



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

**Ευφυείς μέθοδοι πρόβλεψης ροής ηλεκτρικού  
φορτίου πάνω  
στην Ευρωπαϊκή αγορά FBMC**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**του**

**Γεωργίου Σ. Ιωάννου**

**Επιβλέπων :** Ανδρέας - Γεώργιος Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π.



Αθήνα, Σεπτέμβριος 2017  
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

**Ευφυείς μέθοδοι πρόβλεψης ροής ηλεκτρικού  
φορτίου πάνω  
στην Ευρωπαϊκή αγορά FBMC**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**ΤΟΥ**

**Γεωργίου Σ. Ιωάννου**

**Επιβλέπων :** Ανδρέας - Γεώργιος Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20η Σεπτεμβρίου 2017.

.....  
Σταφυλοπάτης Ανδρέας-Γεώργιος  
Καθηγητής ΕΜΠ

.....  
Στάμου Γεώργιος  
Επικουρος Καθηγητής ΕΜΠ

.....  
Τσανάκας Παναγιώτης  
Καθηγητής ΕΜΠ

Αθήνα, Σεπτέμβριος 2017

.....  
Γεώργιος Σ. Ιωάννου

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright ©Γεώργιος Ιωάννου, 2017.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



## Περίληψη

Η ηλεκτρική ενέργεια αποτελεί έναν από τους πιο σημαντικούς παράγοντες της σύγχρονης ζωής που επηρεάζει από απλές καθημερινές δραστηριότητες μέχρι τις παραγωγικές διαδικασίες ολόκληρων κρατών. Είναι λοιπόν αναγκαίο να βελτιστοποιήσουμε την παραγωγή και διανομή ηλεκτρικής ενέργειας όσο το δυνατό μπορούμε. Για να το καταφέρουμε θα πρέπει να προβλέπουμε διάφορες πλευρές τις, όπως την ζήτηση, την ροή, την τιμή της και πολλά άλλα που χρησιμεύουν στα εργοστάσια παραγωγής ενέργειας να παράγουν ενέργεια αποδοτικότερα. Στην παρούσα διπλωματική θα εξετάσουμε την περίπτωση του ευρωπαϊκού συστήματος FBMC και την πρόβλεψη της ροής του ηλεκτρικού ρεύματος ανάμεσα στις ευρωπαϊκές χώρες που εφαρμόζουν. Θα δοκιμάσουμε ευφυείς τεχνικές αλλά και κλασσικές στατιστικές τεχνικές ώστε να βρούμε ποιες μέθοδοι προβλέπουν καλύτερα την ροή με το λιγότερο δυνατό σφάλμα.

### Λέξεις Κλειδιά

πρόβλεψη, ηλεκτρικού φορτίο, day-ahead αγορά, FBMC, ευφυείς μέθοδοι, νευρωνικά δίκτυα, στατιστικές μέθοδοι, Ensco

# **Abstract**

Electrical energy is one of the most important factors of modern lifestyle that influences from simple everyday activities to every complicated process of production of a country. Therefore, it is undeniably significant to improve the production and distribution of electrical energy to the best of our capability. In order to achieve that, we must predict many characteristics, such as the demand of electricity, the flow of the network, the price and many other that help the power plants work more efficiently. In the current thesis, i will examine the case of the European system FBMC and i will forecast the values of the flow of electrical charge between the European countries that implement this system. Various intelligent methods will be tested as well as classical statistical techniques, in order to find which models and methods perform the best and predicting with the least error possible.

## **Key Words**

prediction, electric charge, day-ahead market, FBMC, intelligent methods, neural networks, SVMs, statistical methods, Ensco

## Πρόλογος

Η παρούσα διπλωματική εκπονήθηκε κατά το ακαδημαϊκό έτος 2016-2017 στο Εθνικό Μετσόβιο Πολυτεχνείο στο εργαστήριο των Ευφών Υπολογιστικών Συστημάτων, κάτω από την επίβλεψη του καθηγητή Ανδρέα Σταφυλοπάτη. Το αντικείμενο που πραγματεύεται είναι η πρόβλεψη της ροής ηλεκτρικού φορτίου ανάμεσα σε ευρωπαϊκές χώρες που συμμετέχουν στο σχέδιο FBMC. Στο σημείο αυτό, προτού ξεκινήσει η παρουσίαση της διπλωματικής, θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν κατά την διάρκεια της εκπόνησης της διπλωματικής. Συγκεκριμένα, θα ήθελα να ευχαριστήσω τον ερευνητή κ. Γεώργιο Σιόλα για την υποστήριξή του, το ενδιαφέρον του και τις συμβουλές του σε όλη τη διάρκεια της συνεργασίας μας. Τέλος, πρέπει να δώσω ένα μεγάλο ευχαριστώ στον καθηγητή μου Ανδρέα Σταφυλοπάτη που μου έδωσε την ευκαιρία να ασχοληθώ με αυτή την εργασία και την καθοδήγησή του.

Αθήνα, Σεπτέμβριος 2017

Ιωάννου Γεώργιος

# Περιεχόμενα

<b>Περίληψη</b>	<b>5</b>
<b>Πρόλογος</b>	<b>7</b>
<b>Κεφάλαιο 1 - Εισαγωγή</b>	<b>10</b>
<b>Κεφάλαιο 2 - Αγορά και Ορίζοντες</b>	<b>15</b>
2.1 Day ahead Αγορά	15
2.2 Ορίζοντες πρόβλεψης και σφάλματα	16
<b>Κεφάλαιο 3 - Σύνοψη μεθόδων πρόβλεψης</b>	<b>20</b>
<b>Κεφάλαιο 4 - Στατιστικά Μοντέλα</b>	<b>23</b>
4.1 Εισαγωγή στα στατιστικά μοντέλα	23
4.2 Μοντέλο Παλινδρόμησης	25
4.3 Αυτοπαλινδρομικό μοντέλο (AR)	27
4.4 Μοντέλο Κινούμενου μέσου (MA)	28
4.5 Μοντέλα ARMA, ARIMA και παραλλαγές	30
4.6 Μοντέλα ARCH και παραλλαγές	34
<b>Κεφάλαιο 5 - Ευφυείς Τεχνικές</b>	<b>38</b>
5.1 Εισαγωγή στις ευφυείς τεχνικές	38
5.2 Τεχνητά Νευρωνικά Δίκτυα	42
5.2.1 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης	44
5.2.2 Αναδρομικά Νευρωνικά Δίκτυα	53
5.3 Support Vector Machines	55
<b>Κεφάλαιο 6 - Δεδομένα</b>	<b>57</b>



<b>Κεφάλαιο 7 - Πειραματικά Αποτελέσματα</b>	<b>61</b>
7.1 Στατιστικές Τεχνικές	61
7.1.1 Απλή Γραμμική Παλινδρόμηση	61
7.1.2 Μοντέλο Arima	72
7.1.3 Μοντέλο GARCH	75
7.1.4 Μοντέλο EGARCH	77
7.1.5 Μοντέλο GJR	78
7.2 Ευφυείς Τεχνικές	81
7.2.1 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης	81
7.2.2 Αναδρομικά Νευρωνικά Δίκτυα	96
7.3 Support Vector Machines	98
7.4 SVRArima	101
<b>Συμπεράσματα</b>	<b>104</b>
1. Συμπεράσματα	104
2. Μελλοντικές Επεκτάσεις	109
<b>Βιβλιογραφία</b>	<b>110</b>

## Κεφάλαιο 1ο - Εισαγωγή

Από την στιγμή που ανακαλύφθηκε ο ηλεκτρισμός και εφευρέθηκαν τρόποι με τους οποίους θα μπορούσε ο άνθρωπος να τον χρησιμοποιήσει για το όφελος του, η παραγωγή και η διανομή ηλεκτρικής ενέργειας έχουν γίνει αναπόσπαστο κομμάτι της σύγχρονης ζωής μας. Η ραγδαία ανάπτυξη της ηλεκτρικής τεχνολογίας μεταμόρφωσε την κοινωνία και τη βιομηχανία. Κάθε σπίτι, κάθε επιχείρηση και κάθε δραστηριότητα εξαρτάται από την χρήση του ηλεκτρισμού και χωρίς αυτόν πολλά από αυτά αδυνατούν να λειτουργήσουν. Δραστηριότητες, όπως μεταφορές, θέρμανση, φωτισμός, επικοινωνίες και υπολογισμοί βασίζονται αποκλειστικά από την χρήση του ηλεκτρικού ρεύματος.

Η ηλεκτρική ενέργεια παράγεται σήμερα από μεγάλα εργοστάσια μέσω ηλεκτρομηχανικών γεννητριών καταναλώνοντας διάφορες μορφές ενέργειας, όπως ορυκτά καύσιμα, θερμότητα από πυρηνικές αντιδράσεις ή χρησιμοποιώντας την κινητική ενέργεια από τον αέρα ή το νερό.

Η παραγόμενη ενέργεια είναι δυνατό να αποθηκευτεί σε ειδικές μπαταρίες με περιορισμένη χωρητικότητα, οι οποίες μόλις χρησιμοποιηθούν θα πρέπει να αλλαχθούν ή να επαναφορτιστούν, το οποίο κοστίζει αρκετά. Επομένως για μεγάλου μεγέθους ζήτηση ηλεκτρικού φορτίου, η ενέργεια είναι αδύνατο να αποθηκευτεί και πρέπει να μεταφερθεί συνεχόμενα ύστερα από την παραγωγή της. Αυτό ωθεί τις εταιρίες παραγωγής ενέργειας να κάνουν ακριβείς προβλέψεις για την ζήτηση του ηλεκτρικού φορτίου που απαιτεί το δίκτυο.

Σκοπός αυτής της διπλωματικής είναι να ασχοληθεί με την πρόβλεψη της ζήτησης του φορτίου στο ευρωπαϊκό σύστημα FBMC στην κεντρική Ευρώπη.

Η πρόβλεψη της ζήτησης αλλά και της μετακίνησης ηλεκτρικού φορτίου ανήκει στον γενικό κλάδο του energy forecasting (πρόβλεψη ενέργειας) και αποτελεί ένα από τα

πιο δημοφιλή πεδία πρόγνωσης μαζί με πρόβλεψη αιολικής ενέργειας, ηλιακής ενέργειας και , τα τελευταία χρόνια, πυρηνικής ενέργειας.

Ο γενικός κλάδος του Forecasting ασχολείται με την πρόγνωση ή πρόβλεψη μελλοντικών πληροφοριών βασισμένα σε παρελθοντικά ή υπάρχοντα δεδομένα αναλύοντας τις διάφορες τάσεις. Σύνηθες παράδειγμα είναι η εκτίμηση κάποιας μεταβλητής σε μία συγκεκριμένη στιγμή στο μέλλον. Οι προβλέψεις αναφέρονται σε δεδομένα από time series data, από cross-sectional data και από longitudinal data. Time series (χρονικές ακολουθίες) είναι μία σειρά από διακριτά δεδομένα καταταγμένα σε χρονική σειρά, όπως το ύψος της παλίρροιας, η τιμή μίας μετοχής στο χρηματιστήριο, οι τιμές της θερμοκρασίας σε μία περιοχή και άλλα πολλά. Σε αυτήν την κατηγορία ανήκει το πρόβλημα που θα εξετάζουμε σε αυτή τη διπλωματική.

Το πρόβλημα της πρόβλεψης της ζήτησης του φορτίου της ηλεκτρικής ενέργειας αλλά και της τιμής της είναι ευρέως διαδεδομένο στην σημερινή κοινωνία μας. Από τις αρχές του 1990, με την όλο και μειούμενη ρύθμιση και έλεγχο της αγοράς και σε συνδυασμό με την εισαγωγή της ανταγωνιστικής ηλεκτρικής αγοράς, το τοπίο του παλαιού μονοπωλιακού καθεστώτος του τομέα της παραγωγής ενέργειας έχει αλλάξει δραστικά.

Η ηλεκτρική ενέργεια σε πολλές χώρες σήμερα ανταλλάσσεται με spot και derivative συμβόλαια μέσα σε ένα χρηματιστήριο. Όμως, λόγω της φύσεως της ηλεκτρικής ενέργειας, που είναι αδύνατο να αποθηκευτεί, υπάρχει η ανάγκη της διαρκούς εξισορρόπησης μεταξύ της παραγωγής και της κατανάλωσης. Το κόστος που προκύπτει από την υπερπαραγωγή ή την υποπαραγωγή ενέργειας είναι πολύ υψηλό για τις εταιρίες παραγωγής ενέργειας. Αυτός είναι ο λόγος που είναι απαραίτητες ακριβείς προβλέψεις στον αριθμό του φορτίου ζήτησης αλλά και στη τιμή του.

Στους τομείς παραγωγής και διανομής ηλεκτρικής ενέργειας, η πρόβλεψη του φορτίου παίζει έναν πολύ σημαντικό ρόλο στον σχεδιασμό επιχειρηματικό ενεργειών και στην εξάπλωση των εργοστασίων. Βοηθά τα εργοστάσια ηλεκτρικού ρεύματος να

πάρουν σημαντικές αποφάσεις όσον αφορά την παραγωγή, διανομή, διαμοιρασμό και αγορά ηλεκτρικής ενέργειας καθώς και τον οικονομικό προγραμματισμό και ανάπτυξη.

Συγκεκριμένα, το πρόβλημα που προσπαθούμε να ερευνήσουμε σε αυτή τη διπλωματική είναι το FBMC (Flow Based Market Coupling) στην CWE (Central West Europe). Το FBMC είναι ένα σύστημα που εφαρμόζεται στην κεντρική Ευρώπη από τα μέσα του 2015 σε μία προσπάθεια της Ευρωπαϊκής Ένωσης να βελτιστοποιήσει την οικονομική απόδοση αλλά και να δημιουργήσει μία πιο φιλελεύθερη αγορά ηλεκτρικής ενέργειας μεταξύ των ευρωπαϊκών χωρών.

Σήμερα, το FBMC καλύπτει πάνω από 18 χώρες της Ευρωπαϊκής Ένωσης, μεταξύ των οποίων είναι το Βέλγιο, η Δανία, η Εσθονία, η Φινλανδία, η Γαλλία, η Γερμανία, το Λουξεμβούργο, η Αυστρία, η Αγγλία, η Λετονία, η Λιθουανία, η Ολλανδία, η Νορβηγία, η Πορτογαλία, η Σλοβενία, η Ισπανία και Σουηδία.

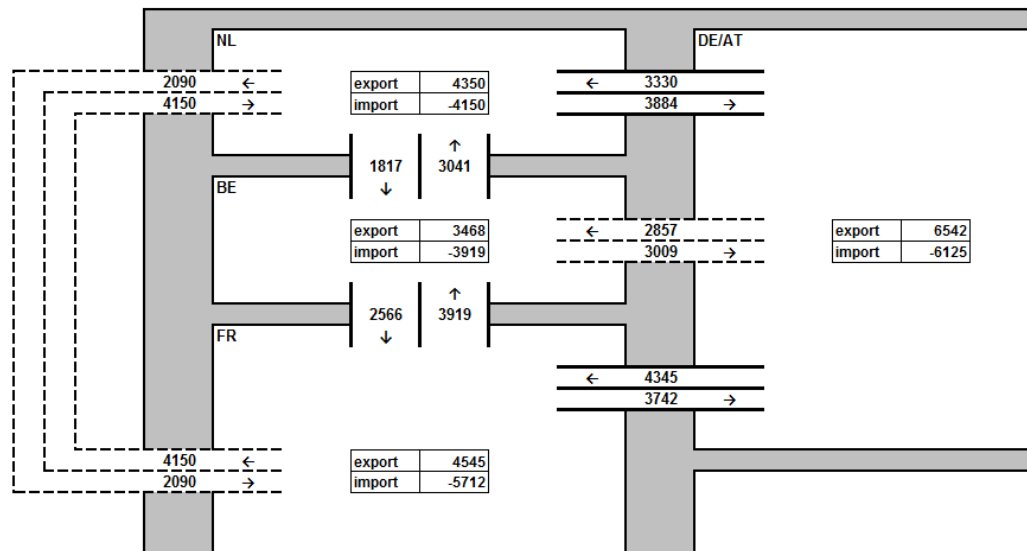
Στη συγκεκριμένη διπλωματική εργασία θα ασχοληθούμε κυρίως με 4 χώρες (ομάδες χωρών) που επικοινωνούν μέσω του FBMC και επηρεάζει η μία την άλλη καθημερινά στα ηλεκτρικά φορτία που ανταλλάσσονται.

Οι χώρες είναι η Γαλλία, η Γερμανία/Αυστρία, το Βέλγιο και η Ολλανδία.

Παρακάτω φαίνεται ένα παράδειγμα από ένα στιγμιότυπο ανταλλαγής ηλεκτρικού φορτίου ανάμεσα σε αυτές τις χώρες.

data shown for hour:  
1

## CWE max net positions and bilateral exchanges



Το FBMC βασίζεται σε δύο αρχές:

- 1) Δημιουργία τοπικής αγοράς μόνο με ενέργεια
- 2) Σύζευξη αγορών μεταξύ γειτονικών χωρών

Η σύζευξη που αναφέρεται παραπάνω συμβάλλει στην καλύτερη απόδοση αλλά, παράλληλα, δημιουργεί πολύπλοκες καταστάσεις. Κάθε ζώνη αγοράς έχει διαφορετική χωρητικότητα μετάδοσης, το οποίο πρέπει να ληφθεί υπόψη. Επίσης, επειδή το ρεύμα δεν μεταφέρεται από τον παραγωγό στον καταναλωτή άμεσα αλλά περνά μέσα από περίπλοκα μονοπάτια ενός μεγάλου δικτύου, δυσκολεύει την πρόβλεψη της ροής του ρεύματος ακόμη περισσότερο.

Την ίδια στιγμή η ζήτηση του ηλεκτρικού ρεύματος εξαρτάται από τις καιρικές συνθήκες, όπως θερμοκρασία, άνεμος, υγρασία, κλπ, καθώς και από την ένταση των καθημερινών δραστηριοτήτων και επιχειρήσεων. Έτσι, δημιουργούνται διάφορα peak κατά την διάρκεια της μέρας, όπως επίσης από βδομάδα σε βδομάδα ή ακόμα

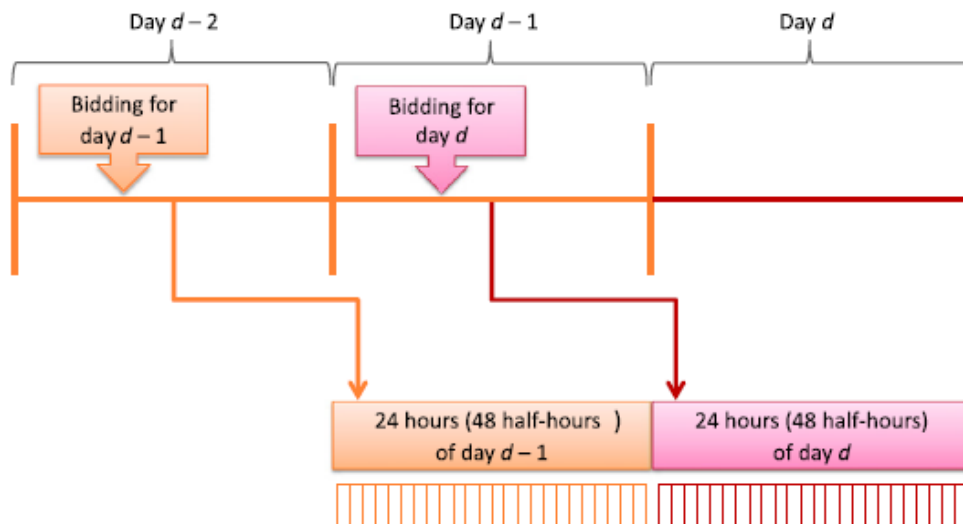
από εποχή σε εποχή, που δεν συναντώνται σε άλλες αγορές κάνοντας το έργο των προβλέψεων αρκετά δυσκολότερο. Από την άλλη αυτό έχει οδηγήσει σε μεγάλη ανάπτυξη και ενασχόληση με την εύρεση καλύτερων και πιο αποδοτικών μεθόδων πρόβλεψης.

## **Κεφάλαιο 2 - Αγορά και Ορίζοντες**

### **2.1 Day-ahead αγορά**

Η αγορά ηλεκτρικού ρεύματος δεν λειτουργεί σαν άλλες αγορές. Πρώτα από όλα απαγορεύει τις ανταλλαγές και είναι day-ahead αγορά, το οποίο σημαίνει ότι οι συναλλαγές μίας μέρας καθορίζονται την προηγούμενη. Αυτό συμβαίνει διότι οι διαχειριστές του συστήματος χρειάζονται ειδοποίηση εκ των προτέρων για το πρόγραμμα που θα εφαρμοστεί ώστε να εξετάσουν αν είναι εφικτό μέσα στο πλαίσιο των περιορισμών που υπάρχουν.

Στην day-ahead αγορά κάθε πράκτορας υποβάλλει τις προσφορές του για συγκεκριμένες ώρες τις επόμενης μέρας πριν την καθιερωμένη ώρα κλεισίματος της αγοράς. Έπειτα, οι τιμές της επόμενης καθορίζονται όλες μαζί με τα δεδομένα τα οποία είναι διαθέσιμα κατά την διάρκεια της προηγούμενης.



Στόχος της αγοράς είναι, λοιπόν, να ισορροπήσει την προσφορά και την ζήτηση, έτσι ώστε να μην μείνει υπόλοιπο ούτε στο ένα ούτε στο άλλο. Η τιμή στη οποία η προσφορά και η ζήτηση είναι ίσες ονομάζεται market-clearing price (MCP) και υπολογίζεται από την ένωση των καμπυλών της προσφοράς και της ζήτησης που διαμορφώνονται μετά από τις συναλλαγές.

Για να μπορούμε να προβλέψουμε με μεγάλη ακρίβεια την market-clearing τιμή, θα πρέπει να έχουμε προβλέψει πρωτύτερα την ζήτηση σε ηλεκτρική ενέργεια από τους καταναλωτές αλλά και πως θα διαμοιραστεί.

## 2.2 Ορίζοντες πρόβλεψης και σφάλματα

Η πρόβλεψη χωρίζεται σε τρεις υποκατηγορίες πρόβλεψης :

α) Βραχυπρόθεσμη (short term): Αποτελείται από προβλέψεις των επόμενων λεπτών ή λίγων ημερών και συναντάται συχνά στις day-to-day συναλλαγές που είναι τύπου χρηματιστηρίου.

β) Μεσοπρόθεσμη (medium term): Αποτελείται από προβλέψεις μερικών ημερών μέχρι λίγων μηνών και χρησιμοποιείται κυρίως σε υπολογισμούς ισολογισμών, διαχείριση κόστους και ρίσκου. Παρατηρείται στην πρόβλεψη της τιμής της ηλεκτρικής ενέργειας και υπολογίζει συνήθως κατανομές μεταβλητών πάνω σε κάποιο χρονικό διάστημα κάποιων μηνών.

γ) Μακροπρόθεσμη (long term): Αποτελείται από προβλέψεις σε βάθος χρόνου, από μήνες έως και μερικά χρόνια στο μέλλον. Συναντάται σε μελέτες σχεδιασμού και προγραμματισμού αλλά και σε αναλύσεις επικείμενων επενδύσεων.

Στην παρούσα διπλωματική θα επικεντρωθούμε στην βραχυπρόθεσμη πρόβλεψη λόγω της φύσης του προβλήματος που απαιτεί καθημερινές προβλέψεις. Βέβαια, από την άλλη είναι σημαντικό να προσθέσουμε ότι σε καμία περίπτωση οι άλλες υποκατηγορίες πρόβλεψης δεν έχουν και αυτές την χρησιμότητα τους στην περίπτωσή μας. Η μεσοπρόθεσμη πρόβλεψη μας βοηθά να κατανοήσουμε καλύτερα τις τάσεις μεταξύ διαφόρων μηνών και εποχών, ενώ από την άλλη η μακροπρόθεσμη πρόβλεψη μας δίνει πληροφορίες για διαφορετικές χρονιές. Συγκεντρώνοντας αυτά τα δεδομένα είναι πιο εύκολο να αποκρυπτογραφήσουμε την φαινομενικά χαώδη και τυχαία φύση των τιμών που προσπαθούμε να προβλέψουμε.

Έχοντας επισημάνει την ιδιαιτερότητα αλλά και υψηλή σημασία των προβλέψεων, είναι φυσικό να εξετάσουμε πως μπορούμε να αξιολογήσουμε τα αποτελέσματα αυτών. Η ακρίβεια των προβλέψεων μπορεί να αποφέρει μείωση κόστους, αποφυγή



σπαταλών και καλύτερη απόδοση στον διαμοιρασμό του ηλεκτρισμού. Επομένως, υπάρχουν και συνεχώς αναπτύσσονται νέες μέθοδοι αξιολόγησης των προβλέψεων. Παρακάτω θα αναφερθούν οι κυριότερες μετρικές που είναι ευρέως διαδεδομένες:

Απόλυτο Σφάλμα (Absolute Error):

$$AE_h = |P_h - P'_h|$$

όπου  $P_h$  είναι η αληθινή μέτρηση και  $P'_h$  και είναι η προβλεπόμενη

Το  $h$  συμβολίζει την χρονική περίοδο και στην προκειμένη περίπτωση αναπαριστά χρονικές ώρες σε μία μέρα.

Το MAE (Mean Absolute Error) είναι ο μέσος όρος των ωριαίων σφαλμάτων για μία μέρα ή μία εβδομάδα και είναι μία χρήσιμη μετρική. Όμως, έχει το μειονέκτημα επειδή σε διαφορετικά dataset είναι δύσκολο να συγκρίνουμε τα αποτελέσματα. Για αυτό σε πολλές μελέτες χρησιμοποιείται το:

Απόλυτο Ποσοστιαίο Σφάλμα (Absolute Percentage Error):

$$APE_h = AE_h / P_h$$

Βάσει αυτού μπορούμε να βρούμε το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα, υπολογίζοντας τον μέσο όρο μέσα σε μία συγκεκριμένη διάρκεια. Είναι από τις πιο διαδεδομένες μετρικές στον τομέα της πρόβλεψης φορτίου αλλά παρουσιάζει προβλήματα και ανακρίβειες σε προβλέψεις των οικονομικών τιμών του ηλεκτρισμού.

Ένα άλλο μέτρο σύγκρισης που έχει προταθεί σαν ένα εναλλακτικό των ποσοστιαίων σφαλμάτων είναι τα κλιμακωτά σφάλματα. Για non-seasonal χρονικές σειρές, τα κλιμακωτά σφάλματα υπολογίζονται χρησιμοποιώντας ένα-βήμα-εμπρός αφελείς προβλέψεις. Το πιο γνωστό από τα κλιμακωτά σφάλματα είναι το:

Μέσο Απόλυτο Κλιμακωτό Σφάλμα (Mean Absolute Scaled Error):

$$MASE_{T,m} = \frac{1}{T} \sum_{h=1}^T \frac{|P_h - P'_h|}{\frac{1}{T-m} \sum_{h=m+1}^T |P_h - P_{h-m}|}$$

όπου  $m$  είναι το μήκος του κύκλου.

Τα κλιμακωτά σφάλματα έχουν καλύτερη ακρίβεια σε διαφορετικά dataset και κλίμακες αλλά στον τομέα της ηλεκτρικής ενέργειας δεν έχουν χρησιμοποιηθεί ιδιαίτερα από μελέτες. Παρόλα αυτά έχουν προταθεί πολλές εναλλακτικές για την ομαλοποίηση του απόλυτου σφάλματος, με την πιο διαδεδομένη να είναι το σταθμισμένο απόλυτο σφάλμα:

Σταθμισμένο Απόλυτο Σφάλμα (WMAE):

$$DMAE_{T=24} = WMAE_{T=168} = \frac{1}{T} \sum_{h=1}^T \frac{|P_h - P'_h|}{\bar{P}_T}$$

$$\text{όπου } \bar{P}_T = \frac{1}{T} \sum_{h=1}^T P_h$$

με  $T = 24$  για το ημερήσιο (Daily MAE) και για  $T = 168$  για το εβδομαδιαίο (Weekly MAE).

Μία από τις πιο πολυχρησιμοποιημένες μετρικές, κυρίως σε θέματα οικονομικών αλλά και σε άλλους τομείς, είναι το root mean square error ή αλλιώς η ρίζα του μέσου τετραγωνικού σφάλματος. Παρατηρείται συχνά και σε μελέτες του προβλήματος που εξετάζουμε σε αυτή τη διπλωματική και υπολογίζεται ως εξής:

$$RMSE_{T=24 \text{ or } 168} = \sqrt{\frac{1}{T} \sum_{h=1}^T (P_h - P'_h)^2}$$

Στα επόμενα κεφάλαια θα διερευνήσουμε μοντέλα και τεχνικές πρόβλεψης και θα αξιολογήσουμε και θα συγκρίνουμε τα αποτελέσματα μεταξύ τους. Θα ασχοληθούμε με στατιστικά μοντέλα, με μοντέλα ευφών συστημάτων και τεχνητής νοημοσύνης και με υβριδικά μοντέλα που παράγονται από την μίξη των δύο προηγούμενων.

## **Κεφάλαιο 3 - Σύνοψη μεθόδων πρόβλεψης**

Σε αυτό το κεφάλαιο θα αναφέρουμε τις κατηγορίες των μεθοδολογιών για τη δημιουργία μοντέλων κατάλληλων για την πρόβλεψη των τιμών ροής ηλεκτρικού φορτίου, που έχουν χρησιμοποιηθεί ή χρησιμοποιούνται σε δημοσιεύσεις. Κάποιες από αυτές μπορεί να αποδίδουν καλύτερα από άλλες αλλά όλες έχουν την θέση τους στην προσέγγιση του προβλήματος. Κατηγοριοποιούμε τις μεθοδολογίες σε πέντε κατηγορίες: Πολυπρακτορικά μοντέλα(multi-agent models), Θεμελιώδεις κατασκευαστικές μέθοδοι(fundamental structural methods), Reduced-form μοντέλα, Στατιστικά μοντέλα, Υπολογιστικές ευφυείς τεχνικές.

**Πολυπρακτορικά συστήματα**(multi-agent systems) είναι υπολογιστικά συστήματα που αποτελούνται από πολλαπλούς ευφυείς πράκτορες που αλληλεπιδρούν μεταξύ τους. Χρησιμοποιούνται σε προβλήματα που είναι πολύ πολύπλοκα για έναν πράκτορα μόνο να τα επιλύσει. Συνήθως καλούνται να προβλέψουν μεσοπρόθεσμα ή μακροπρόθεσμα αποτελέσματα και δεν είναι φημισμένα σε προβλήματα πρόβλεψης ροής φορτίου. Οι περισσότερες εκδοχές βασίζονται στην αρχή του Equilibrium(ισορροπία) και αποδεικνύονται να είναι χρήσιμες στην πρόβλεψη τιμών στην αγορά αλλά δεν έχουν δοκιμαστεί σε δεδομένα με ροή ηλεκτρικού φορτίου. Μία άλλη τεχνική πολυπρακτορικών συστημάτων είναι η προσομοίωση βασισμένη σε πράκτορες. Αυτή η τεχνική έχει αρκετά λιγότερους περιορισμούς από την μέθοδο equilibrium γιατί δεν προσπαθεί να βρει μία ακριβή ισορροπία σε ένα σύστημα με πολλαπλές δύσκολες εξισώσεις, όπως συνηθίζεται σε δύσκολα προβλήματα.

Βασίζεται σε υπολογιστικές δομές και κανόνες που προσομοιώνουν τις ενέργειες και τις αλληλεπιδράσεις αυτόνομων πρακτόρων με σκοπό την πιο ρεαλιστική προσομοίωση και την εξέταση της επιρροής τους στο σύνολο του περιβάλλοντος που δραστηριοποιούνται.

Γενικότερα, τα μοντέλα που βασίζονται σε ευφυείς πράκτορες έχουν μεγάλη ευελιξία και χρησιμότητα, ειδικότερα σε προβλήματα που έχουν μεγάλη πολυπλοκότητα.

Μπορούμε να δούμε τις ακόλουθες μελέτες των Guerci, Rastegar & Cincotti 2010, ή των Kowalska-Pyzalska, Maciejowska, Suszczynski Sznajd-Weron, Weron 2014 όπως επίσης και των Sun & Tesfatsion 2007 για να κατανοήσουμε τις δυνατότητες

των ευφυών πρακτόρων. Όμως, διαθέτουν και κάποια μειονεκτήματα. Πρώτα από όλα για να γίνουν οι προσομοιώσεις των πρακτόρων χρειάζεται ένας μεγάλος όγκος υποθέσεων και πληροφοριών που πρέπει να βασίζεται σε πραγματικά ή θεωρητικά δεδομένα. Αυτό δεν είναι πάντα εφικτό ειδικά σε προβλήματα που είναι καινούργια στον χώρο της επιστήμης. Επίσης, τα πολυπρακτορικά συστήματα είναι πιο κατάλληλα να επιλύσουν ποιοτικά προβλήματα παρά ποσοτικά. Αυτό σημαίνει ότι η ακρίβεια των προβλέψεων δεν θα είναι ικανοποιητική στη παρούσα περίπτωσή μας αλλά θα μπορούν να δώσουν ίσως ποιοτικές πληροφορίες για την κατεύθυνση της ροής και για τυχόν ξεσπάσματα μέσα σε ένα χρονικό διάστημα.

**Θεμελιώδεις κατασκευαστικές μέθοδοι** είναι μέθοδοι που βασίζονται στις βασικές φυσικές σχέσεις και νόμους που υπάρχουν στην αγορά ηλεκτρικής ενέργειας. Αυτό περιλαμβάνει όλες τις σχέσεις μεταξύ παραγωγής ηλεκτρικής ενέργειας, της ζήτησής της αλλά και της διανομής της. Εκτός αυτών υπάρχουν και άλλοι θεμελιώδεις παράγοντες που συμβάλλουν στην δημιουργία αυτών των σχέσεων. Τέτοιοι είναι οι παράμετροι των συστημάτων, οι καιρικές συνθήκες, όπως για παράδειγμα η συχνότητα βροχόπτωσης, η ταχύτητα του ανέμου, η θερμοκρασία και άλλα. Κάποιες από τις μελέτες που εφαρμόζουν τέτοιες τεχνικές είναι των Gonzales, Contreras and Bunn 2012 ή των Karakatsani and Bunn 2008 ή του Kristiansen 2012. Υπάρχουν δύο γενικές κατηγορίες που χωρίζονται τα θεμελιώδη μοντέλα, τα πλούσια σε παραμέτρους θεμελιώδη μοντέλα (parameter-rich fundamental models) και τα φειδωλά κατασκευαστικά μοντέλα (parsimonious structural models). Τα πρώτα απαιτούν μία πληθώρα πληροφοριών από μεταβλητές που επηρεάζουν άμεσα την τιμή που προσπαθούμε να προβλέψουμε μέχρι εξωτερικές παραμέτρους που μπορεί να επηρεάζουν έμμεσα. Οι Vahvilainen and Pyykkonen έφτιαξαν ένα τέτοιο μοντέλο το 2005 για την σκανδιναβική αγορά. Η δεύτερη κατηγορία από την άλλη δε χρειάζεται τόσες πολλές μεταβλητές και είναι αρκετά πιο απλή. Χρησιμοποιεί μόνο τις παραμέτρους που είναι απόλυτα απαραίτητες χωρίς να υπερφορτώνει το μοντέλο με περισσές λεπτομέρειες. Αυτό αποτελεί μεγάλο πλεονέκτημα διότι σε πολύπλοκα προβλήματα δεν υπάρχει η δυνατότητα για συλλογή μεγάλου όγκου πληροφοριών.

Γενικότερα, όμως, τα θεμελιώδη μοντέλα παρουσιάζουν δύο μεγάλα ζητήματα. Πρώτον και κυριότερο, όπως αναφέρθηκε προηγουμένως, απαιτούν μεγάλες

ποσότητες δεδομένων και αυτό σε νέα προβλήματα που δεν έχουν εξερευνηθεί αρκετά δεν είναι εφικτό. Το δεύτερο ζήτημα είναι η ενσωμάτωση των στοχαστικών διακυμάνσεων και ταλαντώσεων στις θεμελιώδεις μεταβλητές. Εκτός αυτών πολλές φορές πρέπει να έχουμε πάρει υπόψη μας διάφορες προϋποθέσεις που δεν είναι απαραίτητο να ισχύουν πάντα, βάζοντας έτσι περαιτέρω ρίσκο στην ακρίβεια της πρόβλεψης.

**Reduced-form μοντέλα** είναι μοντέλα που χρησιμοποιούνται κυρίως στην στατιστική και στην οικονομετρία. Η reduced μορφή ενός συστήματος εξισώσεων είναι όταν επιλύουμε τις εξισώσεις ως προς τις ενδογενείς μεταβλητές, δηλαδή τις εκφράζουμε μέσω των εξωγενών μεταβλητών εάν υπάρχουν. Τέτοιου είδους μοντέλα δεν είναι ικανά για ακριβείς ωριαίες προβλέψεις αλλά εκεί που διαπρέπουν είναι στο να ανακαλύπτουν τα χαρακτηριστικά της αγοράς της ηλεκτρικής ενέργειας, τις δυναμικές του περιβάλλοντος και είναι ιδανικά για ανάλυση ρίσκου. Υπάρχουν δύο γνωστές κατηγορίες reduced-form μοντέλων, τα jump-diffusion μοντέλα και τα Markov regime-switching(MRS) μοντέλα. Το πρώτο είναι ένα μικτό μοντέλο, που αποτελείται από μία συνάρτηση άλματος(στοχαστική συνάρτηση) και μία διαδικασία διάχυσης(μία Markov διαδικασία). Το δεύτερο αποτελείται από πολλές διαφορετικές στοχαστικές διαδικασίες που ονομάζονται regimes. Το MRS υπερτερεί του Jump-diffusion γιατί έχει την ικανότητα να παρουσιάζει διαδοχικά spikes πράγμα που το καθιστά καλύτερο στ πρόβλημα της πρόβλεψης της ροής ηλεκτρικού φορτίου. Παραδείγματα MRS μοντέλων είναι του Huisman 2009 ή των Janczura & Weron 2009.

Αυτές τις μεθοδολογίες μοντελοποίησης και πρόβλεψης, όμως, δεν θα τις χρησιμοποιήσουμε σε αυτή τη διπλωματική. Θα χρησιμοποιήσουμε και θα εξετάσουμε τις δύο τελευταίες, τα στατιστικά μοντέλα και τις υπολογιστικές ευφυείς τεχνικές στα επόμενα κεφάλαια.

## **Κεφάλαιο 4 - Στατιστικά Μοντέλα**

#### 4.1 Εισαγωγή στα στατιστικά μοντέλα

Στατιστικό μοντέλο είναι μία κλάση μαθηματικού μοντέλου που ενσαρκώνει κάποιες υποθέσεις που αφορούν κάποιο δείγμα δεδομένων από ένα μεγαλύτερο σύνολο. Ουσιαστικά αυτό που αντιπροσωπεύουν τα στατιστικά μοντέλα είναι μία διαδικασία δημιουργίας δεδομένων. Οι υποθέσεις αυτές προσπαθούν να περιγράψουν μία πιθανοτική κατανομή που προσεγγίζει την πραγματική κατανομή των δεδομένων που παρακολουθούμε. Ένα στατιστικό μοντέλο αποτελείται από μαθηματικές εξισώσεις μιας ή πολλών μεταβλητών, οι οποίες μεταβλητές επηρεάζουν σε μικρό ή μεγάλο βαθμό την ζητούμενη κατανομή των δεδομένων μας.

Τα στατιστικά μοντέλα είναι αρκετά χρήσιμα επειδή μπορούν να συσχετιστούν με κάποια φυσική ερμηνεία κάνοντας τα πιο εύκολα στην κατανόηση και στην χρήση τους. Πολλές φορές δέχονται κριτική ότι είναι υπεραπλουστευμένα και έχουν περιορισμένες δυνατότητες στα πιο εξελιγμένα μη-γραμμικά προβλήματα αλλά πολλά πειραματικά δεδομένα και πρακτικές εφαρμογές δείχνουν ότι έχουν καλή απόδοση και ακρίβεια ακόμα και σε σχέση με μη-γραμμικές λύσεις.

Παρακάτω θα παρουσιάσουμε διάφορα στατιστικά μοντέλα και τεχνικές που είναι γνωστές στον κλάδο του forecasting. Θα αναφερθούμε, αρχικά, σε τεχνικές που δεν χρησιμοποιήσα στο πρόβλημα που εξετάζει αυτή η διπλωματική αλλά είναι αξιόλογες και είναι απαραίτητο να αναφερθούν, και έπειτα θα παρουσιάσουμε τις τεχνικές που δοκίμασα και εφαρμόσαμε στην επίλυση του προβλήματος.

Μία αρκετά δημοφιλής στατιστική τεχνική είναι η similar-day (παρόμοια-ημέρα) μέθοδος. Σε αυτή την τεχνική παίρνουμε ιστορικά στοιχεία από προηγούμενες μέρες με παρόμοια χαρακτηριστικά με αυτά της μέρας που θέλουμε να προβλέψουμε. Βάσει αυτών των στοιχείων υποθέτουμε ότι η πρόγνωση του φορτίου θα κυμαίνεται στα ίδια ύψη και κάνουμε την αντίστοιχη πρόγνωση. Τα πιο συνηθισμένα χαρακτηριστικά που χρησιμοποιούνται και δίνουν ενδιαφέροντα αποτελέσματα είναι:

- Ίδια μέρα παλαιότερης εβδομάδας
- Ίδια μέρα παλαιότερης χρονιάς
- Παρόμοιου τύπου μέρα διακοπών
- Μέρες με παρόμοιες κλιματικές συνθήκες, όπως θερμοκρασία, ένταση ανέμου, υγρασία, κλπ.
- Κατανάλωση ενέργειας

Ένα από αυτά τα χαρακτηριστικά ή ένας συνδυασμός αυτών μπορεί να μας δώσει το καλύτερο ταίριασμα που αναζητούμε για να προβλέψουμε με όσο το δυνατόν καλύτερη ακρίβεια.

Μία διαφορετική τεχνική, αρκετά απλή, είναι η εκθετική εξομάλυνση (exponential smoothing). Είναι ευρέως διαδεδομένη στις προβλέψεις φορτίου. Η εκθετική εξομάλυνση προσπαθεί να εξομαλύνει δεδομένα χρονικών σειρών, κυρίως εφαρμόζοντας αναδρομικά βαθυπερατά φίλτρα με συνάρτηση εκθετικού παραθύρου. Αυτή η τεχνική εφαρμόζεται σε διάφορες περιπτώσεις λόγω της ευκολίας της στην χρησιμοποίηση σε πολλά δεδομένα.

Μειονέκτημα, όμως, παρατηρείται στην ακρίβεια της και δεν θεωρείται τόσο αξιόπιστη για κάθε περίπτωση. Κύριος σκοπός της είναι να εξομαλύνει τα δεδομένα με βάρη τα οποία ελαττώνονται εκθετικά με τον χρόνο.

Ένα παράδειγμα μίας απλής μορφής της εκθετικής εξομάλυνσης είναι το ακόλουθο:

Έστω με  $x_t$  παριστάνουμε δεδομένα της χρονικής ακολουθίας και με  $s_t$  παριστάνουμε το αποτέλεσμα της μεθόδου, δηλαδή την πρόβλεψη της επόμενης τιμής. Τότε έχουμε:

$$s_0 = x_0$$

$$s_t = \alpha x_t + (1 - \alpha)s_{t-1}, t > 0$$

όπου  $\alpha$  είναι ο παράγοντας εξομάλυνσης (smoothing factor) και κυμαίνεται από (0, 1)



Ο παράγοντας  $\alpha$  είναι κρίσιμος επειδή καθορίζει πόση εξομάλυνση θα υποστούν τα δεδομένα, με τιμές όσο πιο κοντά στο 0 όσο και περισσότερη.

Πολλές παραλλαγές αυτού του αλγορίθμου έχουν αναπτυχθεί για να προβλέπουν καλύτερα τιμές από δεδομένα που παρουσιάζουν τάσεις καθώς και εποχιακές αλλαγές. Τέτοιες μέθοδοι είναι οι διπλή και τριπλή εκθετική εξομάλυνση (double exponential smoothing και triple exponential smoothing), όπου χρησιμοποιείται το εκθετικό φίλτρο δύο ή τρεις φορές αναδρομικά.

Αξιόλογες αναφορές σε μεθόδους που βασίζονται υπάρχουν σε πολλά επιστημονικά άρθρα και σε διάφορους διαγωνισμούς δείχνουν ότι η μέθοδος είναι ικανή να δώσει ικανοποιητικά αποτελέσματα. Ενδεικτικά, να αναφέρουμε μία ενδιαφέρουσα παραλλαγή, την THETA μέθοδο, όπου απέδωσε αρκετά καλά σε πρόσφατους αγώνες πρόβλεψης.

## 4.2 Μοντέλο Παλινδρόμησης

Παλινδρόμηση (Regression) είναι μία από τις πιο διαδεδομένες στατιστικές τεχνικές, αν όχι η πιο διαδεδομένη, όπου σκοπός της είναι να εκτιμήσει τις μαθηματικές σχέσεις μεταξύ μεταβλητών. Συνήθως υπάρχει μία ζητούμενη εξαρτημένη μεταβλητή και διάφορες ανεξάρτητες που επηρεάζουν την πρώτη. Επομένως, αυτή η τεχνική μελετά πως συμπεριφέρεται η εξαρτημένη μεταβλητή όταν κάποιες ή όλες οι ανεξάρτητες μεταβλητές παρουσιάσουν αλλαγές.

Η ανάλυση παλινδρόμησης (regression analysis) χρησιμοποιείται ευρέως στους κλάδους της πρόβλεψης και πρόγνωσης και η χρήση της είναι παρόμοια με αυτής του κλάδου της μηχανικής μάθησης (machine learning).

Κάθε μοντέλο Παλινδρόμησης αποτελείται από τις ακόλουθες οντότητες:

- Άγνωστες παραμέτρους (παριστάνονται από το  $\beta$ )

- Ανεξάρτητες μεταβλητές (παριστάνονται από το  $\mathbf{X}$ )
- Εξαρτημένη μεταβλητή (παριστάνεται από το  $\mathbf{Y}$ )

Το μοντέλο παρουσιάζεται σε συνολική μορφή κάπως έτσι:

$$\mathbf{Y} \approx \mathbf{f}(\mathbf{X}, \boldsymbol{\beta})$$

Πολλές τεχνικές έχουν αναπτυχθεί για την υλοποίηση της Παλινδρόμησης. Η πιο συνηθισμένη μορφή είναι αυτή της γραμμικής παλινδρόμησης, όπου η εξαρτημένη μεταβλητή είναι ένας γραμμικός συνδυασμός των παραμέτρων.

Αν από την άλλη αναφερόμαστε στην απλή γραμμική παλινδρόμηση τότε και η σχέση μεταξύ της εξαρτημένης μεταβλητής και των ανεξαρτήτων είναι και αυτή γραμμική. Για παράδειγμα:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i$$

με  $p$  ανεξάρτητες μεταβλητές.

Η τεχνική της γραμμικής παλινδρόμησης βασίζεται στην μέθοδο των ελαχίστων τετραγώνων. Δηλαδή, το μοντέλο οριστικοποιείται όταν το άθροισμα των τετραγώνων των διαφορών μεταξύ πραγματικών και προβλεπόμενων τιμών ελαχιστοποιείται.

$$\min \left( \sum_{i=1}^n e_i^2 \right)$$

$$\text{με } e_i = y_i - y_i'$$

Εκτός από τις γραμμικές μεθόδους υπάρχουν πολλές άλλες με την κάθε μία να έχει την δική της χρήση σε διαφορετικούς κλάδους, όπως η παλινδρόμηση Poisson σε μετρικά δεδομένα, η λογιστική παλινδρόμηση σε δεδομένα που είναι κατηγοριοποιημένα - π.χ. όταν οι μεταβλητές είναι μεταξύ 0 και 1-, άλλες μη-γραμμικές τεχνικές αναπτυγμένες για την επίλυση περισσότερο περίπλοκων προβλημάτων.

Παρόλο τον μεγάλο αριθμό των εναλλακτικών μεθόδων της παλινδρόμησης είναι αξιοσημείωτο να ειπωθεί ότι τα μοντέλα της γραμμικής regression παραμένουν οι πιο δημοφιλείς προσεγγίσεις στο πρόβλημα της πρόβλεψης φορτίου αλλά και της πρόβλεψης της χρηματικής τιμής της ηλεκτρικής ενέργειας. Στις περισσότερες μελέτες η γραμμική παλινδρόμηση συνδυάζεται με πιο πολύπλοκες μεθόδους, κάποιες από αυτές θα αναφερθούν παρακάτω σε αυτή την διπλωματική. Επιπλέον, είναι δύσκολο σε πολλές περιπτώσεις να ξεχωρίσεις την απλή παλινδρόμηση με την αυτοπαλινδρόμηση ή αλλιώς autoregression τεχνική λόγω των ομοιοτήτων που τους διακρίνουν. Για την δεύτερη θα μιλήσουμε στο επόμενο υποκεφάλαιο.

### 4.3 Αυτοπαλινδρομικό μοντέλο (AR)

Αυτοπαλινδρομικό (Autoregressive) μοντέλο είναι μία αναπαράσταση μίας τυχαίας διεργασίας, η οποία περιγράφει συγκεκριμένες χρονικές σειρές που παρατηρούνται στη φύση, στην οικονομία και σε άλλους τομείς. Το μοντέλο αυτοπαλινδρόμησης προσπαθεί να παράγει μία έξοδο για μία μεταβλητή βασισμένη στις παλαιότερες τιμές της ίδιας μεταβλητής. Η έξοδος βασίζεται γραμμικά από τις προηγούμενες τιμές συν έναν όρο σφάλματος.

Για παράδειγμα έστω μία μεταβλητή  $y_t$  και  $y_{t-1}$  είναι η τιμή της στην ακριβώς προηγούμενη στιγμή της ακολουθίας. Τότε ένα γραμμικό αυτοπαλινδρομικό μοντέλο παριστάνεται ως:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t$$

Η τάξη του μοντέλου είναι ο αριθμός των αμέσως προηγούμενων τιμών που χρησιμοποιούνται για να προβλεφθεί η επόμενη τιμή. Στη προκειμένη περίπτωση είναι 1. Το μοντέλο αυτό ονομάζεται πρώτης-τάξης αυτοπαλινδρομικό μοντέλο και αναπαριστάται από τα αρχικά AR(1). Γενικότερα, για k-τάξης μοντέλα, δηλαδή χρησιμοποιώντας τις k προηγούμενες τιμές, το μοντέλο συμβολίζεται με AR(k).

Το αυτοπαλινδρομικό μοντέλο είναι από τα πιο διαδεδομένα στον χώρο των προβλέψεων ηλεκτρικού φορτίου αλλά όχι στην απλή του μορφή που παρουσιάστηκε παραπάνω. Σχεδόν πάντοτε συνδυάζεται μαζί με κάποια άλλη μέθοδο και τεχνική για να προσφέρει το καλύτερο αποτέλεσμα. Αυτές τις τεχνικές, οι οποίες μπορούν να σταθούν και μόνες τους σαν εργαλείο πρόγνωσης θα τις δούμε αναλυτικά στις επόμενες σελίδες και, ιδιαίτερα, στο πειραματικό στάδιο της διπλωματικής, όπου παίζουν ένα σημαντικό ρόλο στην προσέγγιση του προβλήματος.

#### 4.4 Μοντέλο Κινούμενου μέσου (MA)

Ένα ακόμη μοντέλο για επεξεργασία και πρόγνωση χρονικών ακολουθιών είναι το μοντέλο του κινούμενου μέσου (Moving Average model). Το μοντέλο βασίζεται στις παλαιές τιμές ενός στοχαστικού όρου και στον μέσο όρο τις πραγματικής ακολουθίας που εξετάζουμε. Για παράδειγμα μία απλή αναπαράσταση του μοντέλου με τάξη 1 είναι:

$$x_t = \mu + w_t + \theta_1 w_{t-1}$$

με τα  $w_t$  να ορίζονται ως τυχαίες ανεξάρτητες μεταβλητές όπου η κάθε μία να υπακούει στην κανονική κατανομή με μέσο όρο 0 και ίδια διακύμανση  $N(0, \sigma_w^2)$ . Οι όροι  $\theta_t$  είναι οι παράμετροι του μοντέλου και  $\mu$  είναι ο μέσος όρος της αρχικής χρονικής σειράς. Αυτό το μοντέλο συμβολίζεται ως MA(1).

Γενικότερα, ένα μοντέλο τάξης- $q$  συμβολίζεται με  $MA(q)$  και έχει την εξής μορφή:

$$x_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q}$$

Επομένως, το μοντέλο του κινητού μέσου αποτελεί ουσιαστικά μία γραμμική παλινδρόμηση του μέσου όρου της αρχικής χρονικής ακολουθίας μαζί με παρόντα και παρελθοντικά στοιχεία λευκού θορύβου ή τυχαίων σπασμωδικών διακυμάνσεων. Αυτές οι τυχαίες εκτινάξεις είναι ανεξάρτητες μεταξύ τους και όπως αναφέραμε πριν ακολουθούν την κανονική κατανομή.

Είναι αρκετά σημαντικό να αναφέρουμε ότι το μοντέλο κινούμενου μέσου είναι πάντα στάσιμο (stationary) σε αντίθεση με το προηγούμενο μοντέλο που αναφέραμε, το αυτοπαλινδρομικό. Σε αυτό το σημείο είναι αναγκαίο να αναφέρουμε και να επεξηγήσουμε τι σημαίνει ο όρος στάσιμο και πως εμπλέκεται στην πρόβλεψη χρονικών σειρών και στο πρόβλημα που εξετάζουμε.

Στα μαθηματικά και στην στατιστική μία στάσιμη διαδικασία είναι μία στοχαστική διαδικασία, της οποίας η κοινή πιθανοτική κατανομή παραμένει σταθερή καθώς μετακινείται μέσα στον χρόνο. Σαν αποτέλεσμα, όροι όπως η μέση τιμή ή η διακύμανση παραμένουν σταθερές με την πάροδο του χρόνου. Συγκεκριμένα ορίζεται μαθηματικώς ως εξής:

$$F_{X(t)}(x) = F_{X(t+\Delta)}(x), \quad \text{for all } t, t + \Delta \in J.$$

με  $F$  να είναι η κοινή κατανομή.

Αυτός ο ορισμός καλύπτει τις αυστηρά στάσιμες διαδικασίες αλλά υπάρχει και ένας ορισμός πιο ελαστικός, ο οποίος είναι περισσότερο πρακτικός και χρησιμοποιείται πολλές φορές αντί του κανονικού ορισμού. Μία τυχαία διαδικασία ονομάζεται αδύναμα στάσιμη (weak-sense stationary) ή ευρεία στάσιμη (wide-sense stationary)

όταν η μέση τιμή της και η συσχέτισή της δεν αλλάζει με την πάροδο του χρόνου. Συγκεκριμένα έχουμε:

$$\begin{aligned} 1. & E[X(t_1)] = E[X(t_2)], \\ 2. & E[X(t_1)X(t_2)] = E[X(t_1 + \Delta)X(t_2 + \Delta)]. \end{aligned}$$

Στην μελέτη χρονικών ακολουθιών είναι σύνηθες και αναγκαίο να εξετάζουμε αν η παρούσα σειρά είναι stationary ή όχι, διότι αυτό αλλάζει την συνολική προσέγγιση μας και περιορίζει τα μοντέλα που μπορούμε να ταιριάζουμε. Πολλά μοντέλα για να εφαρμοστούν απαιτούν η ακολουθία να είναι τουλάχιστον αδύναμα στάσιμη για να επεξεργαστεί.

Κύριος λόγος παραβίασης του stationarity είναι η ύπαρξη κάποιας τάσης στα δεδομένα ή η ύπαρξη μιας μοναδιαίας ρίζας στην χαρακτηριστική εξίσωση της διαδικασίας. Αυτό έχει οδηγήσει στην αναγκαιότητα μετατροπής των δεδομένων σε stationary. Τα δεδομένα τα οποία διαθέτουν μοναδιαία ρίζα θα πρέπει να παραγωγιστούν μία η περισσότερες φορές μέχρι να γίνουν stationary. Από την άλλη τα δεδομένα που παρουσιάζουν κάποια κλίση ή τάση πρέπει να επεξεργαστούν κατάλληλα ώστε να αφαιρεθεί η συγκεκριμένη τάση. Συνήθως αυτά τα trend είναι συναρτήσεις ως προς τον χρόνο και είναι εύκολες στην επεξεργασία και στην αφαίρεσή τους.

#### 4.5 Μοντέλα ARMA, ARIMA και παραλλαγές

Ένα υβριδικό μοντέλο των δύο παραπάνω μεθόδων, το οποίο απαιτεί το stationarity των δεδομένων είναι το μοντέλο ARMA (Auto Regressive-Moving Average). Το μοντέλο ARMA συνδυάζει δύο πολυώνυμα, το ένα είναι από το αυτοπαλινδρομικό μοντέλο και το άλλο πολυώνυμο από το μοντέλο κινούμενου μέσου. Επειδή κάθε ένα από τα δύο πολυώνυμα έχει την δική του τάξη, όπως αναφέραμε παραπάνω, έτσι το μοντέλο ARMA καθορίζεται από δύο αριθμούς, τις τάξεις των επιμέρους μοντέλων. Γενικότερα, το μοντέλο ARMA(p,q) συμβολίζεται ως εξής:

$$X_t = \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \dots + \beta_q Z_{t-q}$$

όπου  $X_t$  είναι οι όροι της χρονικής ακολουθίας την χρονική στιγμή  $t$ ,  
 $Z_t$  είναι οι ανεξάρτητες μεταβλητές κανονικής κατανομής και  $\alpha, \beta$  είναι οι  
παράμετροι του μοντέλου.

Εκφράζει, λοιπόν, την εκτιμώμενη τιμή τη χρονική στιγμή  $t$  σαν ένα άθροισμα από  $q$   
όρους, που παριστάνουν την τυχαία διακύμανση των  $q$  προηγούμενων περιόδων, συν  
ένα σταθμισμένο άθροισμα των  $p$  προηγούμενων τιμών της ακολουθίας, που  
υπολογίζει την παρούσα τιμή.

Το ARMA έχει επιτύχει σε σχέση με τα στοιχεία/μοντέλα που αποτελείται επειδή  
είναι ικανό να περιγράψει ένα μοντέλο με λιγότερους όρους συγκριτικά με όσους  
όρους τα AR και MA χρειάζονται από μόνα τους. Όμως, για να λειτουργήσει αυτού  
του είδους το μοντέλο, όπως έχει προαναφερθεί, θα πρέπει τα δεδομένα να είναι  
στάσιμα ή αδύναμα στάσιμα. Πρακτικά, είναι αρκετά δύσκολο να συναντάμε μόνο  
στάσιμα δεδομένα για αυτό υπάρχουν τεχνικές που αφαιρούν τις υπάρχοντες τάσεις  
και ετοιμάζουν τα δεδομένα για να επεξεργαστούν από το μοντέλο. Παρακάτω, θα  
μιλήσουμε για το μοντέλο ARIMA που υλοποιεί αυτή τη δουλειά.

Η πιο συνηθισμένη μέθοδος για την αφαίρεση των τάσεων είναι να ενσωματώσουμε  
στο μοντέλο ένα αρχικό στάδιο όπου παραγωγίζουμε τα δεδομένα μία, δύο ή τρεις  
φορές, ώστε να γίνουν στάσιμα ή σχεδόν στάσιμα (δηλαδή να μην παρουσιάζουν  
τάσεις ή κάποια περιοδικότητα). Έπειτα, εφαρμόζουμε κανονικά το μοντέλο ARMA,  
όπως περιγράψαμε από πάνω. Σαν αποτέλεσμα, δημιουργήθηκε το μοντέλο ARIMA  
(Auto Regressive Integrated Moving Average) και χαρακτηρίζεται από τρεις  
παραμέτρους  $p, d, q$ . Οι παράμετροι  $p$  και  $q$  αντιπροσωπεύουν τις παραμέτρους από τα  
μοντέλα AR και MA αντίστοιχα, ενώ η παράμετρος  $d$  αντιπροσωπεύει πόσες φορές  
παραγωγίστηκαν τα δεδομένα.

Ειδικότερα, το μοντέλο  $ARIMA(1,0,0)$  είναι ταυτόσημο με το μοντέλο  $AR(1)$  και ομοίως το  $ARIMA(0,0,1)$  είναι ίδιο με το μοντέλο  $MA(1)$ . Το  $ARIMA(1,0,1)$  είναι ίδιο με το  $ARMA(1,1)$ , ενώ το  $ARIMA(0,1,0)$  παριστάνει έναν τυχαίο περίπατο.

Τα μοντέλα τύπου  $AR$  στις σύγχρονες μελέτες έχουν αποδειχθεί ότι αποτελούν την σπονδυλική στήλη της έρευνας και πρόγνωσης των χρονικών σειρών στον τομέα της πρόβλεψης ηλεκτρικού φορτίου. Υπάρχουν πολλές εφαρμογές σε σύγχρονα paper που χρησιμοποιούν την μέθοδο  $ARIMA$  αλλά και καινούργιες παραλλαγές βασισμένες σε αυτή.

Μερικές από αυτές τις νέες τεχνικές είναι το  $SARIMA$ (seasonal  $ARIMA$ ) και οι μέθοδοι τύπου  $ARX$ .

Το  $SARIMA$  είναι ένα μοντέλο που εφαρμόζει το  $ARIMA$  ανά συγκεκριμένες χρονικές περιόδους. Αυτό είναι χρήσιμο επειδή πολλά δεδομένα παρουσιάζουν εποχικότητα, για παράδειγμα αν υπάρχουν διάφορες τάσεις σε συγκεκριμένους μήνες του χρόνου και οδηγούν σε ψηλές τιμές ενώ σε άλλους οδηγούν σε χαμηλές. Έτσι, χωρίζουμε τα δεδομένα σε  $S$  περιόδους και παίρνουμε ως παλαιότερες τιμές στο μοντέλο τιμές που είναι πολλαπλάσια του  $S$ . Για παράδειγμα αν παρατηρούμε εποχικότητα μηνιαία τότε το  $S$  είναι ίσο με 12, τότε για να προσδιορίσουμε το  $X_t$  θα χρησιμοποιήσουμε το  $X_{t-12}$  και άλλα πολλαπλάσια του 12.

Τα μοντέλα τύπου  $ARX$  καλύπτουν ένα από τα μειονεκτήματα που εμφάνιζαν τα μοντέλα  $AR$ ,  $ARMA$  και  $ARIMA$  και αυτό είναι ότι εξετάζουν μόνο την δεδομένη χρονική ακολουθία και δεν επεξεργάζονται περαιτέρω δεδομένα. Όμως, όπως έχει ήδη αναφερθεί προηγουμένως στην αγορά του ηλεκτρικού ρεύματος αλλά και στην πρόβλεψη του ηλεκτρικού φορτίου πολλοί παράγοντες παίζουν σημαντικό ρόλο. Παράγοντες όπως η κατάσταση άλλων αγορών, η κατάσταση των γεννητριών και της χωρητικότητάς τους, οι καιρικές συνθήκες και πολλές ακόμα εξωγενείς μεταβλητές αποτελούν βασική πηγή πληροφοριών για να βασιστεί οποιαδήποτε πρόγνωση που θα επιχειρήσουμε. Για αυτό τον λόγο δημιουργήθηκαν τα  $ARX$ ,  $ARMAX$ ,  $ARIMAX$  και  $SARIMAX$  μοντέλα που αντιστοιχούν στα προαναφερθέντα με την διαφορά ότι



χρησιμοποιούν και εξωγενείς παράγοντες (exogenous factors), το οποίο επεξηγεί την προσθήκη του γράμματος X.

Πολλές φορές, όμως, αυτά τα μοντέλα συγχέονται με μοντέλα παλινδρόμησης ειδικά όταν βρίσκονται στην πιο απλή μορφή τους ή όταν διαθέτουν πολλές μεταβλητές. Γενικότερα, είναι ασφαλές να ορίσουμε ότι τα μοντέλα ARX διαθέτουν μία πιο περίπλοκη αυτοπαλινδρομική μορφή, το οποίο τα ξεχωρίζει από μία παλινδρόμηση με παραπάνω μεταβλητές.

Για παράδειγμα, παρακάτω παραθέτω ένα απλό μοντέλο ARMAX με p lags, q error terms και  $X_t$  την εξωτερική μεταβλητή:

$$y_t = \beta x_t + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \dots - \theta_q z_{t-q} + z_t$$

Ένα άλλο χαρακτηριστικό των μοντέλων ARX είναι ότι προϋποθέτουν ομοσκεδαστικότητα των δεδομένων. Αυτό σημαίνει ότι η ακολουθία των μεταβλητών που εξετάζουμε παρουσιάζουν την ίδια πεπερασμένη διακύμανση, φαινόμενο γνωστό και ως ομογένεια της διακύμανσης. Το να προϋποθέτεις την ομοσκεδαστικότητα απλοποιεί σημαντικά τα μαθηματικά μοντέλα και υπολογισμούς που απαιτούνται για την δημιουργία και το ταίριασμα του μοντέλου αλλά σε πολλές περιπτώσεις αυτό δεν αποτελεί αρνητικό. Όμως, η εμπειρία δείχνει ότι σε διάφορα δεδομένα χρονοσειρών υπάρχουν πολύπλοκες συμπεριφορές και δυναμικές μεταξύ των μεταβλητών, συνήθως μη γραμμικές, έχοντας σαν αποτέλεσμα η διακύμανση να μην παραμένει σταθερή αλλά να εξαρτάται από το παρελθόν της σειράς. Έτσι, αυτού του είδους τα δεδομένα χαρακτηρίζονται από απότομα τυχαία ξεσπάσματα ή σοκ σαν αυτά που παρατηρούμε στο πρόβλημα της πρόβλεψης φορτίου. Το φαινόμενο ονομάζεται ετεροσκεδαστικότητα.

#### 4.6 Μοντέλα ARCH και παραλλαγές

Το 1982 βρέθηκε για πρώτη φορά ένα μοντέλο από τον Engle που επιτυχώς απευθύνθηκε στο πρόβλημα της ετεροσκεδαστικότητας, το Auto Regressive Conditional Heteroskedastic(ARCH). Το μοντέλο βασίζεται στην αναπαράσταση της διακύμανσης από μία αυτοπαλινδρομική διαδικασία και ειδικότερα ένα άθροισμα από το τετράγωνο σταθμισμένων παρατηρήσεων. Από τότε το μοντέλο ARCH έχει χρησιμοποιηθεί σε για την επεξεργασία χρονικών σειρών, κυρίως στον τομέα των οικονομικών.

Παρακάτω θα παρουσιάσουμε τη μορφή του μοντέλου ARCH αλλά και την μορφή της γενικευμένης μορφής του, η οποία λέγεται GARCH(Generalized ARCH):

Έστω, η σχέση των ανεξάρτητων και εξαρτημένων μεταβλητών μαζί με το σφάλμα για το ARCH(q), με q να είναι οι καθυστερήσεις των ARCH σφαλμάτων, είναι η εξής:

$$y_t = x_t' b + \epsilon_t$$

Το σφάλμα στο μοντέλο ARCH αποτελείται από δύο στοιχεία. Το πρώτο είναι μία τυχαία μεταβλητή που είναι λευκός θόρυβος, ενώ το δεύτερο στοιχείο είναι μία χρονικά εξαρτημένη τυπική απόκλιση. Ο μαθηματικός τύπος είναι ο εξής:

$$\epsilon_t = \sigma_t z_t$$

Η χρονική σειρά που απεικονίζει το  $\sigma_t^2$  είναι:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$$

με  $\alpha_0 > 0$  και  $\alpha_i \geq 0, i > 0$

Για το ταίριασμα του μοντέλου στα δεδομένα χρησιμοποιείται η απλή μέθοδος των ελαχίστων τετραγώνων ενώ για να εκτιμηθεί πόσες θα πρέπει να είναι οι

καθυστερήσεις των σφαλμάτων  $q$ , χρησιμοποιείται η μεθοδολογία του Lagrange multiplier test, που προτάθηκε από τον ίδιο τον Engle.

Από την άλλη το γενικευμένο μοντέλο GARCH( $p,q$ ) έχει κάποιες διαφορές σε σχέση με την απλή μορφή του, κυρίως στον ορισμό της τυπικής απόκλισης  $\sigma_t^2$ . Αποτελείται από δύο αθροίσματα με όρους από τυπικές αποκλίσεις παλαιότερων χρονικών στιγμών τάξης  $p$  και όρους σφαλμάτων τάξης  $q$ . Παρακάτω φαίνεται σε μορφή μαθηματικής σχέσης:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2$$

Το μοντέλο GARCH έχει αρκετές παραλλαγές και εναλλακτικές μορφές. Κάποιες από αυτές είναι το NGARCH, IGARCH, EGARCH, GARCH-M, QGARCH, GJR-GARCH, TGARCH, fGARCH, COGARCH, όπου το καθένα από αυτά έχει τις δικές του ιδιαιτερότητες. Σε αυτή την διπλωματική θα χρησιμοποιηθούν από αυτές μόνο η EGARCH και η GJR-GARCH, τις οποίες θα παρουσιάσουμε συνοπτικά παρακάτω.

Το EGARCH (Exponential GARCH) περιγράφηκε πρώτη φορά από τον Nelson το 1991 εκφράζεται από δύο αριθμούς ( $p,q$ ) σαν το GARCH. Η διαφορά τους βρίσκεται στον υπολογισμό της τυπικής απόκλισης  $\sigma_t^2$ , η οποία ορίζεται ως εξής:

$$\log \sigma_t^2 = \omega + \sum_{k=1}^q \beta_k g(Z_{t-k}) + \sum_{k=1}^p \alpha_k \log \sigma_{t-k}^2$$

με την συνάρτηση  $g(Z_t)$  να ορίζεται ως εξής:

$$g(Z_t) = \theta Z_t + \lambda(|Z_t| - E(|Z_t|))$$

όπου  $Z_t$  μπορεί να είναι μία μεταβλητή που ακολουθεί κανονική κατανομή ή κάποια γενικευμένη κατανομή σφάλματος (generalized error distribution).

Η συνάρτηση του  $g(Z_t)$  εμπλέκει στο μοντέλο το μέτρο και το πρόσημο του  $Z_t$ , έτσι ώστε να έχουν ξεχωριστά αποτελέσματα, το οποίο έχει παρατηρηθεί να είναι αρκετά

χρήσιμο σε προβλήματα πρόβλεψης τιμών, είτε χρηματικών, είτε άλλων ειδών, πράγμα που είναι ενθαρρυντικό να το έχουμε σαν μία διαφορετική προσέγγιση στο πρόβλημά μας.

Το επόμενο εναλλακτικό GARCH μοντέλο που χρησιμοποιήσα είναι το GJR-GARCH. Το GJR (τα αρχικά είναι από τα ονόματα αυτών που το βρήκαν Glosten-Jagannathan-Runkle) παρουσιάζει και αυτό την διαφορά του με το GARCH στον ορισμό του  $\sigma_t^2$ , το οποίο ορίζεται με τον ακόλουθο ενδιαφέρον τρόπο:

$$\sigma_t^2 = K + \delta \sigma_{t-1}^2 + \alpha \epsilon_{t-1}^2 + \phi \epsilon_{t-1}^2 I_{t-1}$$

με  $I_{t-1} = 0$  εάν  $\epsilon_{t-1} \geq 0$  και  $I_{t-1} = 1$  εάν  $\epsilon_{t-1} < 0$ .

Στο πέρασμα των χρόνων έχουν διατυπωθεί πολλά υβριδικά μοντέλα που έχουν σαν βάση το GARCH. Όταν το GARCH χρησιμοποιείται μόνο του, δεν έχει τα καλύτερα αποτελέσματα στον τομέα την πρόβλεψης φορτίου, λόγω του ότι η μοντελοποίηση της ετεροσκεδαστικότητας φυλάσσει πολλά ρίσκα. Τα δεδομένα που επεξεργαζόμαστε διαθέτουν απότομα ξεσπάσματα πράγμα που σημαίνει ότι σίγουρα επιδεικνύουν ετεροσκεδαστικότητα αλλά πιθανόν στην κλίμακα που εξετάζει το μοντέλο GARCH ίσως όχι. Παρόλα αυτά συνδυάζοντας το GARCH με τα προαναφερθέντα μοντέλα τύπου AR έχει παρατηρηθεί σημαντική βελτίωση στην ακρίβεια των προβλέψεων. Έχει παρατηρηθεί, για παράδειγμα, ότι ένα υβριδικό μοντέλο ARIMA-GARCH συμπεριφέρεται πιο αποδοτικά σε σχέση με μία απλή μορφή του ARIMA σε δεδομένα με παρόμοια συμπεριφορά με τα δικά μας.

Αυτές, λοιπόν, είναι μερικές από τις στατιστικές τεχνικές και μεθόδους που είναι αρκετά αξιόπιστες για την επίλυση του προβλήματος της πρόβλεψης του ηλεκτρικού φορτίου. Οι περισσότερες έχουν δοκιμαστεί σε πολλές μελέτες και έχουν περάσει πολλά τεστ με την πάροδο του χρόνου αποδεικνύοντας την αποδοτικότητά τους. Άλλες από την άλλη είναι καινούργιες τεχνικές και απαιτούν χρόνο και συνεχές τρέστινγκ για να αποφανθούμε για την αξία τους.

Γενικότερα, οι στατιστικές τεχνικές έχουν αποδείξει την χρησιμότητά τους στον τομέα της αγοράς ηλεκτρικού ρεύματος, μιας αγοράς που χαρακτηρίζεται από εποχικότητα, πράγμα που διαχειρίζονται καλά οι στατιστικές μέθοδοι. Οι στατιστικές μέθοδοι, όμως, προσπαθούν να μοντελοποιήσουν τα δεδομένα βάσει προτύπων που εμφανίζονται παρά να μετρήσουν και να μιμηθούν την ουσιαστική υπόσταση των δεδομένων. Αυτό, δυστυχώς, οδηγεί στο να αδυνατούν πολλές φορές να προβλέψουν ικανοποιητικά τιμές που παρουσιάζουν απότομες αιχμές και ξεσπάσματα.

Τέτοιες δυσκολίες ώθησαν τους επιστήμονες στην έρευνα νέων τρόπων αντιμετώπισης των προβλημάτων. Ένας από αυτούς τους τρόπους είναι οι ευφυείς τεχνικές και συστήματα στον γενικό κλάδο της μηχανικής μάθησης. Στο επόμενο κεφάλαιο θα παρατηρήσουμε γνωστές μεθόδους και θα αναφερθώ στις τεχνικές που χρησιμοποιήσαμε για να προσεγγίσω το πρόβλημα της παρούσας διπλωματικής.

## **Κεφάλαιο 5 - Ευφυείς Τεχνικές**

### **5.1 Εισαγωγή στις ευφυείς τεχνικές**

Σε αυτό το κεφάλαιο θα μιλήσουμε για τα ευφυή υπολογιστικά συστήματα και τεχνικές και πως με την βοήθεια αυτών θα μπορούμε να δώσουμε μία πιο ικανοποιητική και αποδοτικότερη λύση στο πρόβλημα της πρόβλεψης ηλεκτρικού φορτίου. Η έκφραση ευφύες σύστημα ή υπολογιστική ευφυΐα διαφέρει στην σημασία της από άνθρωπο σε άνθρωπο. Συνήθως αναφέρεται στην ικανότητα του υπολογιστή να μαθαίνει κάποια συγκεκριμένη εργασία παρατηρώντας δεδομένα και εμπειρικά στοιχεία.

Γενικότερα, είναι εύκολο να σκεφτόμαστε την υπολογιστική ευφυΐα σαν ένα σύνολο από μεθοδολογίες, οι οποίες έχουν εμπνευστεί από διεργασίες της φύσης, και έχουν ως στόχο την επίλυση πραγματικών περίπλοκων προβλημάτων που οι καθιερωμένες μέθοδοι και τεχνικές αδυνατούν να λύσουν. Οι συνηθισμένοι λόγοι που παρουσιάζουν τέτοια δυσκολία είναι είτε εξαιτίας μεγάλης υπολογιστικής πολυπλοκότητας, είτε λόγω αβεβαιότητας διαφόρων δεδομένων ή απλά η διαδικασία είναι στοχαστική.

Πολλές τεχνικές και εφαρμογές υπολογιστικής ευφυΐας έχουν αναπτυχθεί τα τελευταία χρόνια. Αυτές οι μέθοδοι έχουν κατασκευαστεί με τέτοιο τρόπο ώστε να θυμίζει ανθρώπινη ευφυΐα στον τρόπο με τον οποίο προσπαθεί να επεξεργαστεί τα δεδομένα. Για παράδειγμα, όπως ένας άνθρωπος χρησιμοποιεί τις περιορισμένες γνώσεις και εμπειρίες του και προσπαθεί να φτάσει σε ένα συμπέρασμα, έτσι και ο υπολογιστής προσπαθεί με έναν συνεχές προσαρμοστικό τρόπο να μιμηθεί την λειτουργία του ανθρώπου. Υπάρχουν διάφορες τεχνικές μηχανικής ευφυΐας και σε αυτό το κεφάλαιο θα μιλήσουμε για τις πιο γνωστές και αυτές που χρησιμοποιήσαμε για την πρόβλεψη του ηλεκτρικού φορτίου στο πρόβλημα του FBMC.

Η υπολογιστική ευφυΐα μπορεί να χωριστεί σε κάποιους κύριους κλάδους, οι οποίοι είναι οι εξής:

## **Ασαφής λογική(Fuzzy logic)**

Η ασαφής λογική είναι μία μορφή λογικής που αποτελείται από πολλές τιμές ανάμεσα στο 0 και 1 σε αντίθεση με την Boolean λογική που αποτελείται αυστηρά από δύο τιμές. Έτσι, μπορούμε να αναπαριστούμε την μερική αλήθεια κάνοντας ευκολότερη την μοντελοποίηση σε πραγματικά περίπλοκα προβλήματα. Η ασαφής λογική είναι ιδιαίτερα χρήσιμη όταν θέλουμε να πάρουμε κάποια απόφαση κατά προσέγγιση χωρίς να έχουμε ακριβείς υπολογισμούς ή να γνωρίζουμε όλα τα δεδομένα που επηρεάζουν το πρόβλημα. Δεν έχει δυνατότητες μάθησης, δηλαδή να μάθει από παλαιότερα δεδομένα.

Αυτή η τεχνική βρίσκει εφαρμογή σε πολλούς κλάδους σήμερα, όπως στην επεξεργασία εικόνας, στη λήψη αποφάσεων, στην ιατρική διάγνωση, σε οικονομικές στρατηγικές και πολλά άλλα. Σε αυτή την διπλωματική δεν έχω εφαρμόσει πουθενά ασαφή λογική λόγω της περιορισμένης ακρίβειας της στον υπολογισμό τιμών αλλά και λόγω της φύσεως του προβλήματος της πρόβλεψης του ηλεκτρικού φορτίου. Παρόλα αυτά δεν σημαίνει ότι δεν είναι εφικτό να υλοποιηθεί και να έχει επιτυχία. Αρκετές μελέτες έχουν δείξει ότι το μοντέλο ασαφής λογικής ANFIS μαζί πολλές φορές με τον συνδυασμό του με άλλα μοντέλα έχει ενθαρρυντικά αποτελέσματα συγκρίσιμα ακόμα και με άλλες ευφυείς αλλά και στατιστικές μεθόδους. Αυτό δίνει στο μέλλον του κλάδου της πρόγνωσης νέους ορίζοντες για νέες μελέτες και βελτίωση των ήδη υπαρχόντων.

## **Νευρωνικά Δίκτυα**

Από την ονομασία και μόνο είναι αρκετό να καταλάβει κάποιος από που είναι εμπνευσμένος αυτός ο κλάδος των ευφύων συστημάτων. Η ανάπτυξη των τεχνητών νευρωνικών δικτύων (Artificial Neural Networks) έχει βασιστεί στην λειτουργία των βιολογικών νευρώνων και όπως αυτά έτσι και τα νευρωνικά δίκτυα δέχονται πληροφορίες, τις μεταφέρουν από νευρώνα σε νευρώνα, τις επεξεργάζονται και τέλος καταλήγουν με ένα αποτέλεσμα.

Επομένως, τα νευρωνικά συστήματα διαχειρίζονται διαμοιρασμένες πληροφορίες, τις επεξεργάζονται και μαθαίνουν από εμπειρικά δεδομένα. Στη διπλωματική θα

ασχοληθούμε ιδιαίτερα με τα νευρωνικά δίκτυα και με πολλά είδη τους και θα δούμε πως συμπεριφέρονται και ανταποκρίνονται στις ιδιαιτερότητες του προβλήματός μας. Παρακάτω θα παρουσιάσω αναλυτικότερα τα είδη που υπάρχουν, τις επιμέρους τεχνικές και τη χρησιμοποίησα.

### **Εξελικτικός Υπολογισμός**

Βασισμένος στην φυσική εξέλιξη του Δαρβίνου ο εξελικτικός υπολογισμός αποτελείται από μεθόδους που βασίζονται στην αναπροσαρμογή και στην εξέλιξη παρόμοια με αυτή των ζώα στην πάροδο του χρόνου. Συμπεριλαμβάνει εξελικτικούς αλγόριθμους και στρατηγική με τέτοιο τρόπο ώστε να επιλύει διάφορα προβλήματα, όπως προβλήματα βελτιστοποίησης, προβλήματα με πολλαπλούς στόχους, προβλήματα scheduling και άλλα.

### **Θεωρία της Μάθησης**

Η Θεωρία της μάθησης είναι ένας από τους σημαντικότερους κλάδους της υπολογιστικής ευφυΐας και ο κύριος στόχος της είναι η έρευνα του να βρεθεί τρόπος να έρθουμε όλο και πιο κοντά στην ανθρώπινη σκέψη και λογική συνοχή και αιτιολόγηση. Με βάση εμπειρίες, γνώσεις, συναισθηματικές και περιβαλλοντικές επιρροές μαθαίνουμε θεωρίες και τις αναπροσαρμόζουμε συνεχώς με την εμφάνιση νέων δεδομένων.

### **Support Vector Machines**

Στην μηχανική μάθηση τα support vector machines (η όπως θα ονομάζουμε από εδώ και πέρα σε αυτή την διπλωματική SVM) είναι μοντέλα επιβλεπόμενης μάθησης που



σχετίζονται με ιδιαίτερους αλγόριθμους μάθησης και χρησιμοποιούνται κυρίως σε προβλήματα ομαδοποίησης ή σε προβλήματα regression analysis(ανάλυση παλινδρόμησης). Σαν γενικό παράδειγμα, μοντελοποιούν τα δεδομένα σε σημεία στο επίπεδο ή στο χώρο και έπειτα βρίσκουν μία γραμμή, η οποία χωρίζει βέλτιστα τον χώρο σε δύο υποχώρους, οι οποίοι συμβολίζουν ξεχωριστές κατηγορίες. Έτσι, η διαδικασία εκπαίδευσης, ουσιαστικά, βρίσκει την βέλτιστη γραμμή για την καλύτερη ομαδοποίηση των δεδομένων. Όμως, έχει αναπτυχθεί το 1996 και μία τεχνική με SVM regression, που είναι κατάλληλο για να εφαρμοστεί στο δικό μας πρόβλημα. Αυτή την μέθοδο θα την δούμε αναλυτικότερα πιο μετά που την δοκιμάζουμε πειραματικά.

## **Πιθανοτικές Μέθοδοι**

Οι πιθανοτικές μέθοδοι στοχεύουν να επεξεργαστούν και να αξιολογήσουν τα αποτελέσματα των ευφυών υπολογιστικών συστημάτων χρησιμοποιώντας κυρίως τυχαιότητα. Πάνω σε αυτές βασίστηκε και η ανάπτυξη της ασαφής λογικής(fuzzy logic) και συνεχίζει να εξελίσσεται μέχρι σήμερα. Ουσιαστικά, προσφέρει ένα σύνολο πιθανών λύσεων στο κάθε πρόβλημα βασισμένες πάνω προηγούμενη γνώση και παρατήρηση.

Παρακάτω θα παραθέσουμε λεπτομέρειες για τις τεχνικές που χρησιμοποιήσα σε αυτήν τη διπλωματική, η οποίες ανήκουν όπως προανέφερα στον κλάδο των νευρωνικών δικτύων.

## **5.2 Τεχνητά Νευρωνικά Δίκτυα**

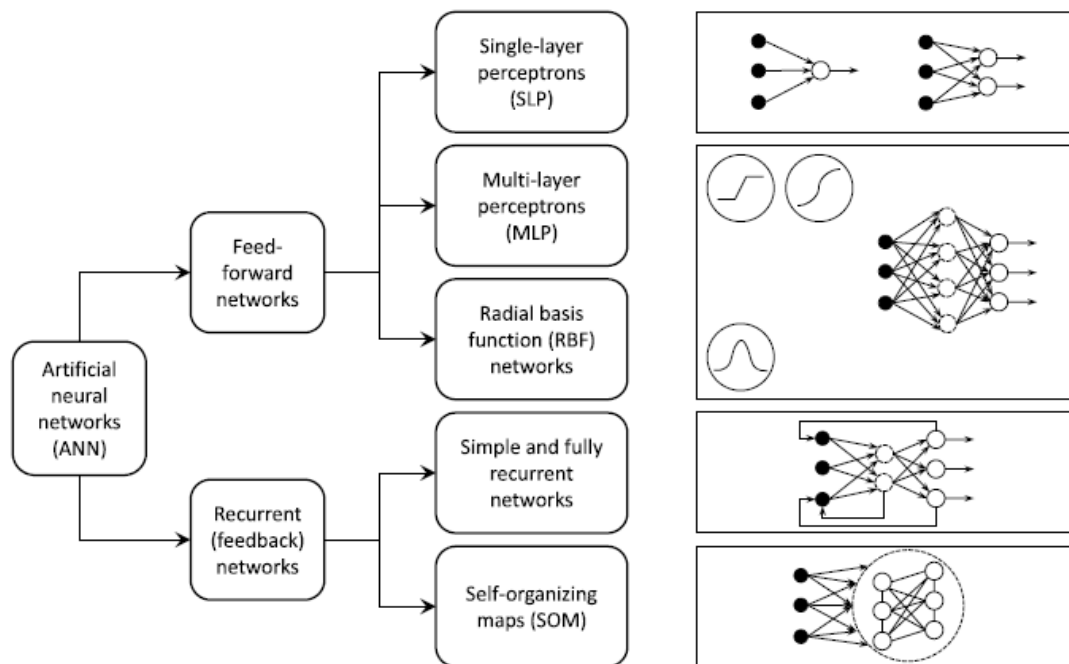
Τεχνητά Νευρωνικά Δίκτυα είναι ευφυή συστήματα εμπνευσμένα από τα βιολογικά νευρικά δίκτυα όπου είναι φτιαγμένα να εκτελούν περίπλοκες διαδικασίες μαθαίνοντας σταδιακά από παραδείγματα. Για παράδειγμα, συνήθης χρήση των

νευρωνικών είναι στην αναγνώριση εικόνας, όπου μπορούν να μάθουν να αναγνωρίζουν ένα αντικείμενο σε φωτογραφία αφού πάρουν ως δεδομένα πολλές φωτογραφίες και "μάθουν" ποιες από αυτές το περιέχουν το συγκεκριμένο αντικείμενο.

Ένα νευρωνικό σύστημα αποτελείται από μία συλλογή μονάδων που ονομάζονται νευρώνες, παρόμοιοι με τους βιολογικούς. Κάθε νευρώνας συνδέεται με άλλους με κατάλληλη σύνδεση που μεταφέρει τα δεδομένα. Όταν ένας νευρώνας λάβει data από την είσοδό του τότε τα επεξεργάζεται και με βάση παραμέτρων και μιας συνάρτησης βγάζει στην έξοδο το αποτέλεσμα. Γενικά, οι νευρώνες παρατάσσονται σε στρώματα και κάθε στρώμα ενώνεται με το προηγούμενο και το επόμενο. Υπάρχει το στρώμα εισόδου που μεταφέρει τα δεδομένα στο ενδιάμεσο ή ενδιάμεσα κρυμμένα στρώματα, που επεξεργάζονται τα δεδομένα περαιτέρω και τέλος το στρώμα εξόδου. Για να εκπαιδευτεί το κάθε νευρωνικό σύστημα ώστε να μπορεί να επιλύσει κάποιο συγκεκριμένο πρόβλημα που επιθυμούμε θα πρέπει να το εκπαιδεύσουμε με δεδομένα εκπαίδευσης μαζί με έναν αλγόριθμο μάθησης.

Κάθε τεχνητό νευρωνικό δίκτυο μπορεί να κατηγοριοποιηθεί με βάση την αρχιτεκτονική του και τον αλγόριθμο μάθησής του. Η αρχιτεκτονική ενός νευρωνικού δικτύου περιγράφει την φυσική του δομή. Συγκεκριμένα περιγράφει τον αριθμό των εισόδων, τον αριθμό των εξόδων και τον αριθμό των κρυμμένων επιπέδων. Σε κάθε επίπεδο διευκρινίζει τον αριθμό των νευρώνων που το αποτελούν και το είδος των νευρώνων, δηλαδή για παράδειγμα ποια είναι συνάρτηση μεταφοράς του κάθε νευρώνα. Συναρτήσεις μεταφοράς υπάρχουν πολλές, γραμμικές και μη-γραμμικές, μερικές από τις οποίες θα τις δοκιμάσουμε στο πρόβλημά μας. Επίσης, περιγράφει αναλυτικά τις συνδέσεις μεταξύ διαφορετικών επιπέδων αλλά και πολλές φορές συνδέσεις νευρώνων μεταξύ του ίδιου επιπέδου.

Τώρα θα ήταν καλή στιγμή να παραθέσουμε ένα σχήμα, το οποίο περιγράφει τα διαφορετικά είδη νευρωνικών δικτύων, τις κύριες κατηγορίες αυτών θα τις αναφέρουμε αναλυτικά και θα τις μελετήσουμε στην συνέχεια:



Δύο κύριες κατηγορίες υπάρχουν: τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (feed-forward) και τα αναδρομικά νευρωνικά δίκτυα(feedback). Η διαφορά που τα κάνει να ξεχωρίζουν το ένα από το άλλο είναι ότι τα πρόσθιας τροφοδότησης δεν έχουν loops, δηλαδή δεν υπάρχει κάποια σύνδεση μεταξύ εξόδου και εισόδου για ανατροφοδότηση, ενώ τα αναδρομικά τροφοδοτούν την έξοδο στην είσοδο για επεξεργασία. Πολλές μελέτες έχουν δείξει ότι, στατιστικά, τα πρόσθιας τροφοδότησης δείχνουν καλύτερα αποτελέσματα σε πρόγνωση τιμών, ενώ τα αναδρομικά εκδηλώνουν καλύτερη απόδοση σε προβλήματα κατηγοριοποίησης και ομαδοποίησης προτύπων.

### 5.2.1 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης

Το πιο απλό δίκτυο που θεωρείται νευρωνικό είναι το single-layer perceptron (perceptron ενός επιπέδου), το οποίο δεν διαθέτει κρυμμένα επίπεδα και είναι όμοιο με μία γραμμική παλινδρόμηση. Η έξοδος είναι δηλαδή ένας γραμμικός συνδυασμός από τις εισόδους και τα βάρη που αντιστοιχούν στο perceptron. Τα βάρη καθορίζονται από τα δεδομένα εκπαίδευσης και τον αλγόριθμο μάθησης που εφαρμόζεται στο νευρωνικό.

Αν πάρουμε αυτό το απλό μοντέλο και το εμπλουτίσουμε με επιπλέον επίπεδα με κρυμμένους κόμβους νευρώνων θα λάβουμε το μη-γραμμικό multi-layer perceptron (MLP ή αλλιώς το πολυεπίπεδο perceptron). Το MLP έχει αναδειχθεί ως το πιο γνωστό και πολυχρησιμοποιημένο από την οικογένεια των δικτύων πρόσθιας τροφοδότησης και δίκαια. Το MLP αποτελείται από επίπεδα, που επικοινωνούν μεταξύ τους, δηλαδή η έξοδος του ενός είναι η είσοδος του επόμενου.

Είναι σημαντικό να αναφέρουμε ότι το MLP διακρίνεται για την χρησιμότητά του στην αναγνώριση προτύπων, στην υπολογιστική νευροεπιστήμη και στην παράλληλη επεξεργασία. Στην έρευνα και στις μελέτες έχουν αποδειχθεί σημαντικό εργαλείο επειδή μπορούν να λύνουν ένα πρόβλημα στοχαστικά, πράγμα που αποτελεί μεγάλη υπόθεση όταν βρισκόμαστε αντιμέτωποι με αρκετά πολύπλοκα προβλήματα που η μόνη λύση που μπορούμε να βρούμε είναι μία εκτίμηση. Εκτός από ιδιαίτερα αποδοτικά σε regression ανάλυση, όπως έχουμε επαναφέρει προηγουμένως, τα MLP είναι αρκετά καλοί ομαδοποιητές. Γενικότερα, από το 1980 χρησιμοποιούνται ευρύτατα στον κλάδο της μηχανικής μάθησης με μόνο σκληρό αντίπαλο τα SVMs (support vector machines).

Τα δίκτυα πρόσθιας τροφοδότησης που χρησιμοποιήσαμε είναι ποικίλα και διαφέρουν στην αρχιτεκτονική τους όσο και στον αλγόριθμο μάθησης τους. Παρακάτω θα περιγράψουμε με συντομία τους αλγόριθμους μάθησης που χρησιμοποιήσαμε στο πειραματικό κομμάτι της διπλωματικής μου.

Βασική μέθοδος για την ανανέωση των βαρών είναι η οπισθοδιάδοση (backpropagation), η οποία λειτουργεί ως εξής:

Αφού προσδιοριστεί το σφάλμα για ένα δείγμα από τα δεδομένα εκπαίδευσης τότε υπολογίζουμε τα νέα βάρη του κρυμμένου επιπέδου με συγκεκριμένο μαθηματικό τύπο, ο οποίος διαφέρει από μέθοδο σε μέθοδο. Έπειτα ανανεώνονται τα βάρη της εισόδου και προχωρούμε με την ίδια διαδικασία στο επόμενο δείγμα των δεδομένων εκπαίδευσης μέχρι όλα να προβλέπονται σωστά ή κάποιος άλλος περιορισμός ικανοποιηθεί.

Αξίζει να σημειωθεί ότι όπου αναφερθεί το όνομα *perf* σαν μεταβλητή, παριστάνει το error, το οποίο στην προκειμένη περίπτωση είναι το mse, δηλαδή το τετράγωνο της διαφοράς της πραγματικής με την προβλεπόμενη τιμή. Επίσης, κάθε μία από τους παρακάτω αλγόριθμους μάθησης λειτουργεί σε όλα τα νευρωνικά δίκτυα, με την προϋπόθεση ότι οι συναρτήσεις μεταφοράς, τα βάρη και οι είσοδοι του δικτύου είναι παραγωγίσιμα.

#### Gradient Descent Backpropagation:

Η μέθοδος Gradient Descent Backpropagation είναι από τις πιο γνωστές που έχουν χρησιμοποιηθεί σε νευρωνικά δίκτυα και η συνάρτηση ανανέωσης των βαρών είναι η ακόλουθη:

$$W_{new} = W_{old} + lr * \frac{dperf}{dW}$$

με *lr* να είναι ο ρυθμός μάθησης (learning rate), δηλαδή πόσο γρήγορα γίνονται οι αλλαγές. Με μεγάλο ρυθμό μάθησης το σύστημα "μαθαίνει" γρήγορα και κάνει μεγάλα βήματα προς την σύγκλιση, πράγμα που φυλάσσει κινδύνους και κάνει το σύστημα ασταθές. Από την άλλη με μικρό ρυθμό ο αλγόριθμος κάνει μεγάλο διάστημα να συγκλίνει.

Ο αλγόριθμος σταματά όταν κάποιο από τα παρακάτω κριτήρια ικανοποιηθούν:

- Όταν επιτευχθεί ο μέγιστος αριθμός εποχών. (Εποχές είναι οι επαναλήψεις του αλγορίθμου).
- Όταν περάσει το μέγιστο χρονικό όριο.
- Όταν η απόδοση φτάσει συγκεκριμένη τιμή. (Απόδοση εννοούμε συνήθως την τιμή του σφάλματος).
- Όταν η κλίση του performance στον παραπάνω τύπο πέσει κάτω από κάποιο όριο.
- Όταν το Validation Performance αυξήθηκε περισσότερο από κάποιο όριο από την τελευταία φορά που μειώθηκε (Όταν χρησιμοποιηθεί Validation).

#### Gradient Descent with Momentum Backpropagation:

Το Gradient Descent with Momentum είναι μία εναλλακτική του απλού Gradient Descent όπου χρησιμοποιεί την έννοια του momentum, δηλαδή ορμής, για να ξεφεύγει από τοπικά ελάχιστα όταν εγκλωβίζεται σε αυτά. Ο τύπος για την ανανέωση των βαρών είναι λίγο διαφορετικός:

$$W_{new} = mc * dW_{prev} + W_{old} + lr * (1 - mc) * \frac{dperf}{dW}$$

με mc να είναι το momentum που παίρνει τιμές από 0 έως 1.

Ο αλγόριθμος σταματά στα ίδια κριτήρια με τον Gradient Descent.

#### Variable Learning Rate Gradient Descent:

Αυτή είναι μία παραλλαγή του προηγούμενου αλγορίθμου, μόνο που εδώ αλλάζει ο ρυθμός μάθησης. Στους προηγούμενους αλγορίθμους ο ρυθμός κρατιόταν σταθερός πράγμα που αποδεικνύεται επικίνδυνο. Το σύστημα έτσι μπορεί να μάθαινε πολύ γρήγορα και να γινόταν ασταθές ή να μάθαινε πολύ αργά και να καθυστερούσε να τελειώσει λόγω του ότι διαφορετικά δεδομένα εκπαίδευσης χρειάζονται και διαφορετικό ρυθμό μάθησης. Αναπτύχθηκε για αυτό τον λόγο αυτός ο αλγόριθμος

που αναπροσαρμόζει το learning rate. Για κάθε εποχή, αν η απόδοση μειωθεί προς τον στόχο τότε το learning rate αυξάνεται κατά ένα σταθερό ποσό. Αν η απόδοση μεγαλώσει (δηλαδή μεγαλώσει το error), τότε ο ρυθμός μάθησης μικραίνει και η αλλαγή που επηρέασε την απόδοση δε γίνεται. Ο τύπος είναι ο εξής:

$$W_{new} = mc * dW_{prev} + W_{old} + lr * mc * \frac{dperf}{dW}$$

Ο αλγόριθμος σταματά στα ίδια κριτήρια με πριν.

#### Resilient Backpropagation:

Πολλά νευρωνικά δίκτυα έχουν σαν συναρτήσεις μεταφοράς σιγμοειδείς συναρτήσεις, οι οποίες συμπιέζουν τις τιμές σε μία πεπερασμένη ζώνη. Έτσι, η κλίση αυτών των συναρτήσεων σε μεγάλες τιμές προσεγγίζει το 0. Όμως, όταν χρησιμοποιούμε steepest descent, δημιουργείται πρόβλημα στην ανανέωση των βαρών επειδή η αλλαγές στην κλίση είναι πολύ μικρές και τα βάρη είναι μακριά από τις βέλτιστες τιμές τους. Ο σκοπός αυτού του αλγόριθμου είναι να μειώσει όσο γίνεται αυτό το αρνητικό. Για αυτό, λοιπόν, στον τύπο της ανανέωσης των βαρών μόνο το πρόσημο της κλίσης της απόδοσης παίρνεται ως δεδομένο και όχι η τιμή της. Η τιμή των βαρών επηρεάζεται από μία άλλη μεταβλητή deltaX. Η μεταβλητή deltaX αυξάνεται από έναν σταθερό όρο όταν η κλίση έχει το ίδιο πρόσημο δύο φορές συνεχόμενα, ενώ μειώνεται από ένα σταθερό όρο όταν το πρόσημο της κλίσης αλλάξει. Όταν η κλίση είναι 0, η μεταβλητή δεν αλλάζει. Ο τύπος του αλγορίθμου παρουσιάζεται παρακάτω:

$$W_{new} = W_{old} + deltaX * sign\left(\frac{dperf}{dW}\right)$$

Ο αλγόριθμος σταματά στα προαναφερθέντα κριτήρια.

#### Levenberg-Marquardt Backpropagation:

Η μέθοδος Levenberg-Marquardt έχει σχεδιαστεί για να προσεγγίζει δεύτερης τάξης ταχύτητα εκπαίδευσης χωρίς να υπολογίζει τον πίνακα Hessian. Πρώτα, υπολογίζουμε τον Ιακωβιανό πίνακα  $J$  που αποτελείται από τις παραγώγους ως προς κάθε βάρος και παράμετρο. Έτσι, μπορούμε να εκτιμήσουμε τον Hessian πίνακα κάνοντας μόνο Backpropagation που είναι λιγότερο περίπλοκος. Ο τύπος ανανέωσης των βαρών είναι ο ακόλουθος:

$$W_{new} = W_{old} - [J^T J + \mu I]^{-1} * J^T e$$

όπου η προσέγγιση του πίνακα Hessian είναι

$$H = J^T J$$

και η κλίση υπολογίζεται από τον όρο

$$J^T e$$

$\mu$  είναι μία παράμετρος όπου όταν είναι μεγάλη ο αλγόριθμος γίνεται σαν gradient descent με μικρό βήμα. Το  $\mu$  μικραίνει με κάθε επιτυχημένο βήμα(δηλαδή όταν μικραίνει το σφάλμα) και αυξάνεται όταν κάποιο βήμα θα αύξανε το error. Με αυτόν τον τρόπο πάντα η performance συνάρτηση πάντα μειώνεται σε κάθε επανάληψη του αλγορίθμου. Αυτός ο αλγόριθμος μοιάζει αρκετά με την μέθοδο Newton και αν το  $\mu$  είναι 0 τότε είναι ταυτόσημη.

Ο αλγόριθμος σταματά όταν κάποιο από τα κριτήρια της gradient descent ικανοποιηθούν ή το  $\mu$  περάσει κάποια max καθορισμένη τιμή.



### Bayesian Regularization Backpropagation:

Η μέθοδος Bayesian Regularization χρησιμοποιεί τον αλγόριθμο Levenberg-Marquardt. Παίρνει το γραμμικό άθροισμα των σφαλμάτων και των βαρών και το ελαχιστοποιεί με τέτοιο τρόπο ώστε στο τέλος της εκπαίδευσης το δίκτυο να είναι καλά γενικευμένο. Οι τύποι υπολογισμού των βαρών είναι ίδιοι χρησιμοποιώντας τον Ιακωβιανό πίνακα και ο αλγόριθμος σταματά στα ίδια κριτήρια με τον αλγόριθμο Levenberg-Marquardt.

### BFGS Quasi-Newton Backpropagation:

Η μέθοδος του Newton είναι μία εναλλακτική στις μεθόδους conjugate gradient και πολλές φορές συγκλίνει γρηγορότερα από αυτές. Χρησιμοποιεί τον Hessian πίνακα για να βρει το επόμενο βήμα ανανέωσης του αλγορίθμου. Όμως, αυτό κοστίζει πολύ στα feed-forward δίκτυα. Για αυτόν τον λόγο αναπτύχθηκε μία νέα τεχνική, ονομαζόμενη Quasi-Newton, που δεν απαιτεί τον υπολογισμό των παραγώγων δευτέρας τάξης και σύγχρονες μελέτες τον καθιστούν από τους πιο επιτυχημένους αλγορίθμους. Ο τύπος των βαρών είναι ο εξής:

$$W_{new} = W_{old} + a * deltaX$$

και

$$deltaX = -\frac{gX}{H}$$

όπου  $gX$  είναι η κλίση και  $H$  είναι η προσέγγιση του Hessian πίνακα. Η παράμετρος  $a$  χρησιμοποιείται για να ελαχιστοποιεί το performance κατά την κατεύθυνση της έρευνας.

Ο αλγόριθμος τερματίζει στα ίδια κριτήρια με την μέθοδο gradient descent.

### One Step Secant:

Ο αλγόριθμος One Step Secant είναι μία προσπάθεια να γεφυρωθεί η απόσταση μεταξύ των conjugate descent αλγορίθμων και του quasi-Newton. Αυτό αποτελεί σημαντική ανάγκη διότι ο BFGS αλγόριθμος απαιτεί περισσότερους υπολογισμούς και αποθηκευτικό χώρο από τους conjugate αλγορίθμους. Η τεχνική του one step secant είναι ότι δεν αποθηκεύει ολόκληρο τον Hessian πίνακα και υποθέτει ότι ο προηγούμενος Hessian είναι ο identity matrix (δηλαδή ο μοναδιαίος πίνακας). Επιπλέον, η νέα τεχνική έχει το πλεονέκτημα ότι δεν χρειάζεται να υπολογιστεί ο αντίστροφος πίνακας, που πάντα κοστίζει αρκετά σε χρόνο. Ο τύπος ανανέωσης των βαρών είναι ίδιος. Το μόνο που αλλάζει είναι πως υπολογίζουμε την κατεύθυνση αναζήτησης, δηλαδή την μεταβλητή  $\Delta X$ , η οποία υπολογίζεται ως εξής:

$$\Delta X = -gX + Ac * \Delta W_{prev} + Bc * \Delta gX_{prev}$$

με  $gX$  να είναι το gradient,  $\Delta W_{prev}$  να είναι η αλλαγή των βαρών στην προηγούμενη επανάληψη και  $\Delta gX_{prev}$  είναι η αλλαγή στην κλίση στην προηγούμενη επανάληψη. Οι όροι  $Ac$  και  $Bc$  αποτελούν παραμέτρους που ελέγχουν την συνεισφορά των παραπάνω όρων. Ο αλγόριθμος σταματά στα ίδια κριτήρια με την gradient descent.

### Conjugate Gradient Backpropagation with Fletcher-Reeves updates:

Η φιλοσοφία των Conjugate αλγορίθμων είναι να γίνεται μία έρευνα προς την κατεύθυνση της steepest descent για να βρεθεί πιο γρήγορα το ελάχιστο και το επιτυγχάνει καθιστώντας αυτή τη μέθοδο γρηγορότερη από learning rate backpropagations. Όμως, τα αποτελέσματα διαφέρουν από πρόβλημα σε πρόβλημα

για αυτό δεν είναι πάντα η καλύτερη λύση. Χρησιμοποιούν λίγη μνήμη πράγμα το οποίο τα κάνει να είναι ιδανικά για νευρωνικά δίκτυα με μεγάλο αριθμό βαρών. Κάθε βάρος αλλάζει σύμφωνα με τον ακόλουθο τύπο:

$$W_{new} = W_{old} + a * dX$$

με  $dX$  να είναι η κατεύθυνση που ψάχνουμε και με  $a$  να είναι μία παράμετρος που έχει επιλεγεί για μικρύνει την συνεισφορά της κατεύθυνσης. Η κατεύθυνση της έρευνας υπολογίζεται από τον τύπο:

$$dX_{new} = -grad + dX_{old} * Z$$

όπου η πρώτη κατεύθυνση είναι η αρνητική κλίση ( $-grad$ ). Η μεταβλητή  $Z$  υπολογίζεται με πολλούς διαφορετικούς τρόπους. Στην προκειμένη περίπτωση της Fletcher-Reeves παραλλαγής υπολογίζεται από τον τύπο:

$$Z = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}$$

που παριστάνει το τετράγωνο της παρούσας νόρμας της κλίσης προς το τετράγωνο της νόρμας της προηγούμενης κλίσης.

Ο αλγόριθμος σταματά στα ίδια κριτήρια με αυτά της απλής gradient descent.

#### Conjugate Gradient Backpropagation with Powell/Beale Restarts:

Η μέθοδος αυτή είναι σαν τις άλλες conjugate gradient μεθόδους με την διαφορά ότι εδώ χρησιμοποιείται ο αλγόριθμος Powell/Beale για την επαναφορά της κατεύθυνσης. Συνήθως, η κατεύθυνση επαναφέρεται στην αρχική της τιμή, δηλαδή στην αρνητική τιμή της κλίσης, μετά από έναν αριθμό επαναλήψεων όσο και ο

αριθμός των βαρών και παραμέτρων του συστήματος. Αυτό μπορεί να βελτιωθεί με αυτόν τον αλγόριθμο και να γίνεται επαναφορά όταν υπάρχει λίγη ορθογωνικότητα (δεν είναι conjugate πια) μεταξύ της παρούσας κλίσης και της προηγούμενης. Το κριτήριο που πρέπει να ικανοποιηθεί για να γίνει επαναφορά είναι:

$$|g_{k-1}^T g_k| \geq 0.2 \|g_k\|^2$$

Ο αλγόριθμος σταματά στα ίδια κριτήρια που σταματά και η απλή gradient descent.

#### Conjugate Gradient Backpropagation with Polak-Ribiere updates:

Ο αλγόριθμος Polak-Ribiere μοιάζει αρκετά με αυτόν των Fletcher και Reeves με την μόνη διαφορά να είναι ο τύπος υπολογισμού της παραμέτρου  $Z$ , η οποία χρησιμοποιείται για να βρεθεί η κατεύθυνση έρευνας. Ο νέος τύπος είναι:

$$Z = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}} = \frac{(g_k^T - g_{k-1}^T) g_k}{g_{k-1}^T g_{k-1}}$$

Η απόδοση αυτού του αλγόριθμου σε σχέση με τους άλλους conjugate gradient είναι δύσκολος να εκτιμηθεί. Ο αλγόριθμος σταματά όπως η απλή μέθοδος gradient descent.

#### Scaled Conjugate Gradient Backpropagation:

Ο αλγόριθμος ScaledCG είναι βασισμένος στους προηγούμενους Conjugate αλγορίθμους με μόνη διαφορά ότι δεν υπολογίζει μία γραμμική αναζήτηση σε κάθε επανάληψη του αλγορίθμου. Χρησιμοποιεί το model-trust region προσέγγιση που

υπάρχει στον Levenberg-Marquardt αλγόριθμο σε συνδυασμό με τις τεχνικές της conjugate gradient backpropagation.

### 5.2.2 Αναδρομικά Νευρωνικά Δίκτυα

Τα Αναδρομικά Νευρωνικά δίκτυα είναι η δεύτερη μεγάλη κατηγορία δικτύων μετά τα πρόσθια τροφοδότησης. Σε αντίθεση με αυτά, τα αναδρομικά διαθέτουν μνήμη και επεξεργάζονται μία σειρά από συνεχόμενες εισόδους. Όταν μία νέα είσοδος εισέλθει στο σύστημα, τότε υπολογίζονται οι εξόδοι και λόγω της ανάδρασης η είσοδος κάθε νευρώνα αλλάζει καθιστώντας μία νέα κατάσταση στο δίκτυο. Για αυτό τον λόγο τα αναδρομικά χαρακτηρίζονται και ως δυναμικά συστήματα.

Υπάρχουν διάφορες εκδοχές αναδρομικών δικτύων. Δύο από τις απλές μορφές είναι τα δίκτυα του Elman και του Jordan. Το νευρωνικό δίκτυο του Elman αποτελείται από 3 επίπεδα μαζί με την προσθήκη ενός συνόλου από 'context units'. Τα 'context units' συνδέονται με το μεσαίο κρυφό επίπεδο, έχουν σταθερά βάρη και δεν ανανεώνονται κατά την διάρκεια της εκπαίδευσης. Σαν αποτέλεσμα κάθε νευρώνας στο κρυφό επίπεδο επεξεργάζεται εξωτερικά σήματα αλλά και σήματα από την ανατροφοδότηση. Από την άλλη τα σήματα από το επίπεδο εξόδου δε συμμετέχουν στη διαδικασία ανατροφοδότησης. Στο άλλο αναδρομικό νευρωνικό, του Jordan, τα 'context units' παίρνουν σήμα από το επίπεδο εξόδου αντί από το κρυφό επίπεδο και, επίσης, έχουν αναδρομική σύνδεση και με τον εαυτό τους.

Μία πιο γενική κλάση είναι αυτή των ολικά επαναλαμβανόμενων (fully recurrent) δικτύων, που ονομάζονται αλλιώς και αναδρομικά πραγματικού χρόνου (RTRN). Σε αυτή την δομή, όπως υπαγορεύει και το όνομά της, η έξοδος κάθε νευρώνα είναι συνδεδεμένη σαν είσοδος σε όλους τους υπόλοιπους νευρώνες. Τέτοια νευρωνικά δίκτυα συνήθως εκπαιδεύονται με gradient αλγορίθμους, όπως και τα πρόσθια τροφοδότησης, αλλά με την διαφορά σε αυτήν την περίπτωση η διαδικασία είναι πιο πολύπλοκη.

Τα αναδρομικά νευρωνικά δίκτυα λόγω της φύσης τους είναι αρκετά αποδοτικά στο να λύνουν κάποιου είδους προβλήματα, όπως φωνητική αναγνώριση, αναγνώριση γραφής και διάφορα άλλα. Δυστυχώς, όμως τα αναδρομικά δίκτυα δεν έχουν ικανοποιητικά αποτελέσματα σε προβλήματα πρόβλεψης χρονοσειρών και ως αποτέλεσμα ούτε στο πρόβλημά μας. Αυτό συμβαίνει επειδή αυτές οι χρονικές σειρές εξαρτώνται από μακροχρόνιες εξαρτήσεις ανάμεσα σε πληθώρα δεδομένων, όπου τα recurrent δε μπορούν να αξιοποιήσουν.

Εμφανίζεται μία συμπεριφορά λήθης (forgetting behaviour), όπως ονομάζεται, που συμβαίνει λόγω του φαινομένου της εξαφανιζόμενης κλίσης που υφίσταται με τον χρόνο κάτω από κατάλληλες συνθήκες. Έτσι, το κομμάτι του σφάλματος που οφείλεται σε παλιές τιμές της χρονοσειράς αρχίζει και ξεχνιέται και η επίδρασή του μειώνεται εκθετικά καθώς περνάει περισσότερο στο παρελθόν. Για να ξεπεραστεί αυτό το πρόβλημα δημιουργήθηκε ένα νέο είδος αναδρομικού δικτύου, το NAR και το NARX. Αυτά τα νευρωνικά θα χρησιμοποιήσουμε σε αυτή τη διπλωματική επειδή είναι πιο αποδοτικά στα προβλήματα πρόβλεψης και έχουν μεγάλες δυνατότητες μάθησης.

Τα NAR και NARX (Nonlinear AutoRegressive Models with exogenous inputs) συνήθως αποτελούνται από σιγμοειδής συναρτήσεις στα κρυμμένα επίπεδα και καθυστερήσεις της γραμμής για να αποθηκεύουν παλαιότερες τιμές της χρονικής σειράς. Έχουν χρησιμοποιηθεί τα τελευταία χρόνια με ενθαρρυντικά αποτελέσματα και, ειδικά, σε προβλέψεις ηλεκτρικού φορτίου και τιμής της ηλεκτρικής ενέργειας.

### **5.3 Support Vector Machines**

Τα SVMs, όπως αναφέραμε και νωρίτερα, είναι μία τεχνική που υπάγεται στον κλάδο των ευφυών συστημάτων και ειδικεύεται σε προβλήματα ομαδοποίησης. Όμως,

μπορεί να τροποποιηθεί και να είναι ικανή να λύσει και προβλήματα regression, δηλαδή προβλήματα πρόβλεψης σαν το δικό μας. Έτσι, αναπτύχθηκε μία ειδική κατηγορία SVM, το support vector regression ή εν συντομία SVR. Το SVR που χρησιμοποιήσουμε σε αυτήν την διπλωματική ονομάζεται  $\epsilon$ -SVR και ο στόχος του είναι να βρει μία συνάρτηση  $f(x)$ , που η απόκλισή της στα δεδομένα εκπαίδευσης είναι μικρότερη από  $\epsilon$ . Πιο συγκεκριμένα, έστω έχουμε μία σειρά από δεδομένα  $(x_1, y_1) \dots \dots (x_n, y_n)$ , όπου τα  $x_i$  είναι οι εισόδοι και  $y_i$  είναι οι έξοδοι. Το  $\epsilon$ -SVR μετατρέπεται, έτσι, σε πρόβλημα βελτιστοποίησης, αφού προσπαθούμε να βρούμε την καλύτερη συνάρτηση που με τις δεδομένες εισόδους, οι έξοδοι δεν θα απέχουν από την πραγματική τιμή πάνω από  $\epsilon$ .

Παρακάτω παραθέτουμε τους μαθηματικούς τύπους που περιγράφουν το πρόβλημα βελτιστοποίησης:

$$\min_{\omega, b, \xi, \xi^*} \quad \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n (\xi_i + \xi_i^*)$$

και ισχύει ότι

$$\begin{cases} y_i - ((\omega, x_i) + b) \leq \epsilon + \xi_i \\ ((\omega, x_i) + b) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

όπου  $n$  είναι ο αριθμός των δειγμάτων,  $\xi_i$  και  $\xi_i^*$  αντιπροσωπεύουν τα άνω και κάτω σφάλματα εκπαίδευσης αντίστοιχα,  $C > 0$  είναι μία σταθερά που κανονικοποιεί το εμπειρικό σφάλμα.

Εκτός από αυτό δοκίμασα και ένα υβριδικό μοντέλο βασισμένο στο  $\epsilon$ -SVR, που ονομάζεται SVRARIMA και αποτελεί μία ένωση του SVR και της στατιστικής τεχνικής ARIMA. Χρησιμοποιούμε την αποτελεσματικότητα του SVR σαν ευφυής τεχνική ώστε να προβλέψουμε τα μη-γραμμικά πρότυπα που είναι δύσκολο να μοντελοποιηθούν. Συνδυάζοντάς τη με τη μέθοδο ARIMA για να μοντελοποιήσουμε τα σφάλματα παίρνουμε ένα πολλά υποσχόμενο μοντέλο. Αναλυτικότερα, αφού έχουμε εκπαιδεύσει το σύστημα με δεδομένα του τύπου  $(x_1, y_1) \dots \dots (x_n, y_n)$  και

έστω η έξοδος του SVR στα δεδομένα εκπαίδευσης είναι της μορφής  $y_1^{**}, \dots, y_n^{**}$ , τότε μοντελοποιούμε με την μέθοδο ARIMA την μεταβλητή  $e_i^{**} = y_i - y_i^{**}$ . Έτσι, προσπαθούμε να προβλέψουμε το error που θα δώσει το SVR στα δεδομένα ελέγχου. Αυτό λειτουργεί στη προκειμένη περίπτωση επειδή τα δεδομένα είναι χρονοσειρές και τυχαίνει τα δεδομένα που θέλουμε να προβλέψουμε να είναι οι χρονικές στιγμές που έπονται τα δεδομένα εκπαίδευσης. Η έξοδος του SVRARIMA συνδυάζει την πρόβλεψη του SVR με την πρόβλεψη σφάλματος του ARIMA για να έρθουμε πιο κοντά στην πραγματική τιμή. Ο τύπος δίνεται από:

$$out_i = y_i^{**} + e_i^{**}$$

Αυτή η μέθοδος δεν είναι πολλή γνωστή και δεν υπάρχουν κριτικές για τα αποτελέσματά της. Παρακάτω στο πειραματικό μέρος θα δούμε πως συμπεριφέρεται στα δεδομένα του προβλήματός μας.

Κλείνοντας το κεφάλαιο με τις ευφυείς τεχνικές είναι σημαντικό να αναφέρουμε την χρησιμότητά τους και την ικανότητά τους στο να χειρίζονται προβλήματα με μεγάλη πολυπλοκότητα και μη-γραμμικότητα με μεγάλη ευκολία. Έχουν την δυνατότητα να προβλέπουν spikes και να προσαρμόζονται σε μη-γραμμικές συμπεριφορές. Αν και είναι πλεονέκτημα, αυτή η μεγάλη ευελιξία πολλές φορές είναι αρνητικό. Όταν υπερεκπαιδεύεται το δίκτυο σε απότομες εκρήξεις και σε συμπεριφορές με ξεσπάσματα, τότε αυτό δεν συνεπάγεται πάντα σε καλά αποτελέσματα. Η ποικιλία τους, επίσης, είναι ευρεία και για κάθε πρόβλημα συνήθως αλλάζει ποιο είδος νευρωνικού είναι το καλύτερο. Αυτό αποτελεί δυσκολία στον να συγκρίνουμε μεθόδους γιατί πολλές φορές η απόδοση διαφορετικών μοντέλων αλλάζει με άλλο πρόβλημα ακόμα και με αλλαγή δεδομένων εκπαίδευσης. Παρακάτω, στο πειραματικό μέρος θα δούμε αναλυτικότερα τα αποτελέσματα των τεχνικών που περιγράψαμε παραπάνω και θα τις σχολιάσουμε.

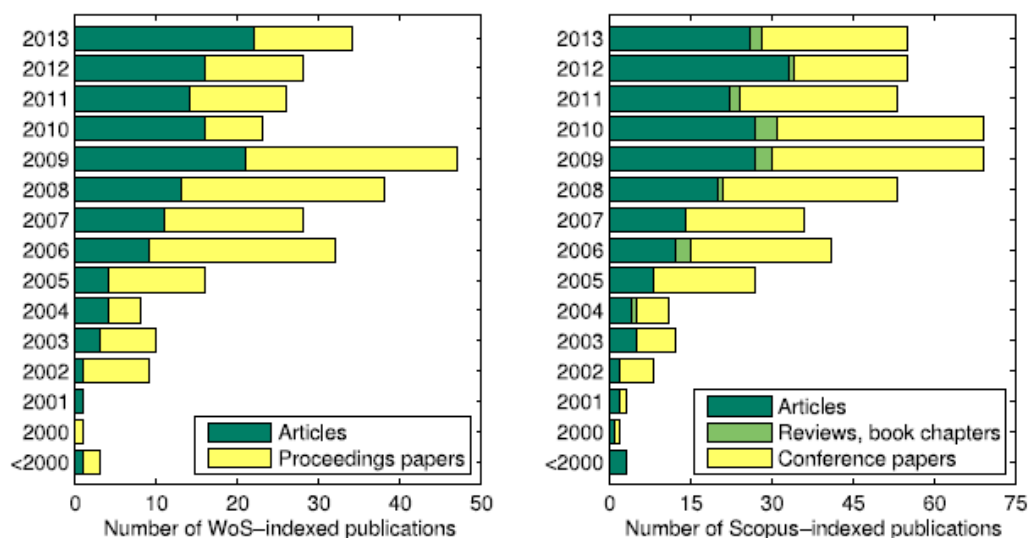
## **Κεφάλαιο 6 - Δεδομένα**



Προτού περάσουμε στο πειραματικό κομμάτι που θα δοκιμάσουμε και θα εξετάσουμε τις μεθόδους που παρουσιάσαμε στα προηγούμενα κεφάλαια, θα αφιερώσουμε αυτό το κεφάλαιο για να αναφερθούμε στα δεδομένα που υπάρχουν για μελέτη.

Το πρόβλημα του FBMC είναι σχετικά νέο στον χώρο της επιστήμης αφού έχει ξεκινήσει από τα μέσα του 2015. Οπότε τα δεδομένα, δυστυχώς, θα περιορίζονται σε αυτό το διάστημα του ενάμισι περίπου χρόνου. Παρόλα αυτά υπάρχουν διάφορες πηγές για να αναζητήσει κανείς πληροφορίες και δεδομένα που αντιπροσωπεύουν το πρόβλημά μας. Αρχικά, είναι σημαντικό να αναφερθούν δύο μεγάλες βάσεις δεδομένων που λειτουργούν εδώ και χρόνια τροφοδοτώντας πολλές μελέτες με χρήσιμα δεδομένα. Οι συγκεκριμένες βάσεις δεδομένων είναι το Web of Science(WoS) και το Scopus. Δεν έχουν πολλές διαφορές μεταξύ τους και αξίζουν και τα δύο να τα συμβουλευόμαστε για νέα δεδομένα που εμφανίζονται στο χώρο.

Η μηχανή αναζήτησης του Scopus προσφέρει καλύτερη εξυπηρέτηση και πιο φιλικές στον χρήστη αναζητήσεις ενώ η βάση δεδομένων του WoS έχει καλύτερη ποιότητα σε κάποιου είδους δεδομένα. Από το 1989 μέχρι το 2013 έχουμε έναν ικανοποιητικό αριθμό δημοσιεύσεων και στις δύο database για τον κλάδο της πρόγνωσης στον τομέα της ηλεκτρικής ενέργειας. Παρακάτω παραθέτουμε διαγράμματα για τους ακριβείς αριθμούς δημοσιεύσεων:



Όμως, τα δεδομένα που χρησιμοποιήσαμε δεν προερχόντουσαν από αυτές τις βάσεις δεδομένων. Τα δεδομένα και γενικά την πρώτη επαφή με το πρόβλημα του FBMC την απόκτησα μέσω της εταιρίας Ensco. Η Ensco είναι μία διεθνής εταιρία που παρέχει ενεργειακές υπηρεσίες καθώς και προμήθειες και εργαλεία σε όλο τον κόσμο. Αφού ήρθαμε σε επαφή με την Ensco είχα στην διάθεση μου μία σειρά από χρήσιμα δεδομένα που με βοήθησαν στο πειραματικό κομμάτι της διπλωματικής.

Πρώτα από όλα μας δόθηκαν δεδομένα μετακίνησης ηλεκτρικού φορτίου ανάμεσα σε τέσσερεις χώρες που συμμετέχουν στο πρόγραμμα του FBMC, τις εξής: Γαλλία, Γερμανία, Βέλγιο και Ολλανδία. Τα δεδομένα ξεκινάνε από τις 22/11/2015 στις 0:00 και συνεχίζουν ανά ώρα μέχρι τις 5/11/2016 στις 23:00 και μας δείχνουν τα MW τεσσάρων ροών. Πρώτη ροή είναι Γαλλίας-Γερμανίας, δεύτερη είναι Ολλανδίας-Γερμανίας, τρίτη είναι Ολλανδίας-Βελγίου και, τέλος, τέταρτη ροή είναι αυτής της Βελγίου-Γαλλίας. Είναι εμφανές ότι κάποιες τιμές είναι αρνητικές. Αυτό εξηγείται αν για παράδειγμα η ροή Γαλλίας-Γερμανίας κατευθύνεται από την Γερμανία προς την Γαλλία, οπότε χρησιμοποιούμε αρνητικό πρόσημο.

Εκτός αυτών των δεδομένων, είχαμε την ευκαιρία να χρησιμοποιήσουμε και άλλου είδους δεδομένα που θα βοηθήσουν την πρόβλεψη του ηλεκτρικού φορτίου. Μας δόθηκε πάλι για το ίδιο χρονικό διάστημα οι ωριαίες τιμές κατανάλωσης ηλεκτρικού φορτίου(παρακάτω θα αναφερόμαστε σε αυτό σαν Consumption) στις χώρες που προαναφέρθηκαν. Επίσης, για τις ίδιες χώρες δόθηκαν οι ωριαίες τιμές παραγωγής ηλεκτρικής ενέργειας που οφείλονται σε τρεις διαφορετικές πηγές. Οι πηγές ενέργειας είναι η αιολική(όπου θα αναφερόμαστε σαν Wind), δηλαδή πόση ηλεκτρική ενέργεια παράχθηκε με την χρήση της αιολικής ενέργειας, η πυρηνική(όπου θα αναφερόμαστε σαν Nuclear) και η ενέργεια από τα φωτοβολταϊκά(όπου θα αναφερόμαστε σαν PV) ή αλλιώς η ηλιακή. Αυτές οι τιμές αποτελούν σημαντικές πληροφορίες που μπορούν να βοηθήσουν την πρόβλεψη της μετακίνησης του ηλεκτρικού φορτίου. Είτε έμμεσα είτε άμεσα κάποια από αυτές τις μεταβλητές μπορεί να επηρεάζει ή να επηρεάζεται με άγνωστο τρόπο προς στιγμήν το ζητούμενο του προβλήματός μας και για αυτό έχουν υψηλή σημασία.

Η Enesco, τέλος, μας παρείχε και ένα τελευταίο σύνολο δεδομένων που αποδείχτηκε αρκετά χρήσιμο στη μελέτη του προβλήματος. Αυτό είναι ένα σύνολο προβλέψεων που κυκλοφορούν στην αγορά τις ενέργειας, έτσι ώστε να το χρησιμοποιήσω σαν ένα μέτρο σύγκρισης στα δικά μου αποτελέσματα.

Η διάρκεια των προβλέψεων ξεκινά από τις 20/9/2016 στις 0:00 μέχρι τις 31/10/2016 στις 23:00. Οι προβλέψεις ήταν ένα αναπόσπαστο κομμάτι της μελέτης μου για να μπορώ να καταλαβαίνω ποιες προβλέψεις δεν ήταν αποδοτικές και ποιες είχαν ικανοποιητικές προοπτικές, αλλά θα αναφερθούμε συγκεκριμένα για αυτές στο επόμενο κεφάλαιο που θα μιλήσουμε αναλυτικότερα για το πειραματικό μέρος της διπλωματικής μου.

Είναι άξιο λόγου να αναφέρουμε ότι όλα αυτά τα δεδομένα που ανέφερα παραπάνω είναι διαθέσιμα μέσω ενός εξαιρετικά χρήσιμου εργαλείου, που ονομάζεται JAO Utility Tool. Αυτή η πλατφόρμα μπορεί να βρεθεί στο διαδίκτυο μέσω δύο διαφορετικών URL, όπου μπορεί κάποιος να αποκτήσει πρόσβαση μέσω πρωτοκόλλων HTTP και SOAP αντίστοιχα.

Το interface της πλατφόρμας είναι αρκετά φιλικό προς τον χρήστη και μοιάζει κάπως έτσι:

**JAO**  
Joint Association Office

## CWE Flow Based Utility Tool

Reference time: 2016-11-08  
date: 2016-11-08  
hour: 1

**1) Check volume (interactive module)**

Here you can check the simultaneous execution of trading volumes of the markets involved in the CWE Market Coupling

HUB TO HUB EXCHANGES	Hub-to-Hub trade in MWh/h (please insert values)	Test 1: hub to hub inside FB space
DE/AT=>BE	0	<b>Trades feasible</b>
DE/AT=>NL	0	
DE/AT=>FR	0	
NL=>BE	0	
NL=>FR	0	
BE=>FR	0	

**2) Max volume (information module)**

Here you can find the maximal trade volumes (MWh/h) which can be physically transported between two Hubs under the condition that no other trade is executed between other Hubs.

	direction -->	direction <--
DE/AT=>BE	4171	4598
DE/AT=>NL	3130	4108
DE/AT=>FR	4663	4477
NL=>BE	2694	4155
NL=>FR	3179	4015
BE=>FR	3664	4500

**HUB POSITION**

	Hub Positions trade in MWh/h (please insert values)	Test 1: sum hub positions = 0	Test 2: hub positions inside FB space
DE/AT	0	<b>OK</b>	<b>Trades feasible</b>
BE	0		
FR	0		
NL	0		

	export	import
DE/AT	6405	-4995
BE	5949	-4500
FR	6101	-4663
NL	4204	-4155

Όπως θα μπορεί να παρατηρήσει κανείς, ο χρήστης μπορεί να συμπληρώσει την ημερομηνία και ώρα που επιθυμεί και έπειτα θα παραλάβει τις τιμές ηλεκτρικού φορτίου της αντίστοιχης χρονικής στιγμής. Εκτός αυτού το εργαλείο παρέχει και γραφήματα του δικτύου αλλά και της αγοράς ηλεκτρικής ενέργειας που αποδεικνύονται αρκετά χρήσιμα και εύχρηστα.

## **Κεφάλαιο 7 - Πειραματικά Αποτελέσματα**

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τα πειραματικά αποτελέσματα που λάβαμε από τα μοντέλα που δοκιμάσαμε στο πρόβλημα της πρόβλεψης ηλεκτρικού φορτίου. Τα πειράματα διεξάχθηκαν στο προγραμματιστικό περιβάλλον της MATLAB και την έκδοση R2016. Τα αποτελέσματα θα παρουσιαστούν ανάλογα με την κατηγορία των μοντέλων που ανήκουν και για κάθε ίδιο μοντέλο θα διευκρινιστούν οι διαφορές που παρατηρούμε σε αλλαγή παραμέτρων του συστήματος. Θα επικεντρωθούμε κυρίως στην ροή Γαλλίας-Γερμανίας και όταν χρησιμοποιήσουμε διαφορετική ροή θα αναφερθεί κατάλληλα. Επίσης, επειδή έχουμε στη διάθεση μου έτοιμες έγκυρες

προβλέψεις θα χρησιμοποιήσουμε αυτό το διάστημα για έλεγχο. Επομένως, τα δεδομένα εκπαίδευσης των μεθόδων θα είναι από τις 22/11/2015 μέχρι τις 19/9/2016 και τα δεδομένα για έλεγχο είναι από τις 20/9/2016 μέχρι 31/10/2016.

## 7.1 Στατιστικές Τεχνικές

Πρώτα θα δοκιμάσουμε τις στατιστικές τεχνικές που έχουμε προαναφέρει σε προηγούμενο κεφάλαιο.

### 7.1.1 Απλή Γραμμική Παλινδρόμηση

Με την χρήση της συνάρτησης `fitlm` δημιουργήσαμε τέσσερα μοντέλα γραμμικής παλινδρόμησης. Έστω  $X_1, X_2, \dots, X_n$  είναι οι ανεξάρτητες μεταβλητές εισόδου και  $Y$  είναι η εξαρτημένη μεταβλητή εξόδου. Το πρώτο ήταν καθαρά γραμμικό μοντέλο:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$$

Το δεύτερο μοντέλο ονομάζεται interactions και εμπεριέχει όρους που είναι ζευγάρια από παράγοντες μεταβλητών εισόδου και είναι:

$$Y = a_1X_1 + \dots + a_nX_n + \sum_{i=0, j=0}^n X_iX_j + b$$

Το τρίτο και το τέταρτο μοντέλο μοιάζουν μεταξύ τους. Ουσιαστικά, το τρίτο, που ονομάζεται purequadratic, είναι μία πιο απλή μορφή του τέταρτου μοντέλου, το οποίο ονομάζεται quadratic. Και τα δύο εμπεριέχουν στον τύπο τους τετραγωνισμένους όρους, με την μόνη διαφορά το μοντέλο quadratic να εμπεριέχει και τους όρους από το μοντέλο interactions. Οι μαθηματικοί τύποι είναι οι εξής:

Purequadratic:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b + c_1X_1^2 + \dots + c_nX_n^2$$

Quadratic:

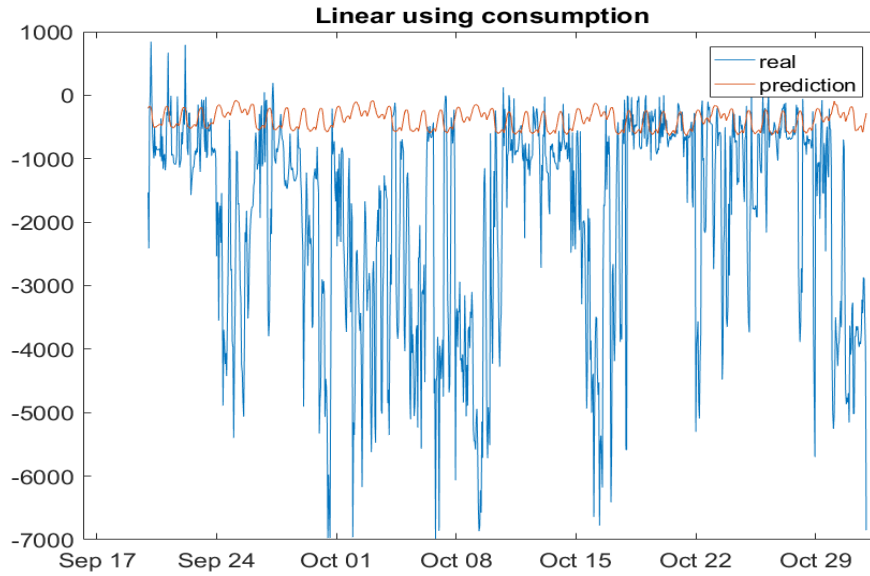
$$Y = a_1X_1 + \dots + a_nX_n + \sum_{i=0, j=0}^n X_iX_j + b + c_1X_1^2 + \dots + c_nX_n^2$$

Εδώ όμως είναι σημαντικό να αναφέρω οι τιμές του Consumption που χρησιμοποιήθηκαν είναι οι τιμές της επόμενης μέρας που σε πραγματικά δεδομένα δεν θα ήταν εφικτό γιατί η πρόβλεψη γίνεται την μέρα πριν. Για αυτό χρησιμοποιήσαμε τα ίδια.

Εκπαιδεύσαμε αυτά τα μοντέλα με τρεις διαφορετικούς τρόπους που διαφέρουν στις μεταβλητές εισόδου:

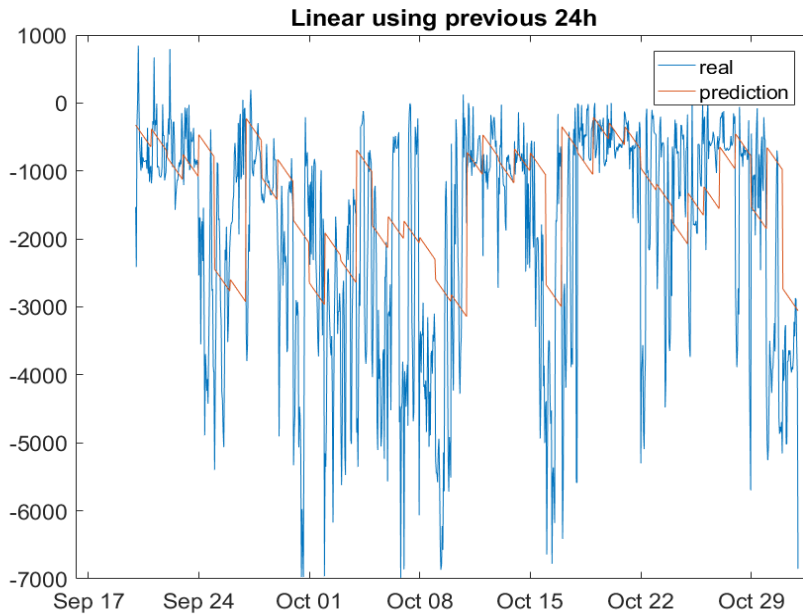
α) Χρησιμοποιούμε σαν μεταβλητές εισόδου το Consumption, δηλαδή αντιστοιχούμε την ωριαία τιμή κατανάλωσης με την ωριαία τιμή ηλεκτρικού μεταφοράς ηλεκτρικού φορτίου.

Όπως, μπορεί κανείς να παρατηρήσει στα επόμενα διαγράμματα χρησιμοποιώντας μόνο το Consumption δεν αποφέρει πολύ καλά αποτελέσματα σε αυτά τα τέσσερα μοντέλα που εξετάζουμε. Όλα έχουν την ίδια μορφή και δεν μπορούν να προβλέψουν την ασταθή αλλαγή της τιμής του ηλεκτρικού φορτίου. Ενδεικτικά, το RMSE αυτών των μοντέλων κυμάνθηκε γύρω στα 2300 περίπου σε όλα.

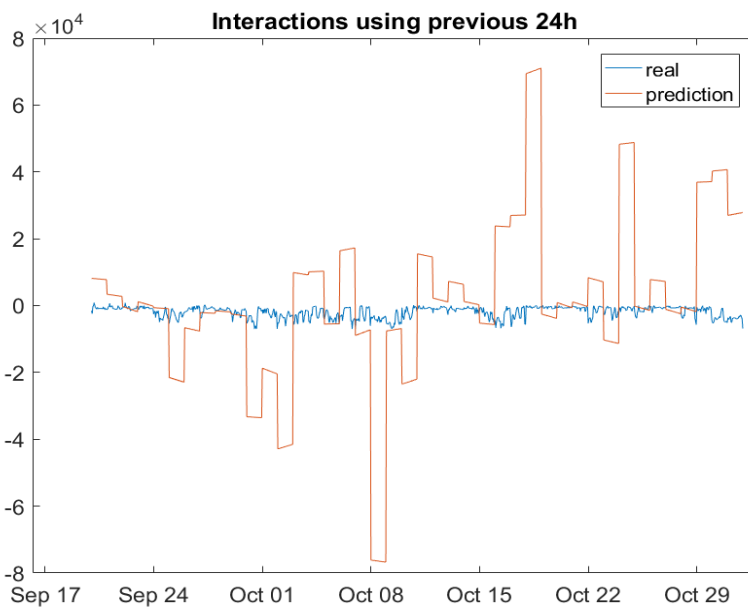


Η μορφή των υπολοίπων μοντέλων(interactions, pure quadratic και quadratic) είναι αρκετά παρόμοια με την από πάνω και παρατηρούμε ότι δεν αντιπροσωπεύει μία καλή πρόβλεψη.

β) Χρησιμοποιούμε σαν μεταβλητές εισόδου τις τιμές των προηγούμενων 24 ωρών μαζί με την ώρα(δηλαδή 0-23), την ημερομηνία(1-30~31) και τον μήνα(1-12), σύνολο επομένως 27 ανεξάρτητων μεταβλητών. Πήραμε τα ακόλουθα αποτελέσματα:

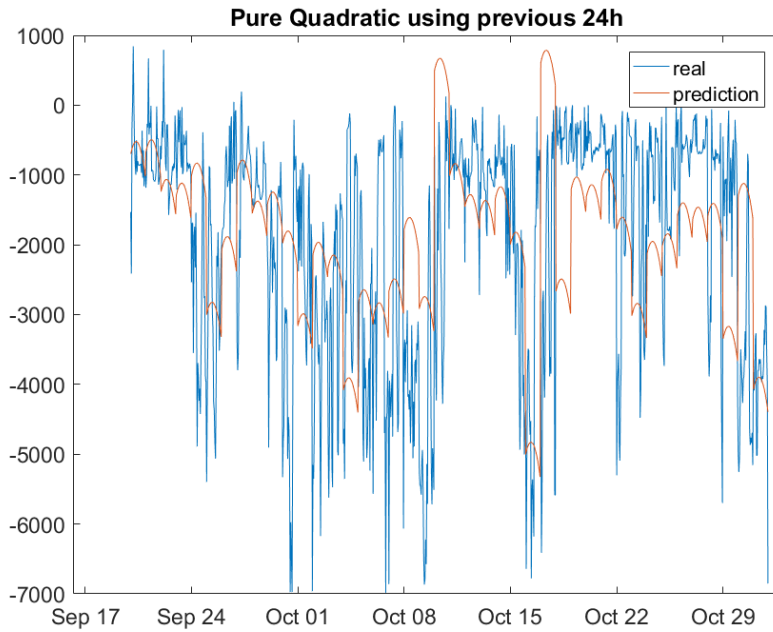


Παρατηρούμε ότι στο παραπάνω διάγραμμα με το linear μοντέλο ( $rmse = 1616$ ) η πρόβλεψη έχει μία αρκετά καλή μορφή για την κατεύθυνση των διακυμάνσεων αλλά δεν μπορεί να προβλέψει τα ξεσπάσματα και τα spikes που υπάρχουν στις πραγματικές τιμές. Σίγουρα, όμως, είναι μία βελτίωση από το προηγούμενο ταίριασμα με το Consumption.

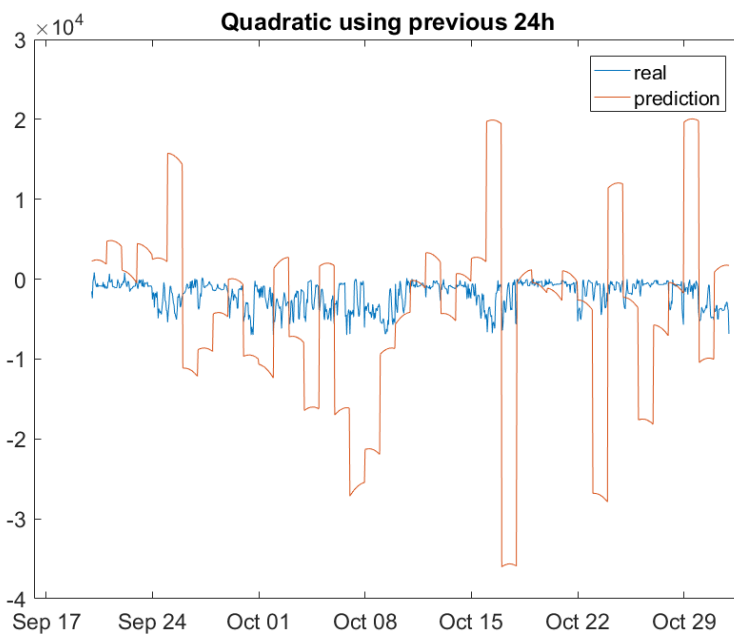


Στο μοντέλο interaction ( $rmse = 23751$ ) παρατηρούμε μεγάλα σφάλματα, το μοτίβο δεν είναι σωστό και γενικά μπορούμε να αποφανθούμε ότι δεν είναι κατάλληλο για την πρόβλεψη.





Το pure quadratic (rmse = 1721) σίγουρα αποτελεί βελτίωση καθώς βλέπουμε ότι σε πολλά σημεία ακολουθεί το γενικό σχήμα της χρονοσειράς χωρίς όμως ακόμα να μπορεί να αναπαράγει τα διάφορα ξεσπάσματα ή spikes που υπάρχουν.

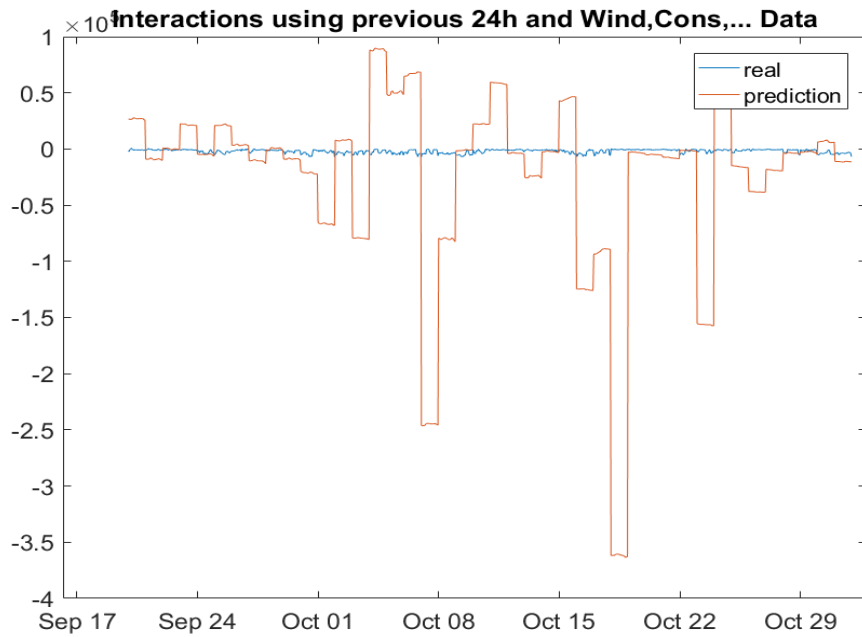


Το Quadratic (rmse = 11395) μοντέλο έχει τα ίδια προβλήματα με το μοντέλο interactions, δηλαδή απέχει πάρα πολύ από την μορφή των πραγματικών τιμών και είναι ακατάλληλο για πρόβλεψη.

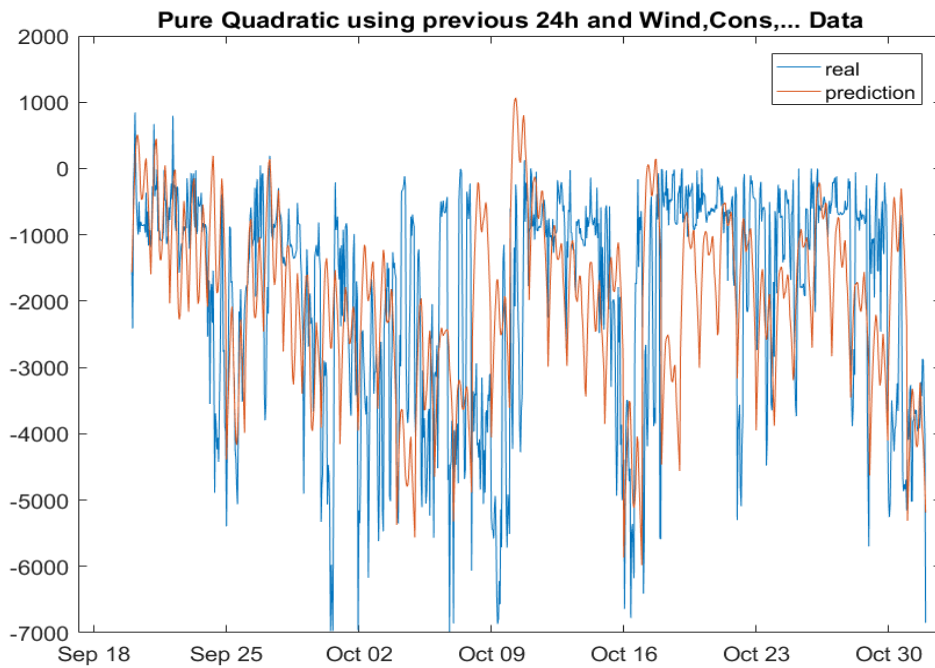
γ) Εκτός από τις προηγούμενες 24 ώρες χρησιμοποιώ και τα side data που έχω στη διάθεσή μου, δηλαδή το Consumption, Wind, PV, Nuclear.



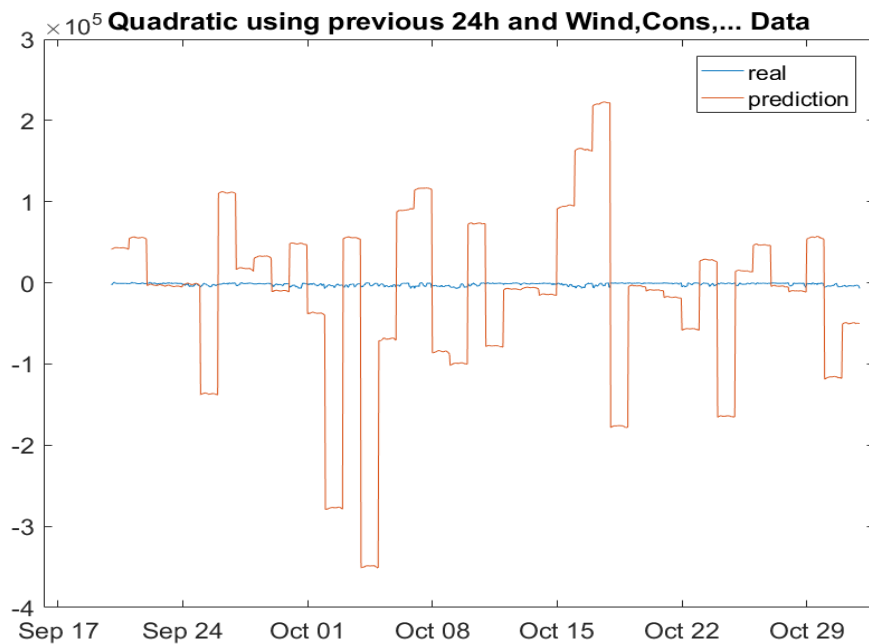
Στο  $\text{linear}(\text{rmse} = 1593)$  μοντέλο παρατηρούμε ότι με την προσθήκη των έξτρα μεταβλητών βλέπουμε καλύτερη συμπεριφορά στα spikes χωρίς όμως να χάνεται η συνολική μορφή της γραφικής παράστασης.



Στο interactions (rmse = 82026) όπως και προηγουμένως δεν έχουμε καλές προβλέψεις και φαίνεται να είναι αδύνατο να προβλέψει και την μέση τάση των τιμών, πόσο μάλλον τα απρόβλεπτα fluctuations που έχουμε παρατηρήσει στα πραγματικά δεδομένα.



Στο pure quadratic (rmse = 1739) παρατηρούμε ότι δεν υπάρχει ιδιαίτερη βελτίωση στο rmse με τις επιπλέον μεταβλητές αλλά σίγουρα υπάρχει καλύτερη απόκριση στα spikes.



Όπως και πριν δε μας ξαφνιάζει η απόδοση του quadratic (rmse = 105191) μοντέλου που αποκλίνει σημαντικά από τις πραγματικές τιμές και συγκεκριμένα εδώ έχουμε την χειρότερη απόδοση μέχρι στιγμής.

Φαίνεται γενικά πως τα μοντέλα, όπως το quadratic και το interactions, που έχουν όρους της μορφής  $X_i * X_j$ , δηλαδή όρους που αποτελούνται από παράγοντες των ανεξάρτητων μεταβλητών να μην ταιριάζουν στο πρόβλημα μας.

Παρόλα αυτά είναι σημαντικό να ειπωθεί ότι στις προβλέψεις που χρησιμοποιήθηκαν τα επιπλέον δεδομένα πήραμε την πραγματική τους τιμή για την ώρα  $k$  σαν είσοδο στο μοντέλο για προβλέψουμε την τιμή της μεταφοράς του φορτίου για την ίδια ώρα. Όμως, αυτό δεν είναι ρεαλιστικό γιατί όταν καλούμαστε να προβλέψουμε στην

πραγματικότητα δεν θα έχουμε στην διάθεσή μας τα πραγματικά δεδομένα. Για αυτό τον λόγο χρησιμοποιήσαμε τα παραπάνω μοντέλα (linear, interactions, pure quadratic, quadratic) ώστε να προβλέψω τις τιμές των side data ώστε να χρησιμοποιούμε αυτές όπου χρειαστεί. Παρακάτω θα παραθέσουμε τα αποτελέσματα και ποιο μοντέλο είναι καλύτερο για ποιο δεδομένο(Consumption, Nuclear, PV, Wind). Στο υπόλοιπο της διπλωματικής θα χρησιμοποιούνται αυτές οι προβλέψεις σαν εξωγενή δεδομένα χωρίς να χρειαστεί να αναφερθεί ξανά ότι είναι προβλέψεις αυτών των μοντέλων.

Σαν μετρική ποιο μοντέλο θα χρησιμοποιήσουμε για κάθε κατηγορία δεδομένων είναι το RMSE.

#### Consumption:

Στο consumption παρατηρούμε ότι το linear μοντέλο έχει τα καλύτερα αποτελέσματα με 8101 rmse.

#### Nuclear:

Στο nuclear παρατηρούμε μεγάλες διαφορές μεταξύ των μοντέλων. Τα μοντέλα linear και pure quadratic έχουν ελάχιστο σφάλμα και αυτό φαίνεται από rmse τους που είναι 189 και 260 αντίστοιχα. Ενώ τα μοντέλα interactions και quadratic ξεπερνούν τα 17000 rmse. Επομένως, επιλέγουμε το linear μοντέλο.

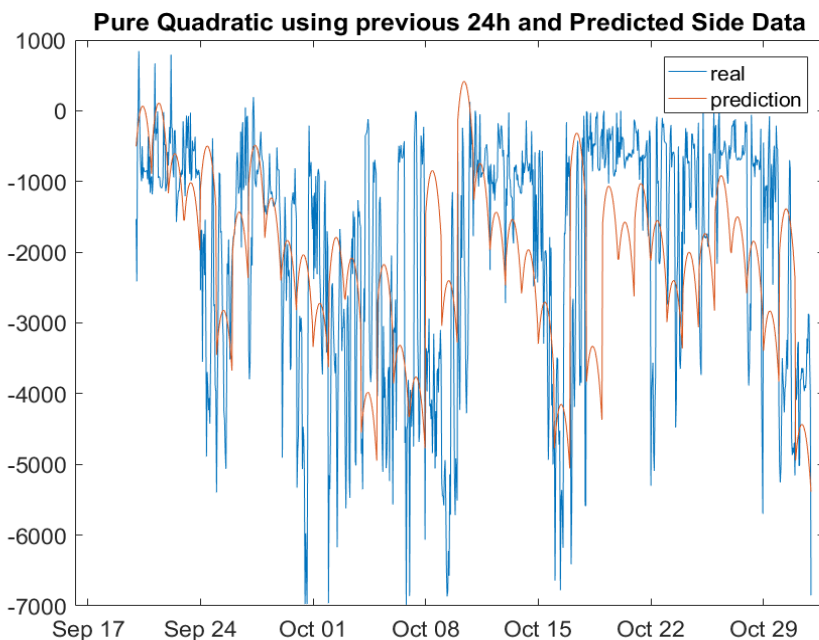
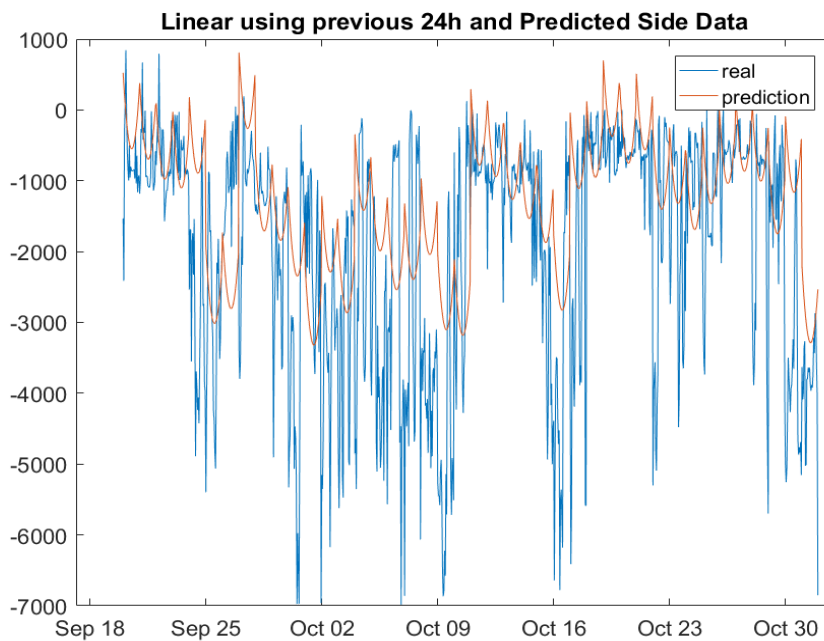
#### PV:

Εδώ βλέπουμε πάλι κοντινές τιμές από linear και pure quadratic με το δεύτερο μοντέλο να είναι λίγο καλύτερο.

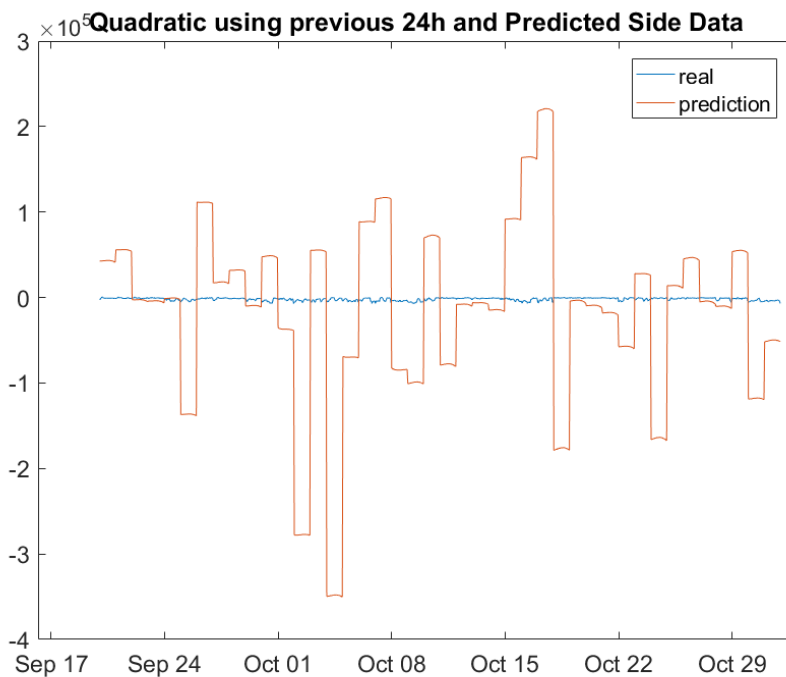
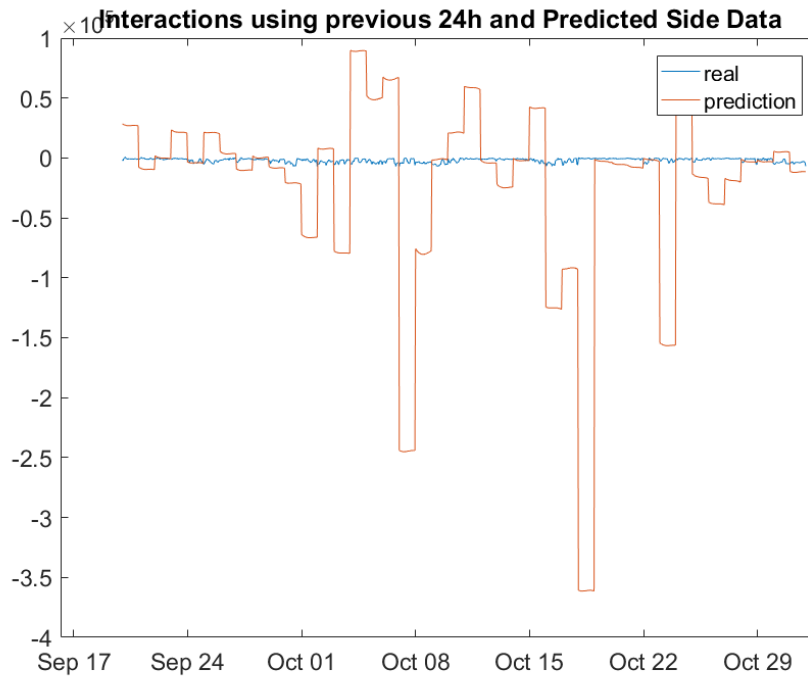
#### Wind:

Τέλος, στο Wind παρατηρούμε το ίδιο μοτίβο όπως και στα προηγούμενα και ανάμεσα στα linear και pure quadratic υπερτερεί το δεύτερο.

Τώρα θα παρουσιάσω τα αποτελέσματα των παραπάνω μοντέλων με είσοδο τις προβλεπόμενες τιμές των εξωγενών δεδομένων. Στο rmse δεν υπάρχουν διαφορές από τις προβλέψεις με τις πραγματικές τιμές αλλά μπορούμε να παρατηρήσουμε την μορφή των linear και pure quadratic μοντέλων ότι έχει αλλάξει αισθητά αλλά όχι απαραίτητα προς το χειρότερο.



Από την άλλη τα μοντέλα interactions και quadratic παραμένουν να μην αποδίδουν τόσο καλά.

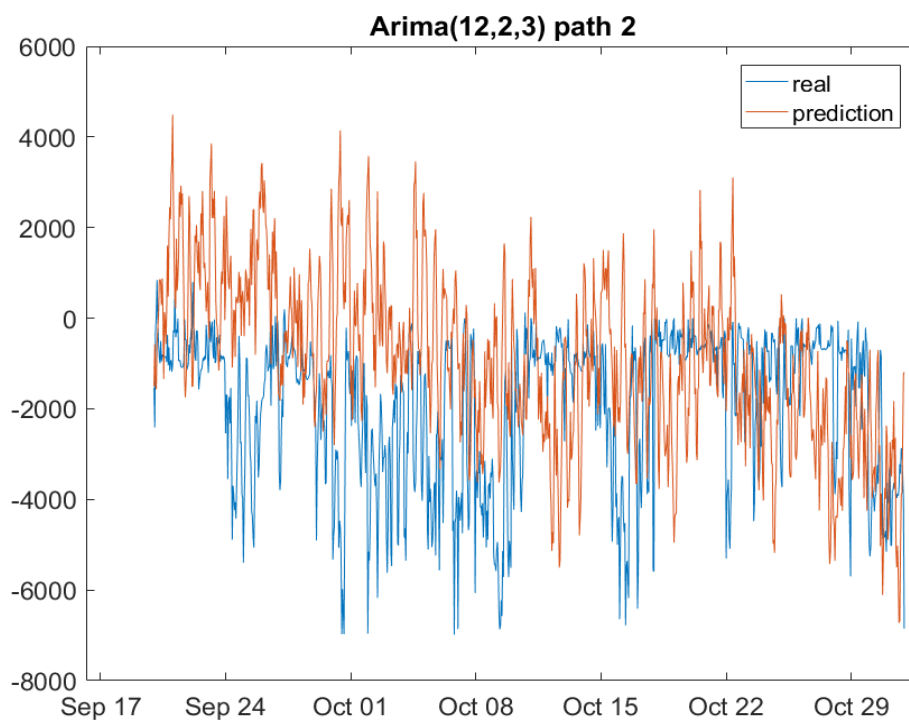


Έτσι, κλείνοντας τις γραμμικές regression τεχνικές εκλάβαμε ότι τα μοντέλα linear και pure quadratic είναι τα πλέον κατάλληλα μέχρι στιγμής για μία καλή πρόβλεψη στο πρόβλημα του FBMC.

### 7.1.2 Μοντέλο Arima

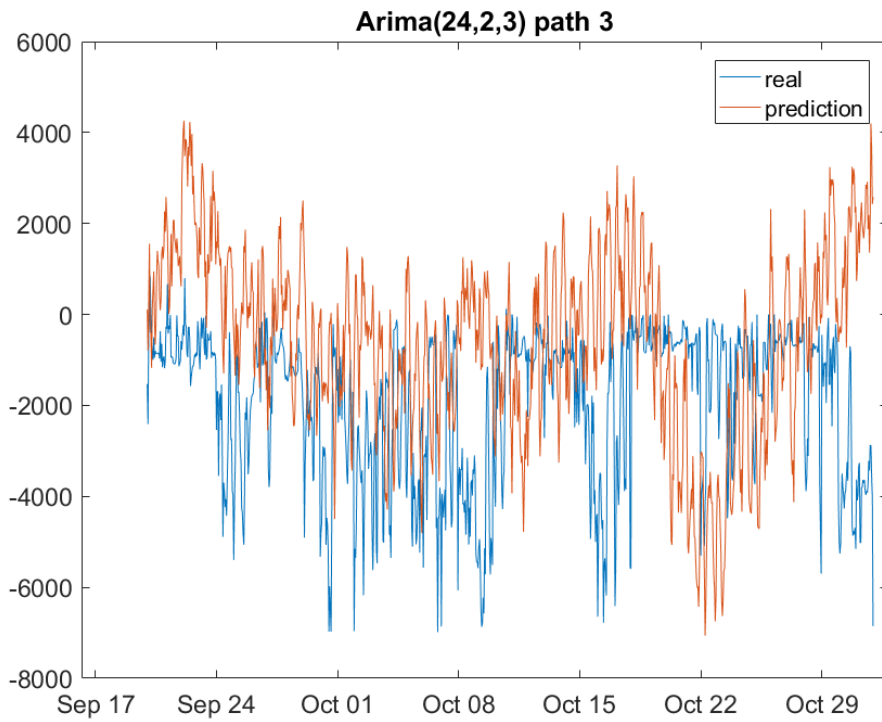
Μοντελοποιήσαμε διάφορες μορφές του μοντέλου  $arima(p,d,q)$ . Δοκιμάσαμε όλους τους συνδυασμούς που μπορούν να γίνουν με  $p = 12,24$ ,  $d = 0,1,2$  και  $q = 1,2,3$ . Δηλαδή χρησιμοποίησα 12 ή 24 προηγούμενες τιμές σαν μεταβλητές, παραγωγίσαμε 0 ή 1 ή 2 φορές και όρισα από 1 έως 3 όρους σφάλματος. Ενδεικτικά παρακάτω θα παρουσιάσουμε τα αποτελέσματα άξια λόγου και γενικές παρατηρήσεις πάνω σε αυτά.

Είναι απαραίτητο να τονίσουμε ότι κάθε μοντέλο  $arima$  το έτρεξα με τρία διαφορετικά μονοπάτια, τα οποία βγάζουν διαφορετικά αποτελέσματα μεταξύ τους. Παρακάτω βλέπουμε τα τρία καλύτερα μοντέλα σύμφωνα με το  $rmse$  τους αλλά και με την γενική μορφή που παρουσιάζουν.

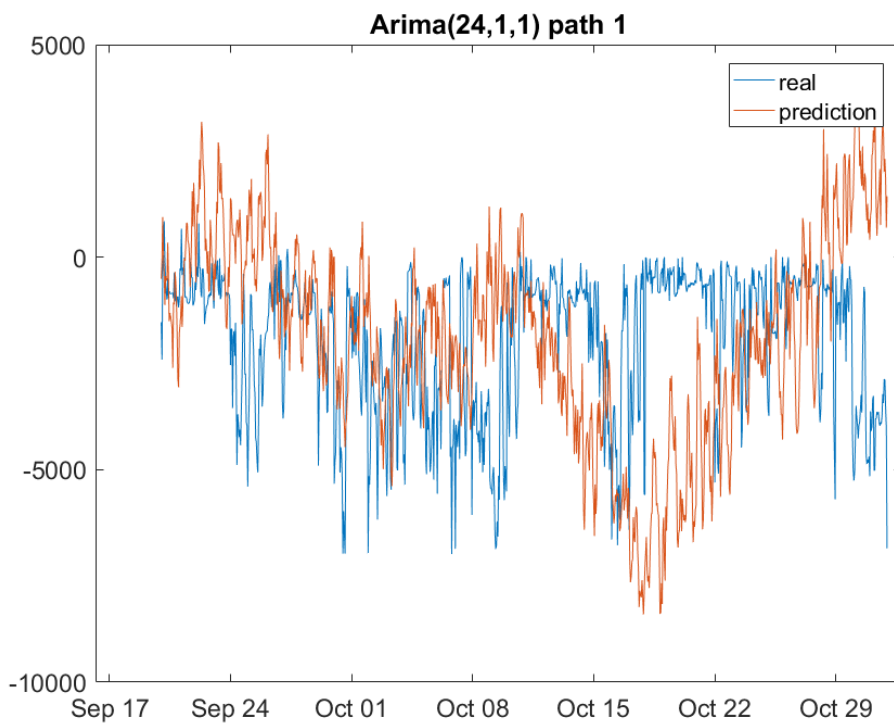


με  $rmse = 2669$





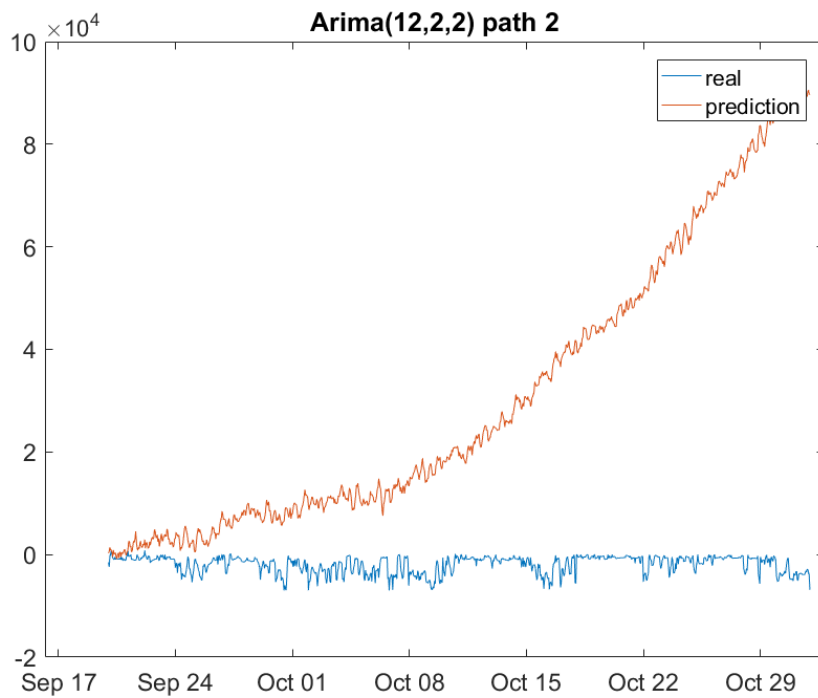
με  $rmse = 3040$

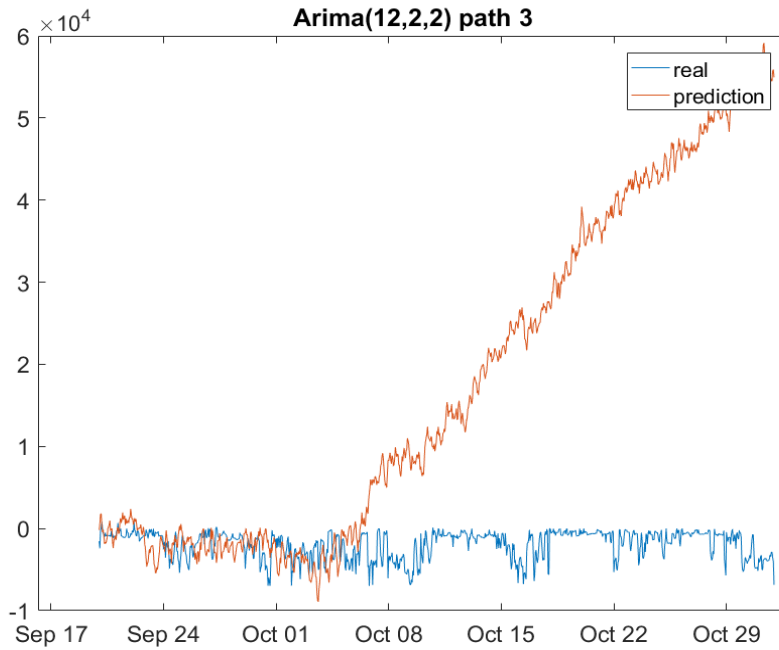


με  $rmse = 3051$

Με μία γρήγορη ματιά μπορούμε να καταλάβουμε ότι το μοντέλο Arima μπορεί να κατασκευάσει καλύτερα τα ξεσπάσματα που υπάρχουν στις πραγματικές τιμές αλλά

αυτό είναι χρήσιμο όταν γνωρίζουμε ή όταν μπορούμε να προβλέψουμε την γενική τάση και μέση τιμή της χρονοσειράς. Διαφορετικά αποτελεί μεγάλο κίνδυνο να προβλέψουμε τα spikes πάνω σε λάθος μορφές συνάρτησης επειδή θα αποκλίνει ακόμα περισσότερο από την πραγματική τιμή. Αυτό μπορούμε να το δούμε σε μερικά από τα μοντέλα arima που αν και είναι αρκετά κοντά στην αρχιτεκτονική τους με κάποια από τα παραπάνω, έχουν τεράστια σφάλματα και σε καμία περίπτωση δε μπορούν να χρησιμοποιηθούν για έγκυρες προβλέψεις. Τέτοια παραδείγματα είναι τα ακόλουθα:



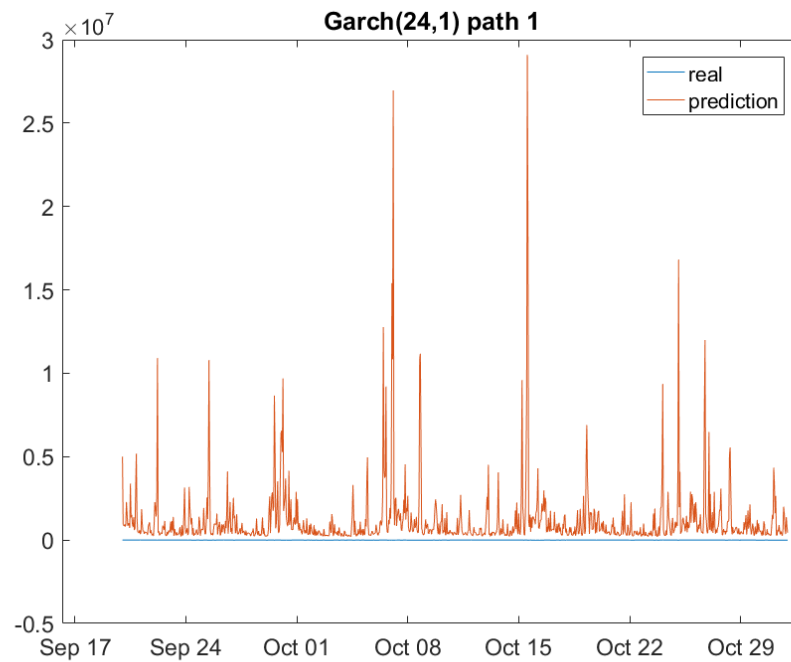


### 7.1.3 Μοντέλο Garch

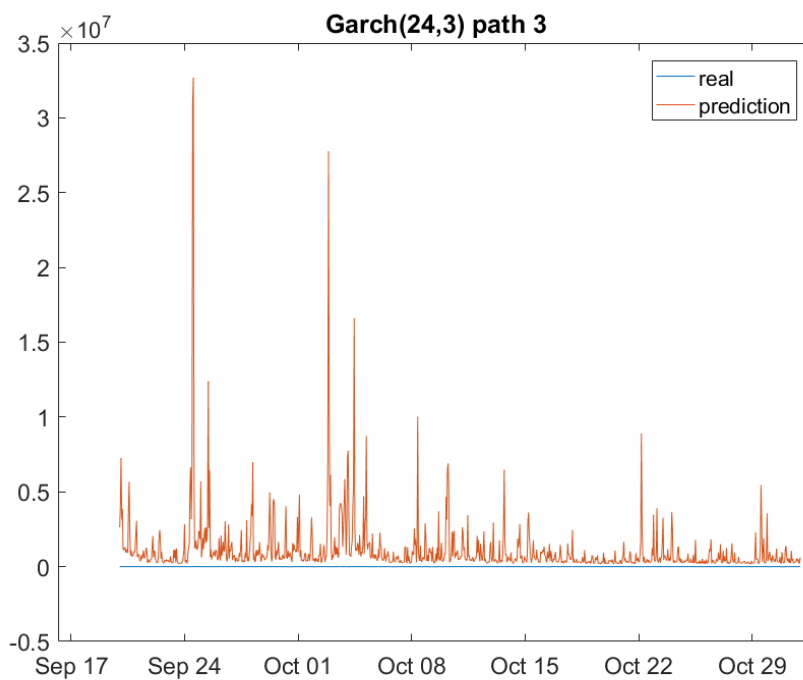
Έπειτα από το arima δοκιμάσαμε το Garch(p, q) και όπως πριν κάθε φορά κάναμε simulate τρία διαφορετικά path για να δούμε τις διαφορές και τις αποκλίσεις που μπορούν να συμβούν. Έτσι, μπορούμε να παρατηρήσουμε την γενική συμπεριφορά του μοντέλου και να μην εξαρτόμαστε από τυχόν περιστασιακά καλά ή κακά αποτελέσματα. Υλοποίησα το garch για p=12,24 και q=1,2,3.

Σαν γενική παρατήρηση στα αποτελέσματα του garch είναι ότι αν και είναι κατασκευασμένο να προσομοιώνει καλύτερα τα απότομα ξεσπάσματα, η απόδοση του ήταν χειρότερη από το arima. Όσον αφορά την εκτίμηση σφάλματος (rmse, wmae) είναι αυστηρά κατώτερο από όλες τις προηγούμενες μεθόδους.

Τα καλύτερα δείγματα είναι τα δύο



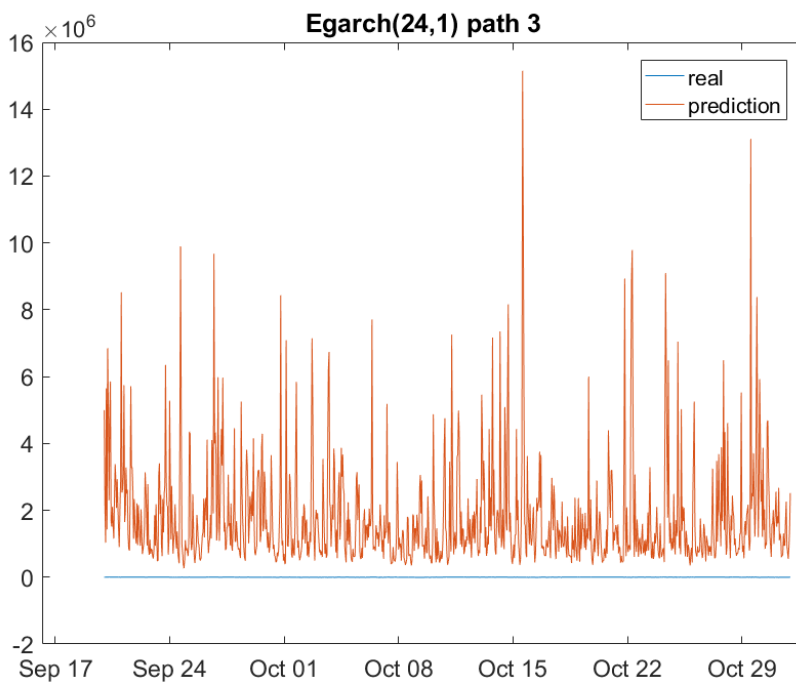
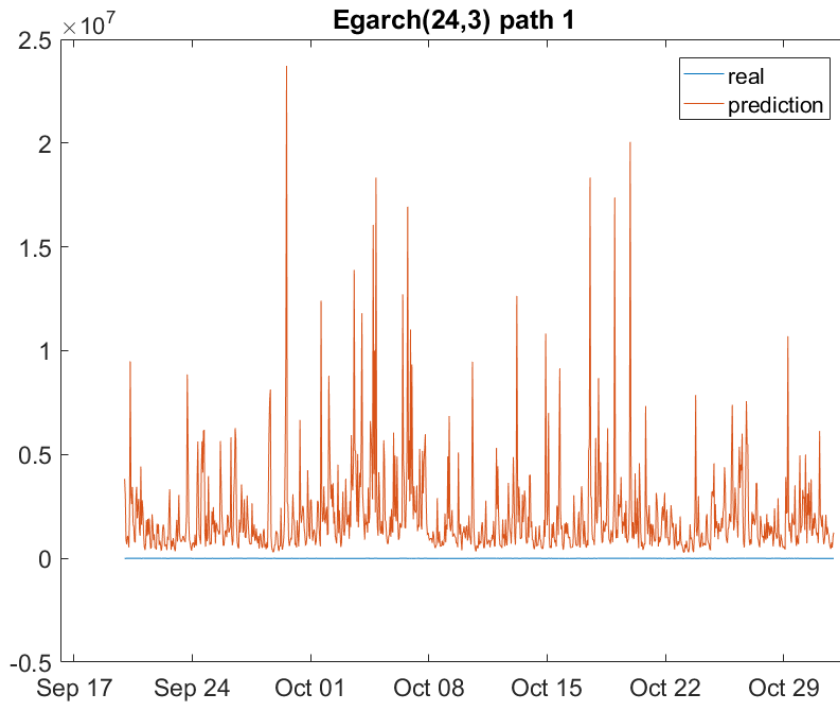
ακόλουθα:



Και στα δύο παρατηρούμε μεγάλη έμφαση στα spikes αποκλίνοντας από την πραγματική τιμή ακόμα περισσότερο. Τα υπόλοιπα παραδείγματα garch είναι στην ίδια μορφή που τελικά αποδεικνύεται ακατάλληλη για προβλέψεις τέτοιου είδους.

### 7.1.4 Μοντέλο EGarch

Όπως και προηγουμένως, έτσι και εδώ τεστάρουμε το μοντέλο EGarch( $p, q$ ) για  $p=12,24$  και  $q=1,2,3$ . Τα αποτελέσματα του egarch μοιάζουν αρκετά με αυτά του garch, το οποίο είναι λογικό. Παρακάτω βλέπουμε δύο χαρακτηριστικά παραδείγματα:

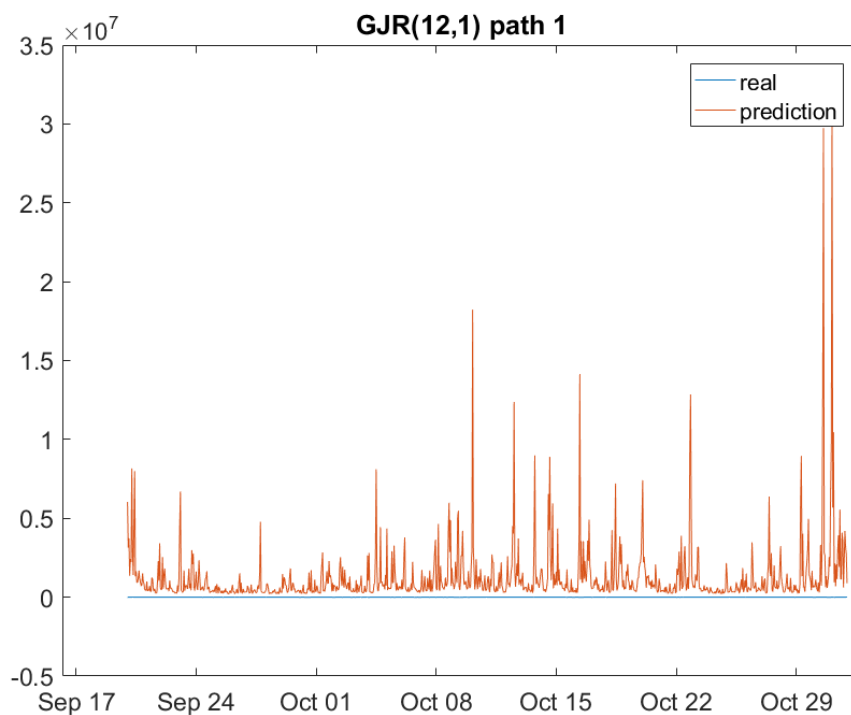


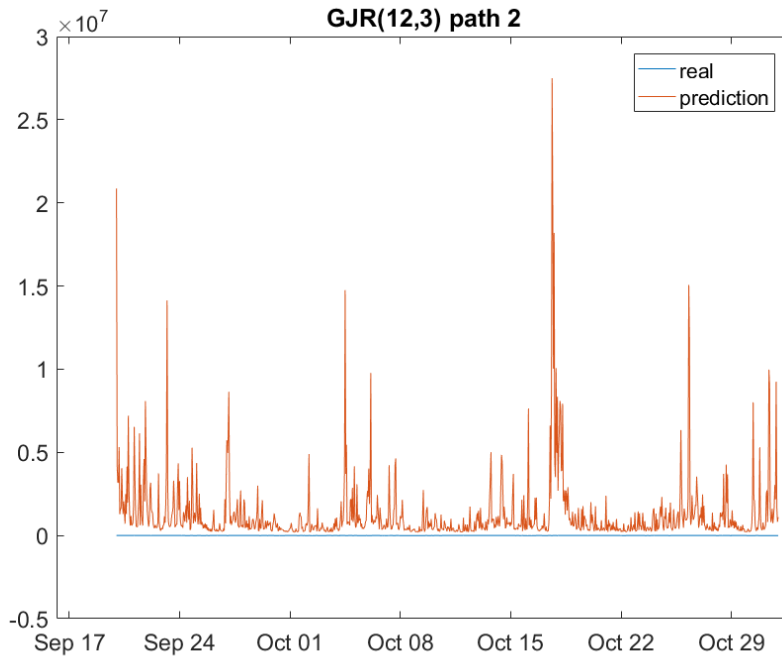
Μπορούμε να παρατηρήσουμε ότι υπάρχουν πάλι τεράστια σφάλματα όπως πριν αλλά εδώ οι προβλεπόμενες τιμές παίρνουν συνήθως αρκετά μεγαλύτερες τιμές από τις πραγματικές. Σαν σύνολο όλες οι προσομοιώσεις των egarch μοντέλων είχαν την ίδια μορφή και σαν μέσο όρο σφάλματος (rmse) της τάξης  $10^6$ .

### 7.1.5 Μοντέλο GJR

Τεστάρουμε το μοντέλο GJR( $p, q$ ) για  $p=12,24$  και  $q=1,2,3$  και, όπως περιμέναμε, τα αποτελέσματα είναι παρόμοια με τα υπόλοιπα garch μοντέλα, δηλαδή έχουμε κακή αναπαράσταση της χρονοσειράς και τα ξεσπάσματα είναι τόσο μεγάλα που ουσιαστικά ξεπερνούν τις πραγματικές τιμές κατά πολλές τάξεις μεγέθους.

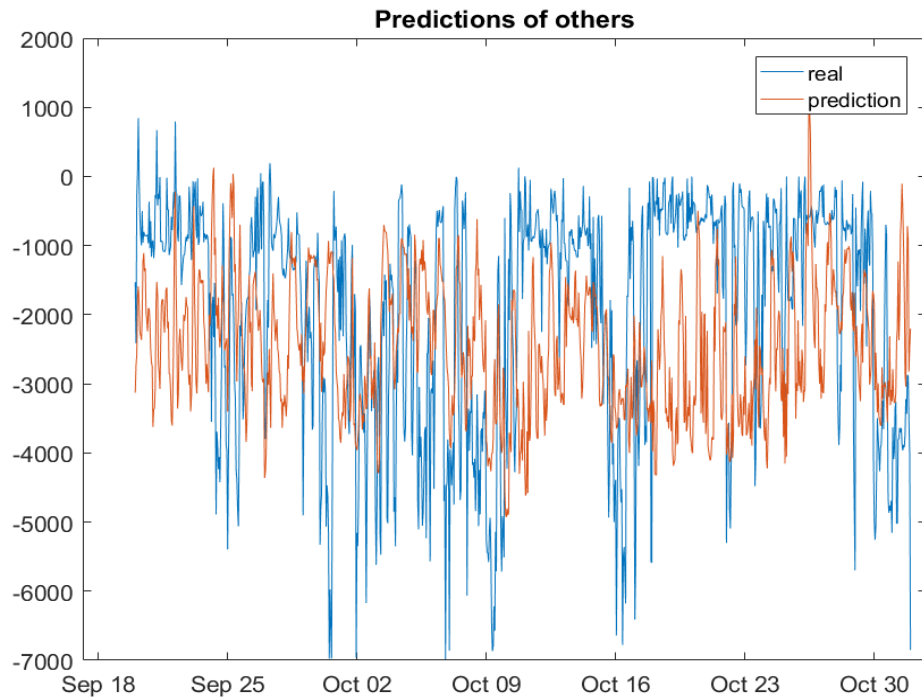
Παρακάτω είναι ένα δείγμα του GJR μοντέλου:





Κλείνοντας τα στατιστικά μοντέλα θα θέλαμε να παρουσιάσουμε τις επαγγελματικές προβλέψεις που μας έχουν δοθεί ώστε να παρατηρήσουμε την δική τους μορφή και τα σφάλματα τους. Το RMSE τους κυμαίνεται γύρω στα 1920, που αποτελεί αρκετά καλό σφάλμα αν αναλογιστούμε την δυσκολία του προβλήματος και το μέγεθος των τιμών. Ειδικότερα, όσο μεγαλώνουν οι τιμές και το εύρος τους, όλο και περισσότερο αυξάνεται και το ρίσκο μεγαλύτερου λάθους. Το εύρος των τιμών μας είναι αρκετά μεγάλο λόγω και του ότι επιτρέπουμε αρνητικές τιμές για να εκφράσουμε την αντίθετη κατεύθυνση του ηλεκτρικού φορτίου. Αυτό έχει ως αποτέλεσμα τον διπλασιασμό του εύρους, ενώ ήταν ήδη αρκετά μεγάλο.

Παρακάτω παραθέτω το διάγραμμα της πρόβλεψης:



Όπως μπορούμε να παρατηρήσουμε, τα μοντέλα linear και pure quadratic είχαν καλύτερα αποτελέσματα αλλά αυτό δεν είναι καθησυχαστικό γιατί μπορεί να ήταν τυχαίο σε αυτά τα δεδομένα. Για να είμαστε πιο σίγουροι ότι βρήκαμε κάτι που συμπεριφέρεται στις περισσότερες περιπτώσεις καλύτερα πρέπει να βρούμε ένα μοντέλο που να έχει όσον το δυνατό λιγότερα σφάλματα.

Παρακάτω θα εξετάσουμε τις ευφείς μεθόδους των πρόσθιας τροφοδότησης και αναδρομικών νευρωνικών αλλά και το SVR και τις παραλλαγές του, τα οποία θα τα συγκρίνουμε με όσα έχουμε παρατηρήσει προηγουμένως.



## 7.2 Ευφυείς Τεχνικές

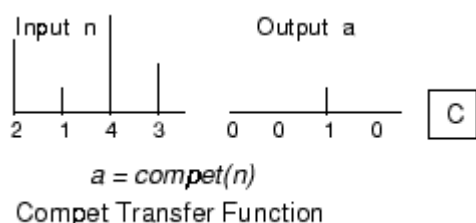
### 7.2.1 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης

Τα έχουμε χωρίσει σε τέσσερις κατηγορίες. Στις πρώτες τρεις χρησιμοποιούμε σαν είσοδο τις τιμές της προηγούμενης μέρας και είναι ενός επιπέδου, δύο επιπέδων, τριών επιπέδων ενώ στην τελευταία κατηγορία χρησιμοποιώ τις εξωτερικές μεταβλητές της ίδιας ώρας με την ώρα που προσπαθούμε να προβλέψουμε με νευρωνικό τριών επιπέδων.

Σε προηγούμενο κεφάλαιο αναφέραμε και επεξηγήσαμε τις συναρτήσεις εκπαίδευσης που θα χρησιμοποιήσουμε (`trainlm(Levenberg-Marquardt)`, `trainbr(Bayesian Regularization)`, `trainbfg(BFGS Quasi-Newton)`, `trainrp(Resilient Backpropagation)`, `trainscg(Scaled Conjugate Gradient)`, `traincbg(Conjugate Gradient with Powell/Beale Restarts)`, `traincgf(Fletcher-Powell Conjugate Gradient)`, `traincgp(Polak-Ribiere Conjugate Gradient)`, `trainoss(One Step Secant)`, `traingdx(Variable Learning Rate Gradient Descent)`, `traingdm(Gradient Descent with Momentum)`, `traingd(Gradient Descent)`), όπως αναφέρονται στη MATLAB και έτσι θα τις αναφέρουμε και έπειτα στη διπλωματική).

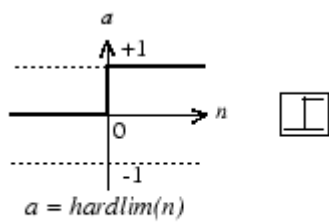
Εκτός αυτών δοκιμάσαμε τις εξής συναρτήσεις μεταφοράς:

- `Compet(Competitive transfer function)`: όπου επιστρέφει 1 στη στήλη που έχει την max τιμή και 0 αλλού.



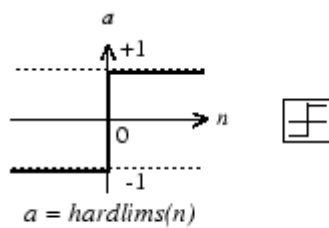
- `Elliposig(Elliot symmetric sigmoid transfer function)`: όπου συμπίπτει τα στοιχεία στο διάστημα [-1, 1] με μία σιγμοειδή μορφή συνάρτησης.

- Hardlim(Hard-limit transfer function):



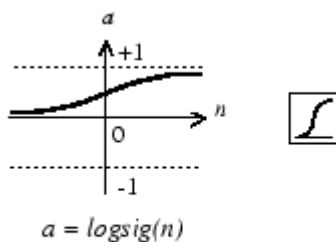
Hard-Limit Transfer Function

- Hardlims(symmetric Hard-limit transfer function):



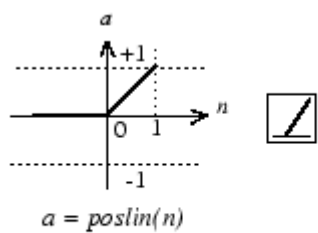
Symmetric Hard-Limit Transfer Function

- Logsig(Log-sigmoid transfer function):



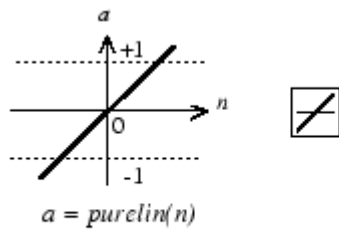
Log-Sigmoid Transfer Function

- Netinv(Inverse transfer function): επιστρέφει τον αντίστροφο αριθμό.
- Poslin(Positive linear transfer function):



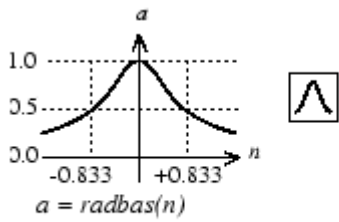
Positive Linear Transfer Function

- Purelin(Linear transfer function):



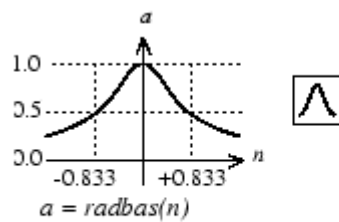
Linear Transfer Function

- Radbas(Radial basis transfer function):



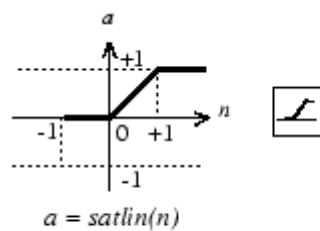
Radial Basis Function

- Radbasn(Normalized radial basis transfer function): η διαφορά με την προηγούμενη είναι ότι κανονικοποιούνται με την διαίρεσή τους με το άθροισμα των μη κανονικοποιημένων τιμών.



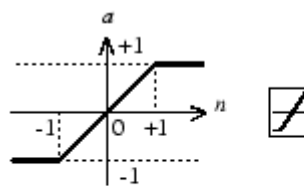
Radial Basis Function

- Satlin(Saturating linear transfer function):



Satlin Transfer Function

- Satlins(Symmetric saturating linear transfer function):

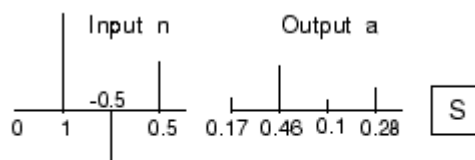


$$a = \text{satlins}(n)$$

Satlins Transfer Function

- Softmax(Soft max transfer function): Υπολογίζεται από τη παρακάτω συνάρτηση

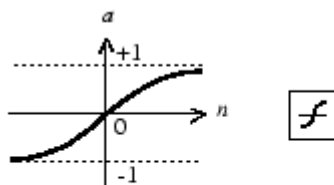
$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K.$$



$$a = \text{softmax}(n)$$

Softmax Transfer Function

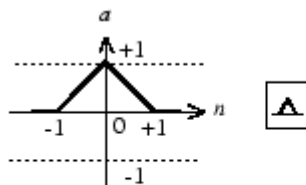
- Tansig(Hyperbolic tangent sigmoid transfer function):



$$a = \text{tansig}(n)$$

Tan-Sigmoid Transfer Function

- Tribas(Triangular basis transfer function):



$$a = \text{tribas}(n)$$

Triangular Basis Function

Χρησιμοποιήσαμε όλους τους δυνατούς συνδυασμούς μεταξύ των αλγορίθμων μάθησης και των συναρτήσεων μεταφοράς πάνω στη κάθε από τις τέσσερις

αρχιτεκτονικές νευρωνικών που περιγράψαμε πριν και πήραμε τα ακόλουθα αποτελέσματα:

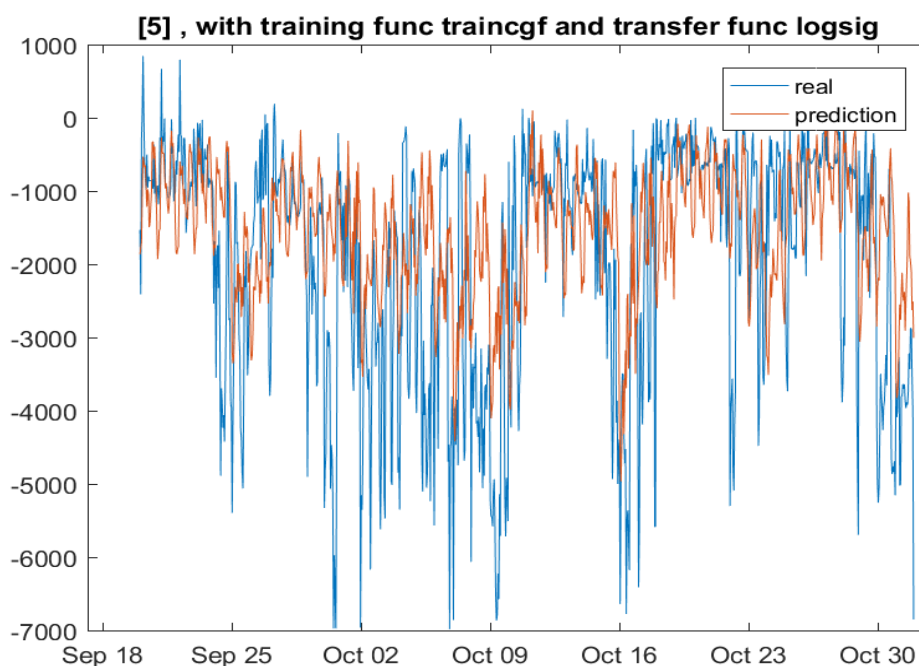
α)

Δοκιμάσαμε διαφορετικά νευρωνικά πρόσθιας τροφοδότησης με 1 επίπεδο και με 5, 10, 15, 18, 20, 24 νευρώνες.

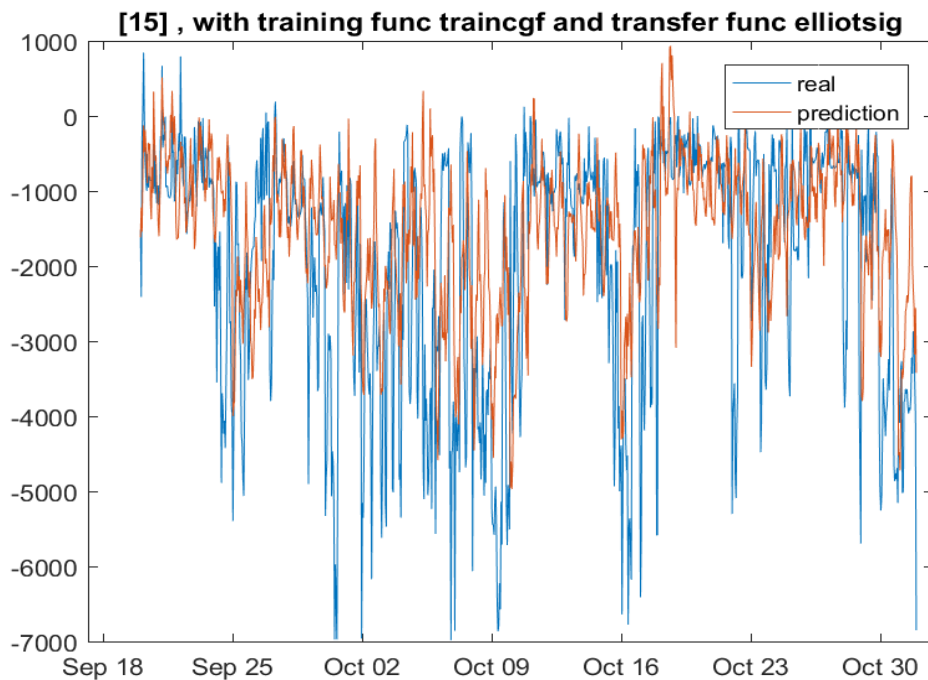
Παρατηρούμε ότι ο αλγόριθμος μάθησης παίζει έναν πάρα πολύ σημαντικό ρόλο στην απόδοση των προβλέψεων. Ειδικότερα, οι αλγόριθμοι `traincgf`, `traincgp`, `trainoss`, `trainidx` και `trainbfg` βρίσκονται στις υψηλότερες θέσεις με το μικρότερο σφάλμα που κυμαίνεται γύρω στα 6000-6200.

Από πλευράς συνάρτησης μεταφοράς οι πιο αποδοτικές που συνδυάστηκαν με τους παραπάνω αλγορίθμους είναι οι `logsig`, `elliotsig`, `softmax`, `satlins`. Γενικά, το `rmse` όλων των προβλέψεων κυμαίνεται από 6000 μέχρι  $3.7 \cdot 10^7$  ενώ το `wmae` κυμάνθηκε από -0.60 (που ήταν και η καλύτερη) μέχρι -785.7, που πρακτικά βλέπουμε εδώ ότι κάποιοι συνδυασμοί είναι τελείως άχρηστοι στο πρόβλημά μας.

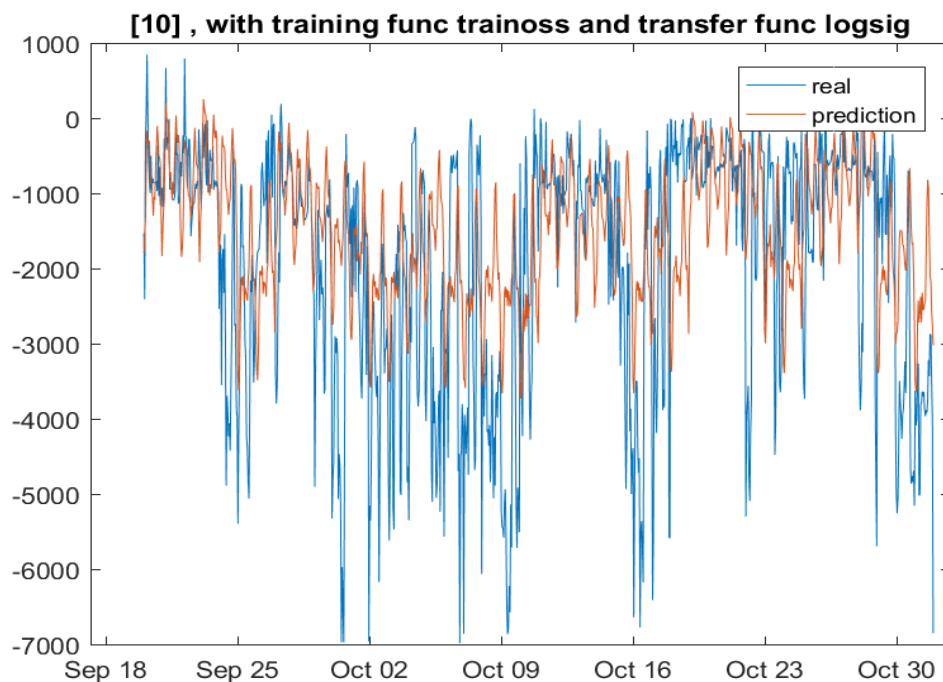
Παρακάτω θα παραθέσουμε ενδεικτικά διαγράμματα αυτών των πιο αποδοτικών συνδυασμών. Πάνω στο διάγραμμα θα φαίνεται ο αριθμός των νευρώνων κάθε φορά. Πρώτα θα παρατηρήσουμε τον αλγόριθμο `traincgf` με συναρτήσεις μεταφοράς τις `logsig` και `elliotsig`:



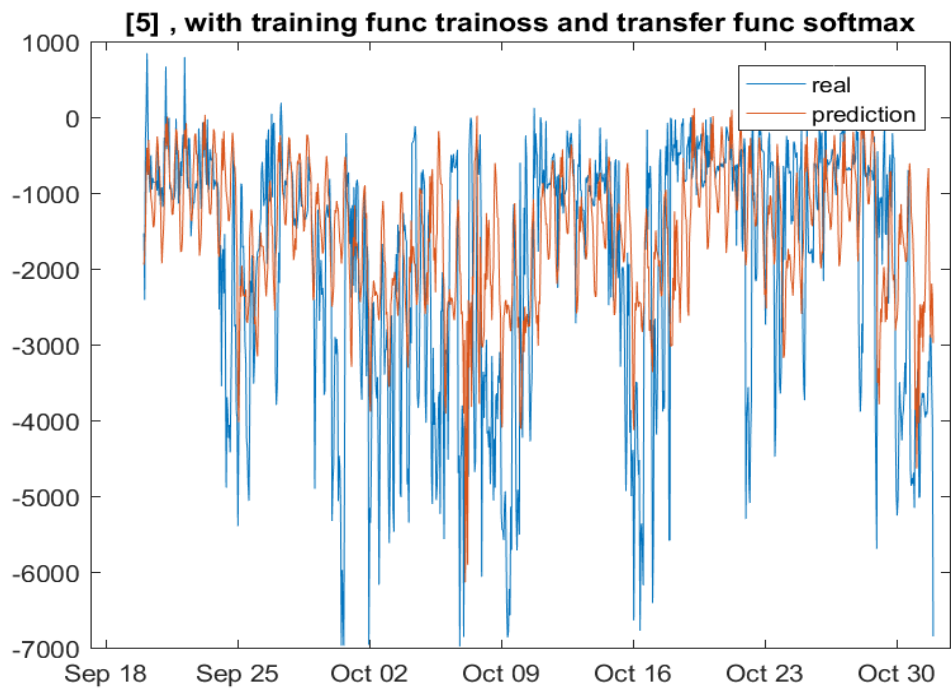
με  $rmse = 6023$  και  $wmae = -0.602$ , και είναι η καλύτερη πρόβλεψη από πλευράς και των δύο μετρικών σφάλματος. Η επόμενη περίπτωση είναι με την χρήση της `elliotsig`:



με  $rmse = 6080$  και  $wmae = -0.6041$  έρχεται δεύτερη στη κατάταξη με λιγότερο σφάλμα. Έπειτα, ο επόμενος αλγόριθμος με κάποια επιτυχία είναι ο `trainoss` με συναρτήσεις μεταφοράς τις `softmax` και `logsig`:

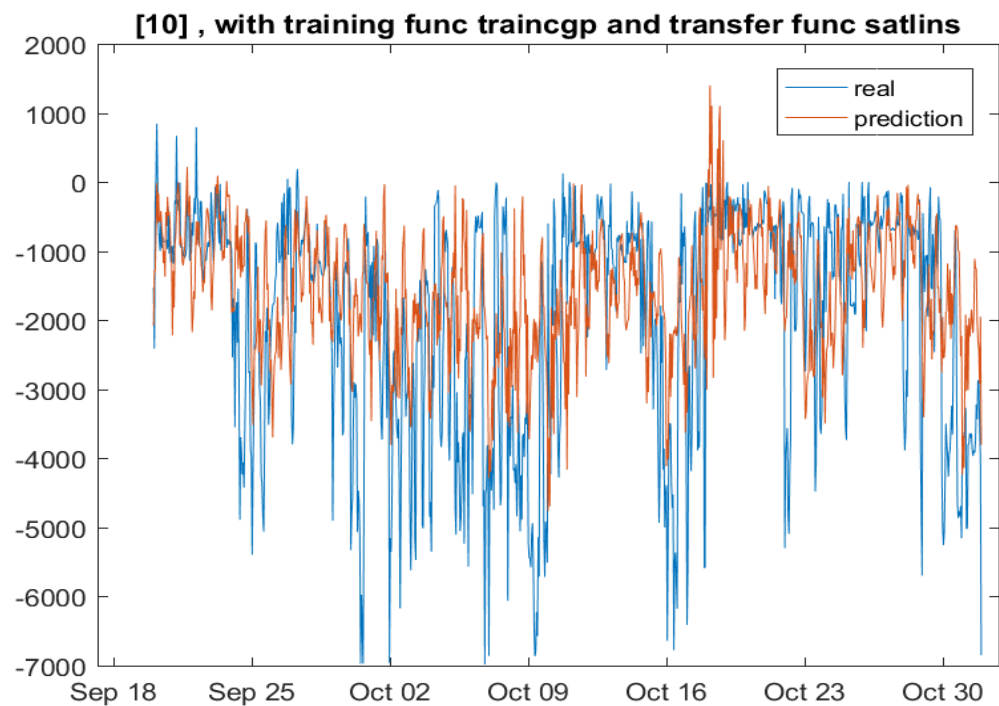


Με την logsig παρατηρούμε μία πολύ καλή πρόβλεψη συγκρίσιμη με τις προηγούμενες με  $rmse = 6090$  και  $wmae = -0.613$



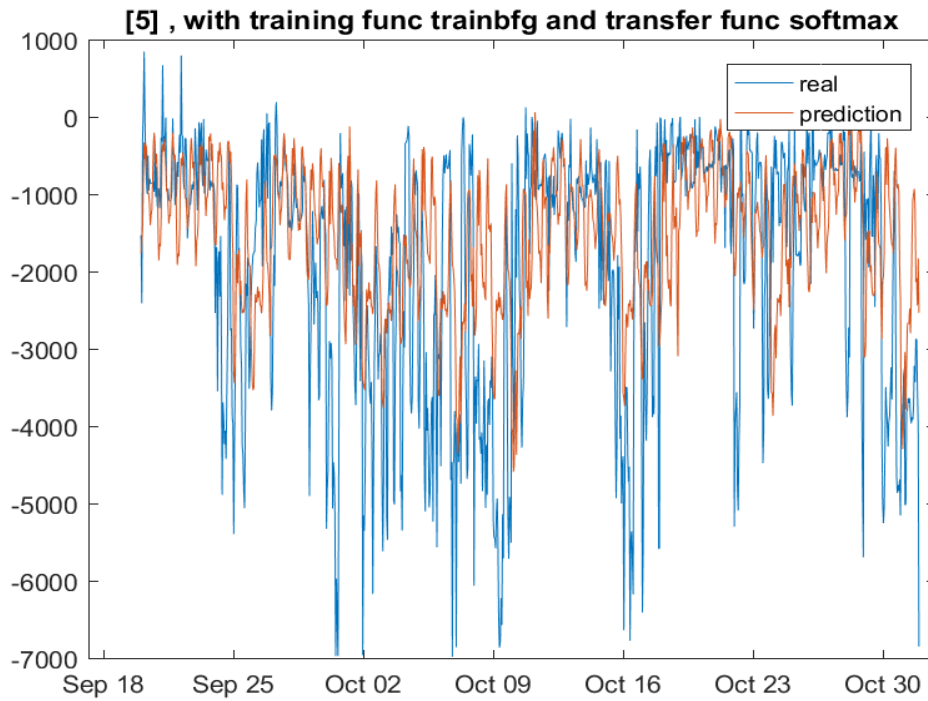
με  $rmse = 6106$  και  $wmae = -0.6108$

Παρακάτω βλέπουμε τον αλγόριθμο `traincgp`, οποίος απέδωσε αρκετά καλά:



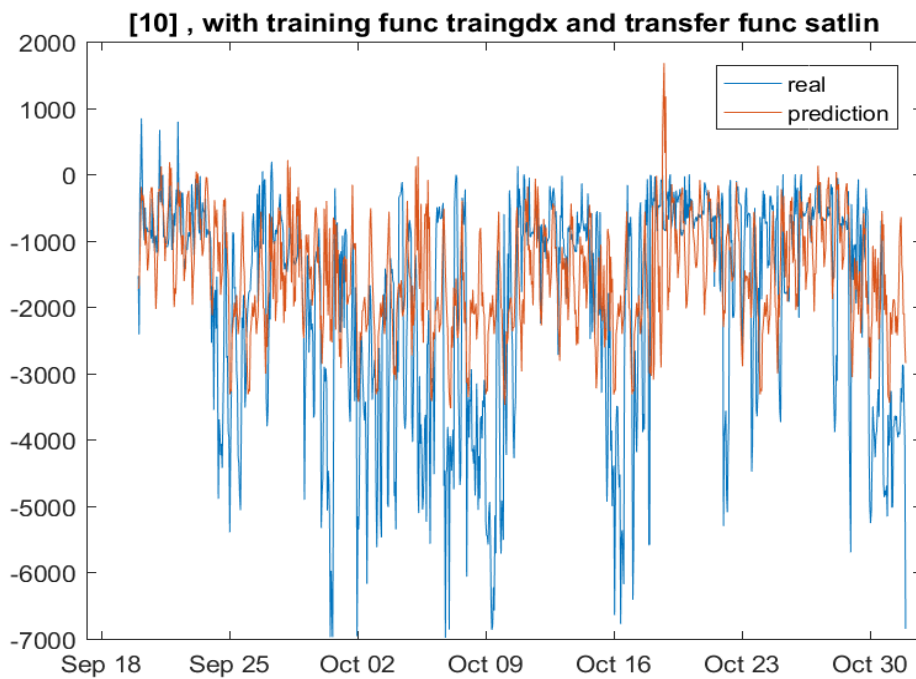
με  $rmse = 6091$  και  $wmae = -0.613$

Εδώ παρουσιάζουμε την καλύτερη περίπτωση του αλγορίθμου trainbfg:

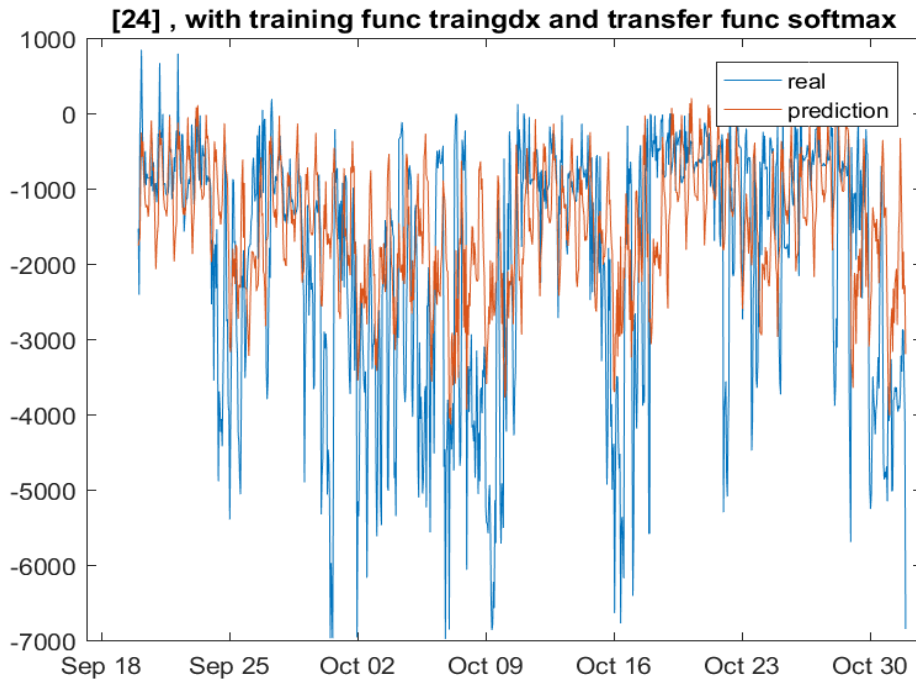


rmse = 6139 και wmae = -0.614

Παρακάτω βλέπουμε τις δύο καλύτερες προβλέψεις με τον αλγόριθμο traingdx:







με  $rmse = 6118$  και  $wmae = -0.62$  για το πρώτο και  $rmse = 6118$  και  $wmae = -0.612$  για το δεύτερο.

Μέχρι στιγμής όλα τα δείγματα των νευρωνικών έχουν ένα κοινό, ότι η συνολική τους μορφή είναι ικανοποιητική όπως και τα ξεσπάσματά τους. Αλλά δεν καταφέρνουν να έχουν μικρότερο σφάλμα από τις πιο αποδοτικές από τις στατιστικές τεχνικές. Επίσης, ο αριθμός των νευρώνων φαίνεται να μην παίζει τόσο σημαντικό ρόλο σε σχέση με τον αλγόριθμο μάθησης και τις συναρτήσεις μεταφοράς. Αυτό θα το επαληθεύσουμε και αργότερα όπου θα μεγαλώσει ο αριθμός των κρυφών επιπέδων.

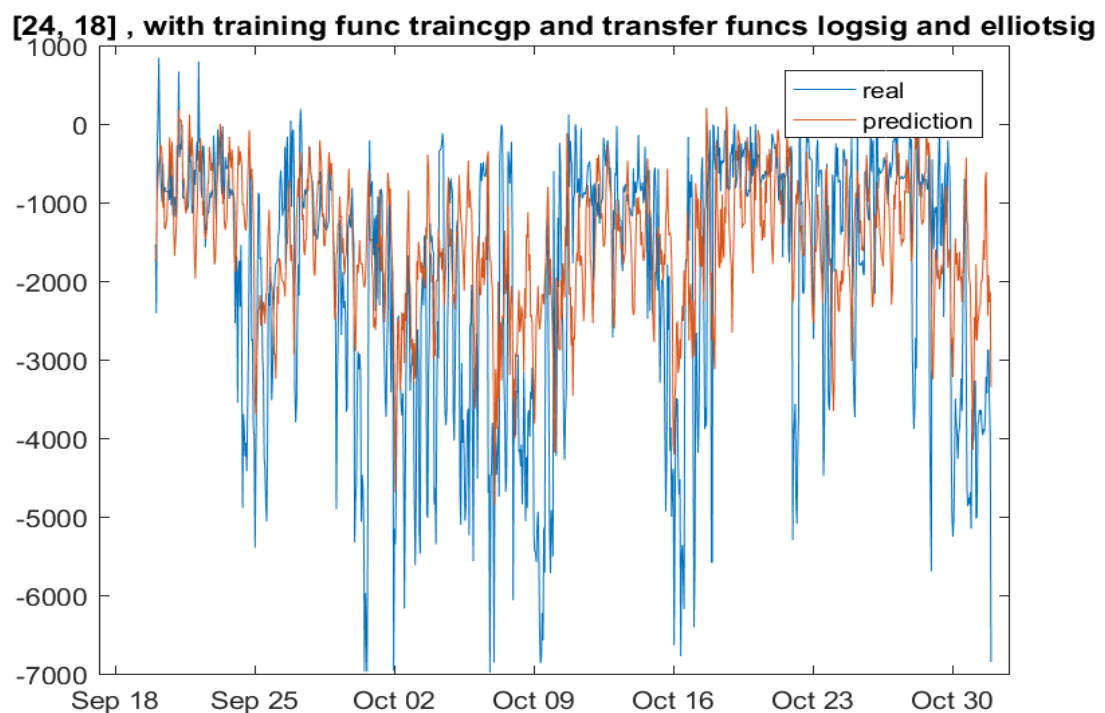
Είναι σημαντικό να επισημάνουμε και τους αλγορίθμους ή τις συναρτήσεις μεταφοράς που απέδωσαν λιγότερο από όλες τις άλλες και δεν θα ήταν συνετό να χρησιμοποιηθούν σε κάποια πρόβλεψη. Από τους αλγορίθμους μάθησης οι χειρότεροι είναι οι `traingd`, `traingdm` και `trainbr` με σφάλμα ( $rmse$ ) να ξεπερνά την τάξη  $10^4$  άσχετα με τον αριθμό των νευρώνων και την συνάρτηση μεταφοράς. Από συναρτήσεις μεταφοράς είναι προφανές ότι η `netinv` είναι τελείως ακατάλληλη γιατί σε όλα τα νευρωνικά που ήταν είχαν  $rmse$  της τάξης  $10^6 \sim 10^7$  και  $wmae$  από -11 μέχρι -700. Άλλες συναρτήσεις που δεν είχαν αρκετά καλά αποτελέσματα είναι οι

hardlims, radbasn και poslin, πράγμα που δείχνει ότι οι σιγμοειδείς συναρτήσεις ίσως ταιριάζουν περισσότερο στο πρόβλημά μας.

β)

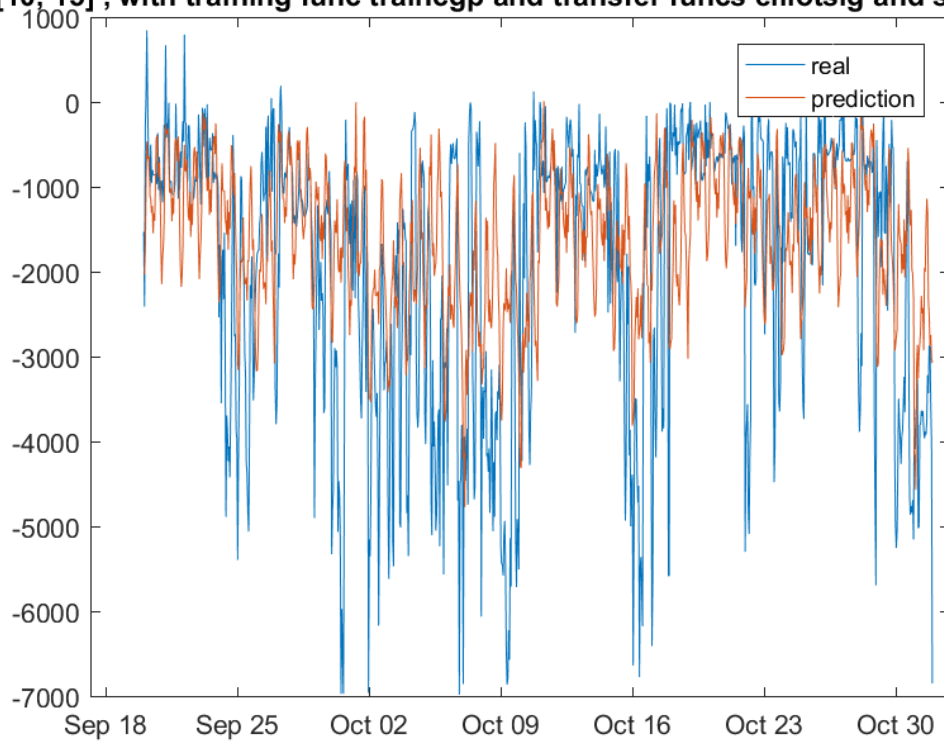
Εδώ θα εξετάσουμε τα νευρωνικά πρόσθια τροφοδότησης με 2 επίπεδα. Ο αριθμός των νευρώνων σε κάθε επίπεδο είναι μεταξύ των αριθμών 5, 10, 15, 18, 20, 24. Σαν συναρτήσεις εκπαίδευσης δοκίμασα μόνο τις πιο αποδοτικές που πήρα από τα προηγούμενα πειράματα (traincgf, trainoss, traincgp, traingdx) και το ίδιο και τις συναρτήσεις μεταφοράς (elliotsig, logsig, satlins, softmax, tansig).

Σαν γενική παρατήρηση που έχω είναι ότι τα σφάλματα δεν μειώνονται ιδιαίτερα και πολλές φορές δεν είναι αισθητά σε σχέση με τα νευρωνικά του ενός επιπέδου αλλά υπάρχουν διαφορές. Ακολουθούν τα καλύτερα δείγματα προβλέψεων:



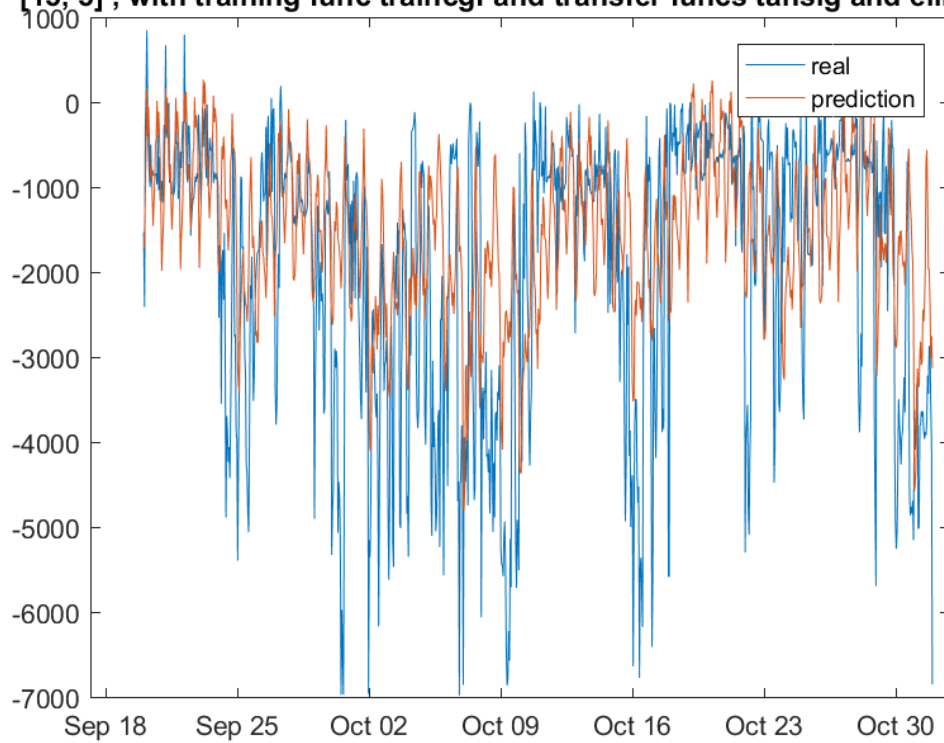
με  $rmse = 5973$  και  $wmae = -0.598$

**[10, 15] , with training func traincgp and transfer funcs elliotsig and softmax**

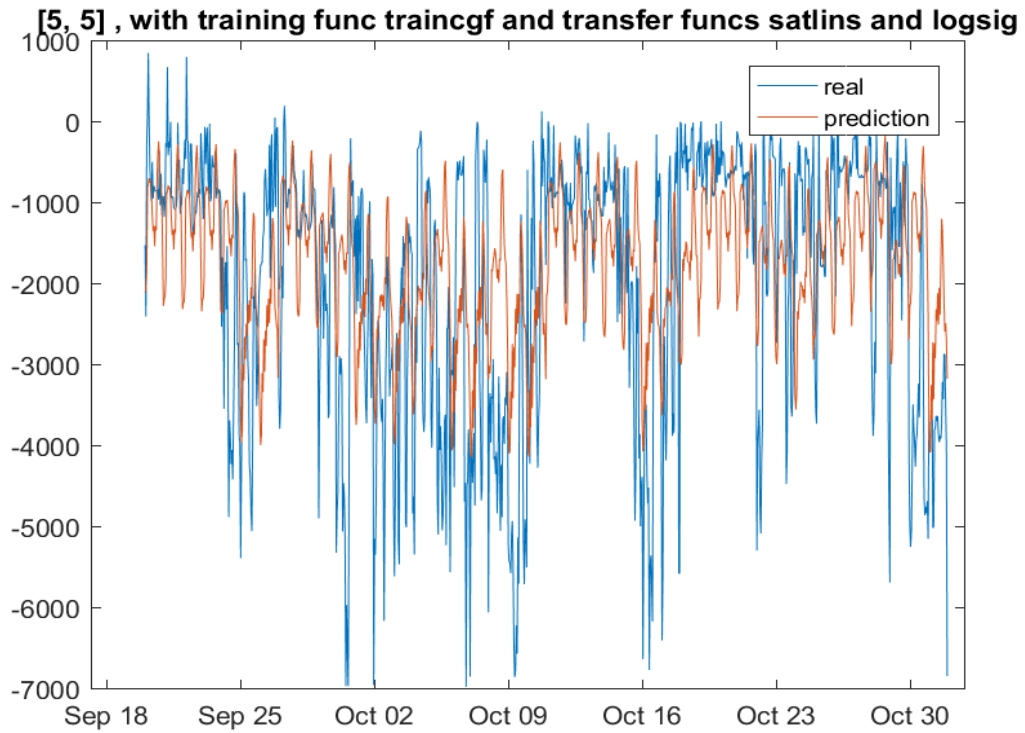


$\mu\epsilon$  rmse = 5991 και wmae = -0.606

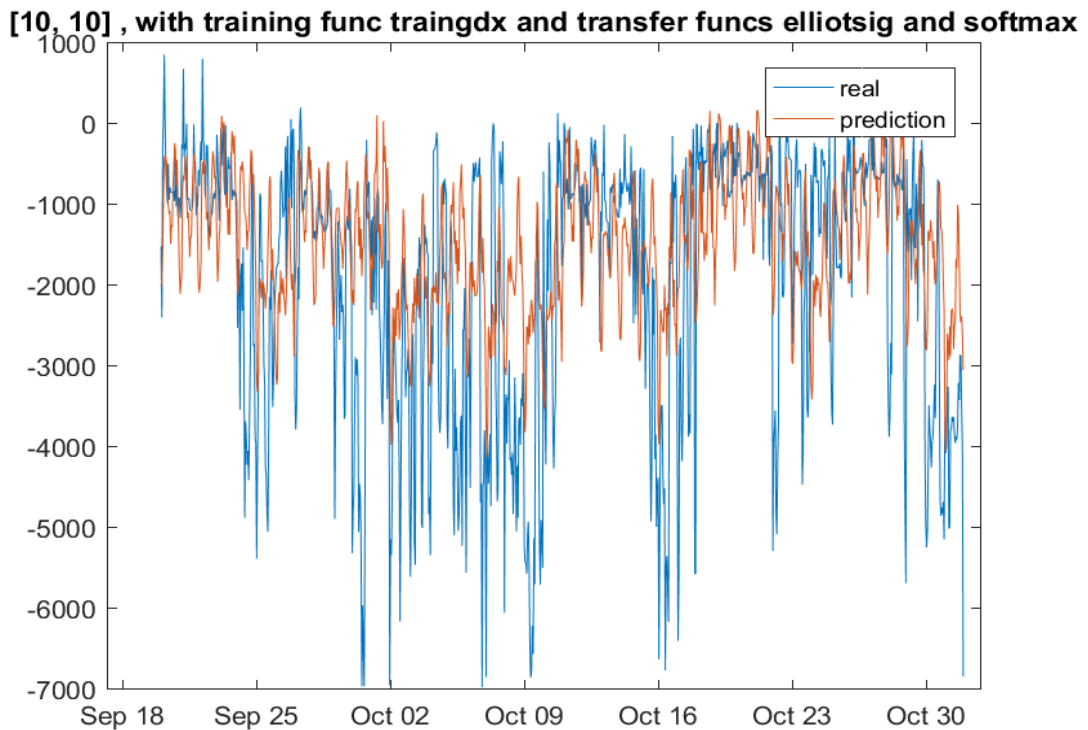
**[15, 5] , with training func traincgf and transfer funcs tansig and elliotsig**



$\mu\epsilon$  rmse = 6006 και wmae = -0.60



$\mu\epsilon$  rmse = 6032 και wmae = -0.622



$\mu\epsilon$  rmse = 6051 και wmae = -0.606

Από πλευράς σφάλματος παρατηρούμε κάποιες βελτιώσεις, βέβαια όχι ικανοποιητικές. Η μορφή της παράστασης δεν έρχεται πιο κοντά στο πραγματικό και αυτό το καθιστά λιγότερο αποδοτικό σε σχέση με προηγούμενα μοντέλα.

γ)

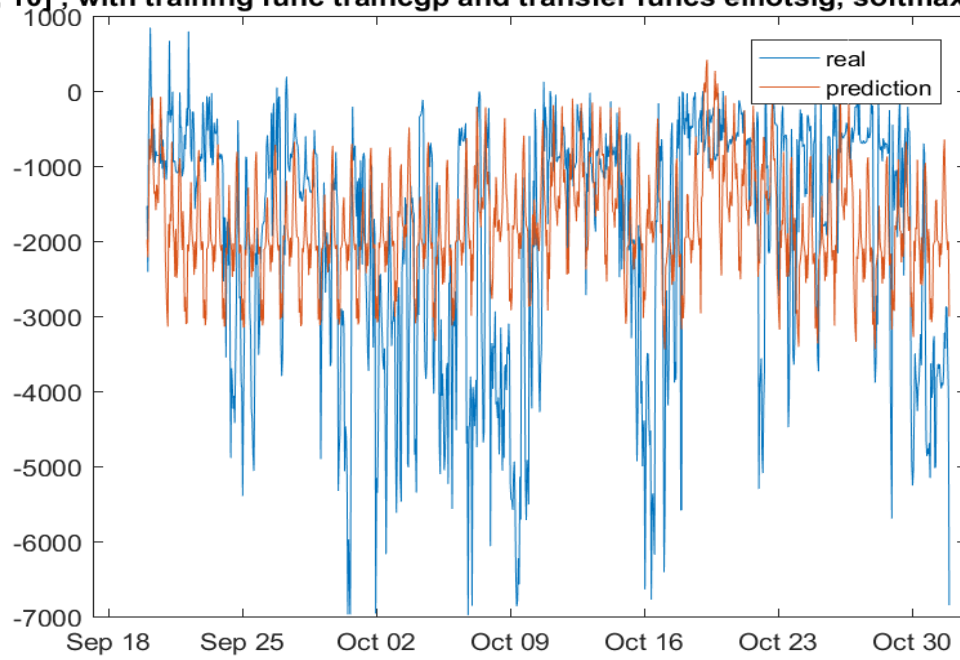
Εδώ θα εξετάσουμε τα νευρωνικά πρόσθιας τροφοδότησης με τρία επίπεδα. Στις γραφικές παραστάσεις δε παρατηρούμε καμία σημαντική αλλαγή σε σχέση με τα νευρωνικά με 2 επίπεδα αλλά μπορούμε να δούμε χαρακτηριστικές αλλαγές στα σφάλματα όπου πέφτουμε σε rmse της τάξης 5900 που είναι το πιο αποδοτικό μέχρι στιγμής στα νευρωνικά. Ενδεικτικά, να αναφέρουμε ότι ο αλγόριθμος traincgr απέδωσε αρκετά καλύτερα από όλους τους υπόλοιπους κρατώντας τις πρώτες θέσεις με τα λιγότερα σφάλματα (rmse και wmae). Αμέσως μετά έρχεται η traincgf με αρκετά κοντινές επιδόσεις.

Όσον αφορά τις συναρτήσεις μεταφοράς στο πρώτο επίπεδο, ανεξάρτητα με τη συνάρτηση εκπαίδευσης, φαίνεται να είναι πιο αποδοτικές οι satlins, elliotssig και tansig. Στο δεύτερο επίπεδο η softmax φαίνεται να επικρατεί με τις satlins και elliotssig να ακολουθούν. Τέλος, στο τρίτο και τελευταίο επίπεδο υπάρχει μεγαλύτερη ποικιλία με όλες τις προηγούμενες να επικρατούν στις πιο αποδοτικές επιλογές, αλλά αυτή που είχε πιο καλά αποτελέσματα ήταν η logsig.

δ)

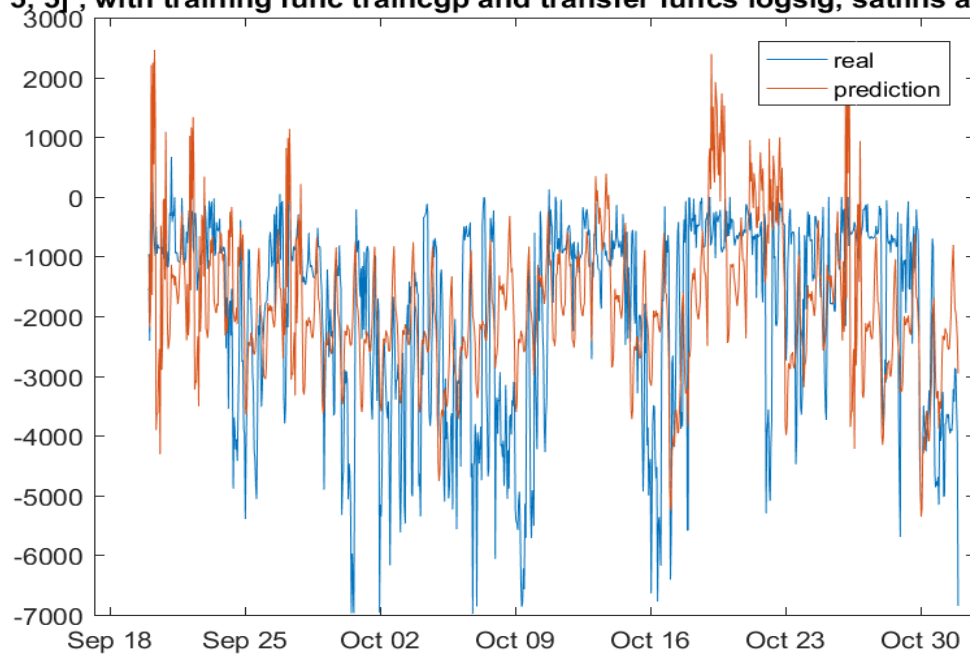
Εδώ θα εξετάσουμε νευρωνικά πρόσθιας τροφοδότησης με είσοδο τα επιπλέον δεδομένα σε συνδυασμό με τις τιμές της προηγούμενης μέρας με αρχιτεκτονική 3 επιπέδων. Παρακάτω μπορούμε να δούμε τις τρεις καλύτερες περιπτώσεις και πως συμπεριφέρονται:

**[0, 5, 10]**, with training func traincgp and transfer funcs elliotsig, softmax and so



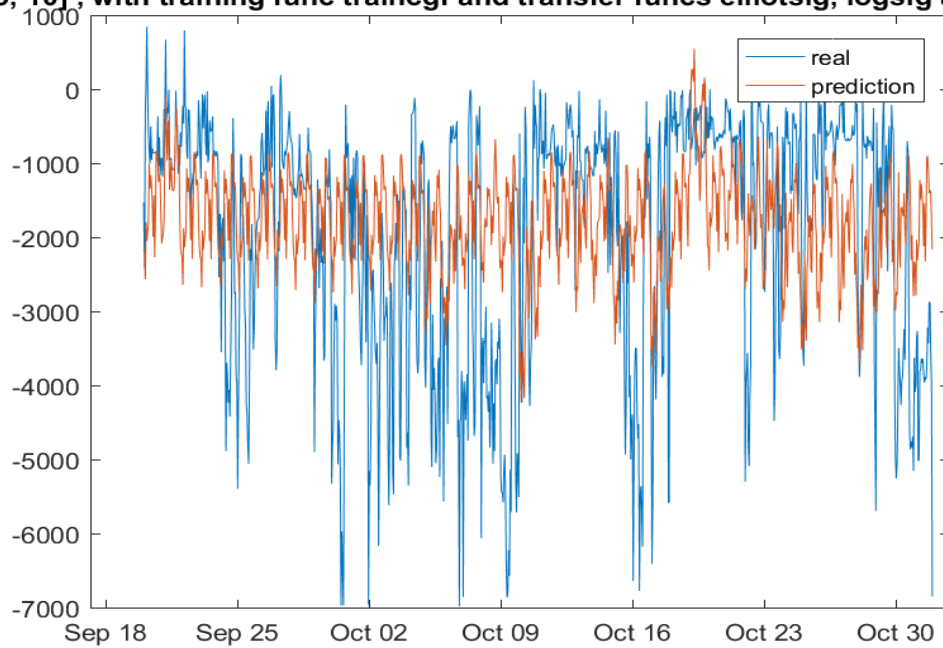
$\mu\epsilon$  rmse = 6348 και wmae = -0.674

**[5, 5, 5]**, with training func traincgp and transfer funcs logsig, satlins and satlii



$\mu\epsilon$  rmse = 6392 και wmae = -0.691

25, 5, 10], with training func traincgf and transfer funcs elliotsig, logsig and soft



με  $rmse = 6449$  και  $wmae = -0.678$

Κλείνοντας το κομμάτι των δικτύων πρόσθιας τροφοδότησης υπάρχουν πολλές παρατηρήσεις που απορρέουν από την εξέταση των διαγραμμάτων και των σφαλμάτων της κάθε πρόβλεψης. Μπορούμε να αποφανθούμε ότι ο αριθμός των νευρώνων στα αντίστοιχα επίπεδα δεν έχει μεγάλη σημασία, με την έννοια του ότι δεν επηρεάζει τόσο την τελική πρόβλεψη όσο οι υπόλοιπες παράμετροι. Ο αλγόριθμος μάθησης σε συνδυασμός με τις συναρτήσεις μεταφοράς κάνουν την μεγαλύτερη διαφορά. Όμως, όταν συγκρίνουμε τα αποτελέσματα με τα προηγούμενα αποτελέσματα των στατιστικών μεθόδων δεν υπάρχει σημαντική βελτίωση και όπου υπάρχει φαίνεται τυχαία. Αυτό φαίνεται στο ότι οι προβλέψεις από τα πρόσθια τροφοδότησης προβλέπουν τα spikes λίγο καλύτερα για αυτό μπορεί η πρόβλεψη να τύχει να είναι πιο κοντά σε μερικές χρονικές περιόδους, αλλά σαν συνολικό αποτέλεσμα, το οποίο φαίνεται και από τις μετρικές σφαλμάτων, οι πιο αποδοτικές στατιστικές μέθοδοι είναι καλύτερες. Παρακάτω θα εξετάσουμε τα αναδρομικά νευρωνικά δίκτυα και θα συγκρίνουμε τις προβλέψεις τους με τα πρόσθια τροφοδότησης.

## 7.2.2 Αναδρομικά Νευρωνικά Δίκτυα

Σε αυτή την ενότητα θα εξετάσουμε τα αναδρομικά νευρωνικά δίκτυα.

Χρησιμοποιήσαμε τα αναδρομικά δίκτυα της MATLAB, τα `nnnet`. Δοκιμάσαμε δύο κατηγορίες `nnnet`, ενός επιπέδου και δύο επιπέδων. Το μοντέλο `nnnet` δέχεται σαν είσοδο μία μεταβλητή καθυστέρησης(`delay`) όπου καθορίζει πόση θα είναι η καθυστέρηση που θα εφαρμόζεται στα δεδομένα. Αυτήν την καθυστέρηση την θέσαμε να είναι μία από τις ακόλουθες τιμές κάθε φορά, 5, 10, 15, 20 και 24. Ο αριθμός των νευρώνων βρίσκεται ανάμεσα στις επόμενες τιμές, 5, 10, 15, 20 και 25. Ένα άλλο χαρακτηριστικό των νευρωνικών `nnnet` είναι η επιλογή για το αν θα είναι κλειστό το κύκλωμα της ανατροφοδότησης ή όχι. Υπάρχουν τρεις επιλογές:

- 1) ανοιχτός βρόγχος(`open loop`)
- 2) κλειστός βρόγχος(`closed loop`)
- 3) κανένας βρόγχος(`none loop`)

Τέλος, σαν αλγόριθμο μάθησης δοκιμάσαμε όλους τους ακόλουθους:  
`trainlm`, `trainrp`, `trainsecg`, `traincgb`, `traincgf`, `traincgp`, `trainoss`, `traingdx`

Τα αποτελέσματα, δυστυχώς, απέδειξαν τα λεγόμενα ότι τα αναδρομικά δίκτυα δεν είναι φτιαγμένα για πρόβλεψη χρονοσειρών. Φαίνεται ότι ο κλειστός βρόγχος δεν είναι ικανό να παράγει κοντινές τιμές στις πραγματικές αλλά και οι άλλες δύο επιλογές, αν και καλύτερες, καταλήγουν σε μεγάλα σφάλματα. Αυτά τα σφάλματα ανήκουν στην τάξη της  $10^4$ , το οποίο επιβεβαιώνει την ακαταλληλότητα των μοντέλων αυτών.

Παρόλα αυτά μπορούμε να παρατηρήσουμε ότι οι αλγόριθμοι `traincgb`, `traincgf`, και `traincgp` βρίσκονται πρώτοι στη λίστα με μικρότερα σφάλματα. Πράγμα που επιβεβαιώνει ότι αυτοί οι αλγόριθμοι είναι οι πιο αποδοτικοί στο πρόβλημα που εξετάζουμε αφού είχαν τις καλύτερες προβλέψεις άσχετα με το είδος του νευρωνικού δικτύου αλλά και της αρχιτεκτονικής του.

Παρακάτω παραθέτω ενδεικτικά τις καλύτερες είκοσι αρχιτεκτονικές από δίκτυα `nnnet` ενός επιπέδου και δύο επιπέδων και το αντίστοιχο `rmse` τους.



Ενός

επιπέδου:

1	2.9209e+04	'Narnet with 24 delays, [5,25] neurons, open loop and traincgb'
2	2.9210e+04	'Narnet with 24 delays, [5,5] neurons, open loop and traincgb'
3	2.9260e+04	'Narnet with 15 delays, [15,20] neurons, none loop and traincgb'
4	2.9304e+04	'Narnet with 24 delays, [20,5] neurons, open loop and traincgb'
5	2.9363e+04	'Narnet with 15 delays, [10,10] neurons, none loop and traincgb'
6	2.9367e+04	'Narnet with 20 delays, [5,20] neurons, open loop and traincgb'
7	2.9377e+04	'Narnet with 20 delays, [5,5] neurons, open loop and traincgb'
8	2.9398e+04	'Narnet with 10 delays, [5,10] neurons, open loop and traincgb'
9	2.9398e+04	'Narnet with 24 delays, [15,20] neurons, none loop and traincgb'
10	2.9400e+04	'Narnet with 20 delays, [10,5] neurons, none loop and traincgb'
11	2.9412e+04	'Narnet with 15 delays, [5,25] neurons, none loop and traincgb'
12	2.9436e+04	'Narnet with 20 delays, [5,15] neurons, none loop and traincgb'
13	2.9439e+04	'Narnet with 10 delays, [15,20] neurons, none loop and traincgb'
14	2.9449e+04	'Narnet with 10 delays, [5,5] neurons, open loop and traincgb'
15	2.9452e+04	'Narnet with 20 delays, [5,10] neurons, none loop and traincgb'
16	2.9463e+04	'Narnet with 20 delays, [5,20] neurons, open loop and traincgb'
17	2.9471e+04	'Narnet with 24 delays, [20,10] neurons, open loop and traincgb'
18	2.9472e+04	'Narnet with 20 delays, [20,20] neurons, none loop and traincgb'
19	2.9485e+04	'Narnet with 10 delays, [5,10] neurons, none loop and traincgb'
20	2.9489e+04	'Narnet with 15 delays, [10,5] neurons, none loop and traincgb'

Δύο επιπέδων:

1	2.9429e+04	'Narnet with 10 delays, 5 neurons, none loop and traincgb'
2	2.9446e+04	'Narnet with 10 delays, 5 neurons, none loop and traincgp'
3	2.9477e+04	'Narnet with 24 delays, 10 neurons, open loop and traincgp'
4	2.9494e+04	'Narnet with 15 delays, 5 neurons, none loop and traincgp'
5	2.9499e+04	'Narnet with 10 delays, 10 neurons, open loop and traincgb'
6	2.9504e+04	'Narnet with 15 delays, 5 neurons, open loop and trainscg'
7	2.9523e+04	'Narnet with 10 delays, 5 neurons, open loop and traincgp'
8	2.9533e+04	'Narnet with 24 delays, 5 neurons, open loop and traincgb'
9	2.9543e+04	'Narnet with 15 delays, 20 neurons, open loop and traincgb'
10	2.9546e+04	'Narnet with 20 delays, 10 neurons, none loop and trainlm'
11	2.9552e+04	'Narnet with 20 delays, 10 neurons, open loop and traincgp'
12	2.9569e+04	'Narnet with 10 delays, 20 neurons, open loop and trainlm'
13	2.9583e+04	'Narnet with 24 delays, 5 neurons, open loop and traincgp'
14	2.9584e+04	'Narnet with 10 delays, 15 neurons, none loop and traincgb'
15	2.9595e+04	'Narnet with 15 delays, 5 neurons, none loop and traincgb'
16	2.9606e+04	'Narnet with 24 delays, 5 neurons, none loop and traincgb'
17	2.9613e+04	'Narnet with 10 delays, 10 neurons, none loop and traincgp'
18	2.9613e+04	'Narnet with 20 delays, 15 neurons, none loop and trainlm'
19	2.9624e+04	'Narnet with 24 delays, 5 neurons, open loop and traincgp'
20	2.9626e+04	'Narnet with 24 delays, 5 neurons, open loop and trainlm'

### 7.3 Support Vector Machines

Σε αυτή την ενότητα θα ασχοληθούμε με τα SVM(Support Vector Machines) και ειδικότερα μία υποκατηγορία τους τα SVR, που εξειδικεύονται σε παλινδρόμηση, πράγμα που ταιριάζει απόλυτα στο πρόβλημα της πρόβλεψης ηλεκτρικού φορτίου. Οι παράμετροι που θα δοκιμάσουμε στο SVM είναι οι εξής:

α) Η Kernel Function, όπου μπορεί να είναι linear(γραμμική), rbf(Gaussian or Radial Basis Function) ή polynomial(πολυώνυμο τάξης p).

β) Kernel Scale: Αυτή η παράμετρος χωρίζει τα δεδομένα σύμφωνα με τον αριθμό που θα την ορίσουμε ή αν την ορίσουμε 'auto' χρησιμοποιεί μία ευριστική μέθοδο για να κάνει scale τα δεδομένα.

γ) Standardize: Είναι η παράμετρος που καθορίζει αν τα δεδομένα θα κανονικοποιηθούν σύμφωνα με τον μέσο όρο και την τυπική απόκλιση των σταθμισμένων στηλών των δεδομένων.

δ) Σαν Solver χρησιμοποιήσαμε τον SMO(Sequential Minimal Optimization) που χρησιμοποιείται για την βελτιστοποίηση που απαιτείται από το πρόβλημα.

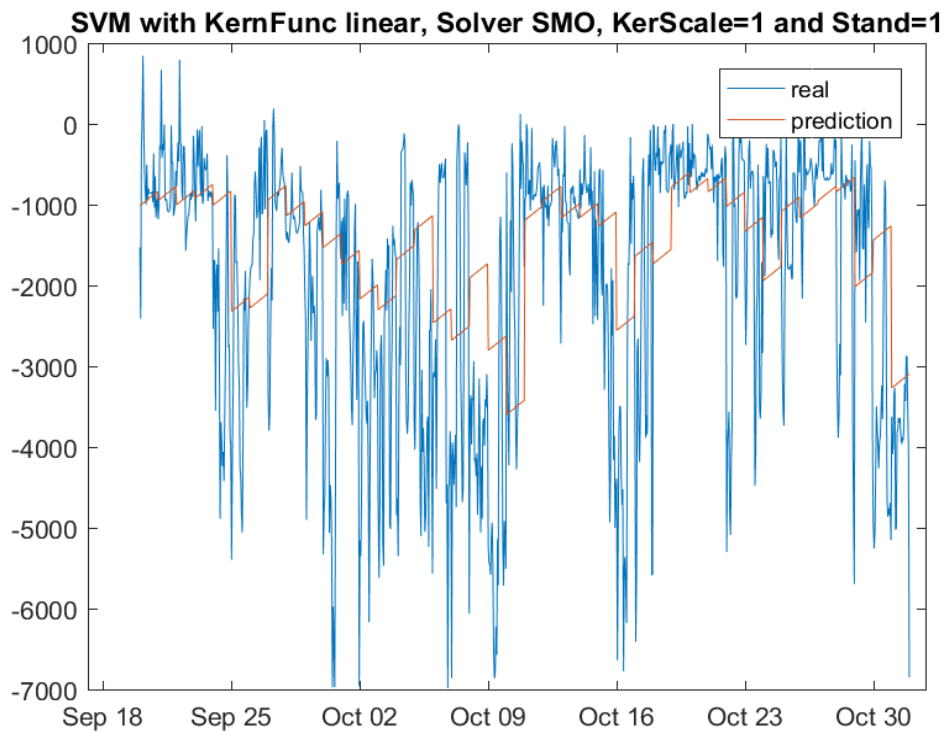
Σαν είσοδο χρησιμοποιήσαμε την ώρα, την μέρα τον μήνα που προσπαθούμε να προβλέψουμε αλλά και τις τιμές των προηγούμενων 24 ωρών. Με αυτά φτιάξαμε 20 διαφορετικά SVR μοντέλα και παρακάτω μπορούμε να δούμε τα σφάλματά τους:

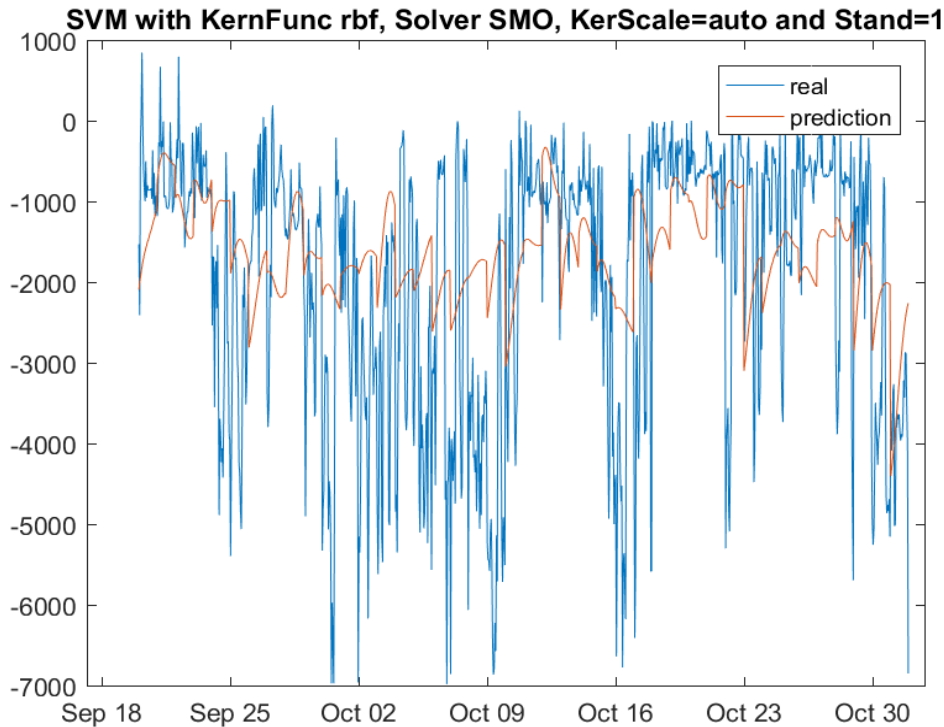
	1 RMSE	2 WMAE
1 SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1	1.5720e+03	-0.5922
2 SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=1 and Stand=1	1.5720e+03	-0.5922
3 SVM with KernFunc rbf, Solver SMO, KerScale=auto and Stand=1	1.5903e+03	-0.6417
4 SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=auto and Stand=0	1.6094e+03	-0.6049
5 SVM with KernFunc linear, Solver SMO, KerScale=auto and Stand=0	1.6097e+03	-0.6049
6 SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=auto and Stand=1	1.6157e+03	-0.6071
7 SVM with Polynomial KerFunc 2 order, Solver SMO, KerScale=auto and Stand=0	1.6200e+03	-0.5997
8 SVM with KernFunc linear, Solver SMO, KerScale=auto and Stand=1	1.6219e+03	-0.6086
9 SVM with Polynomial KerFunc 2 order, Solver SMO, KerScale=auto and Stand=1	1.6250e+03	-0.6069
10 SVM with KernFunc rbf, Solver SMO, KerScale=auto and Stand=0	1.7175e+03	-0.6608
11 SVM with Polynomial KerFunc 3 order, Solver SMO, KerScale=auto and Stand=1	1.7336e+03	-0.6293
12 SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1	1.7528e+03	-0.6718
13 SVM with Polynomial KerFunc 3 order, Solver SMO, KerScale=auto and Stand=0	1.8452e+03	-0.6649
14 SVM with Polynomial KerFunc 2 order, Solver SMO, KerScale=1 and Stand=1	1.8921e+03	-0.6961
15 SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=0	1.9019e+03	-0.6858
16 SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=1 and Stand=0	2.6630e+03	-1.0100
17 SVM with Polynomial KerFunc 3 order, Solver SMO, KerScale=1 and Stand=1	3.4343e+03	-1.1360
18 SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=0	3.8143e+03	-1.4237
19 SVM with Polynomial KerFunc 2 order, Solver SMO, KerScale=1 and Stand=0	3.4215e+04	-12.2489
20 SVM with Polynomial KerFunc 3 order, Solver SMO, KerScale=1 and Stand=0	8.6050e+04	-41.4597

Μπορούμε να παρατηρήσουμε ότι τα σφάλματα είναι τα μικρότερα μέχρι στιγμής από όλα τα μοντέλα που δοκιμάσαμε. Πρώτα σε απόδοση βρίσκονται οι γραμμικές συναρτήσεις πυρήνα (kernel), ενώ φαίνεται ότι η κανονικοποίηση των δεδομένων βοηθά. Από την άλλη οι πολυωνυμικές συναρτήσεις δευτέρου και μεγαλύτερου βαθμού δεν είναι τόσο ακριβείς στις προβλέψεις.

Από την άλλη η rbf συνάρτηση έχει κάποια καλά παραδείγματα αλλά φαίνεται ότι δεν ταιριάζει τόσο με μη κανονικοποιημένα δεδομένα.

Ακολουθούν μερικές από τις γραφικές παραστάσεις από τα καλύτερα μοντέλα SVR:





Μπορούμε πολύ εύκολα να παρατηρήσουμε ότι η μορφή των γραφικών παραστάσεων των προβλέψεων δεν μοιάζει με την μορφή των πραγματικών δεδομένων. Σε προηγούμενα μοντέλα, παρά το αρκετά μεγαλύτερο σφάλμα τους, η μορφή τους αν και λίγο ξέφευγε σε σημεία έμοιαζε αρκετά με την επιθυμητή. Τότε μπορεί να αναρωτηθεί κανείς πως είναι δυνατό να έχει το SVR καλύτερο σφάλμα. Αυτό συμβαίνει επειδή δεν προσπαθούμε να προβλέψουμε ολόκληρες συναρτήσεις ή γραφικές παραστάσεις αλλά απλά τα σημεία στο χώρο που αντιπροσωπεύουν τις τιμές του ηλεκτρικού φορτίου. Έτσι, το μοντέλο SVR μπορεί να έχει παράξενη μορφή και να φαίνεται πιο μακριά από την πραγματικότητα ενώ απλά ο τρόπος που προσεγγίζει τις επιθυμητές τιμές είναι διαφορετικός.

#### 7.4 SVRArima

Τέλος, προτού κλείσουμε με τις ευφύεις τεχνικές θα θέλαμε να παρουσιάσουμε την υλοποίηση του υβριδικού μοντέλου SVRArima, που αποτελείται από την προηγούμενη μέθοδο SVR και μία στατιστική, την τεχνική Arima, όπως εξηγήσαμε σε προηγούμενο κεφάλαιο. Δοκιμάσαμε τη ίδια αρχιτεκτονική του SVR μοντέλου όπως πριν, αλλά χρησιμοποιήσαμε το μοντέλο arima με παραμέτρους  $p = 12, 24$ ,

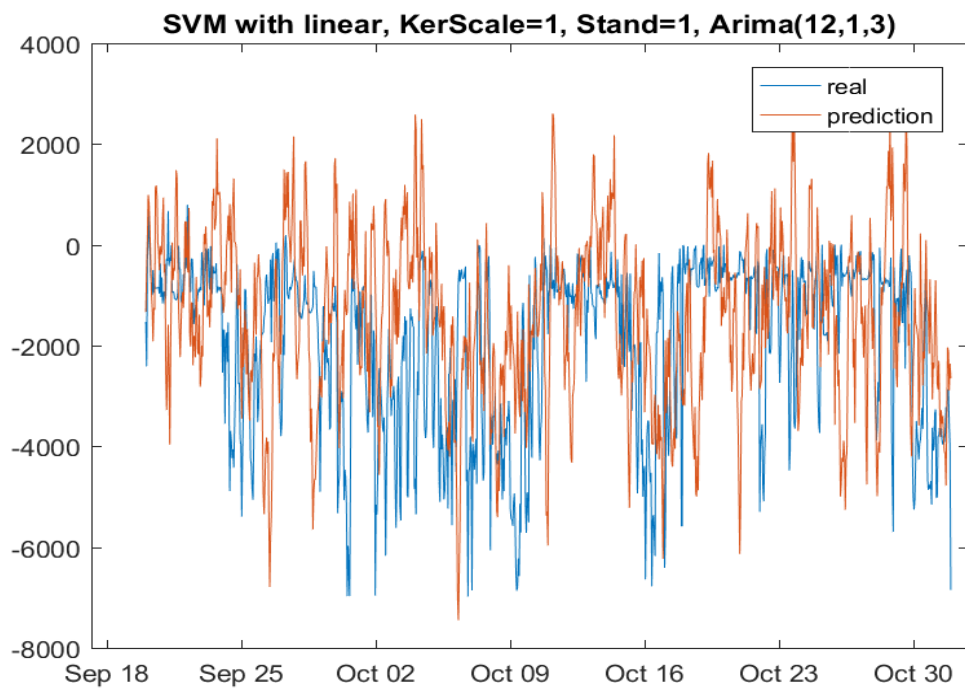
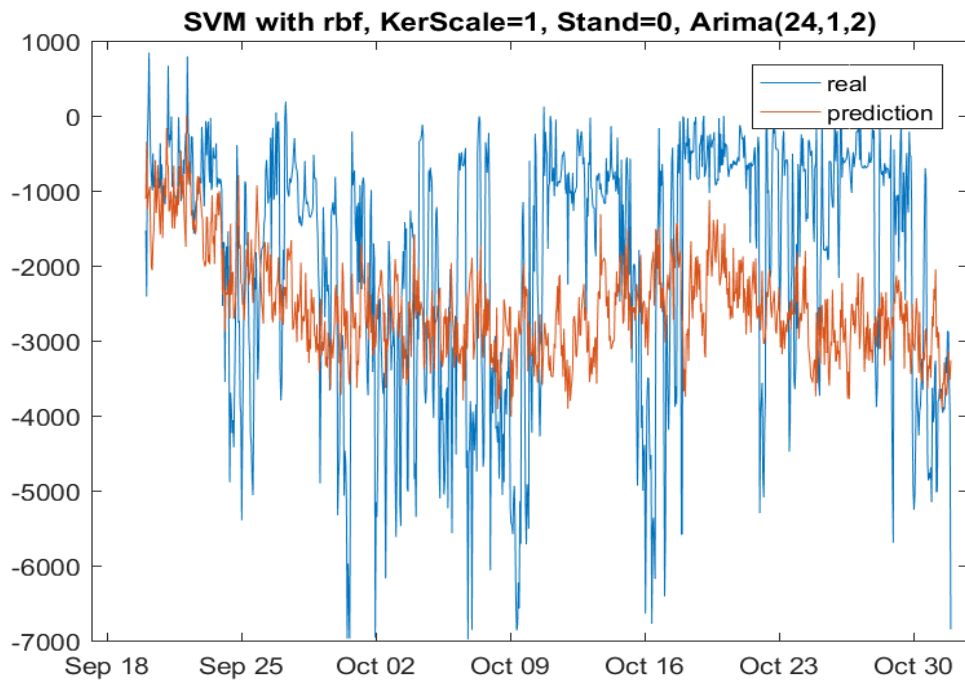
$D = 0,1$  και  $q = 1,2,3$  για να μοντελοποιήσουμε το σφάλμα που υποκύπτει το αρχικό μοντέλο.

Τα αποτελέσματα ήταν μικτά. Το συνολικό μεγάλωσε ελάχιστα χωρίς όμως να περάσει τις άλλες ευφυείς ή στατιστικές μεθόδους. Ενδιαφέρουσα παρατήρηση είναι ότι το σφάλμα rmse έδειξε ότι η συνάρτηση rbf είναι καλύτερη ενώ το σφάλμα wmae έδειξε ότι οι συναρτήσεις linear ή Polynomial είναι αρκετά πιο αποδοτικές. Αυτό μπορούμε να το δούμε στους δύο παρακάτω πίνακες που ο ένας είναι ταξινομημένος για να φαίνεται το ελάχιστο rmse ενώ αντίστοιχα ο δεύτερος για να φαίνεται το ελάχιστο wmae.

1	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=0 \newline Arima(24,1,2)'	1.7499e+03	-0.6858
2	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,1,3)'	1.8475e+03	-0.6718
3	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,0,2)'	1.8733e+03	-0.6718
4	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=0 \newline Arima(12,0,1)'	1.8734e+03	-0.6858
5	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=0 \newline Arima(12,1,1)'	1.8756e+03	-0.6858
6	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,1,3)'	1.8767e+03	-0.6718
7	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,1,2)'	1.8794e+03	-0.6718
8	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,0,1)'	1.8844e+03	-0.6718
9	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,0,3)'	1.8936e+03	-0.6718
10	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,0,2)'	1.8946e+03	-0.6718

1	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,1,3)'	2.1934e+03	-0.5922
2	'SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,1,3)'	2.3887e+03	-0.5922
3	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,1,1)'	2.9553e+03	-0.5922
4	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,0,2)'	3.2459e+03	-0.5922
5	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,0,1)'	3.2663e+03	-0.5922
6	'SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,0,2)'	3.2796e+03	-0.5922
7	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,1,3)'	3.2839e+03	-0.5922
8	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,0,1)'	3.3019e+03	-0.5922
9	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 \newline Arima(24,0,2)'	3.3222e+03	-0.5922
10	'SVM with Polynomial KerFunc 1 order, Solver SMO, KerScale=1 and Stand=1 \newline Arima(12,0,1)'	3.3349e+03	-0.5922

Μπορούμε να δούμε ξεκάθαρα ότι οι μετρικές έχουν διαφορές για το ποια μοντέλα αποδίδουν καλύτερα. Επίσης, παρακάτω παραθέτω δύο γραφικές παραστάσεις. Η πρώτη για το μικρότερο rmse σφάλμα και η δεύτερη αντίστοιχα για το wmae.



# Συμπεράσματα και Μελλοντικές Επεκτάσεις

## 1. Συμπεράσματα

Σε αυτό το κεφάλαιο θα αναφέρουμε όλα τα συμπεράσματα της διπλωματικής καθώς και πράγματα που θα επιθυμούσαμε να δοκιμάσουμε περαιτέρω πάνω στο πρόβλημα στο μέλλον.

Αρχικά, είναι εμφανές ότι τα μοντέλα που επεξεργάζονταν περισσότερες πληροφορίες είχαν συνήθως καλύτερα αποτελέσματα, δηλαδή πληροφορίες εκτός των τιμών φορτίου προηγούμενων ωρών. Τέτοιες πληροφορίες ήταν η κατανάλωση αλλά και παραγωγή ενέργειας από διάφορες πηγές. Εκτός αυτών, δεδομένα που βοήθησαν την μοντελοποίηση σαν μεταβλητές είναι η ώρα, η μέρα και ο μήνας. Ειδικότερα, θα αναφερθούμε στα μοντέλα που παρουσίασαν ενθαρρυντική συμπεριφορά αλλά και ποια μοντέλα φαίνεται ότι δεν απέδωσαν βέλτιστα.

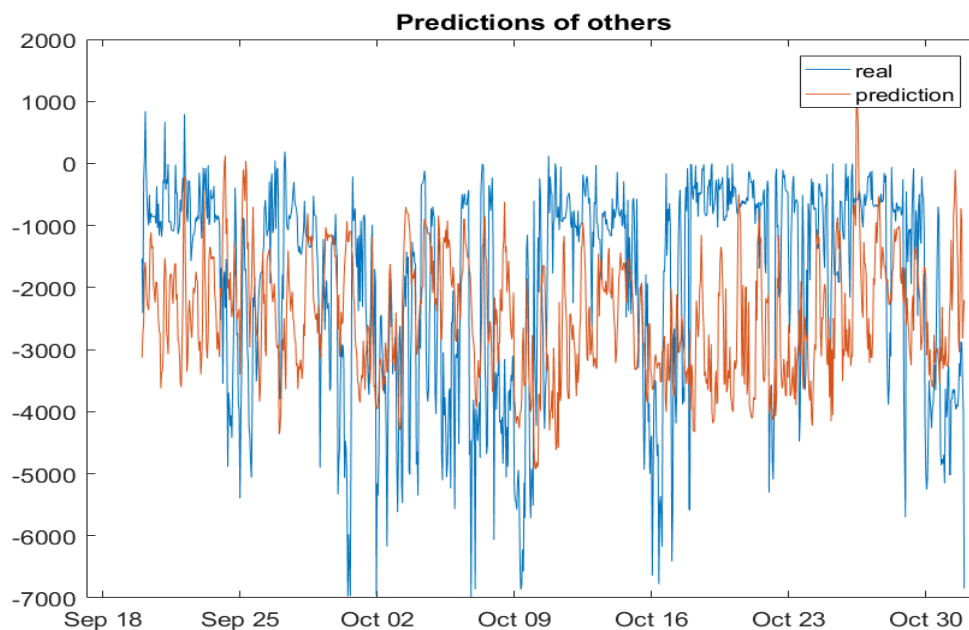
Πρώτα από όλα ξεχώρισαν τα μοντέλα SVM Regression και το υβριδικό με την συνεργασία του Arima. Σε αυτά παρατηρήσαμε τα μικρότερα σφάλματα ανάμεσα σε όλα τα μοντέλα που δοκιμάστηκαν. Το απλό SVR μοντέλο δεν παρουσίασε τόσο ικανοποιητική μορφή αλλά και το rmse και το wmae (γύρω στο 1570 και -0.59 αντίστοιχα) ήταν τα μικρότερα δυνατά σε αντίθεση με το SVRarima που είχε καλύτερη μορφή λόγω των ταλαντώσεων που πρόσφερε το μοντέλο arima αλλά μεγάλωσαν λίγο τα σφάλματα. Όμως, με αυτόν τον τρόπο ίσως είναι πιο εφικτό να προβλέψουμε τα απότομα ξεσπάσματα καλύτερα με την θυσία κάποιων έξτρα σφαλμάτων. Οι καλύτερες συναρτήσεις Kernel ήταν η γραμμική και η rbf.

Από την άλλη τα γραμμικά μοντέλα, κυρίως με τις συναρτήσεις linear και pure quadratic μας έδωσαν αρκετές σταθερές επιδόσεις καθιστώντας τα αρκετά χρήσιμα παρά την χαώδη φύση του προβλήματος. Το linear(χωρίς τα επιπλέον δεδομένα) είχε



το πρόβλημα του SVR, όπου πρόβλεπε κατά μέσο όρο σχετικά ακριβείς τιμές αλλά τα peaks δεν τα πρόβλεπε σωστά, έχοντας όμως μεγαλύτερο σφάλμα rmse αλλά και wmae. Από την άλλη τα γραμμικά μοντέλα με την συμβολή των εξωτερικών δεδομένων είχα μία καλύτερη προσέγγιση στα ξεσπάσματα χωρίς να υπάρχουν διαφορές στις τιμές των errors.

Συγκρίνοντας τα παραπάνω αποτέλεσμα με τις δοσμένες προβλέψεις από την Enscο μπορούμε να καταλάβουμε ότι υπάρχει κάποια βελτίωση.



Παραπάνω βλέπουμε τις προβλέψεις που μας δόθηκαν, με rmse γύρω στο 1923 και wmae γύρω στο -0.84, τα οποία κατατάσσονται πιο χαμηλά σε σχέση με τα προαναφερθέντα μοντέλα. Είναι απαραίτητο να προσέξουμε ότι αν και αυτές οι προβλέψεις έχουν ταλαντώσεις μεγάλου πλάτους και σαν αποτέλεσμα έχουν καλύτερες πιθανότητες να πιάσουν τις συχνές εκρήξεις των δεδομένων, κρύβουν μεγάλο ρίσκο γιατί είναι ευκολότερο να αποκλίνουν από τις υπόλοιπες τιμές περισσότερο, προκαλώντας μεγαλύτερο αθροιστικό σφάλμα. Αυτό μπορούμε να το κατανοήσουμε καλύτερα στα μοντέλα Garch, EGarch και GJR που είναι μοντέλα που έχουν απότομες αλλαγές και ανεβοκατεβάσματα.

Αυτά τα μοντέλα βασισμένα στην ετεροσκεδαστικότητα παρατηρήσαμε ότι ήταν ακατάλληλα για αυτό το πρόβλημα, επειδή οι γραφικές παραστάσεις τους παρέκλιναν αρκετά από το πραγματικό και αυτό φαίνεται και από τις μετρικές σφάλματος. Το rmse κυμαινόταν στην τάξη των  $10^6$  ενώ το wmae ξεπερνούσε τα -200, πράγμα που δείχνει ότι έχουν πολύ πιο συχνά και σε μεγάλου μεγέθους ξεσπάσματα.

Εκτός αυτών, μοντέλα που δεν απέδωσαν ικανοποιητικά ήταν και τα αναδρομικά νευρωνικά δίκτυα που αν και εμφανώς καλύτερα από τα ετεροσκεδαστικά μοντέλα παραμένουν μη αποδοτικά για το πρόβλημα της πρόβλεψης της ροής φορτίου. Αυτό αποδεικνύεται και από την μορφή των γραφικών παραστάσεών τους αλλά και από τα αρκετά μεγάλα σφάλματα (rmse και wmae) που παρουσίασαν με οποιαδήποτε αρχιτεκτονική κι αν δοκιμάστηκε.

Δύο μοντέλα που είχαν ομοιότητες ήταν τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης και το μοντέλο arima. Αρχικά, η μορφή τους ταιριάζει αρκετά γιατί αποτελείται από συνεχόμενα fluctuations που κυμαίνονται γύρω από τις μέσες τιμές των πραγματικών τιμών. Αυτό συνήθως θα οδηγούσε σε παρόμοιες τιμές σφαλμάτων αλλά εδώ συμβαίνει κάτι διαφορετικό. Το rmse του arima κυμαίνεται από 2660 μέχρι 3800 στις πιο αποδοτικές μορφές του, ενώ τα νευρωνικά πρόσθιας τροφοδότησης όλα έχουν πάνω από 6000 rmse. Παρατηρούμε δηλαδή σχεδόν διπλάσιο σφάλμα. Το πιο παράξενο είναι όμως ότι το wmae των κορυφαίων από τα πρόσθιας τροφοδότησης νευρωνικά βρίσκεται γύρω στα -0.60 που είναι εμφανώς καλύτερο από αυτό των μοντέλων arima, αλλά και ανταγωνίζεται τα SVR και γραμμικά μοντέλα για την πρώτη θέση, πράγμα που το rmse δεν το έδειχνε καθόλου. Όσον αφορά την αρχιτεκτονική των νευρωνικών έχουμε ότι με περισσότερα κρυμμένα επίπεδα και νευρώνες υπάρχει βελτίωση αλλά αρκετά μικρή. Από συναρτήσεις μάθησης διακρίνουμε τις traincgp και traincgf ενώ από συναρτήσεις μεταφοράς μπορούμε να δούμε ότι ξεχωρίζουν οι logsig, elliotssig και softmax. Από τα μοντέλα arima δεν παρατηρούμε κάποιο συγκεκριμένο πρότυπο που αποδίδει καλύτερα από κάποιο άλλο, αλλά συνήθως τα μοντέλα arima που παραγωγίζουν μία ή δύο φορές έχουν λιγότερο σφάλμα.

Πως μπορούμε να καταλάβουμε, λοιπόν, ποια από τις δύο μετρικές είναι πιο σωστή και αντιπροσωπεύει την πιο ακριβή πρόβλεψη. Αυτή η απάντηση είναι δύσκολη και

δεν υπάρχει πάντα μία ξεκάθαρη λύση. Συνήθως φαίνεται ότι το  $wmae$  είναι πιο ακριβές επειδή όταν δουλεύουμε με μεγάλα νούμερα, ο τετραγωνισμός που υφίστανται από το  $rmse$  μπορεί να μεγαλώνει το σφάλμα ενώ στην πραγματικότητα το ποσοστιαίο σφάλμα να είναι μικρότερο. Τέτοιες λεπτομέρειες καθιστούν το  $wmae$  καταλληλότερο σε πολλές περιπτώσεις. Επομένως, όταν υπάρχει μία παρατήρηση με πολύ μεγάλη διαφορά σε σχέση με τις υπόλοιπες το  $rmse$  την μετράει περισσότερο. Αυτό συνήθως είναι αρνητικό αλλά υπάρχουν περιπτώσεις που θέλουμε να ελέγχουμε τέτοιες μεγάλες τιμές και να βλέπουμε ποια μοντέλα τις παρουσιάζουν. Στην μελέτη των T.Chai και R.R.Draxler (Root mean square error RMSE or mean absolute error MAE? Arguments against avoiding RMSE in the literature) μπορούμε να δούμε το συγκεκριμένο πρόβλημα και να παρατηρήσουμε αυτές τις ιδιαιτερότητες μεταξύ  $mae$  και  $rmse$ . Σε αυτή την διπλωματική, όμως, χρησιμοποιούμε το  $wmae$ , που είναι το  $mae$  διαιρεμένο με τον μέσο όρο των πραγματικών τιμών της περιόδου που εξετάζουμε. Επιλέχθηκε το  $wmae$  επειδή είναι πιο αντιπροσωπευτικό σε διαφορετικά δεδομένα. Αυτό ισχύει επειδή μία ίδια τιμή σφάλματος σε διαφορετικά δεδομένα δε αντικατοπτρίζει πάντα την αλήθεια, δηλαδή ότι η απόδοση στις δύο διαφορετικές περιπτώσεις είναι ίδια. Για αυτό διαιρούμε με τον μέσο όρο των τιμών, για να είναι τα σφάλματα κανονικοποιημένα στα αντίστοιχα δεδομένα. Έτσι, κρατάμε τα θετικά του  $mae$  και προσθέτουμε και άλλα.

Όμως, το  $wmae$  έχει και αυτό τους κινδύνους του. Ο παρονομαστής του κλάσματος που δημιουργεί το  $wmae$  αποτελείται από τον μέσο όρο των πραγματικών τιμών στο διάστημα που προσπαθούμε να προβλέψουμε. Όταν οι τιμές είναι όλες ομόσημες τότε δεν υπάρχει κάποιο πρόβλημα. Όταν όμως υπάρχουν θετικές, αρνητικές αλλά και μηδενικές τιμές τότε εμφανίζεται το εξής πρόβλημα. Ο μέσος όρος του διαστήματος που εξετάζουμε μπορεί να τείνει στο 0. Εκτός λοιπόν από το προφανές αλλά σπάνιο πρόβλημα που είναι η διαίρεση με το μηδέν, ο παρονομαστής βρίσκεται κοντά στο μηδέν μεγαλώνοντας το συνολικό κλάσμα κατά πολύ. Αυτό μπορεί να αποκρύπτει τυχόν καλές προβλέψεις πίσω από έναν μικρό μέσο όρο που μεγαλώνει το  $wmae$  χωρίς λόγο. Για αυτό η μετρική του  $rmse$  δεν είναι απαραίτητα χειρότερη αλλά σε δεδομένα με θετικά και αρνητικά ίσως αποτελεί μία πιο ασφαλής λύση.

Παρακάτω θα παραθέσω δύο πίνακες με τα καλύτερα μοντέλα από την κάθε κατηγορία σύμφωνα με το rmse και σύμφωνα με το wmae αντίστοιχα. Εκεί που αναφέρει Predictions of others εννοούμε τις εξωτερικές προβλέψεις που μας δόθηκαν από την Ensco.

Καλύτερες επιδόσεις από την κάθε κατηγορία σύμφωνα με το rmse τους:

1	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1'	1.5720e+03
2	'Linear using previous 24h and Wind,Cons,... Data'	1.5936e+03
3	'Linear using previous 24h and Predicted Side Data'	1.6493e+03
4	'SVM with KernFunc rbf, Solver SMO, KerScale=1 and Stand=0 \ Arima(24,1,2)'	1.7499e+03
5	'Predictions of others'	1.9239e+03
6	'Arima(12,2,3) path 2'	2.6691e+03
7	'Neural Network with 3 hidden layers [20, 24, 5] , with training func traincgf and transfe...	5.8910e+03
8	'Neural Network with 2 hidden layer [24, 18] , with training func traincgp and transfer f...	5.9737e+03
9	'Neural Network with 1 hidden layer [5] , with training func traincgf and transfer func l...	6.0235e+03
10	'Neural Network with 3 hidden layers [20, 5, 10] , with training func traincgp and transf...	6.3487e+03
11	'Narnet with 24 delays, [5,25] neurons, open loop and traincgb'	2.9209e+04
12	'Narnet with 10 delays, 5 neurons, none loop and traincgb'	2.9429e+04
13	'Garch(24,1) path 1'	2.3663e+06
14	'Egarch(24,1) path 3'	2.4006e+06
15	'GJR(12,1) path 1'	2.4057e+06

Καλύτερες επιδόσεις από την κάθε κατηγορία σύμφωνα με το wmae τους:

1	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1'	-0.5922
2	'SVM with KernFunc linear, Solver SMO, KerScale=1 and Stand=1 Arima(12,0,1)'	-0.5922
3	'Neural Network with 2 hidden layer [24, 18] , with training func traincgp and transfer funcs logsig and elliotstsig'	-0.5988
4	'Neural Network with 1 hidden layer [5] , with training func traincgp and transfer func logsig'	-0.6029
5	'Linear using previous 24h and Predicted Side Data'	-0.6140
6	'Neural Network with 3 hidden layers [5, 5, 15] , with training func traincgp and transfer funcs elliotstsig, softmax and softmax'	-0.6593
7	'Pure Quadratic using previous 24h and Predicted Side Data'	-0.7229
8	'Predictions of others'	-0.8400
9	'Narnet with 10 delays, [5,20] neurons, open loop and traincgp'	-0.8764
10	'Narnet with 20 delays, 10 neurons, none loop and trainlm'	-0.9432
11	'Arima(12,2,3) path 2'	-1.1019
12	'Arima(24,1,1) path 1'	-1.2565
13	'Arima(24,2,3) path 3'	-1.2899
14	'Garch(24,3) path 3'	-579.2500
15	'Garch(24,1) path 1'	-611.2420
16	'GJR(12,1) path 1'	-649.8609
17	'GJR(12,3) path 2'	-654.8391
18	'Egarch(24,1) path 3'	-937.5718
19	'Egarch(24,1) path 2'	-979.6174

## 2. Μελλοντικές Επεκτάσεις

Σε αυτό το σημείο θα αναφέρουμε πράγματα που θα θέλαμε να δοκιμάσουμε στο μέλλον για την διερεύνηση της αντιμετώπισης του παρόντος προβλήματος. Είναι αδιαμφισβήτητο ότι η μεγαλύτερη ώθηση στην καλύτερη κατανόηση του προβλήματος είναι η πάροδος μεγάλου χρονικού διαστήματος με σκοπό την περισυλλογή περισσότερων δεδομένων. Με αυτό τον τρόπο θα μπορούσαμε να μελετήσουμε καλύτερα τις τάσεις με την πάροδο πολλών χρόνων και την συσχέτιση μεταξύ μηνών διαφορετικών ετών. Στην περίπτωσή μας το πρόβλημα βρίσκεται σε πρώιμο στάδιο και έχουμε μόνο δεδομένα διάρκειας ενός χρόνου. Μέσα σε ένα χρόνο δεν είναι δυνατό να παρατηρηθούν όλες οι τάσεις και τα ξεσπάσματα που κυριαρχούν σε ένα περίπλοκο όπως αυτό πρόβλημα.

Πέραν αυτού θα έχουμε και την δυνατότητα για περισυλλογή άλλων εξωτερικών δεδομένων, δηλαδή όχι δεδομένων ροής ηλεκτρικού φορτίου αλλά άλλων μεταβλητών που μπορεί να επηρεάζουν το πρόβλημά μας και να μην το γνωρίζουμε ή να μην έχουμε αντιληφθεί το μέγεθος της επιρροής τους.

Επίσης, θα ήθελα να εξετάσω περισσότερα υβριδικά μοντέλα και να δημιουργήσω έτσι περισσότερους συνδυασμούς κατανοώντας με αυτόν τον τρόπο πόσο κοντά

μπορεί η πρόβλεψη να φτάσει στην πραγματική τιμή. Μελετώντας τις συμπεριφορές υβριδικών μοντέλων μπορούμε να παρατηρήσουμε καλύτερα και τα ίδια τα μοντέλα καλύτερα επειδή μπορούμε να εξετάσουμε τα πλεονεκτήματά τους και τις αδυναμίες τους που μπορεί να γίνουν πιο εμφανείς.

Επιπλέον, υποσχόμενο μέλλον φαίνεται να έχουν και τα μοντέλα με ευφυείς πράκτορες αλλά και τα υβριδικά μοντέλα που απορρέουν από αυτά. Θα ήταν πολύ σημαντικό να δοκιμάσουμε τέτοιου είδους μοντέλα γιατί έχουν να προσφέρουν καινούργια πράγματα στην επίλυση του προβλήματος.

## Βιβλιογραφία

1. Rafal Weron (2014), Electricity price forecasting: A review of the state-of-the-art with a look into the future.
2. Alberto Cruz, Antonio Munoz, Juan Luis Zamora, Rosa Espinola (2011), The effect of wind generation and weekday on Spanish electricity spot price forecasting.
3. Hongzhan NIE, Guohui LIU, Xiaoman LIU, Yong WANG (2012), Hybrid of Arima and SVMs for Short-Term Load Forecasting.
4. Kenneth Van den Bergh, Jonas Boury, Erik Delarue (2015), The Flow-Based Market Coupling in Central Western Europe: concepts and definitions.
5. Jesus Crespo Cuaresma, Jaroslava Hlouskova, Stephan Kossmeier, Michael Obersteiner (2004), Forecasting electricity spot-prices using linear univariate time-series models.
6. Peter J. Brockwell, Richard A. Davis, Introduction to Time Series and Forecasting.

7. Mosad Alkhatami (2015), Introduction to Electric Load Forecasting Methods.
8. Jinxing Che, Jianzhou Wang (2010), Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average.
9. Zhongfu Tan, Jinliang Zhang, Jianhui Wang, Jun Xu (2010), Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models.
10. Huisman, R. (2009). An introduction to models for the energy market. Risk Books.
11. Janczura, J., & Weron, R. (2010). An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics*, 32, 1059-1073
12. Vahvilainen, I., & Pyykkonen, T. (2005). Stochastic factor model for electricity spot price - the case of the Nordic market. *Energy Economics*, 27(2), 351-367.
13. Gonzales, V., Contreras, J., & Bunn, D.W. (2012). Forecasting power prices using a hybrid fundamental-econometric model. *IEEE Transactions on Power Systems*, 27(1), 363-372.
14. Karakatsani, N.V., & Bunn, D.W. (2008). Forecasting electricity prices: the impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, 24(4), 764-785.
15. Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*, 49, 328-332.
16. Guerci, E., Rastegar, M. A., & Cincotti, S. (2010). Agent-based modeling and simulation of competitive wholesale electricity markets. In S. Rebennack, et al.(Eds.), *Handbook of power systems 2 - energy systems*(pp. 241-286). Springer.
17. Kowalska-Pyzalska, A., Maciejowska, K., Suszczynski, K., Sznajd-Weron, K., & Weron, R. (2014). Turning green: Agent-based modeling of the adoption of dynamic electricity tariffs. *Energy Policy*, 72, 164-174.
18. Sun, J., & Tesfatsion, L. (2007). Dynamic testing of wholesale power market designs: an open-source agent-based framework. *Computational Economics*, 30, 291-327.
19. Chai, T., & Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature.
20. <http://www.statsref.com/HTML/index.html?arima.html>