Predictive Values
PV	Photovoltaic
PNNL	Pacific Northwest National Laboratory
PPV	Positive Predictive Values
PI	Predictive Interval
RES	Renewable Energy Sources
RNN	Recurent Neural Network
RNN	Recurent Neural Network
RMSE	Root Mean Square Error
RH	Relative Humidity
SO	System Operator
SVM	Support Vector Machine
SLP	Single Layer Perceptron
T	Temperature
UPFC	Unified Power Flow Controller
V	Exhaust Vacuum

Chapter 1

Motivation

The wellbeing of our society is strongly related to electricity production that may be characterized as the fuel of our prosperity. High quality energy has to be generated and delivered to end users on a 24/7 basis. To that sense, the uninterrupted generation, transmission and delivery of high quality power is essential. The responsible entity for the incessant electricity supply is the System Operator (SO). The operator has to be continuously aware of the status of the system to effectively coordinate supply and demand ensuring the stability of the system. In the ideal scenario, the SO has to to anticipate rare events ahead of time in order to a priori take the necessary measures and avoid severe side effects such as load shedding.

Furthermore, this task is rendered real challenging because of the existence of loop flows and the fact that electricity obeys only on Kirchhoff's laws, and thus, cannot be directed. Moreover, the uncertainty introduced by the new generation electricity markets because of the advent of new concepts like Demand-Response (DR) programs [6], where entities can buy or sell power at different time intervals based solely on the price of electricity, burden furthermore an already difficult problem. Especially, DR programs that have been designed to alleviate congestion problems through pricing signals, eventually can lead to unpredictable and unprecedented demand peaks; a phenomenon well known as ”rebound effect” [7] [8].

Congestion can be expressed in two forms: i) as voltage limit violations, or ii) as line overloading. There are two major issues emerging when congestion arises. First of all, the stability and the status of the grid are in high risk; brownouts or even worst blackouts can occur with unprecedented financial effects for all the parties of the power system. For instance, cheap generated energy may not be delivered to every geographical area, and thus higher prices for electricity usage may be imposed to the end-users at some locations. Second, the new generation energy markets and the new concepts that were introduced have brought up all the endeavors to predict or deal with congestion in big data environments. In particular, the amount of data coming from a grid source tend to be more and more heterogeneous given that new entities are continuously connected to the grid.

1.1 Contribution of Current Dissertation

The contribution of the current dissertation, presented in detail in the next chapters, are summarized in this section.

A detailed description of the GridLAB-D simulation platform used for all the experiments conducted in the following studies is presented. Moreover, the key modules of the aforementioned platform; namely the market module, the residential module, the tape module, the powerflow module, the climate module and the generators module are analyzed in depth. These modules have been extensively used in order to set our simulation environment and establish a distribution network that resembles the real life distribution grids.

In the third chapter, two approaches based on Fuzzy Logic (FL) are presented. The first one deals with the ampacity level monitoring of the distribution grid which is essential in order to be able to guarantee its uninterrupted operation. The second one, provides a new concept called ”Virtual Budget” and copes with the load management of electric vehicles charging. The goal of that particular study is twofold. On the one hand, we try to ensure system's stability by avoiding congestion, especially during the critical times of the day. On the other hand we try to maximize the benefit of the end user regarding the amount of money he spends for electricity consumption.

In chapter number four, 4 different approaches, based on AI techniques, are presented. These approaches deal with the monitoring of the distribution system and the anticipation of line overloadings, providing useful tools and insights to the SO reagarding the safety of the system and the actions that has to be taken in order to ensure that safety. A hierarchical method based on ANNs for the power output prediction of a combined cycle power plant is also presented.

The last chapter summarizes the efforts and the contributions of the current dissertation along with some future plans.

Chapter 2

Simulation Platform

2.1 GridLAB-D Simulation Platform

The main platform used for all the experiments and simulations during this dissertation was the GridLAB-D [2] simulation platform. GridLAB-D, written in C++, is an open source, agent-based simulation platform entirely focusing in the power distribution level. It has been developed at the Pacific Northwest National Laboratory (PNNL) by the US Department of energy (DOE), while academia and industry also participated in this project. Among others, it provides four key tools necessary for the experiments conducted under the current dissertation.

- Information and agent based tools enabling researchers to develop models focusing on the interaction among energy markets, end-use technologies, distribution automation and Distributed Energy Resources (DER).
- Enable researchers to identify, examine and ratify end-users interactions, dependencies and reactions related with the introduction of new technologies on established or suggested power markets.
- Provides models and interfaces to established power system tools as well as, existed analysis systems.
- It provides a wide variety of data collection, processing and analysis tools.

Getting into more detail, GridLAB-D consist of a core module responsible for the control and the management of each independent object that participates in the simulation process as well as, the monitoring of all the information procured by each of the aforementioned objects during the simulation. We can say that this is the heart of GridLAB-D. Among others, there are four important independent modules/components connected to the core which represent the financial and technical aspects of the distribution system operations. A detailed depiction of this architecture can be seen in Figure 2.1.

More preciesely, these components, namelly buildngs, markets, power systems, contol systems are further sudivided into more specific modules performing key functionalities regarding the simulation process and the acquisition of accurate results. These modules are responsible, as mentioned earlier, for important operations such as building the market framework into which the simulation will take place, the power flow, the load profiles of each individual residency participating into the simulation, the power generation etc.

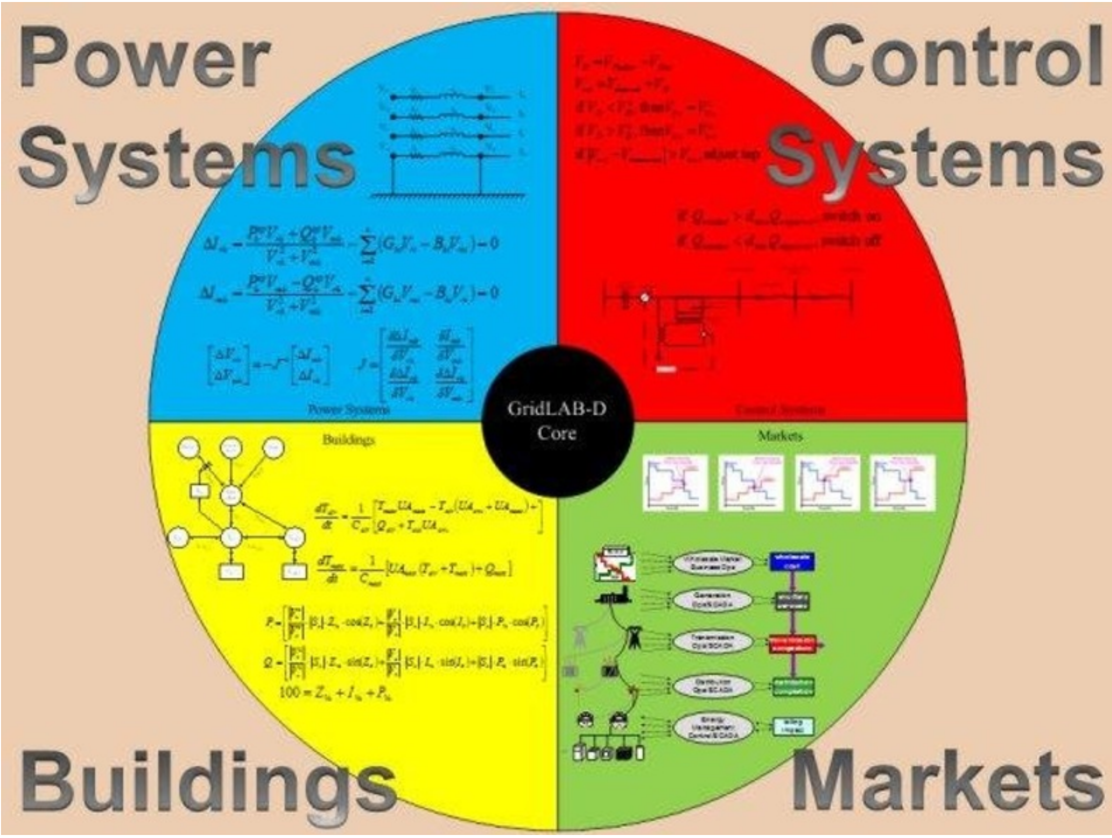

Figure 2.1: GridLAB-D Architecture.

https://sourceforge.net/projects/gridlab-d

Focusing on the most important modules as well as on the modules used in this dissertation, a short description follows:

- **Market Module** [9]. It provides the capabilities of an auction-based power marked as the one described by Hogan in [1] where producers (Power plants, RES etc) and end-users bid in order to sell/buy energy trying to maximize their profit.

- **Residential Module** [10]. This module is responsible for simulating the load demand of the power grid by implementing all the necessary functionality of the end-user's appliances.

- **Tape Module** [11].This module enables researchers to record or modify the conditions of a simulation by manipulating the objects participating in the simulation.

- **Powerflow Module** [12]. The distribution grid is consisted of a set of nodes. For each of these nodes we need to know the voltage, the power loses, the currents and go on. The powerflow module is responsible for all of these calculations. To that end, the entire model regarding the distribution system is provided by this module. The method used by GridLAB-D in order to solve the power flow equations is the Newton-Raphson (NR) technique.

- **Climate Module** [13]. Based on .tmy type files the researcher can provide to the model all the necessary information about the weather condition prevailing during the execution of the simulation. The climate module provides all the necessary functionality to that cause.

- **Generators Module** [14]. The functionality for implementing generators and storage resources such as RES and batteries is provided by the generators module. Theses resources are connected to the main distribution grid using meters and triplex meters which are objects provided by this module.

Building a model in GridLAB-D is realised using the so-called .glm files where the researcher writing pseudo-code can construct the distribution system, the generation and storage facilities, the residencies equipped with appliances, end-users habits etc. Inside these files the researcher should give all the necessary information about the initial state of each node and the settings of the end-user's appliances like the temperature that the air-condition system should maintain according to resident's preferences or the time the resident leaves home to go to work. The outcome of the simulation and the whole information regarding the results is provided through .xml files. Furthermore, a server where the researcher can pass into the system information in real time is also provided.

2.2 GridLAB-D Objects

This section is dedicated to provide further information and analysis regarding the objects that constitutes the five main GridLAB-D modules used for the purposes of this dissertation.

2.2.1 Residential Module

The most important object of that particular module is the house object that enables researchers to simulate the load (i.e the demand) on the distribution grid. That is realized through two different types of appliances the so-called thermostatically and non-thermostatically controlled appliances. In the first category we can find appliances like the water heater and the Heating, Ventilation, and Air Conditioning (HVAC) system. In the latter we have the kitchen, the lights etc.

2.2.2 Tape Module

As mentioned earlier, this module enables the researcher to record or modify the conditions of a simulation in real time. This is realized through four different objects namely the collector, the recorder, the shaper and finally the player object. The collector and the recorder, as their names imply, are used to record information regarding all the parameters of the simulation at predefined periods of time. On the other hand, the shaper and the player objects are used to update simulation parameters at predefined periods of time via files.

2.2.3 Generators Module

The Generators module replicates all the necessary power generation and storage of the system. To that end, this modules provides the researcher with tools to simulate synchronous generators (diesel_dg object), solar panels (solar object), wind turbines (windturbine_dg object) and batteries (battery object).

2.2.4 Powerflow Module

There are three main objects that constitute the powerflow module and play an important role both for designing the distribution system and for acquiring the simulation results.

- The **node** object with which the researcher can design all the necessary connections and buses of the distribution system.

- The **meter** and the **triplex_meter** objects enable the researcher to acquire measurements regarding the load from predefined points of the distribution system. This load stands for the consumption of the end-users connected to each particular meter. Moreover the researcher can acquire information regarding the magnitude of the current flow through each meter. The only difference between those two objects is that the first is used for ordinary three-phase systems while the latter for triplex systems.

2.2.5 Market Module

The most important, at least for the simulations of the current dissertation, module is the market module. The functionality of that module depends on four objects; namely the controller object, the passive_controller object, the auction object and stub_bidder object. A short description of that objects follows.

The responsibility of the controller object is to compare the average price of the market with the clearing price and bid on behalf of the appliances for consumption forming that way the load of the system. These bids are sending to the auction object. the same stands for the passive_controller object. The difference between these two objects is that the passive_controller does not bid back. The reader is referenced in [15] for more information.

The auction object, mentioned earlier, enables researchers to design an electricity market where producers and consumers exchange power trying to maximize their profit. The producers try to sell energy in the highest price, while end-users try to buy in the lowest price in order to feed their devices. There is a particular time period into which both parties submit their bids. This time period is called ”market period”. After this period passes, the bids of both parties are stack up from the highest to the lowest as to price. The two curves that formed by that process are intersecting at a specific point that determines the quantity of the power to be exchanged and at which price will be exchanged. This quantity and price are called clearing quantity and price respectively. To that end, the clearing quantity and price represent the preferences of the market participants. What follows is a process called market dispatch. It is the determination of the consumers and the producers that will be actually involved in the next market period. The producers will produce and sell the clearing quantity while the end-users will buy that quantity at the clearing price.

A representative example of that concept can be seen in Figure 2.2. The reader can see two different clearing market examples in that figure. The sellers stand for the producers while the buyers stand for the end-users' appliances like the water-heater or the HVAC system. The two curves are formed as explained earlier while their intersection point represents the clearing price and quantity of the market. More precisely, the clearing quantity is defined as the aggregation of the quantities of all end-users that have submitted bids in a price higher than the cleared price.

Figure 2.2: Clearing price and quantity

http://gridlab-d.shoutwiki.com/wiki/Spec:Market

As mentioned in the first paragraph, the appliances participate into the market through the controller object. Actually this is true only for the thermostatically controlled devices like the water-heater and the HVAC system. All the other devices, the lights for instance, participate through the stub_bidder object which is responsible of submitting bids for the non-thermostatically controlled devices as well as, for the conventional generators and it is realized through bids in fixed prices and quantities respectively. The non-conventional generators on the other hand e.g. solar panels, wind-turbines etc. participate into the market through a different type of controller called generator_controller. In the same manner, the submitted price is fixed. The difference is that the submitted quantity is differentiated according the weather conditions and the particular type of each generator. It is obvious for example that the generation of a wind-turbine during a day with moderate winds will not be the same with the generation during a windy day.

Chapter 3

Fuzzy Logic Approaches

3.1 Motivation and Related Work

Power system's main responsibility is to constantly provide high quality power to end users. Any aspect of our daily activities, ranging from entertainment to critical operations like surgeries, is based on electricity. We leave in the era where the ability to have light or to cook by simply turning on a switch is taken for granted. The economic consequences for any engaging party in the opposite scenario, especially in case of a lengthy blackout, will be enormous. Hence, it is more than obvious that monitoring the power system and preventing undesirable situations is paramount.

During the last decades, the process of monitoring was based on the so-called rules of thumbs. People empirically knew the critical times or the events that might put the system at risk. Based on that they were able to take actions. However, the constantly growing demand for electricity, as well as the restructuring and the unbundling of the vertically integrated form of the power sector constitutes these practises obsolete. Fortunately, the introduction of IT into the power sector gives us the capability to closely and effectively monitor the power system with a plethora of tools.

To that end, a lot of studies using Artificial Intelligence (AI) techniques and more precisely Fuzzy Logic techniques concerning monitoring and predictive applications have been conducted. The authors in [16] developed a fuzzy based congestion management method for the identification of the optimal sizing and location of a Unified Power Flow Controller (UPFC) at the transmission level. Ray et al. in [17] combine a fuzzy logic system and wavelet transforms to perform fault classification of a transmission line. More precisely, 11 possible fault types are classified through a fuzzy logic based system. The authors in [18] try to address congestion issues, in terms of line overloading, in the power system directing their attention on the transmission level. Based on fuzzy logic and more precisely using fuzzy modeling, they develop a method for determining the Available Transfer Capacity based on three input variables; namely the sink bus injection, the neighboring bus injection and the loading index. The most common use of fuzzy logic in power systems is related to the implementation of controllers called fuzzy controllers. Rezaei and Esmaeili in [19] develop a fuzzy controller to control the reactive power produced by photovoltaic and wind generation distributed units. A fuzzy logic controller has been implemented in [20] regarding the active power insertion of large scale photovoltaic power plants addressing the problem of voltage dips. The control of a photovoltaic system's converter through a fuzzy controller is proposed as well by Basaran and Cetin in [21]. The authors in [22] focus their attention on a microgrid equipped with photovoltaic and wind energy system, as well as battery storage. They propose a fuzzy based control approach towards efficient power sharing. Their scheme is tested both on islanding and grid-connected conditions. Selim et al. in [23] cope with the problem of aggregation and active power control of electric vehicles at the distribution level. A fuzzy controller using as input parameters the electricity price, the maximum allowed to charge time of the vehicle, as well as the initial state of the battery of each vehicle, is proposed. A combination of two AI techniques, namely neural networks and fuzzy logic, is proposed by Dong et al. in [24] The authors develop and evaluate a fuzzy neural network approach for wind power forecasting. The results of the study demonstrated the effectiveness of the approach as well as the improved, compared with conventional approaches, forecasting accuracy. Schouten et al. in [25] have developed a fuzzy logic controller utilizing driver's preferences, the battery's state of charge and the speed of motor or generator of Hybrid Vehicles (HVs). The aim of the research is the determination of the split of these two power units in order to achieve the maximum fuel economy along with the engine's efficiency. In [26], Singh et al. develop and present two fuzzy logic controllers aiming to address two basic issues related with the advent of EVs. The first is related with the flattening of the voltage profile on the respected distributed node. They achieve that with the proper charging coordination of the EVs. The second is related with the use of EVs as distributed energy storage devices in the sense that if the EVs are fully charged, then the excessive energy can be sent back to the grid in a coordinated way. Datta and Senjyu in [27] cope with the issue of frequency control in modern power grids equipped with distributed PV panels and EVs. To that end, a fuzzy based frequency control system regarding EVs, PV inverters and Energy Storage Systems (ESSs) is proposed and simulated in that research work. Hemi et al. [28] present a fuzzy logic based controller towards the power management of hybrid electric vehicles as well as the overcharging protection of the battery related with energy that accumulated due to the braking of the vehicle. In [29], Li and Lie also deal with the issue of hybrid electric vehicles aiming to maximize the efficiency of the vehicle by determining the degree of hybridization (DOH) as well as by developing a novel approach related with the power control of the automobile. They propose a novel approach based on fuzzy logic able to guarantee the overall efficiency, as well as the fuel economy of the vehicle. Ma and Mohammed in [30] present the development of a fuzzy logic controller using the output of PVs placed in a parking garage. They use the demand of the EVs parked at this garage and the price of electricity usage in order to balance the demanded load from the main grid as well as the resulted cost for the EV users.

3.2 Fuzzy Logic Preliminaries

Fuzzy logic basically founded in 1965 when the theory about the fuzzy sets first introduced by Dr. Lofti A. Zadeh. The classic set theory says that an element strictly belongs to a set or not. In fuzzy logic, though, an element can belong to more than one set to a different degree. In that sense, fuzzy logic is closest to human reasoning since when we asked to answer about the height of another human, for instance, we respond by saying tall, short, very short etc. In other words, it's very intuitive to us to define fuzzy sets such as tall, short and so on, where each numerical value of height belongs to a different degree.

More specifically, in crisp set theory an element strictly belongs to a set or not. This can be modelled using a function, called characteristic function, as depicted in equation 3.1.

$$f(x) = \begin{cases} 0, & \wedge x \notin A \\ 1, & \wedge x \in A \end{cases} \tag{3.1}$$

In equation 3.1 A is a crisp set and x is an element. The corresponding equation in fuzzy logic theory, depicted in equation 3.2 is the so-called membership function which assigns to each element a degree of membership related to each fuzzy set and constitutes the cornerstone of fuzzy logic.

$$\mu(x) : X \to [0, 1] \tag{3.2}$$

In equation 3.2 μ stands for the membership function, A is a fuzzy set, x is the element for which we are looking the degree of its membership in the fuzzy set A, while X stands for the universe of discourse. Finally, in order to follow the notation used in classical set theory, equation 3.1 can be transformed into equation 3.3:

$$f(x) = \begin{cases} 0, & \wedge x \notin A \\ \mu(x), & \wedge x \in A \\ 1, & \wedge x \in A \end{cases} \tag{3.3}$$

Graphically, there are different representations of the membership function, including triangular shape, trapezoidal shape, Gaussian shape and so on. The selection depends on the problem as well as on the personal intuition of the expert who designs the system. The reader can see in Figure 3.1 three triangular shape membership functions modelling the fuzzy variable of temperature into three fuzzy sets; namely ”Low”, ”Medium” and ”High”.

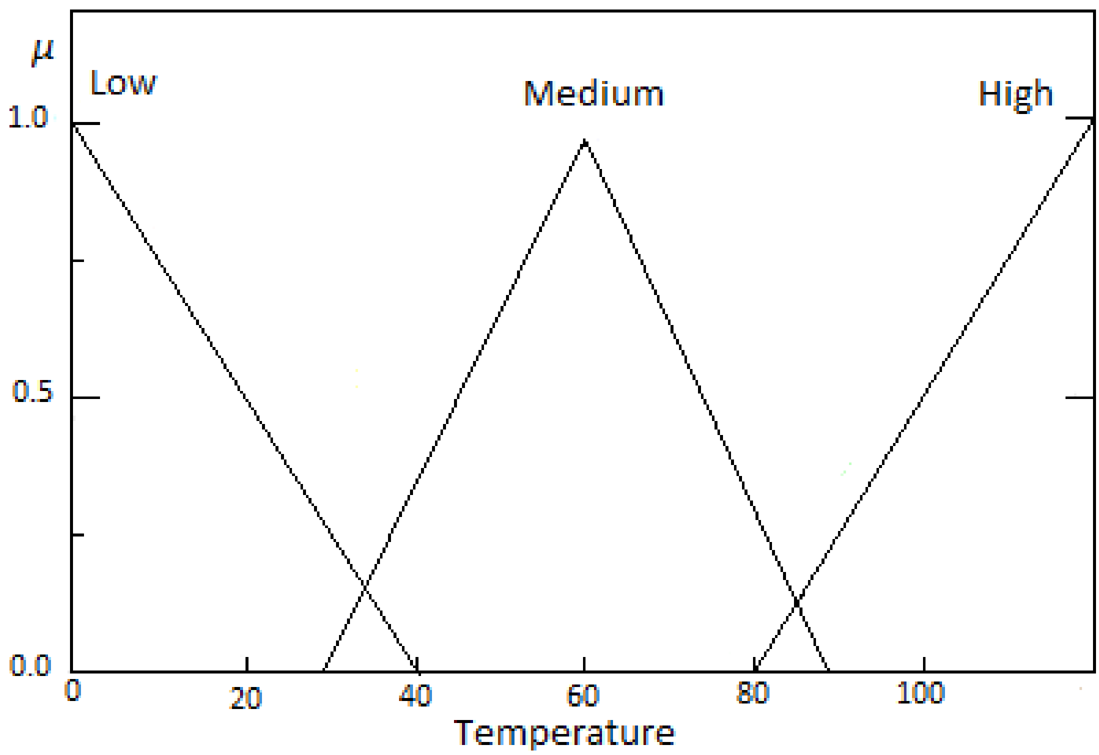

Figure 3.1: Triangular shape membership functions modelling temperature

In Figure 3.1, the x-axis stands for the temperature in Fahrenheit, while the y-axis for the degree of membership of each specific numerical value of temperature on each fuzzy set. A value of 60 degrees Fahrenheit for example, belongs to the fuzzy set ”medium” with 1.0 degrees of membership (the maximum) and with 0 degrees (the minimum) of membership to the other two sets. Further explanation of these concepts is beyond the scope of this dissertation. Hence, the reader is referenced in [31] for information about the fuzzy logic theory and its use in engineering.

3.3 Ampacity Level Monitoring Utilizing Fuzzy Logic Theory

3.3.1 Experimental Configuration

The purpose of that particular study was to develop a method close to human reasoning using fuzzy logic theory in order to monitor the ampacity level on each phase of a distribution line. Given that, the reader can see in Figure 3.2 the flowchart of the actions to be made to develop a fuzzy monitoring system.

As it can be seen, the inputs of the method would be crisp values. The selected input parameters for this study are the ”temperature”, the ”time” of the day and finally the ”price” at that particular time. The use of the parameters ”temperature” and ”price” is more or less obvious since we focus our attention in a competitive electricity market where the operation of the devices strongly depends on the price. Of course, in case of severe weather conditions people will use heating or cooling no matter the height of the electricity price. The use of the parameter ”time” on the other hand is because of the need to simulate the presence of the residents at their residencies and consequently the use of water heater primarily, as well as other devices.

Subsequently, different fuzzy sets are designed for each parameter. For the fuzzy parameter ”temperature” we have the fuzzy sets ”**VeryLow**”, ”**MediumLow**”, ”**Low**”, ”**Comfort**” and ”**High**”. Each fuzzy set has a different range and it is up to the designer and its personal intuition to design properly each of them. Similarly, for the fuzzy parameter ”time” the corresponding fuzzy sets are the ”**morning**”, ”**lateMorning**”, ”**aroundNoon**”, ”**afternoon**”, ”**earlyNight**” and ”**night**”. For the fuzzy variable ”price” the corresponding fuzzy sets are the ”**veryLow**”, ”**low**”, ”**moderate**”, ”**high**”, and ”**very-High**”. In the same manner, we design

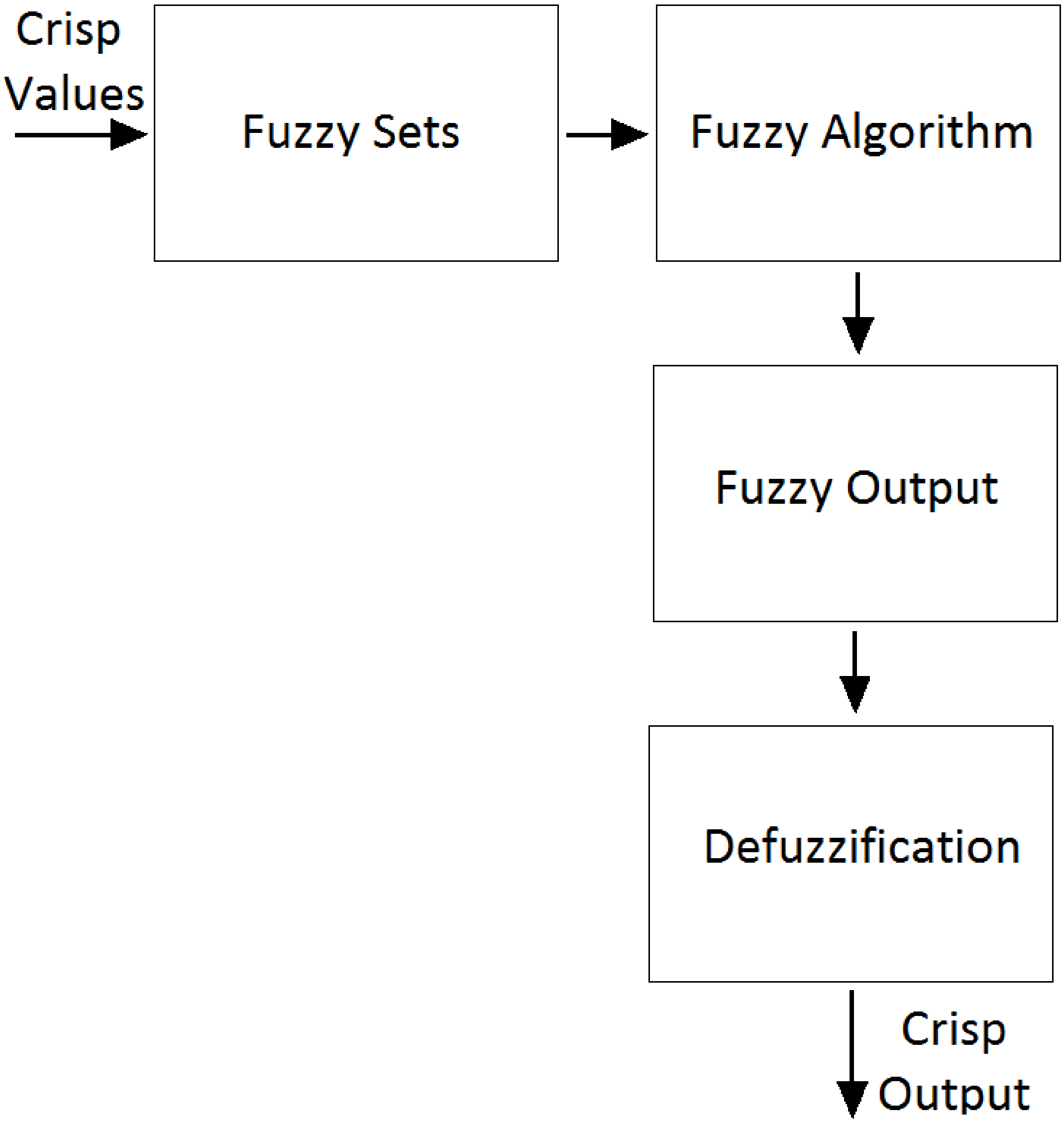

Figure 3.2: Flowchart of actions of a fuzzy expert system

the fuzzy sets of the output, called ”congestion”. The designed fuzzy sets in this case are the ”**veryLow**”, ”**low**”, ”**limit**”, ”**lowCongestion**”, ”**congestion**” and ”**highCongestion**”. The first three indicate absence of congestion; i.e. the ampacity level is below the thermal limit of each phase. Though, the fuzzy set ”**limit**” indicates that the ampacity level approaches the thermal limit and even if we don’t face congestion, some precautionary measures have to be taken. The last three fuzzy sets indicate presence of congestion. The ampacity level is beyond the thermal limit of the line and measures have to be taken immediately before the total collapse of the line.

These parameters are then fuzzified through membership functions. The relations of the input parameters and their degree of membership into prespecified fuzzy sets is determined that way. Taking as an example Figure 3.1 there are three membership functions fuzzyfying the parameter "temperature" into three different fuzzy sets, namely "**Low**", "**Medium**" and "**High**". The fuzzification process takes place by using an implication operator, called mandani-min [31], that can be seen in equation 3.4.

$$\phi_C(x)[\mu(x), \mu(y)] = \mu(x) \wedge \mu(y) \tag{3.4}$$

In equation 3.4 A and B are two fuzzy sets, while μ_A,μ_B are their corresponding membership functions.

Since we want our monitoring method to produce a numeric value as output, i.e. the ampacity level of the phase, and not a fuzzy one, the last step is to defuzzify the fuzzy output. In other words, we seek a crisp output. This process is called defuzzification. Similar to the fuzzification process, there are different approaches depending on the accuracy one wants to accomplish. To that end, the centroid method as presented in equation 3.5 was used in this study.

$$\mu^* = \frac{\int \mu(x) \cdot x dx}{\int \mu(x) dx} \tag{3.5}$$

In equation 3.5 μ^* stands for the crisp output, χ stands for the fuzzy output and $\mu_A(\chi)$ stands for the membership function.

It must be also mentioned that two different monitoring fuzzy systems were designed; one for summer and one for the rest of the seasons. The reason for that is that there are a lot of differences between summer and all the other seasons primarily on the time of electricity usage and secondarily in the height of the electricity usage price. For example, we face two critical, in terms of line overloading, periods during winter. During morning where people wake up and prepare to go to their jobs and afternoon where they return back. During summer on the other hand, we are facing just one. During afternoon where the temperature reaches its maximum values.

The distribution grid, the competitive electricity market mentioned in the introduction, as well as the data for both the input parameters and the evaluation of the method, were implemented in the GridLAB-D [2] simulation platform.

The simulated distribution feeder used in this study is the IEEE radial distribution feeder [32] with 1306 residencies as depicted in Figure 3.3. Each residency is occupied by four

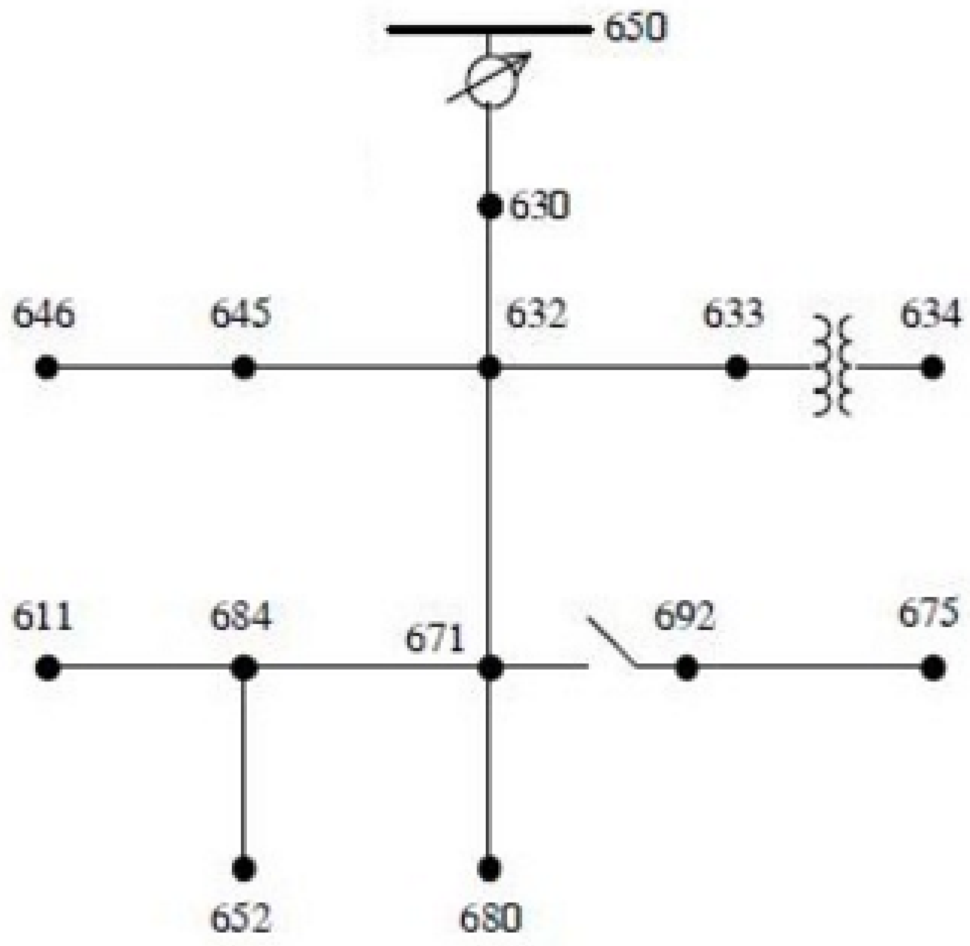

Figure 3.3: IEEE-13 radial distribution feeder

individuals with slightly different energy usage patterns. The energy needs of each residency are constituted of one water heater, one Heating Ventilation and Air Conditioning (HVAC) system and lights. Furthermore, the feeder contains 4 distributed power generators. Generator 1 has capacity of 244 kW and marginal cost of 0.3 $/kWh. Generator 2 has capacity of 500 kW and marginal cost of producing real power of 0.11 $/kWh, while generator 3 and 4 have 400 and 100 kW capacity and 0.25 and 0.15 $/kWh marginal cost respectively.

More descriptively, regarding the form of the competitive electricity market used in this study, it is an auction based market where buyers and sellers submit bids for buying or selling real power in the context defined by William W. Hogan in [1]. This type of market is implemented in GridLAB-D through automated controllers that bid on behalf of the buyers for a certain amount of real power at a certain price, based on prespecified thresholds. These thresholds actually represent the willingness of each buyer to allow the devices of the residency to operate depending on the market price, as well as the weather conditions prevailing at each particular time. Moreover, the controller is responsible for adjusting the operation set points of the devices as a response to the height of the market price. The heating set point of an HVAC system would be increased in case of high prices, for instance, or the opposite in case of low prices. These actions target to leverage the market price for the benefit of the buyer. This concept has been thoroughly explained in the previous chapter.

3.3.2 Results

This sub section is devoted to the presentation of the experimental results. We focus our attention on the phase A of the distribution line 630-632 (see Figure 3.3) where we have intentionally reduced the thermal limit of the line in order to create overloading. At first we present congested times during summer. Since the peak hours, and consequently the possible congested ones, are during the afternoon, only these times are depicted here. Subsequently, congested times for the rest of the seasons are presented. During these months we observe peaks around morning and afternoon. It must be mentioned that because of the competitive nature of the market, there is no specific pattern of congestion presence. Because of that only some selected results are presented. Moreover, trying to create realistic scenarios we didn't decrease the thermal limit too much and thus we have only few congested days. Most of them are presented here.

A. Summer

The reader can see in Table 3.1 a comparison between the real, as obtained by the GridLAB-D simulation platform, and the predicted ones as obtained by our method ampacity levels on the phase A of the line 630-632. Four different days are presented in this table; namely July 20th, July 23th, July 24th and August 6th. The first column of the table stands for day and the month, the second for the temperature, the third for the time while the fourth for the prevailing at that time electricity usage price. The last two columns of the table contain the real and the predicted ampacity level.

Moreover, the reader can see in Figure 3.4 a graphical representation of the whole congested period for the 20th of July. The y-axis stands for the ampacity level, while the x-axis for the time of the day. The blue solid line stands for the real ampacity level, while the red dashed line for the predicted one. Obviously, a positive value either in Table 3.1 or at Figure

3.4 indicates line overloading i.e. congestion.

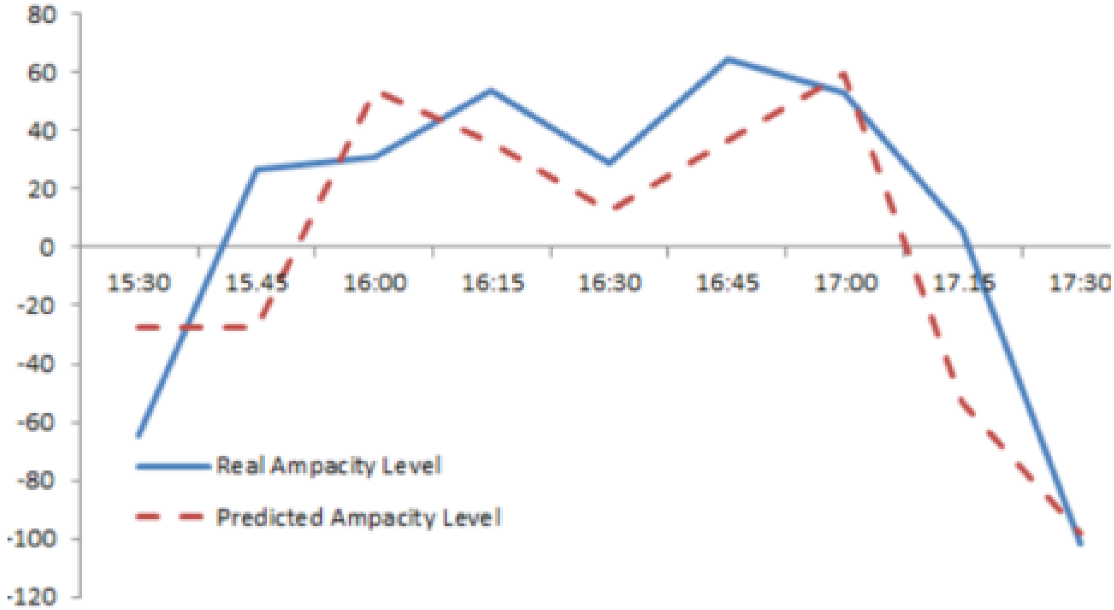

Figure 3.4: Real against predicted ampacity level for the July 20th for the whole congested period of that day (x:amperes,y:amperes)

As it can be seen from both the table and the figure, our method is able to follow the real ampacity level on the phase and more importantly, able to correctly identify line overloading. Even when misses to identify the overloading or falsely identifies its existence, it is very marginal. In both situations, measures for the relief of the system have to be taken since either there is indeed a line overloading and the method indicates that there is not for a while or the opposite.

B. Winter, spring, autumn

The same findings are presented for the rest of the seasons in Figure 3.5 and Table 3.2. During this period of the year, especially during winter, we can face line overloading either during the morning or during the afternoon and maybe early night. Since we are simulating a residential area, at these hours of the day the residents are present in their homes. Consequently, there is no congestion around noon since most of the people work at that time.

The presented results prove the findings and the impression we had previously obtained. The method is able, in most of the times, to correctly identify line overloading when indeed exists and closely follow the real ampacity level when not. Once again, any divergence from

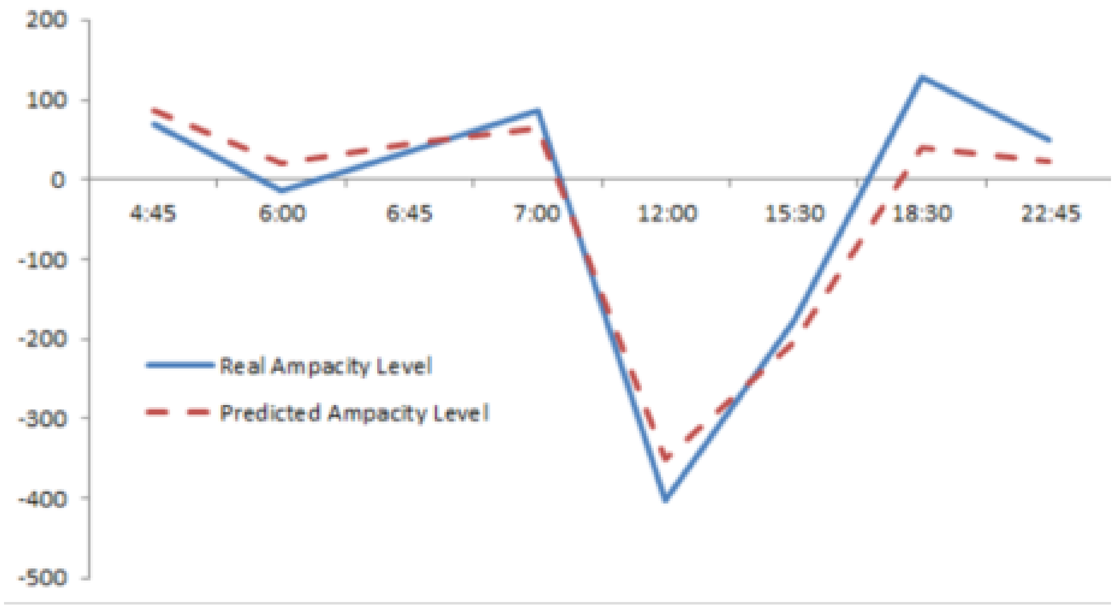

Figure 3.5: Real against predicted ampacity level for the July 20th for the whole congested period of that day (x:amperes,y:amperes)

the real ampacity level is very small. In other words, the method is capable to indicate risky time periods where precautionary measures have to be taken.

It must be finally mentioned here, that the main reason for any fault prediction, regarding the height of the ampacity level, is lying on the design of the fuzzy sets related to the fuzzy parameter of “price” and more precisely on the range of these fuzzy sets. In this study, these ranges kept constant all time. For instance, during winter, spring and autumn the fuzzy set named ”veryLow” is ranging from 421 to 440 cents/kwh. The same stands for all the other sets. The problem is that the height of the prices differs from day to day. As a consequence, a price considered as high for one day may be medium or even low for another. This difference is a source of inconsistencies and the main cause of fault predictions. A solution towards this problem would be the implementation of a dynamic monitoring system where the ranges of the fuzzy sets change from day to day following the system’s dynamics.

3.3.3 Conclusions

The aim of this study was to develop and evaluate a phase monitoring method based on fuzzy logic theory to monitor and alert in time about ampacity overloading at the distribution system. The use of a method so close to the human reasoning and far from strict programming languages is by itself very challenging. Moreover, fuzzy logic has been vastly used for control purposes, but according to our knowledge very few times for pure forecasting applications.

The presented results show that the use of fuzzy logic in this area is indeed very promising and further research is worthwhile. The proposed method was able to correctly identify phase overloading in the most of the cases. Moreover, it was able to closely follow the real ampacity level during the whole simulated day providing that way useful information to all responsible for the safety of the power system parties. Even when overloading has been faulty identified, or the opposite, it was for a while. So, even in that case, it was able to catch the trend of the ampacity level and correctly indicate risky times where precautionary measures are needed.

As also indicated, the main reason for most of the fault predictions was the static nature of the fuzzy sets related with the fuzzy parameter ”price”. The range of these fuzzy sets kept constant for the whole simulation. On the contrary, the height of the prices differs from day to day. As explained, these differences are the main cause of fault predictions. A dynamic system able to adjust the range of each fuzzy set and follow that way the changing, from day to day, dynamics of the electricity market, sounds very promising and must be under serious consideration.

3.4 Load Management of Electric Vehicles Charging in New Generation Power Markets Based on Fuzzy Logic and the Concept of Virtual Budget

3.4.1 Experimental Configuration

Fuzzy Logic Inference system

The aim of this work was to develop and evaluate a fully automated controller based on fuzzy logic in order to schedule the charging time of EVs located on a distribution feeder. This action will result in a feeder resilient to faults and congestion issues. The controller accepts two crisp values as input parameters, namely ”Time_of_the_Day” and ”Virtual_Budget” and determines the charge rate of the EV in a procedure depicted in Figure: 3.6.

The term ”virtual_budget” is used to declare not the specific amount of money the user is able to pay for electricity usage, but the amount of money he is willing to pay. To that end, the ”virtual budget” of each user is not binding for the controller. Instead the con-

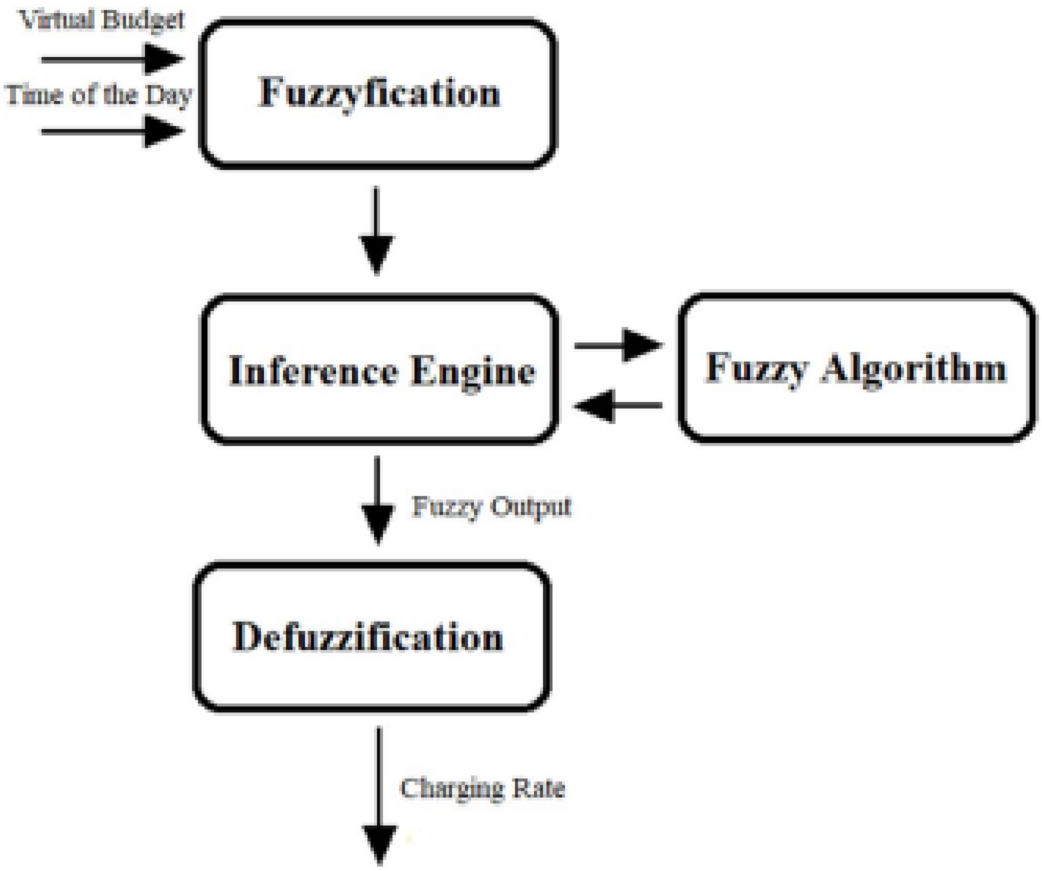

Figure 3.6: Fuzzy Logic Inference System for EV controller

troller will try to keep the final monthly electricity charge around this amount. The parameter ”time_of_the_day” on the other hand is used in order to model the congestion on the distribution feeder. Of course, congestion is something tangible but the time of its occurrence is not, although there is an intuition about it. Special events like New Years Eve, the Super Bowl final and, in our case, the time the users decide to plug-in their EVs can cause excessive demand with devastating effects for the system.

These two parameters are fuzzified into two different fuzzy sets through equation 3.4. Both the fuzzy sets used in this study follow the form depicted in Figure 3.1. To that end, we have the fuzzy sets ”Very_Low”, ”Low”, ”High” and ”Very_High” for the input parameter ”Time_of_the_Day” that correlate the time the user plugs-in his EV with an amount of congestion. Respectively, the fuzzy sets ”Vey_Small”, ”Small”, ”Medium” and ”Big” have been used for the input parameter ”virtual_budget” to declare the amount of money the user is willing to pay for electricity. Returning to equation 3.1, it represents an implication operator, which in our case is the mandani-min implication operator, through which the fuzzification process takes place. There are approximately 40 different operators in the literature. The choice of the appropriate one is at the discretion of the designer and depends on the functionality of the application.

The heart of the fuzzy inference system, is the fuzzy algorithm which is the place where the fuzzy output is produced. It has the form of ”if ... then ... ” rules, called linguistic descriptions. As the term fuzzy output implies, this output is again a degree of membership into prespecified fuzzy sets. In our case these fuzzy sets, having the triangular shape form described earlier, indicate the amount of adjustment that has to be made at the charge rate of each EV to properly manage the demand; namely ”Nothing”, ”Small”, Medium”, ”Big” and ”Very_Big”.

Finally, the crisp amount of the adjustment to be made is computed by a process called defuzzification. Likewise the fuzzification process there is more than one approach depending on the functionality of the application. The centroid method which is depicted in equation 3.5 has been chosen in this study.

Simulation Platform and Distribution System

The implementation of the fuzzy logic controller, as well as, all the simulations and the experiments of this study have been conducted using the GridLAB-D simulation platform. As mentioned earlier the powerflow module allows us to accurately solve the power flow equations and the market module allows us to implement a market environment where producers and consumers interact with each other in an auction based environment to form the price of electricity usage. More specifically, the producers offer the power they produce in a price that represent the marginal cost of producing the offered amount and the consumers submit bids reflecting their willingness to pay in order to feed their devices. The devices, are controlled by controllers which bid on their behalf and adjust their operational set points to reflect their behavior in the new market conditions formed after the auction process. It must be mentioned here that the devices that bid in a price lower than the cleared one, are not cut-off. They operate for less or more time in a try to minimize or maximize the loss or the benefit of the customer respectively.

The IEEE 13-node radial distribution feeder [32] with 1247 residencies depicted in Figure 3.3 was used in this study. Each residency is equipped with a water heater, a Heating Ventilation and Air Conditioning (HVAC) system and lights with slightly different time schedules in order to form the energy needs of their occupants. Each residency accommodates four individuals, while 559 among them own an EV. The feeder is also equipped with 4 distributed power generation units with production capacity of 244 , 500 400 and 100 kW offered in the price of 0.3, 0.11, 0.25 and 0.15$/kWh$.

3.4.2 Results

Before we actually discuss the results, Figure 3.7 is given in order to indicate the necessity of taking load management measures in the presence of EVs.

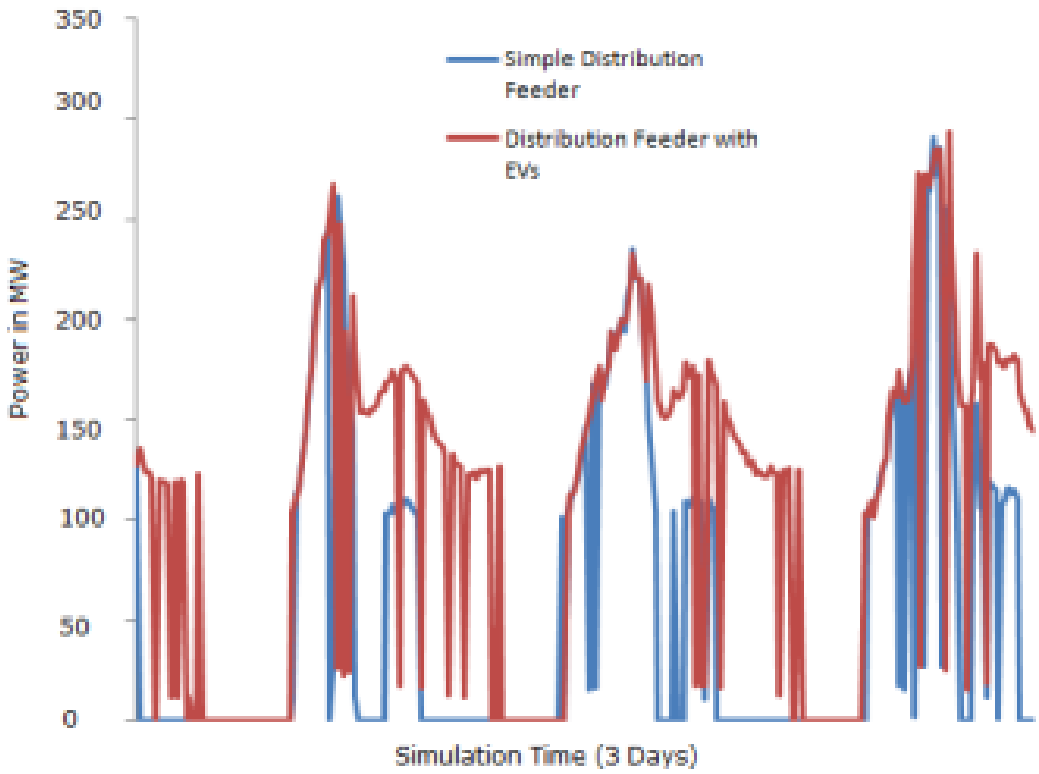

Figure 3.7: Measured Real Power with and without EVs (x:time,y:MW)

Figure 3.7 shows the measured real power of the distribution feeder with and without the presence of EVs. All the measurements have been taken at the node 632. This node is placed at the line that connects the feeder with the transmission level and thus we have a clear picture of the load of the whole feeder. The x-axis stands for the simulation time, while the y-axis for the demand in MWs. The blue line represents the load of the feeder without EVs, while the red line the opposite.

It is clear that with the presence of EVs a new peak is started to formed at 5:00 pm just after the already formed peak around 13:45 pm. In addition, a great deal of demand, which did not exist before, begins to form until 4:30 am when people un-plug their EVs to go to their jobs. It is remarkable the fact that in this experiment only a portion of the whole population possesses EVs; 559 out of 4,988 individuals. Clearly, in the near future following the expansion of EVs this issue will be of great importance.

In Figure 3.8 the reader can see a comparison of the measured real power of the feeder when the EVs are charged with the typical way without any exterior intervention and when the fuzzy approach described in the previous section is applied for the first 4 days of August. In Figure 3.9, only the 2nd day of the same month is presented for clarity. In both figures the blue line represents the load of the feeder when the EVs are charged at will, while the green line when the charging process is controlled via the proposed fuzzy logic controller.

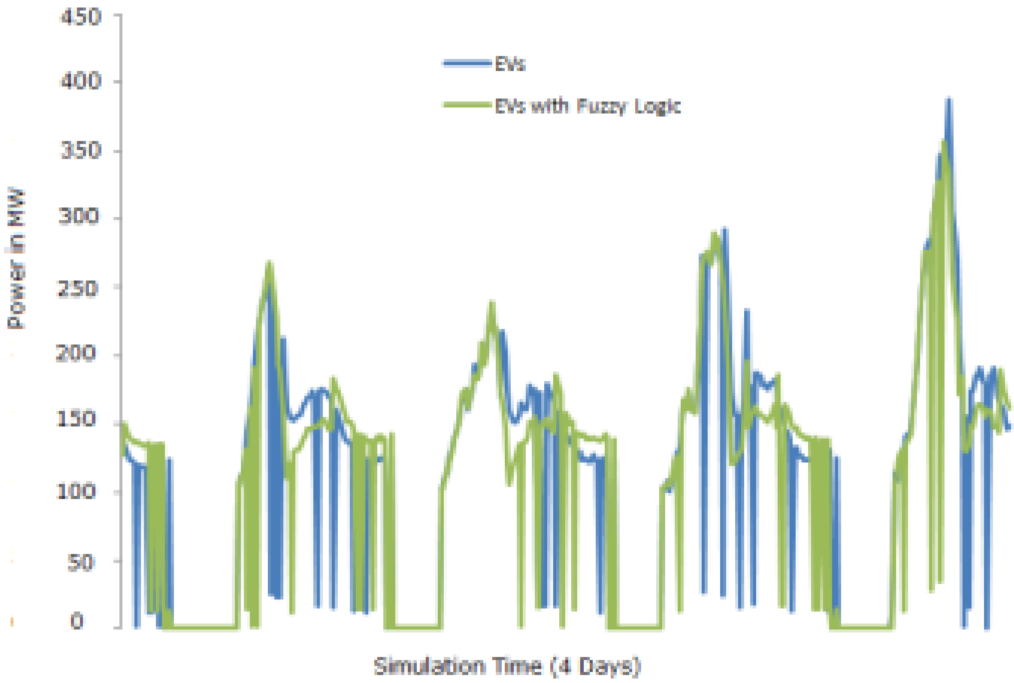

Figure 3.8: Measured Real Power with EVs controlled by fuzzy logic controller (green line) and not (blue line) for the 4 First Days of August (x:time,y:MW)

As it can be seen, the demand formed at 5:00 pm and lasted until 11:00 pm is properly managed as the EVs are scheduled to charge after 11:00 pm when the overall demand on the feeder is much lower. This is of great importance since the time period between 5:00 pm and 11:00 pm is the time when people return home, plug-in their EVs and of course use the electric appliances of the residence to cook, cool their house etc. It is clear that with the spread of the EVs in the near future this time period will be a threatening bottleneck point for the system. It

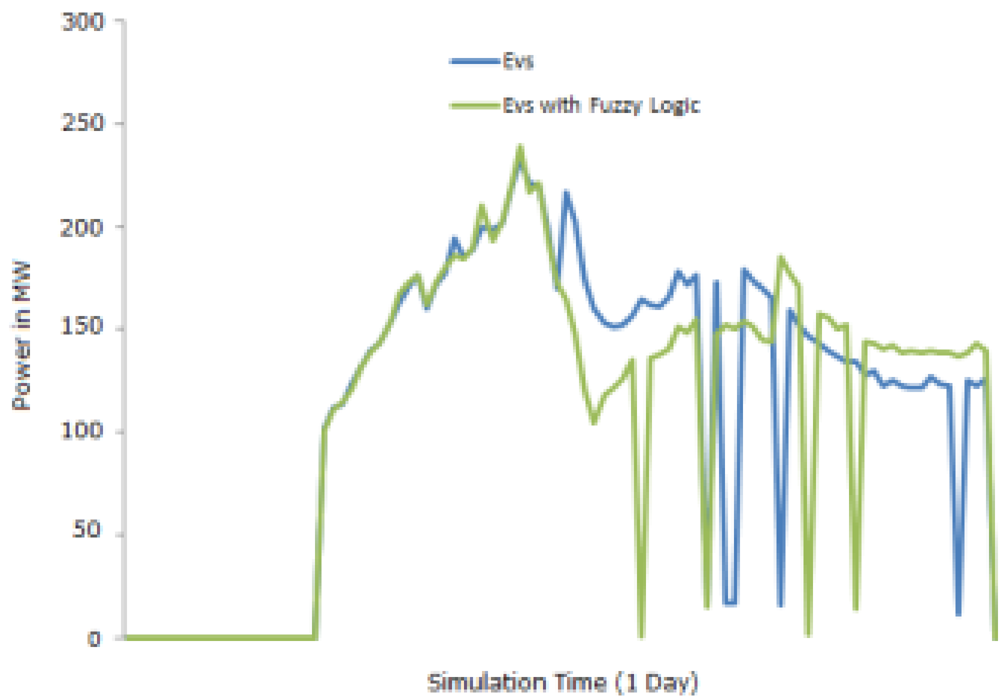

Figure 3.9: Measured Real Power with EVs controlled by fuzzy logic controller (green line) and not (blue line) for the 2nd Day of August (x:time,y:MW)

is worth mentioning that there was no EV that failed to charge and serve its purpose that is to transport its user. The fuzzy logic rules, based on which the controller performs its task, were formed in a way that balances and distributes evenly the charging time. In other words, the amount of demand that was reduced during the critical times was approximately increased during the no critical period.

The same findings observed for the whole simulation. The first 4 days of June are also presented in Figure 3.10. These cases, were selected because of the excessive demand that observed during the summer period.

It must be mentioned that there were a few days, let's name them special from now on, where the pattern of consumption didn't exactly follow the usual pattern of "Gaussian type" distribution. They had a constant peak for a long period of time like the one depicted in Figure 3.11. These particular days were of course less congested than the two cases presented earlier, and consequently no congestion related issues were noticed. On the other hand and due to the specific type of "if...then..." rules the controller did not operate as expected. After all, there was no a single point of excessive demand to be evenly distributed in less congested periods and thus, there is no need for any type of scheduling.

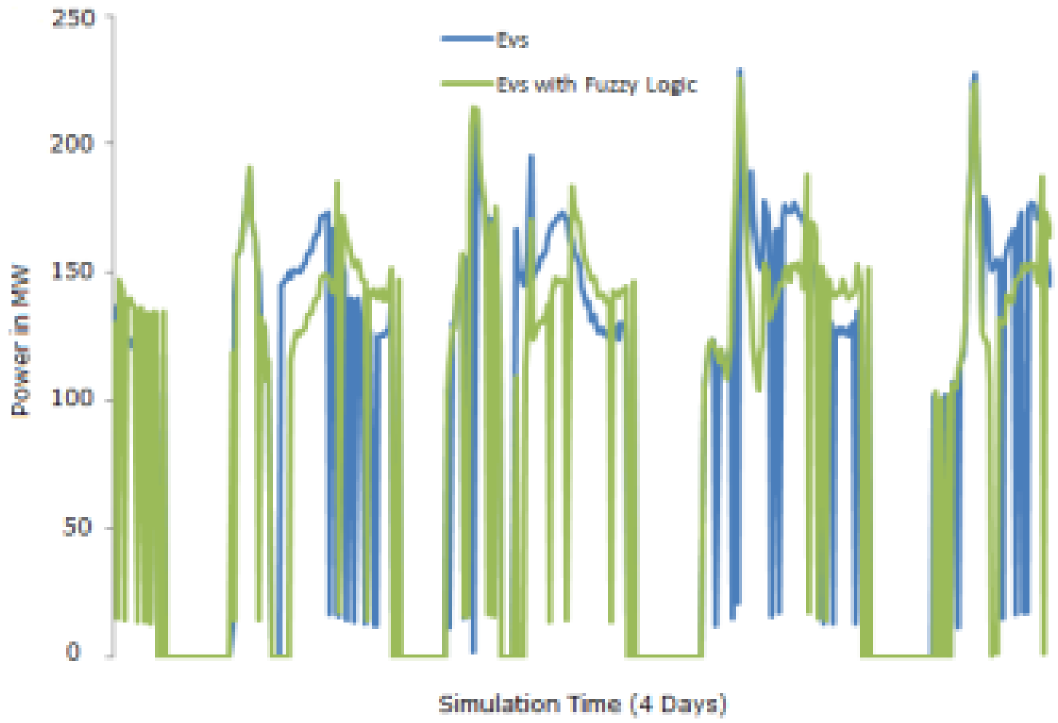

Figure 3.10: Measured Real Power with EVs controlled by fuzzy logic controller (green line) and not (blue line) for the 4 First Days of July (x:time,y:MW)

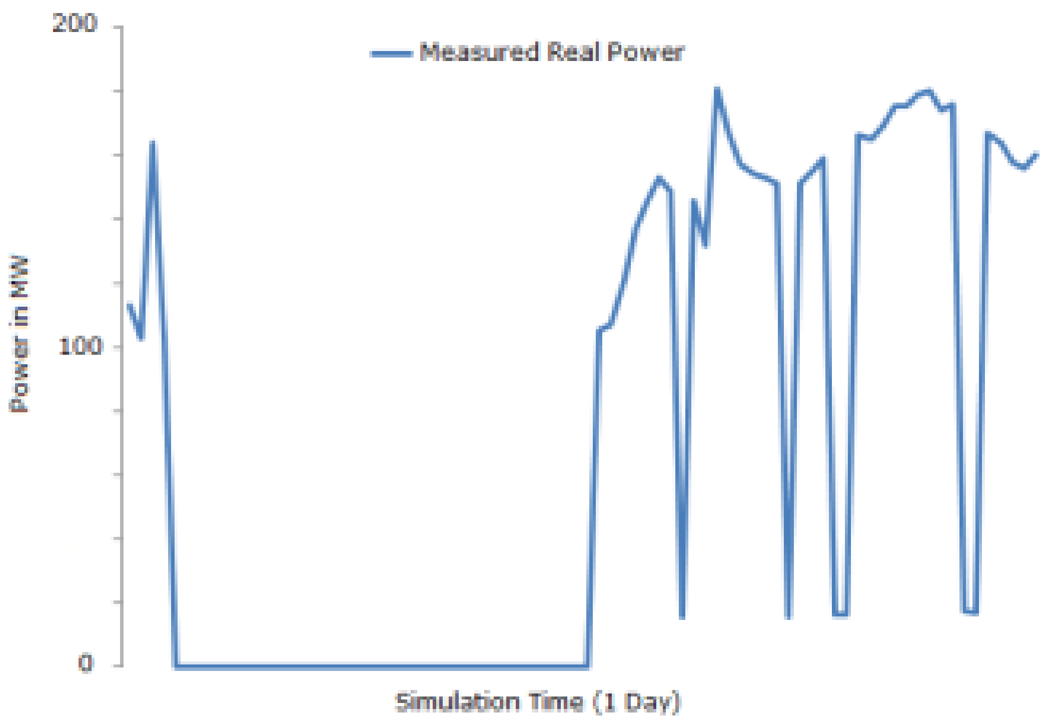

Figure 3.11: Measured Real Power for special days (x:time,y:MW)

3.4.3 Conclusions

The aim of the current study was to develop and evaluate a fuzzy logic controller in order to schedule the charging time of EVs in a power market that follows the rules of competition. This is of quite importance since, as it was observed, the excessive demand that is introduced from the EVs is able to form congestion in periods where issues of this kind had never previously seen. Moreover, we sought for a fully automated controller where the only responsibility of the user is to define his willingness to pay for electricity usage. Thus, the concept of ”virtual budget” was introduced and analyzed.

The presented results indicate the capability of fuzzy logic to assist in this type of applications. The developed controller was able to properly manage the excessive demand during the critical time and to schedule the EVs to be charged in less congested periods. Moreover, it was proved that all the EVs were charged enough in order to safely transport the users to their daily activities.

Table 3.1: Predicted against real ampacity levels on phase A (summer)

Month/Day	Temperature	Time	Price	Real ampacity level	Predicted ampacity level
07/20	82.5	15:30	21.78	-64.46	-27.5
07/20	82.51	15:45	21.48	27.02	-27.5
07/20	82.27	16:00	21.18	31.08	53.9
07/20	81.63	16.15	20.88	53.81	36.2
07/20	80.71	16.30	20.58	28.53	12.8
07/20	77.59	17.15	19.65	5.88	-53.4
07/20	76.46	17:30	19.34	-102	-98.4
07/23	81.38	15.30	21.21	-67	-65.2
07/23	80.87	15:45	20.91	-61	-77.3
07/23	80.41	16:00	20.60	-31	-4.93
07/23	80.05	16:15	20:30	-18	21.7
07/23	79.95	16:30	19.99	39	59.5
07/23	78.87	17:15	19.04	-24.32	-26.6
07/23	78.09	17:30	18.72	-35.39	-31.1
07/24	83.47	15:30	26.88	7.94	3.27
07/24	83.16	15.45	26.63	27.12	-18.8
07/24	82.98	16:00	26.39	47.03	59.5
07/24	83.06	16.15	26.15	90.54	59.7
07/24	83.58	16.30	25.90	94.74	59.7
07/24	84.04	17.15	25.10	106.88	59.8
07/24	83.95	17:30	24.80	64.78	59.8
08/06	82.72	15:30	23.04	-23.32	-23.2
08/06	82.49	15.45	22.75	-48.88	-26.7
08/06	82.20	16:00	22.47	-13.57	60
08/06	81.77	16.15	22.18	-18.63	19.5
08/06	81.28	16.30	21.89	-9.76	23.3
08/06	80.78	16:45	21.60	10.59	12.7
08/06	79.86	17.15	21.01	-17.3	-26.6
08/06	79.43	17:30	20.71	-86.40	-27

Table 3.2: Predicted against real ampacity levels on phase A (winter, spring, autumn)

Month/Day	Temperature	Time	Price	Real ampacity level	Predicted ampacity level
01/02	27.15	6:45	46.84	34.70	44.3
01/02	37.75	12:00	46.04	-401.74	-350
01/02	31.13	18:30	44.81	128	40.3
01/02	30.24	22:45	43.81	50	22.3
12/10	30.18	6:45	16.94	-48	0
12/10	31.05	7:30	46.71	-103	-21.5
12/10	34.09	17:30	44.97	51	87.7
12/10	32.03	20:00	44.42	-1	3.82
12/10	30.99	21:15	44.13	21.7	11.9
12/11	26.79	6:45	47.39	20.35	69
12/11	26.71	7:30	47.31	19.47	55.6
12/11	33.84	16:30	46.10	53.19	55.4
12/11	30.03	23:00	44.8	12.53	-17.8
01/30	37.04	6:00	46.91	-210.66	-238
01/30	36.93	7:30	46.71	-225.20	-135
01/30	35.06	21:15	44.19	-75.21	-21.5
01/30	34.52	18:30	44.81	6.94	32.5
03/01	35.69	17:45	45.07	273.3	13.4
03/01	35.07	20:00	44.48	4.54	-21.6
03/01	33.50	6:45	46.81	-152.78	-127
03/01	33.55	23:00	43.75	-98.67	-131
03/01	35.06	18:30	44.81	-55.82	32.9
11/18	30.92	6:00	46.91	34.14	11.9
11/18	30.57	6:15	46.88	24.61	11.9
11/18	30.27	6:30	46.84	-1.09	11.9
11/18	42.94	12:00	46.04	-397	-359
11/18	34.18	21:45	44.06	-35.83	-91.6

Chapter 4

Artificial Neural Network Approaches

4.1 Motivation and Related Work

As mentioned in the first chapter, the well-being of our society is strongly related to electricity production. High quality energy has to be generated and delivered constantly. The stability and the safety of the power system as a whole is paramount.

Generally speaking, Artificial Intelligence (AI) has been significantly utilized in smart grids, among all purposes, for predicting energy demand [33] and estimating energy production [34]. Power grid congestion issues in particular, are not something new in the literature. Many researchers have dealt with this aspect. Zhou et al. in [35] developed an algorithm by using tools such as linear-affine mapping, convex subsets and convex hull algorithms in order to perform transmission congestion predictions. In [36] Min et al. used sequential Monte Carlo in order to develop a method for transmission congestion forecasting. The forecasting of shadow prices was tested by Li et al. in [37]. High shadow prices, and consequently, the prediction of their time of occurrence could lead to accurate congestion prediction. In [38], Li and Bo also dealt with congestion prediction; in that study they indirectly predict congestion by finding the load level where system constrain will raise a flag pertained to line capacity. Sharma and Srivastava in [39] and [40] used AI techniques for line overloading predictions. More precisely, they used cascade neural networks in the former study and hybrid neural networks in the latter. In [41], Balaraman and Kamaraj addressed exactly the same issue and also used cascade neural networks for congestion prediction. Ivanov et al. in [42] and [43] coped with congestion prediction in terms of voltage limit violations. Multilayer perceptron neural networks were used in both of their studies with remarkable results. Glazunova in [44] developed a state estimator for short-term estimations regarding the state variables of a power system by utilizing Kalman filter-based algorithms and artificial neural networks. Finally, the work of Teeparthi and Vinod Kumar presented in [45] conducts security constrained optimal power flow analysis with the use of fuzzy adaptive artificial physics optimization algorithm.

4.1.1 Preliminaries

Training

Very briefly, a Neural Network consists of neurons. Each neuron receives an input, transforms that input using a weight and a bias parameter and produces an output. The training is a recursive procedure that updates the weight of each neuron in order to approximate the output with the highest precision.

The backpropagation method and specifically the Levenberg-Marquardt algorithm combined with bayesian regularization in order to avoid overfitting [31], [46], is able to lead fast to very efficient results. In equation 4.1, the analytical form of the weight parameter update is presented:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{4.1}$$

In equation 4.1, the vector e includes the network errors, while J is the Jacobian matrix that consists of the first derivatives of the network errors. Also, x_k and x_{k+1} are the vectors of the current and updated weights and biases respectively while I is the identity matrix. At this point, we give a short presentation of gradient descent and Quasi-Newton algorithms in order to make clear the meaning of the variable μ.

In gradient descent algorithms, which are algorithms used to find the local minimum of a function by moving towards the negative gradient at each particular point, the biases and weights are updated according to the following function:

$$x_{k+1} = x_k - \alpha_k g_k \tag{4.2}$$

where x_k and x_{k+1} are the vectors of the current and updated weights and biases respectively, while α_k and g_k represent the learning rate and the gradient of the performance function. Similarly, in the Newton's algorithms weight update is given by the following equation:

$$x_{k+1} = x_k - A_k^{-1} g_k \tag{4.3}$$

where the vectors x_k and x_{k+1} are the vectors of the current and updated weights and biases respectively and g_k is the gradient of the performance function. A_k is the Hessian matrix each entry of which is the performance index of the network's biases and weights. Comparing equation 4.1 with equation 4.3 we observe that the Jacobian matrix is used to approximate the Hessian matrix as follows:

$$H = J^T J \tag{4.4}$$

Moreover, the Jacobian matrix is also used to approximate the gradient of the performance function as it is given in equation 4.5:

$$H = J^T e \tag{4.5}$$

where, J is the Jacobian matrix and e stands for the vector that contains the network errors.

If the parameter μ in equation 4.1 is set to a value larger than zero, then equation 4.1 is turned into a gradient descent algorithm. If μ is set to zero, then equation 4.1 is coincides with Quasi-Newton. More information about these concepts can be found in [31] and [46].

Bayesian Regularization

It is well known that during the training procedure of a neural network, overfitting is very likely to occur. In this case, the neural network even if it appears to be well trained, it fails to provide an accurate output for an unknown input. In other words, it fails to generalize. To that end, it is essential to use techniques to overcome that problem. The approach used throughout all the studies of the current dissertation was the Bayesian Regularization (BR).

Briefly, given a set of inputs $\{i_1,t_1\},\{i_2,t_2\},....,\{i_n,t_n\}$ each training procedure tries to reduce the sum of squared errors (equation 4.6) where E_D is given by equation 4.7.

$$F = E_D \tag{4.6}$$

$$E_D = \sum_{i=1}^{n} (t_i - \alpha_i)^2 \tag{4.7}$$

In equation 4.7 t_i stands for the targets of the training procedure, while α_i stands for network's responses.

When it comes to regularization, the equation 4.6 is transformed into

$$F = \beta E_D + \alpha E_w \tag{4.8}$$

where E_w stands for the sum of squares of network's weights while α and β represent the objective function parameters. To that end, when dealing with regularization, the main issue is to set the proper values of the objective function parameters. The weights in that case can be updated using equation 4.9

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \tag{4.9}$$

In equation 4.9, D stands for the data set, M for the particular NN model, while w represents the vector of network's weights. P(w|α,M) is called prior density and represents the information regarding NN's weights before any data has been collected. P(D|w,β,M) represents the likelihood function that is the probability of data given the particular weights w. Finally, P(D|α,β,M) represents a normalization factor. That factor ensures that total probability is equal to 1. Returning to Bayesian approach, maximizing equation 4.9 guarantees the optimal weights which is exactly the same as τo minimize equation 4.8. It is well documented that this method prevents overfitting. All the information has been retrieved from [47]. The reader is referenced there for more information and details regarding the aforementioned equations.

4.2 Three–Phase Congestion Prediction Utilizing Artificial Neural Networks

The aim of this study was to develop and test a congestion predictor that monitors each phase of a line of the distribution system and predicts congestion line overloading. In particular, our predictor should predict/indicate the point(s) in time where the ampacity of each phase is above of its emergency rating. That is very crucial for the stability of the system because if the load remains above this limit for longer than two hour period, the line will fail. A secondary goal is to assess the amount of overloading. This information may be proved essential for the Distribution System Operator (DSO) in order to take a priori preventive actions to ensure the stability of the system. For this purpose, a feed-forward backpropagation artificial neural network (ANN) trained for a whole year by adopting the Levenberg-Marquardt algorithm has been used. The training data have been acquired through the GridLAB-D [2] simulation tool, while the MySQL database has been utilized in order to handle the vast amount of the produced data.

4.2.1 Methodology

Neural Network Architecture

A power system is inherently a complex non-linear system, and thus we adopt a feed-forward backpropagation artificial neural network, whose architecture is depicted in Figure 4.1, for implementing a congestion prediction mechanism capable of providing highly accurate predictions.

The ANN is consisting of three layers; one input with five neurons, one hidden layer with fifty neurons and one output with three neurons representing each phase of a distribution line. The output will be ”1” for each phase, when the current flow is above its emergency rating and ”0” otherwise.

Training

The Levenberg-Marquardt algorithm as described earlier was used in this study in order to train our Neural Network. Moreover, in order to improve our neural network’s efficiency and to avoid the overfitting of the input data the Bayesian regularization method has been utilized in order to identify the appropriate regularization parameters. The reader is referenced to [47] and [48] for more information.

4.2.2 Experimental Configuration

The IEEE 13-bus radial distribution feeder in Figure 3.3 was used for experimentation. We used 1306 residencies, with each residency to accommodate 4 individuals. Each residence is equipped with a water-heater, an HVAC system and lights. We also used four power generators with maximum capacity of 2400, 500, 400 and 100 kW and marginal cost 0.3, 0.11, 0.25 and 0.15 $/kWh respectively.

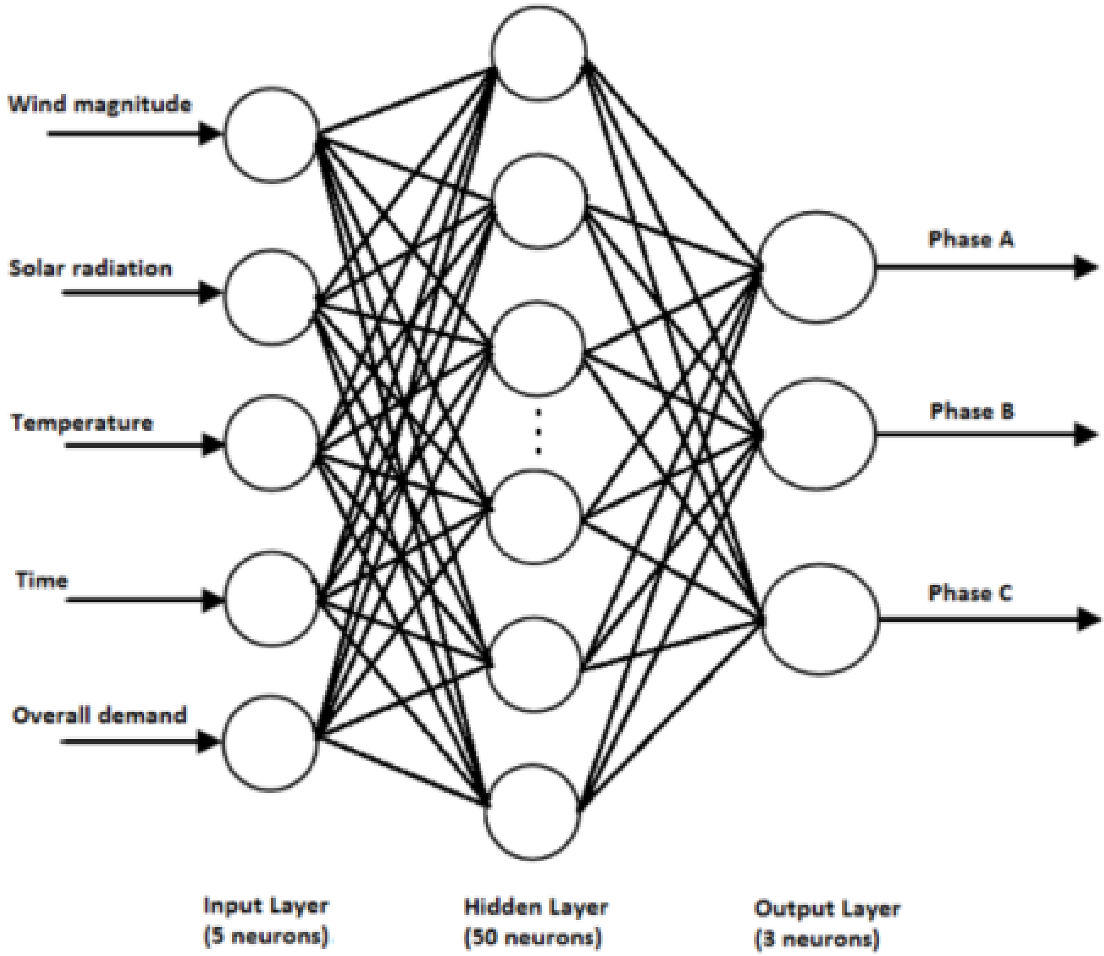

Figure 4.1: Feed-Forward Artificial Neural Network Architecture

In order to add credibility to our outcomes we adopted for our experiments a robust simulator in order to acquire both the training and the evaluation data for our ANN predictor. For this reason we used the GridLAB-D platform which is a power distribution system simulator developed by the Pacific Northwest National Laboratory (PNNL) [2]. GridLAB-D allows us to implement an energy market organized as a double-auction market. In addition, it allows implementation of Demand-Response (DR) program by encompassing the thermostatically controlled devices (HVAC systems and water-heaters). This DR program is implemented through the GridLAB-D's object "transactive_controller" which automatically adjusts the temperature set-points of the thermostatically controlled devices and bids on behalf of the devices in intervals of 15 minutes. In that environment, the customer has only to determine his/her desired temperature set point and to identify a maximum and minimum temperature defining this way a comfort zone [49].

The neural network was trained using the first two weeks of every month of the year 2000 and its performance was tested on data taken for the following year, i.e., 2001. The input data, collected every fifteen minutes, are consisted of the following parameters:

- Wind magnitude
- Solar radiation
- Temperature
- Time
- Overall demand

The implementation and training of the neural network is performed in Matlab, while the MySQL software is used for storing and handling the large volume of generated data.

4.2.3 Results

In order to evaluate the accuracy of our model, we lowered the emergency limit of the line 630-632 (see Figure 3.3) to artificially induce congestion in this line. This is the line that actually connects the distribution feeder with the transmission system, so the stability of this particular line in order to be able to continuously support the overall load of the feeder is of high priority. Moreover, in order to evaluate the performance of our predictor, we used two indices; namely the positive and negative predictive values (PPV and NPV) that given by equations 4.10 and 4.11.

$$PPV = \frac{number of true positives}{number of positives} \tag{4.10}$$

$$NPV = \frac{number of true negatives}{number of negatives} \tag{4.11}$$

By true positive we mean that congestion was actually existed and correctly identified by the predictor. Accordingly, true negative means that congestion was not existed and correctly not identified by the predictor.

In Tables 4.1 and 4.2, the comparison of the predicted congestion time points against the simulated ones are presented. Furthermore, Tables 4.3 and 4.4 provide the results regarding the PPV and NPV indices. It must be mentioned that in order to create a scenario as realistic as

Table 4.1: Real Vs Predicted Congestion Time in Phase A

Phase A		
Day	Congestion presence at	Congestion prediction at
1/1/2001	00:30 to 02:00 06:15 to 06:45 16:00 to 18:45 19:30 to 21:00	16:15 to 18:15 19:45 to 21:45
1/2/2001	04:30 to 05:30 06:15 to 07:15 16:15 to 21:15 21:45 22:15 to 22:45	04:30 to 05:15 06:30 to 07:30 16:15 to 21:15
1/8/2001	16:30 to 17:30	16:45 to 17:15
1/17/2001	17:00 to 17:30	16:45 to 17:00
8/6/2001	16:45	16:15 to 17:00
11/18/2001	05:45 to 07:15	06:00 to 7:00
12/1/2001	05:15 to 08:45	05:15 to 08:15
12/10/2001	16:30 to 18:45 21:15 to 21:45 22:30 to 22:45	16:15 to 17:45 20:15 to 21:45

possible, we set the line limit to a specific set point and not too high. As a result, we observed that there were not so many congested days. However, in Tables 4.1-4.4 we present results obtained for some of the few congested days. It must be also mentioned that the dates in this paper are written in the US format (month/day/year).

The distribution of residences on each phase is not balanced. The majority of the residencies are distributed on the phase A, and thus, congestion observed in most of the cases pertaining this phase and secondarily on the phase B. There wasn't any case in which congestion occurred or predicted on phase C.

The highest error, meaning that our mechanism missed to detect a long time interval of congestion, is observed for the date 1/1/2001. This day is the new Years' Eve and it can

Table 4.2: Real Vs Predicted Congestion Time in Phase B

Phase B		
Day	Congestion presence at	Congestion prediction at
1/2/2001	20:00 to 20:15	20:15 to 20:45
11/18/2001	06:00 to 06:30	06:30 to 06:45
12/1/2001	05:30 to 07:00	05:45 to 07:45

Table 4.3: PPV and NPV Values for Phase A

Phase A		
Day	PPV	NPV
1/1/2001	83	80
1/2/2001	93	89
1/8/2001	100	97
1/17/2001	33	97
8/6/2001	25	100
11/18/2001	57	100
12/1/2001	92	97
12/10/2001	64	92

Table 4.4: PPV and NPV Values for Phase B

Phase B		
Day	PPV	NPV
1/2/2001	33	96
11/18/2001	50	97
12/1/2001	77	98

be considered as a special day in the sense that it cannot exactly follow system dynamics regarding the training procedure that have been selected. These days have to be handled separately: to that end, a neural network is trained for the respective days of three previous years for dates 30/12, 1/1 and 1/2.

Our predictor even if it doesn't predict the exact time of congestion it efficiently captures the trend of the line loading as it can be seen in Figure 4.2, 4.3 and 4.4. This becomes more apparent by observing the PPV values. In the aforementioned figures, the x-axis represents the time of the day, while the y-axis represents the loading of the line in amperes. A negative value means that the current of the line is "y" amperes below the emergency rating of the line, while a positive value indicates the opposite.

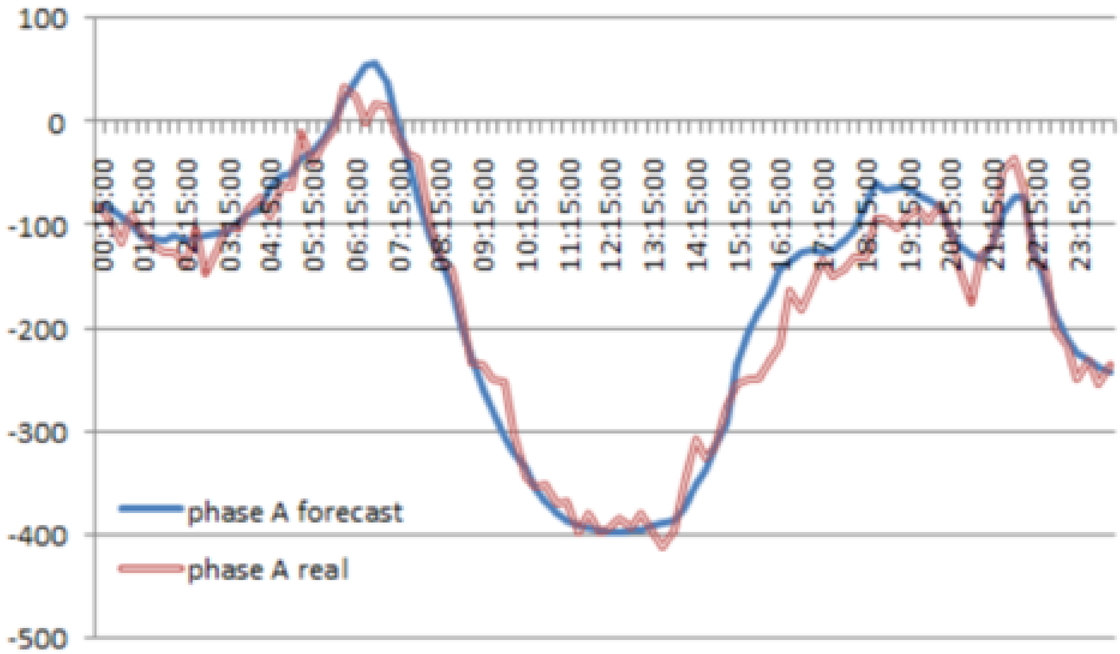

Figure 4.2: 11/18/2001 Trend of line loading on phase A (x:time,y:amperes)

This should be very helpful for the system operator in order to categorize specific periods as periods of increased risk in terms of congestion presence. As it can be also seen by these figures, the reason for which our mechanism does not predict the actual time of congestion occurrence is primarily because of the noise on the real line loading signal.

As it can be seen in Tables 4.5 and 4.6, our methodology gives an approximation of magnitude and overloading in the line. In addition, it provides a close estimation of the time where this overloading will occur; reaffirming our allegations that it will be a useful tool in the hands of System Operator. The operator will be able to identify high risk periods assessing at the same time the scale of danger as well. In these tables, the time of peak occurrence as well as the predicted peak, expressed in terms of amperes, and the time are depicted.

A positive value means that the line loading exceeds the emergency limit by the amount of the corresponding value, while negative value means that the line loading is below the

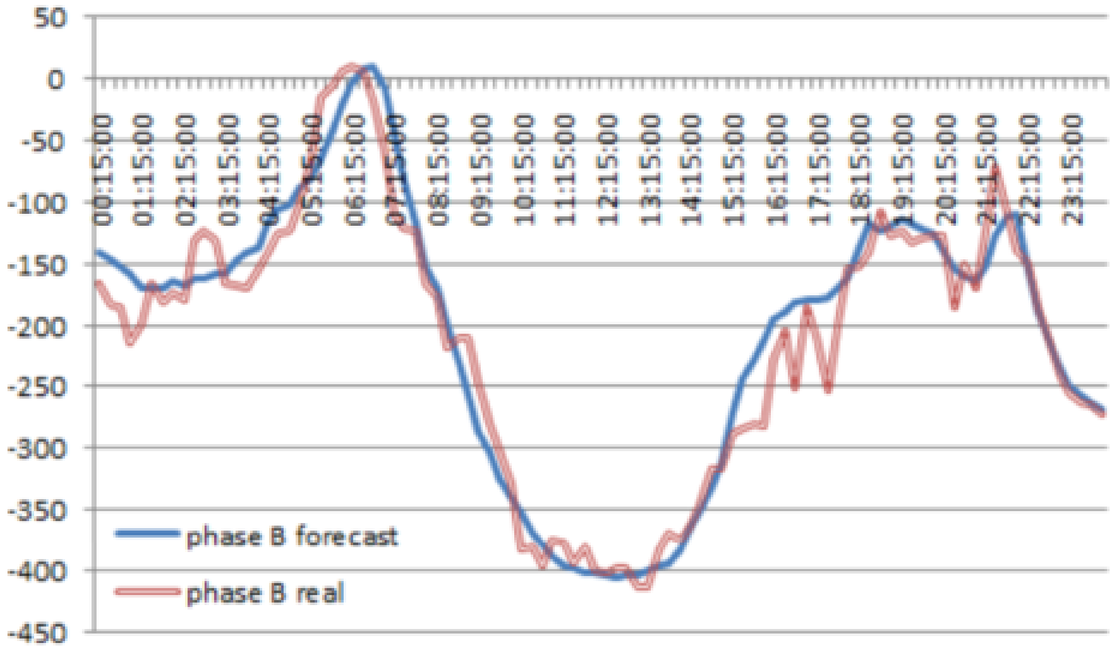

Figure 4.3: 11/18/2001 Trend of line loading on phase B (x:time,y:amperes)

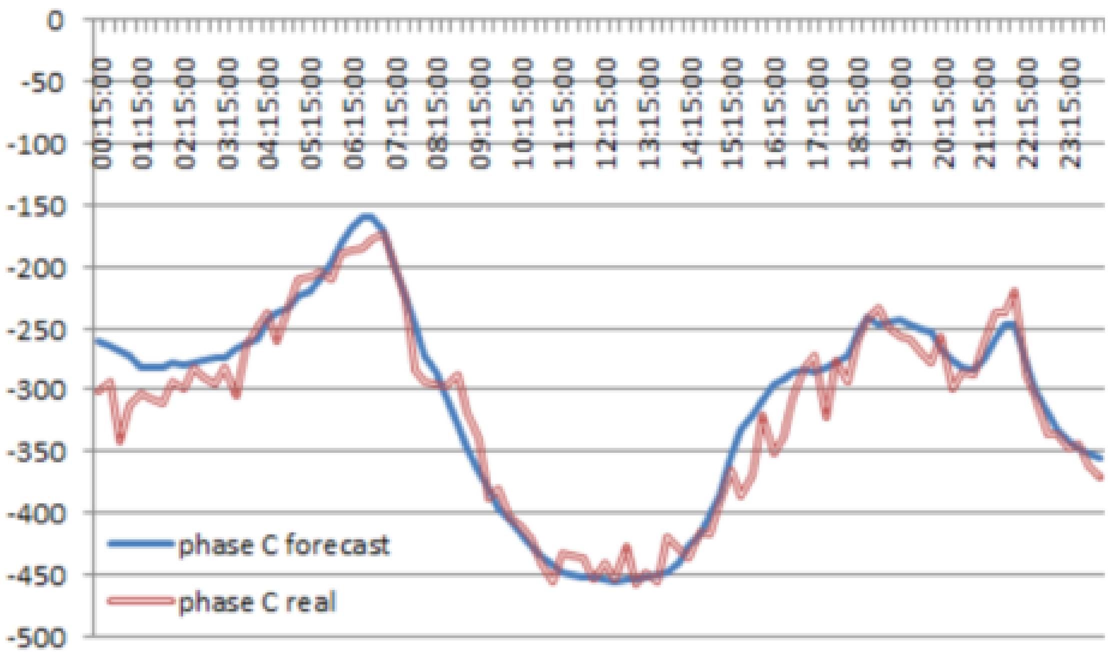

Figure 4.4: 11/18/2001 Trend of line loading on phase C (x:time,y:amperes)

emergency limit of the line by the corresponding amount. Again, as in the previous results, our method fails to provide an accurate approximation regarding the 1st of January for the reasons explained previously.

Table 4.5: Real and Predicted Peak on Phase A

Phase A				
Day	Time	Peak (Amp)	Predicted Time	Predicted Peak (Amp)
1/1/2001	00:45	+297	17:15	+118
1/2/2001	19:15	+130	17:30	+90
1/8/2001	16:45	+25	17:00	+17
1/17/2001	17:00	+16	16:45	+27
8/6/2001	16:45	+10	16:45	+29
11/18/2001	06:00	+34	06:45	+55
12/1/2001	06:00	+132	06:45	+164
12/10/2001	17:30	+51	16:45	+81

Table 4.6: Real and Predicted Peak on Phase B

Phase B				
Day	Time	Peak (Amp)	Predicted Time	Predicted Peak (Amp)
1/2/2001	07:00 20:15	+29 -1	07:00 20:15	+18 +8
11/18/2001	06:15	+9	06:45	+10
12/1/2001	06:45	+86	06:45	+106

4.2.4 Conclusions

The main goal of this study was the development of a three-phase line predictor capable of providing accurate predictions inside the big data environment. This was realized through a feedforward backpropagation neural network developed and trained in Matlab. The data regarding the training and the evaluation were obtained by the GridLAB-D simulation tool and handled through the MySQL software.

As it came out, our method always follows the trend of line loading and indicates the right time of congestion presence giving the possibility to the system operator to accurately define high risk periods and properly react to prevent it. It is also able to approximate the true overloading with high enough accuracy providing a very useful tool to that direction.

4.3 Backpropagation Neural Network for Interval Prediction of Three-phase Ampacity Level in Power Systems

The aim of this study, which is actually an extension of the work presented in the previous section, is to develop a method for accurate ampacity level predictions with the ultimate goal and effort to be towards accurate line overloading predictions ahead of time. Monitoring of the ampacity level and a precise indication about high risk time periods where congestion may occur is of high importance. A current flow beyond the line limit, called thermal limit, for more than two hours will result in a line collapse with severe and unpredictable consequences. Moreover, the high volatility of load in power systems with DR programs renders any predictive task related to the overall load on the grid difficult and challenging. Thus, an approach of constructing Predictive Intervals (PIs) in order to quantify any uncertainty should be more appropriate. The idea is to use a NN that does not only predict the ampacity level of each phase of a line, but also a PI related to each particular prediction.

To that end, a feedforward backpropagation neural network was developed and trained with the Levenberg-Marquardt training function. We also used Bayesian regularization in order to avoid the data overfitting during the training stage. Regarding the training and the evaluation datasets, both sets were acquired from the GridLAB-D [2] simulation engine. The same platform was used in order to design a distribution feeder and a market framework that incorporates active end-users that participate in DR programs. Given that nowadays new generation power systems are operating in a big data environment, the MySQL software was chosen as the most suitable for storage and retrieval of the vast amount of data generated by our simulations.

4.3.1 Methodology

Uncertainty Quantification

The load volatility introduced in new generation energy markets, as a result of the use of DR programs, is high as depicted in Figure 4.5. In every market clearing period the smart devices, responding to price signals, change their operation status trying to maximize their benefit, according to some user defined parameters and thresholds. In Figure 4.5 the ampacity level of a phase is depicted as a red curve. The x-axis represents the time of the day, while the y-axis represents the ampacity level. A positive value signifies overloading, while a negative one the opposite. It can be easily observed that the ampacity level continuously changes over time (in the current work every fifteen minutes) without any specific pattern. Moreover, due to the fact that the trend is to reduce as much as possible the time period of each market clearing process, this volatility can be encountered every five, or even less, minutes [1]. It is obvious that any accurate predictive method in such an area has to deal with and quantify these uncertainties.

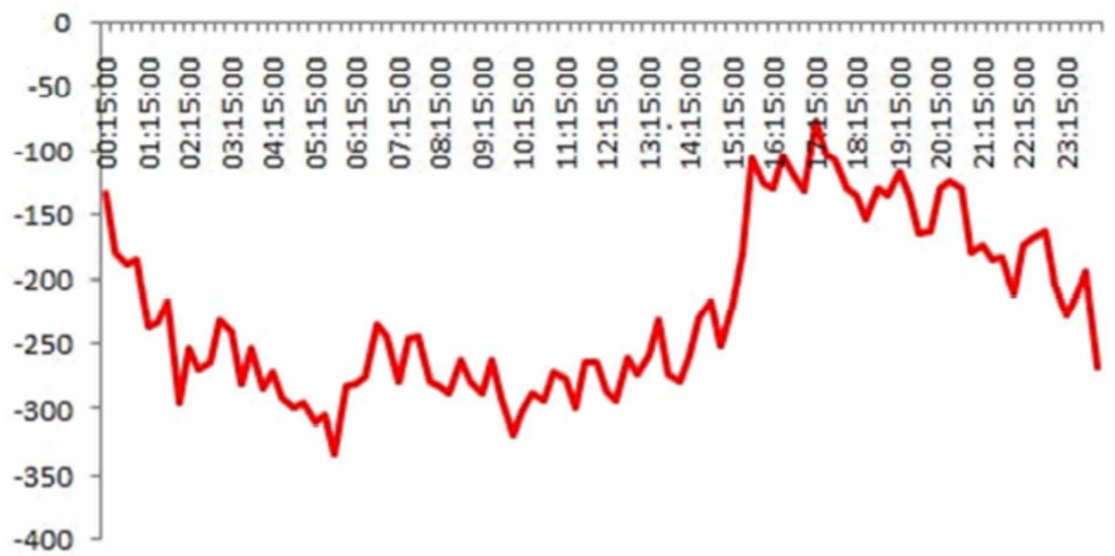

Figure 4.5: Ampacity level of a phase during a day (x:time,y:amperes)

To that end, in this study, we used the standard deviation of the target data with the assumption that the standard deviation of the predicted values will be similar for suchlike conditions. Therefore, the target data, that is the ampacity level of each phase, are pre-processed in order to obtain their standard deviation for a specific period of time. After this procedure, the target data have been expanded and contain not only the ampacity level of each phase, but also its standard deviation for a certain period of time and then are used to train the NN. The whole process is depicted in Figure 4.6

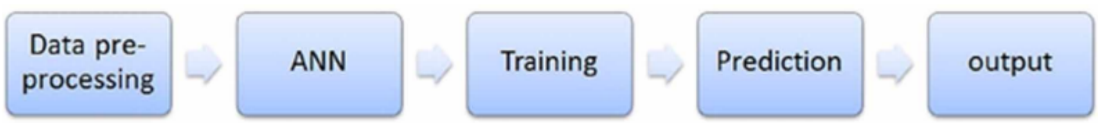

Figure 4.6: Flowchart of the proposed method

After the training, the NN predicts not only the ampacity level of each phase but also its standard deviation at that particular time. The PI is then defined by adding and subtracting the standard deviation by the predicted value as shown in equation 4.12:

$$PI = predictedvalue \pm standarddeviation \tag{4.12}$$

Neural Network's Architecture and training

Power systems and especially the ones that provide DR capabilities to the end users are by nature non-linear systems. NNs, which have already proven their capabilities in this domain, selected as the most appropriate approach for our purposes. The selection of the proposed architecture, which is depicted in Figure 4.7, was based on the trial-and-error procedure where a plethora of different NN architectures were tested regarding their accuracy and effectiveness to both anticipate individual points, as well as their ability to define precise PIs.

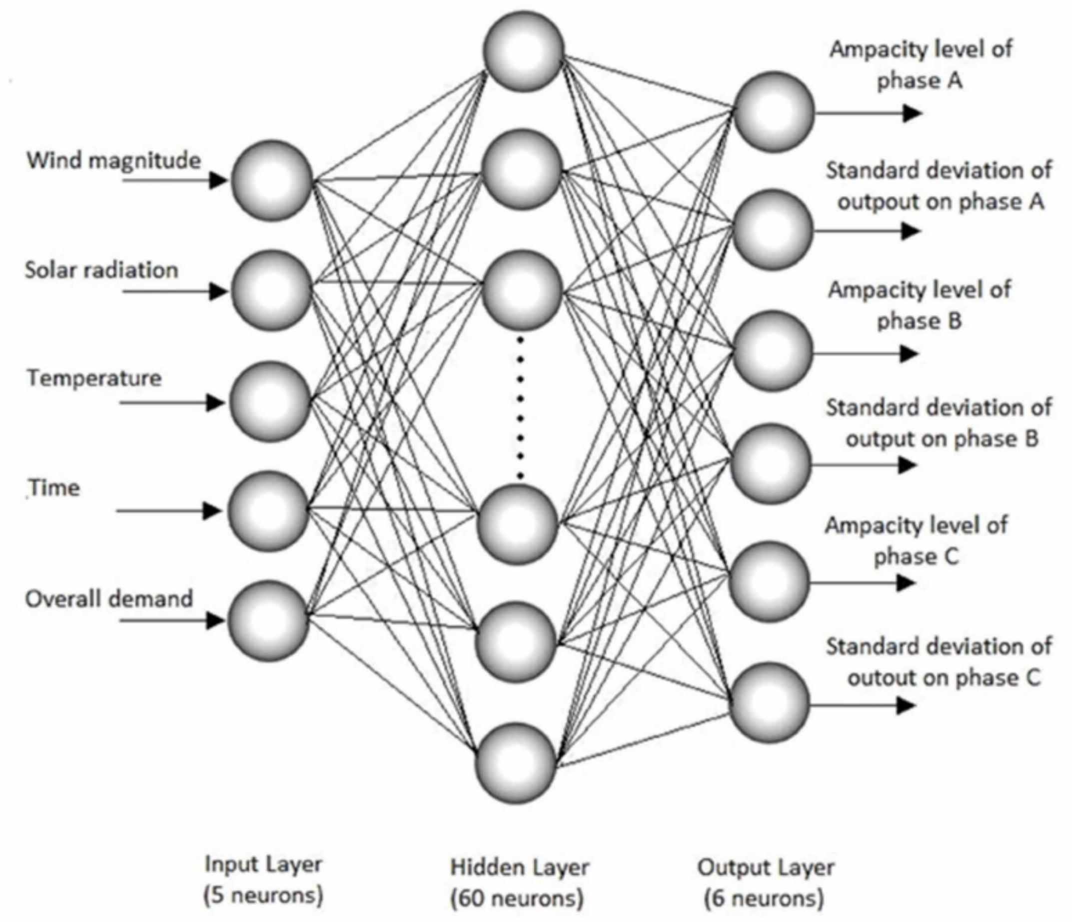

Figure 4.7: Neural Network architecture

The selected NN is a three-layer backpropagation neural network with five inputs and six outputs (5-60-6). As mentioned earlier, the outputs stand for the ampacity level of each

Table 4.7: Input/output NN's parameters

Inputs	Outputs
Wind magnitude	Ampacity level of phase A
Solar radiation	Standard deviation of output of phase A
Temperature	Ampacity level of phase B
Time	Standard deviation of output of phase B
Overall demand	Ampacity level of phase C
	Standard deviation of output of phase C

phase, as well as its standard deviation prevailing at that particular time interval. Both input and output parameters are aggregated on Table 4.7.

The Levenberg-Marquardt method was selected as training function. As mentioned earlier, a big issue regarding nonparametric predictive methods, suchlike NNs, is the overfitting of the training data during the training stage. The method, in such a case, fails to produce a good approximation of the real value of unknown input, although its performance when it is fed with known input is remarkable. The Bayesian regularization method was selected in order to tackle with this issue.

4.3.2 Experimental Configuration

The focus of this work, as mentioned earlier, is to provide accurate predictions of phase ampacity overloading in new generation power markets where end-users are eligible to participate in demand response programs. This task is very difficult in this type of market frameworks because the customers participate and form energy usage prices in every market clearing period. Moreover, customers differentiate their demand management based on each prevailing market price and thus there is no more a specific pattern of energy usage. Consequently, there is high load volatility that affects any predictive applications because of the high uncertainty introduced in the system. Furthermore, the shorter the market clearing period is, the greater the load volatility.

Demand-Response Programs

Based on the above observations, the GridLAB-D [2] simulation platform was selected as the most appropriate for our intentions. It was used to build the distribution feeder, as well as a market framework organized as pool-based double auction market that offers demand-response programs to the end-users.

The capability of demand-response in GridLAB-D is provided through the transactive_controller object which is responsible for the operational management of the thermostatically controlled devices; namely the Heating Ventilation and Air-condition (HVAC) system and water heater. The control of these devices is based on the price signals received from the market. The controller, based on the prevailing electricity usage price (cleared price) adjusts the thermostat set point of the devices trying to maximize their benefits. The set point of an HVAC system during summer (cooling mode) for instance, will be moved higher than the desirable if the cleared price is high in a try to reduce consumption and save money. One the other hand, if the price is low, the set point will be moved lower trying to exploit the low prevailing price. Moreover, the transactive_controller is responsible for the bidding strategy of these devices. Based on pre-specified values set by the user and express his/her willingness to pay for different levels of price and his/her comfort zone regarding the operation of the device, defines the price on which will bid in the market. The reader can find more information in [49].

Distribution Feeder

The distribution system designed for this work is depicted in Figure 3.3 [32]. It is the IEEE 13-bus radial distribution test feeder which was designed to accommodate 1306 residencies and 4 Distributed Energy Resources (DER) each with a different production capacity and different production cost; the so-called marginal cost of producing real power.

Each residency has been designed to have different demand needs, is dwelled by four people, while its energy needs are comprised by three different appliances that is a water heater, lights and an HVAC system. All the information about the specific components of the feeder can be seen in Tables 4.8 and 4.9.

In all the test cases, which will be presented in the next section, the thermal limit of all the phases of the line 630-632 was reduced in order to create congestion conditions. Our intention was to create these conditions in the morning where people get up to go to work and late afternoon where they return back home. That particular line was selected due to its

Table 4.8: Production cost and capacity of DER

	Gen. 1	Gen. 2	Gen. 3	Gen. 4
Production Cost ($/kWh)	0.3	0.11	0.25	0.15
Production Capacity (kW)	2400	500	400	100

Table 4.9: Residencies' characteristics

	Total Number	Occupants	Devices
Residencies	1306	4	HVAC, lights, water heater

importance on the distribution feeder. Since it's the only line connecting the feeder to the rest of the transmission grid, a collapse there, will isolate it and the DERs alone should not be able to serve the underlying load. In other words, the serving of the energy needs of the 1306 residences is jeopardized.

Data Acquisition, Storage and Retrieval

The clearing market period was set to fifteen minutes. Consequently the data, provided in Table 4.7, were collected every fifteen minutes for a period of one year (2000) via the GridLAB-D simulation platform. Moreover, they were stored in a database developed in MySQL, mainly for retrieval purposes and also in order to be able to manage the vast amount of obtained data streams. At this point it should be noted that not all of the data used during the training of the neural network. Our intention was to develop a method primarily focusing in efficiently identifying ampacity overloading. In that sense only the most representative data, corresponding to 10269 entries, were selected and used.

4.3.3 Results

The results of our experiments are presented in this section. The goal of this study is to develop a method of constructing efficient PIs, primarily targeting in effectively identifying high risk periods of congestion. Of course, the width of the PI is of high importance. A broad PI will result in high efficiency but also in lower accuracy. A narrow PI on the other hand will exhibit better accuracy, but most probably low efficiency level.

To that end four different test cases are presented. In each test case we use different stan-

Table 4.10: Details regarding the std used in each case to construct the PIs

	std of target value every
Case 1	1 hour
Case 2	2 hours
Case 3	3 hours
Case 4	4 hours
Case 5	5 hours
Case 6	6 hours
Case 7	7 hours
Case 8	8 hours

dard deviation of the target value (ampacity level of each phase) in order to create the PIs. In the first one we use the standard deviation of the target value every one hour, in the second every two hours, in the third one every three and so on. Our intention was to assess which is the most proper in terms both of efficiency, as well as of accuracy. In other words, we aimed to create an efficient but also an as narrow as possible PI. In the rest of the section, ten randomly selected days are presented in order to assess the validity of our assumptions. The evaluation has been carried out using data obtained in the year 2001 through the GridLAB-D platform. It must be mentioned here that eight different test cases have been examined. Here, the first four are presented in detail since no improvement was observed in the last ones. For greater clarity regarding the test cases, the reader can also see Table 4.10.

Test Case 1

In the first case scenario we used the standard deviation of the target value every one hour in order to create the PIs.

On the left part of Figure 4.8, the reader can see the point predictions of ampacity overloading, while on the right part the interval prediction approach. The predictions regarding the 11^{th} of December 2001 are presented in this particular figure. The x-axis represents the time of the day. The y-axis represents the ampacity level of the phase with the zero point to stand for the thermal limit of the line. The red line represents the real value of the ampacity level of each phase, obtained via the GridLAB-D simulation tool, while the blue line represents the predicted one. A positive value indicates ampacity overloading in that particular phase by magnitude equal to the value. The opposite holds for a negative value. On the right part, the black line stands for the lower bound of the predictive interval, the red line for the upper bound and finally, the green line stands for the real ampacity level on each phase.

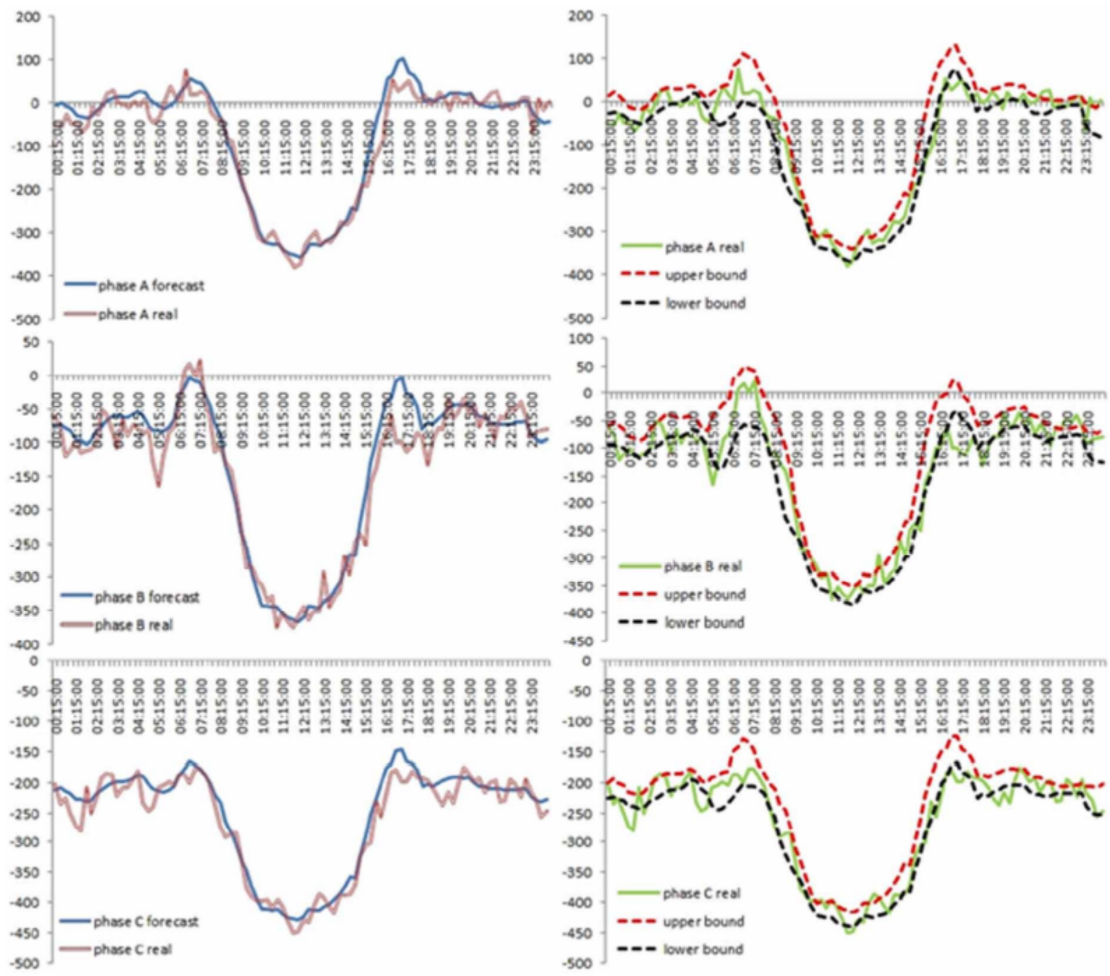

Figure 4.8: Point prediction against PI prediction of the 11th of December (1 hour std) (x:time,y:amperes)

As it can be clearly observed, cases where the plain point prediction method failed to identify overloading, the prediction interval method predicted it correctly. For instance, there is overloading in phase B from 6:15 to 7:15 totally missed from the first method. The PI method, on the other hand, correctly defines a prediction interval around that time that signifies ampacity overloading and consequently a potentially risky period of time.

As it can be seen in Figure 4.9 there are still some cases where even this approach misses some congested points. In this figure the real ampacity level, as well as the predicted by our method interval is depicted. As it appears, two congested points, the first one at 17:30 and

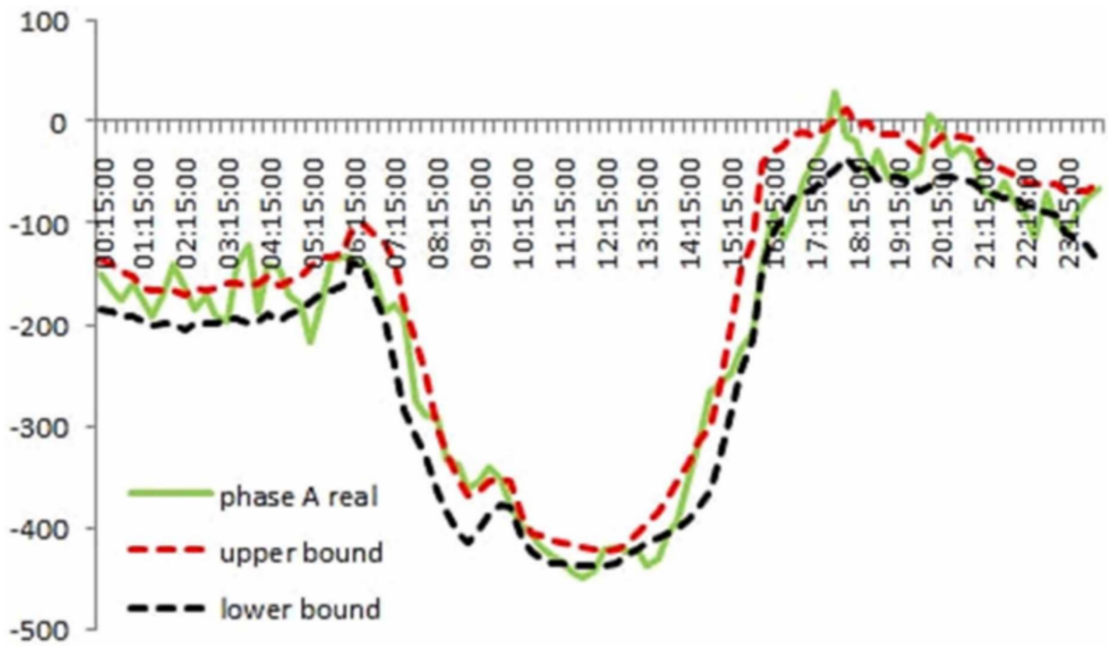

Figure 4.9: PIs of the 1st of March on phase A (1 hour std) (x:time,y:amperes)

the second one at 19.30, were missed. The reason for that is that the produced PI is relatively narrow. The standard deviation was collected every one hour and thus the variance of the ampacity level is low for such a small period of time. In Figure 4.9 the predictions regarding the 1st of March are presented.

The above observation is confirmed by the results contained in Tables 4.11, 4.12, 4.13 and 4.14. As it is depicted in Table 4.11, the performance of the method in this case is low. The average performance for these ten days on phase A is 57.82%, for phase B is 54.89% and in phase C is 50.83%; results that indicate that the use of the standard deviation of the target value every one hour is not adequate enough for a reliable predictive method. The constructed PIs, as it appears in Tables 4.12, 4.13 and 4.14, are very narrow. The average width for the PI of phase A is 55.19 amperes, 43.18 amperes for phase B and 32.27 amperes for phase C. At the same time, the average standard deviation of PI's width is 24.50 amperes for phase A, 22.85 for the phase B and 14.75 for the phase C.

Test Case 2

In the second case scenario the PIs were constructed based on the standard deviation of the target value every two hours. As it can be seen in Figure 4.10 that depicts the output of the method for the 11th of December, the width of the PIs is now broader than the one in the

Table 4.11: Performance of PI method for all phases (Test Case 1)

Day	Phase A	Phase B	Phase C
1/3/2001	61.45%	51.04%	67.07%
2/1/2001	55.20%	50%	55.20%
6/8/2001	46.87%	50%	55.20%
8/1/2001	71.87%	57.29%	58%
10/12/2001	46.87%	43.75%	47.91%
11/12/2001	53.33%	61.45%	54.16%
12/12/2001	62.50%	59.37%	54.16%
13/12/2001	59.37%	59.37%	56.25%
15/2/2001	61.45%	56.25%	62.50%
18/11/2001	59.37%	60.41%	60.41%
Average	**57.82%**	**54.89%**	**50.83%**

Table 4.12: Width average and std on Phase A (Test Case 1)

Day	Width average	std
1/3/2001	43.65	21.52
2/1/2001	43.42	51.48
6/8/2001	38.75	39.75
8/1/2001	45.70	9.51
10/12/2001	42.38	18.66
11/12/2001	50.38	28.22
12/12/2001	43.13	21.41
13/12/2001	48.34	17.65
15/2/2001	42.01	18.38
18/11/2001	43.69	18.43
Average	**55.19**	**24.50**

Table 4.13: Width average and std on Phase B (Test Case 1)

Day	Width average	std
1/3/2001	41.10	18.75
2/1/2001	49.80	45.45
6/8/2001	35.70	36.68
8/1/2001	43.82	8.07
10/12/2001	40.14	17.96
11/12/2001	49.82	28.26
12/12/2001	42.19	20.90
13/12/2001	48	17
15/2/2001	40.55	17.46
18/11/2001	40.69	18.01
Average	**43.18**	**22.85**

Table 4.14: Width average and std on Phase C (Test Case 1)

Day	Width average	std
1/3/2001	32.22	12.74
2/1/2001	37.17	27.26
6/8/2001	26.91	26.97
8/1/2001	33.21	6.28
10/12/2001	31.63	11.43
11/12/2001	34.96	16.20
12/12/2001	30.01	13.24
13/12/2001	33.76	10.27
15/2/2001	31.13	11.26
18/11/2001	32.73	11.86
Average	**32.27**	**14.75**

first case but not too broad to be prohibitive.

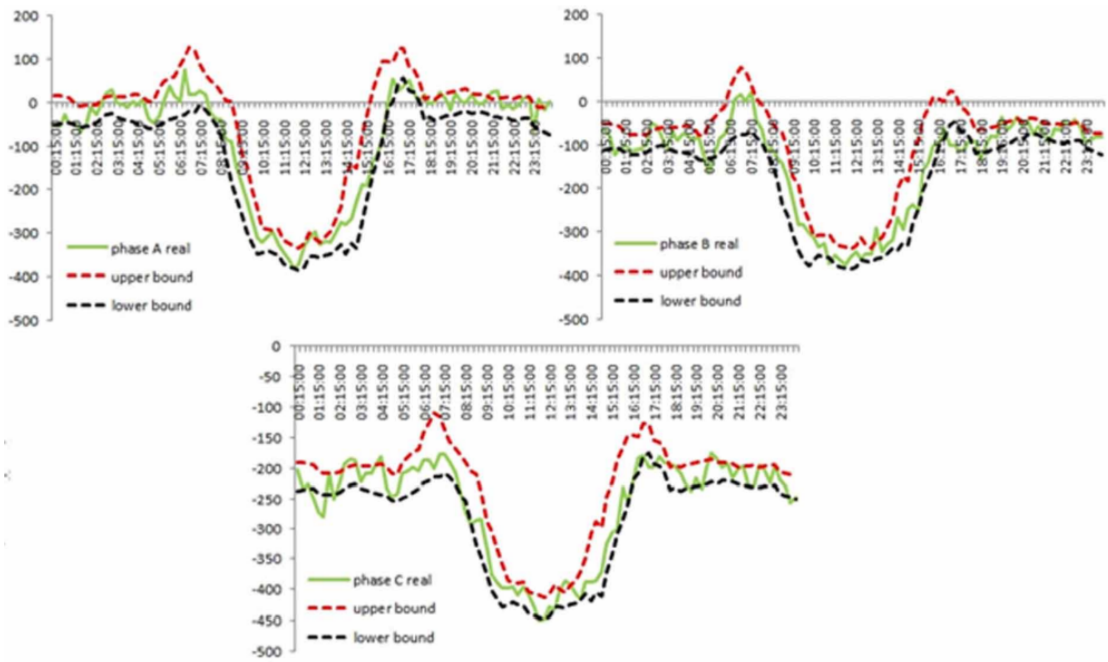

Figure 4.10: Constructed PIs of the 11th of December (2 hours std) (x:time,y:amperes)

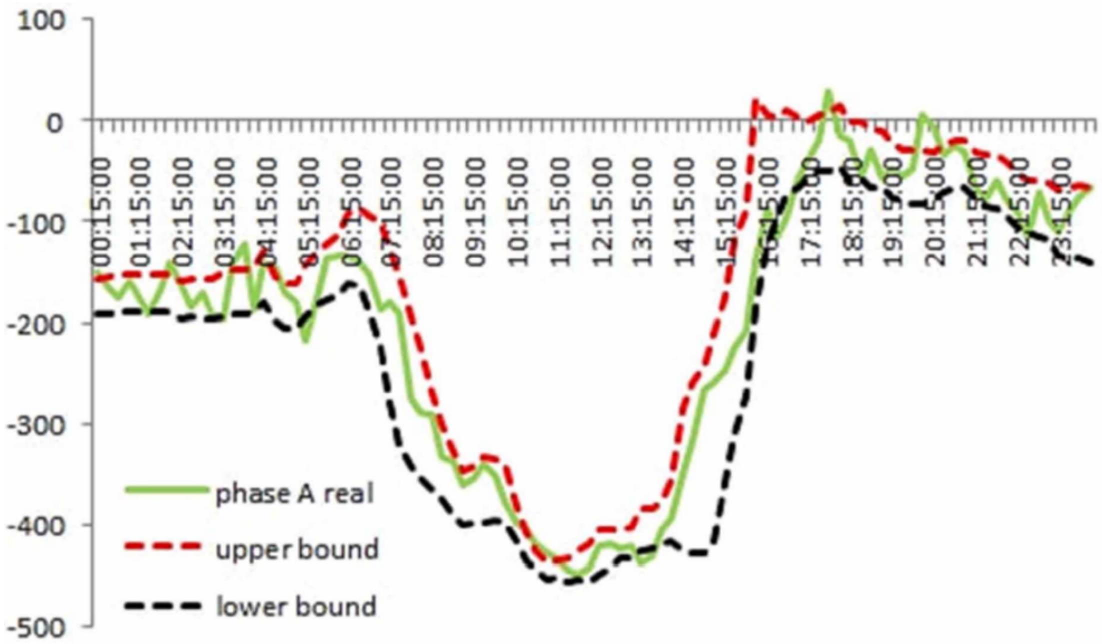

Figure 4.11: PIs of the 1st of March on phase A (2 hours std) (x:time,y:amperes)

Moreover, there are still cases where not all the congested points are included inside the constructed PI. As it can be seen in Figure 4.11 (1st of March), the results are improved, compared to the first case, since this approach indicates the possibility of congestion at time 17.30. Nevertheless, both the peak of overloading, as well as the second congested point

Table 4.15: Performance of PI method for all phases (Test Case 2)

Day	Phase A	Phase B	Phase C
1/3/2001	79.16%	69.79%	77.08%
2/1/2001	77.08%	68%	72.91%
6/8/2001	68.75%	67%	71.87%
8/1/2001	69.71%	58.33%	63%
10/12/2001	72.91%	68.75%	79.66%
11/12/2001	81.25%	78.12%	70.83%
12/12/2001	79.16%	77.08%	76.04%
13/12/2001	83.33%	73.95%	78.12%
15/2/2001	78.12%	70.13%	80.20%
18/11/2001	79.16%	71.87%	68.75%
Average	**76.86%**	**69.80%**	**73.64%**

(19:30), is not included in the PI. As in the previous case, Tables 4.15 and 4.16 contain the performance of the method, as well as the average PI's width and standard deviation for the three phases of each day.

As it is observed in Tables 4.15, 4.16, 4.17 and 4.18 the performance has been improved very much; 76.86% against 57.82% for phase A, 69.80% against 54.89% in phase B and 73.64% against 50.83% in phase C. Moreover, as mentioned earlier, the constructed PI is not prohibitively broad since it is 69.69 amperes for phase A, 64.90 amperes for phase B and 47.81 amperes for phase C.

Test Case 3

In the third case, the PIs were constructed based on the standard deviation of the target value every three hours. The 11^{th} of December is depicted in Figure 4.12, as in the previous cases. The width of the constructed PI is of course broader than the one constructed in the previous two cases, but still allows us to make useful decisions regarding the overloading or not of a phase.

In Figure 4.13, the constructed PIs for the 3 hour std approach for the 1st of March are depicted once again for comparison purposes. It can be observed an improvement regarding

Table 4.16: Width average and std on Phase A (Test Case 2)

Day	Width average	std
1/3/2001	67.74	44.75
2/1/2001	77.47	48.57
6/8/2001	68.02	76.69
8/1/2001	57.48	23.18
10/12/2001	65.24	29.07
11/12/2001	78.64	45.80
12/12/2001	74.36	41.80
13/12/2001	72.53	32.01
15/2/2001	66.53	42.86
18/11/2001	69.47	30.71
Average	**69.69**	**41.54**

Table 4.17: Width average and std on Phase B (Test Case 2)

Day	Width average	std
1/3/2001	60.53	34.41
2/1/2001	70.78	44.35
6/8/2001	61.09	66.29
8/1/2001	50.63	19.95
10/12/2001	60.54	26.01
11/12/2001	73.02	39.90
12/12/2001	69.10	34.73
13/12/2001	68.64	27.74
15/2/2001	62.13	34.62
18/11/2001	67.24	31.94
Average	**64.90**	**45.99**

Table 4.18: Width average and std on Phase C (Test Case 2)

Day	Width average	std
1/3/2001	45.57	25.24
2/1/2001	52.42	30.66
6/8/2001	43.51	46.37
8/1/2001	40.63	14.29
10/12/2001	45.71	18.34
11/12/2001	53.53	27.63
12/12/2001	51.24	25.09
13/12/2001	49.71	19.22
15/2/2001	46.73	25.81
18/11/2001	49.11	21.83
Average	**47.81**	**25.44**

Table 4.19: Performance of PI method for all phases (Test Case 3)

Day	Phase A	Phase B	Phase C
1/3/2001	84.37%	80.20%	86.45%
2/1/2001	72.91%	73%	71.87%
6/8/2001	72.91%	73%	71.8%
8/1/2001	77.08%	77.08%	67%
10/12/2001	87.5%	78.12%	86.45%
11/12/2001	85.41%	86.45%	80.20%
12/12/2001	81.25%	86.45%	84.37%
13/12/2001	84.37%	82.29%	79.16%
15/2/2001	83.33%	85.41%	85.41%
18/11/2001	84.37%	85.41%	81.25%
Average	**81.35%**	**80.74%**	**79.39%**

Table 4.20: Width average and std on Phase A (Test Case 3)

Day	Width average	std
1/3/2001	85.88	51.92
2/1/2001	91.71	51.49
6/8/2001	85.86	86.86
8/1/2001	66.59	25.36
10/12/2001	89.18	35.14
11/12/2001	92.71	40.99
12/12/2001	84.59	38.51
13/12/2001	81.13	31.15
15/2/2001	84.95	50.68
18/11/2001	95.47	41.60
Average	**85.80**	**45.37**

Table 4.21: Width average and std on Phase B (Test Case 3)

Day	Width average	std
1/3/2001	77.42	42
2/1/2001	79.26	47.10
6/8/2001	82.05	79.05
8/1/2001	60.49	18.31
10/12/2001	78.41	31.50
11/12/2001	80.86	35.23
12/12/2001	73.15	29.94
13/12/2001	72.66	25.64
15/2/2001	78.54	45.77
18/11/2001	88.82	44.12
Average	**77.16**	**39.86**

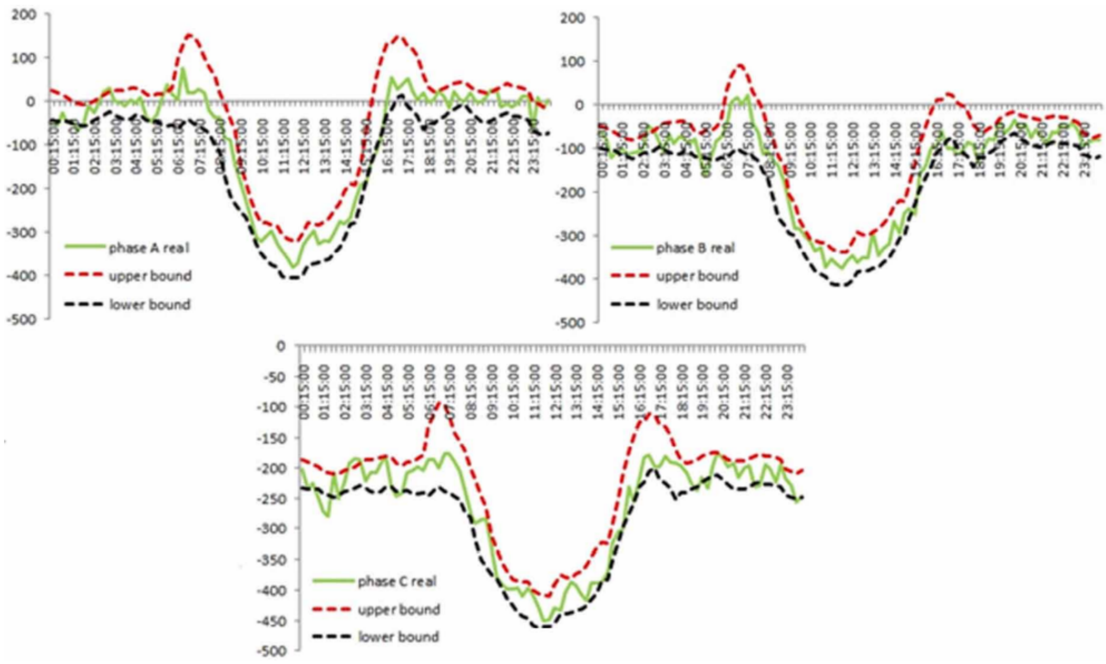

Figure 4.12: . Constructed PIs of the 11th of December (3 hours std) (x:time,y:amperes)

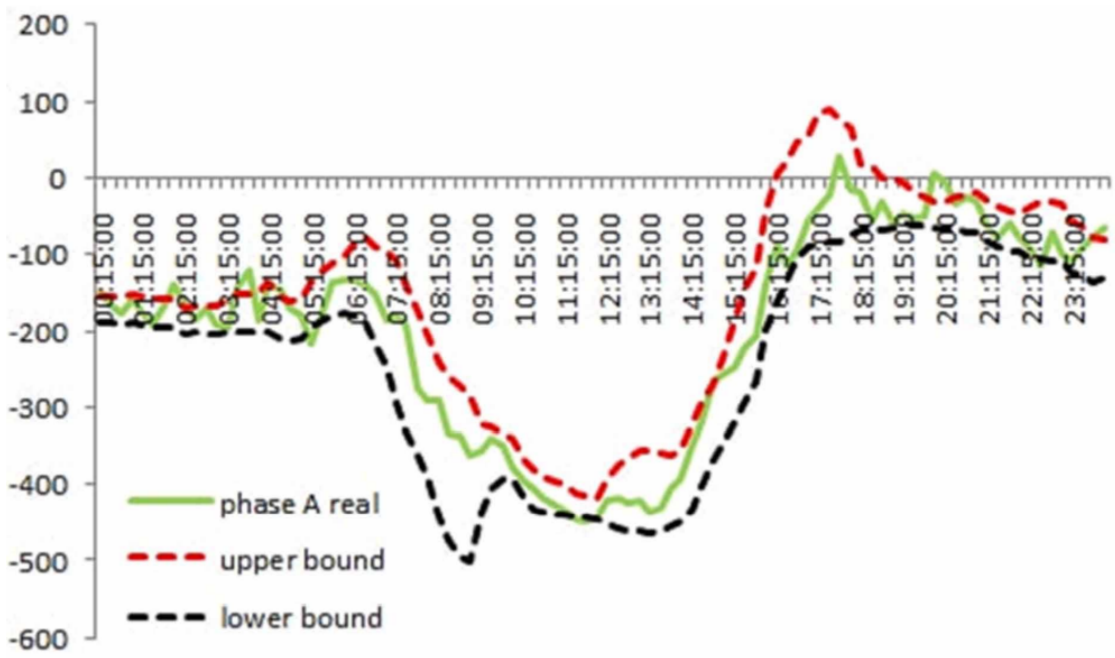

Figure 4.13: PIs of the 1st of March on phase A (3 hours std) (x:time,y:amperes)

the previous cases since the first overloading, observed at 17:30, is included in the PI. Nevertheless, the second overloading, observed in 19.30, is not. Yet, this is one of the very few cases that a congested point was not included in the predictive interval.

Moreover, as it is appeared by Tables 4.19, 4.20, 4.21 and 4.22 there is a big improvement regarding the efficiency compared to the 2 hour std approach; 81.35% against 76.86% for the

Table 4.22: Width average and std on Phase C (Test Case 3)

Day	Width average	std
1/3/2001	57.53	31.27
2/1/2001	60.71	33.23
6/8/2001	54.17	53.28
8/1/2001	46.45	16
10/12/2001	59.66	22.25
11/12/2001	62.09	25.76
12/12/2001	56.36	23.67
13/12/2001	55.19	20.15
15/2/2001	57.13	39.19
18/11/2001	66.43	28.81
Average	**57.57**	**28.56**

phase A, 80.74% against 69.80% in the phase B and 79.39% against 73.64% in the phase C. On the other hand the width of the PI has not expanded too much; 85.80 amperes against 69.69 amperes for the phase A, 77.16 against 64.90 for the phase B and finally 57.57 amperes against 47.81 amperes for the phase C.

Test Case 4

In the fourth case, the standard deviation of the target value every four hours were used in order to create the PIs. Of course we expect improvement on what concerns the efficiency of the method. What has to be proved is whether the width of the constructed PIs will be capable to provide useful information or not.

In Figure 4.14, the reader can see the constructed PIs for the 11th of December. At a first glance, the reader may perceive that PI's width is large enough although this time the constructed PIs are capable to capture congested points in the majority of the cases. To illustrate that, the output of the method regarding the phase A of March 1st is once again presented in Figure 4.15. This is the first time where all the congested points are included in the PI.

Tables 4.23, 4.24, 4.25 and 4.26 confirm our allegations. The performance of the approach that uses the standard deviation of the target value every four hours is the better compared to

Table 4.23: Performance of PI method for all phases (Test Case 4)

Day	Phase A	Phase B	Phase C
1/3/2001	81.25%	81.25%	86.45%
2/1/2001	82.29%	81%	82.29%
6/8/2001	90.62%	91.66%	90.62%
8/1/2001	76.04%	79.12%	70%
10/12/2001	84.37%	81.25%	85.41%
11/12/2001	87.5%	86.45%	83.33%
12/12/2001	86.45%	82.29%	84.37%
13/12/2001	94.79%	84.37%	86.45%
15/2/2001	90.65%	86.45%	88.54%
18/11/2001	88.45%	91.66%	84.37%
Average	**86.24%**	**84.55%**	**84.18%**

Table 4.24: Width average and std on Phase A (Test Case 4)

Day	Width average	std
1/3/2001	100.72	61.92
2/1/2001	106.67	42.74
6/8/2001	130.85	105.07
8/1/2001	75.01	24.10
10/12/2001	99.79	38.28
11/12/2001	113.36	48.32
12/12/2001	112.53	53.43
13/12/2001	97.29	35.21
15/2/2001	92.80	30.72
18/11/2001	99.62	29.72
Average	**102.86**	**46.95**

Table 4.25: Width average and std on Phase B (Test Case 4)

Day	Width average	std
1/3/2001	89.77	49.40
2/1/2001	94.31	37.11
6/8/2001	119.35	92.16
8/1/2001	71.91	20.31
10/12/2001	89.59	32.47
11/12/2001	101.16	40.91
12/12/2001	101.45	45.93
13/12/2001	88.34	29.09
15/2/2001	86.07	25.45
18/11/2001	93.39	29.24
Average	**93.53**	**40.20**

Table 4.26: Width average and std on Phase C (Test Case 4)

Day	Width average	std
1/3/2001	66.10	35.55
2/1/2001	68.79	25.25
6/8/2001	81.79	64.53
8/1/2001	51.18	14.29
10/12/2001	65.01	22.26
11/12/2001	73.51	27.00
12/12/2001	73.59	31.70
13/12/2001	63.29	19.08
15/2/2001	63.06	17.14
18/11/2001	68.33	19.34
Average	**67.46**	**27.67**

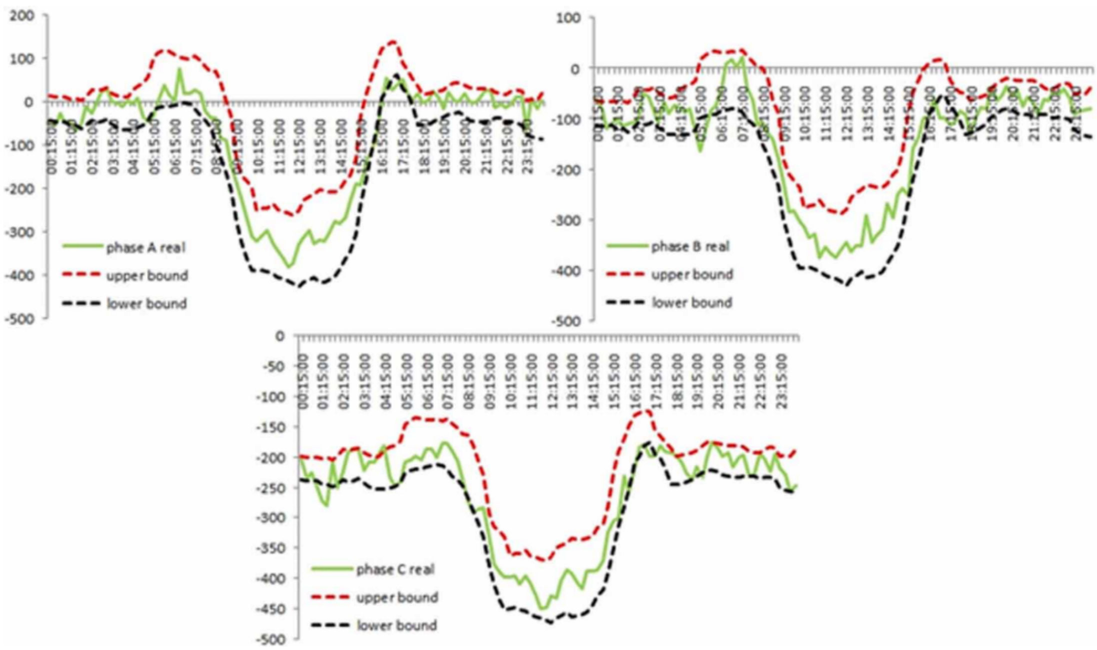

Figure 4.14: Constructed PIs of the 11th of December (4 hours std) (x:time,y:amperes)

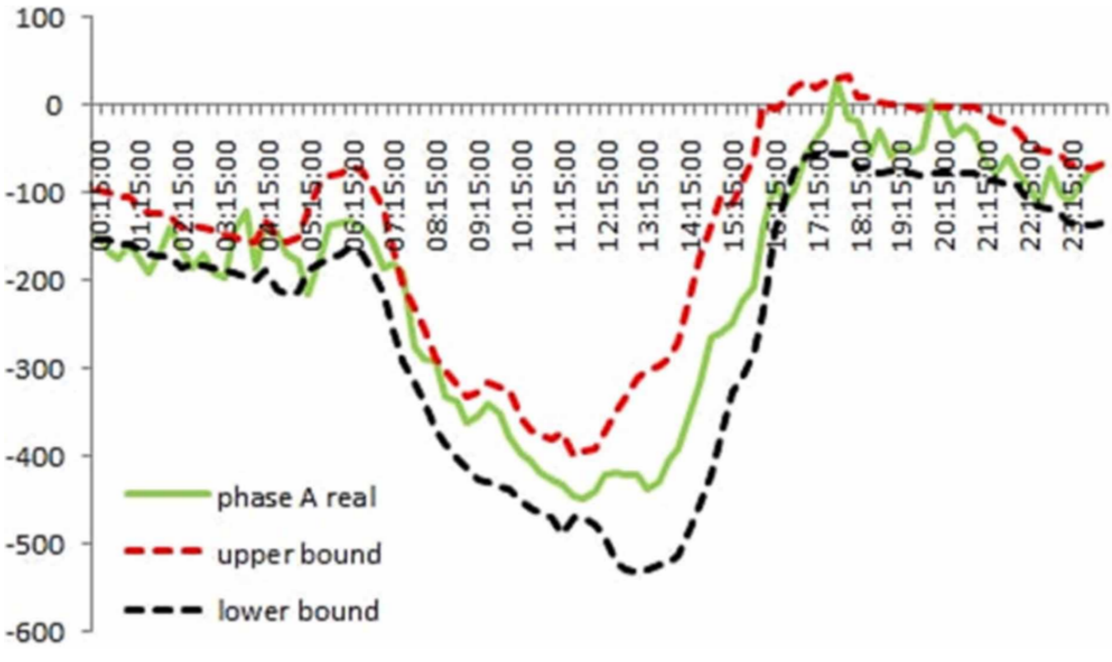

Figure 4.15: PIs of the 1st of March on phase A (4 hours std) (x:time,y:amperes)

the three previous regarding the performance; above 84% for all the three phases. PIs' width on the other hand started to be broad enough so as to provide useful information. The average width on phase A, for instance, is 102.86 amperes, 93.53 for phase B and 67.46 amperes for phase C.

Regarding the rest four cases where the standard deviation of the target value used for the construction of PIs was ranged from five to eight hours, both the performance as well as the width of the constructed PI continued to increase. nevertheless, based on the fact that a very broad PI has nothing to offer regarding the extraction of useful information, the authors conclude that the 3 hour std approach is preferable since it combines a fair enough level of performance with the narrowest possible PI.

4.3.4 Conclusions

This study was conducted in order to deploy and evaluate a method for constructing predictive intervals for enhanced ampacity overloading predictions. The advent of smart grids and the evolvement of consumers into entities that actually participate in the energy market by forming energy usage prices and by adjusting their energy needs so as to maximize their welfare has to totally change the way we look and act to sustain system's stability. The load of the system changes constantly from time to time and thus a plain point prediction method may be proved incapable to follow the system's dynamics.

To that end, a method based on artificial neural networks was developed. The NN was trained by using the Levenberg-Marquardt training function for selected data aiming primarily to efficiently capture ampacity level peaks. The NN trained so as to produce a point prediction regarding the ampacity level of each phase, as well as a value which by added and subtracted by the point prediction defines a prediction interval. This particular value was defined by using the standard deviation of the target value of ampacity level for known input data.

The method was assessed on eight different test cases. In every test case we used the standard deviation of the target value for different time period. For instance, in the first case we used the standard deviation of the target value every one hour, in the second case every two hours and so on. We concluded that the three hour approach is the most appropriate since it combines a fair enough level of performance with the narrowest possible prediction interval. In all the three phases, the average of performance was above 79%, while the average width of the constructed prediction interval was 85.80 amperes in phase A, 77.16 on phase B and 57.57 amperes on phase C.

4.4 Distribution Congestion Prediction using Artificial Neural Networks for Big Data

In this study we focus our attention on the distribution system aiming at developing a congestion predictor that monitors part of the electricity distribution system and predicts whether any line of that specific part will be overloaded or not in a prespecified ahead of time horizon. Congestion prediction could be beneficial i) for system designers to optimally place distributed generation sources, ii) for Distribution System Operator (DSO) to take immediate actions before an emergency actually occurs or to optimize the energy imports/exports of each region of the system, and iii) for customers and producers to properly plan their consumption/generation schedules. Our intention is to use artificial neural networks (ANN) as our congestion predictor, since ANN exhibit superior performance in nonlinear problems [46].

4.4.1 Methodology

Proposed Architecture and Training

As mentioned earlier, we propose the use of ANN for the development of a predictive mechanism with the goal of monitoring part of the grid and promptly warn System Operators for periods of time during which congestion is high likely to occur. In this direction, we used a feedforward backpropagation neural network created and trained in Matlab software [50].

The neural network, whose architecture is depicted in Figure 4.16, was comprised of one input layer with five inputs, one hidden composed of fifty neurons and one output. The output should be ”1” if at least one line of the feeder is overloaded and ”0” otherwise.

It must be mentioned here that a lot of different architectures regarding the number of hidden layers and neurons in them have been tested in order to select the one that fits to our purposes. The selection was based on accuracy and training time in an attempt to achieve the best accuracy with as short as possible training time.

The ANN was trained by using the back-propagation method. The training algorithm used in this study, is the Levenberg-Marquardt. Moreover, in order to confront overfitting issues, the Levenberg-Marquardt method was used combined with Bayesian regularization.

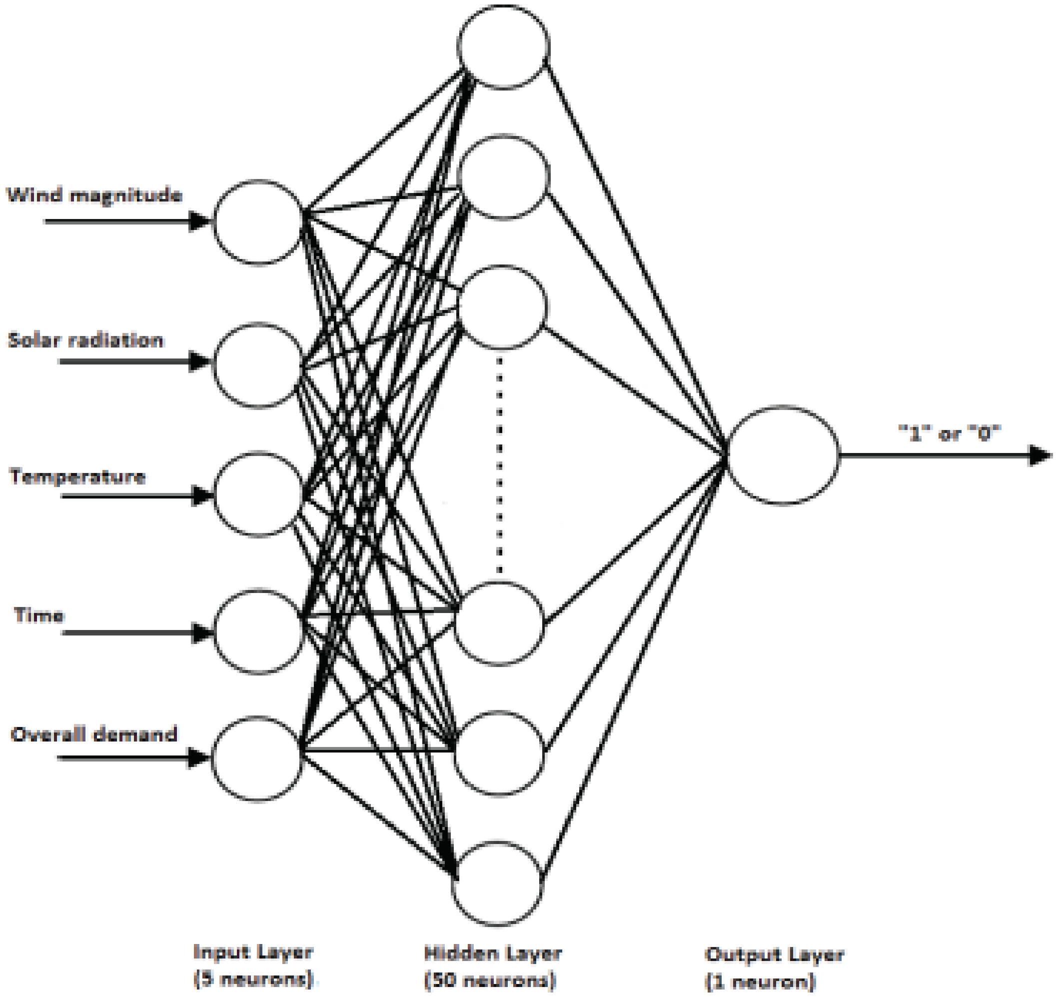

Figure 4.16: Architecture of the neural network where the grid parameters are given as inputs

4.4.2 Configuration

All the experiments have been contacted on the IEEE 13-bus test feeder which is depicted in Figure 3.3 [32]. Our feeder is equipped with 1306 residencies. Each residency can accommodate 4 individuals and with respect to its energy needs, includes lights, HVAC system and a water heater. The feeder is also equipped with 4 producers that provide 2400, 500, 400 and 100 kW in the price of 0.3, 0.11, 0.25 and 0.15 \$/kWh respectively. The electricity market follows the principles defined by W. W. Hogan in [1], while appliances and producers bid into the market to exchange power every fifteen minutes.

The data was obtained via the GridLAB-D [2] simulation platform, and data points were collected every fifteen minutes for a period of a year. Obtained data consist of the following parameters:

- Wind magnitude
- Solar radiation
- Temperature
- Time
- Overall demand

Based on this data we trained our network for the whole year of 2000 and then we used the next year (2001) for testing purposes.

It must be also mentioned that except for the non-responsive loads, i.e., the lights in our experiments, the thermostatically controlled devices i.e. the HVAC system and the water heater, are participating into demand response programs. GridLAB-D, as explained many times before, offers this capability through an object called ”transactive-controller” which is responsible of bidding on behalf of a device for a specified amount of power in a certain price. On one hand, the end-user has just to define his desired temperature set point and his willingness to pay depending on the price levels. The controller on the other hand, undertakes to operate on his behalf. It is responsible of adjusting the temperature set points of these devices by responding to the current price and to determine into which price it will bid as a function of the house's internal temperature; seeking to indirectly control the loading level of the grid. For an HVAC system for instance, when the price is too high, implying there is excessive demand, the set points are moved higher expecting reduction of demand in order to preserve system stability. More information about transactive controllers and the implementation of demand response programs in GridLAB-D can be found in [49].

Finally, two indices were used in order to further evaluate the accuracy of our method. The first index is defined as the ratio between the true positive predictions to the total of positive predictions and is called Positive Predictive Value (PPV). The mathematical formulation of the index is as follows:

$$PPV = \frac{number of true positives}{number of positives} \tag{4.13}$$

The meaning of true positive in equation 4.13 is that congestion was truly identified by the predictor, while the number of positives stands for the total number of points where congestion has been anticipated. Similarly, the second index is defined as the ratio between true negative predictions to the total of negative predictions. It is called Negative Predictive values (NPV) and is given by

$$NPV = \frac{numberoftruenegatives}{numberofnegatives} \tag{4.14}$$

In this case, the meaning of true negatives is that the predictor truly did not identify congestion, while the number of negatives stands for the total number of points for which the predictor did not anticipate congestion occurrence.

4.4.3 Results

In this section we present the results of our method. We reduced the limits of the lines in order to intentionally induce congestion on the distribution system. Our predictor's task is to monitor a part of the distribution system and to predict the time that at least one line of this part will be overloaded. This should be very helpful for System Operator to identify and prevent events that may be crucial for system stability ahead of time.

In Table 4.27 the real points in time at which there is congestion, as obtained by the GridLAB-D simulation platform, against the predicted points in time are presented. The presented days were arbitrarily selected among the few congested; we didn't reduce line limits too much in order to produce a realistic scenario.

As it turned out, out method does not capture the exact time of congestion occurrence. Instead it effectively captures a narrow time window inside which congestion actually appears. Congestion appears early morning when people wake up and prepair themselves to go to their jobs and late afternoon when they are returning back to their homes or at noon when the use of HVAC systems reaches the maximum level; during summer months especially. For instance, in the 2th of December 2001 congestion appeared for one and a half hour during morning from 05:30 to 07:00 and during the afternoon from 18:30 to 21:15. Our predictor on the other hand predicted congestion from 05:45 to 07:30 during morning and from 17:00 to 21:30 during afternoon and night. Based on this predicted time window the system operator can effectively take preventive actions or to be in high alert for an upcoming crisis.

These findings are also confirmed by Table 4.28 and Figures 4.17, 4.18, 4.19 and 4.20. The figures show the loading trend of the feeder for four different days; namely the 27th of February 2001, the 18th of November 2001, the 12th of April 2001 and the 6th of December 2001. The x-axis represents the time of the day, while the y-axis represents the load of the most

Table 4.27: Predicted Vs Congested points

Day	Congestion presence at	Congestion prediction at
1/4/2001	16:00 to 18:00	16:30 to 17:30
3/2/2001	05:15 to 05:30 06:15 to 07:15 20:15 to 21:15	04:15 to 06:45 20:15 21:15 to 21:45
8/6/2001	15:45 to 17:15	14:15 to 14:30 15:45 to 17:15
2/27/2001	15:45 to 20:45	16:15 to 21:00
10/6/2001	14:00 to 17:30	13:30 to 14:45 16:00 to 17:15
11/18/2001	05:00 to 07:45 18:45 to 19:30 21:30 to 22:00	04:30 to 07:45 18:30 to 20:15 21:30 to 22:00
12/2/2001	05:30 to 07:00 18:30 to 21:15	05:45 to 07:30 17:00 to 21:30

loaded line of this part of the distribution system in kW, compared with the line's limit. It must be mentioned again that it is not necessarily the same line; just the most loaded line. A value of -30 kW for example means that the most loaded line on the system operates by 30 kW below its capacity limit. A positive value on the other hand indicates line overloading by the amount indicated by the y-axis. The first two days represent days with congestion, while the third and fourth ones represent days without congestion. It appears that our mechanism indeed closely follows the real loading of the system for both the congested and the uncongested cases.

During spring especially, as you can see in Figure 4.19, the temperatures aren't either too high or too low, but in an intermediate level. Because of that, the need for cooling or heating and consequently the need to use the HVAC system is reduced significantly. This is why we cannot observe line overloading in any line of the feeder in this particular case.

The same can be also observed in Figures 4.21 and 4.22 where the error, related to the misclassification rate on the congested days, and the accuracy are presented in bar diagrams. These measurements correspond to the findings presented in Figures 4.17 and 4.18. The error for both cases expressed in percentages was 5.15% while the accuracy was 94.85%.

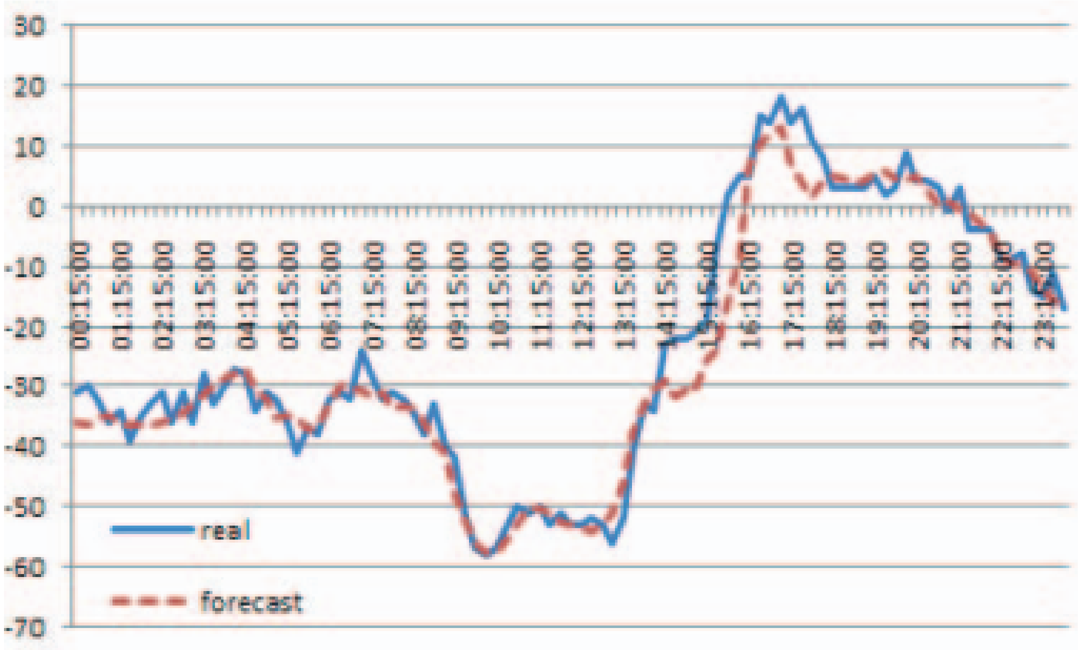

Figure 4.17: 2/27/2001 Trend of grid loading (x:time,y:amperes)

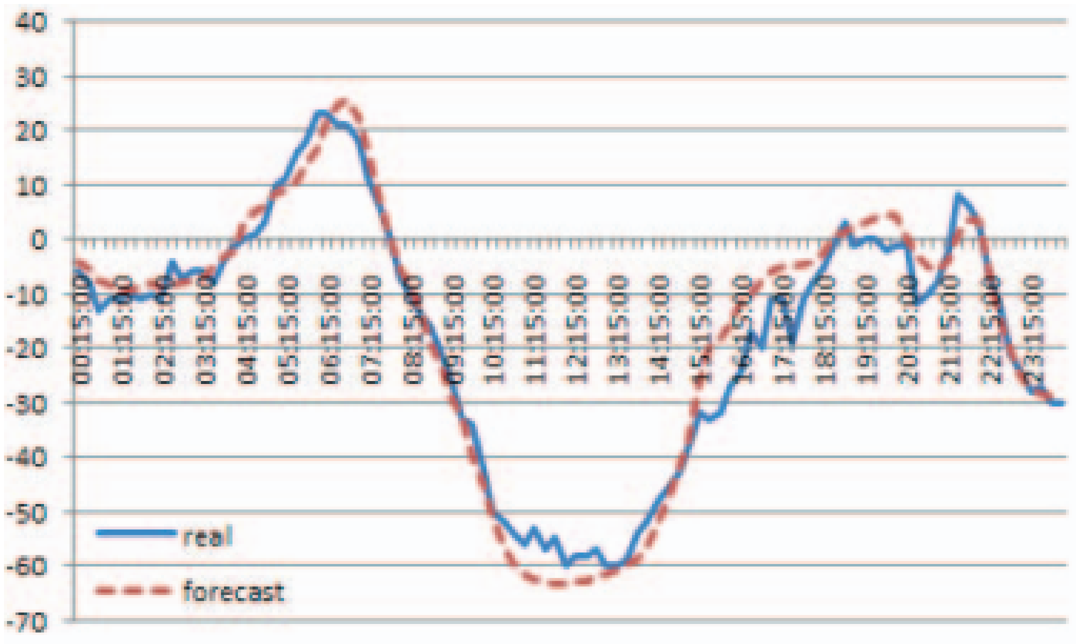

Figure 4.18: 11/18/2001 Trend of grid loading (x:time,y:amperes)

Moreover, Table 4.28 contains the aggregated misclassification error and accuracy values for the seven selected days. It appears that our method achieves high rates of accuracy and low rates of error. The highest accuracy and consequently the lowest error were observed on the 4^{th} of January 2001 with 96.91% and 3.09% respectively. The worst rates on the other hand, appeared on the 2^{nd} of March 2001 with rates 83.50% and 16.50% respectively. Though poor,

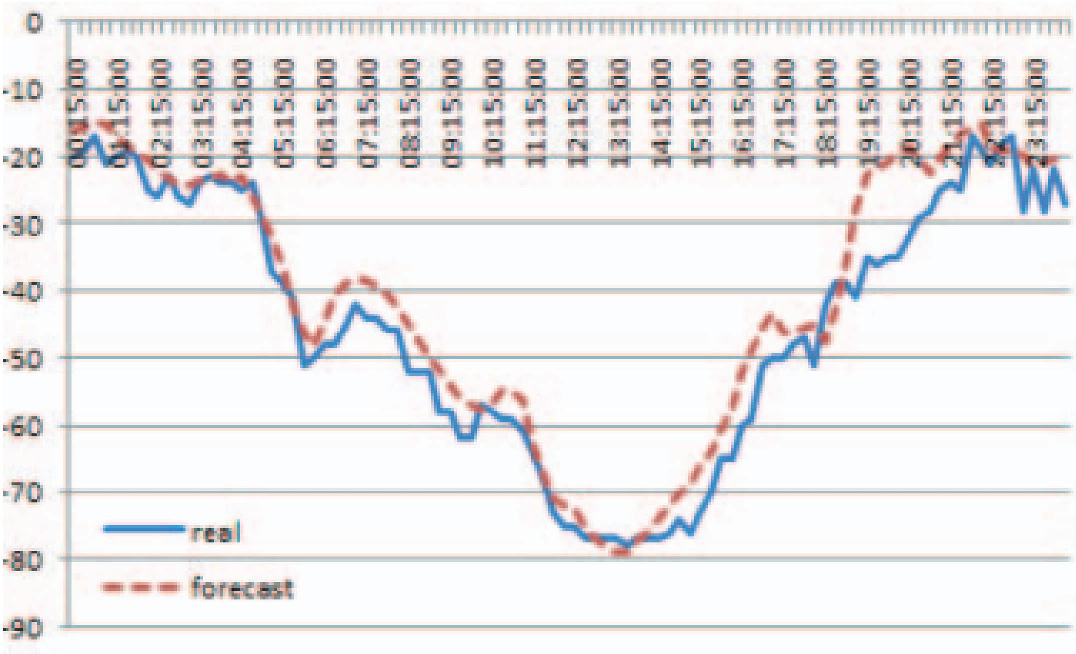

Figure 4.19: 4/12/2001 Trend of grid loading (x:time,y:amperes)

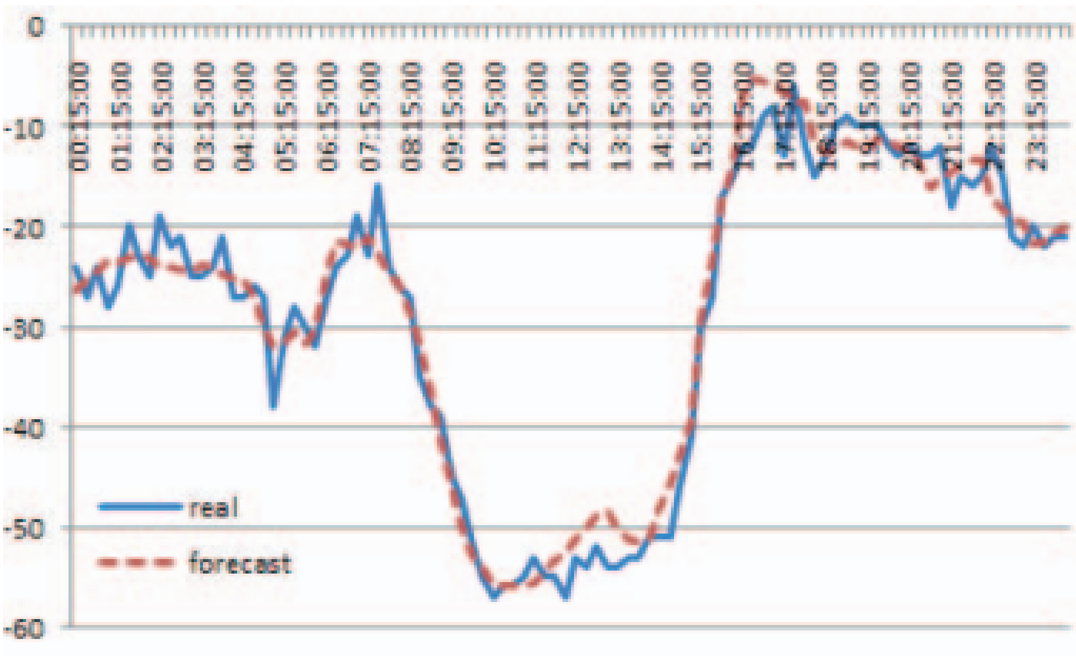

Figure 4.20: 12/6/2001 Trend of grid loading (x:time,y:amperes)

but acceptable rates, can be explained by the noise in the signal that represents the loading of the feeder. The noise on the other hand is due to the fact that we are not monitoring a single line but a whole feeder and record each time the most loaded line; consequently the loading signal is not smooth. Moreover, as mentioned in Section 3, the household appliances participate into demand-response programs. As a consequence, their operation status is constantly changing

Table 4.28: Misclassification Error/Accuracy

Day	Error	Accuracy
1/4/2001	3.09%	96.91%
3/2/2001	16.50%	83.50%
8/6/2001	3.09%	96.91%
2/27/2001	5.15%	94.85%
10/6/2001	7.20%	92.80%
11/18/2001	5.15%	94.85%
12/2/2001	10.30%	89.70%
Average	7.21%	92.79%

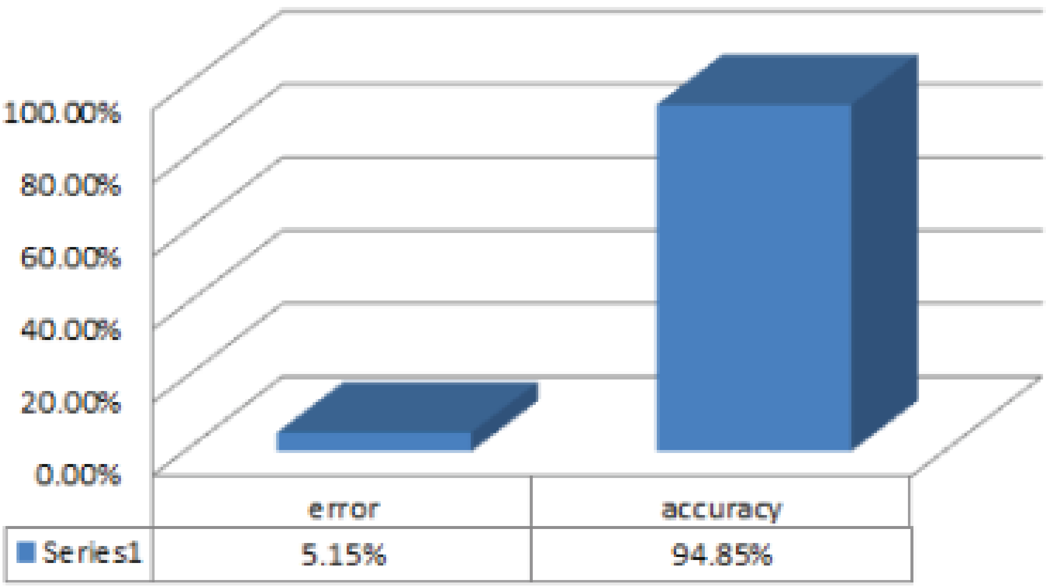

Figure 4.21: Misclassification Error/Accuracy for 2/27/2001

during time (every fifteen minutes) and because of that, the noise in the loading signal of the feeder is further maximized. Finally, in the last row of the table you can see the average accuracy and error values of the selected days with rates 92.79% and 7.21% respectively.

Table 4.29 on the other hand, contains the PPV and NPV values of the days presented in Table 4.28. The PPV and NPV indices represent how trustful a prediction is, since they are related to the number of truly predicted situations.

In particular the PPV value for the prediction regarding the 6th day of August 2001 was

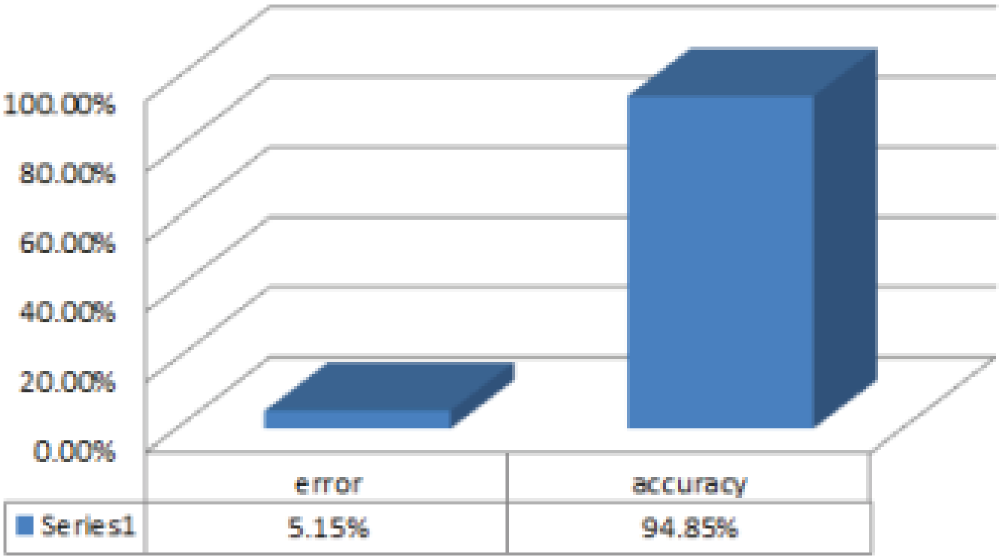

Figure 4.22: Misclassification Error/Accuracy for 11/18/2001

Table 4.29: PV and NPV values

Day	PPV	NPV
1/4/2001	100%	97%
3/2/2001	55%	89%
8/6/2001	67%	100%
2/27/2001	95%	94%
10/6/2001	82%	94%
11/18/2001	78%	100%
12/2/2001	67%	98%

67%. This means that in all of the tested cases, where our method predicted congestion on a line of the feeder for this day, 67% of them were indeed congested. The NPV index for the same day was 96.91%. Correspondingly, this means that in the cases where our mechanism predicted that no line of the feeder was congested, in the 96.91% was indeed not congested. This day and the 2nd of March 2001 were the days where our ANN achieved the lowest PPV rates. The corresponding values for the 2nd of March were 55% and 89%. The highest rates on the other hand, achieved for the 4th of January 2001 with PPV and NPV values 100% and 97% respectively.

As it turned out, in most of the cases, our ANN achieves high accuracy. The lower levels achieved in the 2nd of March 2001 and in the 6th of August 2001 can be explained by the noise in the loading signal, as explained earlier (see Figures 4.17 to 4.20) and by the fact that in our attempt to create a realistic scenario we didn't induce high congestion, so the congested days in the training data were only few. The few days exhibiting congestion in the training data resulted in an additional difficulty related to the efficient training of the ANN and consequently to its ability to be able to generalize more accurately when it is fed with new data. The always high rate of the NPV value confirms this allegation.

4.4.4 Conclusions

The aim of this study was to develop and evaluate a congestion predictor eligible to act in a big data environment. The goal was to allow a predictor to be able to monitor a part of the distribution system and to provide useful predictions about the congestion status of this part. This may be proven useful for every entity participating in a new generation energy market; for customers and generators in order to organize a more appropriate portfolio, and for system planers in order to design a most effective power grid. However, such a tools is crucial for the System Operator to monitor and take proper action before an irreversible situation prevails.

In this direction, we developed a predictor by utilizing a Feedforward Neural Network. Its responsibility was to monitor a portion of the distribution system and to provide forecasts about the time where congestion will occur at least on one of the lines of this portion. It trained by using the Levenberg-Marquardt algorithm combined with Bayesian regularization and tested on the IEEE 13-bus distribution feeder. All the training and evaluation data were obtained by the GridLAB-D simulation platform.

It was proved that our method was able to define narrow time windows into which congestion actually occurred. This can be seen clearly by Figures 4.17, 4.18, 4.19 and 4.20 that depict that the predictions always followed the real loading signal. Indeed, there were cases for which our method predicted the exact time of congestion occurrence or cases where the time of congestion presence predicted with high accuracy. Based on these forecasts, events that otherwise would have enormous social and economic impact can be prevented or at least reduce their effects.

4.5 Three–Phase Line Overloading Predictive Monitoring Utilizing Artificial Neural Networks

In the current study, we direct our efforts to the distribution level. Our ultimate goal is to develop and evaluate a monitoring method that in real time monitors each phase of a distribution line and provides useful information regarding the overloading or not of a phase. Moreover, except for the line overloading status, only the information about the loading level of each phase can be proved beneficial. The SO can be always aware at real time about the status of each line and coordinate the operation of the grid in a more effective way. Actions that aim to retain system stability can be taken in advance avoiding that way extreme measures taken as last resort with high financial impact like load shedding. More precisely, our method monitors the ampacity level of each phase and its proximity to the emergency rating of the phase. It must be mentioned here that loading level higher than the emergency rating for longer than a two hour period will result in a line failure with devastating results. At the same time, we also aim at examining the effectiveness of the AI methods, and more precisely of ANNs on line overloading predictions. This effectiveness can not be guaranteed as easy as it first appears in a market framework where the devices participate into demand response programs because of the noise in the loading signal that appears in this type of markets [1].

To that end, a neural network which is repeatedly trained every hour with recently observed events to improve its efficiency, was used. The GridLAB-D [2] platform was utilized in order to simulate the distribution grid, the market framework into which it operates, and to acquire both the simulation and the evaluation data for testing the ANN. Given that our simulations generated a big data environment with multiple heterogeneous data streams, the MySQL software was selected as the most appropriate for data storage and retrieval.

4.5.1 Methodology

ANN-based Monitoring Method

Our main goal was to create a method that will support operators in real time monitoring of grid stability. This method should have been able not only to predict accurately the loading level of each phase, focusing primarily on alerting for line overloading, but also to continuously adapt to the current status of the grid.

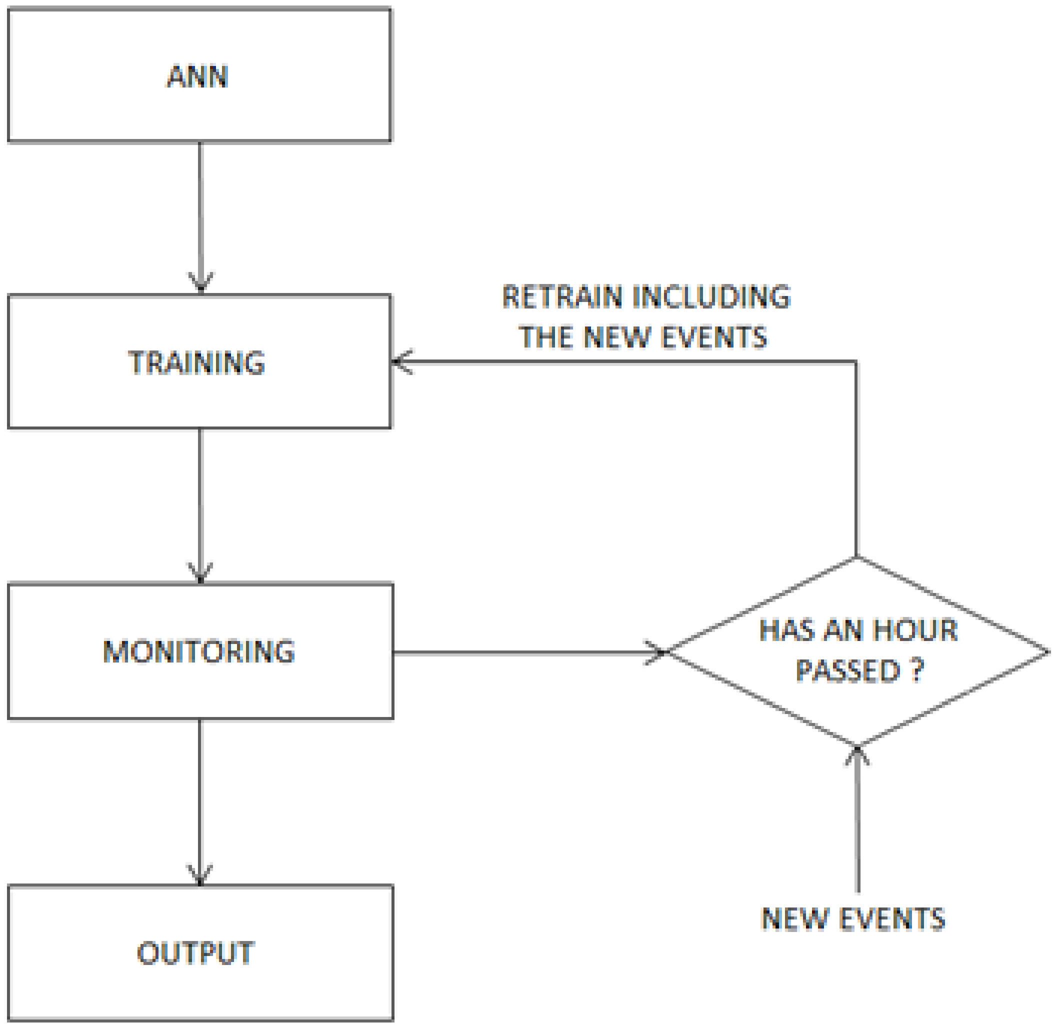

Figure 4.23: Flow chart of the proposed method

Power grids are dynamic and complex systems that depend on various parameters whose values may not easily anticipated. For instance, a random line failure or even just the passage of a big cloud on a very hot and sunny day can totally change system's dynamics. Huge variability in system parameter may be observed on price driven power systems that also integrate Renewable Energy Sources (RES).

The block diagram of the proposed method is depicted in Figure 4.23. The proposed method is consisted of a neural network that is initially trained to monitor the ampacity level of each phase of a distribution line, and then is repeatedly trained with new events every one hour. New events coincide with the values of the system parameters, which are are most recently obtained measurement pertained to the ampacity level of each phase.

Five different types of neural networks, namely a Feed-Forward Neural Network (FFNN),

Table 4.30: Most prominent ANN's and their particular efficiencies as expressed through the RMSE index

	FFN	MFFNN	CFNN	RNN	ENN
Phase A	25.1369	26.7234	26.3495	26.3568	27.0950
Phase B	26.8966	26.3576	25.8302	27.2004	28.8594
Phase C	18.9054	19.2489	19.1925	19.0309	19.1940

a multilayer FFNN (MFFNN), a Cascade-Forward Neural Network (CFNN), a Recurrent Neural Network (RNN) and an Elman Neural Network (ENN), were developed and evaluated using the RMSE index as efficiency measure to pick the most efficient one. Each ANN has 5 inputs, namely, "wind magnitude", "solar radiation", "temperature", "time_of_the_day" and "overall_demand". It also has 3 outputs concerning the ampacity level of each phase (A, B and C). More information can be found in [31] and [46].

The reader can see in Table 4.30 the most prominent candidates of each category and their efficiencies that are expressed through the RMSE index for each phase. In particular, there is a comparison among a three layer FFNN with 5 inputs, 25 neurons in the hidden layer and 3 outputs (5-25-3), a four layer (2 hidden) FFNN (5-15-15-3), a three layer CFNN (5-35-3), a three layer RNN (5-15-3) and an ENN (5-15-3). It must be mentioned that the RMSE values presented in the table are the average values of 12 randomly selected congested days. We were particularly interested to achieve the best performance on phase A, since the most congested cases are observed there, and secondarily on phase B since less congested cases are observed on that phase. No congested cases observed on phase C; at least on what concerns the way we have configured the distribution feeder. As it can be seen the FFNN (depicted in Figure 4.24) exhibited the best performance and was selected for the implementation of our method.

Training

All the five types of neural networks used in this study were trained using the algorithm described in the preliminaries section which is a widely used numerical optimization technique. In this particular study, because of the fact that we wanted to produce as realistic as possible test cases, the dataset comprised of a few congested days, and consequently, we were not able to divide the data in an efficient way. This is why we adopted an automated

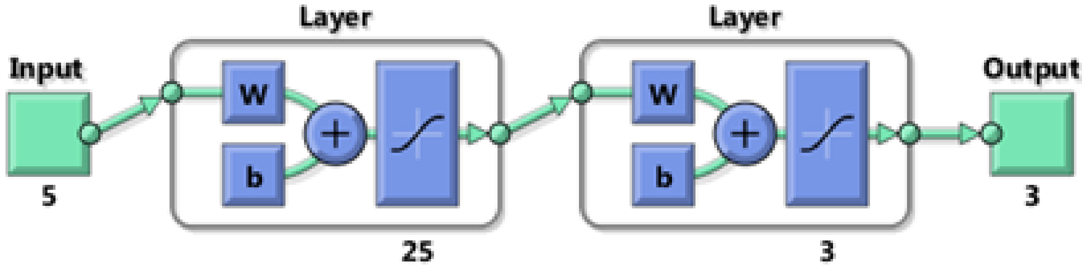

Figure 4.24: Selected ANN

correcting method, known as Bayesian regularization. A description of Bayesian regularization has been given earlier. The interested reader though, is referred to [48] and [47] for more information.

4.5.2 Experimental Configuration

As explained earlier, the aim of this study is to develop and test a monitoring method built on ANNs in a power market framework whose consumers participate into demand-response programs. This type of markets introduces uncertainty, as well as, many and sudden disturbances on the total demanded load within short time intervals. Consequently, a "successful" monitoring or predictive method is not a priori guaranteed.

The GridLAB-D [2] simulation platform was selected in order to develop the market framework and to obtain the data that we needed for training and evaluating the neural network. It must be mentioned here that GridLAB-D allows us to modify or add new capabilities and moreover to connect it with third-party tools. Most importantly, it provides us the capability to create a market framework where all the thermostatically controlled devices are able to take part into demand-response programs. It is consisted of several software modules each one responsible for the functionality of a general framework like market, powerflow, climate etc. Furthermore, each module is consisted of certain smaller objects, where each object is entrusted to perform a particular task. The thermostatically controlled devices, for instance, participate into demand-response programs through the object "transactive_controller", which is part of the market module.

The "transactive_controller" controls the operation of these devices in two ways. At first, it adjusts the operation set point by responding to the market status and specifically to the current market price. For example, consider a system that operates near its limits and a thermostatically controlled device like an HVAC system in the cooling mode. The market module in this case will increase the market price in an attempt to reduce the demand. The “transactive_controller” in response to this price increase will also increase the temperature set point of the HVAC system in an attempt to balance the monthly expenses of this system and the comfort of the user. The user has to define his/her desired temperature set point and a temperature zone into which he/she is willing to forfeit his comfort because of the high prices. In a different scenario where the prices are low, the “transactive_controller” will reduce the temperature set point in order to exploit the low prices. Secondly, the ”transactive_controller” manages the bidding strategy of the device for a specific operational period of a market. For instance, in the aforementioned case of the HVAC system, this happens by taking into consideration the internal air temperature of the house and the user’s preferences [49].

Regarding the simulated distribution system, we used the IEEE 13-bus test feeder which is depicted in Figure 3.3 [32]. By using the GridLAB-D simulation tool, we designed the feeder so as to contain 1306 residencies where each residency is inhabited by four people and contains three different devices that constitute its energy needs; namely lights, HVAC system and water heater. Moreover, we equipped the feeder with four different power producers with different capacities and marginal costs. Generator 1 has capacity of 2400 kW and marginal cost 0.3$/kWh, while the generators 2, 3 and 4 have capacity of 500, 400 and 100 kW respectively and marginal cost of 0.11, 0.25 and 0.15$/kWh.

Regarding the training of the ANNs, all the data that has been utilized were collected for a whole year period (2000) at fifteen minute intervals and stored in a database implemented through the MySQL software. The use of a database was of high importance for the handling of this high volume of information. Then, the most appropriate data points were extracted and used for the training procedure. Their selection was based on the fact that we wanted to ensure that our monitoring method will be mainly able to capture the loading peaks of each phase in order to assist to the efforts toward a safe and reliable operation of the power grid.

Finally, in order to identify the effectiveness of our method to capture loading peaks and especially line overloading, we lowered the thermal limit of the line 630-632 (see Figure 3.3) in all the three phases. This is the most heavily loaded line of the feeder because it connects it to the distribution network, and it is consequently of high importance with respect to system stability. A failure on this line is going to put out of service the whole feeder; in our scenario that is interpreted in 1306 residencies without electricity.

Table 4.31: A comparison between the proposed method and the selected ANN expressed through RMSE index for each phase

	Selected NN	Proposed Method
Phase A	25.1369	23.0053
Phase B	26.8966	25.6070
Phase C	18.9054	19.5664

4.5.3 Results

This section is devoted to the presentation of some selected results. As mentioned earlier, our main intention was to create a method able to be constantly adjusted to the new dynamics formed during a day.

To that end, the selected ANN is retrained with new data - observed in real time - every hour. There was an extensive try-and-error procedure in order to find the best combination of the data to be used for the retraining process. If we use the whole amount of the old data (10269 entries) the new data will have zero impact. It turned out that the lowest RMSE value on each phase is observed when we keep the 10% of the old data (the data that the selected ANN was first trained - randomly selected) and enrich this data set with the new ones observed in real time. To further maximize the impact of the new data in the retraining process, since these data express the new dynamics, we included them 10 times in the training set.

The findings presented in Table 4.31, as well as in Figure 4.25 and 4.26, show that the proposed method is indeed capable of improving the efficiency of the predictions. As it can be seen, there is a 2.1316 decrease in the RMSE value on phase A (23.0053 instead of 25.1369) and 1.2896 on phase B (25.6070 instead of 26.8966). There was a small increase on phase C (19.5664 instead of 18.9054 - 0.661) though, but since no congestion has observed there, we can accept it.

In Figure 4.25 the reader can see a graphic representation of the ampacity level of phase A for the 18 th of November 2001 using both a simple ANN (below) and the proposed method (above). The blue solid line stands for the predicted values, while the red dashed one for the real values. The x-axis shows the time of the day, while the y-axis the ampacity of the phase in amperes. A positive value indicates that the phase is overloaded, while a negative value the opposite.

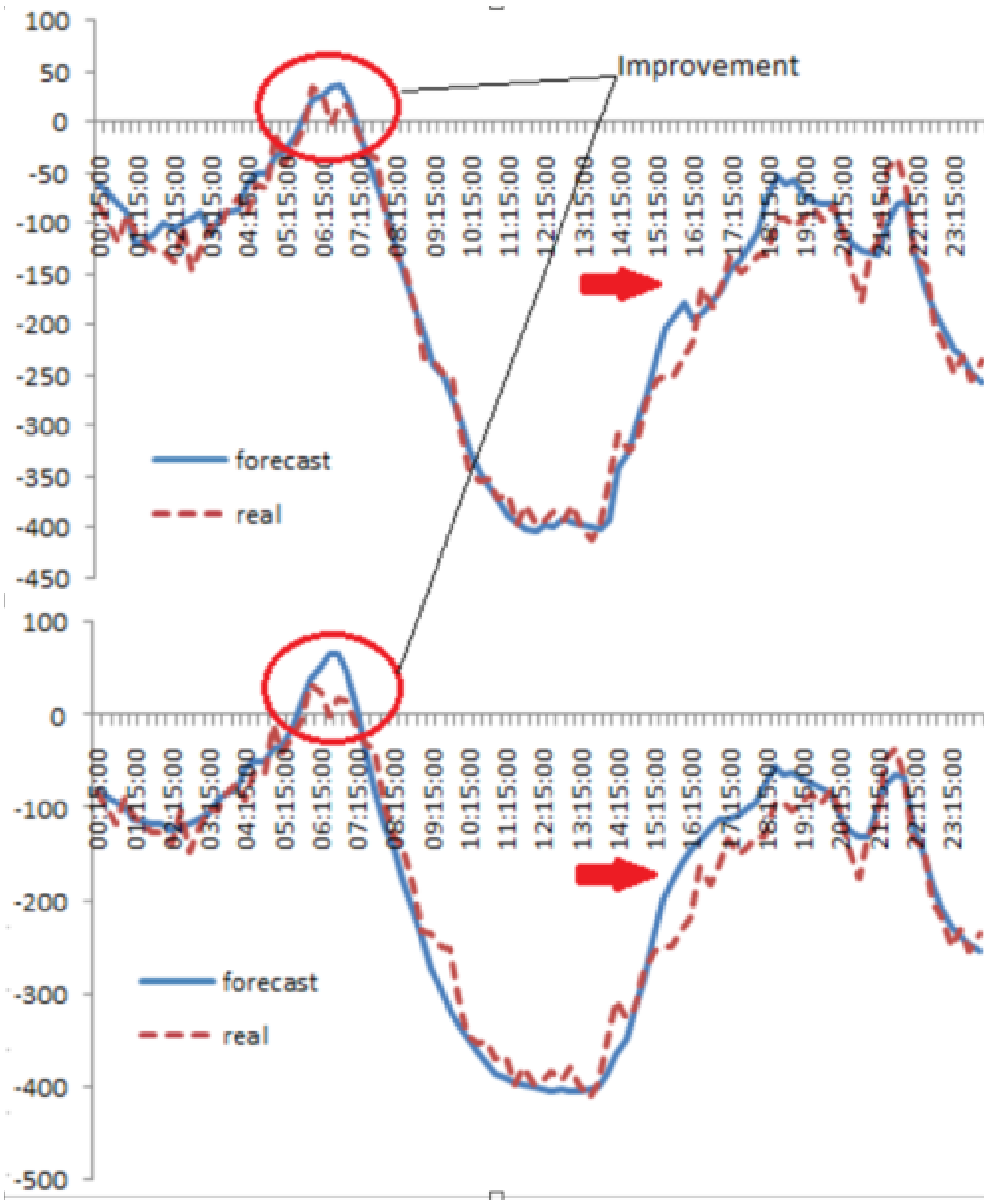

Figure 4.25: Predicted Vs real ampacity level on phase A when using a simple ANN (above) and when using the proposed method (below) (November 18th 2001) (x:time,y:amperes)

In Figure 4.26 the reader can see a graphic representation of the ampacity level of phase A for the 1st of January 2001 using both a simple ANN (below) and the proposed method (above). Generally speaking, the 1st of January can be considered as a special day since it does not follow system's dynamics by any means. These days like New Year's Eve, Thanksgiving etc. must be handled in a different way if we want to be precise. A separate ANN has to be used only for these days for instance. However, as it can be seen here, the proposed method

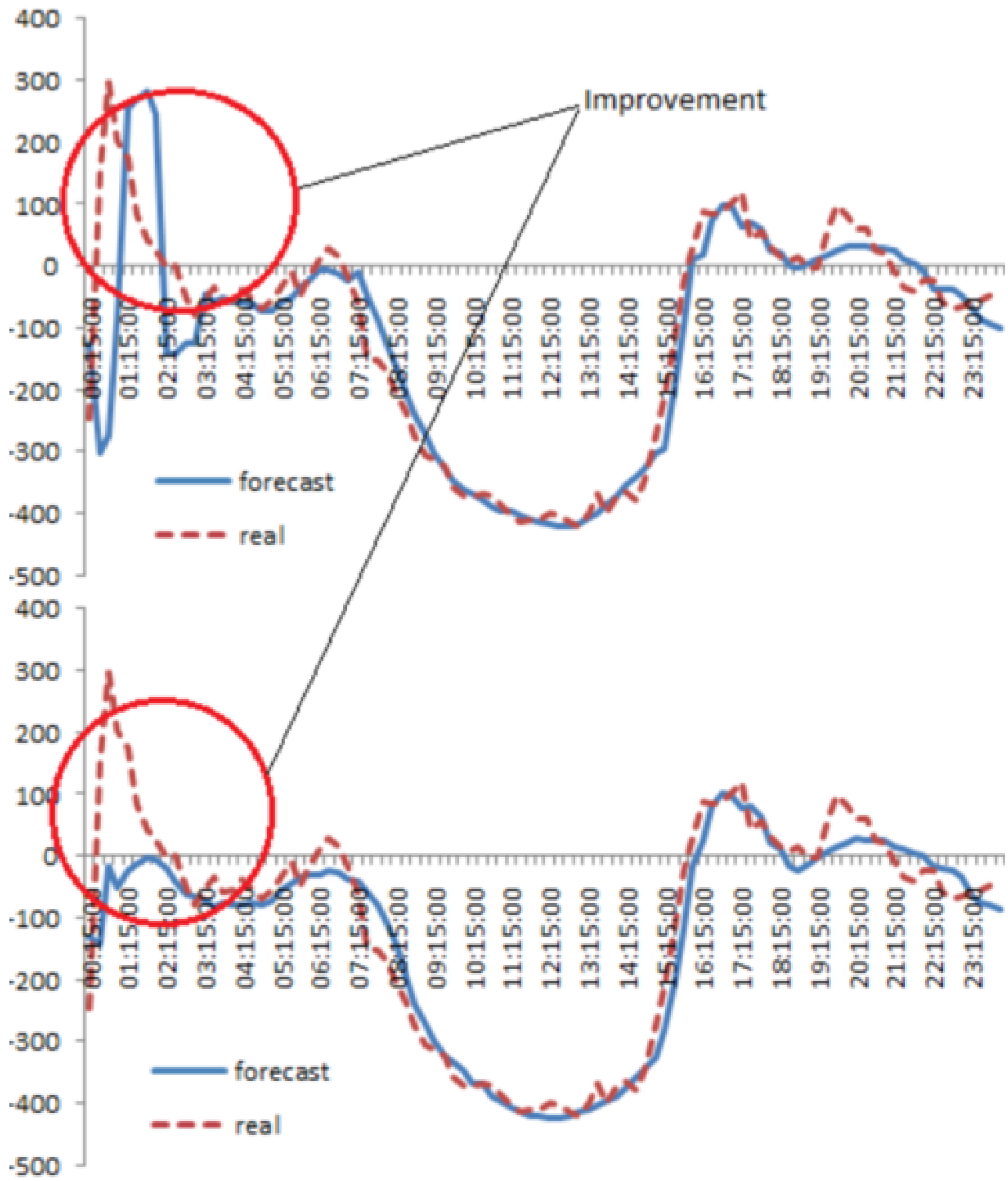

Figure 4.26: Predicted Vs real ampacity level on phase A when using a simple ANN (above) and when using the proposed method (below) (January 18th 2001) (x:time,y:amperes)

is able to indicate the possibility of congestion. Maybe this is the example where the user can clearly see how the proposed method is able to be adjusted, as the time passes, and improve its predictions; there is a high error, compared to the real values. Nevertheless, the predicted values are certainly improved, related to a simple ANN, and able to indicate high risk time intervals.

4.5.4 Conclusions

A three-phase line overloading monitoring method based on ANNs was developed and evaluated with respect to accuracy and efficiency in this study. The neural network should be repeatedly trained every hour with new events in order to increase its accuracy and to be always updated regarding the system dynamics. A secondary but equally important goal was the evaluation of ANNs in new generation power systems where the customers are eligible of participating into demand-response programs. Although neural networks have been utilized a lot the past decades, their effectiveness into this type of markets cannot be guaranteed a priori. The reason is the uncertainty and the disturbances on the load of the system introduced by the participation of customers on the formation of the market price and the overall demand.

It was proved that the proposed method is able to closely follow the real loading of the line with high accuracy. Although it did not in some instances predict the exact time of overloading, it was able to define a certain time period of high risk regarding the presence of overloading in all the three phases of a line. Moreover, it is able to adjusts to the new dynamics of the system and further improve its accuracy.

4.6 Hierarchical Method Based on Artificial Neural Networks for Power Output Prediction of a Combined Cycle Power Plant

4.6.1 Introduction

The combined cycle power plants are utilized more and more all around the world due to the more efficient power production that these units can offer. A combined cycle power plant is the result of a combination of two or more thermodynamic cycles aiming to exploit wasted forms of energy in order to produce more electricity - compared to the one produced by a conventional power plant - from the same fuel source [51]. Moreover, the deregulation of the power system and the structure of new generation power markets, formed as a two settlement system that combines day-ahead, as well as, real-time transactions and balancing actions, renders the development of anticipation applications more pressing than ever. In new generation power markets, the customers and the power generators participate in a bid-based market seeking and offering real power for specific time intervals. This is the way that the supply and the demand of the system are formed. Then, a central entity, the System Operator, is responsible to match supply and demand in the most economic way trying at the same time to retain the power system safe and uninterrupted. In other words, it is responsible to settle on which power units will be dispatched in order to meet the demand for both types of transactions; the day ahead and the real-time [1] [8]. Hence, the accurate prediction of the power generation of such type of power plants, plays a significant role helping the system operator to make precise decisions and take the necessary actions towards the proper operation and coordination of all the engaged parties. The reliability and the economic viability of the power system as a whole is consequently strengthened.

In parallel, it is highly recognized that the power output of these units, strongly depends on features related to the hourly average ambient variables of Temperature (T), Ambient Pressure (AP), Relative Humidity (RH) and finally, Exhaust Vacuum (V). Any sharp variation even on one of these features may affect the power output of the generation system and decrease its overall performance. Hence, any predictive effort has to deal with these features. To that end, the present paper proposes a new method for predicting the power output of a CCPP based on artificial neural networks regarding the aforementioned parameters. In fact, the contribution and the main goal of this study is to develop a predictive method that decreases as much as possible the forecast error; a common side effect of all regression applications. Towards that direction, a two layer hierarchical architecture that combines different types of ANNs is developed and evaluated. A detailed description follows in Section 4.6.3, but in a nutshell, through this stratification we aim to reduce the error that would be produced if only one neural network was used for the same purpose. A parallel, but of equal importance goal, is the competence assessment of neural networks in such type of applications.

4.6.2 Literature Review

A plethora of Artificial Intelligence (AI) techniques has been proposed in the literature towards the development of anticipatory methods in order to confront the challenges related to new generation power systems. A medium term load forecasting method is developed in [52]. More precisely, two different models based on kernel machines, namely Gaussian process and relevance vector regression combined with Gaussian kernels, were developed and compared against the ARMA (2,2) model with very promising results. Support Vector

Machines (SVMs) and fuzzy inference techniques are combined to produce an anticipatory system regarding the states of a complex power system in [53].

Getting closer to anticipatory methods that have to deal with generators' power output, an adaptive neuro-fuzzy inference system is utilized in [54] to predict the turbine generator power output of nuclear power plants. The authors in this study use the operating parameters as inputs in a Gaussian function. The validity of this approach is assessed using the root means square error. As stated earlier, the variation in the ambient temperature, pressure and the relative humidity affects the power production of the power plants. Therefore, their effects are commonly known and studied [55] [56]. In [57], the output of a combined gas and steam turbine is predicted using a dataset gathered for a period of 6 years. The data points of the data set have information retrieved hourly from various sensors when the power plant works with full load. Local and global learning methods are used to predict the power output. The predictive model is formed using additive and multivariate regression, k-NN, ANN and k-means clustering. For a different type of power plant, a combined Heat Power (CHP) plant, the power output is predicted using as data points the measurements of a one month period and as inputs the ambient temperature, the relative humidity and the ambient pressure. The Bagging REP Tree (BREP) algorithm is used in [58] for the prediction of the electrical power output of a base load operated combined cycle power plant. Moreover, in [59] a particle swarm optimization trained Feed-Forward Neural Network (FNN) is utilized in order to make predictions regarding the power production of CCPPs. The neural network takes as inputs the ambient temperature, the atmospheric pressure, the relative humidity and the exhaust vacuum. A Matlab simulator has been developed by the authors in [60] in order to study and analyze the effects of different input features of both gas and steam cycle of the power output of a CCPP. Besides the prediction of a gas-steam based power plant in [61] the authors propose a prediction model based on artificial neural networks to predict the power output of a coal–fired power plant focusing on investigating the feasibility of a model based on neural networks to such a power plant. The proposed artificial neural network model consists of two ANN models with the one to be associated with the boiler and the second one with the turbine. The data set for the prediction procedure consists of real power plant measurements. A same model using ANNs has been also studied in [62] [63] using real measurements for the steam process of a biomass and coal co-fired power plant.

In addition, there are studies in the literature concentrated on power plants that do not solely focus on the prediction of the power output of a generating unit. For example, a lot of researchers have coped with techniques and methods in order to predict the performance and the operational parameters of gas turbine power plants as described in [64], or the thermal performance of a biomass based integrated gasification power plant as discussed in [65].

4.6.3 Methodology

In this section, the proposed method and the different types of neural networks – namely Feed-Forward, Cascade-Forward and Recurrent neural network - used in this study are presented and described. Moreover, the motivation behind the development of the method, as well as the proposed architecture is also presented and discussed.

Artificial Neural Networks

An artificial neural network is a system that consists of a set of units called neurons. In predictive applications, a neural network is used to map already existed input and output values of a system in order to make predictions for unseen input values in the future [66]. In a process analog to the process taking place in a biological neuron of a human brain, the input signals are weighted and forwarded to the neurons that are connected to, being processed by an activation function and forwarded to the next layer of neurons, until reach the output layer. In particular, each connection between two neurons has a specific weight that "modifies" (weighs) the output of the corresponding neuron. During the training stage, the neural network is fed with some inputs and the corresponding for those inputs' target values. Then, the weight of each connection is adjusted according to the produced error, which is the difference between the output of the neural network and the target value. This procedure, well-known as supervised learning, is repeated for the same inputs and target values, until a convergence criterion, i.e. an acceptable error level or a maximum number of iterations for example, is met [31].

Feed Forward Neural Networks

The main characteristic of a Feed Forward Neural Network is that the input and the output signals do not form a cycle. More specifically, the information flows only to the one direction; from the input nodes to the hidden ones (i.e. the neurons between the input and the output layer of neurons) and at the end to the output nodes. It is the most used and well-studied category of neural networks. According to [31], it is used in the 80% of the applications based on neural networks and it is able to map any non-linear relationship using at least one hidden layer of neurons. The most common, in terms of simplicity, Feed-Forward Neural Networks are the Single Layer Perceptron (SLP) and the Multi-Layer Perceptron (MLP). In Figure 4.27 architecture of such a neural network is depicted.

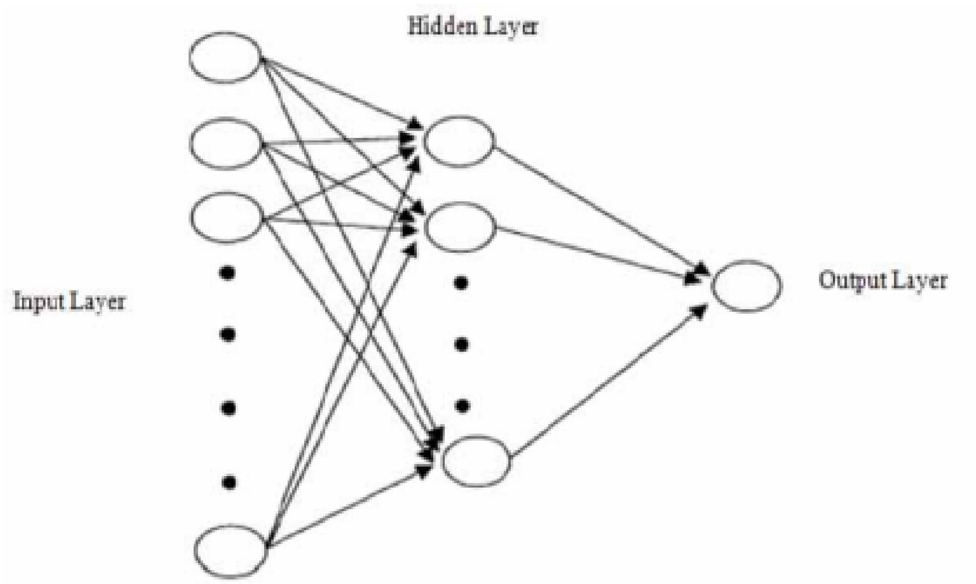

Figure 4.27: A typical Feed Forward Neural Network with one hidden layer [3]

Recurrent Neural Networks

In contrast to the Feed-Forward Neural Networks, in the Recurrent Neural Networks (RNNs), there is a bi-directional flow of information. A virtual internal memory is created that way, giving the capability to model systems with temporal behavior. The most commonly used types of this category are the Hopfield Network and the Simple Recurrent Network. The reader can see in Figure 4.28 a neural network of this category. It is a simple Recurrent Neural Network with one input, one output connection, as well as, connections from the input layer to itself forming three different directed cycles.

Cascade Neural Networks

The Cascade Forward Neural Networks (CFNNs) have similar characteristics with the Feed-Forward Neural Networks. The difference is that a connection between the layers exists as well. In particular, there is a weighted connection between the input layer and every

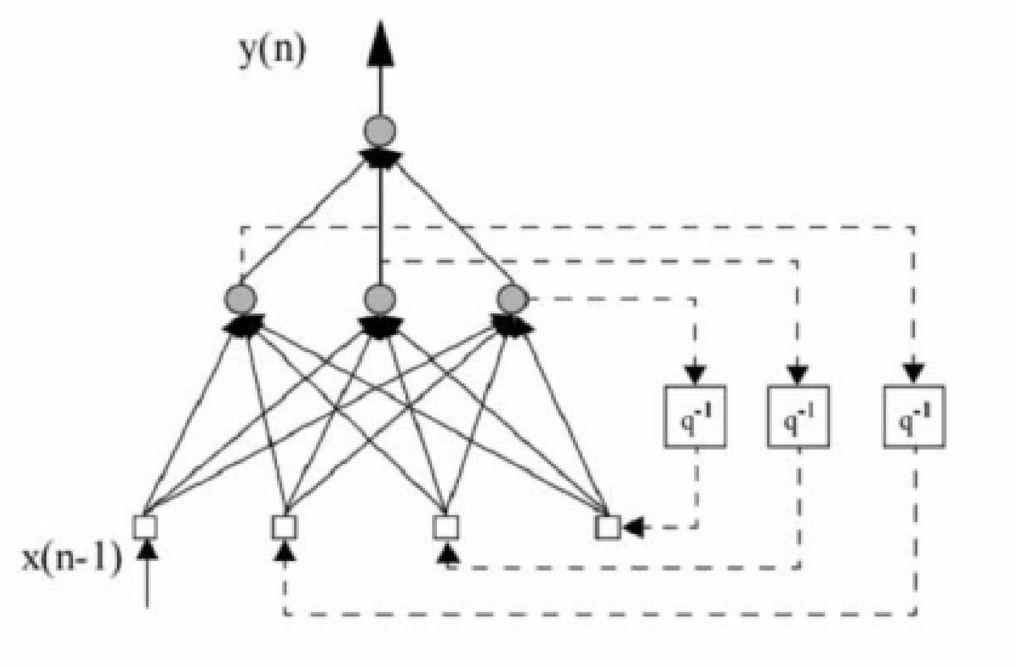

Figure 4.28: A Simple Recurrent Network with one input and one output [4]

previous layer as well as, to the layers that follow. A big advantage of this type of neural networks is that according to [46], these connections might benefit the training process of the neural network reducing the speed of the whole procedure. The reader can see in Figure 4.29 the architecture of a Cascade Neural Network.

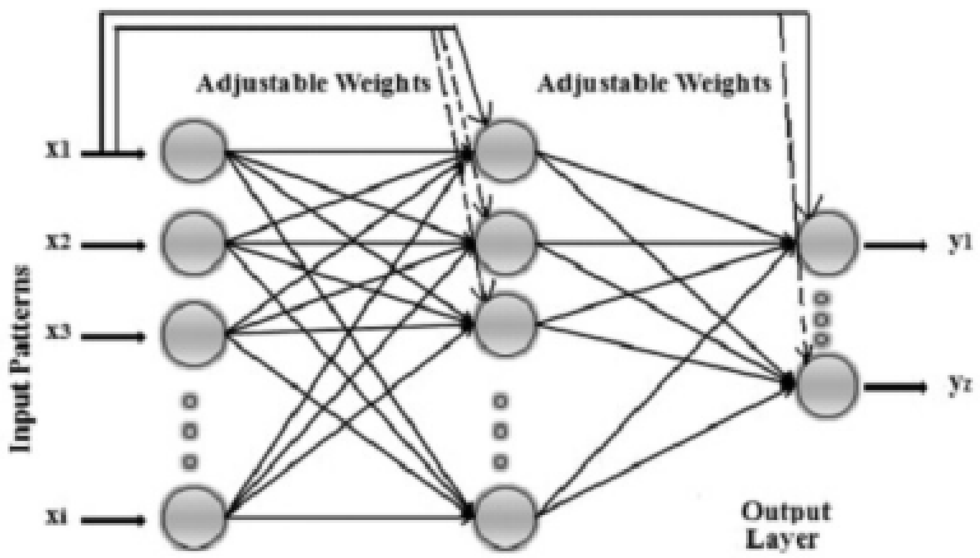

Figure 4.29: A Cascade Forward Neural Network architecture [5]

4.6.4 Proposed Methodology

Predictive applications and especially those that have to deal with safety and reliability tasks have to be precise and effective. On the other hand, the biggest problem of any application related to regression is the produced error. To that end, a method based on neural networks is developed and evaluated aiming to reduce as much as possible that error. More precisely, three different neural networks are combined in order to form a two layer architecture as the one depicted in Figure 4.30.

The idea behind this architecture is simple. Both the neural networks of the first layer are fed with the same input, that is the hourly measurements of temperature, exhaust vacuum, relative humidity and ambient pressure, and produce a slightly different output, that is the power output of the CCPP. Moreover, they also produce an error related to the efficiency of each neural network that, as told in the previous section, is the difference between the produced output from the neural network and the target value during the training stage. Then, these outputs are weighted through a third neural network in a try to reduce as much as possible the previously produced error.

The selected for the training of the neural network algorithm was the Levenberg-Marquardt [47]. The arlogirithm has been described in detail previously. It was also combined with Bayesian regularization in order to overcome the problem of data over-fitting [48]. The over-fitting is a very unpleasant situation where the neural network, although it exhibits a very small error for known inputs, fails to generalize for unknown ones.

4.6.5 Data

All the neural networks used in this study were trained with real data acquired from the UCI machine learning repository [67]. As stated earlier, these data consist of the hourly measurements of Temperature (T), Exhaust Vacuum (V), Relative Humidity (RH) and Ambient Pressure (AP) that represent the input values, while the related energy production of the CCPP constitutes the target value which was used during the training, as well as, during the evaluation stage. All the data were collected for a period of six years spanning from 2006 to 2011. From the total number of 9568 data points, the 9519 data points were used in order to train each neural network, while the last 49 in order to evaluate the performance of each individual neural network, as well as, the performance of the proposed method. Finally, all the codes regarding the development, the training and the evaluation of the neural networks have been

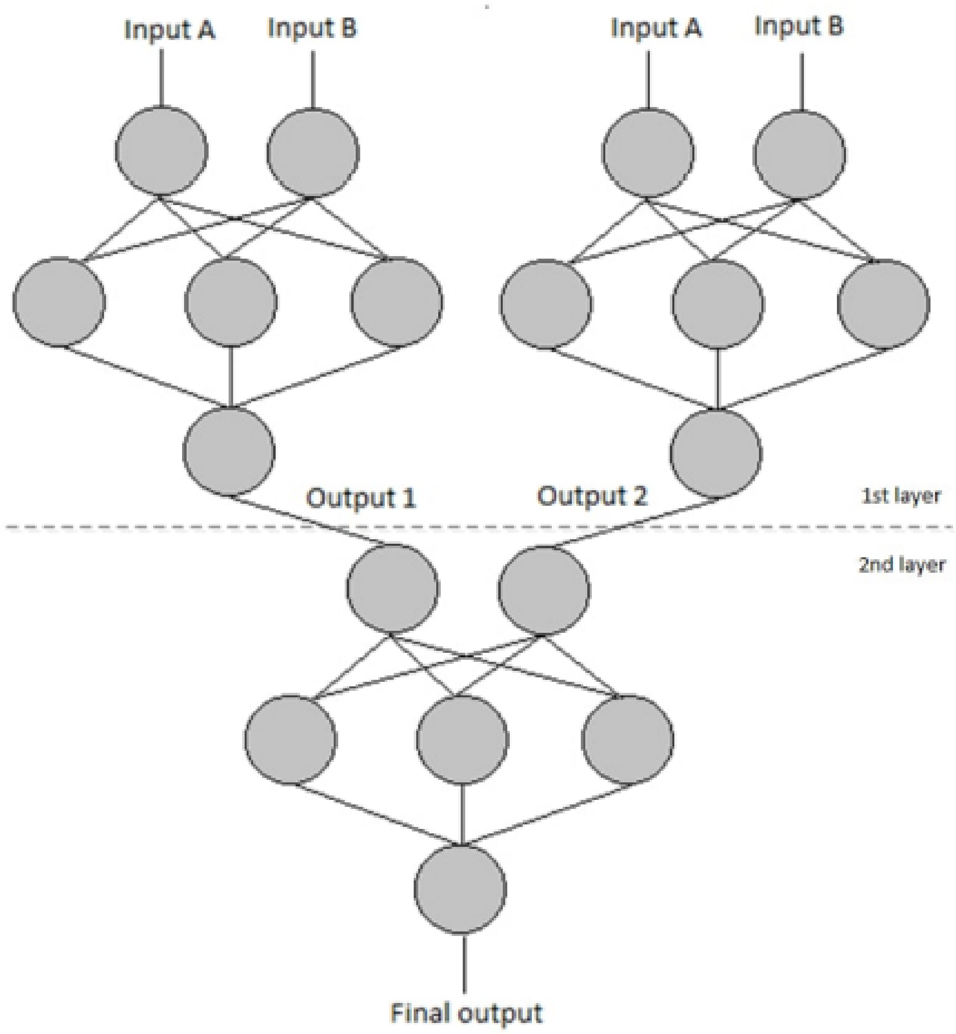

Figure 4.30: The two-layer hierarchical architecture of the proposed method

implemented by using the MATLAB (Matlab R2014a, Version 8.3.0.532) software.

4.6.6 Results

In this section, the experimental results of the study are presented and discussed. At first, three different ANNs, namely a Feed Forward, a Cascade Forward and a Recurrent neural network were trained and evaluated regarding their efficiency with the Root Mean Square Error (RMSE) to be used as a validity measure.

More descriptively, the first neural network, which is depicted in Figure 4.31, is a three layer Feed Forward Neural Network with 4 inputs, 70 neurons in the hidden layer and 1 output (4-70-1). The other two, a three layer Cascade Forward and a three layer Recurrent Neural

Network, depicted in Figure 4.32 and Figure 4.33 respectively, have 4 inputs, 40 neurons in the hidden layer and 1 output (4-40-1) each. It must be mentioned here that these particular neural networks, as well as their particular architectures and number of neurons in the hidden, were selected through a thorough trial and error procedure before we ended up with the most efficient ones.

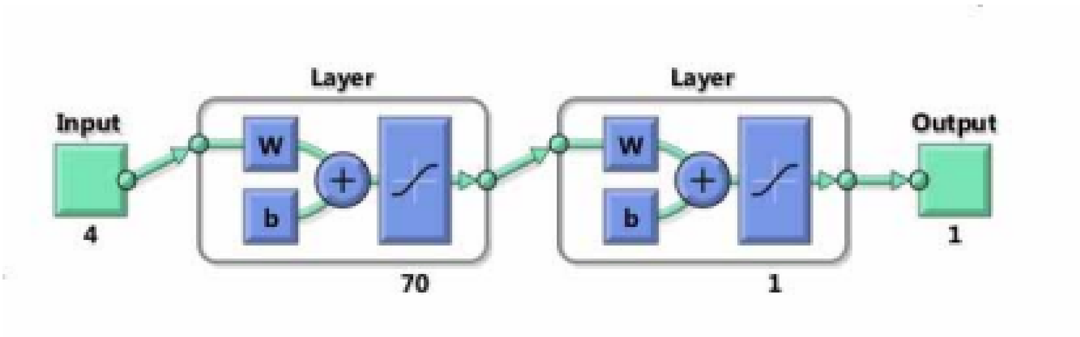

Figure 4.31: Architecture of the utilized Feed Forward Neural Network (neurons 4-70-1)

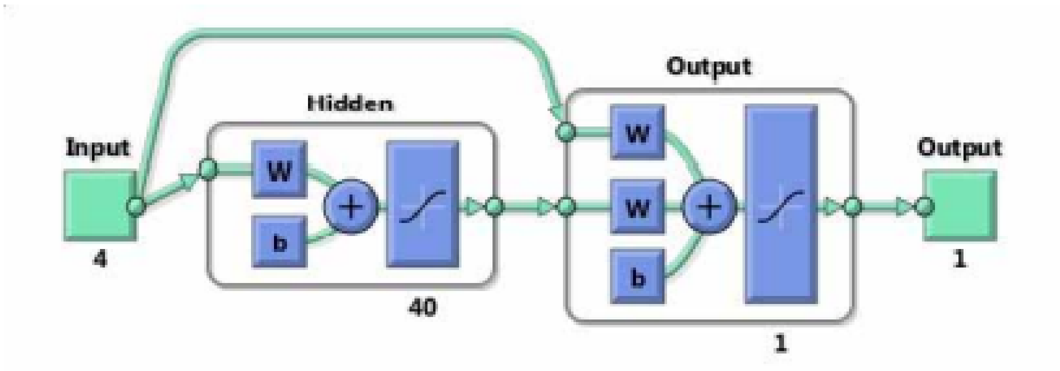

Figure 4.32: Architecture of the utilized Cascade Forward Neural Network (neurons 4-40-1)

The reader can see in Tables 4.32, 4.33 and 4.34 the performance and a comparison between the predicted by each neural network energy output and the actual one. The reader can see in the first column the real amount of the produced power in kW, in the second column the predicted one, while in the third the associated error. As mentioned in the previous section, the last 49 simulation hours were kept for evaluation purposes. Nevertheless, in Tables 4.32 through 4.34 only the first 10 simulation hours are presented.

As it can be seen by comparing the results contained in the tables, but more clearly, the RMSE of each neural network, the best performance achieved by the Feed Forward Neural Network with RMSE of 2.78. The Cascade Forward Neural Network closely follows with

Table 4.32: Real Vs Predicted power output of a CCPP for the Feed Forward Neural Network and the related error

Feed Forward Neural Network		
Real power output (kW)	Predicted power output (kW)	Error
445.49	442.3428069	3.147193
464.95	463.5198307	1.430169
472.16	474.4807357	-2.32074
456.57	460.2112978	-3.6413
475.13	474.2478537	0.882146
465.42	465.1347506	0.285249
445.6	447.5822983	-1.9823
440.25	443.1624597	-2.91246
439.03	438.1106933	0.919307
488.65	487.7701227	0.879877

Table 4.33: Real Vs Predicted power output of a CCPP for the Cascade Forward Neural Network and the related error

Cascade-Forward Neural Network		
Real power output (kW)	Predicted power output (kW)	Error
445.49	443.1110335	2.378967
464.95	464.2752949	0.674705
472.16	474.8199555	-2.65996
456.57	462.4458603	-5.87586
475.13	474.2020233	0.927977
465.42	464.5165018	0.903498
445.6	448.3570633	-2.75706
440.25	444.2954533	-4.04545
439.03	438.2306431	0.799357
488.65	487.5428992	1.107101

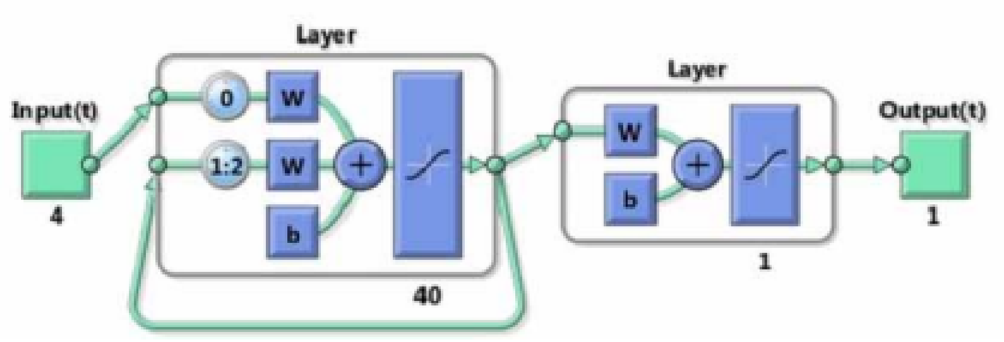

Figure 4.33: Architecture of the utilized Recurrent Neural Network (neurons 4-40-1)

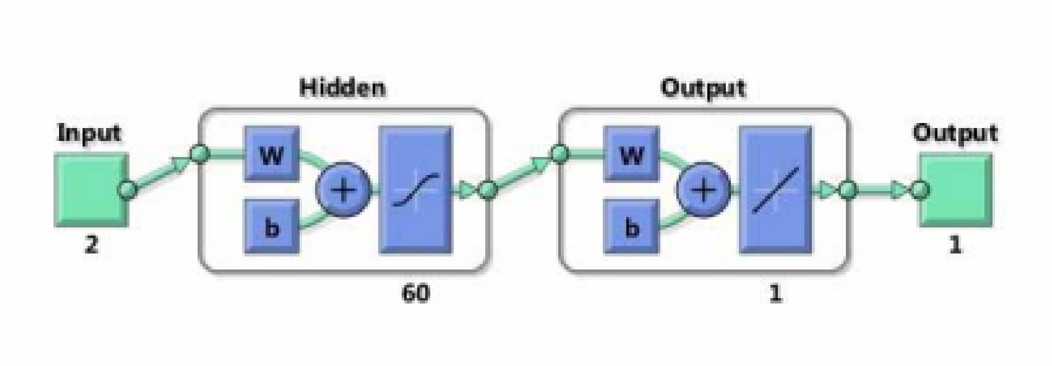

Figure 4.34: The architecture of the Feed Forward Neural Network used in the second layer (neurons 2-60-1)

3.25, while the Recurrent Neural Network comes last with RMSE of 3.37. The best two, with regard to their efficiency neural networks, namely the Feed Forward and the Cascade Forward neural network, were selected and their output was used as input to a third one which was trained regarding the same target value, i.e. the power output of the CCPP, as well. On the other hand, the root mean square error, as enlisted below, was computed for the whole set of 49 data points.

- RMSE = 2.78 for the Feed Forward Neural Network
- RMSE = 3.25 for Cascade Forward Neural Network
- RMSE = 3.37 for the Recurrent Neural Network

This third neural network, which actually constitutes the second layer of the architecture, (see Figure 4.34) was selected based on a trial and error procedure as well. Different types

Table 4.34: Real Vs Predicted power output of a CCPP for the Recurrent Neural Network and the related error

Cascade-Forward Neural Network		
Real power output (kW)	Predicted power output (kW)	Error
445.49	443.0176727	2.472327
464.95	463.8664231	1.083577
472.16	473.6074588	-1.44746
456.57	462.6330089	-6.06301
475.13	473.3270796	1.80292
465.42	464.4199436	1.000056
445.6	447.52339	-1.92339
440.25	444.6366228	-4.38662
439.03	438.6519933	0.378007
488.65	487.0126941	1.637306

of neural networks (categories, architectures, activation function, etc.) were trained and evaluated regarding their efficiency on mapping the outputs of the previous layer with the real output of the under consideration CCPP. The selected one was a three layer Feed Forward Neural Network with 2 inputs, 60 neurons in the hidden layer and 1 output (2-60-1). The two inputs are the output of each neural network of the previous layer. The difference in this neural network is that the activation function used in the output layer was a linear one instead of the sigmoid used in all the previous cases.

On one hand, in Table 4.35, the reader is able to observe the performance and a comparison between the predicted by our method power output of the CCPP and the real one as obtained from the repository for the first 10 hours of the simulation. On the other hand, it is observed that the achieved RMSE for the entire period of 49 hours was 2.71. As it can be seen, the reduction on the RMSE compared to the RMSE of the best neural network of the first layer, is not significant; 2.71 instead of 2.78. Nevertheless, it turns out that this method has indeed some merit and further research on this field is worthwhile.

The reader can also see in Figure 4.35 a graphical representation of the predicted against the real power plant's output for these 10 hours. The x-axis stands for the time, while the y-axis for the produced amount of power in kW. The blue solid line corresponds to the real power output, while the red dashed line stands for the predicted one.

In order to further evaluate the competence of the proposed method, we tried to weight the outputs of the ANNs of the first layer by using a simple linear regression model instead of a third neural network to see if we were going to achieve better performance. Moreover, we also tried to predict the final power output of the power plant by using a simple linear regression model alone to check if this method produces more accurate results. The outcome of both approaches is presented in Table 4.36 and 4.37 respectively.

Generally speaking, a linear regression prediction model attempts to model the correlation of a variable, called criterion variable, for which we make the prediction and a variable, called predictor variable, on which the predictions are based on. When the predictor variable is just one, then the linear regression is called simple linear regression [68].

The RMSE for the simple linear regression model was 3.91, and 2.79 when the outputs of the first layer were weighted using a simple linear regression model instead of a third neural network. Clearly, our method provides better, even slightly, performance indicating the potential of neural networks for being applied for CCCP power output predictive applications. Moreover, it turned out that the pure simple linear regression model exhibits the lowest performance among all the other techniques used in this study.

4.6.7 Conclusions

Summarizing, the aim of this study was twofold. To develop and evaluate an accurate predictive method for the power output of a CCPP. This is very important since a failure or a misinterpretation of the weather conditions prevailing that time and affect the produced amount of power, may put at risk not only the power plant, but the whole power system. Moreover, new generation power markets are organized as a two-settlement systems where real-time and day-ahead market operations take place. The system operator which is responsible for keeping the whole system in balance must be always and in time aware of every aspect that can emerge unforeseen events that might put the stability of the system in risk. The evaluation of the capability of neural networks to carry out this task was a second, but equally important goal of this research.

To that end, a two layer hierarchical predictive method based on neural networks was developed and evaluated. The proposed method, was aiming to reduce as much as possible

Table 4.35: Real Vs Predicted power output of a CCPP of the proposed method and the related error

Two Layer Hierarchical Method		
Real power output (kW)	Predicted power output (kW)	Error
445.49	442.4547479	3.035252
464.95	463.8262386	1.123761
472.16	474.8197756	-2.65978
456.57	460.758921	-4.18892
475.13	474.5172618	0.612738
465.42	465.1182771	0.301723
445.6	447.5815609	-1.98156
440.25	443.2947447	-3.04474
439.03	438.2326129	0.797387
488.65	488.3187657	0.331234

Table 4.36: Real Vs Predicted power output of a CCPP and the related error when the outputs of the first layer are weighted through linear regression

Two Layer Hierarchical Method		
Real power output (kW)	Predicted power output (kW)	Error
445.49	442.3892562	3.100744
464.95	463.5872465	1.362753
472.16	474.5298208	-2.36982
456.57	460.3806676	-3.81067
475.13	474.269267	0.860733
465.42	465.1059708	0.314029
445.6	447.6346258	-2.03463
440.25	443.2357418	-2.98574
439.03	438.1065873	0.923413
488.65	487.7925819	0.857418

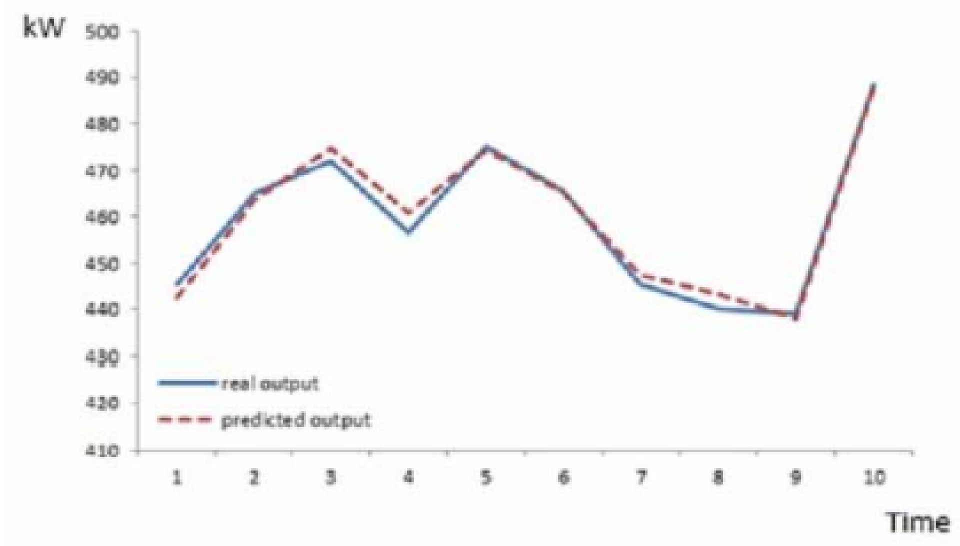

Figure 4.35: . A graphical representation of the predicted power output against the real power output of the CCPP. The x-axis stands for the time, while the y-axis corresponds to the produced amount of power in kW (x:time,y:kW)

the produced by any predictive application error. Obviously, this error has to be detained as small as possible, especially in real time applications where a wrong decision can have great impact and put at risk the whole system for which the predictions are made. It must be also mentioned, that the achieved results were also compared with a pure simple linear regression model in order to evaluate whether this particular method is able to produce better results.

The findings showed a slightly reduced root mean square error compared to the predicted values obtained with methods that use only one neural network, as well as, to the simple linear regression model. Getting into more details, the RMSE taken by our method was 2.71 against the 2.78 obtained by a Feed Forward Neural Network, which was the one that exhibited the best performance among all the other methods used in this study. These findings show that further research in this field is necessary and may be required. Moreover, it turns out that neural networks are more than capable to pave the way and be the cornerstone for predictive methods and actions operating in new generation power systems which require fast and accurate responses.

Table 4.37: Real Vs Predicted power output of a CCPP of a simple linear regression model and the related error

Two Layer Hierarchical Method		
Real power output (kW)	Predicted power output (kW)	Error
445.49	443.5980485	1.891952
464.95	464.3852715	0.564728
472.16	470.6863855	1.473614
456.57	462.7396022	-6.1696
475.13	474.8186594	0.311341
465.42	467.4462403	-2.02624
445.6	450.0418232	-4.44182
440.25	444.5755809	-4.32558
439.03	434.6937103	4.33629
488.65	486.7046941	1.945306

Chapter 5

Conclusions

The current dissertation aimed to examine the feasibility of AI techniques, as well as to develop novel models based on AI in order to assist the DSO in the effort to ensure the stability of the distribution grid. This is essential due to the role of electricity in the modern societies. The uninterrupted operation of the power grid and the 24/7 access to electricity is paramount. Moreover, the advent of EVs and the constantly increasing introduction of the RES in the power network further complicates stability issues. The development of anticipation techniques is essential in order to mitigate and smooth electricity consumption. Demand is constantly changing. We need to find and evaluate ways to transfer electricity demand from peak times to times that electricity consumption is lower.

In the third chapter two studies regarding the anticipation of congestion have been presented:

- Ampacity Level Monitoring Utilizing Fuzzy Logic Theory
- Load Management of Electric Vehicles Charging in New Generation Power Markets Based on Fuzzy Logic and the Concept of Virtual Budget

In the first work, a method as close as possible to human reasoning in order to detect line overloading was developed using the fuzzy logic theory. Parameters such as ”temperature”, ”time” of the day and ”price” at each particular time were used in order to built our predictor.

The results were indeed very promising showing that the fuzzy logic area can be used for accurate predictions and measures in order to sustain systems’ stability. The predicted ampacity level closely followed the real ampacity level during the whole simulated day providing that way useful information to all responsible for the safety of the power system parties. Even when there was a faulty prediction, it was very close to the real one. To that end, even in that cases, the predictor was able to form a trend regarding the ampacity level, providing useful information to the SO.

In the second work, we tried to alleviate congestion by scheduling the charging of EVs in a power market the follows the rules of competition. This is a very important task because EVs can introduce a big load in the system and consequently, more congestion on peak times. Moreover, it has been observed that EVs can cause congestion on times that had never been observed previously. To that end, a fuzzy logic controller was developed and evaluated.

The concept of the ”virtual budget”, translated as the willingness of customer to pay for electricity in the established prices, was also introduced in that work seaking to develop a fully automated controller.

The results proved that fuzzy logic is indeed capable to support such concepts. The controller was able to schedule the charging of EVs in less congested times. It was also proved that all the EVs were properly charged and thus able to support users needs regarding their transportation.

In the fourth chapter four studies regarding the anticipation of congestion using ANNs have been developed and evaluated. Moreover, a study regarding the output prediction of a Combined Cycle Power Plant based on ANNs was also presented:

- Three-Phase Congestion Predictions Utilizing Artificial Neural Networks
- Backpropagation Neural Network for Interval Prediction of Three-Phase Ampacity Level in Power Systems
- Distribution Congestion Prediction using Artificial Neural Networks for Big Data
- Three-Phase line Overloading Predictive Monitoring Utilizing Artificial Neural Networks
- Hierarchical Method Based on Artificial Neural Networks for Power Output Prediction of a Combined Cycle Power Plant

The goal of the first study was to develop and evaluate a three-phase line predictor able to provide accurate predictions inside the big data environment. A feedforward backpropagation neural network developed and trained in Matlab was used for this purpose. All the data regarding the training and the evaluation of the model were retrieved from the GridLAB-D simulation tool and handled using MySQL.

It has been proved that the predicted results was always able to follow the trend of line loading indicating that way the precise time of congestion presence. The model was also able to provide a good prediction of the true overloading with high enough accuracy.

The second study aimed to develop and evaluate a NN based method for constructing predictive intervals for enhanced ampacity overloading predictions. The dynamics of the new generation markets where consumers actively participate into the market and form energy usage prices are totally different from what we knew. Thus, models that provide plaint point predictions it might be incapable of providing accurate insight of line overloadings.

To that end, the Levenberg-Marquardt training method was used to train a NN aiming to primarily capture ampacity level peaks. The NN was designed to provide a prediction regarding the ampacity level of each particular phase and a value providing a predictive window. This particular value was defined by using the standard deviation of the target value of ampacity level for known input data.

Eight different test case scenarios were used in order to assess our model. For each scenario the standard deviation of the target value for a particular time period was used in order to construct the aforementioned predictive interval. It has been concluded that the three hour approach provides a good level of accuracy with the narrowest predictive window.

In the third study, we aimed to develop a congestion predictor able to perform in a big data environment. The predictor should be able to monitor a part of the distribution system and provide congestion and time of congestion presence predictions regarding that particular part.

A Feedforward Neural Network trained by the Levenberg-Marquardt algorithm combined with Bayesian regularization was used to that end. It was turned out that our method was able to provide narrow time windows of congestion presence and was always followed the real ampacity level.

A three-phase line overloading monitoring method based on ANNs was developed and evaluated in the fourth study. The idea was to develop a repeatedly trained neural network every hour with new data points in order to achieve the highest possible accuracy regarding congestion prediction. The biggest challenge in that consept was that in new generation power markets consumers are actively participate in the market forming energy usage prices and changing their habits according to the established price at each particular time.

The results showed that the predictor was able to capture the ampacity level of the under consideration distribution line with high accuracy. Moreover, although it was not able to provide the exact time of congestion presence at some times, it was always able to provide accurate time windows of congestion presence and thus, to indicate critical times where the SO should take measures to mitigate the load. It was also proved that the predictor was adjustable, meaning that it was able to improve its performance as time passed.

The goal of the last study was to develop and evaluate a predictive method for a power output production of a Combined Cycle Power Plant. The new generation power markets are organized as a two-settlement systems where real-time and day-ahead market operations take place. The system operator should always be aware of system dynamics in order to provide its safe and uninterrupted operation which is essential for the modern societies.

A two layer hierarchical predictive method based on neural networks was developed and evaluated. Moreover, our predictor was compared with different NN architectures in order to prove its ability to achieve better performance. The root mean square error was used for that purpose. Finally, the results were also compared to a pure linear regression model in order to evaluate whether this particular method is able to produce better results.

The results indicated that our method achieved better results, slightly reduced root mean square error, compared to architectures that use only one neural network, as well as, to the simple linear regression model.

Publications

[1] E. Tsoukalas R. Fainti, A. Nasiakou and M. Vavalis. Simulations of all electric vehicles on fixed routes. *International Journal of Electric and Hybrid Vehicles*, 5(4):296–316, 2013.

[2] M. Vavalis R. Fainti, A. Nasiakou and E. Tsoukalas. Design and early simulations of next generation intelligent energy systems. *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, 2014.

[3] M. Alamaniotis R. Fainti and L. H. Tsoukalas. Backpropagation neural network for interval prediction of three-phase ampacity level in power systems. *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, 2016.

[4] M. Alamaniotis R. Fainti, A. Nasiakou and L. H. Tsoukalas. Hierarchical method based on artificial neural networks for power output prediction of a combined cycle power plant. *International Journal of Monitoring and Surveillance Technologies Research, IGI-Globa*, 4(4):20–32, 2016.

[5] E. Tsoukalas R. Fainti, A. Nasiakou and M. Vavalis. Large scale simulations for electric energy markets. *5th International Conference on Information, Intelligence, Systems and Applications (IISA 2014)*, pages 160–165, 2014.

[6] A. Nasiakou R. Fainti E. Houstis G. Papavasilopoulos E. Sarri C. Nikolaou E. Tsoukalas, M. Vavalis and G. Koutras. Towards next generation intelligent energy systems: Design and simulations engines. *5th International Conference on Information, Intelligence,Systems and Applications (IISA 2014)*, pages 412–418, 2014.

[7] M. Alamaniotis R. Fainti and L. H. Tsoukalas. Three-phase congestion prediction utilizing artificial neural networks. *Proceedings of IISA 2016 - 7th International Conference on Information, Intelligence, Systems and Applications*, 2016.

[8] R. Fainti K. Mao L. H. Tsoukalas R. Bean M. Alamaniotis P. L. Lagari, A. Nasiakou. Evaluation of human machine interface (hmi) in nuclear power plants with fuzzy logic method. *Proceedings of IISA 2016 - 7th International Conference on Information, Intelligence, Systems and Applications*, 2016.

[9] M. Alamaniotis R. Fainti and L. H. Tsoukalas. Distribution congestion prediction using artificial neural networks for big data. *Proceedings of the 10th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion*, 2016.

[10] L. H. Tsoukalas M. Karasimou R. Fainti, M. Alamaniotis and I. Tsionas. Ampacity level monitoring utilizing fuzzy logic theory in deregulated power markets. *Proceedings of IISA 2017 - 8th International Conference on Information, Intelligence, Systems and Applications*, 2017.

[11] L. H. Tsoukalas M. Karasimou R. Fainti, M. Alamaniotis and I. Tsionas. Three-phase line overloading predictive monitoring utilizing artificial neural networks. *Proceedings of 2017 Intelligent System Applications to Power Systems (ISAP) Conference*, 2017.

[12] I. Tsionas M. Alamaniotis R. Fainti, M. Karasimou and L. H. Tsoukalas. Load management of electric vehicles charging in new generation power markets based on fuzzy logic and the concept of virtual budget. *Proceedings of IISA 2018 - 9th International Conference on Information, Intelligence, Systems and Applications*, 2018.

References

[1] W. W. Hogan. Competitive electricity markets: A wholesale primer. *http://www.hks.harvard.edu/fs/whogan/empr1298.pdf*, 1998.

[2] K. Schneider D.P. Chassin and C. Gerkensmeyer. Gridlab-d: An open-source power systems modelling and simulation environment. *Transmission and Distribution Exposition Conference, IEEE PES Powering Toward the Future, PIMS*, 2008.

[3] Egrioglu E. Aladag C. H. Tiring E. Alpaslan, F. An statistical research on feed forward neural networks for forecasting time series. In *American Journal of Intelligent Systems*, pages 2(3), 21–25, 2012.

[4] Lehtokangas M. Saarinen J. Kaski K. Koskela, T. Time series prediction with multilayer perceptron. In *Proceedings of the FIR and Elman Neural Networks. World Congress on Neural Networks*, pages pp. 1–5, 2012.

[5] Al-Allaf A. Omaima, N. Cascade-forward vs. function fitting neural network for improving image quality and learning time in image compression system. In *Proceedings of the World Congress on Engineering 2012*, pages Vol. 2, pp. 1–6, 2012.

[6] M. Alamaniotis V. Chrysikou and L.H. Tsoukalas. A review of incentive based demand response methods in smart electricity grids. In *International Journal of Monitoring and Surveillance Technologies Research*, pages 61–72, 2015.

[7] E. Tsoukalas M. Vavalis R. Fainti, A. Nasiakou. Design and early simulations of next generation intelligent energy systems. In *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, pages pp. 58–82, 2014.

[8] M. Vavalis M. Foti. Intelligent bidding in smart electricity markets. In *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, pages 68–90, 2015.

[9] David P. Chassin, Jason C. Fuller, and Ned Djilali. Gridlab-d: an agent-based simulation framework for smart grids.(research article)(report). *Journal of Applied Mathematics*, 2014, 2014.

[10] z. T. Taylor. *GridLAB-D Technical Support Document Residential End-Use Module Version 1.0*. United States. Dept. of Energy ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy, Washington, D.C : Oak Ridge, Tenn., 2008.

[11] D. P. Chassin. *GridLAB-D Technical Support Document Tape Modules Version 1.0*. United States. Dept. of Energy ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy, Washington, D.C : Oak Ridge, Tenn., 2008.

[12] K. P. Schneider, D. Chassin, Y. Chen, and J. C. Fuller. Distribution power flow for smart grid technologies. In *2009 IEEE/PES Power Systems Conference and Exposition*, pages 1–7, March 2009.

[13] Nathan D. Tenney. Gridlab-d technical support document: Climate module version 1.0 - pnnl-17679. Technical report, 2008.

[14] J.C. Fuller. *Modeling Wind Turbines in the GridLAB-D Software Environment*. United States. Dept. of Energy. Office of Science ; distributed by the Office of Scientific and Technical Information, U.S. Dept. of Energy, Washington, D.C. : Oak Ridge, Tenn., 2009.

[15] D. Aliprantis, S. Penick, L. Tesfatsion, and H. Zhao. Integrated retail and wholesale power system operation with smart-grid functionality. In *IEEE PES General Meeting*, pages 1–8, July 2010.

[16] P. Lakshmi V. Uma and J.D. Anunciya. Congestion management in deregulated power system by fuzzy based optimal location and sizing of upfc. In *WSEAS Transactions on Power Systems*, page 258–266, 2014.

[17] D.P. Mishra P. Ray and S. Mohaptra. Fault classification of a transmission line using wavelet transform fuzzy logic. In *1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES*, 2016.

[18] M.W. Mustafa A.A.M. Zin A.B. Khairuddin, S.S. Ahmed and H. Ahmad. A novel method for atc computations in a large-scale power system. In *IEEE Transactions on Power Systems*, page 1150–1158, 2004.

[19] F. Rezaei and S. Esmaeili. Decentralized reactive power control of distributed pv and wind power generation units using an optimized fuzzybased method. In *International Journal of Electrical Power and Energy Systems*, page 87:27–42, 2017.

[20] M.K. Hossain and M.H. Ali. Fuzzy logic controlled power balancing for low voltage ride-through capability enhancement of large-scale gridconnected pv plants. In *IEEE Texas Power and Energy Conference, TPEC 2017*, 2017.

[21] K. Basaran and N.S. Cetin. Designing of a fuzzy controller for grid connected photovoltaic system's converter and comparing with pi controller. In *2016 IEEE International Conference on Renewable Energy Research and Applications, ICRERA 2016*, page 102–106, 2016.

[22] K.R. Suraparaju and G. Pillai. Type 2 fuzzy logic–based robust control strategy for power sharing in microgrids with uncertainties in operating conditions. In *International Transactions on Electrical Energy Systems*, page 27(4), 2017.

[23] M. Abdel-Akher A. Selim and M.M. Aly. Plug-in hybrid electric vehicles aggregation and real-time active power control simulation analysis in distribution systems. In *2016 18th International Middle-East Power Systems Conference, MEPCON 2016 - Proceedings*, page 723–728, 2016.

[24] Y. Sun Q. Dong and P. Li. A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in china. In *Renewable Energy*, page 102:241–257, 2017.

[25] N.A. Kheir N.J. Schouten, M.A. Salman. Fuzzy logic control for parallel hybrid vehicles. In *IEEE Transactions on Control Systems Technology*, pages 10 (3), pp. 460–468, 2002.

[26] I. Kar M. Singh, P. Kumar. Implementation of vehicle to grid infrastructure using fuzzy logic controller. In *2012 IEEE Transactions on Smart Grid*, pages 3 (1), pp. 565–577, 2012.

[27] T. Senjyu M. Datta. Fuzzy control of distributed pv inverters/energy storage systems/electric vehicles for frequency regulation in a large power system. In *2013 IEEE Transactions on Smart Grid*, pages 4 (1), pp. 479–488, 2013.

[28] A. Cheriti H.Hemi, J. Ghouili. A real time fuzzy logic power management strategy for a fuel cell vehicle. In *2014 Energy Conversion and Management*, pages 80, pp. 63–70, 2014.

[29] G.P. Liu C.Y. Li. Optimal fuzzy power control and management of fuel cell/battery hybrid vehicles. In *2009 Journal of Power Sources*, pages 192(2), pp. 525–533, 2009.

[30] O. A. Mohammed T.Ma. Optimal charging of plug-in electric vehicles for a car-park infrastructure. In *2014 IEEE Transactions on Industry Applications*, pages 50(4), pp. 2323–2330, 2014.

[31] L.H. Tsoukalas and R.E. Uhrig. Fuzzy and neural approaches in engineering. In *John Wiley Sons, Inc. NY*, 1997.

[32] W.H. Kerting. Radial distribution test feeders. In *IEEE distribution planning working group report. IEEE Transactions on Power Systems*, page 6(3):975–985, 1991.

[33] Z. Baharudin B. Islam and P. Nallagownden. Development of chaotically improved meta-heuristics and modied bp neural network-based model for electrical energy demand prediction in smart grid. In *Neural Computing and Applications*, pages pp. 1–15, 2016.

[34] C. Karakus I. Mert and F. Unes. Estimating the energy production of the wind turbine using artificial neural network. In *Neural Computing and Applications*, pages 27(5), pp.1231–1244, 2016.

[35] L. Tesfatsion Q. Zhou and C.C. Liu. Short-term congestion forecasting in wholesale power markets. In *IEEE Transactions on Power Systems*, pages 26(4), pp.2185–2196, 2011.

[36] P. Zhang V. Rose L. Min, S.T. Lee and J. Cole. Short-term probabilistic transmission congestion forecasting. In *3rd International Conference on Deregulation and Restructuring and Power Technologies, DRPT*, page pp. 764–770, 2008.

[37] C.C. Liu G. Li and H. Salazar. Forecasting transmission congestion using day-ahead shadow prices. In *IEEE PES Power Systems Conference and Exposition, PSCE 2006 - Proceedings*, page pp. 1705–1709, 2006.

[38] F. Li and R. Bo. Congestion and price prediction under load variation. In *IEEE Transactions on Power Systems, 24(2)*, page pp. 911–922, 2006.

[39] S. Sharma and L. Srivastava. Prediction of transmission line overloading using intelligent technique. In *Applied Soft Computing Journal*, page pp. 626–633, 2008.

[40] S. Sharma and L. Srivastava. Hybrid-neural-network-based estimation of line overloading. In *International Journal of Computational Intelligence and Applications*, pages 7(4), pp. 385–400, 2008.

[41] S. Balaraman and N. Kamaraj. Cascade bpn based transmission line overload prediction and preventive action by generation rescheduling. In *Neurocomputing*, pages 94, pp. 1–12, 2012.

[42] M. Gavrilas O. Ivanov and B. Neagu. Intelligent monitoring and control in transmission and distribution networks. In *International Conference on Optimization of Electrical and Electronic Equipment*, page pp. 185–191, 2014.

[43] M. Gavrilas O. Ivanov. State estimation for power systems with multilayer perceptron neural networks. In *11th Symposium on Neural Network Applications in Electrical Engineering, NEUREL 2012 - Proceedings*, page pp. 243–246, 2012.

[44] A. M. Glazunova. Forecasting power system state variables on the basis of dynamic state estimation and artificial neural networks. In *Proceedings - 2010 IEEE Region 8 International Conference on Computational Technologies in Electrical and Electronics Engineering, SIBIRCON*, page pp. 470–475, 2010.

[45] K. Teeparthi and D.M. Vinod Kumar. Security-constrained optimal power flow with wind and thermal power generators using fuzzy adaptive artificial physics optimization algorithm. In *Neural Computing and Applications*, pages pp. 1–17, 2016.

[46] M. Beale H. Demutl and M. Hagan. Neural network toolbox for use with matlab. In *Natick, MA: The MathWorks, Inc*, 1998.

[47] F.D. Foresee and M.T. Hagan. Gauss-newton approximation to bayesian learning. In *IEEE International Conference on Neural Networks - Conference Proceedings*, pages vol. 3, pp. 1930–1935, 1997.

[48] D.J.C. MacKay. Bayesian interpolation. In *Neural Computation*, pages vol. 4, no. 3, pp. 415–447, 1992.

[49] K.P. Schneider J.C. Fuller and P.D. Chassin. Analysis of residential demand response and double-auction markets. In *IEEE Power and Energy Society General Meeting*, 2011.

[50] Matlab-mathworks. In *http://www.mathworks.com/products/matlab*.

[51] O. K. Singh. Performance enhancement of combined cycle power plant using inlet air cooling by exhaust heat operated ammonia-water absorption refrigeration system. In *Applied Energy*, pages 180, 867–879, 2016.

[52] Bargiotas D. Tsoukalas L. H. Alamaniotis, M. Towards smart energy systems: Application of kernel machine regression for medium term electricity load forecasting. In *Springer Plus*, pages 5(1), 1–15, 2016.

[53] Agarwal Alamaniotis, M. Fuzzy integration of support vector regression models for anticipatory control of complex energy systems. In *International Journal of Monitoring and Surveillance Technologies Research*, pages 2(2), 26–40, 2014.

[54] Gu J Chan, Y. Modeling of turbine cycles using a neuro-fuzzy based approach to predict turbine-generator output for nuclear power plants. In *Energies*, pages 5(12), 101–118, 2012.

[55] Kang D. W. Kim T. S. Lee, J. J. Development of a gas turbine performance analysis program and its application. In *Energy*, pages 36(8), 5274–5285, 2011.

[56] Al Zubaidy S. De, S. A. Gas turbine performance at varying ambient temperature. In *Applied Thermal Engineering*, pages 31(14–15), 2735–2739, 2011.

[57] Tüfekci P. Gürgen F. S. Kaya, H. Local and global learning methods for predicting power of a combined gas steam turbine. In *Proceedings of the International Conference on Emerging Trends in Computer and Electronics Engineering (ICETCEE)*, page pp. 13–18, 2011.

[58] P. Tüfekci. Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. In *Electrical Power and Energy Systems*, pages 60, 126–140, 2014.

[59] Kamal K. Zafar T. Sheikh Z. Shah A. Mathavan S. Rashid, M. Energy prediction of a combined cycle power plant using a particle swarm optimization trained feedforward neural network. In *Proceedings of the International Conference on Mechanical Engineering, Automation and Control Systems (MEACS)*, pages pp. 1–5, 2015.

[60] Venkatesh V. Mohanty, D. M. Performance analysis of a combined cycle gas turbine under varying operating conditions. In *Mechanical Engineering: An International Journal*, pages 1(2), 11–25, 2014.

[61] Pandit D. Fast M. Assadi M. De S. Smrekar, J. Prediction of power output of a coal-fired power plant by artificial neural network. In *Neural Computing Applicationsl*, pages 19(5), 725–740, 2010.

[62] Kaiadi M. Fast M. Assadi M. De, S. A. Development of an artificial neural network model for the steam process of a coal biomass cofired combined heat and power (chp) plant in sweden. In *Energyl*, pages 32(11), 2099–2109, 2007.

[63] Assadi M. Smrekar J. Fast, M. Application of artificial neural network to the condition monitoring and diagnosis of a chp plant. In *Proceedings of the 21th international conference on efficiency, cost, optimization, simulation and environmental impact of energy systems*, page pp 981–988, 2008.

[64] Assadi M. Deb S. Fast, M. Development and multi-utility of an ann model for an industrial gas turbine. In *Applied Energyl*, pages 86(1), 9–17, 2009.

[65] Reddy B. V. Gupta A. V. S. S. K. S. Srinivas, T. Thermal performance prediction of a biomass based integrated gasification combined cycle plant. In *Journal of Energy Resources Technology*, pages 134(2), 1–9, 2012.

[66] Pretorius P. Sibanda, W. Artificial neural networks- a review of applications of neural networks in the modeling of hiv epidemic. In *International Journal of Computers and Applications*, pages 44(16), 1–9, 2012.

[67] Uci machine learning repository. In *Combined Cycle Power Plant Data Set. Retrieved from http://archive.ics.uci.edu/ml/datasets/Combined+Cycle+Power+Plant*, 1987.

[68] S. Bhattacharya. A linear regression model to detect user emotion for touch input interactive systems. In *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015.