



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

Πρόβλεψη οριακής τιμής ηλεκτρικής ενέργειας
με τη χρήση τεχνητών νευρωνικών δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΜΑΡΙΝΑΣ ΓΙΑΝΝΟΥΚΟΥ

Επιβλέπων: Νικόλαος Χατζηαργυρίου
Καθηγητής Ε.Μ.Π.

ΕΡΓΑΣΤΗΡΙΟ ΣΥΣΤΗΜΑΤΩΝ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ
Αθήνα, Ιούλιος 2020



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρικής Ισχύος
Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας

Πρόβλεψη οριακής τιμής ηλεκτρικής ενέργειας με τη χρήση τεχνητών νευρωνικών δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΜΑΡΙΝΑΣ ΓΙΑΝΝΟΥΚΟΥ

Επιβλέπων: Νικόλαος Χατζηαργυρίου
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 23η Ιουλίου 2020.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Νικόλαος Χατζηαργυρίου
Καθηγητής Ε.Μ.Π.

.....
Σταύρος Παπαθανασίου
Καθηγητής Ε.Μ.Π.

.....
Πάυλος Γεωργιλάκης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2020

(Υπογραφή)

.....

ΜΑΡΙΝΑ ΓΙΑΝΝΟΥΚΟΥ

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

© 2020 – All rights reserved



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρικής Ισχύος
Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας

Copyright ©–All rights reserved Μαρίνα Γιαννούκου, 2020.

Με επιφύλαξη παντός δικαιώματος.

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.

Ευχαριστίες

Εκφράζω τις θερμές ευχαριστίες μου στον υποψήφιο διδάκτορα Θεόδωρο Κωνσταντίνου, για την καθοδήγησή του στην εκπόνηση της παρούσας διπλωματικής εργασίας.

Ευχαριστώ τον κ. Νικόλαο Χατζηαργυρίου, που μου έδωσε τη δυνατότητα να αναλάβω το συγκεκριμένο θέμα.

Τέλος, θα ήθελα να ευχαριστήσω ιδιαίτερα την οικογένειά μου και τους φίλους μου για τη στήριξή τους καθ' όλη τη διάρκεια των σπουδών μου στο Εθνικό Μετσόβιο Πολυτεχνείο.

Περίληψη

Στην παρούσα διπλωματική εργασία επιδιώκεται να πραγματοποιηθεί βραχυπρόθεσμη πρόβλεψη ορίζοντα 24 ωρών για την οριακή τιμή της ηλεκτρικής ενέργειας στην Προ-Ημερήσια Αγορά ενέργειας ή Αγορά Επόμενης Ημέρας. Η ανάγκη για τέτοια πρόβλεψη αναδείχτηκε με την απελευθέρωση της αγοράς και τη διείσδυση πολλών συμμετοχόντων σε αυτήν. Στο περιβάλλον του ανταγωνισμού και της μεταβλητότητας που χαρακτηρίζει την αγορά, μια εύστοχη πρόγνωση της τιμής της ηλεκτρικής ενέργειας της επόμενης ημέρας επωφελεί όλους τους συμμετέχοντες, για διαφορετικούς λόγους έκαστο.

Η προσέγγισή μας για αυτή την πρόγνωση γίνεται με δύο τρόπους, αρχικά με στατιστικές μεθόδους, και ειδικότερα με αυτοπαλίνδρομο μοντέλο κινητού μέσου όρου (ARMA), και στη συνέχεια με τη χρήση τεχνητών νευρωνικών δικτύων. Συγκεκριμένα, εξετάζονται δύο κατηγορίες εξειδικευμένων νευρωνικών μοντέλων: τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks ή CNN) και τα Αναδρομικά Νευρωνικά Δίκτυα που αποτελούνται από στρώματα νευρώνων Μακράς και Βραχείας Μνήμης (Long and Short Term Memory ή LSTM). Προκειμένου να μελετηθεί η συνεισφορά των δύο κατηγοριών μοντέλων υλοποιούνται τρία διαφορετικά μοντέλα: Ένα νευρωνικό δίκτυο μονοδιάστατης συνέλιξης (1D-CNN) και δύο αναδρομικά δίκτυα νευρώνων μακράς και βραχείας μνήμης, εκ των οποίων το δεύτερο χρησιμοποιεί ως είσοδο τα χαρακτηριστικά που εξάγει ένα μονοδιάστατο συνελικτικό νευρωνικό δίκτυο (CNN-LSTM). Στη συνέχεια, εκτιμάται η απόδοση των μοντέλων, συγκρίνονται τα σφάλματα που προκύπτουν από τις διαφορετικές μεθόδους πρόβλεψης, διαπιστώνεται η υπεροχή των νευρωνικών έναντι των στατιστικών μοντέλων και αξιολογείται η βελτίωση που προσφέρει η χρήση μηχανικής μάθησης στην πρόγνωση των οριακών τιμών ηλεκτρικής ενέργειας.

Λέξεις Κλειδιά

Προ-Ημερήσια Αγορά ηλεκτρικής ενέργειας, πρόβλεψη Οριακής Τιμής Συστήματος, οικονομικές χρονοσειρές, Αυτοπαλίνδρομο μοντέλο, Μοντέλο κινητού μέσου όρου, Συνελικτικά Νευρωνικά Δίκτυα, Νευρώνες Μακράς και Βραχείας Μνήμης, Αναδρομικά Νευρωνικά Δίκτυα

Abstract

The present diploma thesis presents an attempt to forecast the System Marginal Price (SMP) of electricity in the Day-Ahead market. The importance of this problem emerged after the deregulation of the electricity markets around the world, and the increase of the market participants. Considering the competitive framework and the volatile market environment, a good prediction over the marginal price of the day-ahead benefits each and every participant.

This prediction is carried out firstly by using statistical forecasting techniques, and specifically Autoregressive Moving Average Model (ARMA), and then by implementing Artificial Neural Networks (ANN), and more in detail, the two categories examined are the Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM). In order to evaluate the contribution of these types of neural models, we build a 1-D CNN, an LSTM and a combined architecture, in which the output of a CNN layer proceeds through an LSTM layer to produce the final forecasts. Afterwards, we estimate the performance of the models, we compare the errors of the forecasts induced by each method and, after verifying that ANN outperformed, we evaluate the improvement that the use of machine learning offers in the fields of short-term energy price forecasting.

Keywords

Day-Ahead electricity market, System Marginal Price short-term forecast, financial timeseries, Autoregressive Model, Moving Average Model, Artificial Neural Networks, Convolutional Neural Networks, Long Short-Term Memory Networks, Recurrent Neural Networks

Περιεχόμενα

Ευχαριστίες	1
Περίληψη	3
Abstract	5
Περιεχόμενα	8
Κατάλογος Σχημάτων	11
Κατάλογος Πινάκων	13
1 Εισαγωγή	15
1.1 Εισαγωγή στην αγορά ενέργειας	15
1.1.1 Τα ιδιαίτερα χαρακτηριστικά της ηλεκτρικής ενέργειας	15
1.1.2 Μετάβαση από το μονοπώλιο στην ελεύθερη αγορά	16
1.1.3 Μοντέλα Οργάνωσης των Αγορών Ηλεκτρικής Ενέργειας	17
1.2 Αξία της πρόβλεψης τιμής της ηλεκτρικής ενέργειας	20
1.3 Παράγοντες που επηρεάζουν την τιμή της ηλεκτρικής ενέργειας	22
1.4 Επισκόπηση της σύγχρονης βιβλιογραφίας- Στόχοι παρούσας διπλωματικής . .	23
2 Εισαγωγή στις χρονοσειρές	27
2.1 Βασικά χαρακτηριστικά και ανάλυση χρονοσειρών	27
2.2 Στατιστικά μεγέθη χρονοσειρών	30
2.3 Στασιμότητα-μη στασιμότητα	31
3 Εισαγωγή στις μεθόδους προβλεψεων	35
3.1 Στατιστικές μέθοδοι πρόβλεψης χρονοσειρών- μοντέλα ARIMA	35
3.1.1 Αυτοπαλίνδρομο μοντέλο (AR)	36
3.1.2 Μοντέλο κινητού μέσου (MA)	36
3.1.3 Αυτοπαλίνδρομο μοντέλο κινητού μέσου όρου (ARMA)	37
3.1.4 Μικτό ολοκληρωμένο υπόδειγμα (ARIMA)	38
3.2 Τεχνητα Νευρωνικά Δίκτυα (ΤΝΔ)	38

3.2.1	Ιστορική αναδρομή και εφαρμογές των τεχνητών νευρωνικών δικτύων	38
3.2.2	Το μοντέλο του τεχνητού νευρώνα και η αρχιτεκτονική των τεχνητών νευρωνικών δικτύων	39
3.2.3	Λειτουργία τεχνητών νευρωνικών δικτύων	44
3.2.4	Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)	47
3.2.5	Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks) και νευρώνες Μακράς-Βραχείας Μνήμης (Long Term Short Term Memory)	49
3.3	Σφάλματα πρόβλεψης - Μέτρα απόδοσης	52
4	Πρόβλεψη οριακής τιμής ηλεκτρικής ενέργειας της επόμενης ημέρας	55
4.1	Προεπεξεργασία δεδομένων	55
4.2	Εφαρμογή στατιστικών μεθόδων	60
4.3	Εφαρμογή νευρωνικών δικτύων	65
4.3.1	LSTM	65
4.3.2	CNN	68
4.3.3	CNN & LSTM	71
4.4	Σχολιασμός και σύγκριση μεθόδων πρόβλεψης	74
4.4.1	Σχολιασμός μεθόδων ARIMA και σύγκριση των διαφορετικών μοντέλων	74
4.4.2	Σχολιασμός των ΤΝΔ και σύγκριση μεταξύ των διαφορετικών αρχιτεκτονικών	79
4.4.3	Σύγκριση των ΤΝΔ με τις στατιστικές μεθόδους πρόβλεψης	83
5	Επίλογος	85
5.1	Σύνοψη και συμπεράσματα	85
5.2	Μελλοντικές επεκτάσεις	86
	Βιβλιογραφία	88
	Γλωσσάριο	93

Κατάλογος Σχημάτων

1.1	Χρονοδιάγραμμα συναλλαγών στην αγορά ηλεκτρικής ενέργειας (Πηγή: [1], σελ 18)	17
1.2	Χρονοδιάγραμμα των προθεσμιακών συναλλαγών ηλεκτρικής ενέργειας, που αναφέρονται και ως Συναλλαγές Αγοράς Ηλεκτρικής Ενέργειας (Πηγή: [1], σελ 21)	18
1.3	Μοντέλα Οργάνωσης της Προ-ημερήσιας Αγοράς Ηλεκτρικής Ενέργειας (Πηγή: [2])	18
1.4	Ποιοτικό σχήμα καθορισμού ΟΤΣ από καμπύλες προσφοράς (S) και ζήτησης (D)	19
1.5	Οριακή Τιμή Συστήματος στο Χρηματιστήριο Ενέργειας (Πηγή: [1], σελ 24)	20
1.6	(α)Σχέση μεταξύ ζήτησης-τιμής ενέργειας και (β) ενεργειακού ελλείμματος/πλεονάσματος -τιμής για την περίοδο 01/05/2003 έως 22/09/2003 στο Οντάριο του Καναδά. Η τιμή μετράται σε Δολάριο Καναδά. (Πηγή: [3])	23
1.7	Προβλεπόμενη έναντι πραγματική ωριαία τιμή ηλεκτρικής ενέργειας για το διάστημα 27 Απρ - 5 Μαΐου 1997. Για το διάστημα αυτό, το σχετικό λάθος κατα μέσο όρο ήταν 4.3%. (Πηγή: [4])	24
2.1	Μεσοσταθμική τιμή εισαγωγής φυσικού αερίου, ανά μήνα, για το διάστημα Ιούλιος 2016 – Ιούλιος 2017 (Τα στοιχεία των τιμών ανακοινώνονται στην ιστοσελίδα του Διαχειριστή του Εθνικού Συστήματος Φυσικού Αερίου)	27
2.2	Οπτικοποίηση των ποιοτικών χαρακτηριστικών μιας χρονοσειράς (Πηγη [5])	29
2.3	Παράδειγμα μη στάσιμης (αριστερά) και στάσιμης (δεξιά) χρονοσειράς (Πηγή: [6])	31
2.4	Χρονοσειρά με τάση πριν (αριστερά) και μετά (δεξιά) τη διαφύση πρώτης τάξης. Η νέα χρονοσειρά που προκύπτει είναι εμφανώς πιο στάσιμη.(Πηγή: [6])	32
3.1	Μοντέλο βιολογικού νευρώνα (Πηγή: [7])	39
3.2	Νευρώνας τεχνητού νευρωνικού δικτύου (Πηγή: [8])	40
3.3	Συνήθεις συναρτήσεις ενεργοποίησης (Πηγή: [9])	41
3.4	Συναρτήσεις ενεργοποίησης (Πηγή: [10])	42
3.5	Παράδειγμα αρχιτεκτονικής νευρωνικού δικτύου εμπρόσθιας τροφοδότησης (Πηγή: [11])	43

3.6	Μάθηση με επίβλεψη (Πηγή: [12])	45
3.7	Μάθηση χωρίς επίβλεψη (Πηγή: [12])	45
3.8	Ενισχυτική μάθηση (Πηγή: [12])	45
3.9	Αναπαράσταση φαινομένων υποπροσαρμογής, υπερπροσαρμογής και καλής προ- σαρμογής (Πηγή: [13])	47
3.10	Απεικόνιση 1D CNN (Πηγή: [14], σελ 83)	48
3.11	Παράδειγμα αρχιτεκτονικής CNN (Πηγή: [7])	48
3.12	«Ξεδίπλωμα» ενός αναδρομικού δικτύου. (Πηγή: [9])	49
3.13	Δομή κρυφού κυττάρου LSTM. (Πηγή: [15])	50
4.1	Οπτικοποίηση δεδομένων	56
4.2	Μορφή δεδομένων, κινούμενου μέσου και κινούμενης τυπικής απόκλισης	57
4.3	Ανάλυση της χρονοσειράς στις συνιστώσες της τάσης και της εποχικότητας. Αφορά την φθινοπωρινή περίοδο 01/10/2011 - 33/11/2011	58
4.4	Ανάλυση της χρονοσειράς στις συνιστώσες της τάσης και της εποχικότητας. Αφορά την φθινοπωρινή περίοδο 01/10/2011 - 07/11/2011	58
4.5	Συνάρτηση αυτοσυσχέτισης και μερικής αυτοσυσχέτισης, συναρτήσε της υ- στέρησης, για την εαρινή περίοδο 01/07/2012- 30/08/2012	59
4.6	Άνοιξη - Διαγράμματα προβλέψεων με μοντέλα ARIMA για το διάστημα 01/06/2013- 30/06/2013	61
4.7	Μέσο Απόλυτο Ποσοστιαίο Σφάλμα ανά ώρα, για τις προβλέψεις του διαστήμα- τος 01/06/2013- 30/06/2013	62
4.8	Τυπική απόκλιση του απόλυτου ποσοστιαίου σφάλματος ανά ώρα, για τις προ- βλέψεις του διαστήματος 01/06/2013- 30/06/2013	62
4.9	Χειμώνας, Καλοκαίρι, Φθινόπωρο: διαγράμματα προβλέψεων με μοντέλα ARI- MA, $p = 24$	63
4.10	Χειμώνας, Καλοκαίρι, Φθινόπωρο: διαγράμματα σφαλμάτων σε μέσο όρο ανά ώρα, με μοντέλα ARIMA	64
4.11	αρχιτεκτονική LSTM νευρωνικού δικτύου	66
4.12	Φθινόπωρο - Πρόβλεψη Δεκ 2011 με LSTM	67
4.13	Καλοκαίρι - Πρόβλεψη Σεπ 2012 με LSTM	67
4.14	Χειμώνας - Πρόβλεψη Μαρ 2013 με LSTM	67
4.15	Άνοιξη - Πρόβλεψη Ιουν 2013 με LSTM	68
4.16	Άνοιξη - Πρόβλεψη Ιουν 2013 με LSTM και προσθήκη Ιουνίου 2011&2012 στην εκπαίδευση	68
4.17	αρχιτεκτονική CNN νευρωνικού δικτύου	69
4.18	Φθινόπωρο - Πρόβλεψη Δεκ 2011 με CNN	70
4.19	Καλοκαίρι - Πρόβλεψη Σεπ 2012 με CNN	70
4.20	Χειμώνας - Πρόβλεψη Μαρ 2013 με CNN	70
4.21	Άνοιξη - Πρόβλεψη Ιουν 2013 με CNN	71
4.22	αρχιτεκτονική νευρωνικού δικτύου	71

4.23 Φθινόπωρο - Πρόβλεψη Δεκ 2011 με CNN & LSTM	72
4.24 Καλοκαίρι - Πρόβλεψη Σεπ 2012 με CNN & LSTM	73
4.25 Χειμώνας - Πρόβλεψη Μαρ 2013 με CNN & LSTM	73
4.26 Άνοιξη - Πρόβλεψη Ιουν 2013 με CNN & LSTM	73
4.27 Χρονοσειρά δεδομένων τιμής ηλεκτρικής ενέργειας για τους μήνες που υλοποιείται η πρόβλεψη.	75
4.28 Κατανομή της οριακής τιμής ηλεκτρικής ενέργειας για τους μήνες Μάρτιο 2013 και Σεπτέμβριο 2012.	75
4.29 Κατανομή της οριακής τιμής ηλεκτρικής ενέργειας για το μήνα Ιούνιο 2013 - Εαρινή περίοδος	76
4.30 Επίδοση μοντέλων ARIMA για τις διαφορετικές περιόδους πρόβλεψης	76
4.31 Κατανομή του σφάλματος της πρόβλεψης για το μήνα Ιούνιο 2013 με ARIMA(24,0,1) για διάφορες ώρες της ημέρας	77
4.32 Κατανομή του σφάλματος της πρόβλεψης για την ώρα 5:00 κατά το μήνα Ιούνιο 2013 με τα διάφορα μοντέλα ARIMA	77
4.33 MAPE των προβλέψεων με ARIMA(12,0,0) και ARIMA(6,0,2)	78
4.34 MAPE των προβλέψεων με ARIMA(12,0,2) και ARIMA(24,0,0)	78
4.35 Προβλέψεις και σφάλματα με μοντέλο CNN για την εαρινή περίοδο (Ιουν 2013), χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης	80
4.36 Προβλέψεις και σφάλματα με μοντέλο CNN & LSTM για την εαρινή περίοδο, χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης	81
4.37 Προβλέψεις και σφάλματα με μοντέλο CNN για την εαρινή περίοδο, χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης	82

Κατάλογος Πινάκων

4.1	Στατιστικά στοιχεία δεδομένων	55
4.2	Έλεγχος στασιμότητας δεδομένων με Augmented Dickey - Fuller test	59
4.3	Υπερπαραμέτροι του LSTM μοντέλου	66
4.4	Σύγκριση του MAPE του LSTM σε σχέση με το στατιστικό μοντελο	66
4.5	Υπερπαραμέτροι του CNN μοντέλου	69
4.6	Σύγκριση του MAPE του CNN σε σχέση με το στατιστικό μοντελο	69
4.7	Υπερπαραμέτροι του CNN & LSTM μοντέλου	72
4.8	Σύγκριση του MAPE του νευρωνικού δικτύου με CNN & LSTM αρχιτεκτονική σε σχέση με το στατιστικό μοντελο	72
4.9	MAPE των προβλέψεων εαρινής περιόδου των διαφόρων νευρωνικών μοντέλων, ανάλογα με την πραγματική οριακή τιμή της ηλεκτρικής ενέργειας. Το «+» αναφέρεται στην περίπτωση εκπαίδευσης με επιπλέον δεδομένα παρελθόντος . .	81
4.10	Σύγκριση του MAPE των υλοποιηθέντων μοντέλων πρόβλεψης	83

Κεφάλαιο 1

Εισαγωγή

1.1 Εισαγωγή στην αγορά ενέργειας

Η ηλεκτρική ενέργεια συνιστά τη βασικότερη μορφή ενέργειας για καθημερινή χρήση, με τον άνθρακα, το πετρέλαιο, το φυσικό αέριο και τους άλλους ενεργειακούς πόρους να χρησιμοποιούνται κυρίως για την παραγωγή της. Η εύρεση υποκατάστατου της ηλεκτρικής ενέργειας αποτελεί μεν μεγάλη πρόκληση, χωρίς ωστόσο, να υπάρχει ακόμα εφαρμόσιμο και συνδεδεμένο με την αγορά πλάνο.[16] Έτσι, η αγορά ηλεκτρικής ενέργειας (electricity market) αποτελεί αναπόσπαστο κομμάτι της σύγχρονης οικονομίας, και οι σύγχρονες κοινωνίες εξαρτώνται απόλυτα από την αδιάκοπη και άφθονη παροχή ηλεκτρικής ενέργειας.

1.1.1 Τα ιδιαίτερα χαρακτηριστικά της ηλεκτρικής ενέργειας

Η ηλεκτρική ενέργεια ως αγαθό διαφέρει από τα υπόλοιπα προϊόντα, και τα ιδιαίτερα χαρακτηριστικά της καθορίζουν τις εμπορικές ρυθμίσεις που διέπουν την αγορά ηλεκτρικής ενέργειας. Κυριότερο αυτών είναι η αδυναμία της να αποθηκευτεί οικονομικά σε μεγάλες ποσότητες. Η τεχνολογία των μπαταριών δεν μπορεί να διαχειριστεί τα ποσά της ενέργειας που απαιτούνται σε μια ανεπτυγμένη κοινωνία με οικονομικό και αποδοτικό τρόπο. Αυτό έχει ως αποτέλεσμα η παραγωγή της να πρέπει να ακολουθεί συνεχώς τις διακυμάνσεις της ζήτησης κατά τη διάρκεια της ημέρας, προκειμένου να εξασφαλίζεται η διαρκής ισορροπία παραγωγής-ζήτησης.[17] Μία άλλη ιδιαιτερότητα της ηλεκτρικής ενέργειας είναι η μεταφορά της, καθώς δεν πρόκειται για ένα προϊόν που μπορεί να μεταφερθεί σε πακέτα από την περιοχή παραγωγής στην περιοχή κατανάλωσης σε κάθε δεδομένη χρονική στιγμή. Η μεταφορά της υπόκειται στους περιορισμούς του δικτύου μεταφοράς, αφού η ροή της στις γραμμές του δικτύου υπακούει στους νόμους του Kirchhoff, υπόκειται σε απώλειες και το σύστημα μεταφοράς χαρακτηρίζεται από ένα όριο στην ηλεκτρική ισχύ που μπορεί να μεταφέρει σε μία δεδομένη χρονική στιγμή, το οποίο καθορίζεται από τα θερμικά όρια των γραμμών μεταφοράς και τα όρια τάσης και ευστάθειας. Επιπλέον, μικρές διαταραχές μπορεί να έχουν σημαντικές συνέπειες στο σύστημα ηλεκτρικής ενέργειας, αφού διακυμάνσεις στην τάση ή τη συχνότητα μπορεί να προκαλέσουν διακοπές της παροχής. Τέλος, η μη ύπαρξη υποκατάστατου της ηλεκτρικής ενέργειας ως αγαθού καθιστά τη βραχυχρονίως προσδιοριζόμενη ζήτηση ηλεκτρικής

ενέργειας εντόνως ανελαστική.[2, 18]

1.1.2 Μετάβαση από το μονοπώλιο στην ελεύθερη αγορά

Για το μεγαλύτερο μέρος του 20ου αιώνα, η βιομηχανία της ηλεκτρικής ενέργειας ήταν οργανωμένη σε γεωγραφικά μονοπώλια. Μια επιχείρηση που, στις περισσότερες περιπτώσεις, αποτελούσε κάθετα ολοκληρωμένο μονοπώλιο είχε υπό τον έλεγχο της το σύνολο των υπηρσιών παραγωγής, μεταφοράς και διανομής της ηλεκτρικής ενέργειας. Σε άλλες περιπτώσεις, οι επιχειρήσεις που προμήθευαν την ηλεκτρική ενέργεια στους καταναλωτές ήταν υπεύθυνες μόνο για τη διανομή της και ήταν υποχρεωμένες με τη σειρά τους να αγοράζουν την ενέργεια από την επιχείρηση που είχε το μονοπώλιο για την παραγωγή και τη μεταφορά της σε μία ευρύτερη γεωγραφική περιοχή.

Οι λόγοι του μονοπωλιακού καθεστώτος συνοψίζονται στους εξής:

1. Την μονοπωλιακή φύση της διανομής, καθώς μόνο ένα σετ γραμμών - καλωδίων θα έπρεπε να υπάρχει σε κάθε δρόμο, γεγονός που στηρίχτηκε σε λόγους χώρου, εμφάνισης, αλλά κυρίως στο υψηλό οικονομικό κόστος που διέπει τη διανομή της ηλεκτρικής ενέργειας, και καθιστά απαγορευτικό τον ανταγωνισμό σε αυτό το επίπεδο.
2. Την μονοπωλιακή φύση της μεταφοράς, που στηρίχτηκε τόσο στο πολύ υψηλό οικονομικό της κόστος, όσο και στις υψηλές απαιτήσεις χωροθέτησης και τα ιδιαίτερα χαρακτηριστικά του δικτύου μεταφοράς.
3. Τις τεχνικές προκλήσεις για το συντονισμό μιας μονάδας παραγωγής με το δίκτυο μεταφοράς, που καθιστούσαν σχεδόν αδύνατο να διαχωριστούν οι δύο διαδικασίες, για αυτό και επικράτησε η άποψη ότι το καλύτερο θα ήταν να υπόκεινται και οι δύο στην ίδια εταιρεία.
4. Τέλος, το υψηλό κόστος για την κατασκευή των μονάδων παραγωγής σε συνδυασμό με τις χαμηλές τιμές της ηλεκτρικής ενέργειας που παρήγαγαν, που ενίσχυαν την επικράτηση του φυσικού μονοπωλίου και στην παραγωγή της ηλεκτρικής ενέργειας.

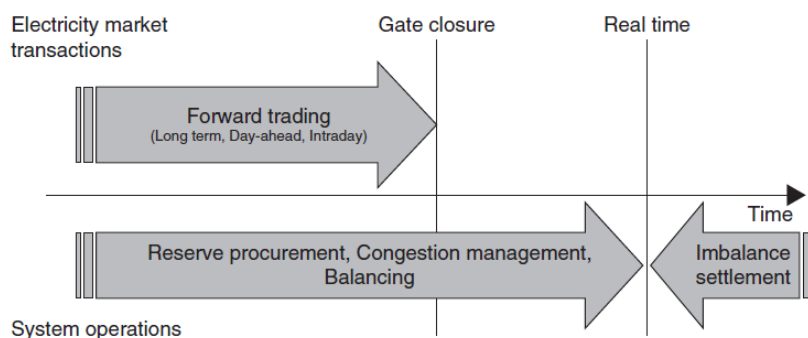
Οι τομείς της μεταφοράς και της διανομής της ηλεκτρικής ενέργειας εξακολουθούν να θεωρούνται φυσικά μονοπώλια, όμως η επικράτηση της αντίληψης ότι η παραγωγή και η πώληση ενός προϊόντος θα πρέπει να διαχωριστούν από τη μεταφορά του, σε συνδυασμό με τη ραγδαία τεχνολογική πρόοδο και την είσοδο του ανταγωνισμού σε άλλους τομείς της οικονομίας (τηλεπικοινωνίες, μεταφορές, αερογραμμές, αγορά φυσικού αερίου), υπήρξε καθοριστική για τη μετάβαση στις ανταγωνιστικές αγορές ηλεκτρικής ενέργειας.[2]

Τα κύρια, βέβαια, επιχειρήματα εναντίον της μονοπωλιακής διαχείρισης υπήρξαν οικονομικής φύσης. Ένα κεντρικά/μονοπωλιακά διαχειριζόμενο Σύστημα Ηλεκτρικής Ενέργειας (ΣΗΕ) αδυνατεί να εξασφαλίσει την ταυτόχρονη υλοποίηση δύο βασικών στόχων για την οικονομική λειτουργία του συστήματος: την τιμολόγηση κοντά στο πραγματικό κόστος παραγωγής και την παροχή ισχυρών κινήτρων για την ελαχιστοποίηση του κόστους παραγωγής. Αντίθετα, μία ανταγωνιστικά οργανωμένη αγορά μπορεί να πετύχει ταυτόχρονα και τους δύο

αυτούς στόχους. Ο κρατικός έλεγχος στον τομέα του ηλεκτρισμού υπήρξε επίσης προβληματικός, καθώς πολύ συχνά η διαχείριση των επιχειρήσεων ηλεκτρισμού υπάκουε περισσότερο σε πολιτικές επιδιώξεις παρά σε οικονομικά και λειτουργικά κριτήρια.[19] Από τις αρχές του 1980 μέχρι και σήμερα πολλές χώρες έχουν εδραιώσει καθεστώς μερικώς ή πλήρους απελευθερωμένης αγοράς ενέργειας. Οι παραγωγοί και οι προμηθευτές στοχεύουν στη μεγιστοποίηση του κέρδους τους, ενώ οι καταναλωτές προσδοκούν χαμηλό κόστος, αξιόπιστες υπηρεσίες και λογικά προβλέψιμες χρεώσεις.[20] Από την άλλη πλευρά, ο Λειτουργός της Αγοράς (ΛΑ ή Market Operator - MO), ο Διαχειριστής του Συστήματος (ΔΣ ή System Operator - SO) και οι Ρυθμιστικές Αρχές Ενέργειας (ΡΑΕ ή Regulatory Authorities for Energy - RAE) αποτελούν ανεξάρτητες οντότητες, οι οποίες καλούνται να διασφαλίσουν όλες τις απαραίτητες προϋποθέσεις εύρυθμης λειτουργίας των αγορών, με στόχο τη μεγιστοποίηση της κοινωνικής ευημερίας.

1.1.3 Μοντέλα Οργάνωσης των Αγορών Ηλεκτρικής Ενέργειας

Οι συναλλαγές που λαμβάνουν χώρα στην χονδρική αγορά ηλεκτρικής ενέργειας μπορούν να διακριθούν σε δύο κατηγορίες: στις συναλλαγές της αγοράς ενέργειας (Electricity Market Transactions) και συναλλαγές σχετιζόμενες με τη διαχείριση του συστήματος (System Operations). Οι πρώτες αναφέρονται στις προθεσμιακές συναλλαγές (forward transactions) μεταξύ των συμμετεχόντων στην αγορά, αυτές δηλαδή που πραγματοποιούνται μέχρι την στιγμή της παράδοσης, ενώ οι δεύτερες λαμβάνουν χώρα σε διάφορους χρόνους γύρω από την στιγμή της παράδοσης και αφορούν την αγορά εξισορρόπησης (Balancing Market)¹. Το τυπικό χρονοδιάγραμμα των συναλλαγών φαίνεται στο παρακάτω σχήμα. [1]

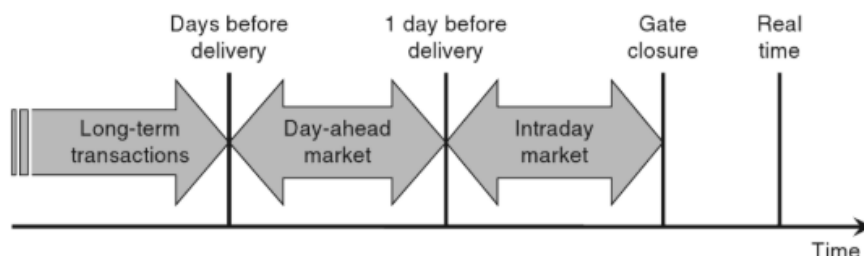


Σχήμα 1.1: Χρονοδιάγραμμα συναλλαγών στην αγορά ηλεκτρικής ενέργειας (Πηγή: [1], σελ 18)

Εμείς θα ασχοληθούμε με μία από τις βασικότερες αγορές που αφορούν προθεσμιακές συναλλαγές, την Προ-Ημερήσια Αγορά ηλεκτρικής ενέργειας ή Αγορά Επόμενης Ημέρας (Day-

¹Αγορά Εξισορρόπησης: Η Αγορά Ηλεκτρικής Ενέργειας που τελεί υπό τη διαχείριση του Διαχειριστή του ΕΣΜΗΕ και περιλαμβάνει τις Αγορές Ισχύος Εξισορρόπησης και Ενέργειας Εξισορρόπησης ως και τη διαδικασία εκκαθάρισης αποκλίσεων σύμφωνα με τον Κανονισμό Αγοράς Εξισορρόπησης του άρθρου 18 του ν. 4425/2016.[21]

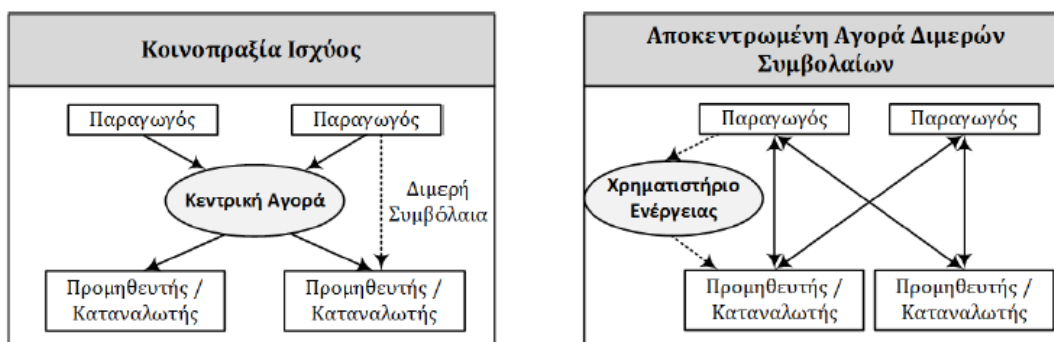
Ahead Market), δηλαδή την οργανωμένη αγορά στην οποία πραγματοποιούνται συναλλαγές αγοράς και πώλησης ηλεκτρικής ενέργειας με υποχρέωση φυσικής παράδοσης την επόμενη ημέρα (Ημέρα Εκπλήρωσης Φυσικής Παράδοσης) [21]



Σχήμα 1.2: Χρονοδιάγραμμα των προθεσμιακών συναλλαγών ηλεκτρικής ενέργειας, που αναφέρονται και ως Συναλλαγές Αγοράς Ηλεκτρικής Ενέργειας (Πηγή: [1], σελ 21)

Διακρίνονται δύο γενικοί τύποι οργάνωσης, ως ακολούθως:

- α) Κεντρικά Οργανωμένη Αγορά ή Κοινοπραξία Ισχύος (Power Pool)
- β) Αποκεντρωμένη Αγορά Διμερών Συμβολαίων, σε συνδυασμό με την προαιρετική Λειτουργία ενός Χρηματιστηρίου Ενέργειας (Bilateral Contracts with Voluntary Power Exchange Model).

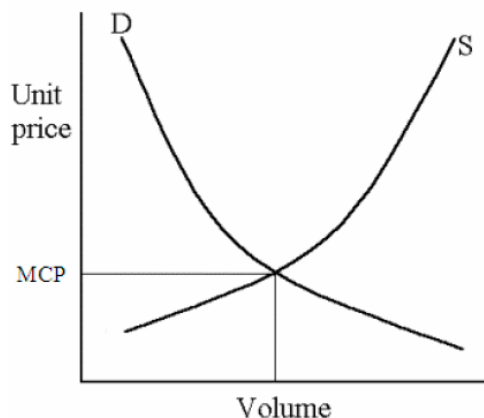


Σχήμα 1.3: Μοντέλα Οργάνωσης της Προ-ημερήσιας Αγοράς Ηλεκτρικής Ενέργειας (Πηγή: [2])

Όπως φαίνεται στο Σχήμα 1.3, η Κοινοπραξία Ισχύος αντιπροσωπεύει μια κεντρικά οργανωμένη αγορά, όπου όλη η ενέργεια πωλείται και αγοράζεται από τον Διαχειριστή της αγοράς. Αυτό επιτυγχάνεται με τον κεντρικό προγραμματισμό όλων των συναλλαγών από τον Διαχειριστή, στον οποίο είναι υποχρεωμένοι όλοι οι παραγωγοί να καταθέτουν τις προσφορές τους για παραγωγή ηλεκτρικής ενέργειας και όλοι οι προμηθευτές με τη σειρά τους, τις προσφορές τους για κατανάλωση ηλεκτρικής ενέργειας.

Συγκεκριμένα, ο Διαχειριστής της αγοράς συγκεντρώνει όλες τις προσφορές έγχυσης και τις προσφορές κατανάλωσης.[2] Συνακόλουθα, σχηματίζεται μια καμπύλη προσφοράς και μία καμπύλη ζήτησης, από την τομή των οποίων προκύπτει η Οριακή Τιμή Συστήματος (ΟΤΣ ή

Marginal Clearing Price - MCP), η οποία αποτελεί την τιμή εκκαθάρισης, και είναι δημόσια διαθέσιμη.



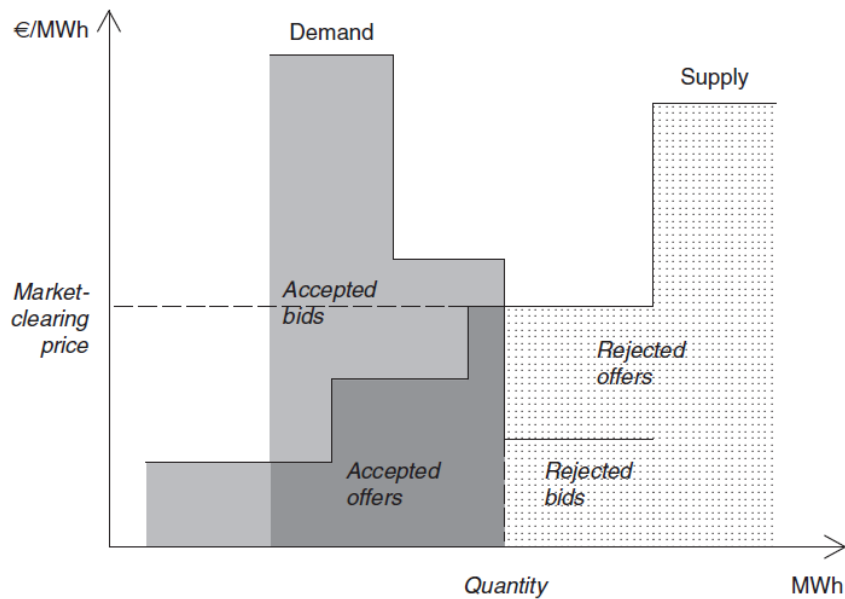
Σχήμα 1.4: Ποιοτικό σχήμα καθορισμού ΟΤΣ από καμπύλες προσφοράς (S) και ζήτησης (D)

Οι προσφορές από τις μονάδες παραγωγής κατατάσσονται με αύξουσα σειρά ως προς το κόστος, ώστε να δημιουργηθεί μια “σειρά προτεραιότητας κατανομής”, για κάθε περίοδο κατανομής, με σκοπό την ελαχιστοποίηση του κόστους του συστήματος. Η ΟΤΣ ορίζεται ως η τιμή προσφοράς της οριακής (τελευταίας) μονάδας που χρειάζεται, ώστε να εξασφαλιστεί ότι καλύπτεται η προβλεπόμενη ζήτηση για μια συγκεκριμένη περίοδο κατανομής. Οι διαθέσιμες μονάδες παραγωγής που υποβάλλουν προσφορές στο pool σε τιμές ανώτερες της ΟΤΣ δεν κατανέμονται και κατά συνέπεια δεν λαμβάνουν εισόδο στην αγορά. Έτσι, δημιουργείται ανταγωνισμός μεταξύ των διαθέσιμων μονάδων ώστε να κατανεμηθούν και να λάβουν εισόδο βάσει ΟΤΣ. [22] Οι χρονικές περίοδοι κατανομής είναι συνήθως ωριαίες, επομένως για κάθε μέρα προκύπτουν 24 οριακές τιμές ηλεκτρικής ενέργειας.

Σε μία αποκεντρωμένη αγορά ηλεκτρικής ενέργειας, οι συναλλαγές βασίζονται στη σύναψη διμερών συμβολαίων (bilateral contracts) μεταξύ παραγωγών, προμηθευτών και μεγάλων καταναλωτών για την παράδοση συγκεκριμένων ποσοτήτων ενέργειας, σε προκαθορισμένες τιμές και προκαθορισμένη περίοδο στο μέλλον. Οι συναλλαγές αυτές πραγματοποιούνται μέσω ηλεκτρονικών πινάκων ανακοινώσεων (electronic bulletin boards), είτε μέσω μεσαζόντων (brokers), οι οποίοι αποσκοπούν στην εύρεση των κατάλληλων εταίρων για την απευθείας ανταλλαγή ηλεκτρικής ενέργειας. Ο συγκεκριμένος τύπος συναλλαγών, αναφέρεται συχνά με τον όρο «Εξωχρηματιστηριακές Συναλλαγές» (Over the Counter - OTC), ο οποίος περιγράφει το σύνολο των συναλλαγών που πραγματοποιούνται απουσία κάποιου επιβλέποντος Χρηματιστηρίου. Τα διμερή συμβόλαια λόγω των προκαθορισμένων όρων δεν υπόκεινται στη ρευστότητα που παρουσιάζουν οι τιμές εκκαθάρισης σε μία κοινοπραξία ισχύος και έτσι αποτελούν μέσο αντιστάθμισης κινδύνου για τους Συμμετέχοντες (hedging tool).

Το κύριο μειονέκτημα των Αγορών Διμερών Συμβολαίων αποτελεί το γεγονός ότι αυτές πραγματοποιούνται σε μεσοπρόθεσμο ή μακροπρόθεσμο χρονικό ορίζοντα, γεγονός που εμποδίζει την εξασφάλιση της ασφάλειας του συστήματος ηλεκτρικής ενέργειας μόνο μέσω των συγκεκριμένων αγορών. Γι' αυτό το λόγο, πολύ συχνά συναντάται η λειτουργία ενός

απλού Χρηματιστηρίου παράλληλα με την Αγορά Διμερών Συμβολαίων, το οποίο έχει τον χαρακτήρα εθελοντικής συμμετοχής των συμμετεχόντων και εκκαθαρίζει την αγορά με βάση μία απλή οικονομική σειρά εκκαθάρισης των προσφορών.[2] Συγκεκριμένα, τα χρηματιστήρια ενέργειας άμεσης παράδοσης (spot exchanges) στα οποία πραγματοποιούνται συναλλαγές της επόμενης μέρας (αλλά και ενδοημερήσιες -intraday), διαθέτουν πλατφόρμα συναλλαγών και υλοποιούν δημοπρασίες κάθε μέρα κατά τις μεσημβρινές, κυρίως, ώρες, με εμπόρευμα την ωριαία, συνήθως, ηλεκτρική ενέργεια προς παράδοση την επόμενη μέρα. Η τελική εκκαθάριση προκύπτει από την αντιστοίχιση ζήτησης και προσφοράς (η οποίες, όπως και στο pool, γίνονται με υποβολή προσφορών αγοράς και πώλησης- buy bids και sell bids) [22]



Σχήμα 1.5: Οριακή Τιμή Συστήματος στο Χρηματιστήριο Ενέργειας (Πηγή: [1], σελ 24)

1.2 Αξία της πρόβλεψης τιμής της ηλεκτρικής ενέργειας

Το γεγονός ότι το πρότερο μονοπωλιακό καθεστώς, που συνδεόταν με παραδοσιακές πρακτικές λήψης αποφάσεων, έχει αντικατασταθεί από τη σύγχρονη ανταγωνιστική αγορά ενέργειας, με τους συμμετέχοντες αυτής να αποσκοπούν ο καθένας στη μεγιστοποίηση του κέρδους τους, καθιστά προφανή την ανάγκη σχεδιασμού ενεργειών και πολυκριτήριας ανάλυσης για λήψη αποφάσεων. Ο χρηματοοικονομικός σχεδιασμός δε, είτε βραχυπρόθεσμος είτε μακροπρόθεσμος, βασίζεται σε μεγάλο βαθμό στις προβλέψεις: πρόβλεψη εσόδων, παραγωγής, πωλήσεων, συνθηκών του περιβάλλοντος, κλπ, οι οποίες θα καθορίσουν και ένα πλήθος επιχειρηματικών διαδικασιών. Η πρόβλεψη είναι μια προσπάθεια οπτικοποίησης του μελλοντικού περιβάλλοντος και αντίληψης του τρόπου με τον οποίο η αγορά θα εξελιχθεί στη διάρκεια του ορίζοντα σχεδιασμού. Στις αγορές ηλεκτρικής ενέργειας, η πρόβλεψη της Οριακής Τιμής Συστήματος έχει μεγάλη σημασία τόσο για τους παραγωγούς και προμηθευτές, όσο και για τους καταναλωτές. Είναι ένα πολύτιμο εργαλείο ενάντια στην αστάθεια και τη μεταβλητότητα

τα που χαρακτηρίζει μια αγορά πλειστηριασμών, όπως την απελευθερωμένη αγορά ενέργειας. Συγκεκριμένα: [23, 24]

- Οι παραγωγοί που συμμετέχουν σε μια κοινοπραξία ισχύος και είναι σε θέση να κάνουν πρόβλεψη, μπορούν να προσαρμόσουν τις δικές τους τιμές και τον προγραμματισμό της παραγωγής, λαμβάνοντας υπόψη την ΟΤΣ ανά ώρα καθώς και τα κόστη παραγωγής. Υψηλής ποιότητας προβλέψεις της ΟΤΣ και του διαστήματος εμπιστοσύνης της επιτρέπουν στους παραγωγούς να υποβάλλουν αποτελεσματικές προσφορές με χαμηλό ρίσκο, και προσφέρουν, έτσι, βραχυπρόθεσμα, προστασία από την αβεβαιότητα. Η προστασία αυτή προσφέρεται επίσης μεσο- και βραχυπρόθεσμα στα διμερή συμβόλαια, με τους παραγωγούς να θέτουν λογικές προσφορές. Στις αγορές που συνυπάρχουν τα δύο μοντέλα, η γνώση της ΟΤΣ επιτρέπει στις εταιρείες να βελτιστοποιήσουν το σχεδιασμό της παραγωγής τους και να αντιμετωπίσουν βραχυπρόθεσμα τη μεταβλητότητα τιμής της κοινοπραξίας ισχύος συνάπτοντας πιο αποδοτικά διμερή συμβόλαια. Μεσο και μακροπρόθεσμα δύνανται να υπολογίσουν την ποσότητα της ηλεκτρικής ενέργειας που θα διαθέσουν στην κοινοπραξία ισχύος και στα διμερή συμβόλαια. Σε κάθε μοντέλο οργάνωσης της αγοράς ενέργειας μακροπρόθεσμα, τα σχέδια για επέκταση παραγωγής επηρεάζονται άμεσα από δείκτες τιμών της αγοράς ενέργειας.
- Στην πλευρά των καταναλωτών και σε όλους τους χρονικούς ορίζοντες αντιμετωπίζεται η άγνοια ή και αδιαφορία σχετικά με την τιμή της ενέργειας, η οποία μπορεί να οδηγήσει σε μεγάλες διακυμάνσεις ή και εκτόξευση της τιμής της ηλεκτρικής ενέργειας. Με τη γνώση της ΟΤΣ της επόμενης μέρας ο καταναλωτής μπορεί να καταστρώσει σχέδιο ώστε να μεγιστοποιήσει τη δική του ωφέλεια και να εντρυφήσει στις στρατηγικές υποβολής προσφορών. Επιπλέον, αν έχει τη δυνατότητα της παραγωγής ηλεκτρικής ενέργειας τοπικά, μπορεί να την διαλέξει έναντι της αγοράς, ανάλογα με το συμφέρον του βραχυπρόθεσμα. Βραχυ- και μεσοπρόθεσμα καταπολεμά το ρίσκο της μεταβλητότητας της τιμής μέσω διμερών συμβολαίων και μακροπρόθεσμα του επιτρέπεται ο σχεδιασμός επενδύσεων σε δυνατότητες κατανεμημένης παραγωγής ενέργειας.
- Οι προμηθευτές, δηλαδή εταιρείες που αγοράζουν ενέργεια είτε από pool είτε με διμερή συμβόλαια και την πωλούν στους πελάτες τους, επωφελούνται αντίστοιχα, επιδιώκοντας μέσω καλών προβλέψεων της τιμής να μεγιστοποιήσουν τα βραχυπρόθεσμα και μακροπρόθεσμα κέρδη τους.

1.3 Παράγοντες που επηρεάζουν την τιμή της ηλεκτρικής ενέργειας

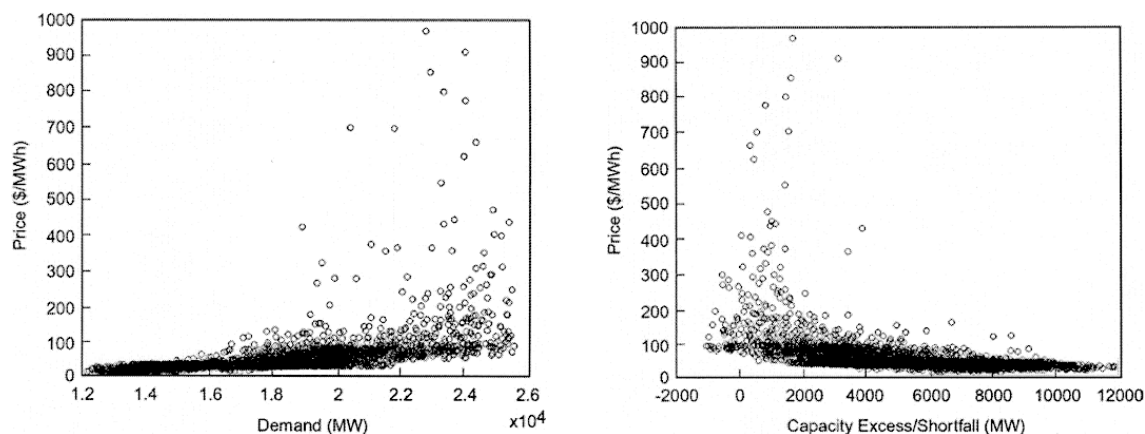
Στις περισσότερες ανταγωνιστικές αγορές ενέργειας οι χρονοσειρές της ωριαίας τιμής παρουσιάζουν τα ακόλουθα χαρακτηριστικά: [24]

- * μεταβαλλόμενη μέση τιμή και διασπορά (μη στάσιμες χρονοσειρές)
- * πολλαπλή εποχικότητα (σε ημερήσιο αλλά και εβδομαδιαίο επίπεδο)
- * ημερολογιακές ανωμαλίες (calendar effect or calendar anomaly), όπως σαββατοκύριακα ή διακοπές
- * μεγάλη αστάθεια
- * μεγάλο ποσοστό ασυνήθιστων τιμών λόγω απρόβλεπτων ή μη ελέγξιμων γεγονότων στις ενεργειακές αγορές

Ποιοί είναι, όμως, οι παράγοντες που καθορίζουν τελικά την τιμή της ηλεκτρικής ενέργειας, και ποια η σχέση τους με αυτή;

Γενικά, ο κυριότερος παράγοντας που επηρεάζει την τιμή της ηλεκτρικής ενέργειας είναι η ζήτηση. Για παράδειγμα, στο Σχήμα 1.6(α) παρουσιάζεται η σχέση μεταξύ ζήτησης και τιμής στο Οντάριο του Καναδά, για την περίοδο Μαΐου-Σεπτεμβρίου 2003. Παρατηρήθηκε ότι η ζήτηση επηρεάζει γραμμικά την τιμή της ενέργειας, αν εξαιρεθούν κάποιες ακραίες τιμές της ενέργειας, που μπορούν να ληφθούν υπόψη ως θόρυβος και προκαλούν την μη γραμμικότητα στη σχέση των δύο μεγεθών. Η θερμοκρασία επίσης παρουσιάζει αντίστοιχη με τη ζήτηση συσχέτιση με την τιμή της ηλεκτρικής ενέργειας. Βέβαια, η επιρροή της θερμοκρασίας ήδη απεικονίζεται στο μέγεθος της ζήτησης, οπότε και δεν συνεισφέρει πολύ στην πρόβλεψη της τιμής ως ξεχωριστός παράγοντας. Το ίδιο συμβαίνει και με την ώρα της ημέρας, που σχετίζεται με την τιμή της ενέργειας, αλλά ο αντίκτυπός της είναι ήδη φανερός στην ζήτηση. Άλλη μεταβλητή που έχει μελετηθεί είναι η ποσότητα αποθεματικών ρύθμισης, που όμως δεν έχει άμεση επίδραση στην τιμή της ενέργειας γενικά, ωστόσο υπάρχει μη γραμμική συσχέτιση, που απεικονίζεται στις περισσότερες αιχμές (peak) στην τιμή της ενέργειας. Αυτή η επιρροή γίνεται πιο φανερή σε διαστήματα αυξημένης ζήτησης ηλεκτρικής ενέργειας, επομένως η συγκεκριμένη μεταβλητή θα μπορούσε να αποτελεί ένδειξη ότι είναι πιθανό να ακολουθήσει αύξηση της τιμής της ενέργειας, αλλά η ακριβής επίδρασή της καταγράφεται εκ των υστέρων. Το πλεόνασμα/ έλλειμμα ενέργειας καθορίζει επίσης την τιμή της ηλεκτρικής ενέργειας, όπως φαίνεται και από τη συσχέτιση τους μέσω του διαγράμματος στο Σχήμα 1.6(β). Σε περιόδους ενεργειακού ελλείμματος ή χαμηλού ενεργειακού πλεονάσματος η τιμή της ηλεκτρικής ενέργειας αυξάνεται. Η σχέση αυτή είναι αντίστοιχη με τη σχέση ζήτησης-τιμής, δηλαδή γραμμική, εκτός των ακραίων τιμών (υψηλή ζήτηση/ μεγάλο έλλειμμα) που προκαλούν μη γραμμικότητες. [3]

Παρόλο που ο αντίκτυπος ορισμένων παραγόντων στην τιμή της ενέργειας είναι τόσο φανερός, η πρόβλεψη της παραμένει δύσκολο εγχείρημα, καθώς οι ακριβείς τιμές της ζήτησης ή οι καιρικές συνθήκες, για παράδειγμα, δεν είναι γνωστές εκ των προτέρων, απαιτούν δηλαδή και αυτές πρόβλεψη. Έτσι, οι αναλυτές καλούνται να συνοδεύουν τις προβλέψεις στην τιμή της ενέργειας με τις προβλέψεις της αβεβαιότητας που τις χαρακτηρίζει. Σημαντικότερη,



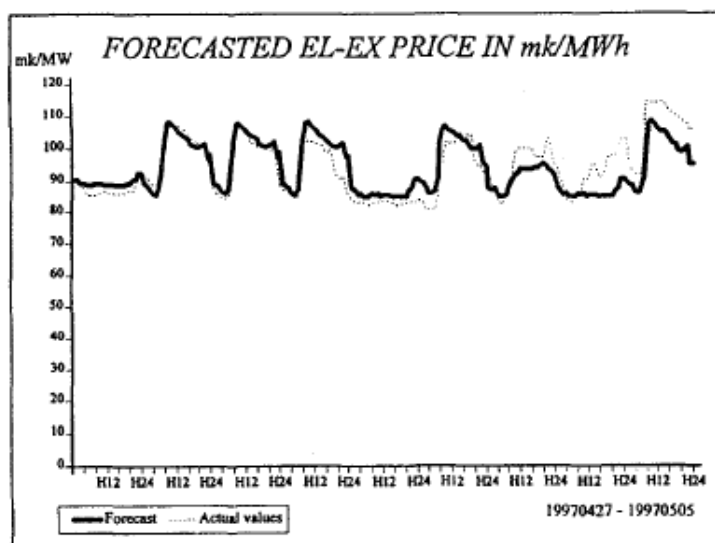
Σχήμα 1.6: (α) Σχέση μεταξύ ζήτησης-τιμής ενέργειας και (β) ενεργειακού ελλείμματος/πλεονάσματος -τιμής για την περίοδο 01/05/2003 έως 22/09/2003 στο Οντάριο του Καναδά. Η τιμή μετράται σε Δολάριο Καναδά. (Πηγή: [3])

βέβαια, αστάθεια στην πρόβλεψη της τιμής της ηλεκτρικής ενέργειας προκαλούν οικονομικοί παράγοντες. Σε ένα πιο μακροσκοπικό επίπεδο, παράγοντες που παίζουν καθοριστικό ρόλο στις μακροχρόνιες μεταβολές της τιμής και στις διαφοροποιήσεις ανά κράτος σχετίζονται με την αύξηση του ανταγωνισμού μετά τη διασύνδεση των αγορών, το διαχωρισμό της παραγωγής ηλεκτρικής ενέργειας από τη λειτουργία του συστήματος, το κόστος μεταφοράς και διανομής και την αύξηση του δυναμικού παραγωγής ηλεκτρικής ενέργειας με χαμηλές λειτουργικές δαπάνες (όπως η αιολική και ηλιακή ενέργεια, σε συνδυασμό με τις υφιστάμενες δυνατότητες της πυρηνικής ή της υδροηλεκτρικής ενέργειας). [25] Πιο βραχυπρόθεσμος είναι ο αντίκτυπος της τιμής του γαιάνθρακα και του φυσικού αερίου [26] και των συνθηκών της αγοράς, και μάλιστα όταν μελετάμε μια δεδομένη αγορά που οι παραπάνω παράγοντες είναι γνωστοί, ο ανταγωνισμός και οι στρατηγικές υποβολής προσφορών αποτελούν αιτίες απρόβλεπτων μεταβολών στην τιμή της ηλεκτρικής ενέργειας.

1.4 Επισκόπηση της σύγχρονης βιβλιογραφίας- Στόχοι παρούσας διπλωματικής

Η ανάγκη για πρόβλεψη της τιμής της ηλεκτρικής ενέργειας γεννάται μαζί με την απελευθέρωση την αγοράς και την εισαγωγή της στον ανταγωνισμό. Είναι φανερό, επομένως, ότι πρόκειται για ένα πρόβλημα των τελευταίων δεκαετιών. Στις αρχές του 1990 διατυπώθηκε ως πρόβλημα η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας, και μελετήθηκαν στατιστικές μέθοδοι για την επίλυση του [27], ενώ τέθηκαν και οι βάσεις για τη θεώρηση των νευρωνικών δικτύων ως χρήσιμου εργαλείου για πρόβλεψη [28]. Το 1998 ο G. Koreneff παρουσιάζει ένα στατιστικό μοντέλο, συγκεκριμένα παλινδρόμηση, που αποσκοπεί στην βραχυπρόθεσμη πρόβλεψη της τιμής της ηλεκτρικής ενέργειας στην φινλανδική αγορά, η οποία έχει πρόσφατα μεταβεί σε καθεστώς απελευθέρωσης και περιλαμβάνει spot-market που αφορά τις επόμενες 24 ώρες [4]. Παρατηρεί ότι η χρονοσειρά της τιμής παρουσιάζει τους ίδιους κύκλους με την

χρονοσειρά της ζήτησης. Αξιοποιεί τις παρελθοντικές τιμές της χρονοσειράς της ωριαίας τιμής ηλεκτρικής ενέργειας και τη θερμοκρασία και διαπιστώνει ότι με τη γνώση των δύο μόνο αυτών παραγόντων είναι εφικτή μια αρκετά ακριβής πρόβλεψη.



Σχήμα 1.7: Προβλεπόμενη έναντι πραγματική ωριαία τιμή ηλεκτρικής ενέργειας για το διάστημα 27 Απρ - 5 Μαΐου 1997. Για το διάστημα αυτό, το σχετικό λάθος κατα μέσο όρο ήταν 4.3%. (Πηγή: [4])

Ένα χρόνο μετά, ο B.R.Szkuta υλοποιεί ένα feed-forward τεχνητό νευρωνικό δίκτυο τριών επιπέδων και, με είσοδο τις χρονοσειρές της προβλεπόμενης ζήτησης και των αποθεμάτων ηλεκτρικής ενέργειας, προβλέπει την ΟΤΣ ανά μισάωρο στην αγορά της Βικτώριας στην Αυστραλία.[29] Ακολούθησε μεγάλη αύξηση του ερευνητικού ενδιαφέροντος πάνω στην πρόβλεψη της τιμής της ηλεκτρικής ενέργειας της επόμενη ημέρας, με νέες μεθόδους να προστίθενται στη βιβλιογραφία. Η C.P. Rodriguez μελέτησε την αγορά ηλεκτρικής ενέργειας του Οντάριο και υλοποίησε πολλαπλά μοντέλα νευρωνικών δικτύων με backpropagation, χρησιμοποιώντας ως είσοδο είτε την ζήτηση, είτε το ενεργειακό έλλειμμα, είτε και τα δύο και πέτυχε μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE) 23,3%. Με νευροασαφή συστήματα και δοκιμή διαφόρων μοντέλων με ποικίλους παράγοντες που επηρεάζουν την τιμή της ενέργειας ως είσοδο, το σφάλμα βελτιώθηκε, με το βέλτιστο μοντέλο να έχει 19,83%.[3] Ο H. Zareipour χρησιμοποιεί Multivariate Adaptive Regression Splines (MARS), που είναι μη παραμετρική τεχνική παλινδρόμησης και, ξεπερνώντας τους περιορισμούς των γραμμικών στατιστικών μοντέλων, πετυχαίνει μέσο απόλυτο ποσοστιαίο σφάλμα για ημερήσια πρόβλεψη σε περίοδο μιας εβδομάδας 11,8%, και διαπιστώνει ότι σε περιόδους μεγάλης ζήτησης το σφάλμα βελτιώνεται, αν στο μοντέλο εισάγεις δεδομένα ζήτησης εκτός από προηγούμενες τιμές ενέργειας.[23] Για την πρόβλεψη της day-ahead (για 24 ώρες) τιμής ηλεκτρικής ενέργειας χρησιμοποιούνται και άλλες γραμμικές στατιστικές μέθοδοι, όπως το μοντέλο ARIMA, με το οποίο η T. Jaksasa πετυχαίνει MAPE 3,55% εργαζόμενη σε δεδομένα της γερμανικής αγοράς ενέργειας από

το 2000-2011.[30] Η X. Chen θα υλοποιήσει μηχανές ακραίας μάθησης (Extreme Learning Machine) σε συνδυασμό με Bootstrapping, για να προβλέψει τις 48 Οριακές Τιμές Συστήματος της επόμενης ημέρας στην αυστραλιανή αγορά ενέργειας (η ΟΤΣ καθορίζεται ανά μισή ώρα) και, χωρίζοντας τα δεδομένα ανά εποχή, πετυχαίνει ακριβέστερες προβλέψεις για το χειμώνα, με MAPE 8,34%, και χειρότερες για το καλοκαίρι με MAPE 21,88%.[31] Στην πορεία, αυξάνονται οι μελέτες που χρησιμοποιούν τεχνητά νευρωνικά δίκτυα για την βραχυπρόθεσμη πρόβλεψη ΟΤΣ της ηλεκτρικής ενέργειας, όπως αναδρομικά νευρωνικά δίκτυα (RNN) [32], ή ειδικότερες αρχιτεκτονικές τους, όπως LSTM (Long short-term memory) που υλοποιούν βαθιά νευρωνικά δίκτυα [33]. Τα νευρωνικά δίκτυα αποδεικνύονται περισσότερο ικανά να μοντελοποιήσουν τις μη γραμμικές συμπεριφορές των χρονοσειρών και να προβλέψουν ραγδαίες αλλαγές της ΟΤΣ, σε σχέση με τις παραδοσιακές τεχνικές γραμμικής μοντελοποίησης.[24]

Αφού αναφέρθηκαν οι ποικίλες τεχνικές που για την πρόβλεψη της Οριακής Τιμής Συστήματος ηλεκτρικής ενέργειας, θα διατυπωθούν συνοπτικά οι στόχοι της παρούσας διπλωματικής, μέσα από το σχεδιάγραμμά της. Πρόκειται να μελετηθεί μια χρονοσειρά από ωριαίες τιμές ηλεκτρικής ενέργειας, άγνωστης αγοράς, και να εφαρμοστούν τόσο στατιστικές μέθοδοι, όσο και τεχνητά νευρωνικά δίκτυα, με σκοπό να προβλεφθούν οι 24 οριακές τιμές της επόμενης μέρας, για διάστημα ενός μήνα. Για κάθε μέρα του μήνα που περνάει, οι τιμές της θεωρούνται γνωστές και χρησιμοποιούνται για την πρόβλεψη της επόμενης μέρας. Συγκεκριμένα, αρχικά θα υλοποιηθούν μοντέλα ARIMA, με ποικίλες παραμέτρους, και θα επιλεγεί αυτό με τη βέλτιστη επίδοση. Η επίδοση του θα αξιολογηθεί με το μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE), η τιμή του οποίου θα διατηρηθεί ως benchmark. Στη συνέχεια θα πραγματοποιηθεί πρόβλεψη για το ίδιο χρονικό διάστημα με τη χρήση τεχνητών νευρωνικών δικτύων. Θα υλοποιηθούν LSTM, CNN καθώς και νευρωνικό δίκτυο που συνδυάζει τις παραπάνω αρχιτεκτονικές. Τέλος, θα συγκριθεί η επίδοση των μοντέλων με το benchmark. Σημειώνεται ότι, σε κάθε περίπτωση, θα χρησιμοποιείται ως μοναδικό δεδομένο εισόδου οι προηγούμενες τιμές της χρονοσειράς της ΟΤΣ ηλεκτρικής ενέργειας, και ότι η μελέτη θα διακριθεί σε τέσσερις περιπτώσεις, οι οποίες καλύπτουν τις εποχές του χρόνου. Στο κεφάλαιο 2 θα γίνει μια σύντομη αναφορά στα χαρακτηριστικά και τις ιδιότητες των χρονοσειρών, με σκοπό να είναι αργότερα κατανοητός ο τρόπος επεξεργασίας τους. Στο κεφάλαιο 3 θα αναλυθεί το θεωρητικό υπόβαθρο των μοντέλων πρόβλεψης χρονοσειρών που χρησιμοποιήθηκαν, ενώ στο κεφάλαιο 4 θα περιγραφεί αναλυτικά η υλοποίηση των μεθόδων, οι αρχιτεκτονικές που χρησιμοποιήθηκαν, τα αποτελέσματά τους και οι επιδόσεις που επιτεύχθηκαν. Κλείνοντας, θα σχολιαστούν και θα συγκριθούν οι διάφορες μέθοδοι, και θα διατυπωθούν τα συμπεράσματά.

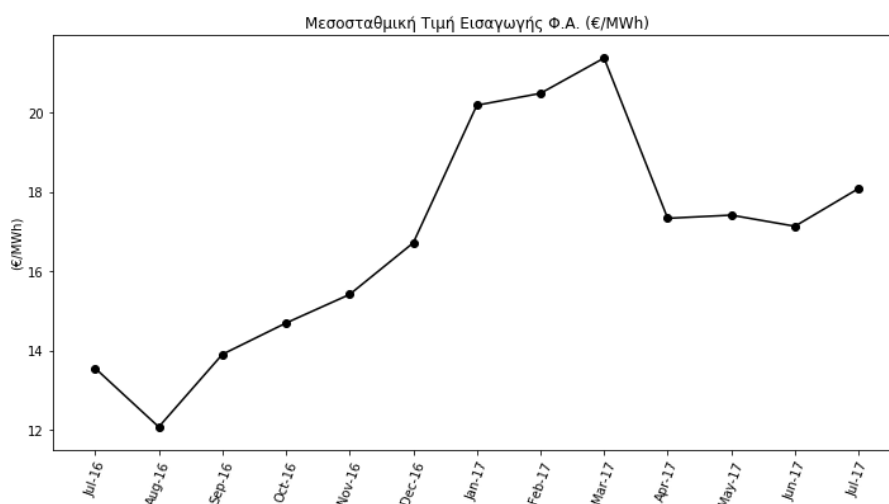
Κεφάλαιο 2

Εισαγωγή στις χρονοσειρές

2.1 Βασικά χαρακτηριστικά και ανάλυση χρονοσειρών

Το σύνολο των δεδομένων, τα οποία συλλέγονται διαχρονικά και εκφράζουν την εξέλιξη των τιμών μιας μεταβλητής κατά τη διάρκεια ίσων διαδοχικών χρονικών περιόδων ονομάζεται χρονοσειρά (ή χρονολογική σειρά, τιμς σειρές). Μαθηματικά, χρονοσειρά είναι ένα σύνολο παρατηρήσεων y_1, y_2, \dots, y_n όπου ο δείκτης $t = 1, 2, \dots, n$ παριστάνει ισαπέχοντα χρονικά σημεία ή διαστήματα. Οι παρατηρήσεις y_1, y_2, \dots, y_n είναι συγκεκριμένες τιμές των τυχαίων μεταβλητών Y_1, Y_2, \dots, Y_n και είναι μέρος μόνο μιας άπειρης ακολουθίας τυχαίων μεταβλητών και συμβολίζεται με $\{Y_t\}$.

Οι χρονοσειρές διακρίνονται σε συνεχείς και σε διακριτές. Συνεχείς χρονοσειρές είναι αυτές που η τιμή του φαινομένου παρατηρείται συνεχώς (π.χ. καταγραφή θερμοκρασίας). Διακριτές χρονοσειρές είναι αυτές όπου η τιμή του φαινομένου καταγράφεται σε ορισμένα χρονικά διαστήματα (π.χ. η τιμή ενός προϊόντος ανά ημέρα). [5]



Σχήμα 2.1: Μεσοσταθμική τιμή εισαγωγής φυσικού αερίου, ανά μήνα, για το διάστημα Ιούλιος 2016 – Ιούλιος 2017 (Τα στοιχεία των τιμών ανακοινώνονται στην ιστοσελίδα του Διαχειριστή του Εθνικού Συστήματος Φυσικού Αερίου)

Τα βασικά ποιοτικά χαρακτηριστικά της χρονοσειράς είναι: η τάση, η κυκλικότητα, η εποχικότητα και οι ακραίες τιμές, και μπορούν συνήθως να γίνουν αντιληπτά από την οπτικοποίηση της σειράς στο πεδίο του χρόνου με γράφημα.

1. Η τάση (trend) είναι η μακροπρόθεσμη μεταβολή του μέσου επιπέδου των τιμών μιας χρονοσειράς. Έτσι, η τάση των τιμών μπορεί να είναι ανοδική, πτωτική ή σταθερή σε ένα συγκεκριμένο χρονικό διάστημα. Συχνά, μπορεί να εκτιμηθεί από διάφορες οικογένειες καμπυλών, όπως μια ευθεία γραμμή ή μια εκθετική καμπύλη. Για να εξαχθούν ασφαλή συμπεράσματα για το αν μια σειρά παρουσιάζει τάση ή όχι, θα πρέπει να έχουμε ένα ικανό αριθμό παρατηρήσεων και να εκτιμηθεί ένα κατάλληλο χρονικό διάστημα.
2. Η κυκλικότητα (cyclic) αντιπροσωπεύει μια μεταβολή που εμφανίζεται λόγω εξωγενών παραγόντων κατά μεγάλες περιόδους. Οι περίοδοι αυτοί είναι μεγαλύτερες του έτους και συνήθως της τάξεως της πενταετίας και δεκαετίας, χωρίς όμως αυτό να σημαίνει ότι είναι σταθερού μήκους. Στις γραφικές παραστάσεις των χρονοσειρών παρουσιάζεται ως μια κυματοειδής γραμμή που κινείται ανάμεσα στην υψηλότερη και χαμηλότερη στάθμη. Η κυκλικότητα εμφανίζεται κυρίως σε οικονομικές χρονοσειρές, λόγω των ανόδων και των υφέσεων που παρουσιάζουν οι οικονομίες.
3. Η εποχικότητα (seasonal) μπορεί να εκφραστεί σαν μια περιοδική διακύμανση που επηρεάζεται από εποχικούς παράγοντες, και η οποία έχει σταθερό και μικρότερο ή ίσο μήκος ενός έτους. Για παράδειγμα, η ποσότητα κατανάλωσης του πετρελαίου θέρμανσης, η οποία είναι μεγαλύτερη κατά τους χειμερινούς μήνες κάθε έτους, παρουσιάζει εποχικότητα. Εφόσον, η εποχική διακύμανση παρουσιάζεται με συστηματικό τρόπο, είναι ένα χαρακτηριστικό εύκολα οπτικά αναγνωρίσιμο που μπορεί να μετρηθεί και να απομονωθεί, ώστε να μην επηρεάζει τα δεδομένα.
4. Οι ακραίες τιμές (outliers) είναι οι απομονωμένες παρατηρήσεις που εμφανίζονται στο γράφημα κάποιας χρονοσειράς ως απότομες αλλαγές στο πρότυπο συμπεριφοράς της και οφείλονται σε κάποιο απρόβλεπτο γεγονός. Οι ακραίες τιμές είναι μη προβλέψιμες και η επίδρασή τους στην χρονοσειρά έχει μικρή χρονική διάρκεια. Η ερμηνεία τέτοιων παρατηρήσεων χρειάζεται ιδιαίτερη προσοχή, διότι απαιτείται θεωρητική γνώση, κριτική ικανότητα και κοινή λογική.

Η ανάλυση χρονοσειρών αποσκοπεί στην ανεύρεση των χαρακτηριστικών εκείνων που συμβάλλουν στην κατανόηση της ιστορικής συμπεριφοράς μιας μεταβλητής και επιτρέπουν την πρόβλεψη μελλοντικών τιμών της. Σε μια συγκεκριμένη χρονοσειρά είναι δυνατόν να μην συνυπάρχουν και οι τέσσερις συνιστώσες, αλλά κάποιες από αυτές.

Για την ανάλυση των χρονοσειρών χρησιμοποιούμε τους ακόλουθους συμβολισμούς:

Y_t = πραγματική τιμή της χρονοσειράς

T_t = Τάση

S_t = Εποχικότητα

C_t = Κυκλικότητα

$E_t =$ Τυχαίο σφάλμα

όπου $t = 1, 2, 3, \dots, n$.

Η εξέταση των στοιχείων αυτών γίνεται με κάποιο μαθηματικό υπόδειγμα, όπως είναι το προσθετικό μοντέλο (addictive model) και το πολλαπλασιαστικό μοντέλο (multiplicative model) και φανερώνει τον τρόπο με τον οποίο οι παρατηρήσεις της χρονοσειράς προσδιορίζονται από τις συνιστώσες της χρονοσειράς. Τα μοντέλα αυτά είναι αντίστοιχα:

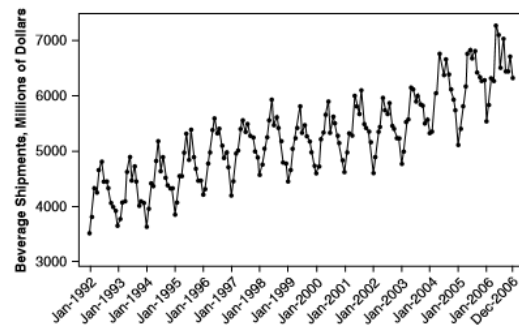
$$Y_t = T_t + S_t + C_t + E_t \quad (2.1)$$

$$Y_t = T_t \cdot S_t \cdot C_t \cdot E_t \quad (2.2)$$

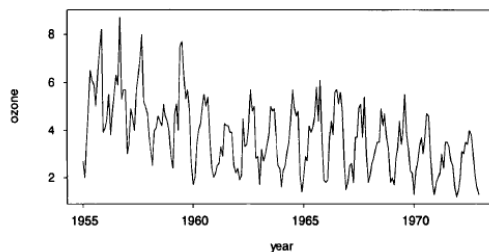
Το κύριο χαρακτηριστικό του προσθετικού μοντέλου είναι ότι όλες οι συνιστώσες είναι ανεξάρτητες μεταξύ τους και εφαρμόζονται στην ίδια μονάδα μέτρησης με εκείνη των παρατηρήσεων της χρονοσειράς. Αντίθετα, στο πολλαπλασιαστικό μοντέλο, μόνο η τάση εκφράζεται στην ίδια μονάδα με εκείνη της χρονοσειράς Y_t , ενώ τα υπόλοιπα στοιχεία S_t , C_t και T_t είναι δείκτες ανεξάρτητοι από μονάδες μέτρησης. [5]



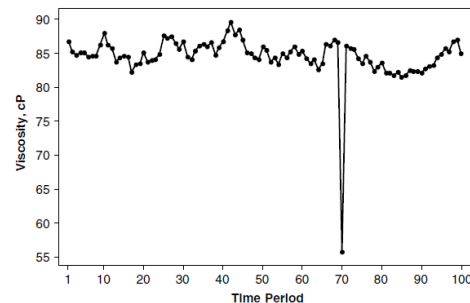
(α') Χρονοσειρά με θετική τάση (Απεικονίζεται η ετήσια παραγωγή των Η.Π.Α. για το μπλε τυρί.)



(β') Χρονοσειρά με κυκλικό μοτίβο ανά έτος (Απεικονίζονται μηνιαίες αποστολές από κατασκευαστή ποτού στις ΗΠΑ.)



(γ') Χρονοσειρά με εποχικότητα (Απεικονίζονται μηνιαίες ενδείξεις όζοντος στο Λος Άντζελες: υψηλές το χειμώνα και χαμηλές το καλοκαίρι.)



(δ') Χρονοσειρά με μια ακραία τιμή (Απεικονίζεται γράφημα αναγνώσεων χημικής διεργασίας.)

Σχήμα 2.2: Οπτικοποίηση των ποιοτικών χαρακτηριστικών μιας χρονοσειράς (Πηγή [5])

2.2 Στατιστικά μεγέθη χρονοσειρών

Παρακάτω θα παρουσιαστούν μερικά από τα κυριότερα στατιστικά μεγέθη μιας χρονοσειράς. [5, 34, 6]

Η μέση τιμή ή αναμενόμενη τιμή μιας χρονοσειράς Y_t δίνεται από την σχέση:

$$\mu = \frac{\sum_{i=t}^n y_t}{n} \quad (2.3)$$

Η μέση τιμή μ σχετίζεται άμεσα με την έννοια της τάσης της χρονοσειράς, εφόσον εκφράζεται ως συνάρτηση της χρονικής στιγμής t της παρατήρησης Y_t . Συγκεκριμένα, αν μια χρονοσειρά παρουσιάζει αυξητική ή πτωτική τάση αντίστοιχα σε ένα χρονικό διάστημα, αυτό θα αποτυπώνεται και στη μέση τιμή ως συνάρτηση του χρόνου.

Η διακύμανση δίνεται από τη σχέση:

$$Var(Y_t) = \sigma^2 = \frac{\sum_{t=1}^n (y_t - \mu)^2}{n} \quad (2.4)$$

και δείχνει πόσο απομακρύνονται οι τιμές της χρονοσειράς από τη μέση τιμή της.

Ο συντελεστής αυτοσυσχέτισης είναι ένας στατιστικός δείκτης ο οποίος χρησιμοποιείται στην ανάλυση χρονοσειρών για τον καθορισμό της τυχαιότητας ή μη της χρονοσειράς. Ο συντελεστής αυτοσυσχέτισης υστέρησης k μας δείχνει κατά πόσο η τιμή της χρονοσειράς σε μία περίοδο εξαρτάται στη γενική περίπτωση από την τιμή της παρατήρησης k περιόδων πίσω. Παίρνει τιμές από $+1$ έως -1 , οι οποίες δηλώνουν απόλυτα θετική και αρνητική γραμμική συσχέτιση αντίστοιχα. Αν ο συντελεστής ισούται με μηδέν τότε δεν υπάρχει καμία συσχέτιση μεταξύ των δύο παρατηρήσεων. Η τιμή της αυτοσυσχέτισης δίνεται από την ακόλουθη σχέση:

$$\rho_k = \frac{\sum_{t=k+1}^n (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^n (Y_t - \mu)^2} \quad (2.5)$$

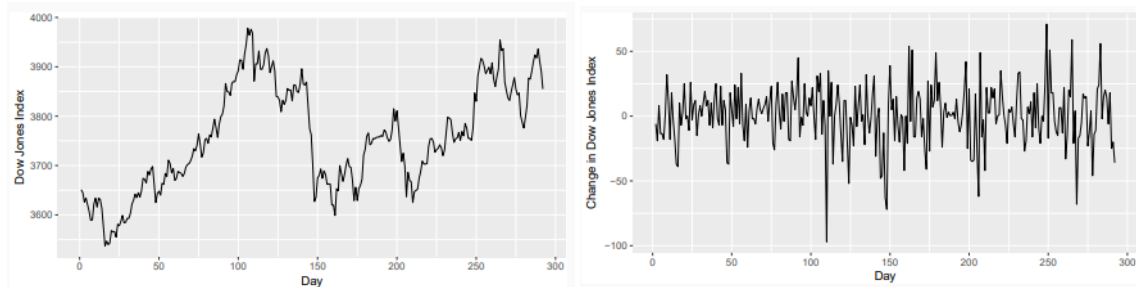
Ο συντελεστής μερικής αυτοσυσχέτισης υστέρησης k δείχνει κατά πόσο η τιμή της χρονοσειράς σε μία περίοδο εξαρτάται από την τιμή της παρατήρησης k περιόδων πίσω, μη λαμβάνοντας υπόψη την επίδραση που μπορεί ενδεχομένως να επιφέρουν οι τιμές που παρεμβάλλονται. Προφανώς, για $k = 1$ ο δείκτης ACF ταυτίζεται με αυτόν του PACF. Και σε αυτήν την περίπτωση, ο συντελεστής παίρνει τιμές από $+1$ έως -1 , οι οποίες δηλώνουν απόλυτα θετική και αρνητική γραμμική συσχέτιση αντίστοιχα. Η τιμή της μερικής αυτοσυσχέτισης δίνεται από τις ακόλουθες σχέσεις:

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \\ \phi_{kk} &= \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \cdot \rho_{k-j}}{1 - \sum_{j=1}^k \phi_{k-1,j} \cdot \rho_j} \quad k = 3, \dots \\ \phi_{kj} &= \phi_{k-1,j} - \phi_{k,k} \phi_{k-1,k-j} \quad k = 2, \dots, j = 1, 2, \dots, k-1 \end{aligned} \quad (2.6)$$

Φτιάχνοντας το γράφημα της αυτοσυσχέτισης συναρτήσει της καθυστέρησης k , το οποίο ονομάζεται συνάρτηση αυτοσυσχέτισης (autocorrelation function – ACF), και το γράφημα της μερικής αυτοσυσχέτισης συναρτήσει της καθυστέρησης k , το οποίο ονομάζεται συνάρτηση μερικής αυτοσυσχέτισης (partial autocorrelation function – PACF) μπορεί κανείς να βγάλει συμπεράσματα για τα χαρακτηριστικά της χρονοσειράς.

2.3 Στασιμότητα-μη στασιμότητα

Μια χρονοσειρά είναι στάσιμη αν η μέση τιμή, η διακύμανση και η συνάρτηση αυτοσυσχέτισης (ACF) διατηρούνται σταθερές στο πέρασμα του χρόνου. Αν ισχύει η υπόθεση της στασιμότητας, τα χαρακτηριστικά της χρονοσειράς δεν εξαρτώνται από τη χρονική στιγμή στην οποία αυτή μελετάται (βλέπε λευκός θόρυβος¹) και, κατά συνέπεια, οποιοδήποτε τυχαίο δείγμα της y_1, y_2, \dots, y_n ταυτίζεται εν γένει με οποιοδήποτε άλλο επιλεγεί $y_{1+t}, y_{2+t}, \dots, y_{n+t}$, όπου t ακέραιος μεγαλύτερος του μηδέν. Έτσι, η χρονοσειρά αποδεσμεύεται από την έννοια του χρόνου και μπορεί να μελετηθεί στοχαστικά.



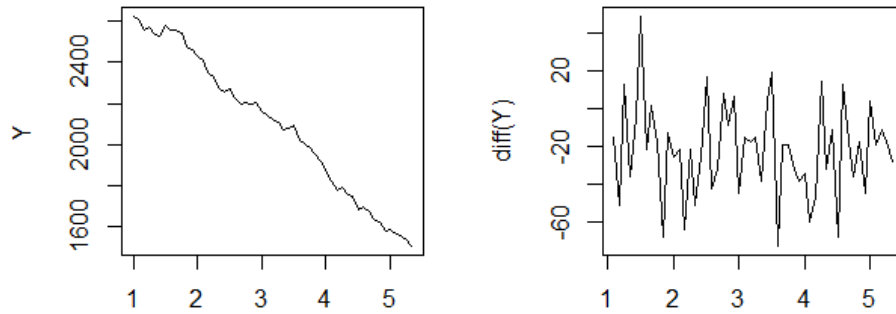
Σχήμα 2.3: Παράδειγμα μη στάσιμης (αριστερά) και στάσιμης (δεξιά) χρονοσειράς (Πηγή: [6])

Η ύπαρξη στασιμότητας θεωρείται προϋπόθεση για την εφαρμογή ορισμένων μεθόδων πρόβλεψης, ωστόσο σπανίως ικανοποιείται εξαρχής, καθώς οι περισσότερες χρονοσειρές χαρακτηρίζονται από τάση, εποχικότητα και ασυνέχειες (π.χ. ειδικά γεγονότα). Κάτι τέτοιο μπορεί να επιτευχθεί με χρήση μετασχηματισμών ή/και διαφορίσης, που αποτελούν τεχνικές επεξεργασίας δεδομένων. Οι μετασχηματισμοί, πχ η λογαρίθμηση, αποτελούν διαδικασίες που περιορίζουν την τυχαιότητα και τις ακραίες τιμές (ασυνέχειες) που τυχόν υπάρχουν, οδηγώντας έτσι σε μία νέα χρονοσειρά μικρότερης και σταθερότερης διακύμανσης. Η διαφορίση, ανάλογα με την μορφή που λαμβάνει, περιορίζει τις διακυμάνσεις επιπέδου αφαιρώντας τάση και εποχικότητα. Έτσι, παράγεται μία νέα χρονοσειρά σταθερότερου επιπέδου και διακύμανσης. Στην ουσία, κατά τη διαφορίση μίας χρονοσειράς n παρατηρήσεων, δημιουργείται μία νέα, $n - 1$ παρατηρήσεων, με στοιχεία της τις διαφορές των παρατηρήσεων της πρώτης. Ανάλογα την τάξη διαφορίσης έχουμε λοιπόν:

$$1\eta \text{ τάξη: } Y_t = Y_t - Y_{t-1}$$

¹Λευκός θόρυβος (White noise): $Y_t = e_t$ Οι παρατηρήσεις μιας χρονοσειράς λευκού θορύβου e_t έχουν μέση τιμή ίση με μηδέν και σταθερή διακύμανση όλες τις χρονικές στιγμές. Ακόμη, όλες οι παρατηρήσεις είναι ασυσχέτιστες μεταξύ τους.[34]

2η τάξη: $Y_t'' = Y_t' - Y_{t-1}' = Y_t - 2Y_{t-1} + Y_{t-2}$, κ.ο.κ.



Σχήμα 2.4: Χρονοσειρά με τάση πριν (αριστερά) και μετά (δεξιά) τη διαφορίση πρώτης τάξης. Η νέα χρονοσειρά που προκύπτει είναι εμφανώς πιο στάσιμη. (Πηγή: [6])

Κατ' αναλογία με την απλή διαφορίση μπορούμε να εφαρμόσουμε και εποχιακή διαφορίση σε περιπτώσεις χρονοσειρών έντονης εποχικότητας. Εδώ η χρονοσειρά που παράγεται είναι αποτέλεσμα της διαφοράς μεταξύ των παρατηρήσεων της αρχικής χρονοσειράς και εκείνων των προηγούμενων αντίστοιχων εποχιακών περιόδων. Ανάλογα με την τάξη διαφορίσης έχουμε λοιπόν:

1η τάξη: $Y_t = Y_t - Y_{t-m}$

2η τάξη: $Y_t'' = Y_t' - Y_{t-m}' = Y_t - 2Y_{t-m} + Y_{t-2m}$, κ.ο.κ.

,όπου m η περίοδος της εποχικότητας.

Αν ορίσουμε τώρα ως B τον τελεστή ολίσθησης, ούτως ώστε $BY_t = Y_{t-1}$ και $B(BY_t) = B^2Y_t = Y_{t-2}$, τότε μπορούμε να αναπαραστήσουμε τη διαφορίση n τάξης ως $(1 - B)^n Y_t$ και την εποχιακή διαφορίση περιόδου m , τάξης N , ως $(1 - B^m)^N Y_t$.

Ο έλεγχος για στασιμότητα περιλαμβάνει μελέτη των γραφικών παραστάσεων ή στατιστικά τεστ που εξετάζουν την ύπαρξη μοναδιαίας ρίζας. Από τη γραφική παράσταση των δεδομένων της χρονοσειράς συναρτήσει του χρόνου, μπορεί να είναι φανερή η ύπαρξη τάσης ή εποχικότητας, οπότε και προκύπτει το συμπέρασμα της μη στασιμότητας. Μελετώντας την γραφική παράσταση της αυτοσυσχέτισης μπορούμε επίσης να εξάγουμε κάποια ποιοτικά συμπεράσματα. Αν η χρονοσειρά παρουσιάζει τάση, τότε η συνάρτηση αυτοσυσχέτισης της (ACF) θα φθίνει πολύ αργά, ενώ αν παρουσιάζει εποχικές διακυμάνσεις, τότε θα εμφανίζονται ταλαντώσεις ίδιας συχνότητας και στην ACF - για παράδειγμα αν μια χρονοσειρά παρουσιάζει ημιτονοειδές μοτίβο, το ίδιο θα ισχύει και για την ACF [35]. Ένας πρακτικός τρόπος για να καταλαβαίνουμε σε ποιες περιπτώσεις οι τιμές των αυτοσυσχετίσεων που παρουσιάζονται στα διαγράμματα ACF και PACF είναι στατιστικά σημαντικές, δηλαδή υποδηλώνουν μη στασιμότητα, είναι να συγκρίνουμε το αποτέλεσμά τους με αυτό ενός λευκού θορύβου ίδιου πλήθους παρατηρήσεων. Για διάστημα εμπιστοσύνης 95%, οι αυτοσυσχετίσεις ενός λευκού θορύβου $N(0, 1/n)$ αναμένεται να λαμβάνουν τιμές στο διάστημα $[-1.96\sqrt{n}, 1.96\sqrt{n}]$. Αν η χρονοσειρά που μελετάται εμφανίζει παρόμοια συμπεριφορά, τότε δεν είναι στατιστικά διάφορη του λευκού θορύβου και συνεπώς είναι στάσιμη. Σε αντίθετη περίπτωση, είναι στατιστικά

διάφορη και απαιτούνται αντίστοιχες ενέργειες. [6] Πιο αξιόπιστοι είναι οι έλεγχοι μοναδιαίας ρίζας, η ύπαρξη της οποίας συνεπάγεται μη στασιμότητα. Οι επικρατέστεροι στατιστικοί έλεγχοι είναι ο Dickey - Fuller (DF), ο επαυξημένος έλεγχος των Dickey - Fuller (Augmented Dickey - Fuller test) και ο έλεγχος Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Αυτές οι τεχνικές έχουν μια μηδενική υπόθεση (null hypothesis), και υπολογίζεται η στατιστική σημαντικότητα (ή p-value) της μηδενικής υπόθεσης και ο στατιστικός δείκτης t-Statistic, ο οποίος συγκρίνεται με μια κρίσιμη τιμή, για να προκύψει το συμπέρασμα επιβεβαίωσης ή μη της υπόθεσης. Συγκεκριμένα, έστω το υπόδειγμα

$$\Delta X_t = \delta X_{t-1} + e_t,$$

όπου: e_t είναι μια ανεξάρτητη και στάσιμη διαδικασία. Η μηδενική υπόθεση του Dickey - Fuller είναι:

H0: Αν $\delta = 0$, υπάρχει μοναδιαία ρίζα (t-Statistic DF > κρίσιμη τιμή), οπότε η χρονική σειρά δεν είναι στάσιμη και η εναλλακτική

Hα: Αν $\delta < 0$, δεν υπάρχει μοναδιαία ρίζα (t-Statistic DF < κρίσιμη τιμή), οπότε η χρονική σειρά είναι στάσιμη.

Αν ο όρος e_t δεν είναι ανεξάρτητος, λόγω πιθανών συσχετίσεων στη χρονική σειρά, τότε χρησιμοποιείται ο επαυξημένος έλεγχος των Dickey - Fuller (Augmented Dickey - Fuller test) που είναι ίδιος με τον απλό έλεγχο D-F, με την μόνη διαφορά ότι η εξίσωση της παλινδρόμησης έχει επαυξηθεί με τις υστερήσεις της εξαρτημένης μεταβλητής. Στον έλεγχο των παραπάνω υποθέσεων χρησιμοποιούμε τους πίνακες των Dickey - Fuller (Dickey & Fuller, 1979) Τέλος, χρησιμοποιούνται οι κρίσιμες τιμές του MacKinnon (MacKinnon, 1991) για επίπεδο σημαντικότητας $\alpha = 1\%$, 5% και 10% , οι οποίες δηλαδή συγκρίνονται με την t-Statistic του ADF. Όταν επιτυγχάνεται στασιμότητα σε επίπεδο: 1% τότε η στασιμότητα της μεταβλητής είναι ισχυρή, 5% τότε η στασιμότητα της μεταβλητής είναι μέτρια, 10% τότε η στασιμότητα της μεταβλητής μας είναι ασθενής ενώ αν υπάρχει σταθερότητα και στις 3 κρίσιμες τιμές (δηλ. 1% , 5% , 10%) τότε η στασιμότητα της μεταβλητής μας είναι ισχυρή.

Σε αντιδιαστολή με την παραπάνω θεωρία, στον έλεγχο KPSS η μηδενική υπόθεση αναφέρει ότι μια χρονική σειρά είναι στάσιμη (Kwiatkowski et al 1992). Για την πραγματοποίηση του ελέγχου των KPSS η μηδενική υπόθεση είναι:

H0: Αν $\delta = 0$, δεν υπάρχει μοναδιαία ρίζα (t-Statistic KPSS > κρίσιμη τιμή), οπότε η χρονική σειρά είναι στάσιμη ενώ η εναλλακτική:

Hα: Αν $\delta < 0$, υπάρχει μοναδιαία ρίζα (t-Statistic KPSS < κρίσιμη τιμή), οπότε η χρονική σειρά δεν είναι στάσιμη.[36]

Δεν θα επεκταθούμε περαιτέρω στα στη θεωρία των στατιστικών ελέγχων για μοναδιαία ρίζα, παρά μόνο θα τους χρησιμοποιήσουμε ως εργαλείο για να διαπιστώσουμε τη στασιμότητα ή μη των χρονοσειρών που θα μελετηθούν.

Κεφάλαιο 3

Εισαγωγή στις μεθόδους προβλεψεων

3.1 Στατιστικές μέθοδοι πρόβλεψης χρονοσειρών- μοντέλα ARIMA

Τα ολοκληρωμένα αυτοπαλίνδρομα μοντέλα κινητών μέσων όρων ARIMA (Auto Regressive Integrated Moving Average) είναι στοχαστικά μοντέλα τα οποία μας βοηθάνε να περιγράψουμε το μηχανισμό εξέλιξης ενός μεγέθους ενδιαφέροντος και ως εκ τούτου να προβλέψουμε την τιμή του στο μέλλον. Μελετήθηκαν εκτεταμένα από τους Box και Jenkins τη δεκαετία του '70 και συχνά αναφέρονται στη βιβλιογραφία με το αντίστοιχο όνομα. Όπως και άλλα στατιστικά μοντέλα πρόβλεψης χρονοσειρών, τα μοντέλα ARIMA παράγουν προβλέψεις βασισόμενα αποκλειστικά στις ιστορικές παρατηρήσεις του υπό εξέταση μεγέθους. Αυτό αποτελεί από μόνο του σημαντικό πλεονέκτημα, ειδικά σε περιπτώσεις που δεν γνωρίζουμε τις εξωτερικές μεταβλητές που το επηρεάζουν ή δεν μπορούμε να τις αξιοποιήσουμε άμεσα (βλ. ντετερμινιστικά μοντέλα). Η ειδοποιός διαφορά ωστόσο μεταξύ των μοντέλων ARIMA και των υπολοίπων μοντέλων χρονοσειρών, είναι η ιδιαίτερη ευελιξία με τον οποίο τα πρώτα διαχειρίζονται τα ιστορικά δεδομένα: δεν υποθέτουν εξαρχής το μηχανισμό με τον οποίο εξελίσσεται το μέγεθος ενδιαφέροντος, αλλά αντίθετα επιλέγουν μέσα από μία ευρεία γκάμα μηχανισμών εκείνον που έχει τη μεγαλύτερη πιθανότητα να αποκαλύπτει τη σχέση που συνδέει την κάθε παρατήρηση με τις προηγούμενες της. Στην πραγματικότητα, το κάθε μοντέλο ARIMA εκφράζει ένα διαφορετικό μηχανισμό εξέλιξης και η επιλογή του καταλληλότερου για την προέκταση μιας χρονοσειράς γίνεται εξετάζοντας παράγοντες όπως η σχέση μεταξύ k διαδοχικών παρατηρήσεων (αυτοσυσχέτιση), η ύπαρξη τάσης, η ύπαρξη εποχικότητας και το σφάλμα πρόβλεψης. Στην γενική τους μορφή τα μοντέλα ARIMA είναι γραμμικός συνδυασμός του τυχαίου παράγοντα (παράγοντας MA), των τιμών που εμφανίστηκαν σε προηγούμενες περιόδους (παράγοντας AR και I), και άλλων στοχαστικών μεταβλητών. Η εφαρμογή των μοντέλων ARIMA προϋποθέτει ότι η χρονοσειρά μήκους n πρέπει να είναι διακριτή, δηλαδή οι παρατηρήσεις της Y_t να αναφέρονται σε ισαπέχουσες χρονικές στιγμές $Y_t, Y_{t+\tau}, \dots, Y_{t+k\tau}$, όπου τ

ακέραιος μεγαλύτερος του μηδέν. Τέλος, η εφαρμογή των μοντέλων ARIMA συνάδει με την παραγωγή βραχυπρόθεσμων προβλέψεων, εφόσον, όπως αναφέρθηκε, πρόκειται για γραμμικούς συνδυασμούς ιστορικών παρατηρήσεων και στοχαστικών παραγόντων. Αυτό σημαίνει πως για να προβλέψουμε την τιμή Y_{n+1} απαιτείται γνώση των τιμών $Y_n, Y_{n-1}, \dots, Y_{n-k}$. Αντίστοιχα, για την πρόβλεψη της τιμής Y_{n+2} απαιτείται γνώση των τιμών $Y_{n+1}, Y_n, \dots, Y_{n-k-1}$, εκ των οποίων, όμως, η Y_{n+1} δεν αποτελεί δεδομένο, αλλά μία πρόβλεψη του μοντέλου που φυσικά εμπεριέχει κάποιο σχετικό σφάλμα. Αν επεκτείνουμε την εν λόγω διαδικασία στο μέλλον, γίνεται κατανοητό πως για μεγάλους ορίζοντες πρόβλεψης η τιμή των προβλέψεων θα βασίζεται αποκλειστικά σε προβλέψεις που πραγματοποιήθηκαν νωρίτερα. Έτσι, η αξιοπιστία και η ακρίβεια πρόβλεψης ενός μοντέλου ARIMA αναμένεται να μειώνεται σημαντικά καθώς αυξάνει ο ορίζοντας πρόβλεψης. [6]

3.1.1 Αυτοπαλίνδρομο μοντέλο (AR)

Μια χρονοσειρά λέγεται αυτοπαλινδρομένη (autoregressive time series) τάξης p (AR(p)), όταν κάθε παρατήρηση y_t εκφράζεται ως ένα σταθμισμένο άθροισμα μιας σταθεράς c και p καθυστερημένων εκδοχών της χρονοσειράς y και ορίζεται από τη σχέση:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} \quad (3.1)$$

όπου η σταθερά c σχετίζεται με τη μέση τιμή της χρονοσειράς και $(\phi_1, \phi_2, \dots, \phi_p)$ είναι παράμετροι αυτοπαλινδρόμησης. Χρησιμοποιώντας τον τελεστή ολίσθησης, ένα μοντέλο AR μπορεί να γραφτεί και ως:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \bar{y}_t = 0 \quad (3.2)$$

, όπου $\bar{y}_t = y - \mu$. Η χρονοσειρά \bar{y}_t έχει τις ίδιες στατιστικές ιδιότητες με την αρχική χρονοσειρά με μηδενική μέση τιμή. Το όνομα «αυτοπαλίνδρομο» προέρχεται από το ότι η τιμή του Y στο χρόνο t εξαρτάται κατά παράγοντα ϕ_i ($i=1,2,\dots,p$) από τις p προηγούμενες τιμές του Y , δηλαδή τις τιμές με χρονική υστέρηση. Επομένως, στα μοντέλα αυτά οι προβλέψεις βασίζονται αποκλειστικά στις παρελθούσες τιμές και τα εμφανισθέντα πρότυπα συμπεριφοράς της χρονοσειράς που εξετάζεται. Το p υποδηλώνει την τάξη του αυτοπαλινδρομου υποδείγματος και αναφέρεται στο μήκος της υστέρησης. [36]

Για ένα μοντέλο AR(p):

Οι τιμές των συντελεστών ACF φθίνουν προς το μηδέν ακολουθώντας εκθετική ή ημιτονοειδή πορεία.

Οι τιμές των συντελεστών PACF μηδενίζονται απότομα μετά από p περιόδους υστέρησης. [6]

3.1.2 Μοντέλο κινητού μέσου (MA)

Τα μοντέλα κινητού μέσου όρου θεωρούν γραμμικές σχέσεις ανάμεσα στην παρατήρηση της χρονοσειράς που εξετάζεται και στα σφάλματα που εμφάνισε το μοντέλο MA σε προηγούμενες περιόδους. Συγκεκριμένα, μια χρονοσειρά Y λέγεται χρονοσειρά κινητού μέσου τάξης q

(MA(q)) όταν κάθε παρατήρηση y_t εκφράζεται ως ένα σταθμισμένο άθροισμα μιας σταθεράς c και q καθυστερημένων εκδοχών της χρονοσειράς λευκού θορύβου. Γράφεται αλγεβρικά ως εξής:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \quad (3.3)$$

όπου $\{e_t\}$ είναι η χρονοσειρά του λευκού θορύβου, q η υψηλότερη υστέρηση των διαταραχών που περιλαμβάνεται σε μία τέτοια διαδικασία και οι παράμετροι $(\theta_1, \theta_2, \dots, \theta_q)$ οι συντελεστές μερικής αυτοσυσχέτισης του μοντέλου MA για υστέρηση $i = 1, \dots, q$. Χρησιμοποιώντας τον τελεστή ολίσθησης, ένα μοντέλο MA μπορεί να γραφτεί και ως:

$$\bar{y}_t = (\theta_1 B + \theta_2 B^2 + \dots + \theta_p B^p) e_t \quad (3.4)$$

Για ένα μοντέλο MA(q):

Οι τιμές των συντελεστών ACF μηδενίζονται απότομα μετά από q περιόδους υστέρησης.

Οι τιμές των συντελεστών PACF φθίνουν προς το μηδέν ακολουθώντας εκθετική η ημιτονοειδή πορεία. [6]

3.1.3 Αυτοπαλίνδρομο μοντέλο κινητού μέσου όρου (ARMA)

Κάποιες στάσιμες χρονοσειρές δεν μπορούν να προτυποποιηθούν αποκλειστικά ως AR ή MA χρονοσειρές, αφού μπορεί να παρουσιάζουν ιδιότητες και από τις δύο κατηγορίες. Έτσι ένα πιο γενικό υπόδειγμα είναι ο συνδυασμός ενός AR(p) υποδείγματος και ενός MA(q) υποδείγματος, το οποίο ονομάζεται αυτοπαλίνδρομο υπόδειγμα κινητού μέσου όρου τάξης (p, q). Κάθε παρατήρηση y_t μιας ARMA(p,q) χρονοσειράς Y εκφράζεται ως εξής [36]:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \quad (3.5)$$

Να σημειωθεί ότι η εφαρμογή των μοντέλων ARMA προϋποθέτει ότι η χρονοσειρά είναι στάσιμη. Παρακάτω δίνονται οι περιορισμοί που υπόκεινται οι παράμετροι ϕ, θ για τα πιο διαδεδομένα μοντέλα ARMA. Για τα υπόλοιπα μοντέλα που είναι και αρκετά πιο σύνθετα στη δομή τους, οι αντίστοιχοι περιορισμοί γίνονται ιδιαίτερος πολύπλοκοι και για αυτό το λόγο δεν παρουσιάζονται.

$$\text{AR}(1): -1 < \phi_1 < 1$$

$$\text{AR}(2): -1 < \phi_2 < 1 \text{ και } \phi_1 + \phi_2 < 1 \text{ και } \phi_2 - \phi_1 < 1$$

$$\text{MA}(1): -1 < \theta_1 < 1$$

$$\text{MA}(2): -1 < \theta_2 < 1 \text{ και } \theta_1 + \theta_2 > -1 \text{ και } \theta_1 - \theta_2 < 1$$

Τέλος, για ένα μοντέλο ARMA(p,q):

Οι τιμές των συντελεστών ACF φθίνουν προς το μηδέν μετά από $q - p$ περιόδους υστέρησης.

Οι τιμές των συντελεστών PACF φθίνουν προς το μηδέν μετά από $p - q$ περιόδους υστέρησης. [6]

3.1.4 Μικτό ολοκληρωμένο υπόδειγμα (ARIMA)

Στις περισσότερες χρονοσειρές παραβιάζεται η ύπαρξη στασιμότητας, και απαιτείται χρήση μετασχηματισμών ή διαφορίσης. Τα ARMA μοντέλα μπορούν να επεκταθούν και σε μη στάσιμες χρονοσειρές με τη χρήση της μεθόδου της διαφορίσης. Τότε ονομάζονται ολοκληρωμένα αυτοπαλίνδρομα υποδείγματα κινητού μέσου τάξεως (p, d, q) (Autoregressive Integrated Moving Average)— και συμβολίζονται με $ARIMA(p, d, q)$. Επομένως, η $ARIMA(p, d, q)$ διαδικασία, είναι μια διαδικασία η οποία διαφορίζεται d φορές και παράγει $ARMA(p, q)$ διαδικασία. Σε περίπτωση που η χρονοσειρά είναι εποχική, τα μοντέλα $ARIMA$ μπορούν να επεκταθούν κατάλληλα προκειμένου να προσομοιώσουν και την εποχική συμπεριφορά των δεδομένων. Σε αυτή την περίπτωση η έκφρασή τους έχει τη μορφή $ARIMA(p,d,q)(P,D,Q)$, όπου τα P, D και Q αναφέρονται αντίστοιχα στην τάξη των εποχιακών παραγόντων $ARIMA$. Το συνολικό μοντέλο αναπαρίσταται με τη χρήση του τελεστή ολίσθησης B ως εξής:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d(1 - B^m)^N y_t = c + (\theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) e_t \quad (3.6)$$

Ο πρώτος όρος του πρώτου μέλους της εξίσωσης αναπαριστά το μοντέλο $AR(p)$, ο δεύτερος την διαφορίση $I(d)$, ενώ ο όρος στο δεύτερο μέλος της εξίσωσης το μοντέλο $MA(q)$. [6, 36]

3.2 Τεχνητα Νευρωνικά Δίκτυα (ΤΝΔ)

3.2.1 Ιστορική αναδρομή και εφαρμογές των τεχνητών νευρωνικών δικτύων

Παρά την πρόσφατη σημαντική αύξηση του ενδιαφέροντος της επιστημονικής κοινότητας σχετικά με τα βαθιά νευρωνικά δίκτυα, το υπολογιστικό πρότυπο των νευρωνικών δικτύων αποτελούσε αντικείμενο έρευνας από τη δεκαετία του 1970, με τις βάσεις να έχουν τεθεί πολύ νωρίτερα. Τα ορόσημα στην εξέλιξη του χώρου των ΤΝΔ είναι τα ακόλουθα:[37]

1943: McCulloch & Pitts – Δημιουργούν το πρώτο μοντέλο ΤΝΔ

1949: Hebb – Δημιουργεί το μοντέλο μάθησης που πήρε το όνομά του στο οποίο κάθε φορά που ενεργοποιείται μια σύναψη αυτή ενισχύεται, με αποτέλεσμα το δίκτυο να μαθαίνει “λίγο περισσότερο” το πρότυπο που του παρουσιάζεται εκείνη τη στιγμή.

1957: Rosenblatt - Προτείνει το στοιχειώδες ΤΝΔ του απλού αισθητήρα που ονόμασε Perceptron.

1969: Minsky & Papert - Αποδεικνύουν μαθηματικά ότι τα ΤΝΔ ενός επιπέδου δεν μπορούν να λύσουν μη γραμμικά προβλήματα.

1982: Μαθηματική απόδειξη ότι ένα ΤΝΔ πολλών επιπέδων μπορεί να αποθηκεύσει οποιαδήποτε πληροφορία.

1986: Werbos & Rumelhart – Προτείνουν τη μέθοδο οπισθοδιάδοσης (backpropagation) για την εκπαίδευση ΤΝΔ.

Σήμερα τα ΤΝΔ είναι δημοφιλή σε προβλήματα τα οποία περιέχουν μη-προβλέψιμες λειτουργίες και δεν είναι πλήρως κατανοητά (κατηγοριοποίηση, αποτίμηση, πρόβλεψη):

- αναγνώριση εικόνας

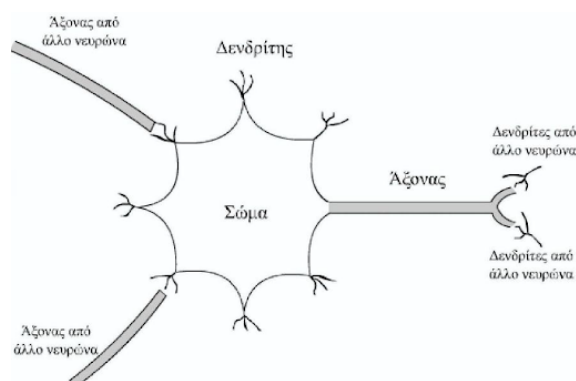
- επεξεργασία φωνής
- σχεδιασμός ενεργειών (planning)
- εξόρυξη πληροφορίας
- χρονοπρογραμματισμός (Scheduling)
- συστήματα ελέγχου και παραγωγής

Όλα τα παραπάνω βρίσκουν εφαρμογή στη Ρομποτική, στην Άμυνα, στην Αεροπορία και Αυτοκίνηση, στην Ιατρική Διαγνωστική, σε γεωλογικές έρευνες, στην Οικονομία και σε τραπεζικές εφαρμογές, στη Βιομηχανία και σε πολλούς ακόμα χώρους παραγωγής.

3.2.2 Το μοντέλο του τεχνητού νευρώνα και η αρχιτεκτονική των τεχνητών νευρωνικών δικτύων

Τα τεχνητά νευρωνικά δίκτυα είναι μια προσπάθεια προσομοίωσης, με την βοήθεια υπολογιστών, του ανθρώπινου νευρικού συστήματος. Τα βιολογικά νευρωνικά δίκτυα που υπάρχουν στον ανθρώπινο εγκέφαλο αποτελούνται από νευρώνες. Ο νευρώνας είναι το μικρότερο τμήμα του εγκεφάλου που είναι ικανό να επεξεργαστεί πληροφορίες, με ταχύτητα 5 με 6 τάξεις μεγέθους πιο αργή από ότι στις σύγχρονες ψηφιακές λογικές πύλες. Ο χρόνος για τις ψηφιακές λογικές πύλες μετριέται σε δισεκατομμυριοστά του δευτερολέπτου (nanoseconds), ενώ για τους νευρώνες σε χιλιοστά του δευτερολέπτου (milliseconds). Όμως ο εγκέφαλος αντισταθμίζει την σχετικά αργή ταχύτητα λειτουργίας νευρώνων με τον πραγματικά τεράστιο αριθμό των μεταξύ τους συνδέσεων. Υπολογίζεται ότι υπάρχουν 10 δισεκατομμύρια νευρώνες και 60 τρισεκατομμύρια συνδέσεις στον φλοιό του ανθρώπινου εγκεφάλου, γεγονός που αποτελεί βασικό λόγο ανεπάρκειας του υπολογιστή για την απόλυτη προσομοίωση του με τον εγκέφαλο. [38]

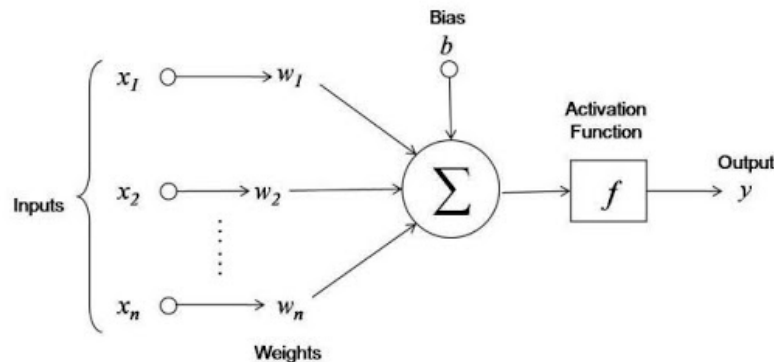
Η βασική δομή ενός νευρώνα μπορεί να περιγραφεί από τρία βασικά μέρη: τους δενδρίτες, το σώμα και τον άξονα. Οι πληροφορίες μεταφέρονται μεταξύ των νευρώνων με μορφή ηλεκτρικών ερεθισμάτων κατά μήκος των δενδριτών. Οι εισερχόμενες πληροφορίες που φθάνουν στους δενδρίτες εισόδου του νευρώνα προστίθενται και υφίστανται επεξεργασία στο σώμα. Έ-



Σχήμα 3.1: Μοντέλο βιολογικού νευρώνα (Πηγή: [7])

στερα μεταφέρονται κατά μήκος του άξονα του νευρώνα στους δενδρίτες εξόδου, όπου μέσω αυτών οι πληροφορίες θα μεταδοθούν σε άλλους νευρώνες, αν το ερέθισμα έχει ξεπεράσει ένα κάτω όριο, οπότε λέμε ότι ο νευρώνας ενεργοποιείται. [7]

Το τεχνητό νευρωνικό δίκτυο, εμπνευσμένο τα βιολογικά νευρικά συστήματα, είναι ένα μοντέλο επεξεργασίας πληροφοριών με βασική δομική μονάδα τους νευρώνες (Processing Units, PUs), οι οποίοι συνδέονται μεταξύ τους, λειτουργούν αρμονικά και είναι οργανωμένοι σε στρώματα (Layers). Κάθε νευρώνας έχει πολλές εισόδους (Inputs) αλλά μόνο μια έξοδο (Output), η οποία με την σειρά της μπορεί να αποτελέσει είσοδο για άλλους νευρώνες. Οι συνδέσεις μεταξύ των νευρώνων διαφέρουν μεταξύ τους και η σπουδαιότητα τους προσδιορίζεται από τον συντελεστή βάρους για κάθε σύναψη. Η επεξεργασία κάθε νευρώνα καθορίζεται από την συνάρτησης ενεργοποίησης, η οποία καθορίζει την έξοδο σε σχέση με τις εισόδους και τους συντελεστές βάρους.[38] Όλα τα παραπάνω οπτικοποιούνται με το ακόλουθο σχήμα.



Σχήμα 3.2: Νευρώνας τεχνητού νευρωνικού δικτύου (Πηγή: [8])

Τα τρία βασικά στοιχεία του μοντέλου του τεχνητού νευρώνα είναι: [38, 9, 7]

1. Ένα σύνολο από συνάψεις ή συνδετικούς κρίκους, κάθε μια από τις οποίες χαρακτηρίζεται από ένα βάρος ή δύναμη. Συγκεκριμένα, κάθε στοιχείο της εισόδου x_i πολλαπλασιάζεται με ένα βάρος σύνδεσης ή βάρος σύναψης w_i , παράγοντας έτσι το γινόμενο $x_i w_i$. Η ιδέα πίσω από αυτό το γινόμενο είναι ότι τα βάρη αυτά είναι εκπαιδύσιμα και ελέγχουν τη δύναμη της επίδρασης του ενός νευρώνα στον άλλον.
2. Ένας αθροιστικός κόμβος για την πρόσθεση των σημάτων εισόδου. Στο αποτέλεσμα προστίθεται ένας όρος πόλωσης (bias) $b \in R$. Η εισαγωγή της πόλωσης προσδίδει στο νευρώνα επιπλέον υπολογιστικές δυνατότητες μάθησης και προσαρμογής στα δεδομένα εκπαίδευσης. Οι μέχρι τώρα υπολογισμοί μαθηματικά περιγράφονται από την παρακάτω σχέση, που δίνει τη συνολική είσοδο:

$$u(\mathbf{x}) = \sum_{i=1}^n x_i w_i + b \quad (3.7)$$

3. Όταν το ζυγισμένο άθροισμα των εισόδων είναι μεγαλύτερο μιας ορισμένης τιμής κατωφλίου (threshold value) θ , δηλαδή όταν

$$\sum_{i=1}^n x_i w_i - \theta > 0 \quad (3.8)$$

η συνολική είσοδος $u(\mathbf{x})$ περνάει από μια συνάρτηση ενεργοποίησης ή συνάρτηση μετάβασης (activation function) $\phi(\cdot)$, και προκύπτει η έξοδος του νευρώνα:

$$y = f(\mathbf{x}) = \phi\left(\sum_{i=1}^n x_i w_i + b\right) \quad (3.9)$$

Συνηθέστερα ως συνάρτηση ενεργοποίησης επιλέγεται μια από τις ακόλουθες[9]:

- Η συνάρτηση ReLU (Rectifier Linear Unit) η οποία ορίζεται ως:

$$\phi(x) = \max\{0, x\} \quad (3.10)$$

- Η σιγμοειδής συνάρτηση (sigmoid), η οποία περιγράφεται από την παρακάτω σχέση:

$$\phi(x) = \frac{1}{1 + e^{-x}} \quad (3.11)$$

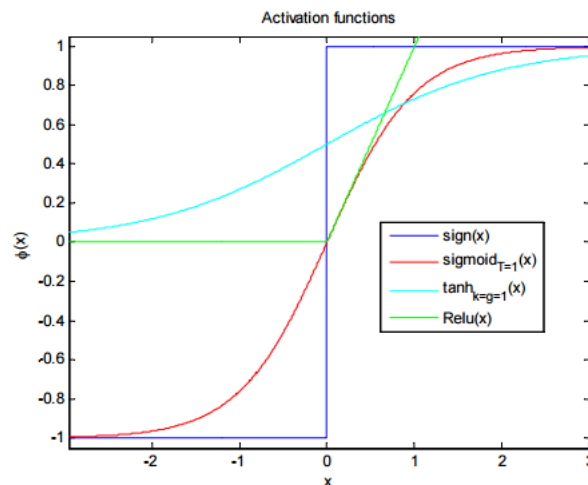
- Η συνάρτηση υπερβολικής εφαπτομένης, η οποία χρησιμοποιεί και τις παραμέτρους k και g και περιγράφεται από την παρακάτω σχέση:

$$\phi(x) = k \tanh(gx) \quad (3.12)$$

- Η συνάρτηση προσήμου



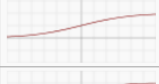
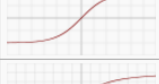

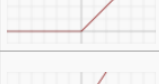
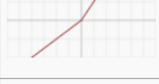


$$\phi(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \quad (3.13)$$

Στο Σχήμα 3.3 παρουσιάζονται συγκεντρωμένες οι αναφερθείσες συναρτήσεις ενεργοποίησης, ενώ στο Σχήμα 3.4 παρουσιάζονται συνοπτικά περισσότερες συναρτήσεις ενεργοποίησης. Στην πράξη η συχνότερα εμφανιζόμενη συνάρτηση ενεργοποίησης είναι η ReLU σε νευρωνι-



Σχήμα 3.3: Συνήθεις συναρτήσεις ενεργοποίησης (Πηγή: [9])

κά μοντέλα εμπρόσθιας τροφοδότησης, λόγω της μικρής υπολογιστικής της πολυπλοκότητας αλλά και μεγάλης ταχύτητας στη διαδικασία μάθησης¹, σε σύγκριση με την σιγμοειδή. Η συνάρτηση υπερβολικής εφαπτομένης πολλές φορές εμφανίζεται σαν συνάρτηση ενεργοποίησης σε νευρώνες αναδρομικών νευρωνικών μοντέλων όπως Long-Term Short-Term νευρώνες.

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Tanh		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) ^[2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) ^[3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

Σχήμα 3.4: Συναρτήσεις ενεργοποίησης (Πηγή: [10])

Αρχιτεκτονική Τεχνητών Νευρωνικών Δικτύων

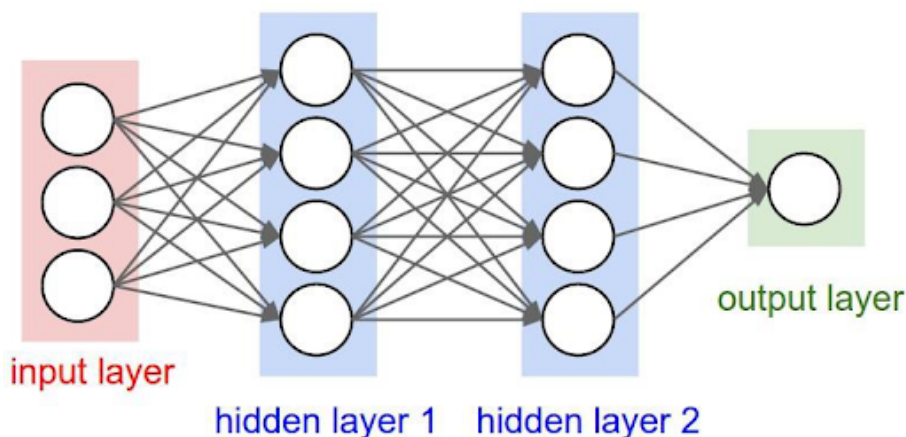
Βασικοί παράμετροι της αρχιτεκτονικής των ΤΝΔ αποτελούν [7, 39, 37]:

- (i) Το πλήθος των επιπέδων του δικτύου. Κάθε ΤΝΔ αποτελείται από τουλάχιστον δύο στρώματα, το στρώμα εισόδου (input layer), από όπου εισέρχονται τα δεδομένα εκπαίδευσης και το στρώμα εξόδου (output layer) από το οποίο παίρνουμε την έξοδο του δικτύου. Αν ένα Νευρωνικό Δίκτυο δεν αποτελείται μόνο από το στρώμα εισόδου και το στρώμα εξόδου τότε το δίκτυο ονομάζεται πολυστρωματικό (multilayer) και τα ενδιάμεσα στρώματα ονομάζονται κρυφά στρώματα (hidden layers). Γενικά όταν λέμε ότι

¹Το γεγονός ότι στο ήμισυ του πεδίου ορισμού της, η τιμή της ReLU είναι μηδέν, διευκολύνει την ύπαρξη μεγαλύτερων τιμών κλίσης (gradient) της εξόδου ενός νευρώνα ως προς τα βάρη των συνάψεών του. Αντίθετα, οι σιγμοειδείς συναρτήσεις ενεργοποίησης παρουσιάζουν κορεσμό είτε σε μία θετική είτε σε μία αρνητική τιμή για το μεγαλύτερο μέρος του πεδίου ορισμού τους, γεγονός που οδηγεί σε πολύ μικρές τιμές κλίσης, καθιστώντας τη μάθηση αργή ή ακόμα και αδύνατη.

ένα Νευρωνικό Δίκτυο είναι N- στρωμάτων συνήθως δε συμπεριλαμβάνουμε το στρώμα εισόδου.

- (ii) Το πλήθος των νευρώνων σε κάθε επίπεδο. Η φύση του προβλήματος υποδεικνύει τον αριθμό των νευρώνων που απαιτούνται. Οι μεταβλητές του προβλήματος ορίζονται ως είσοδοι του δικτύου, ενώ για την πλειοψηφία των ΤΝΔ μία έξοδος αρκεί. Το πλήθος των νευρώνων των κρυφών επιπέδων συνήθως καθορίζεται μέσω δοκιμής και σφάλματος. Τα στρώματα δεν είναι απαραίτητο να έχουν το ίδιο πλήθος νευρώνων μεταξύ τους.
- (iii) Οι συνδέσεις μεταξύ των επιπέδων. Αυτές μπορεί να διακρίνονται σε εμπρόσθια τροφοδότηση (Feed-forward Neural Networks), που τα στρώματα είναι δυνατόν να τροφοδοτούν τις εξόδους τους μόνο προς τα στρώματα που έπονται, και σε ανατροφοδότηση (αναδρομικά νευρωνικά δίκτυα- Recurrent Neural Networks), όπου η έξοδος ενός στρώματος μπορεί τροφοδοτείται τόσο σε στρώματα που έπονται όσο και σε στρώματα που προηγούνται.
- (iv) Οι συνδέσεις μεταξύ των νευρώνων. Αν κάθε νευρώνας ενός στρώματος συνδέεται με όλους τους νευρώνες του επόμενου στρώματος τότε το δίκτυο είναι πλήρως διασυνδεδεμένο (fully interconnected) και τα στρώματα ονομάζονται πλήρως συνδεδεμένα (fully connected layers).
- (v) Η μορφή της συνάρτησης ενεργοποίησης. Οι νευρώνες του ίδιου στρώματος πρέπει να έχουν τις ίδιες συναρτήσεις ενεργοποίησης.
- (vi) Η τιμή ενεργοποίησης (τιμή κατωφλίου θ).
- (vii) Οι τιμές των αρχικών βαρών μεταξύ των μονάδων.
- (viii) Οι αλγόριθμοι (κανόνες εκπαίδευσης) που χρησιμοποιούνται, για να ενισχυθούν οι σύνδεσμοι μεταξύ των μονάδων κατά τη διαδικασία της εκπαίδευσης.



Σχήμα 3.5: Παράδειγμα αρχιτεκτονικής νευρωνικού δικτύου εμπρόσθιας τροφοδότησης (Πηγή: [11])

Το Νευρωνικό Δίκτυο του Σχήματος 3.5 είναι ένα Feed-forward και fully connected δίκτυο. Η παραπάνω αρχιτεκτονική ονομάζεται πολυστρωματικό Perceptron ή MLP (Multilayer Perceptron). Το συγκεκριμένο αποτελείται από ένα στρώμα εισόδου με 3 νευρώνες, 2 κρυφά στρώματα με 4 νευρώνες το καθένα και ένα στρώμα εξόδου με μία μοναδική έξοδο. Αφού, όπως αναφέρθηκε, δεν προσμετράται το στρώμα εισόδου (αναφερόμαστε σε αυτό απλά ως είσοδος) θα λέγαμε ότι το δίκτυο αποτελείται από $4+4+1=9$ νευρώνες και $(3 \times 4)+(4 \times 4)+(4 \times 1)=32$ συνάψεις ή αλλιώς 32 βάρη και $4+4+1=9$ biases, δίνοντας το σύνολο των 41 εκπαιδευσιμων παραμέτρων.

3.2.3 Λειτουργία τεχνητών νευρωνικών δικτύων

Οι λειτουργίες που μπορεί να εκτελέσει ένα Τεχνητό Νευρωνικό Δίκτυο είναι [13, 37]:

(a) Η εκπαίδευση, η οποία ορίζεται ως η διαδικασία τροποποίησης των βαρών των συνδέσεων, ώστε δοθείσης μιας συγκεκριμένης εισόδου να μπορεί να αποδώσει μία συγκεκριμένη έξοδο. Τα δεδομένα εισόδου αποτελούν το σύνολο δεδομένων εκπαίδευσης (training set). Επιπλέον, ο αλγόριθμος εκπαίδευσης, προκειμένου να δημιουργήσει νέα γνώση, μπορεί να χρησιμοποιήσει λιγότερο ή περισσότερο την υπάρχουσα γνώση, είτε να μην τη χρησιμοποιήσει καθόλου.

(b) Η ανάκληση, η οποία ορίζεται ως η διαδικασία υπολογισμού της εξόδου βάσει συγκεκριμένων τιμών εισόδου, των δεδομένων ελέγχου (test data) και βαρών.

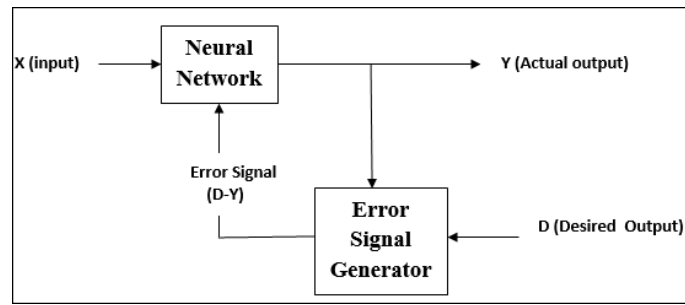
Η διαδικασία εκπαίδευσης ή μάθησης διακρίνεται στα παρακάτω είδη [7, 37]:

- i. Μάθηση υπό επίβλεψη (Supervised learning)
- ii. Μάθηση χωρίς επίβλεψη (Unsupervised learning)
- iii. Βαθμολογημένη ή Ενισχυτική Μάθηση (Graded / Reinforcement learning)

Η εκπαίδευση με επίβλεψη, η οποία υιοθετείται στην πλειοψηφία των εφαρμογών, είναι η διαδικασία όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένες εισόδους (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Συγκεκριμένα, τα δεδομένα εκπαίδευσης αποτελούνται από ζευγάρια δεδομένου εισόδου – επιθυμητής εξόδου. Τα δεδομένα εισόδου εισέρχονται στο δίκτυο και αυτό παράγει μια έξοδο. Η έξοδος του δικτύου συγκρίνεται με την επιθυμητή έξοδο και υπολογίζεται η διαφορά τους. Αυτή η διαφορά αποτελεί το σφάλμα εκπαίδευσης του δικτύου. Στη συνέχεια τα βάρη του δικτύου μεταβάλλονται έτσι ώστε η νέα έξοδος του δικτύου να προσεγγίζει με μεγαλύτερη ακρίβεια την επιθυμητή έξοδο και αυτό θα έχει σαν συνέπεια την μείωση του σφάλματος του δικτύου. Στο παρακάτω σχήμα παρουσιάζεται συνοπτικά η παραπάνω διαδικασία.

Χρησιμοποιείται σε προβλήματα:

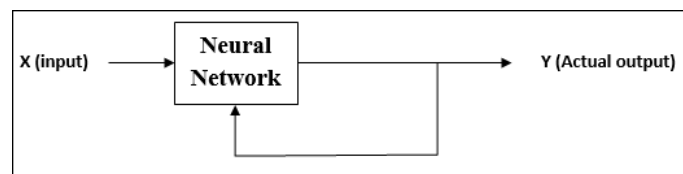
- Ταξινόμησης (Classification)
- Πρόγνωσης (Prediction)
- Διερμηνείας (Interpretation)



Σχήμα 3.6: Μάθηση με επίβλεψη (Πηγή: [12])

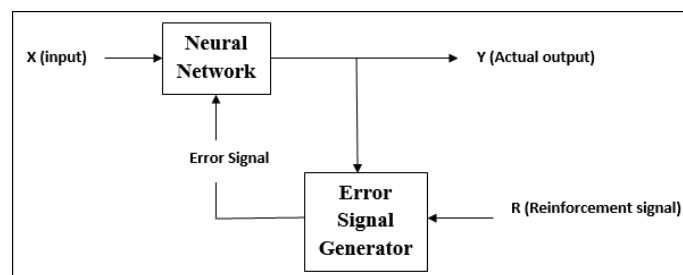
Στη μάθηση χωρίς επίβλεψη ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων, χωρίς να γνωρίζει τις επιθυμητές εξόδους. Δηλαδή, τα βάρη τροποποιούνται έτσι ώστε να ανταποκρίνονται μόνο στα δεδομένα εισόδου, ενώ δεν υπάρχουν δεδομένα που καθορίζουν κάποια επιθυμητή έξοδο. Χρησιμοποιείται σε προβλήματα:

- Ανάλυσης Συσχετισμών (Association Analysis)
- Ομαδοποίησης (Clustering)



Σχήμα 3.7: Μάθηση χωρίς επίβλεψη (Πηγή: [12])

Στην ενισχυτική μάθηση ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning) με περιορισμένα δεδομένα (feedback), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους.



Σχήμα 3.8: Ενισχυτική μάθηση (Πηγή: [12])

Στο σημείο αυτό πρέπει να τονίσουμε την σημαντική επίπτωση που είχε η ανάπτυξη του αλγορίθμου οπισθοδιάδοσης σφάλματος (Backpropagation algorithm), του συνηθέστερου, δηλαδή, τρόπου επιβλεπόμενης μάθησης, στην εξέλιξη των ΤΝΔ. Μέχρι την εξεύρεση του, τα ΤΝΔ αδυνατούσαν να λύσουν απαιτητικά προβλήματα, γεγονός που οδήγησε σε δυσμενή κριτική και κατ' επέκταση στην περιθωριοποίησή τους. Η ανάπτυξη του αλγορίθμου, στις

αρχές της δεκαετίας του 1980, έδωσε νέα ώθηση στην έρευνα στον τομέα των Τεχνητών Νευρωνικών Δικτύων. Η λειτουργία του έχει ως εξής [39, 37]: Τα δεδομένα εκπαίδευσης παρουσιάζονται στο δίκτυο και υπολογίζονται οι έξοδοι σύμφωνα με τους τύπους 3.1 - 3.9 . Για τον κάθε κόμβο i του επιπέδου εξόδου j υπολογίζεται το σφάλμα e σύμφωνα με τον τύπο:

$$e_{ij} = |y_{ij} - \hat{y}_{ij}| \quad (3.14)$$

όπου \hat{y}_{ij} είναι η έξοδος που υπολογίζει το δίκτυο και y_{ij} η επιθυμητή έξοδος.

Στην συνέχεια υπολογίζεται το προσαρμοσμένο σφάλμα του κάθε νευρώνα ως εξής:

$$\delta_{ij} = e_{ij} \cdot \phi'(u_{ij}) \quad (3.15)$$

Η διαδικασία συνεχίζεται με την οπισθοκύλιση του σφάλματος στο προηγούμενο της εξόδου επίπεδο $j - 1$:

$$\delta_{i,j-1} = \phi'(u_{i,j-1}) \cdot \sum w_{ij-1,ij} \cdot \delta_{ij} \quad (3.16)$$

όπου το άθροισμα αναφέρεται σε όλους τους νευρώνες του επόμενου επιπέδου με τους οποίους ο νευρώνας i συνδέεται με βάρη $w_{ij-1,ij}$.

Η διαδικασία συνεχίζεται έως ότου διαδοθεί τα σφάλμα μέχρι το επίπεδο εισόδου.

Η προσαρμογή των βαρών που ενώνουν τον νευρώνα i του επιπέδου $j - 1$ με τον νευρώνα i του επιπέδου j υπολογίζεται με τον τύπο:

$$\Delta w_{ij-1,ij} = -d \cdot \delta_{ij} \cdot y_{ij-1} \quad (3.17)$$

όπου d : ο ρυθμός εκπαίδευσης

Το συνολικό σφάλμα για n νευρώνες εξόδου σε m δεδομένα εκπαίδευσης αθροίζεται, ανάλογα με τη συνάρτηση σφάλματος ή κόστους (loss/cost function) που χρησιμοποιείται, πχ το μέσο τετραγωνικό σφάλμα (MSE) , ή η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE).

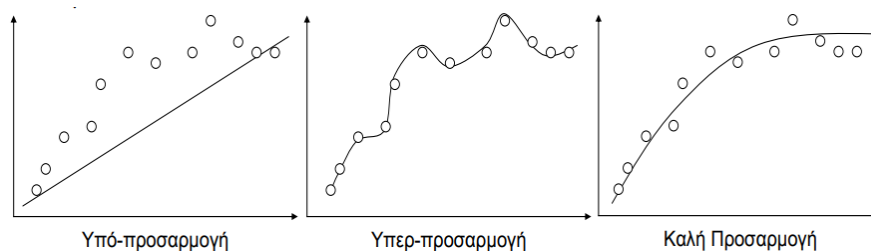
Η διαδικασία της εκπαίδευσης περιλαμβάνει πλήθος κύκλων διόρθωσης σφάλματος και λήγει με τη μείωση του σφάλματος (κόστους) κάτω από ένα επιθυμητό όριο. Εναλλακτικά, ως συνθήκη τερματισμού μπορεί να θεωρηθεί η πραγματοποίηση ενός συγκεκριμένου αριθμού κύκλων εκπαίδευσης ή η πάροδος ενός συγκεκριμένου χρονικού διαστήματος.

Η παρουσίαση όλων των παραδειγμάτων (samples) του training set μια φορά αντιστοιχεί σε έναν κύκλο εκπαίδευσης και ονομάζεται εποχή εκπαίδευσης (training epoch).

Τέλος, να αναφέρουμε ότι η εκπαίδευση ενός ΤΝΔ δεν δίνει πάντα επιθυμητά αποτελέσματα. [13] Παραδείγματα μη επιθυμητής εκπαίδευσης αποτελούν η:

- υποπροσαρμογή (underfitting), δηλαδή αποτυχία μοντελοποίησης των δεδομένων εκπαίδευσης , πχ λόγω απλότητας του ΤΝΔ
- και η υπερπροσαρμογή (overfitting), δηλαδή η υπερβολική μοντελοποίηση των δεδομένων εκπαίδευσης (απομνημόνευση) και αδυναμία πρόβλεψης για άλλα δεδομένα εισόδου.

Μια μέθοδος αντιμετώπισης της υπερπροσαρμογής είναι η προσθήκη ενός Dropout Layer, ενός ειδικού στρώματος, δηλαδή, που αγνοεί νευρώνες κατά την εκπαίδευση στη φάση της εμπρόσθιας τροφοδότησης με βάση μια πιθανότητα, έτσι ώστε να μη δημιουργούνται σχέσεις αλληλεξάρτησης μεταξύ νευρώνων, που θα οδηγήσουν σε overfitting.

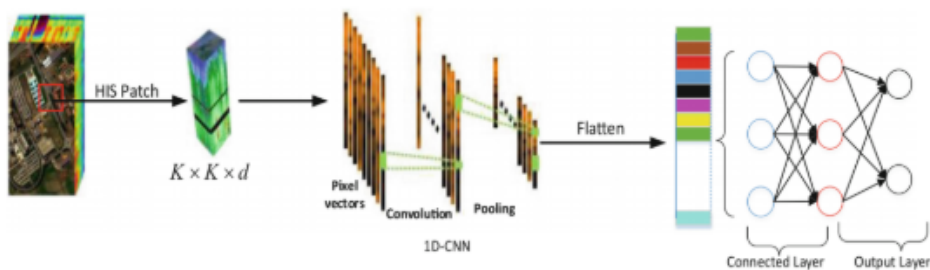


Σχήμα 3.9: Αναπαράσταση φαινομένων υποπροσαρμογής, υπερπροσαρμογής και καλής προσαρμογής (Πηγή: [13])

3.2.4 Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)

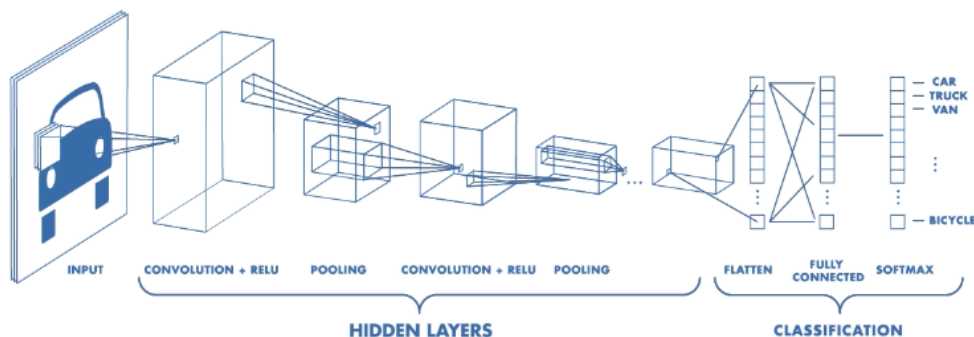
Τα συνελικτικά νευρωνικά δίκτυα (CNN) αποτελούν δημοφιλή κατηγορία νευρωνικών δικτύων εμπρόσθιας τροφοδότησης, κυρίως για πολυδιάστατα δεδομένα (πχ εικόνα, βίντεο). Λειτουργούν με αντίστοιχο τρόπο που περιγράφηκε ήδη, με τη βασική διαφορά ότι εδώ η βασική δομική μονάδα είναι ένα φίλτρο. Συγκεκριμένα, ένα CNN χρησιμοποιεί σε κάθε Συνελικτικό Στρώμα (convolutional layer) έναν αριθμό διαφορετικών φίλτρων με τα οποία συνελίσσεται η είσοδος και παράγεται ένα χάρτης ενεργοποίησης ή χάρτες χαρακτηριστικών (activation / feature map). Ο χάρτης ενεργοποίησης που προκύπτει από την συνέλιξη με κάθε φίλτρο είναι μονοδιάστατος στην περίπτωση που τα δεδομένα είναι χρονοσειρές, δισδιάστατος στην περίπτωση των εικόνων και τρισδιάστατος στην περίπτωση των βίντεο. Ο χάρτης ενεργοποίησης μπορεί να νοηθεί ως εξαγωγή χρήσιμων χαρακτηριστικών. Τα βάρη του κάθε φίλτρου αποτελούν τις υπό μάθηση παραμέτρους του μοντέλου. Μέχρι αυτό το σημείο, η δράση του συνελικτικού στρώματος στην είσοδο είναι μόνο γραμμική καθώς η συνέλιξη αποτελεί γραμμικό τελεστή. Στη συνέχεια, σε κάθε τιμή του χάρτη ενεργοποίησης εφαρμόζεται μία μη γραμμική συνάρτηση ενεργοποίησης και έτσι προκύπτει η τελική έξοδος του στρώματος.

Τα φίλτρα εντάσσουν στο δίκτυο την τοπολογία των αρχικών χαρακτηριστικών τον δεδομένων εισόδου, έχοντας ίδια -αλλά μικρότερη- μορφή, και λόγω κοινής χρήσης παραμέτρων (parameter sharing), μειώνουν το πλήθος των παραμέτρων εκπαίδευσης. Η κοινή χρήση παραμέτρων βασίζεται αποκλειστικά στον τρόπο υπολογισμού της συνέλιξης, έτσι ώστε οι παράμετροι ενός φίλτρου να χρησιμοποιούνται (σχεδόν) σε κάθε σημείο της εισόδου χωρίς να αλλάζουν. Διαισθητικά, για το παράδειγμα των εικόνων ως δεδομένων εισόδου, μπορούμε να περιγράψουμε την ιδιότητα αυτή αν παρομοιάσουμε τη συνέλιξη με την μεταφορά και εφαρμογή του φίλτρου πάνω σε διαφορετικές γειτονιές pixel της εικόνας. Σε αυτή την περίπτωση, οι παράμετροι που χρησιμοποιούνται για την εκτέλεση υπολογισμών χρησιμοποιώντας ένα εκ των φίλτρων σε κάθε γειτονιά σημείων είναι οι ίδιες. Ως εκ τούτου, ένα συνελικτικό δίκτυο έχει τη δυνατότητα να εντοπίζει και να εξάγει τα ίδια ακριβώς οπτικά χαρακτηριστικά από διαφορετικές περιοχές της εισόδου. Αντίθετα, ένα πλήρως συνδεδεμένο μοντέλο θα χρησιμοποιούσε μία διαφορετική παράμετρο για κάθε διαφορετική θέση της εισόδου. Επιπλέον, ο αυτόματη εξαγωγή των σημαντικών χαρακτηριστικών των δεδομένων μειώνει την ανάγκη για χειροκίνητη σχεδίαση του δικτύου και ανθρωπίνη παρέμβαση. [9, 7, 40]



Σχήμα 3.10: Απεικόνιση 1D CNN (Πηγή: [14], σελ 83)

Εκτός των στρώματων συνέλιξης, ένα τυπικό CNN αποτελείται από συγκεντρωτικά επίπεδα (pooling layers) και από πλήρως συνδεδεμένα επίπεδα (fully connected layers). Το συγκεντρωτικό επίπεδο είναι το επίπεδο που επιτυγχάνει την μείωση της διάστασης του μοντέλου, και συνεπώς και των εκπαιδευόμενων παραμέτρων. Συγκεκριμένα, η έξοδος του προηγούμενου στρώματος συνέλιξης υφίσταται τη δράση ενός τελεστή συσσώρευσης - συνήθως μια συνάρτηση τοπικού μεγίστου (max pooling layer) που εφαρμόζεται σε μη επικαλυπτόμενες υποπεριοχές του τανυστή εισόδου - και έτσι διεξάγεται η δειγματοληψία, κατά την οποία μειώνεται η διαστασιοποίηση των ανά επίπεδο χαρτογραφημένων χαρακτηριστικών (feature maps). Αξίζει να σημειώσουμε ότι στο συγκεντρωτικό επίπεδο δεν υπάρχουν παράμετροι προς εκπαίδευση και τον αντίστοιχο ρόλο τον διαδραματίζουν υπερπαραμέτροι, όπως το μέγεθος του φίλτραρίσματος και το βήμα. Τα τελικά χαρτογραφημένα χαρακτηριστικά μετατρέπονται σε ένα μονοδιάστατο διάνυσμα με τη χρήση της συνάρτησης flatten (ένα flatten layer χωρίς παραμέτρους, μόνο για μετατροπή του πολυδιάστατου χάρτη σε μονοδιάστατο διάνυσμα) και συνδέονται σε ένα ή περισσότερα πλήρως συνδεδεμένα επίπεδα (fully-connected layers). Οι νευρώνες αυτών των στρώματων είναι συνδεδεμένοι με όλους τους νευρώνες του προηγούμενου στρώματος όπως ακριβώς και στα τυπικά ΤΝΔ. Η ενεργοποίηση του κάθε νευρώνα λοιπόν μπορεί να υπολογιστεί, κατά τα γνωστά, με πολλαπλασιασμό πινάκων ακολουθούμενο από το άθροισμα ενός bias. Κάθε ένα από τα πλήρως συνδεδεμένα επίπεδα, που καλούνται και πυκνά

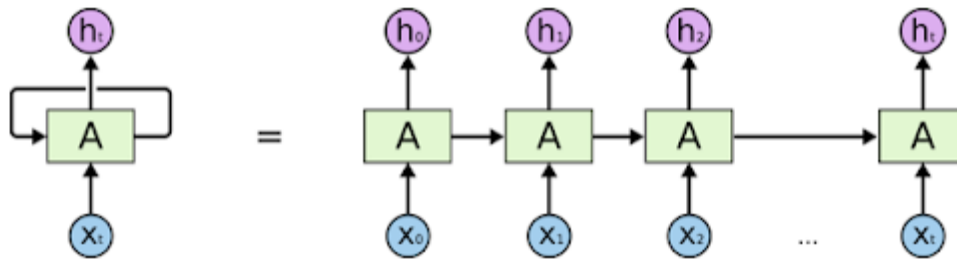


Σχήμα 3.11: Παράδειγμα αρχιτεκτονικής CNN (Πηγή: [7])

(dense) στρώματα, ακολουθείται συνήθως από μία μη γραμμική συνάρτηση ενεργοποίησης. Η επιλογή της συνάρτησης εξαρτάται από το πρόβλημα που αντιμετωπίζουμε, αλλά η πιο συνηθισμένη περίπτωση που εισάγει τους λιγότερους κινδύνους είναι η ReLU. Το τελευταίο πλήρως συνδεδεμένο επίπεδο (fully connected layer) συνήθως έχει τον ίδιο αριθμό κόμβων εξόδου με των αριθμό των κλάσεων, αν πρόκειται για πρόβλημα ταξινόμησης, ή με το ζητούμενο πλήθος τιμών προβλέψεων, αν πρόκειται για πρόβλημα πρόγνωσης. [9, 7, 41]

3.2.5 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks) και νευρώνες Μακράς-Βραχείας Μνήμης (Long Term Short Term Memory)

Τα αναδρομικά νευρωνικά δίκτυα είναι μία ειδική κατηγορία νευρωνικών δικτύων που επεξεργάζονται αποτελεσματικά κάθε είδους ακολουθιακά δεδομένα, με κύριο γνώρισμα τους την ιδιότητα τους να συνδέουν προηγούμενες εισόδους με την εκάστοτε τρέχουσα. Με αυτό τον τρόπο, ένα κόμβος του δικτύου έχει τη δυνατότητα σε κάθε χρονική στιγμή να παίρνει τα τρέχοντα δεδομένα εισόδου, παράλληλα με τις τιμές των κρυμμένων κόμβων, συλλέγοντας έτσι πληροφορίες από προηγούμενες χρονικές στιγμές. Παραδείγματα ακολουθιακών δεδομένων αποτελούν η φωνή, η γραφή, η οπτική πληροφορία που προκύπτει από μία κίνηση ή μία δράση ή ακόμα και τα pixel μίας εικόνας αν τα διατρέξουμε με κάποιο δομημένο τρόπο, και φυσικά οι χρονοσειρές.



An unrolled recurrent neural network.

Σχήμα 3.12: «Ξεδίπλωμα» ενός αναδρομικού δικτύου. (Πηγή: [9])

Αν x_t είναι η είσοδος την χρονική στιγμή t , $f(\cdot)$ η συνάρτηση που περιγράφει την επίδραση του αναδρομικού δικτύου πάνω στην είσοδο και h_t η έξοδος του, τότε:

$$h_t = f(x_t, h_{t-1}) = f(x_t, f(x_{t-1}, h_{t-2})) = \dots = f(x_t, f(x_{t-1}, \dots, f(x_1, h_0) \dots)) \quad (3.18)$$

Η έξοδος h_t συχνά αναφέρεται και ως κατάσταση του δικτύου. Αν ληφθούν υπόψη και οι παράμετροι υπό μάθηση, δηλαδή τα βάρη δύο ειδών: αυτά που επιδρούν πάνω στην είσοδο (W) και αυτά που καθορίζουν την σημασία που δίνεται στην προηγούμενη κατάσταση του μοντέλου για τον υπολογισμό της επόμενης κατάστασης (U), προκύπτει η γενική μορφή των εξισώσεων ενός αναδρομικού δικτύου:

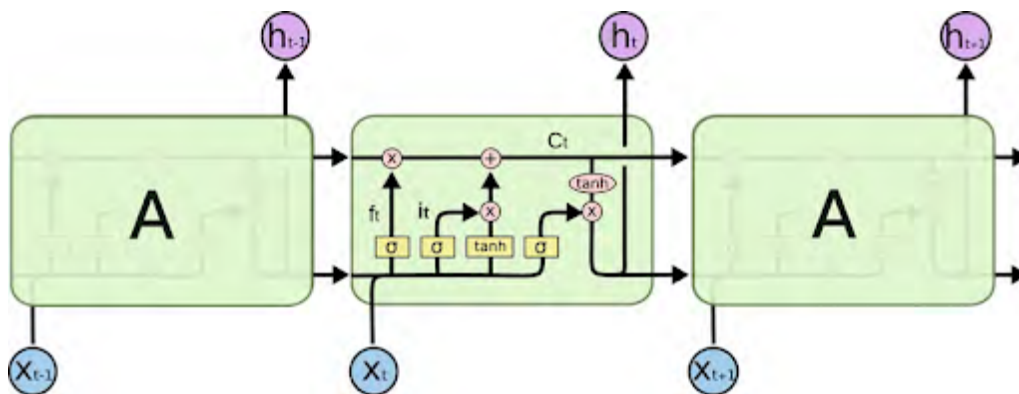
$$h_t = \phi(Wx_t + Uh_{t-1} + b) \quad (3.19)$$

όπου $\phi(\cdot)$ μία μη γραμμική συνάρτηση ενεργοποίησης, W ο πίνακας παραμέτρων που επιδρούν πάνω στην είσοδο x_t , U ο πίνακας παραμέτρων που επιδρούν πάνω στην έξοδο του δικτύου την προηγούμενη χρονική στιγμή και b ένα διάνυσμα πόλωσης.

Λόγω αυτού του μηχανισμού τα αναδρομικά δίκτυα έχουν τη δυνατότητα να μοντελοποιούν χρονικές εξαρτήσεις ακόμα και ανάμεσα σε μη συνεχόμενες παρατηρήσεις, αφού μέσω της αναδρομικότητας υλοποιούν έναν μηχανισμό μνήμης. Ωστόσο, όταν πρόκειται για μακρινές εξαρτήσεις, δηλαδή για εξαρτήσεις μεταξύ δύο χρονικών βημάτων που απέχουν αρκετά μεταξύ τους, τα αναδρομικά δίκτυα αδυνατούν να τις διατηρήσουν. Αυτό οφείλεται στο γεγονός ότι κατά την διάρκεια της εκπαίδευσης, για να καθοριστεί το σφάλμα και επομένως και η εξάρτηση της εξόδου της χρονικής στιγμής t σε σχέση με την είσοδο της χρονικής στιγμής $t - n$, θα παρεμβληθεί γινόμενο n παραγώγων σύμφωνα με τον κανόνα της αλυσίδας. Αυτό έχει ως αποτέλεσμα, όταν οι κλίσεις είναι μικρές, δηλαδή της τάξης 10^{-1} ή και μικρότερες, το συνολικό γινόμενο να τείνει στο μηδέν, με αποτέλεσμα η επιρροή μια αρχικής χρονικής στιγμής σε μια μακρινή μεταγενέστερη στιγμή, ουσιαστικά να εξαφανίζεται. Το πρόβλημα αυτό συνήθως αναφέρεται ως εξαφάνιση ή εκτόξευση της κλίσης (the vanishing or exploding gradient problem) και έχει σαν αποτέλεσμα η διαδικασία εκπαίδευσης να μην εντοπίζει ικανοποιητικά καλές τιμές για τα βάρη του μοντέλου.

Η παραπάνω δυσλειτουργία των αναδρομικών δικτύων, αντιμετωπίστηκε μέσω της εισαγωγής των νευρώνων Μακράς-Βραχείας Μνήμης (Long Term Short Term Memory ή L.S.T.M) που προτάθηκαν πρώτη φορά από τους Hochreiter και Schmidhuber.

Η βασική μονάδα στο κρυμμένο επίπεδο ενός LSTM τεχνητού νευρωνικού δικτύου είναι το μπλοκ μνήμης (memory block), το οποίο αντικαθιστά τις κρυφές μονάδες των κοινών RNN. Ένα μπλοκ μνήμης περιέχει ένα ή περισσότερα κελιά μνήμης (memory cells) και ένα ζευγάρι από προσαρμοστικές, πολλαπλασιαστικές μονάδες πύλης, οι οποίες χορηγούν είσοδο στις εισροές και εκροές όλων των κελιών του μπλοκ. Η δομή ενός κελιού LSTM φαίνεται στο Σχήμα 3.13.



Σχήμα 3.13: Δομή κρυφού κυττάρου LSTM. (Πηγή: [15])

Η βασική ιδέα για τα LSTMs είναι η κατάσταση του κυττάρου c_t και η δυνατότητα αφαίρεσης ή προσθήκης πληροφοριών σε αυτή, μέσα από ορισμένες πύλες. Ένα LSTM έχει τρεις πύλες, για να προστατεύει και να ελέγχει την κατάσταση του κυττάρου.

- Η θύρα λησμόνησης - διαγραφής μνήμης (forget gate) f_t ελέγχει το κατά πόσο πρέπει το περιεχόμενο του κυττάρου μνήμης c_t να ξεχαστεί κατά τη μετάβαση στο επόμενο βήμα λειτουργίας. Συγκεκριμένα, λαμβάνοντας το σήμα ανάδρασης και την τρέχουσα είσοδο x_t , το νευρωνικό δίκτυο με την σιγμοειδή ενεργοποίηση παράγει έναν αριθμό μεταξύ 0 και 1. Η έξοδος αυτή πολλαπλασιάζεται με κάθε αριθμό του διανύσματος C_{t-1} της προηγούμενης κατάστασης, ρυθμίζοντας έτσι το ποια πληροφορία θα «ξεχαστεί».
- Η θύρα εισόδου (input gate) i_t ελέγχει το ποια διανύσματα του νευρώνα θα μεταβληθούν και σε ποιο βαθμό. Η είσοδος στο κύτταρο είναι η τρέχουσα είσοδος του νευρωνικού x_t και η προηγούμενη κρυφή κατάσταση h_{t-1} .
- Η θύρα εξόδου (output gate) o_t ελέγχει την έκθεση της εσωτερικής μνήμης στα υπόλοιπα μέρη του νευρώνα. Συγκεκριμένα, η έξοδος θα βασιστεί στην τρέχουσα κατάσταση του κυττάρου και θα αποτελεί μια φιλτραρισμένη εκδοχή της. Αρχικά συμπιέζουμε την κατάσταση μέσω της συνάρτησης \tanh . Στη συνέχεια η είσοδος περνάει από το νευρωνικό δίκτυο της πύλης εξόδου με σιγμοειδή ενεργοποίηση και το αποτέλεσμα αυτό (o_t) συνδυάζεται μέσω του πολλαπλασιαστή με την συμπιεσμένη κατάσταση κυττάρου, αποφασίζοντας έτσι ποια μέρη της κατάστασης αυτής θα συμμετέχουν στην τελική έξοδο.
- Η κρυφή κατάσταση (hidden state) h_t αναπαριστά μία περιορισμένη και ελεγχόμενη εξωτερίκευση της αποθηκευμένης στο κύτταρο πληροφορίας η οποία εξαρτάται και από την νέα είσοδο x_t και όλες τις προηγούμενες λειτουργίες.

Δεδομένου ότι οι μη γραμμικές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται είναι η σιγμοειδής συνάρτηση και η υπερβολική εφαιπτομένη \tanh , οι μαθηματικές εξισώσεις που χαρακτηρίζουν έναν νευρώνα LSTM είναι οι εξής:

$$i_t = \sigma(W^i x_t + U^i h_{t-1} + b^i) \quad (3.20)$$

$$f_t = \sigma(W^f x_t + U^f h_{t-1} + b^f) \quad (3.21)$$

$$o_t = \sigma(W^o x_t + U^o h_{t-1} + b^o) \quad (3.22)$$

$$u_t = \tanh(W^u x_t + U^u h_{t-1} + b^u) \quad (3.23)$$

$$c_t = i_t \odot u_t + f_t \odot c_{t-1} \quad (3.24)$$

$$h_t = o_t \odot \tanh(c_t) \quad (3.25)$$

όπου x_t η είσοδος τη χρονική στιγμή t , b ένα διάνυσμα πόλωσης και \odot το γινόμενο Hadamard. [9, 42]

3.3 Σφάλματα πρόβλεψης - Μέτρα απόδοσης

Σφάλμα υπάρχει σε κάθε πρόβλεψη που υλοποιείται. Τα σφάλματα των προβλέψεων διακρίνονται σε συστηματικά, που αφορούν το μοντέλο πρόβλεψης και οφείλονται στην κακή εκτίμηση ή παράλειψη παραμέτρων, και σε τυχαία σφάλματα, που οφείλονται σε μη προβλέψιμους παράγοντες. Τα σφάλματα χρησιμοποιούνται ως μέτρο απόδοσης των μεθόδων πρόβλεψης. Αρνητική τιμή σφάλματος υποδηλώνει υπερεκτίμηση των τιμών της χρονοσειράς, ενώ θετική τιμή δείχνει υποεκτίμηση των τιμών της. Στις περισσότερες περιπτώσεις υπολογίζονται απόλυτες τιμές σφαλμάτων, εκτός αν οι ανάγκες του προβλήματος υποδεικνύουν διαφορετικά. Έστω πλήθος δεδομένων N , που αντιστοιχούν σε N περιόδους, για τις οποίες διατίθενται οι τιμές πρόβλεψης F_t και οι αντίστοιχες πραγματικές τιμές Y_t τη χρονική στιγμή t , $t = 1, 2, \dots, N$. Παρακάτω ορίζονται ορισμένα είδη σφαλμάτων για τη δυνατή σύγκριση και κατανόηση των αποκλίσεων των προβλέψεων. [43]

- Σφάλμα Πρόβλεψης (Forecasting Error) ή Υπόλοιπο (Residual)

$$e_t = Y_t - F_t \quad (3.26)$$

- Μέσο Σφάλμα (Mean Error)

$$ME = \frac{1}{N} \sum_{t=1}^N (Y_t - F_t) = \frac{1}{N} \sum_{t=1}^N e_t \quad (3.27)$$

Για να έχουμε αποτελεσματική πρόβλεψη πρέπει $ME \rightarrow 0$, ενώ μεγάλη θετική/αρνητική τιμή του ME εκφράζει υποεκτίμηση/ υπερεκτίμηση των τιμών της χρονοσειράς. Έχει το μειονέκτημα ότι αν τα σφάλματα αλλάζουν πρόσημο, στο άθροισμα τους θα αλληλοαναιρούνται και έτσι είναι πιθανό να προκύψει μέσο σφάλμα κοντά στο 0, χωρίς ωστόσο οι προβλέψεις να είναι ακριβείς, επομένως δεν χρησιμοποιείται ευρέως στην πράξη.

- Μέσο Απόλυτο Σφάλμα (Mean Absolute Error) ή Μέση Απόλυτη Απόκλιση (Mean Absolute Deviation).

$$MAE = \frac{1}{N} \sum_{t=1}^N |Y_t - F_t| = \frac{1}{N} \sum_{t=1}^N |e_t| \quad (3.28)$$

Εκφράζει τη μέση τιμή των απόλυτων αποκλίσεων των προβλεπόμενων τιμών της χρονοσειράς και είναι, επομένως, απαλλαγμένο από το μειονέκτημα του Μέσου Σφάλματος ME . Μετράει το σφάλμα της πρόβλεψης στις ίδιες μονάδες με τις τιμές της χρονοσειράς και έτσι είναι εύκολη η ερμηνεία του. Το MAE βασίζεται στην υπόθεση ότι η σοβαρότητα του σφάλματος ή το κόστος που δημιουργείται από το σφάλμα της πρόβλεψης σχετίζεται γραμμικά με το μέγεθος του σφάλματος.

- Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error) ή Μέση Τετραγωνική Απόκλιση (Mean Squared Deviation)

$$MSE = \frac{1}{N} \sum_{t=1}^N (Y_t - F_t)^2 = \frac{1}{N} \sum_{t=1}^N e_t^2 \quad (3.29)$$

Εκφράζει τη μέση τιμή των τετραγώνων των αποκλίσεων των προβλεπόμενων τιμών της χρονοσειράς από τις αντίστοιχες πραγματικές και μετρά τη διασπορά της κατανομής των σφαλμάτων πρόβλεψης. Παρέχει πληροφορίες παρόμοιες με το Μέσο Απόλυτο Σφάλμα (MAE), κάνοντας όμως περισσότερο αισθητή την ύπαρξη προβλέψεων που απέχουν αρκετά από τις αντίστοιχες πραγματικές τιμές, καθώς μεγεθύνει τις μεγάλες αποκλίσεις λόγω της ύψωσης στο τετράγωνο. Έτσι, μια μέθοδος πρόβλεψης με μέτρια σφάλματα μπορεί να θεωρηθεί καλύτερη σε σχέση με μια άλλη με μικρότερα σφάλματα αλλά και σποραδικά με κάποια μεγάλα. Για αυτό, χρησιμοποιείται όταν επιθυμούμε πολλές μικρές αποκλίσεις έναντι μιας πολύ μεγάλης. Χρησιμοποιείται συχνότερα από το κριτήριο Μέσο Απόλυτο Σφάλμα (MAE) για την αξιολόγηση και τελικά την επιλογή της «κατάλληλη» μεθόδου πρόβλεψης. Έχει ως μονάδα μέτρησης τη μονάδα μέτρησης των τιμών των παρατηρήσεων υψωμένη, όμως, στο τετράγωνο. Για αυτό μπορούμε να χρησιμοποιούμε την τετραγωνική του ρίζα, που ονομάζεται Ρίζα Μέσου Τετραγωνικού Σφάλματος RMSE.

- Ρίζα Μέσου Τετραγωνικού Σφάλματος (Root Mean Squared Error)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (Y_t - F_t)^2} = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2} \quad (3.30)$$

Έχει ίδια μονάδα μέτρησης με εκείνη των τιμών της χρονολογικής σειράς.

- Ποσοστιαίο Σφάλμα (Percentage Error)

$$PE = \frac{Y_t - F_t}{Y_t} \cdot 100\% \quad (3.31)$$

Μετρείται για μια μόνο περίοδο.

- Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{Y_t - F_t}{Y_t} \right| \cdot 100\% = \frac{1}{N} \sum_{t=1}^N \left| \frac{e_t}{Y_t} \right| \cdot 100\% \quad (3.32)$$

Εκφράζει το άθροισμα των απόλυτων τιμών των σφαλμάτων της πρόβλεψης προς τις αντίστοιχες πραγματικές τιμές της χρονοσειράς διαιρούμενο με τον αριθμό των χρονικών περιόδων N , στις οποίες έγιναν προβλέψεις. Μελετά το πώς συμπεριφέρεται η απόλυτη τιμή του σφάλματος της πρόβλεψης σε σχέση με την πραγματική τιμή της χρονοσειράς. Μετράει, δηλαδή, το πόσο έξω πέφτουν οι προβλέψεις ως ποσοστά της πραγματικής τιμής της μεταβλητής. Το κριτήριο αυτό είναι απαλλαγμένο από μονάδες μέτρησης και είναι χρήσιμο όταν για την αξιολόγηση της μεθόδου έχει σημασία το μέγεθος της μεταβλητής πρόβλεψης. Μάλιστα, το χρησιμοποιούμε και για τη σύγκριση της ακρίβειας μιας ή περισσοτέρων μεθόδων πρόβλεψης και για περισσότερες από μια χρονολογικές σειρές.

Κεφάλαιο 4

Πρόβλεψη οριακής τιμής ηλεκτρικής ενέργειας της επόμενης ημέρας

Σε αυτό το κεφάλαιο αναπτύσσονται οι λεπτομέρειες της εφαρμογής των μεθόδων πρόβλεψης της οριακής τιμής της ηλεκτρικής ενέργειας στην Αγορά Επόμενης Ημέρας, που αφορούν τόσο την υλοποίηση των μεθόδων, όσο και τα αποτελέσματα αυτών. Μετράται η επίδοση του κάθε μοντέλου, μέσω του σφάλματος πρόβλεψης και στο τέλος συγκρίνονται οι μέθοδοι. Σε όλους τους απαραίτητους υπολογισμούς για το χειρισμό των δεδομένων, την εφαρμογή των στατιστικών τεχνικών προβλέψεων καθώς και τη σχεδίαση και την υλοποίηση των νευρωνικών μοντέλων χρησιμοποιήθηκε η γλώσσα Python έκδοση 3.7.4 και το προγραμματιστικό περιβάλλον Spyder 3.3.6. Η σχεδίαση των νευρωνικών μοντέλων έγινε με χρήση Keras, ενός high-level API της TensorFlow, κατάλληλου για την κατασκευή και εκπαίδευση μοντέλων βαθιάς μάθησης.

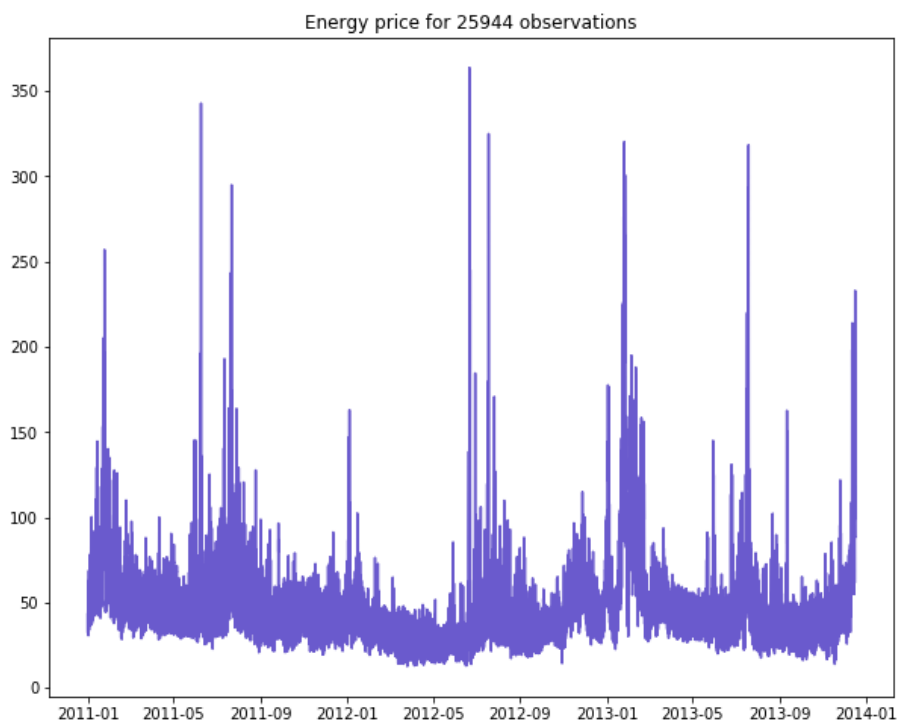
4.1 Προεπεξεργασία δεδομένων

Τα δεδομένα που έχουμε στη διάθεση μας περιλαμβάνουν την τιμή της ηλεκτρικής ενέργειας ανά ώρα για όλες τις ημέρες μεταξύ του διαστήματος 01/01/2011 έως 16/12/2013 (στο σύνολο δηλαδή έχουμε μια χρονοσειρά με 25944 παρατηρήσεις). Οι τιμές αφορούν άγνωστη

mean	48.146034
std	26.142308
min	12.52000
25%	33.467500
50%	42.860000
75%	54.24
max	363.8

Πίνακας 4.1: Στατιστικά στοιχεία δεδομένων

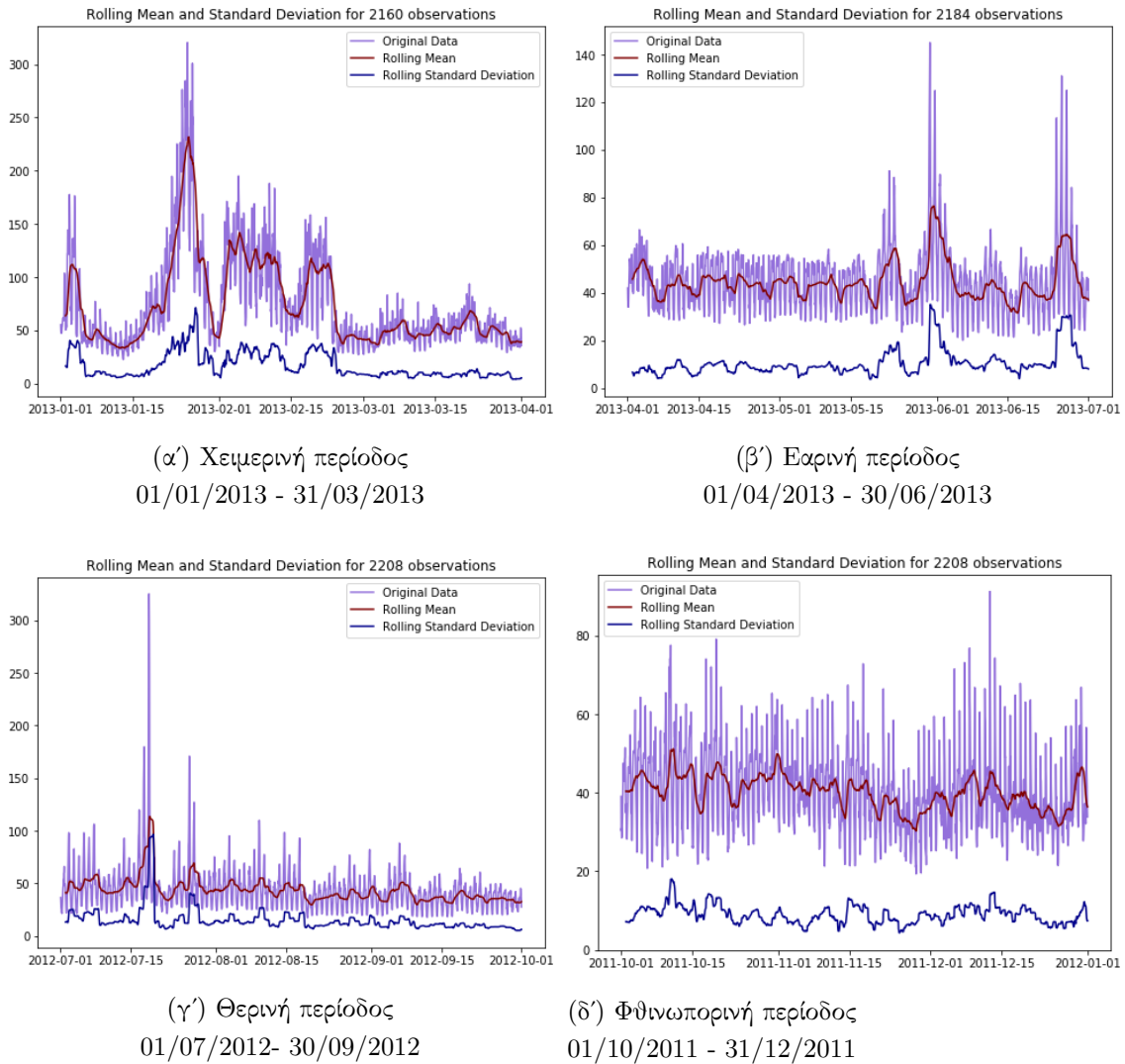
γεωγραφική ζώνη και αγνοούμε και τη νομισματική μονάδα. Η εξέλιξη της οριακής τιμής της ηλεκτρικής ενέργειας συναρτήσει του χρόνου φαίνεται στο παρακάτω σχήμα. Τα δεδομένα



Σχήμα 4.1: Οπτικοποίηση δεδομένων

χωρίστηκαν σε τρίμηνα, με σκοπό στις διάφορες μεθόδους προβλέψεων να χρησιμοποιήσουμε τους δύο πρώτους μήνες για εκπαίδευση και τον τρίτο μήνα για πρόβλεψη των τιμών ενέργειας. Έτσι, προέκυψαν δώδεκα περίοδοι, από τις οποίες επιλέξαμε να μελετήσουμε τέσσερις. Η επιλογή έγινε με πρωταρχικό κριτήριο να μελετηθούν όλες οι εποχές του χρόνου, και επιπλέον σημαντικό ρόλο έπαιξε η μορφή των δεδομένων, καθώς όσο πιο μεγάλη στασιμότητα παρουσίαζαν, τόσο πιο εύστοχες θα ήταν οι προβλέψεις. Στο στατιστικό μοντέλο ARIMA, όπως είδαμε, η στασιμότητα θεωρείται προϋπόθεση, και τα νευρωνικά μοντέλα, τα οποία είναι μεν πιο αποδοτικά στην πρόβλεψη ραγδαίων αλλαγών [24], έχουν καλύτερη επίδοση σε πιο στάσιμα δεδομένα.[44]

Η προεπεξεργασία των δεδομένων περιλάμβανε δύο στάδια, τη μελέτη γραφικών παραστάσεων και την εφαρμογή στατιστικών ελέγχων (ADF και KPSS). Η μορφή των δεδομένων για τα τρίμηνα που επιλέχθηκαν, καθώς και ο κινούμενος μέσος και κινούμενη τυπική απόκλιση φαίνονται στο Σχήμα 4.2. Από την οπτικοποίηση των δεδομένων διαπιστώσαμε ότι ο κινούμενος μέσος και η τυπική απόκλιση δεν είναι σταθερές τιμές, αλλά είναι συνάρτηση του χρόνου. Αυτό βέβαια συμβαίνει συνήθως στις οικονομικές χρονοσειρές στην πράξη, καθώς δεν είναι επιβίβαστο να παρουσιάζεται απόλυτη στασιμότητα.

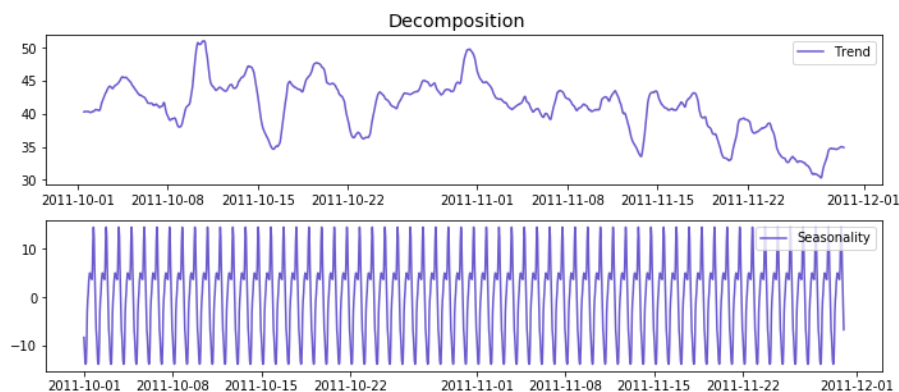


Σχήμα 4.2: Μορφή δεδομένων, κινούμενου μέσου και κινούμενης τυπικής απόκλισης

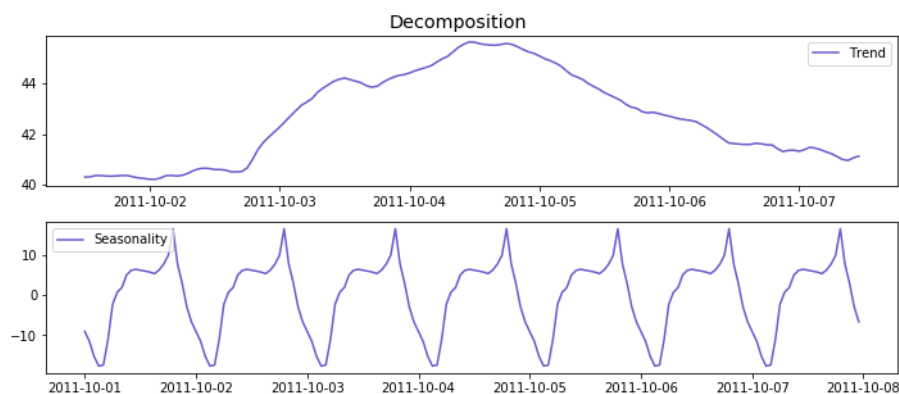
Σχεδιάστηκαν επίσης οι γραφικές παραστάσεις της αποσύνθεσης (Decomposition) - ανάλυσης με προσθετικό μοντέλο και των αυτοσυσχετίσεων, με τη βοήθεια συναρτήσεων του πακέτου statsmodels που παρέχει η ρηθον. Από την ανάλυση στις συνιστώσες της χρονοσειράς παρουσιάζονται η κλίση και η εποχικότητα, ενδεικτικά για την φθινοπωρινή περίοδο, τόσο για διάστημα δύο μηνών - όπου αποτυπώνεται καλύτερα το μοτίβο της κλίσης, όσο και για επτά ημέρες - όπου είναι φανερό το μοτίβο της εποχικότητας. Οι συναρτήσεις αυτοσυσχέτισης και μερικής αυτοσυσχέτισης αφορούν την εαρινή περίοδο, και υπολογίστηκαν για υστέρηση μέχρι 150.

Αποτυπώνεται μέσω αυτών των γραφικών παραστάσεων μια περιοδικότητα στη συμπεριφορά της τιμής της ηλεκτρικής ενέργειας, με συχνότητα 24. Παρόλο που η ακριβής μορφή της συνάρτησης αυτοσυσχέτισης ή της καμπύλης εποχικότητας δεν ήταν η ίδια για τις 4 περιόδους που μελετήθηκαν, όλες παρουσίασαν εποχικότητα με περίοδο μίας ημέρας. Όσον αφορά το πλάτος της αυτοσυσχέτισης, από το σχήμα 4.5 βλέπουμε ότι ακολουθεί ημιτονοειδές

μοτίβο και φθίνει σταδιακά. Εμπειρικά, για αυτοσυσχέτιση μικρότερη από 0.5 η χρονοσειρά δεν χρειάζεται διαφορίση.[6] Επιπλέον, η μερική αυτοσυσχέτιση για όλες τις περιόδους έθινε πολύ γρήγορα μετά τις πρώτες δύο υστερήσεις, παρουσίαζε μερικές τιμές στατιστικά σημαντικές για υστερήσεις έως 24 και είχε μεμονομένα αυξημένη τιμή κυρίως για υστέρηση 24 ($y_{t-1}, y_{t-2}, \dots, y_{t-24}$). Από τις παραπάνω παρατηρήσεις καταλαβαίνουμε ότι αν y_t είναι η τιμή που θέλουμε να προβλέψουμε, τότε είναι σημαντικό να γνωρίζουμε τις προηγούμενες 24 τιμές της χρονοσειράς. Συνεπώς, για την πρόβλεψη των ωριαίων τιμών της επόμενης ημέρας είναι σκόπιμο να χρησιμοποιήσουμε τις προηγούμενες τιμές της ίδιας της χρονοσειράς, και συγκεκριμένα για υστέρηση έως 24.

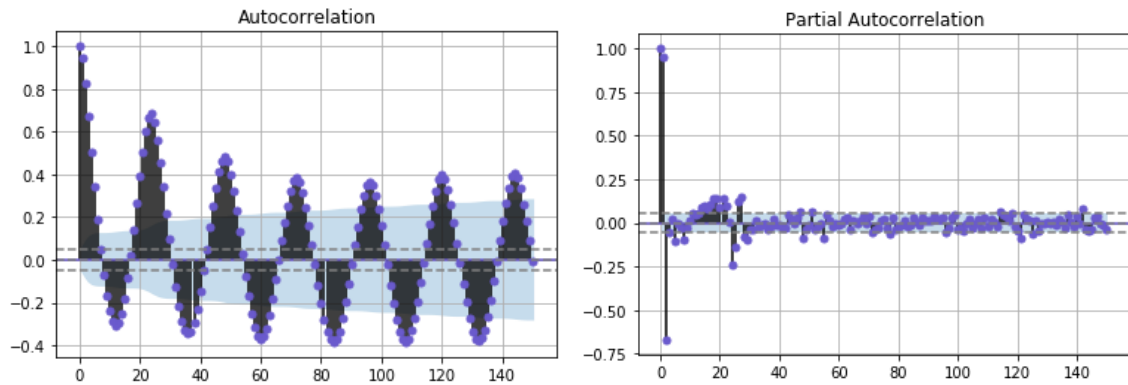


Σχήμα 4.3: Ανάλυση της χρονοσειράς στις συνιστώσες της τάσης και της εποχικότητας. Αφορά την φθινοπωρινή περίοδο 01/10/2011 - 33/11/2011



Σχήμα 4.4: Ανάλυση της χρονοσειράς στις συνιστώσες της τάσης και της εποχικότητας. Αφορά την φθινοπωρινή περίοδο 01/10/2011 - 07/11/2011

Η ύπαρξη στασιμότητας ελέγχθηκε και με τα στατιστικά τεστ, ενώ η κλάση ARIMA απο το πακέτο statsmodels της rython που θα χρησιμοποιηθεί για την υλοποίηση των στατιστικών μοντέλων πρόβλεψης περιέχει εσωτερικό έλεγχο για ύπαρξη ή μη στασιμότητας. Στον πίνακα 4.2 παρατίθενται τα αποτελέσματα του Augmented Dickey - Fuller test για τις 4 περιόδους (υπενθυμίζεται ότι αν $p\text{-value} < 0.05$, ή αλλιώς $T\text{-Statistics} < \text{Critical Value}$, απορρίπτεται η null hypothesis και άρα η χρονοσειρά είναι στάσιμη.) Τα παραπάνω μας οδηγούν στο



Σχήμα 4.5: Συνάρτηση αυτοσυσχέτισης και μερικής αυτοσυσχέτισης, συναρτήσει της υστέρησης, για την εαρινή περίοδο 01/07/2012- 30/08/2012

ADF test	Winter	Spring	Summer	Autumn
Test Statistics	-2.879862	-3.990236	-3.943855	-4.855272
p-value	0.047740	0.001463	0.001735	0.000043
#Lags Used	26	26	27	26
Critical Value (1%)	-3.433419	-3.433385	-3.433353	-3.433352
Critical Value (5%)	-2.862896	-2.862881	-2.862867	-2.862866
Critical Value (10%)	-2.567492	-2.567484	-2.567476	-2.567476

Πίνακας 4.2: Έλεγχος στασιμότητας δεδομένων με Augmented Dickey - Fuller test

συμπέρασμα να μην προβούμε σε διαφορίση της χρονοσειράς.

Σε αυτό το σημείο, πρέπει να επισημάνουμε ότι η υποβολή προσφορών στην Αγορά της Επόμενης Ημέρας κλείνει κατά τις μεσημεριανές ώρες (πχ σε πολλές αγορές ανέργειας στις 12 μμ). Επομένως, σε πραγματικές συνθήκες, έχει νόημα να έχει πραγματοποιηθεί μέχρι εκείνη την ώρα η πρόβλεψη για τις 24 οριακές τιμές ηλεκτρικής ενέργειας της επόμενης ημέρας, χωρίς να διατίθενται, μάλιστα, όλες οι πραγματικές τιμές της ίδιας ημέρας. Στην παρούσα διπλωματική εργασία, για λόγους απλούστευσης, κάνουμε την παραδοχή ότι για την πρόβλεψη της επόμενης ημέρας διαθέτουμε και τις 24 τιμές της τελευταίας ημέρας.

Τέλος, ως μέτρο επίδοσης επιλέχτηκε το μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE), εφόσον η φύση του προβλήματος είναι τέτοια που μας ενδιαφέρει το πόσο μακριά 'έπεσε' η πρόβλεψη συγκριτικά με την πραγματική τιμή.

4.2 Εφαρμογή στατιστικών μεθόδων

Για την υλοποίηση των μοντέλων ARIMA(p,d,q) κληθήκαμε να επιλέξουμε τις παραμέτρους p , d και q . Όπως αναφέρθηκε, η χρονοσειρά μας δεν διαφορίστηκε για καμία από τις περιόδους, επομένως έχουμε $d = 0$ (Μπορούμε, δηλαδή να αναφερόμαστε και σε ARMA(p,q)). Όσο για το p και το q , είδαμε ότι οι συναρτήσεις αυτοσυσχέτισης έφθιναν αργά με το χρόνο, ενώ της μερικής αυτοσυσχέτισης έφθιναν απότομα σύντομα, επομένως περισσότερη σημασία για το μοντέλο έχει ο παράγοντας p παρά ο q . Υλοποιήθηκαν, λοιπόν, πολλά μοντέλα, αφού για κάθε περίοδο υπό μελέτη επιλέχθηκε σταθερό $q = 0$ ή $q = 1$ και μεταβάλλαμε την παράμετρο p , στην οποία δώσαμε τις τιμές: $p = 1, 6, 12, 24$. Όπως αναφέρθηκε, για κάθε ένα τρίμηνο, οι τιμές των δύο πρώτων μηνών χρησιμοποιήθηκαν στην εκπαίδευση του μοντέλου, και για κάθε μια από τις ημέρες του τρίτου μήνα προβλέφθηκαν οι 24 οριακές τιμές ηλεκτρικής ενέργειας. Για κάθε μέρα του τελευταίου μήνα που περνάει, οι πραγματικές οριακές τιμές της ηλεκτρικής ενέργειας θεωρούνται γνωστές, οπότε προστίθενται στο σύνολο των δεδομένων που χρησιμεύουν στην πρόβλεψη της επόμενης ημέρας.

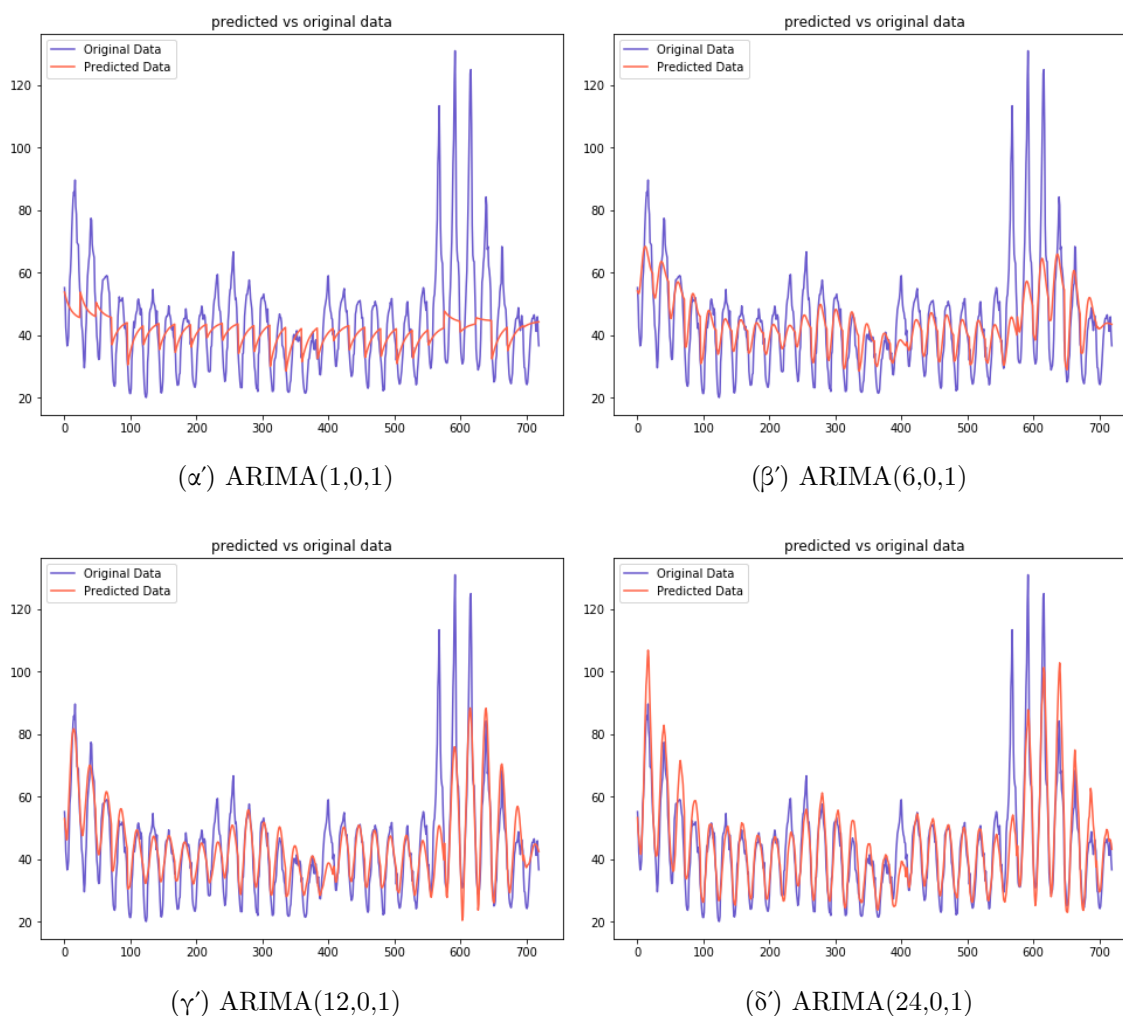
- **Υλοποίηση αλγορίθμου:** Μεταφράζοντας τα παραπάνω σε αλγόριθμο και περιγράφοντας συνοπτικά την υλοποίηση του συγκεκριμένου τμήματος, χωρίσαμε τα δεδομένα σε δύο τμήματα, *train* και *test*, με μέγεθος $2/3$ και $1/3$ επί του αρχικού. Εκτελείται μια επανάληψη 30 ή 31 φορές (όσες και οι μέρες υπό πρόβλεψη). Έστω d_t η μέρα υπό πρόβλεψη. Μέσα στην επανάληψη χρησιμοποιούμε την κλάση `statsmodels.tsa.arima_model.ARIMA` και τις συναρτήσεις της `fit` και `forecast`, για να προβλέψουμε, με βάση τα δεδομένα εκπαίδευσης, τις 24 οριακές τιμές της d_t . Στο τέλος προσθέτουμε στο σετ δεδομένων εκπαίδευσης τις πραγματικές τιμές της d_t , και προχωράμε στην επόμενη επανάληψη που δώσει την πρόβλεψη των 24 τιμών της d_{t+1} ημέρας. Σημειώνεται, τέλος, ότι τα δεδομένα εκπαίδευσης έχουν σταθερό μήκος, δηλαδή σαν ένα παράθυρο 2 μηνών που σύρεται: για νέα κάθε μέρα που προστίθεται, αφαιρείται η παλιότερη του σετ.

Θα παρουστούν οι τιμές ηλεκτρικής ενέργειας της πρόβλεψης, σε συγκριτικό διάγραμμα με τις πραγματικές τιμές ηλεκτρικής ενέργειας, για τις 30 τελευταίες μέρες κάθε περιόδου. Επιπλέον, υπολογίζεται και παρουσιάζεται το μέσο ανά ώρα σφάλμα MAPE που είχαν οι προβλέψεις των 30 αυτών ημερών.

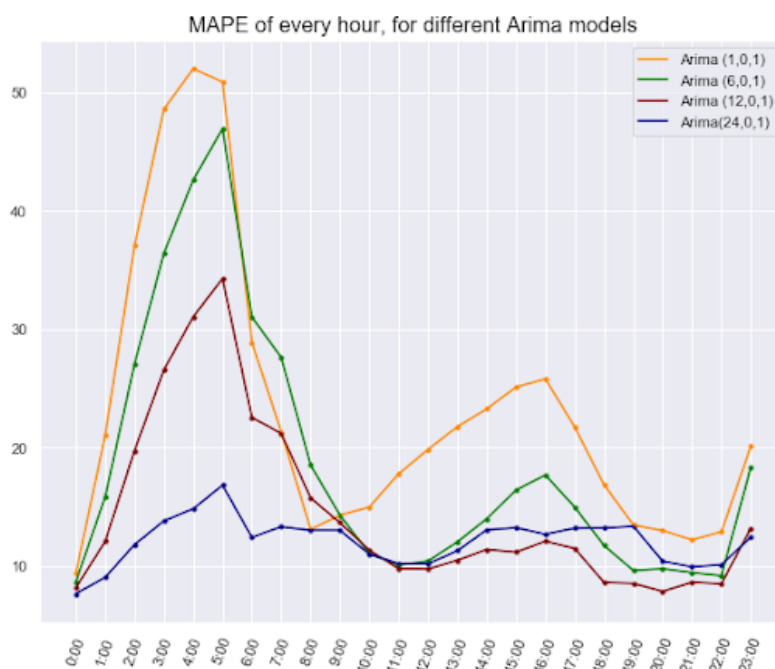
- **Προβλέψεις και σφάλματα**

Το σχήμα 4.6 αφορά την ανοιξιάτικη περίοδο 01/04/2013 - 30/06/2013, και παρατίθενται οι προβλέψεις για τα μοντέλα ARIMA(1,0,1), ARIMA(6,0,1), ARIMA(12,0,1), ARIMA(24,0,1), ώστε να οπτικοποιηθεί η επίδραση του παράγοντα p στην πρόβλεψη του μοντέλου. Παρατηρούμε τη βελτίωση στην πρόβλεψη όσο αυξάνεται το p , και αυτό μπορούμε να το δούμε και στο ενιαίο συγκριτικό διάγραμμα σφαλμάτων ανά ώρα, στο σχήμα 4.7. Συνολικά, το βέλτιστο μοντέλο για την ανοιξιάτικη περίοδο, που θα διατηρήσουμε και ως benchmark, είχε αρκετά καλή επίδοση, με ολικό MAPE = 12.10590418625. Στο σχήμα 4.8 φαίνεται η

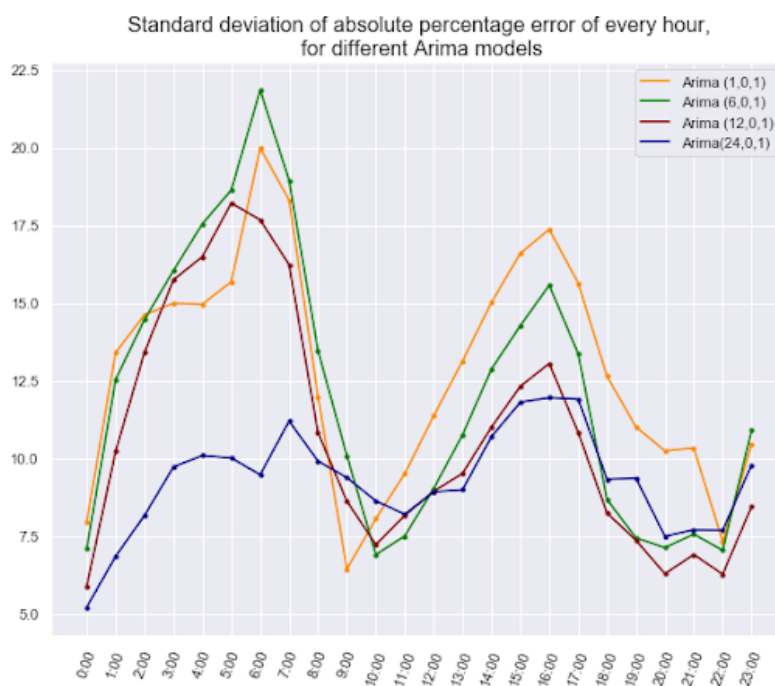
τυπική απόλιση του σφάλματος ανά ώρα για το διάστημα πρόβλεψης. Για τις υπόλοιπες περιόδους θα παρουσιάσουμε μόνο το βέλτιστο μοντέλο ARIMA, το οποίο θα χρησιμοποιηθεί ως benchmark για να συγκρίνουμε στην πορεία με τα νευρωνικά δίκτυα, και το συγκριτικό διάγραμμα του MAPE ανά ώρα για τα διάφορα μοντέλα καθώς μεταβάλαμε την τιμή του p .



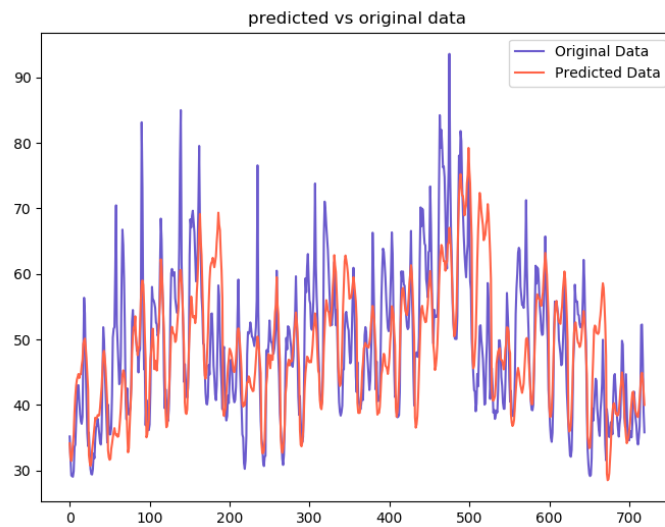
Σχήμα 4.6: Άνοιξη - Διαγράμματα προβλέψεων με μοντέλα ARIMA για το διάστημα 01/06/2013- 30/06/2013



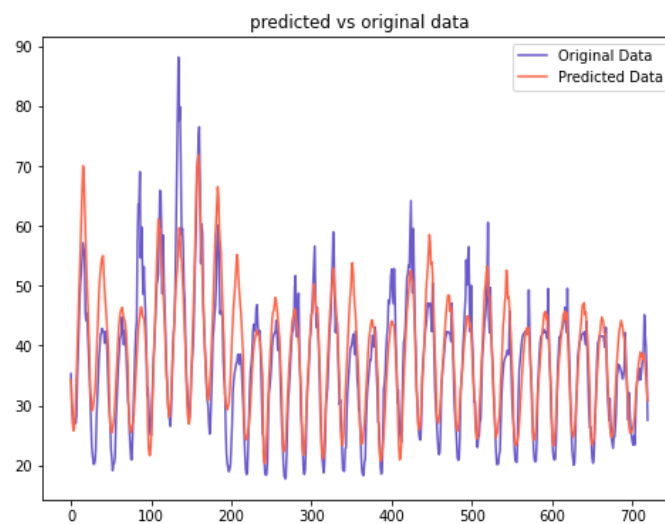
Σχήμα 4.7: Μέσο Απόλυτο Ποσοστιαίο Σφάλμα ανά ώρα, για τις προβλέψεις του διαστήματος 01/06/2013- 30/06/2013



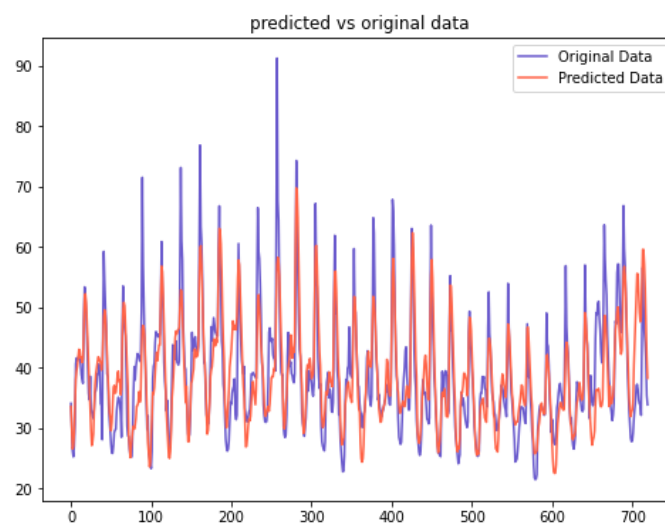
Σχήμα 4.8: Τυπική απόκλιση του απόλυτου ποσοστιαίου σφάλματος ανά ώρα, για τις προβλέψεις του διαστήματος 01/06/2013- 30/06/2013



(α') Χειμώνας ARIMA(24,0,1)

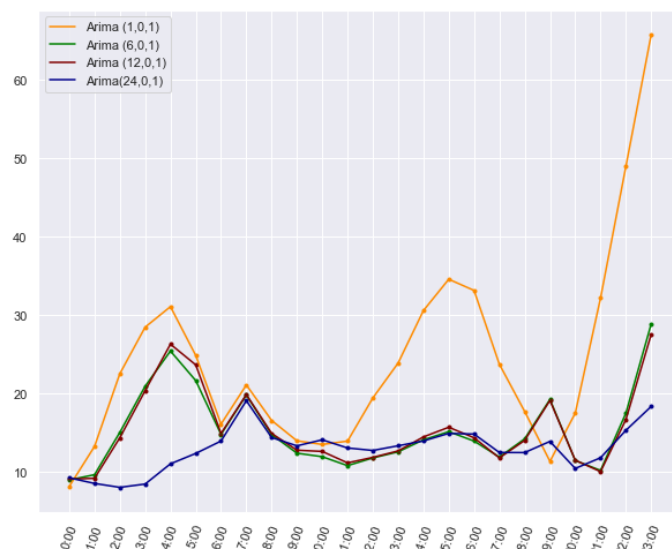


(β') Καλοκαίρι ARIMA(24,0,0)

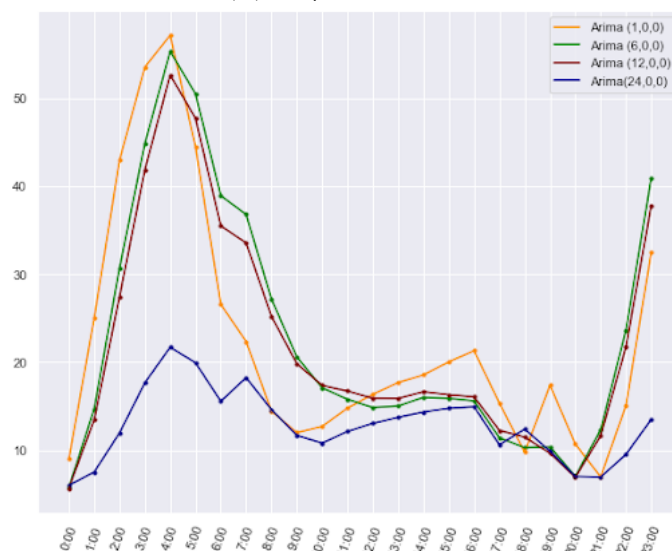


(γ') Φθινόπωρο ARIMA(24,0,0)

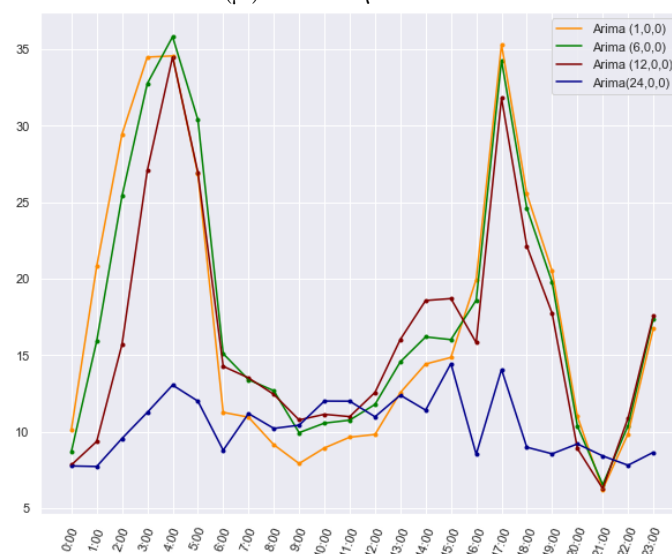
Σχήμα 4.9: Χειμώνας, Καλοκαίρι, Φθινόπωρο: διαγράμματα προβλέψεων με μοντέλα ARIMA, $p = 24$



(α') Χειμώνας MAPE



(β') Καλοκαίρι MAPE



(γ') Φθινόπωρο MAPE

Σχήμα 4.10: Χειμώνας, Καλοκαίρι, Φθινόπωρο: διαγράμματα σφαλμάτων σε μέσο όρο ανά ώρα, με μοντέλα ARIMA

4.3 Εφαρμογή νευρωνικών δικτύων

Για την πρόβλεψη με νευρωνικά δίκτυα συστήνεται κανονικοποίηση των δεδομένων, την οποία και υλοποιήσαμε. Μετά την πρόβλεψη εφαρμόστηκε αντίστροφος μετασχηματισμός, οπότε στα διαγράμματα που ακολουθούν φαίνονται οι προβλέψεις των τιμών σε πραγματική κλίμακα. Όσον αφορά τον αλγόριθμο που υλοποιήθηκε, για κάθε περίοδο τα δεδομένα εκπαίδευσης ήταν, όπως έχει αναφερθεί, οι τιμές των δύο πρώτων μηνών, και τα δεδομένα ελέγχου οι τιμές του τρίτου μήνα.

- **Υλοποίηση αλγορίθμου:** Συγκεκριμένα, δημιουργήθηκαν ζεύγη δεδομένων εισόδου - εξόδου, με γνώμονα ότι οι 24 τιμές ηλεκτρικής ενέργειας μιας μέρας δίνουν επαρκή πληροφορία ώστε να προβλεφθεί η ωριακή τιμή για κάθε ώρα της επόμενης ημέρας. Έτσι, μια είσοδος έχει τη μορφή ακολουθίας 24 ωριακών τιμών και η αναμενόμενη έξοδος είναι η ακολουθία των 24 επόμενων τιμών. Το σύνολο των δεδομένων εκπαίδευσης χωρίστηκε, λοιπόν, σε ζεύγη ακολουθιών εισόδου-εξόδου. Επισημαίνεται ότι τα δεδομένα εκπαίδευσης χωρίζονται σε training και validation σετ με τυχαίο τρόπο, σε ποσοστό 80%-20%. Το σύνολο των δεδομένων ελέγχου, δηλαδή του τρίτου μήνα, σχηματίζεται σε ακολουθίες 24 τιμών, που αντιστοιχούν στις ωριαίες τιμές της ηλεκτρικής ενέργειας σε μία μέρα, που θα αποτελέσουν είσοδο στο εκπαιδευμένο μοντέλο. Το πλήθος αυτών των ακολουθιών είναι φυσικά ίδιο με το πλήθος των ημερών του μήνα. Το σφάλμα που υπολογίζουμε, και σε αυτή την περίπτωση, είναι το MAPE, και μάλιστα η συνάρτηση κόστους (loss function), που το νευρωνικό καλείται να ελαχιστοποιήσει, μέσω της εκπαίδευσης, σε όλες τις αρχιτεκτονικές που υλοποιήσαμε ήταν η mape.

Κάθε περίοδος προς πρόβλεψη αντιμετωπίστηκε ανεξάρτητα, αφού οι χρονοσειρές είχαν διαφορετικά χαρακτηριστικά, και έτσι στις περισσότερες περιπτώσεις μεταβάλλαμε την αρχιτεκτονική και τις υπερπαραμέτρους. Σε κάθε περίπτωση, σχεδιάζονται οι προβλέψεις του νευρωνικού δικτύου σε σύγκριση με τις αναμενόμενες τιμές (δηλαδή τις πραγματικές ωριαίες τιμές ενέργειας), και η επίδοση του μοντέλου συγκρίνεται με αυτή του ARIMA, μέσα από κοινό διάγραμμα MAPE ανά ώρα.

Για την κανονικοποίηση των δεδομένων και τον χωρισμό σε training - validation dataset χρησιμοποιήσαμε συναρτήσεις της βιβλιοθήκης sklearn, ενώ για την υλοποίηση των νευρωνικών δικτύων χρησιμοποιήθηκε η βιβλιοθήκη keras.

4.3.1 LSTM

Η βασική αρχιτεκτονική που υλοποιήθηκε ήταν ένα Bidirectional LSTM στρώμα 100 νευρώνων, ακολουθούμενο από ένα Dropout Layer με πιθανότητα 0.2, και τέλος το στρώμα εξόδου είναι ένα fully connected 24 νευρώνων, όσες και οι επιθυμητές έξοδοι. (βλ. Σχήμα 4.11). Οι υπερπαραμέτροι που επιλέχθηκαν φαίνονται στον πίνακα 4.3.

```

Model: "sequential_20"
-----
Layer (type)                Output Shape          Param #
-----
bidirectional_15 (Bidirectio (None, 200)          81600
-----
dropout_20 (Dropout)        (None, 200)           0
-----
dense_20 (Dense)            (None, 24)            4824
-----
Total params: 86,424
Trainable params: 86,424
Non-trainable params: 0

```

Σχήμα 4.11: αρχιτεκτονική LSTM νευρωνικού δικτύου

	Autumn	Summer	Winter	Spring
Epochs	50	30	50	30
Activation Function	<i>Tanh</i>	ReLU	<i>Tanh</i>	<i>Tanh</i>
Batch size	16			
Optimizer	Adam			
Loss function	mape			

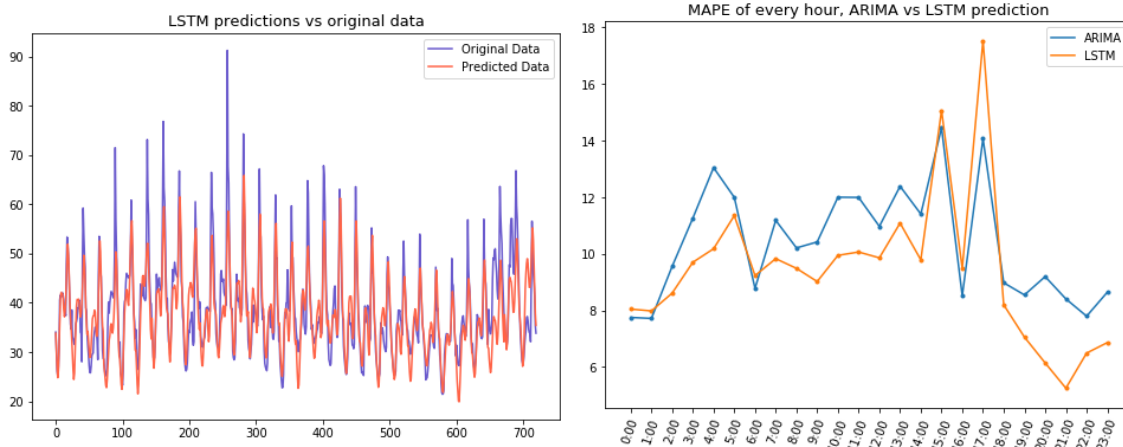
Πίνακας 4.3: Υπερπαραμέτροι του LSTM μοντέλου

- Προβλέψεις και σφάλματα

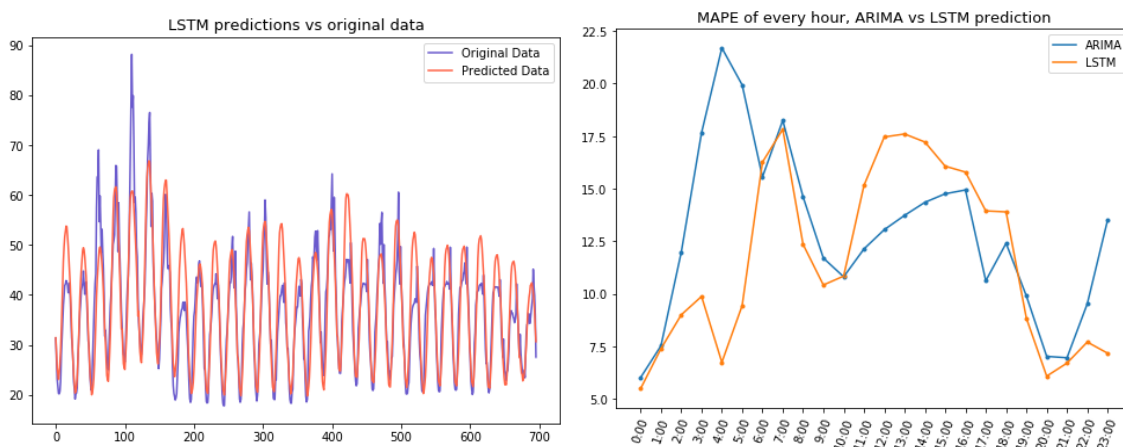
Ακολουθούν οι γραφικές παραστάσεις των προβλέψεων και των σφαλμάτων των μοντέλων ανά ώρα. Για την εαρινή περίοδο (πρόβλεψη του Ιουνίου 2013 με δεδομένα εκπαίδευσης από τον Απρίλιο και Μάιο 2013), η επίδοση του μοντέλου δεν είχε τα επιθυμητά αποτελέσματα. Η προσαρμογή των υπερπαραμέτρων στις ανάγκες του προβλήματος δεν βελτίωσε τα αποτελέσματα, καθώς ο μεγαλύτερος πααράγοντας που συνισφέρει στην βελτίωση ενός νευρωνικού μοντέλου είναι η επάρκεια των δεδομένων εκπαίδευσης. Έτσι, προσθέσαμε στα δεδομένα εκπαίδευσης παραπάνω παρελθοντικές ακολουθίες της χρονοσειράς, που προέρχονταν από τον Ιούνιο 2012 και Ιούνιο 2011. Τα αποτελέσματα για τις δύο περιπτώσεις φαίνονται στα Σχήματα 4.15 & 4.16. Συνολικά το μέσο απόλυτο ποσοστιαίο σφάλμα που παρήγαγε το LSTM για κάθε περίοδο πρόβλεψης, σε σύγκριση με το benchmark φαίνονται στον Πίνακα 4.4.

Model \ Period	ARIMA	LSTM
autumn	10.3814232	9.4238780
summer	12.8588003	11.6351827
winter	12.9432410	11.5993538
spring	12.1059042	10.5890539

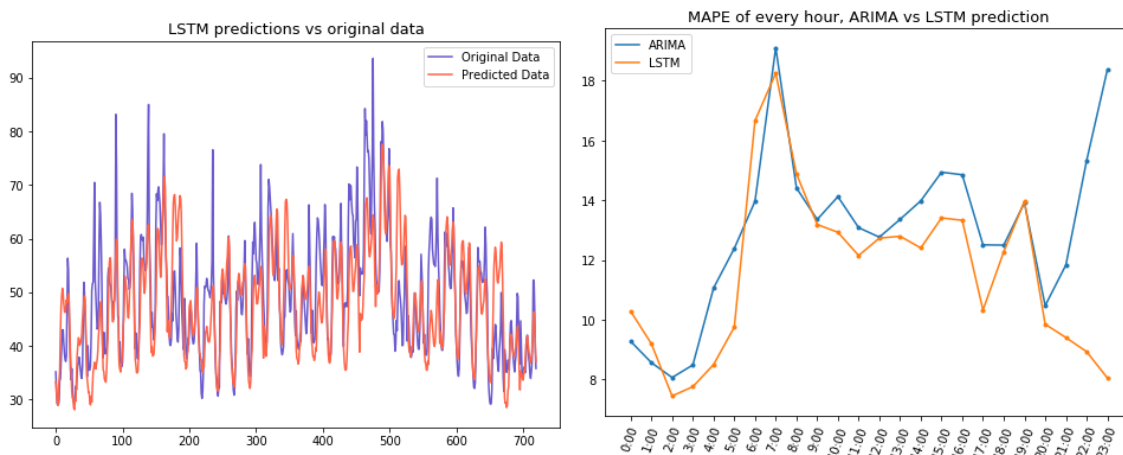
Πίνακας 4.4: Σύγκριση του MAPE του LSTM σε σχέση με το στατιστικό μοντελο



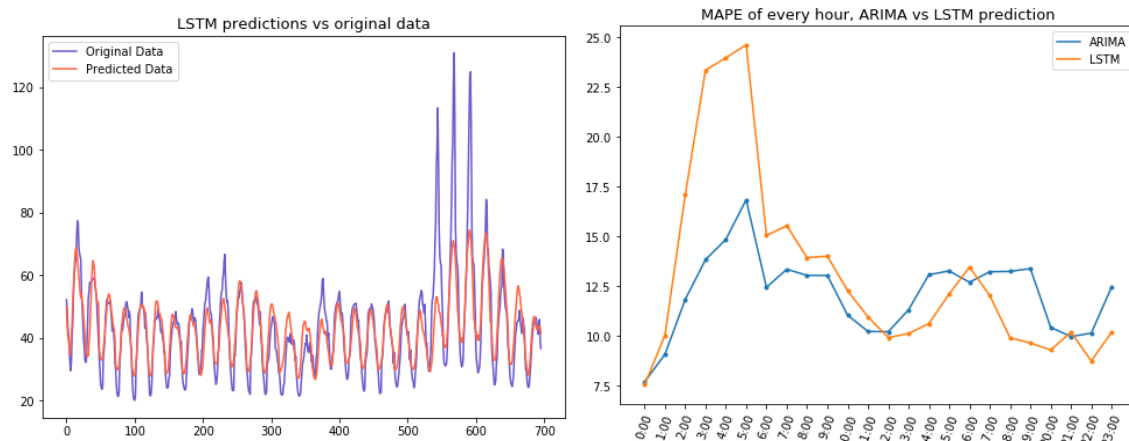
Σχήμα 4.12: Φθινόπωρο - Πρόβλεψη Δεκ 2011 με LSTM



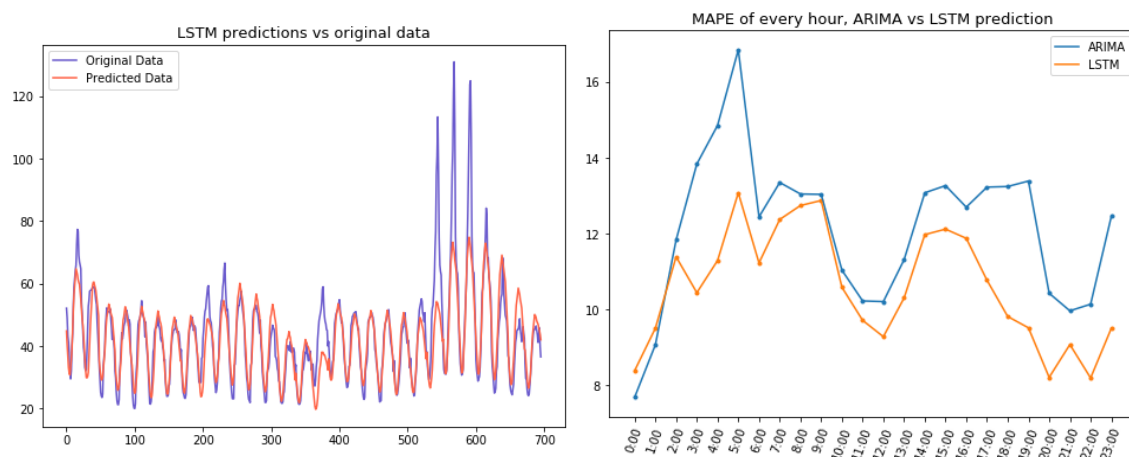
Σχήμα 4.13: Καλοκαίρι - Πρόβλεψη Σεπ 2012 με LSTM



Σχήμα 4.14: Χειμώνας - Πρόβλεψη Μαρ 2013 με LSTM



Σχήμα 4.15: Άνοιξη - Πρόβλεψη Ιουν 2013 με LSTM



Σχήμα 4.16: Άνοιξη - Πρόβλεψη Ιουν 2013 με LSTM και προσθήκη Ιουνίου 2011&2012 στην εκπαίδευση

4.3.2 CNN

Η αρχιτεκτονική που υλοποιήθηκε περιλαμβάνει: (βλ. Σχήμα 4.17)

- ένα μονοδιάστατο convolutional στρώμα 64 φίλτρων, με kernel size, δηλαδή μέγεθος φίλτρου 2,
- ένα MaxPooling στρώμα με μέγεθος παραθύρου 2
- ακολούθως, ένα Dropout layer με πιθανότητα για dropout 0.2
- ένα flatten στρώμα
- και το επίπεδο εξόδου είναι ένα dense, δηλαδή fully connected με 24 νευρώνες, όσες και οι ζητούμενες έξοδοι.

Οι υπερπαραμετροί που επιλέχθηκαν για κάθε περίοδο πρόβλεψης φαίνονται στον Πίνακα 4.5.

```

Model: "sequential_32"
-----
Layer (type)                Output Shape                Param #
-----
conv1d_25 (Conv1D)          (None, 23, 64)             192
max_pooling1d_25 (MaxPooling (None, 11, 64)             0
dropout_28 (Dropout)        (None, 11, 64)             0
flatten_25 (Flatten)        (None, 704)                 0
dense_42 (Dense)            (None, 24)                  16920
-----
Total params: 17,112
Trainable params: 17,112
Non-trainable params: 0

```

Σχήμα 4.17: αρχιτεκτονική CNN νευρωνικού δικτύου

	Autumn	Summer	Winter	Spring
Epochs	30	30	35	30
Activation Function	None	None	ReLU	None
Batch size	16			
Optimizer	Adam			
Loss function	mape			

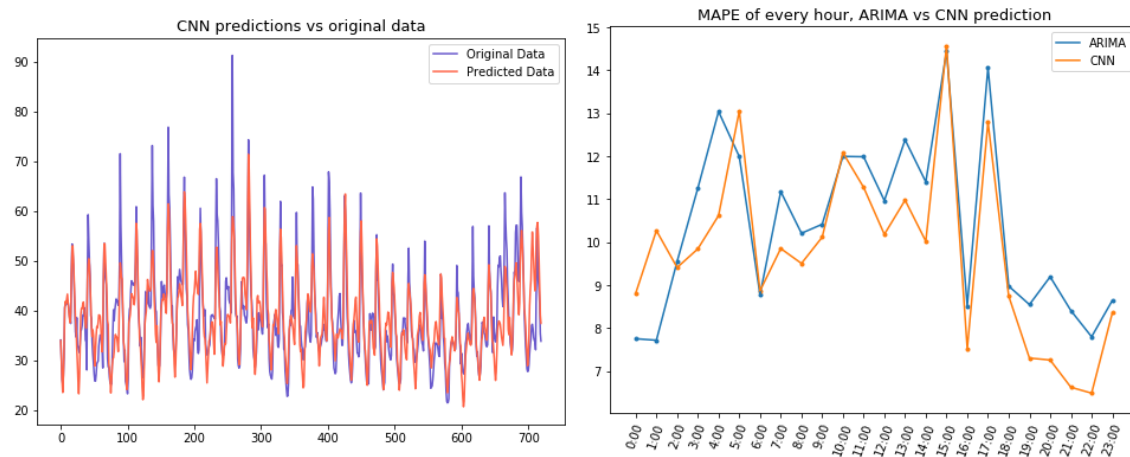
Πίνακας 4.5: Υπερπαράμετροι του CNN μοντέλου

- Προβλέψεις και σφάλματα

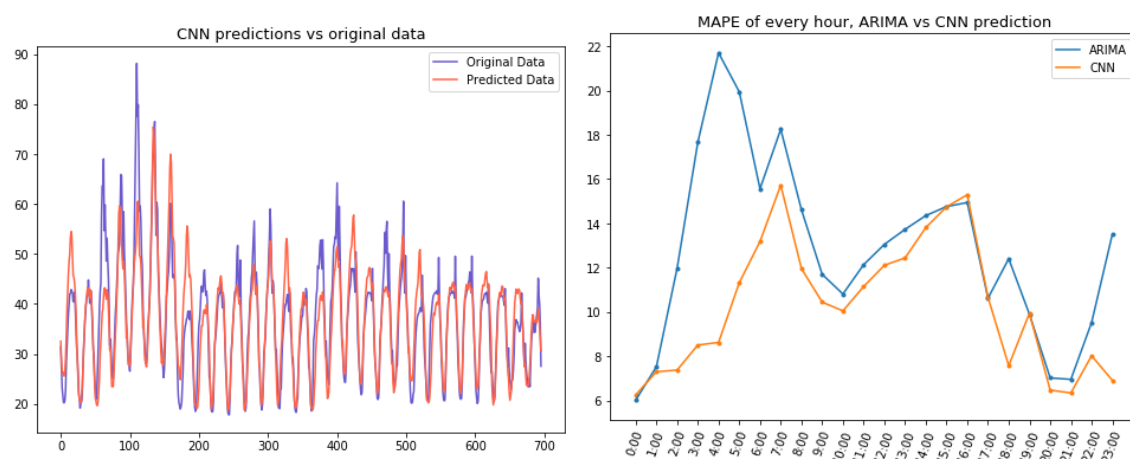
Το συνελικτικό νευρωνικό δίκτυο που υλοποιήθηκε αποδείχτηκε πολύ αποτελεσματικό και γρήγορο στην εκπαίδευση. Συνολικά το μέσο απόλυτο ποσοστιαίο σφάλμα που παρήγαγε το CNN για κάθε περίοδο πρόβλεψης, σε σύγκριση με το benchmark φαίνονται στον Πίνακα 4.6.

Model \ Period	ARIMA	CNN
autumn	10.3814232	9.7694650
summer	12.8588003	10.2568160
winter	12.9432410	11.1548018
spring	12.1059042	10.9365406

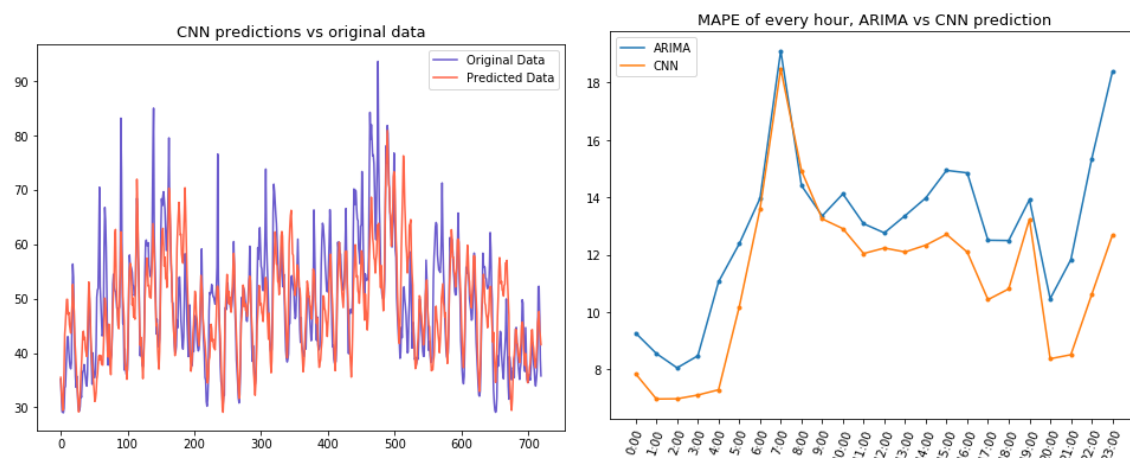
Πίνακας 4.6: Σύγκριση του MAPE του CNN σε σχέση με το στατιστικό μοντέλο



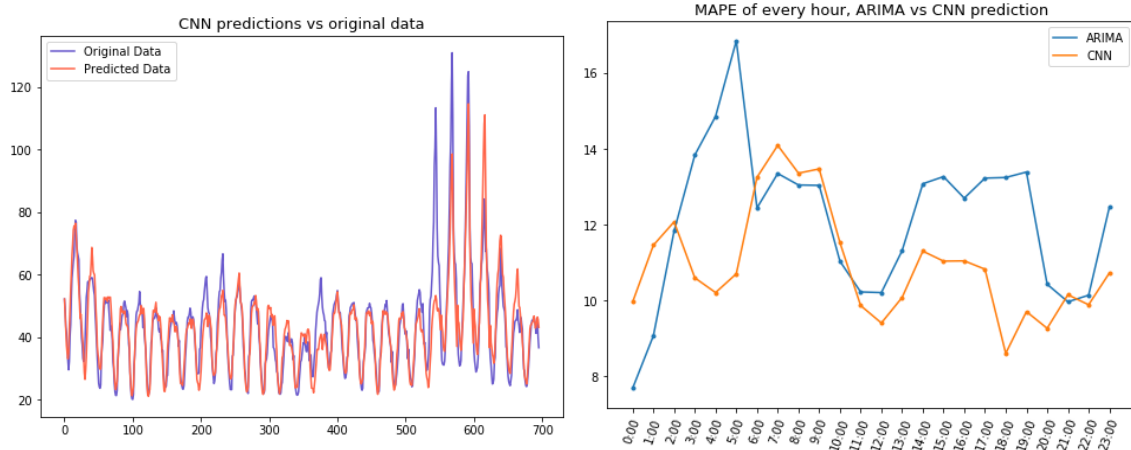
Σχήμα 4.18: Φθινόπωρο - Πρόβλεψη Δεκ 2011 με CNN



Σχήμα 4.19: Καλοκαίρι - Πρόβλεψη Σεπ 2012 με CNN



Σχήμα 4.20: Χειμώνας - Πρόβλεψη Μαρ 2013 με CNN



Σχήμα 4.21: Άνοιξη - Πρόβλεψη Ιουν 2013 με CNN

4.3.3 CNN & LSTM

Εκτός από τις παραπάνω αρχιτεκτονικές, δοκιμάστηκε και ο συνδυασμός των δύο. Συγκεκριμένα, η έξοδος του MaxPooling layer του CNN αποτέλεσε είσοδο για ένα bidirectional LSTM δίκτυο 100 νευρώνων. Ακολούθησε ένα πλήρως συνδεδεμένο επίπεδο 24 νευρώνων που ήταν και το επίπεδο εξόδου. Την ακριβή αρχιτεκτονική του νευρωνικού δικτύου βλέπουμε στο Σχήμα 4.22.

Οι υπερπαραμετροί που επιλέχθηκαν για κάθε περίοδο πρόβλεψης φαίνονται στον Πίνακα 4.7.

```
Model: "sequential_15"
```

Layer (type)	Output Shape	Param #
conv1d_15 (Conv1D)	(None, 23, 64)	192
max_pooling1d_15 (MaxPooling)	(None, 11, 64)	0
dropout_15 (Dropout)	(None, 11, 64)	0
bidirectional_7 (Bidirection)	(None, 200)	132000
dense_11 (Dense)	(None, 24)	4824

Total params: 137,016
Trainable params: 137,016
Non-trainable params: 0

Σχήμα 4.22: αρχιτεκτονική νευρωνικού δικτύου

- Προβλέψεις και σφάλματα

Η επίδοση του νευρωνικού δικτύου φαίνεται στις γραφικές παραστάσεις των προβλέψεων και των σφαλμάτων που ακολουθούν. Αξίζει να σημειωθεί ότι, όπως και στην περίπτωση της LSTM αρχιτεκτονικής, στην εαρινή περίοδο το μοντέλο απέτυχε να δώσει καλύτερες συνολικά προβλέψεις από το benchmark. Συγκεκριμένα, το συνολικό MAPE της της πρόβλεψης που φαίνεται στο Σχήμα 4.26 είναι 13.091996546%, έναντι του 12.105904187% που πέτυχε το

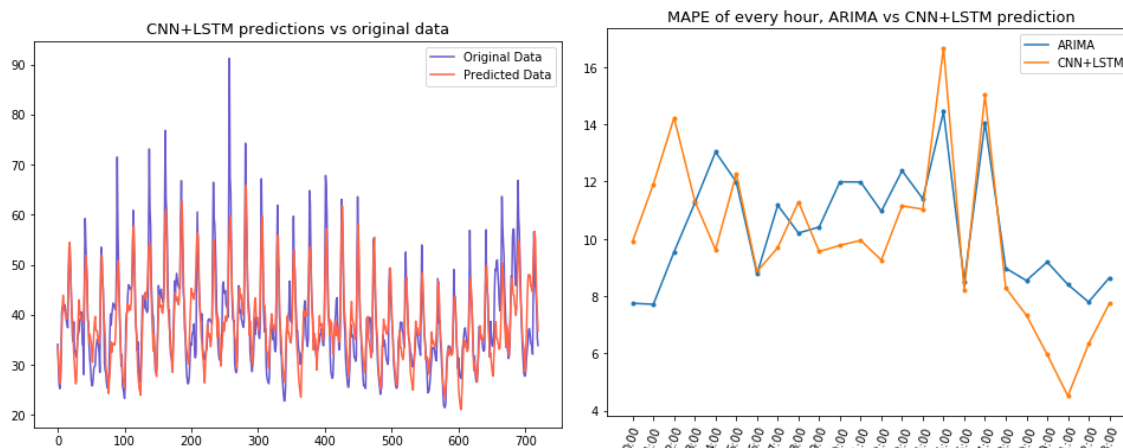
	Autumn	Summer	Winter	Spring
Epochs	40	40	40	35
Activation Function Conv 1D	None			
Activation Function LSTM	<i>tanh</i>			
Batch size	16			
Optimizer	Adam			
Loss function	mape			

Πίνακας 4.7: Υπερπαράμετροι του CNN & LSTM μοντέλου

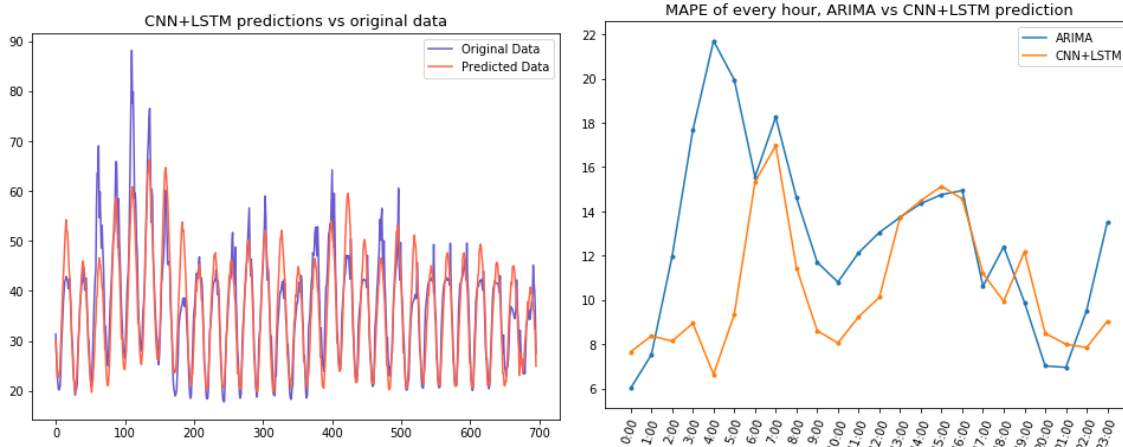
στατιστικό μοντέλο μας. Προσθέτοντας, όμως, στην εκπαίδευση, και τα δεδομένα του Ιουνίου 2012 και Ιουνίου 2011 επιτεύχθη MAPE 10.336497246%. Στον Πίνακα 4.8, που γίνεται η σύγκριση της επίδοσης του νευρωνικού μοντέλου αναφορικά με το στατιστικό, αναφέρονται τα σφάλματα από τα βέλτιστα μοντέλα νευρωνικών δικτύων (με την προσθήκη, δηλαδή, των επιπλέον δεδομένων τιμών ενέργειας).

Model \ Period	ARIMA	CNN & LSTM
autumn	10.3814232	9.9975536
summer	12.8588003	10.5617418
winter	12.9432410	11.4704586
spring	12.1059042	10.3364972

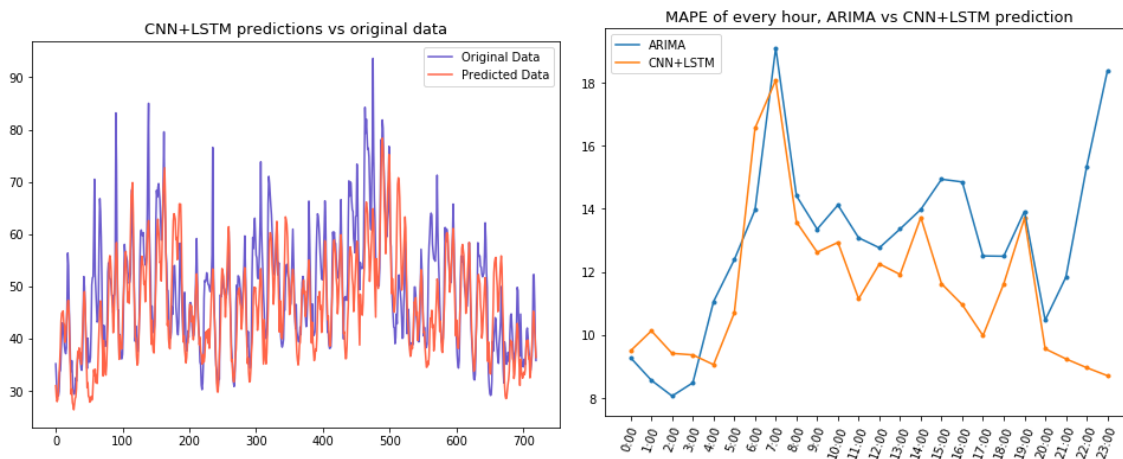
Πίνακας 4.8: Σύγκριση του MAPE του νευρωνικού δικτύου με CNN & LSTM αρχιτεκτονική σε σχέση με το στατιστικό μοντέλο



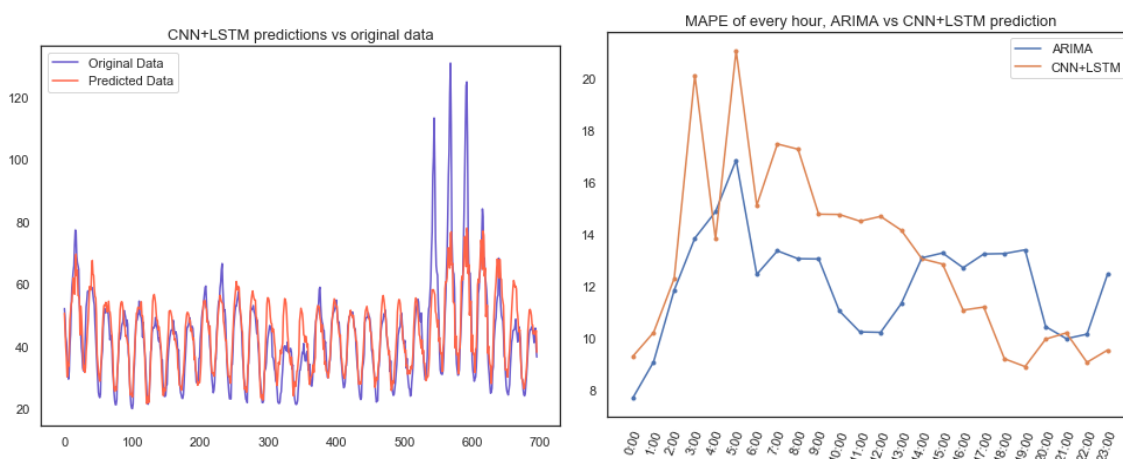
Σχήμα 4.23: Φθινόπωρο - Πρόβλεψη Δεχ 2011 με CNN & LSTM



Σχήμα 4.24: Καλοκαίρι - Πρόβλεψη Σεπ 2012 με CNN & LSTM



Σχήμα 4.25: Χειμώνας - Πρόβλεψη Μαρ 2013 με CNN & LSTM



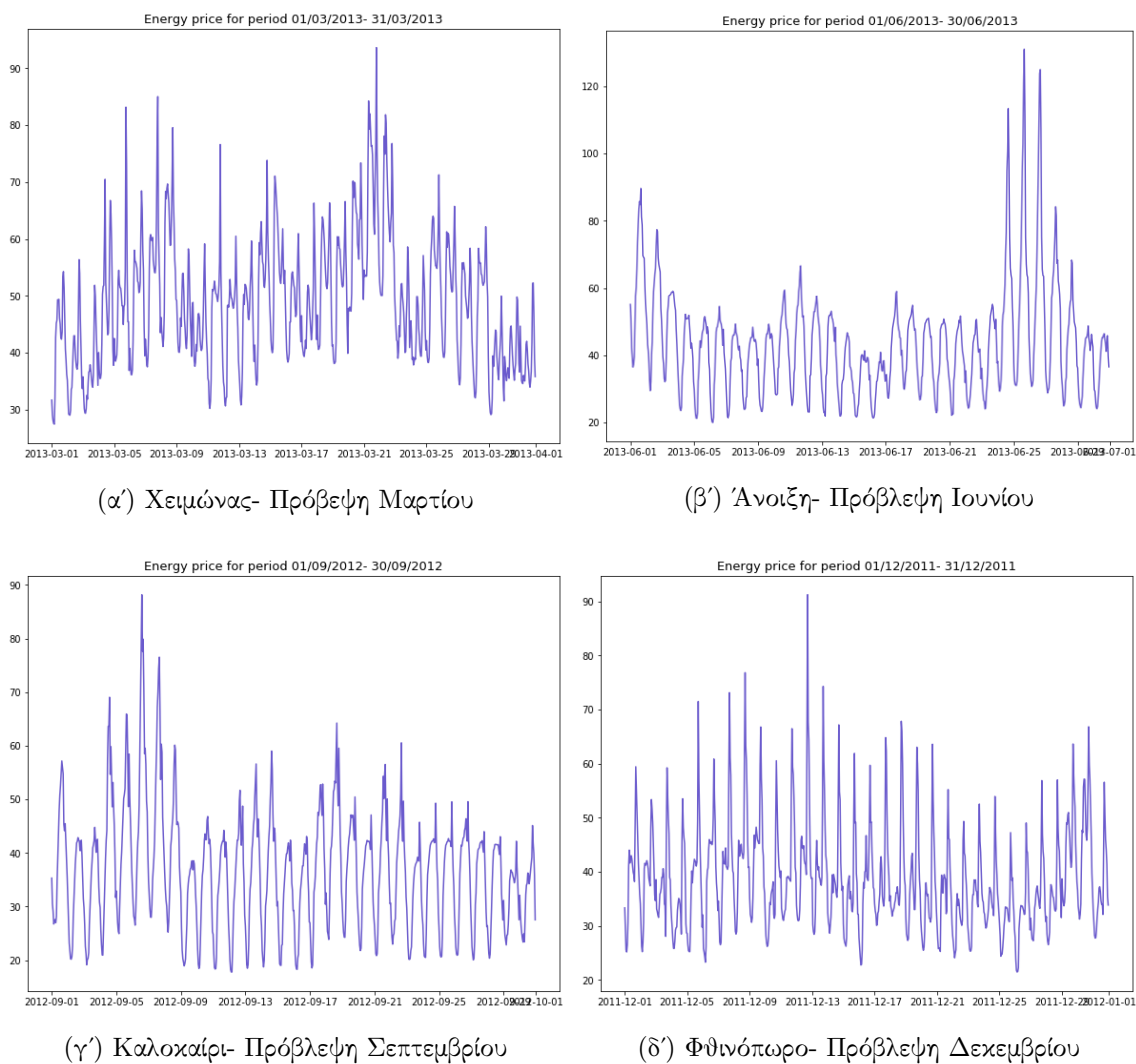
Σχήμα 4.26: Άνοιξη - Πρόβλεψη Ιουν 2013 με CNN & LSTM

4.4 Σχολιασμός και σύγκριση μεθόδων πρόβλεψης

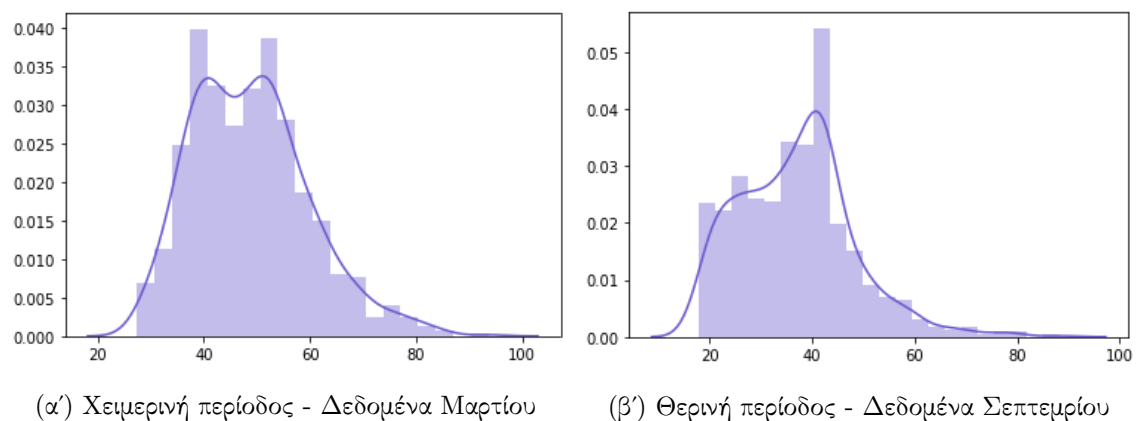
4.4.1 Σχολιασμός μεθόδων ARIMA και σύγκριση των διαφορετικών μοντέλων

Από το σχήμα 4.6 είναι φανερό ότι σε ένα μοντέλο ARMA(p,q) όσο αυξάνουμε την μεταβλητή p , δηλαδή όσο το μοντέλο λαμβάνει υπόψη περισσότερες τιμές υστέρησης για την πρόβλεψή του, τόσο γίνεται πιο εύστοχο. Βέβαια, όσο μεγαλύτερη γίνεται η υστέρηση τόσο αυξάνονται και οι παράμετροι που υπολογίζει το μοντέλο, και αυτό το καθιστά πιο πολύπλοκο. Υπάρχει δηλαδή trade off μεταξύ της επίδοσης του μοντέλου και της πολυπλοκότητάς του. Περαιτέρω αύξηση του p , καθιστά το μοντέλο ασύμφορα πολύπλοκο, καθώς δεν μπορεί να υπάρξει ακραία βελτίωση στην επίδοσή του, εφόσον είδαμε ότι οι 24 τελευταίες τιμές είναι αυτές που παρουσίαζαν ισχυρή αυτοσυσχέτιση.

Μάλιστα, από το σχήμα 4.10 παρατηρούμε ότι είναι διαφορετική η επίδραση της μεταβολής του p στην ευστοχία των προβλέψεων, άρα και στη μείωση του σφάλματος, για τις διαφορετικές εποχές του χρόνου. Συγκεκριμένα, στο χειμώνα η αύξηση του p δεν επέφερε καθοριστική βελτίωση στις προβλέψεις. Εκτός από το ARIMA(1,0,1), που στην ουσία λαμβάνει υπόψη την προηγούμενη τιμή της χρονοσειράς, άρα την τιμή της ηλεκτρικής ενέργειας την προηγούμενη ώρα μόνο, και, όπως είναι αναμενόμενο, παρουσίασε υψηλά σφάλματα, η υστέρηση 6, 12 ή 24 ωρών δεν μετέβαλε ουσιαστικά το σφάλμα για το μεγαλύτερο εύρος των ωρών μέσα στη μέρα. Αντίθετα, στο φθινόπωρο και το καλοκαίρι οι προβλέψεις που χρησιμοποιούν υστέρηση 1,6, και 12 τιμών παρουσιάζουν αντίστοιχα σφάλματα, τα οποία βελτιώνονται εντυπωσιακά με την εισαγωγή των 24 προηγούμενων ωρών στο μοντέλο. Ας θυμηθούμε τη μορφή των δεδομένων που τα μοντέλα κλήθηκαν να προβλέψουν, για να καταλάβουμε την αιτία της παραπάνω παρατήρησης (Σχήμα 4.27). Παρατηρούμε ότι κατά τον μήνα Μάρτιο υπάρχουν πολύ έντονες αυξομειώσεις στο πλάτος της τιμής. Η συχνότητα αυτών μπορεί να είναι μεν περιοδική, αλλά η μορφή τους δεν είναι απόλυτα όμοια, μιας και το πλάτος κυμαίνεται από 27.45 (που είναι η ελάχιστη τιμή) έως 93.62 (που είναι η μέγιστη τιμή) καθόλη τη διάρκεια του μήνα (για κατανομή βλ. Σχήμα 4.28 (α')). Έτσι, η γνώση των 24 προηγούμενων τιμών δεν βοηθάει πολύ σε σχέση με αυτή των 6 ή 12, εφόσον το πλάτος της τιμής είναι περισσότερο απρόβλεπτο. Από τη συμπεριφορά της χρονοσειράς αναμένουμε και μεγάλα σφάλματα πρόβλεψης. Αυτό έρχεται σε αντίθεση με το καλοκαίρι, που παρατηρούμε ότι η τιμή της ενέργειας μεταβάλλεται περιοδικά (με συχνότητα 24) ακολουθώντας όμοιο μοτίβο κίνησης και μάλιστα σε ένα εύρος τιμών κατά κύριο λόγο μεταξύ 20 και 50 (βλ. Σχήμα 4.28 (β')). Είναι φανερό ότι, σε αυτή την περίπτωση, η γνώση των 6 και των 12 προηγούμενων τιμών της χρονοσειράς έχει αντίστοιχο αντίκτυπο, ενώ συνυπολογίζοντας τις 24 προηγούμενες τιμές μπορούμε να πετύχουμε πολύ μεγαλύτερη ακρίβεια στην πρόβλεψη.

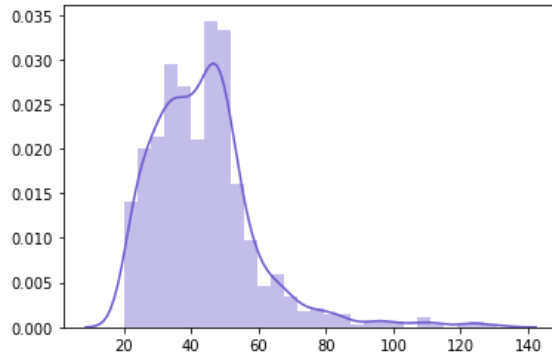


Σχήμα 4.27: Χρονοσειρά δεδομένων τιμές ηλεκτρικής ενέργειας για τους μήνες που υλοποιείται η πρόβλεψη.



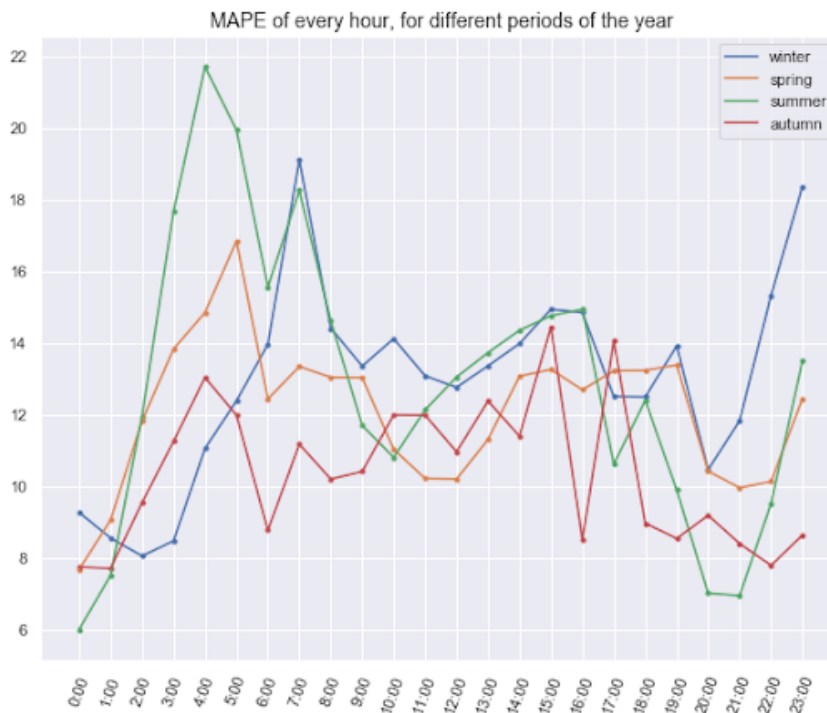
Σχήμα 4.28: Κατανομή της οριακής τιμής ηλεκτρικής ενέργειας για τους μήνες Μάρτιο 2013 και Σεπτέμβριο 2012.

Από την άλλη πλευρά, η τιμή κατά το μήνα Ιούνιο (βλ. Σχήμα 4.27 (β΄)) εξελίσσεται με αρκετά επαναλαμβανόμενο μοτίβο, και εύρος τιμών κυρίως μεταξύ 25-55, έτσι όμως, το μεγάλο πλάτος στην τιμή της ηλεκτρικής ενέργειας αποτελεί ακραίες τιμές, και για αυτό το λόγο το μοντέλο αδυνατεί να κάνει αποδοτική πρόβλεψη. Η ελάχιστη τιμή είναι για αυτό το μήνα 20.0 και η μέγιστη 131.0, και η κατανομή φαίνεται στο παρακάτω σχήμα.



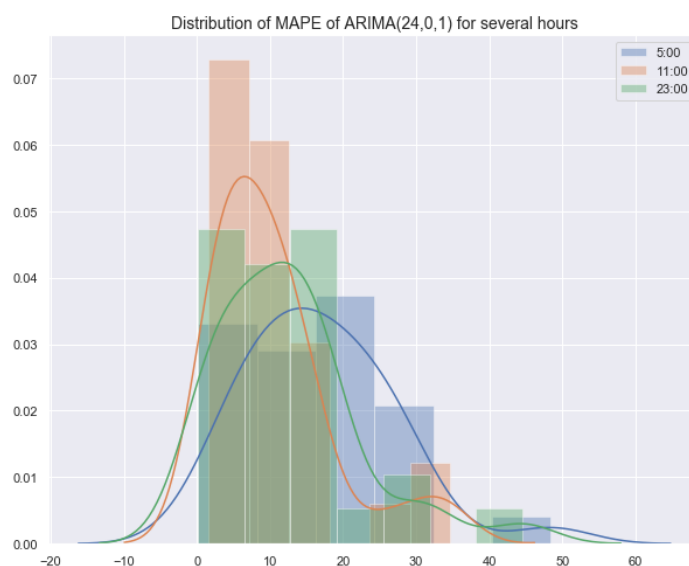
Σχήμα 4.29: Κατανομή της οριακής τιμής ηλεκτρικής ενέργειας για το μήνα Ιούνιο 2013 - Εαρινή περίοδος

Μπορούμε να δούμε συγκριτικά την επίδοση των βέλτιστων μοντέλων ARIMA για τις διαφορετικές περιόδους πρόβλεψης. Από το διάγραμμα διακρίνουμε μικρότερα σφάλματα για το φθινόπωρο και την άνοιξη, και μεγαλύτερα για το χειμώνα και το καλοκαίρι, όπως αναμέναμε από τις παραπάνω παρατηρήσεις.

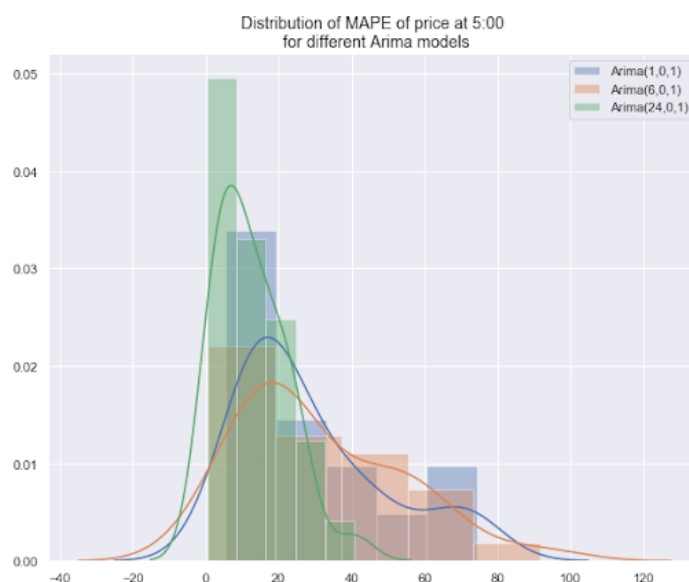


Σχήμα 4.30: Επίδοση μοντέλων ARIMA για τις διαφορετικές περιόδους πρόβλεψης

Εκτός αυτού, παρατηρούμε ότι, ανεξάρτητα από την εποχή, κατά μέσο όρο, κατά τις μεσημεριανές ώρες οι προβλέψεις ήταν πιο ακριβείς, ενώ για τις πρώτες πρωινές ώρες (2:00-7:00) και τις τελευταίες βραδινές (22:00-23:00) το σφάλμα αυξανόταν. Η παρατήρηση ισχύει και στην εφαρμογή μοντέλων χαμηλότερης πολυπλοκότητας και συνολικής ακρίβειας. Μπορούμε να δούμε και πώς μεταβάλλεται η κατανομή του σφάλματος για τρεις ώρες της ημέρας, επιλέξαμε τις 5.00, 11.00 και 23.00, έστω για τα σφάλματα της ανοιξιάτικης περιόδου. (Σχ. 4.31)



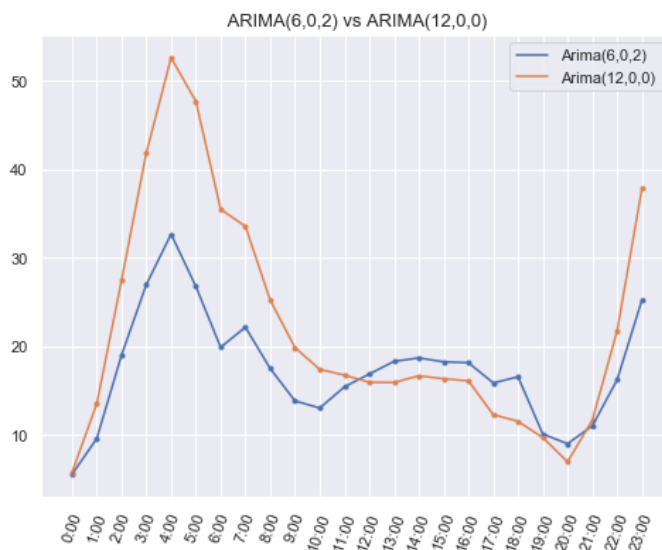
Σχήμα 4.31: Κατανομή του σφάλματος της πρόβλεψης για το μήνα Ιούνιο 2013 με ARIMA(24,0,1) για διάφορες ώρες της ημέρας



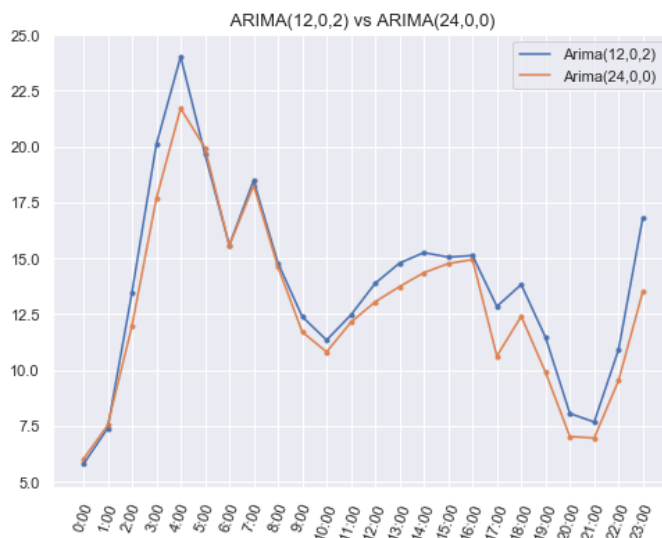
Σχήμα 4.32: Κατανομή του σφάλματος της πρόβλεψης για την ώρα 5:00 κατά το μήνα Ιούνιο 2013 με τα διάφορα μοντέλα ARIMA

Ο οριζόντιος άξονας απεικονίζει τις τιμές του MAPE, και βλέπουμε ότι η καμπύλη που αντιστοιχεί στις 11.00 συγκεντρώνεται στις τιμές 0-20 (δηλαδή MAPE 0%-20%). Συγκρίνοντας για τις 5.00 το πρωί την κατανομή του σφάλματος για τα διάφορα μοντέλα ARIMA, (Σχ. 4.32) βλέπουμε ότι όσο αυξάνεται η τάξη του p τόσο η «καμπάνα» στενεύει και κινείται σε πιο χαμηλές τιμές σφάλματος.

Τέλος, παρατηρήσαμε ότι ένα μοντέλο με μικρότερη πολυπλοκότητα και λιγότερο υπολογιστικό κόστος μπορεί να έχει ίδια ή και καλύτερη απόδοση, χάρη στην προσθήκη του παράγοντα MA στο μοντέλο. Για παράδειγμα, στην θερινή περίοδο που υλοποιήσαμε AR μοντέλο (αφού είχαμε $q = 0$), διαπιστώσαμε με την προσθήκη του $q = 2$ την βελτίωση της επίδοσης. (Για λόγους υπολογιστικού κόστους δεν υλοποιήσαμε ARIMA(24,0,2), καθώς η εκπαίδευση ήδη του AR(24) ήταν πολύ χρονοβόρα.)



Σχήμα 4.33: MAPE των προβλέψεων με ARIMA(12,0,0) και ARIMA(6,0,2)



Σχήμα 4.34: MAPE των προβλέψεων με ARIMA(12,0,2) και ARIMA(24,0,0)

4.4.2 Σχολιασμός των ΤΝΔ και σύγκριση μεταξύ των διαφορετικών αρχιτεκτονικών

Η συνολική επίδοση των ΤΝΔ ήταν αντίστοιχη, με όλες τις αρχιτεκτονικές που υλοποιήθηκαν. Παρόλο που η λογική των LSTM διαφέρει πολύ από αυτή του CNN, το ολικό MAPE ανά περίοδο πρόβλεψης που παρήγαγαν τα 3 υλοποιημένα μοντέλα είχε διαφορά λιγότερο του 0.5% (μόνο για το καλοκαίρι το CNN πέτυχε κατά 1% μικρότερο σφάλμα από το LSTM). Ας επισημάνουμε, λοιπόν, τα σημεία που διέφεραν τελικά οι αρχιτεκτονικές στην πράξη, και πού υστερούσε η κάθε μία.

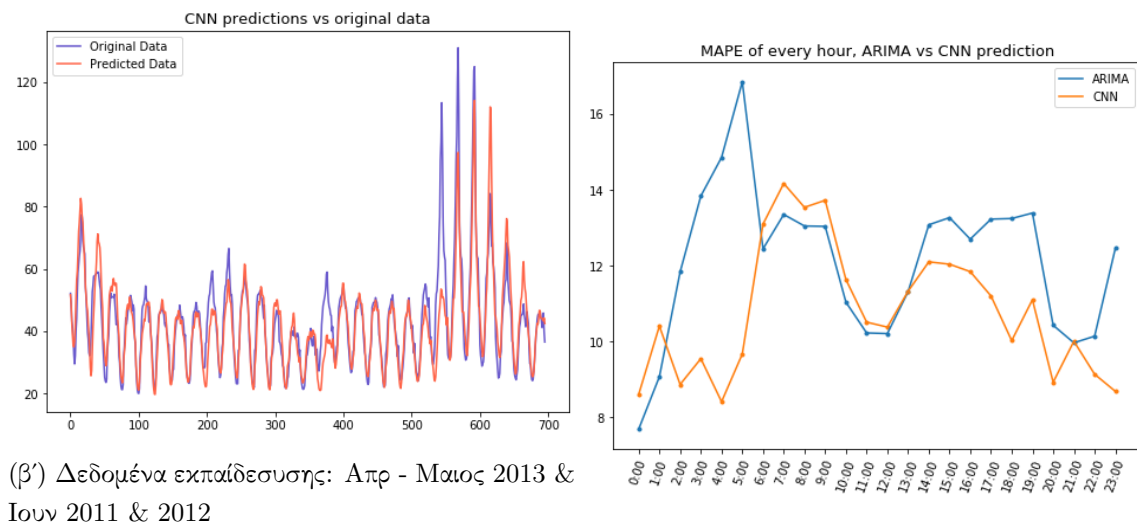
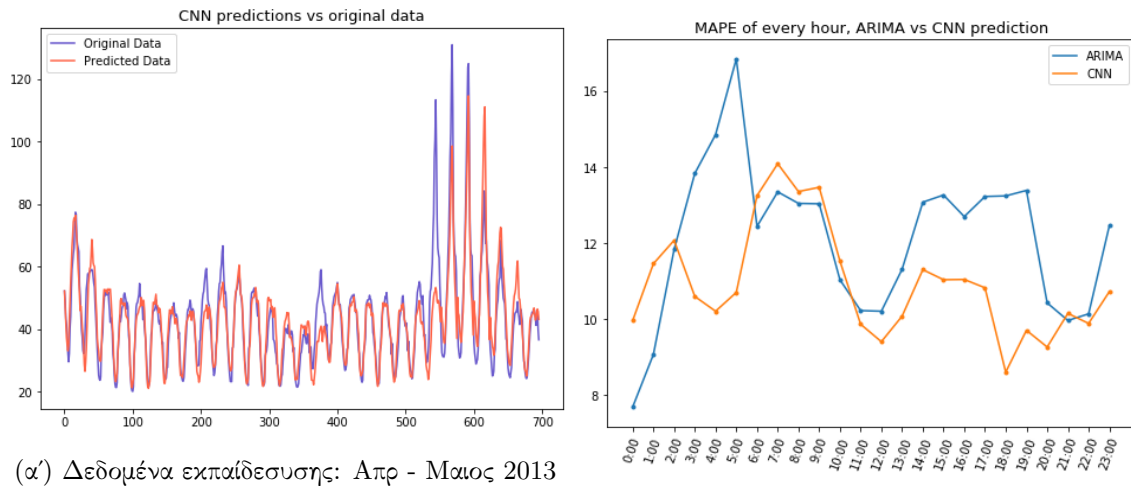
- Πρώτη, και πολύ εμφανής, διαφορά, ήταν ο χρόνος εκτέλεσης των αλγορίθμων στις διάφορες αρχιτεκτονικές. Για ένα σετ 2500 δεδομένων η εκπαίδευση με το LSTM χρειαζόταν περίπου 7s για κάθε εποχή, σε αντίθεση με το <1s του CNN, και τα 4s του μοντέλου CNN & LSTM. Δεν εστιάζουμε το ενδιαφέρον μας στον ακριβή χρόνο, καθώς παίζει ρόλο το περιβάλλον στο οποίο εκτελείται (πχ υπολογιστική δύναμη, χρήση GPU, κ.α.), αλλά αν λάβουμε υπόψη ότι κάθε μοντέλο χρειάστηκε για την εκπαίδευσή του 30-50 εποχές, μπορούμε να καταλάβουμε ότι το CNN ήταν μία ταξη μεγέθους ταχύτερο.
- Επιπλέον, η λογική της λειτουργίας του CNN να εξάγει τα σημαντικά χαρακτηριστικά, αποδείχτηκε ότι απαιτεί και λιγότερες εποχές εκπαίδευσης.
- Μάλιστα, σε συνέχεια με το παραπάνω, το CNN αποδείχτηκε πολύ πιο εύστοχο στην πρόβλεψη της εαρινής περιόδου, που να υπενθυμίσουμε ότι το LSTM βελτίωσε την επίδοσή του μόνο αφού προσθέσαμε περισσότερα δεδομένα παρελθοντικών τιμών στο dataset εκπαίδευσης. Το ίδιο συνέβη και στο τρίτο μοντέλο, που συνδύαζε τις δύο αρχιτεκτονικές, αλλά φάνηκε ότι οι ιδιότητες του LSTM layer υπερίσχυαν.

Ας δούμε όμως τι θα συμβεί και στο δεύτερο μοντέλο μας, το CNN, αν εισάγουμε επιπλέον δεδομένα των ίδιων παρελθοντικών μηνών. Όπως παρατηρούμε στο Σχήμα 4.35, η προσθήκη των δεδομένων των Ιουνίων 2012 και 2011 δεν βελτίωσε την πρόβλεψη, και τα μοντέλα είχαν αντίστοιχη επίδοση. Τα χρήσιμα χαρακτηριστικά έχουν ήδη εξαχθεί επαρκώς από τα λιγότερα δεδομένα εκπαίδευσης και είναι ικανά να προβλέψουν με μεγαλύτερη ακρίβεια, ακόμα και σε σχέση με το βέλτιστο LSTM.

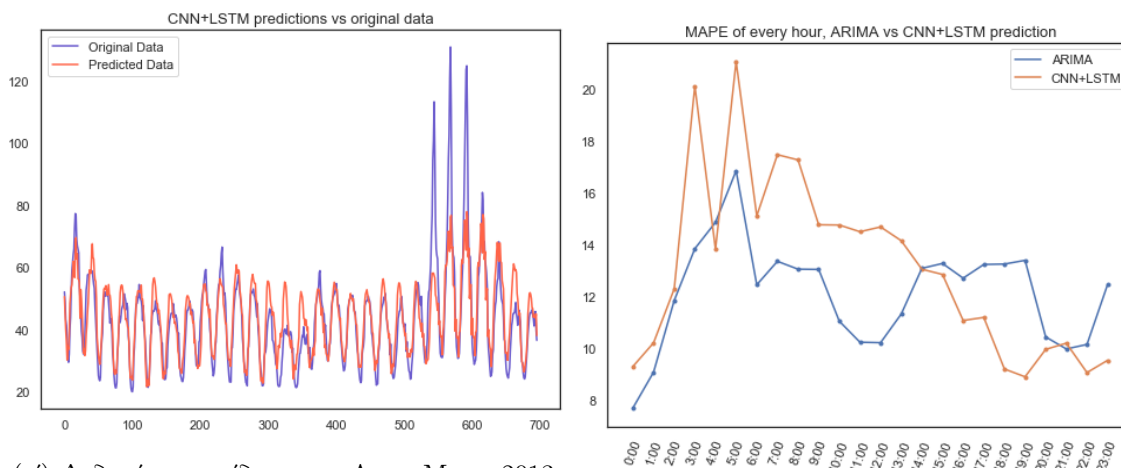
Ας δούμε και τη συμπεριφορά του τρίτου μοντέλου (CNN & LSTM) στην αντίστοιχη περίπτωση. (Σχήμα 4.36) Η βελτίωση με την αύξηση των δεδομένων εκπαίδευσης είναι εμφανής. Βέβαια, το μοντέλο κυρίως βελτίωσε τις προβλέψεις των χαμηλών τιμών, παρά των υψηλών (παρατηρούμε ότι και πάλι τις ακραίες αυξήσεις τιμών αδυνατούσε να τις προβλέψει- το ίδιο είχε συμβεί και με το απλό LSTM). Βλέπουμε ότι, όταν η οριακή τιμή της ενέργειας κυμαίνεται περίπου από 40 έως 80 μονάδες, οι προβλέψεις είναι πολύ εύστοχες. Είναι ενδιαφέρον, λοιπόν, να οπτικοποιήσουμε την κατανομή του σφάλματος ανάλογα με την πραγματική τιμή της ηλεκτρικής ενέργειας, και το σφάλμα σε έναν μέσο όρο με αυτό το κριτήριο ομαδοποίησης. Η κατανομή αυτή φαίνεται στο

Σχήμα 4.37. Ανεξάρτητα με το σετ δεδομένων εκπαίδευσης, όταν η πραγματική τιμή ηλεκτρικής ενέργειας βρίσκεται μεταξύ 40 και 80, το σφάλμα συγκεντρώνεται στο 0 - 10% περίπου, ενώ η καμπύλη της κατανομής των σφαλμάτων στην περίπτωση που η τιμή είναι κάτω από 40 διαφοροποιείται πολύ με την προσθήκη των επιπλέον δεδομένων εκπαίδευσης : η καμπάνα στενεύει και αυτή στο 5-10%.

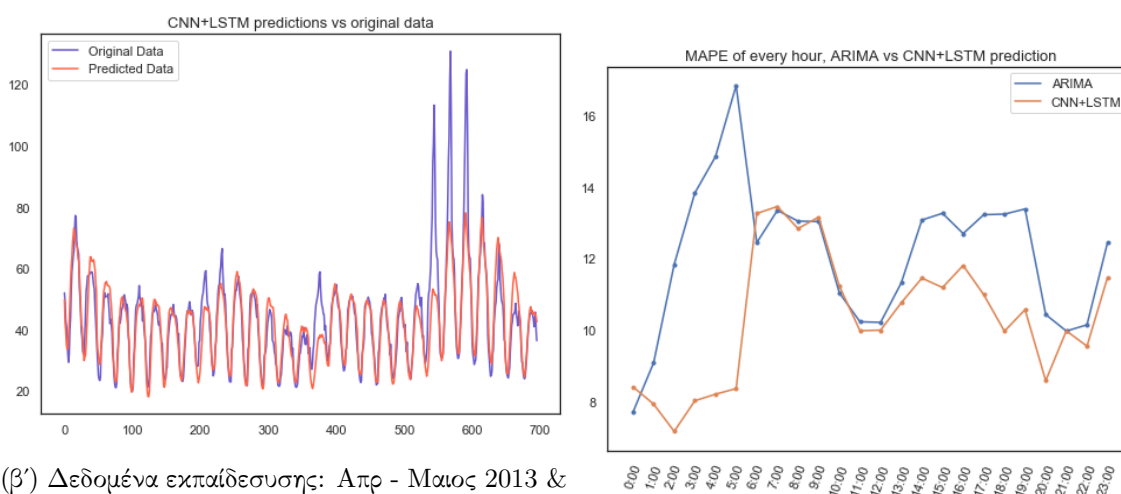
Στον Πίνακα 4.9 παρατίθενται τα μέσα απόλυτα ποσοστιαία σφάλματα των προβλέψεων της εαρινής περιόδου, ομαδοποιημένα ανάλογα με την πραγματική οριακή τιμή της ηλεκτρικής ενέργειας.



Σχήμα 4.35: Προβλέψεις και σφάλματα με μοντέλο CNN για την εαρινή περίοδο (Ιουν 2013), χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης



(α') Δεδομένα εκπαίδευσης: Απρ - Μαιος 2013

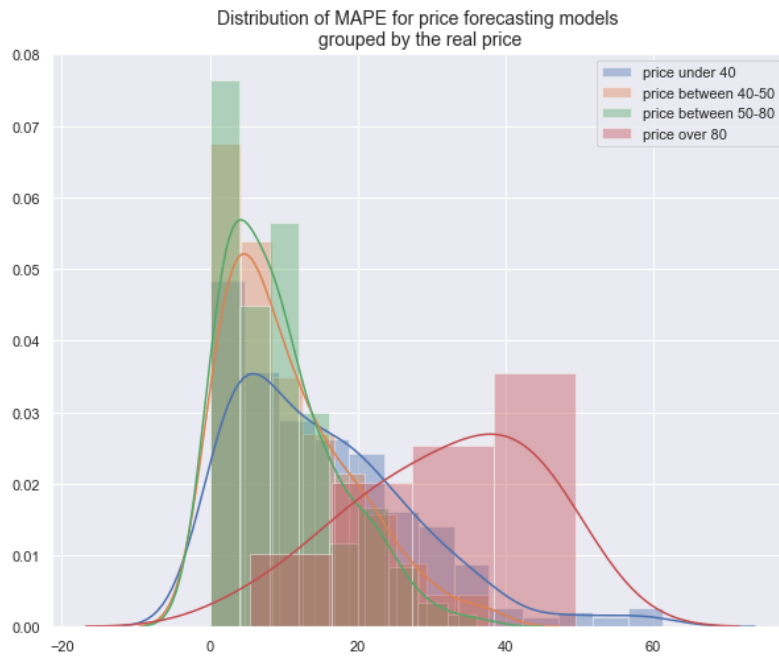


(β') Δεδομένα εκπαίδευσης: Απρ - Μαιος 2013 & Ιουν 2011 & 2012

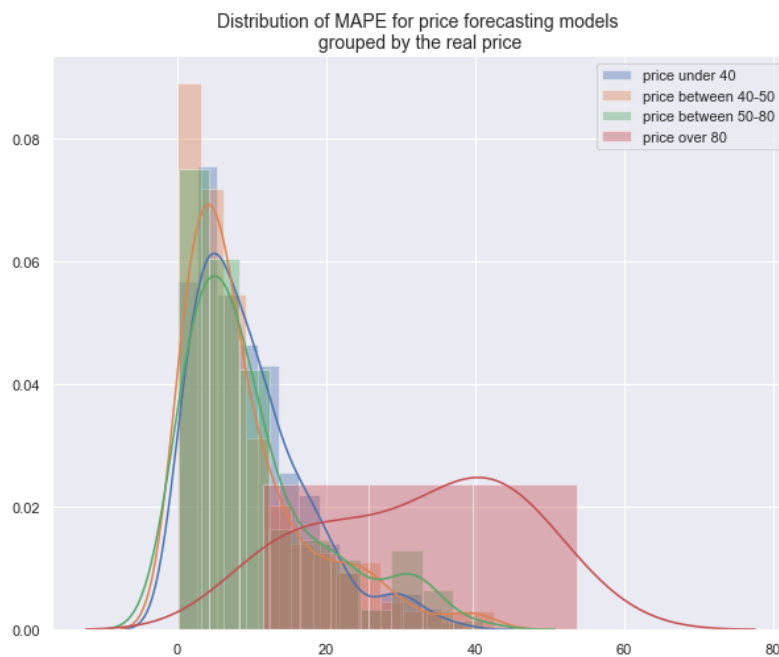
Σχήμα 4.36: Προβλέψεις και σφάλματα με μοντέλο CNN & LSTM για την εαρινή περίοδο, χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης

Model \ Price	LSTM	LSTM+	CNN	CNN & LSTM	CNN & LSTM+
under 40	14.006387	9.3316539	9.8351937	15.4310108	9.8040284
40-50	9.4324480	9.8890973	9.1039099	10.6558634	9.2475724
50-80	10.3810674	10.9128241	12.6996830	9.34409310	10.330720
over 80	24.6412246	34.6812066	26.1403593	31.9540626	32.557538

Πίνακας 4.9: MAPE των προβλέψεων εαρινής περιόδου των διαφόρων νευρωνικών μοντέλων, ανάλογα με την πραγματική οριακή τιμή της ηλεκτρικής ενέργειας. Το «+» αναφέρεται στην περίπτωση εκπαίδευσης με επιπλέον δεδομένα παρελθόντος



(α')



(β')

Σχήμα 4.37: Προβλέψεις και σφάλματα με μοντέλο CNN για την εαρινή περίοδο, χωρίς (α') και με (β') προσθήκη τιμών παλαιότερων ετών στα δεδομένα εκπαίδευσης

4.4.3 Σύγκριση των ΤΝΔ με τις στατιστικές μεθόδους πρόβλεψης

Τελικά, συγκρίνοντας τις προβλέψεις των νευρωνικών μοντέλων με των στατιστικών, τα τεχνητά νευρωνικά δίκτυα υπερτερούν ως προς την ακρίβεια της πρόβλεψης. Ακολουθεί και συγκεντρωτικός πίνακας σφαλμάτων για τα διάφορα μοντέλα, ανά περίοδο πρόβλεψης.

Model \ Period	ARIMA	LSTM	CNN	CNN&LSTM
autumn	10.3814233	9.423878	9.7694650	9.9975536
summer	12.8588003	11.6351827	10.2568160	10.5617418
winter	12.9432410	11.5993538	11.1548018	11.4704586
spring	12.1059042	10.5890539	10.9365406	10.3364972

Πίνακας 4.10: Σύγκριση του MAPE των υλοποιηθέντων μοντέλων πρόβλεψης

Η σύγκριση της επίδοσης ανά ώρα φάνηκε στις γραφικές παραστάσεις των μέσων σφαλμάτων που παρουσιάστηκαν στην ενότητα 4.3, και εκεί αποδεικνύεται επίσης ότι όλα τα μοντέλα με νευρωνικά δίκτυα είχαν χαμηλότερα σφάλματα από το ARIMA, σχεδόν για όλες τις ώρες της μέρας σε κάθε περίοδο. Μεταξύ των νευρωνικών δικτύων, μάλιστα, ξεχωρίσαμε το συνολικτικό ως βέλτιστο.

Στις περισσότερες περιπτώσεις παρατηρήθηκε ότι η καμπύλη του μέσου, ανά ώρα, σφάλματος είχε αντίστοιχη συμπεριφορά για τα στατιστικά και τα νευρωνικά δίκτυα, απλά για τα δεύτερα κυμαινόταν σε χαμηλότερα επίπεδα. (Χαρακτηριστικά παραδείγματα στα Σχήματα 4.18, 4.19, 4.20).

Να επισημάνουμε ότι οι καλύτερες, συνολικά, προβλέψεις έγιναν για το μήνα Δεκέμβριο 2011, με δεδομένα εκπαίδευσης τις τιμές των δύο προηγούμενων μηνών μόνο, κάτι που ήταν αναμενόμενο, αν παρατηρήσουμε απλά τη μορφή της χρονοσειρά: έχει τις λιγότερες απότομες αυξομειώσεις και ακραίες τιμές.

Συνολικά, βέβαια, για τα μοντέλα των νευρωνικών δικτύων δεν απαιτήθηκε ιδιαίτερη προμελέτη της μορφής των δεδομένων και, επιπλέον, μετά την εκπαίδευση του μοντέλου εφαρμόστηκαν σε αυτό άγνωστα δεδομένα, για τα οποία έγιναν προβλέψεις. Αντίθετα, οι στατιστικές τεχνικές απαιτούν προεπεξεργασία των δεδομένων και παραδοχές από τον αναλυτή, και για κάθε νέα πρόβλεψη γίνεται αντίστοιχα μία εκπαίδευση, δηλαδή δεν υπάρχει η δυνατότητα «γενίκευσης».

Άλλο ένα σημείο που αποδείχτηκε να υπερτερούν τα ΤΝΔ ήταν ο χρόνος εκτέλεσης, αφού αλγόριθμοι που υλοποιήθηκαν χρειάζονταν από 30 δευτερόλεπτα έως περίπου 10-12 λεπτά για να εκτελεστούν (και μάλιστα με χρήση απλά της CPU), ενώ τα βέλτιστα και πιο πολύπλοκα μοντέλα ARIMA απαιτούσαν για την εκτέλεσή τους τουλάχιστον 6 ώρες (με εκμετάλλευση της GPU).

Βασικό μειονέκτημα των νευρωνικών δικτύων και των διαδικασιών εκμάθησης συνιστά το ότι το τελικό αποτέλεσμα είναι, σε μικρότερο ή μεγαλύτερο βαθμό, ένα «μαύρο κουτί», και, χωρίς την πλήρη κατανόηση της αρχιτεκτονικής τους και του υπάρχοντος προβλήματος προς

επίλυση, η επιλογή των παραμέτρων και υπερπαραμέτρων και η βελτίωση των επιδόσεων του μοντέλου μπορεί να γίνει τελικά «στην τύχη».

Συνολικά, η επιτυχία των σύγχρονων τεχνικών πρόβλεψης με ΤΝΔ, που έχουν ήδη αποδειχθεί σε πολλούς τομείς, τεκμηριώθηκε και στην πρόβλεψη της οριακής τιμής της ηλεκτρικής ενέργειας της προ-ημερήσιας αγοράς.

Κεφάλαιο 5

Επίλογος

5.1 Σύνοψη και συμπεράσματα

Συνοψίζοντας, μέσα στην ανταγωνιστική αγορά ηλεκτρικής ενέργειας, όλοι οι συμμετέχοντες που αποσκοπούν στην αύξηση του κέρδους τους θα ευνοηθούν από μια εύστοχη πρόβλεψη της ωριακής τιμής της επόμενης ημέρας. Το συγκεκριμένο πρόβλημα πρόβλεψης αποτελεί μεγαλύτερη πρόκληση, συγκριτικά με την πρόβλεψη της παραγωγής ή της ζήτησης φορτίου, αν και σχετίζεται άμεσα με αυτά, περιλαμβάνει όμως περισσότερους αστάθμητους παράγοντες, εφόσον πρόκειται για πρόβλημα όχι μόνο ενεργειακής αλλά και οικονομικής φύσης. Κυριότερος σκοπός της διπλωματικής εργασίας ήταν, μέσα από τη μελέτη και της εφαρμογή των διαφόρων τεχνικών προβλέψεων, να διερευνήσουμε την αξιοπιστία των τεχνητών νευρωνικών δικτύων για την πρόβλεψη της ωριαίας τιμής ενέργειας και την υπεροχή τους έναντι των παραδοσιακών μεθόδων. Ένας παράγοντας που καθιστά μια καλή πρόβλεψη δύσκολη είναι ότι στοχεύουμε σε 24 επόμενες τιμές, και όχι μόνο μία.

Έχοντας στη διάθεσή μας, λοιπόν, μια χρονοσειρά τιμών ηλεκτρικής ενέργειας ανά ώρα για διάστημα περίπου 3 χρόνων, προχωρήσαμε στην επίλυση του προβλήματός μας. Θέσαμε στόχο την πρόβλεψη των τιμών ενέργειας κάθε ημέρας για διάστημα ενός μήνα, και επιλέξαμε τέσσερις διαφορετικούς μήνες, ενδεικτικούς για διαφορετικές εποχές στη διάρκεια του έτους. Σε πρώτη φάση, πραγματοποιήθηκε πρόβλεψη των απαιτούμενων τιμών με στατιστικές μεθόδους, διαδικασία, όμως, που απαιτούσε ιδιαίτερη μελέτη της συμπεριφοράς της χρονοσειράς, και πιθανή προεπεξεργασία των δεδομένων. Αφού επιλέχθηκαν τα κατάλληλα στατιστικά μοντέλα πρόβλεψης, τα λάθη αυτών χρησιμοποιήθηκαν για να κρίνουμε την επίδοση των νευρωνικών δικτύων. Σε δεύτερη φάση, υλοποιήθηκαν αλγόριθμοι με τεχνητά νευρωνικά δίκτυα τριών διαφορετικών αρχιτεκτονικών, τα οποία πραγματοποιούσαν πρόβλεψη των ωριαίων τιμών ενέργειας για τα ίδια διαστήματα. Οι αρχιτεκτονικές που επιλέχθηκαν ως κατάλληλες για την επίλυση ενός προβλήματος $seq2seq$, δηλαδή με είσοδο μια ακολουθία και έξοδο μια άλλη ακολουθία, ήταν τα δίκτυα με νευρώνες μακράς-βραχείας μνήμης (LSTM), τα συνελικτικά νευρωνικά δίκτυα (CNN), και ένας συνδυασμός αυτών. Η ευελιξία των νευρωνικών να προσαρμόζονται στα δεδομένα, χωρίς πρότερη γνώση των συσχετίσεων μεταξύ δεδομένων εισόδου και εξόδου, η καλύτερη επίδοση σε σχέση με τις στατιστικές μεθόδους πρόβλεψης

και το γεγονός ότι δεν απαιτούν τον καθορισμό της μορφής του μοντέλου στο βαθμό που απαιτεί ένα στατιστικό μοντέλο, ανέδειξε την αξιοπιστία των τεχνητών νευρωνικών δικτύων στην πρόβλεψη των χρονοσειρών τιμής ηλεκτρικής ενέργειας. Τέλος, η μη γραμμική φύση τους, εν αντιθέσει με τα στατιστικά μοντέλα, η ανθεκτικότητα στο θόρυβο και η ποικιλία των αρχιτεκτονικών δομών κατάλληλων για την επίλυση των διαφόρων προβλημάτων, θέτουν τα τεχνητά νευρωνικά δίκτυα σε πλεονεκτική θέση, όταν τίθεται το δίλημμα μεταξύ αυτών και των παραδοσιακών στατιστικών τεχνικών πρόβλεψης.

5.2 Μελλοντικές επεκτάσεις

Τα τεχνητά νευρωνικά δίκτυα αποτελούν έναν τομέα που εξελίσσεται, οπότε αυτό το γεγονός από μόνο του δίνει πολλές δυνατότητες για μελλοντικές προεκτάσεις. Ανεξάρτητα από αυτό, μια πιθανή μελλοντική κατεύθυνση, με μικρή προσαρμογή των μοντέλων μας, είναι η ενσωμάτωση εξωτερικών παραγόντων ως δεδομένων εισόδου, για την πρόβλεψη της τιμής της ηλεκτρικής ενέργειας. Τέτοιοι παράγοντες είναι π.χ. τα επίπεδα παραγωγής και ζήτησης φορτίου ανά ώρα, σε αντιστοιχία με τις παρελθοντικές τιμές της ενέργειας, ή η προβλεπόμενη ποσότητα τους για την επόμενη ημέρα. Επιπλέον, τα εκπαιδευμένα μοντέλα θα μπορούσαν να εφαρμοστούν και σε πιο απαιτητικά δεδομένα, που να παρουσιάζουν μεγαλύτερη αστάθεια και ακραίες τιμές. Για τέτοιου είδους δεδομένα θα ήταν χρήσιμη η επαύξηση του συνόλου τους, ώστε να μοντελοποιηθεί η πληροφορία καλύτερα. Τέλος, ιδιαίτερη αξία θα είχε η επέκταση του προτεινόμενου μοντέλου σε ρεαλιστικές συνθήκες, που συνεπάγεται πρόβλεψη των τιμών της επόμενης ημέρας πριν τη λήξη υποβολής προσφορών (περίπου 12 μ.μ. προηγούμενης μέρας δηλαδή), και πιθανώς λαμβάνοντας υπόψη τις στρατηγικές υποβολής προσφορών, από την πλευρά των παραγωγών. Στο πλαίσιο αυτό, σημαντική αλλά και αρκετά απαιτητική θα ήταν η ενσωμάτωση της αβεβαιότητας σχετικά με τις συνθήκες που επικρατούν στην αγορά (αναφορικά με το φορτίο του συστήματος, τις προσφορές των ανταγωνιστών, την παραγωγή των ΑΠΕ, κ.α.).

Βιβλιογραφία

- [1] P. Ranci, Guido Cervigni. *Economics of Electricity Markets: Theory and Policy*. The Loyola de Palacio Series on European Energy Policy. Edward Elgar, 2013.
- [2] Ευάγγελος Γ. Καρδάκος. Προσομοίωση αγοράς & στρατηγική συμπεριφορά παραγωγών ηλεκτρικής ενέργειας. Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Μάρτιος 2016. Διδακτορική Διατριβή.
- [3] C. P. Rodriguez and G. J. Anders. Energy price forecasting in the Ontario competitive power system market. *IEEE Transactions on Power Systems*, 19(1):366–374, 2004.
- [4] G. Koreneff and A. Seppala and M. Lehtonen and V. Kekkonen and E. Laitinen and J. Hakli and E. Antila. Electricity spot price forecasting as a part of energy management in de-regulated power market. In *Proceedings of EMPD '98. 1998 International Conference on Energy Management and Power Delivery (Cat. No.98EX137)*, ολυμε 1, παγες 223–228 ολ.1, 1998.
- [5] Γεωργία Καλαμβόκη. Μέθοδοι Πρόβλεψης Χρονοσειρών: Χρονοσειρές στην Ελληνική Οικονομία. Τμήμα Μαθηματικών, Πανεπιστήμιο Πατρών, Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών "Μαθηματικά των Υπολογιστών και των Αποφάσεων", Φεβρουάριος 2017. Διπλωματική εργασία.
- [6] Ευάγγελος Σπηλιώτης. Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου (ARIMA). Εθνικό Μετσόβιο Πολυτεχνείο. Σημειώσεις για το μάθημα "Τεχνικές Προβλέψεων".
- [7] Κοσμάς Ηλίας. Αυτόματη ταξινόμηση εικόνων βιοψίας μαστικού ιστού με συνελκτικά νευρωνικά δίκτυα. Πανεπιστήμιο Πατρών, 'Σεπτέμβριος 2018. Διπλωματική εργασία.
- [8] Neural Network Models in R. <https://www.datacamp.com/community/tutorials/neural-network-models-r>. Accessed: 2020-06-24.
- [9] Θεόδωρος Μ. Πίσσας. Αναγνώριση Ανθρώπινης Δράσης και Χειρονομιών χρησιμοποιώντας Συνελκτικά και Αναδρομικά Νευρωνικά Δίκτυα. Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων, Εθνικό Μετσόβιο Πολυτεχνείο, Ιούλιος 2017. Διπλωματική εργασία.

- [10] Activation Functions in Neural Networks. <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>. Accessed: 2020-06-24.
- [11] Feedforward Neural Networks. <https://brilliant.org/wiki/feedforward-neural-networks/>. Accessed: 2020-06-24.
- [12] Artificial Neural Network - Building Blocks. https://www.tutorialspoint.com/artificial_neural_network/artificial_neural_network_building_blocks.htm. Accessed: 2020-06-19.
- [13] Ιωάννης Βλαχάβας. Τεχνητά Νευρωνικά Δίκτυα. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 2013. Accessed: 2020-06-18.
- [14] De-Shuang Huang, M. Michael Gromiha, Kyungsook Han, Abir Hussain. Springer International Publishing, 2018.
- [15] Understanding LSTM Networks. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Accessed: 2020-06-20.
- [16] Zhang, Yu and Wu, Peng and Shan, Baoguo and Fang, Chaoxiong and Lin, Yi and Xiong, Ke. An electric energy substitute planning model with economic cost and environmental constraints. *IOP Conference Series: Earth and Environmental Science*, 227:042009, 03 2019.
- [17] Antonio Gomez-Exposito, Antonio J. Conejo, Claudio Canizares. *Electric Energy Systems. Analysis and Operation*. CRC Press, 2018.
- [18] Pelin Ercan, Javier Soto. A model for long term electricity price forecasting for France. School of Electrical Engineering KTH, Royal Institute of Technology Stockholm, 2011. Master thesis.
- [19] Δ. Ι. Χατζηγιάννης. Ολοκλήρωση ενεργειακών αγορών. Εργαστήριο Συστημάτων Ηλεκτρικής Ενέργειας, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Ιούνιος 2014. Διδακτορική Διατριβή.
- [20] S. Hunt. *Making Competition Work in Electricity*. John Wiley & Sons (Wiley), 2002.
- [21] Ρυθμιστική Αρχή Ενέργειας. <http://www.rae.gr/site/>.
- [22] Ευαγγελία Χατζόγλου. Η δομή της χονδρικής αγοράς ηλεκτρικής ενέργειας και ο κανονισμός REMIT. Μεταπτυχιακό πρόγραμμα Ενέργεια: Δίκαιο Οικονομία και Στρατηγική, Πανεπιστήμιο Πειραιώς, Σεπτέμβριος 2016. Διπλωματική εργασία.
- [23] H. Zareipour and K. Bhattacharya and C. A. Canizares. Forecasting the hourly Ontario energy price by multivariate adaptive regression splines. In *2006 IEEE Power Engineering Society General Meeting*, 2006.

- [24] N. Amjady and M. Hemmati. Energy price forecasting - problems and proposals for such predictions. *IEEE Power and Energy Magazine*, 4(2):20–29, 2006.
- [25] European Commission. Οι τιμές και το κόστος ενέργειας στην Ευρώπη, 2014. Ανακοίνωση της επιτροπής προς το Ευρωπαϊκό Κοινοβούλιο, την Ευρωπαϊκή Οικονομική και Κοινωνική Επιτροπή και την Επιτροπή των περιφερειών.
- [26] European Commission. Οι τιμές και το κόστος ενέργειας στην Ευρώπη, 2019. Έκθεση της επιτροπής προς το Ευρωπαϊκό Κοινοβούλιο, την Ευρωπαϊκή Οικονομική και Κοινωνική Επιτροπή και την Επιτροπή των περιφερειών.
- [27] E. Dikaios Tserkezos. Forecasting residential electricity consumption in Greece using monthly and quarterly data. *Energy Economics*, 14(3):226 – 232, 1992.
- [28] Y. H. Fung and V. M. Rao Tummala. Forecasting of electricity consumption: a comparative analysis of regression and artificial neural network models. In *1993 2nd International Conference on Advances in Power System Control, Operation and Management, APSCOM-93.*, παγες 782–787 ολ.2, 1993.
- [29] B. R. Szkuta and L. A. Sanabria and T. S. Dillon. Electricity price short-term forecasting using artificial neural networks. *IEEE Transactions on Power Systems*, 14(3):851–857, 1999.
- [30] Jakasa, Tina and Androcec, Ivan and Sprcic, Petar. Electricity price forecasting — ARIMA model approach. 05 2011.
- [31] X. Chen and Z. Y. Dong and K. Meng and Y. Xu and K. P. Wong and H. W. Ngan. Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping. *IEEE Transactions on Power Systems*, 27(4):2055–2062, 2012.
- [32] S. Anbazhagan and N. Kumarappan. Day-Ahead Deregulated Electricity Market Price Forecasting Using Recurrent Neural Network. *IEEE Systems Journal*, 7(4):866–872, 2013.
- [33] Z. Chang and Y. Zhang and W. Chen. Effective Adam-Optimized LSTM Neural Network for Electricity Price Forecasting. In *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, παγες 245–248, 2018.
- [34] Pickup, M. *Introduction to Time Series Analysis*. Quantitative Applications in the Social Sciences. SAGE Publications, 2014.
- [35] Dr. Sevtap Kestel. Lecture Notes on Time Series Analysis. Albert-Ludwigs University Freiburg, 2011.
- [36] Αρετή Παππά. Μελέτη της επίδρασης ατμοσφαιρικών παραμέτρων στη ρήξη ανευρυσμάτων κοιλιακής αορτής. Τμήμα Φυσικής, Πανεπιστήμιο Πατρών, Μεταπτυχιακό πρόγραμμα στην εφαρμοσμένη μετεωρολογία, Μάρτιος 2019. Διπλωματική εργασία.

- [37] Κατερίνα Γεωργούλη. *Τεχνητή Νοημοσύνη*. Ύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών Εθνικό Μετσόβιο Πολυτεχνείο, 2015.
- [38] Αντωνία Πλέρου. Τεχνητά νευρωνικά δίκτυα προσομοίωσης του ανθρώπινου εγκεφάλου. *Ανοικτή Εκπαίδευση: το περιοδικό για την Ανοικτή και εξ Αποστάσεως Εκπαίδευση και την Εκπαιδευτική Τεχνολογία*, 8(1):128–135, 2012.
- [39] Μπροκαλάκης Ιωσήφ. Πρόβλεψη φαρμακευτικών πωλήσεων με τη χρήση τεχνητών νευρωνικών δικτύων και νευρο-ασαφών συστημάτων. Πολυτεχνείο Κρήτης, 'Μάιος 2014. Διπλωματική εργασία.
- [40] S. Khan and H. Rahmani and S. A. A. Shah and M. Bennamoun and G. Medioni and S. Dickinson. *A Guide to Convolutional Neural Networks for Computer Vision*. 2018.
- [41] Μελάς Χρήστος. Μελέτη νευρωνικών δικτύων και εφαρμογές στην πρόβλεψη χρονοσειρών στο χρηματιστήριο. Πανεπιστήμιο Πειραιώς, 'Φεβρουάριος 2020. Διπλωματική εργασία.
- [42] Θεόδωρος Ι. Φλωράκης. Ανάπτυξη Αρχιτεκτονικών Βαθιάς Μάθησης με Χρήση Δικτύων Μακράς Βραχυπρόθεσμης Μνήμης Για Ανάλυση Συναισθήματος. Εθνικό Μετσόβιο Πολυτεχνείο, 'Νοέμβριος 2018. Διπλωματική εργασία.
- [43] Ευγενία Κ. Κοτζαπαναγιώτου. Χρήση Μεθόδων Μηχανικής Μάθησης για Πρόβλεψη Μελλοντικών Τιμών Χρονοσειρών. Πανεπιστήμιο Πατρών, 'Σεπτέμβριος 2019. Διπλωματική εργασία.
- [44] M. Butler and D. Kazakov. The effects of variable stationarity in a financial time-series on Artificial Neural Networks. In *2011 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFER)*, pages 1–8, 2011.

Γλωσσάριο

Ελληνικός όρος

αναδρομικό νευρωνικό δίκτυο
δεδομένα εκπαίδευσης
δεδομένα αξιολόγησης
δεδομένα εκπαίδευσης
δεδομένα ελέγχου
διαχειριστής αγοράς
διαχειριστής συστήματος
δίκτυο νευρώνων μακράς-βραχείας μνήμης
διμερή συμβόλαια
εμπρόσθια τροφοδότηση
ενδο-ημερήσια αγορά
ενισχυτική μάθηση
κοινοπραξία ισχύος
μάθηση με επίβλεψη
μάθηση χωρίς επίβλεψη
μέσο απόλυτο ποσοστιαίο σφάλμα
οριακή τιμή συστήματος
πλήρως συνδεδεμένο
πόλωση
προ-ημερήσια αγορά
προθεσμιακή αγορά
σημείο αναφοράς
στασιμότητα
στρώμα
συνάρτηση ενεργοποίησης
συνελικτικό νευρωνικό δίκτυο
τεχνητό νευρωνικό δίκτυο, ΤΝΔ
υποβολή προσφορών
χρηματιστήριο ενέργειας

Αγγλικός όρος

recurrent neural network
training dataset
validation dataset
training dataset
test dataset
market operator
system operator
long term short term memory neural network
bilateral contracts
feedforward
intraday or adjustment market
reinforcement learning, RL
power pool
supervised learning
unsupervised learning
mean absolute percentage error, MAPE
marginal clearing price, MCP
fully connected
bias
day-ahead market
forward market
benchmark
stationarity
layer
activation function
convolutional neural network
artificial neural network, ANN
bidding
power exchange

