ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ
ΚΑΙ ΔΙΟΙΚΗΣΗΣ

Πειραματική εφαρμογή, αξιολόγηση και σύγκριση μεθόδων πρόβλεψης της Οριακής Τιμής Συστήματος της ηλεκτρικής ενέργειας: μελέτη στην Ελληνική Αγορά Ενέργειας.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Εμμανουήλ Κ. Μουνταλάς

Επιβλέπων: Ασημακόπουλος Βασίλειος
Καθηγητής Ε.Μ.Π.

Υπεύθυνος: Σπιθουράκης Γεώργιος
Μεταπτυχιακός Φοιτητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

Πειραματική εφαρμογή, αξιολόγηση και σύγκριση μεθόδων πρόβλεψης της Οριακής Τιμής Συστήματος της ηλεκτρικής ενέργειας: μελέτη στην Ελληνική Αγορά Ενέργειας.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Εμμανουήλ Κ. Μουνταλάς

Επιβλέπων: Ασημακόπουλος Βασίλειος

Καθηγητής Ε.Μ.Π.

Υπεύθυνος: Σπιθουράκης Γεώργιος

Μεταπτυχιακός Φοιτητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 12η Ιουλίου 2012.

……………………..	……………………..	……………………..
Βασίλειος Ασημακόπουλος	Ιωάννης Ψαρράς	Γρηγόριος Μέντζας
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012

....

Εμμανουήλ Κ. Μουνταλάς

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

Ευχαριστίες…..

Ευχαριστώ πάρα πολύ το συνάδελφο Σπιθουράκη Γεώργιο, καθώς πέρα από υπεύθυνος για την διπλωματική, είναι και εξαίρετος φίλος, και βοήθησε σημαντικά στην εκπόνηση και τελειοποίηση της παρούσας εργασίας.

Ευχαριστώ επίσης όλα τα παιδιά της μονάδας Προβλέψεων και Στρατηγικής (FSU) καθώς και τον υπεύθυνο του εργαστηρίου και Διδάκτορα του Ε.Μ.Π. Πετρόπουλο Φώτιο για την συμπαράσταση και τη βοήθεια που μου παρείχαν όποτε αυτή ήταν απαραίτητη.

Η παρούσα διπλωματική είναι αφιερωμένη στους κοντινούς μου ανθρώπους και στην οικογένειά μου η οποία μου παρείχε ότι ήταν απαραίτητο καθ' όλη τη διάρκεια των σπουδών μου, αλλά και σε ένα μέλος της που δεν είναι πια κοντά μας.

Περίληψη

Σκοπός της παρούσας Διπλωματικής, είναι η εφαρμογή και αξιολόγηση διάφορων μοντέλων πρόβλεψης της Οριακής Τιμής Συστήματος της ηλεκτρικής ενέργειας στην Ελληνική Αγορά Ενέργειας. Το συγκεκριμένο εγχείρημα επιτυγχάνεται μέσω του σχεδιασμού κατάλληλης διαδικασίας προσομοίωσης η οποία βασίζεται στη λογική των κυλιόμενων προβλέψεων (rolling forecasting). Στα πλαίσια αυτής της διαδικασίας, εφαρμόζονται 6 διαφορετικά μοντέλα πρόβλεψης τα οποία είναι μοντέλα εκθετικής εξομάλυνσης, το μοντέλο απλής γραμμικής παλινδρόμησης, η συνδυαστική μέθοδος Theta και το μοντέλο αναφοράς Naive.

Αρχικά, γίνεται μια γενική θεώρηση των επικρατούντων συνθηκών στην Ελληνική Αγορά Ενέργειας, ενώ παράλληλα εισάγονται οι έννοιες της Οριακής Τιμής Συστήματος (ΟΤΣ) και του Ημερήσιου Ενεργειακού Προγραμματισμού (ΗΕΠ). Στη συνέχεια παρουσιάζονται οι σχετικές με το αντικείμενο έρευνες, οι οποίες διαχωρίζονται με βάση τον ορίζοντα πρόβλεψης σε 3 κατηγορίες: βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη. Για κάθε κατηγορία, γίνεται αναφορά στην αντίστοιχη βιβλιογραφία και αναλύονται ορισμένες διαδεδομένες μέθοδοι πρόβλεψης της ΟΤΣ. Επιπλέον, εισάγονται οι έννοιες των ποιοτικών και ποσοτικών χαρακτηριστικών που αντικατοπτρίζουν τη φύση μίας χρονοσειράς, ενώ παρουσιάζονται οι μέθοδοι πρόβλεψης που αφορούν τα μοντέλα χρονοσειρών (Naive, SES, Holt, Damped) αλλά και τα αιτιοκρατικά μοντέλα, δηλαδή τις μεθόδους γραμμικής παλινδρόμησης (LR), καθώς παρουσιάζεται και η συνδυαστική μέθοδος Theta.

Έχοντας καλύψει τη θεωρητική πτυχή του υπό εξέταση εγχειρήματος, δίνεται ιδιαίτερη έμφαση στην ανάλυση της χρονοσειράς της ΟΤΣ, η οποία αποτελεί το βασικό άξονα μελέτης της παρούσας διπλωματικής, ενώ παρουσιάζονται ορισμένες χρονοσειρές οι οποίες μπορούν να ληφθούν ως ανεξάρτητες μεταβλητές παλινδρόμησης (πρόβλεψη φορτίου, τιμή Φυσικού Αερίου κτλ.). Εν συνεχεία, παρουσιάζονται 4 ομάδες αποτελεσμάτων σχετικές με τον τρόπο προσέγγισης της εποχικότητας, οι οποίες προκύπτουν από τη διαδικασία προσομοίωσης για τα μοντέλα χρονοσειρών. Επιπλέον, παρουσιάζεται ποιοτική και ποσοτική ανάλυση των τριών βέλτιστων αποτελεσμάτων (SES, Theta, Damped) και αφορούν την κατηγορία αντιμετώπισης της διπλής εποχικότητας. καθώς εξάγονται τα τελικά συμπεράσματα για τη μεθοδολογία που αναπτύχθηκε αλλά και για τις επιδόσεις των μεθόδων πρόβλεψης που εφαρμόστηκαν. Τέλος, γίνεται αναφορά σε ορισμένα εγχειρήματα τα οποία διατηρήθηκαν εκτός της σφαίρας ενδιαφέροντος της παρούσας εργασίας, αλλά είναι δυνατό να αποτελέσουν αντικείμενο μελλοντικής έρευνας.

Λέξεις Κλειδιά: Οριακή Τιμή Συστήματος, Ελληνική Αγορά Ηλεκτρικής Ενέργειας, μέθοδοι προβλέψεων, κυλιόμενη πρόβλεψη.

Abstract

The scope of the present thesis is the application and evaluation of various forecasting methods for the Market's Clearing Price (MCP), or System's Marginal Price (SMP), of the electricity for the Greek Electricity Market. This task is accomplished through the designing of a proper simulation process that is based on the logic of rolling forecasting. Within the frame of this simulation process, 6 different forecasting models are applied: various exponential smoothing models, the combinatorial Theta method, the linear regression method (LRL) and Naive as a benchmark method.

To begin with, a general review of the Greek electricity market is given while, simultaneously, the concepts of the System's Marginal Price (SMP) and the Daily Energy Programming (DEP) are introduced. Moreover, existing references relevant to the study's object are presented, which are being categorized based on the forecasting horizon they represent. The three categories refer to short-term, mid-term and long-term forecasts, and there is additional reference to the most widely selected forecasting methods for each category. Additionally, various qualitative and quantitative characteristics of the time series are introduced and mainstream time series forecasting methods (Naive, SES, Holt, Damped and Theta), the deterministic/causal models, such as the linear regression method, and the combinatorial classic Theta method are presented.

Having completed the theoretical aspect of the task, the SMP time series, on which the whole project is based on, is analyzed. Additionally, four supplementary time series which can be used as independent regression variables (demand forecast, natural gas prices etc.) are presented, as well the results that are produced by the application of the 6 forecasting methods, during the simulation process. The results are being categorized into 4 categories based on the approach for tackling with the seasonality pattern which is embedded in the SMP time series. Moreover, a detailed qualitative and quantitative analysis of the three optimal results (SES, Theta and Damped) that refer to the category of double seasonality treatment is elaborated, while final conclusions relevant with the results and the simulation process are deduced. Finally, there is a reference to a couple of issues that were not discussed within the frame of the present thesis, nonetheless they are considered to be of major importance for future studies relevant to the SMP time series forecasts, so guidelines for possible future research are given.

Key words: System's Marginal Price, Market's Clearing Price, Greek Electricity Market, forecasting methods, rolling forecasts.

Περιεχόμενα

Ευρετήριο Σχημάτων-Πινάκων

Κεφάλαιο 1

Σχήματα

Κεφάλαιο 2

Πίνακες

Κεφάλαιο 5

Σχήματα

Κεφάλαιο 6

Σχήματα

Πίνακες

Κεφάλαιο 7

Σχήματα

Πίνακες

Κεφάλαιο 8

Σχήματα

Πίνακες

Ευρεία Περίληψη

Το μέγεθος της Οριακής Τιμής Συστήματος (ΟΤΣ), αποτελεί τον πιο σημαντικό παράγοντα στην Αγορά Ηλεκτρικής Ενέργειας, ιδιαίτερα στην περίπτωση που γίνεται λόγος για απελευθερωμένη Αγορά Ενέργειας. Στην Ελλάδα, η Αγορά Ενέργειας είναι δυνατό να χαρακτηρισθεί ως υπό απελευθέρωση, μετατρεπόμενη σταδιακά από τις αρχές της δεκαετίας του 2000. Οι συμμετέχοντες στην Αγορά Ενέργειας, οφείλουν να πληρούν ορισμένες προϋποθέσεις οι οποίες καθορίζονται από τον Διαχειριστή Ελληνικού Συστήματος Μεταφοράς Ηλεκτρικής Ενέργειας (ΔΕΣΜΗΕ), καθώς και τη Ρυθμιστική Αρχή Ενέργειας (ΡΑΕ).

Το εγχείρημα της πρόβλεψης της ΟΤΣ αποτελεί μείζον ζήτημα για την Αγορά Ηλεκτρικής Ενέργειας, και επηρεάζει κάθε συμμετέχοντα σε αυτήν, είτε είναι Παραγωγός είτε Καταναλωτής. Στην παρούσα Διπλωματική, εφαρμόζονται και αξιολογούνται ποικίλα μοντέλα πρόβλεψης, τα οποία χρησιμοποιούνται στα πλαίσια κατάλληλα διαμορφωμένης διαδικασίας προσομοίωσης, με απώτερο στόχο την εξασφάλιση όσο το δυνατόν ακριβέστερων βραχυπρόθεσμων προβλέψεων.

Στο πρώτο Κεφάλαιο, γίνεται παρουσίαση της Ελληνικής Αγοράς Ενέργειας, δίνοντας βάση στον τρόπο με τον οποίο αυτή λειτουργεί, καθώς αναλύεται και ο ρόλος των σημαντικότερων φορέων (ΔΕΣΜΗΕ, ΡΑΕ, ΔΕΗ κτλ.) που συμμετέχουν στην όλη διαδικασία. Επιπλέον, γίνεται ιδιαίτερη αναφορά στη διαδικασία επίλυσης του Ημερήσιου Ενεργειακού Προγραμματισμού (ΗΕΠ), αλλά και στη σπουδαιότητα του μεγέθους της ΟΤΣ.

Η πρόβλεψη της ΟΤΣ έχει απασχολήσει στο παρελθόν αρκετούς ερευνητές, κάθε ένας από τους οποίους προσέγγισε το ζήτημα από τη δική του διαφορετική σκοπιά, χρησιμοποιώντας περισσότερο ή λιγότερο πολύπλοκες μεθοδολογίες πρόβλεψης. Στο Κεφάλαιο 2, παρουσιάζεται σε πρώτη φάση η σπουδαιότητα πρόβλεψης του συγκεκριμένου μεγέθους, ενώ κατηγοριοποιούνται οι προβλέψεις σε βραχυπρόθεσμες, μεσοπρόθεσμες και μακροπρόθεσμος, με γνώμονα τον επιθυμητό ορίζοντα πρόβλεψης. Στην Ενότητα 2.2 δίνονται ορισμένα βασικά χαρακτηριστικά του μεγέθους της ΟΤΣ, τα οποία σε συνδυασμό με τον ορισμό που δόθηκε στην Ενότητα 1.4, βοηθούν στην περεταίρω κατανόηση της φύσης του συγκεκριμένου μεγέθους.

Στην Ενότητα 2.3, παρουσιάζονται δύο θεμελιώδεις κατηγορίες μεθόδων πρόβλεψης, τα μοντέλα χρονοσειρών και τα αιτιοκρατικά μοντέλα, καθώς και οι διαφορές που υπάρχουν μεταξύ αυτών. Στην τελευταία Ενότητα του δεύτερου Κεφαλαίου, γίνεται αναφορά στη διαθέσιμη βιβλιογραφία και παρουσιάζονται ορισμένες διαδεδομένες μέθοδοι πρόβλεψης για κάθε περίπτωση του ορίζοντα πρόβλεψης (βραχυπρόθεσμος, μεσοπρόθεσμος, μακροπρόθεσμος).

Έχοντας αναλύσει πλήρως τη θεωρητική φύση του εξεταζόμενου προβλήματος, στο Κεφάλαιο 3, εισάγονται ορισμένες σημαντικές έννοιες που αφορούν τις χρονοσειρές. Συγκεκριμένα, στην Ενότητα 3.1, παρουσιάζονται τα ποιοτικά χαρακτηριστικά των χρονοσειρών (Τάση, Εποχικότητα κτλ.), ενώ στην Ενότητα 3.2 γίνεται λόγος για τα ποσοτικά χαρακτηριστικά των χρονοσειρών (μέση τιμή, διακύμανση κτλ.). Επιπλέον, στην Ενότητα 3.3, εισάγονται οι έννοιες ορισμένων σημαντικών στατιστικών δεικτών (σφάλματα) που χρησιμοποιούνται για την αξιολόγηση της ακρίβειας των προβλέψεων. Αξίζει να αναφερθεί πως ο στατιστικός δείκτης του Συμμετρικού Μέσου Απόλυτου Ποσοστιαίου Σφάλματος (sMAPE) είναι εκείνος που χρησιμοποιείται για την αξιολόγηση των παραγόμενων προβλέψεων μέσω της διαδικασίας προσομοίωσης που παρουσιάζεται στο Κεφάλαιο 5.

Η θεωρητική προσέγγιση του ζητήματος συνεχίζεται με την αναλυτική παρουσίαση και περιγραφή των μοντέλων πρόβλεψης στο Κεφάλαιο 4. Οι μέθοδοι πρόβλεψης που παρουσιάζονται αφορούν κυρίως μοντέλα χρονοσειρών, αιτιοκρατικά μοντέλα (μοντέλα παλινδρόμησης), καθώς και τη συνδυαστική μέθοδο Theta.

Σε πρώτη φάση, παρουσιάζεται το μοντέλο αναφοράς Naive, το οποίο αποτελεί το πιο απλό μοντέλο πρόβλεψης από σκοπιά υπολογιστικής πολυπλοκότητας. Στην Ενότητα 4.2, γίνεται λόγος για τις μεθόδους εξομάλυνσης. Συγκεκριμένα παρουσιάζονται 3 μοντέλα εξομάλυνσης, το μοντέλο της Απλής Εκθετικής Εξομάλυνσης (SES), το μοντέλο Γραμμικής Τάσης (Holt) και το μοντέλο Φθίνουσας μη Γραμμικής Τάσης (Damped).

Η ενότητα 4.3, αφορά τα μοντέλα παλινδρόμησης. Συγκεκριμένα, παρουσιάζεται το μοντέλο της απλής γραμμικής παλινδρόμησης (LRL) αλλά και της πολλαπλής γραμμικής παλινδρόμησης. Αξίζει να αναφερθεί πως η παραγωγή προβλέψεων μέσω των μοντέλων παλινδρόμησης δεν ενδιαφέρει στα πλαίσια της παρούσας εργασίας, εντούτοις προτείνεται η εφαρμογή συγκεκριμένης διαδικασίας προσομοίωσης η οποία είναι διαθέσιμη στο Παράρτημα Matlab, ενώ γίνεται ιδιαίτερη αναφορά στο συγκεκριμένο ζήτημα στο Κεφάλαιο 9.

Η τελευταία Ενότητα του τέταρτου Κεφαλαίου, αφορά τη συνδυαστική μέθοδο Theta. Συγκεκριμένα γίνεται λόγος για την κλασική μέθοδο Theta όπως αυτή παρουσιάστηκε από τους Ασημακόπουλος & Νικολόπουλος (2000). Αξίζει να αναφερθεί, πως το συγκεκριμένο μοντέλο δεν έχει εφαρμοστεί ξανά για πρόβλεψη τέτοιου είδους χρονοσειρών, επομένως πρόκειται για ένα πρωτοποριακό εγχείρημα.

Το τελευταίο βήμα για τον πλήρη καθορισμό του θεωρητικού πλαισίου στο οποίο θα κινηθεί η παρούσα εργασία, είναι η σχεδίαση κατάλληλης διαδικασίας προσομοίωσης που αφορά τα μοντέλα χρονοσειρών και βασίζεται στη λογική των κυλιόμενων προβλέψεων (rolling forecasts). Αρχικά, στην Ενότητα 5.1 εισάγεται η έννοια των κυλιόμενων προβλέψεων καθώς αναλύεται και ο λόγος για τον οποίο προτιμάται η συγκεκριμένη μεθοδολογία για την παραγωγή προβλέψεων της ΟΤΣ.

Η Ενότητα 5.2, αφορά την παρουσίαση της διαδικασίας προσομοίωσης, η οποία αποτελείται από 5 βασικά στάδια, τα οποία είναι τα εξής:

1. Συγκέντρωση και τροποποίηση δεδομένων.
2. Γραφική αναπαράσταση δεδομένων.
3. Προσδιορισμός ποιοτικών και ποσοτικών χαρακτηριστικών χρονοσειρών.
4. Εφαρμογή του rolling forecasting.
5. Αξιολόγηση των προβλέψεων.

Το 4^{ο} στάδιο, αναλύεται σε 4 υπό-στάδια τα οποία αφορούν την προεπεξεργασία, την παραγωγή προβλέψεων, τη μετεπεξεργασία και τον υπολογισμό των επιμέρους σφαλμάτων.

Όπως αναφέρθηκε και προηγουμένως, η αξιολόγηση των προβλέψεων γίνεται σύμφωνα με το στατιστικό δείκτη του Συμμετρικού Μέσου Απόλυτου Ποσοστιαίου Σφάλματος (sMAPE).

Στο Κεφάλαιο 6, επιχειρείται η ανάλυση, ποιοτική και ποσοτική, της χρονοσειράς της ΟΤΣ, η οποία αποτελεί το βασικό άξονα μελέτης της παρούσας διπλωματικής. Συγκεκριμένα, στην Ενότητα 6.1 παρουσιάζεται η χρονοσειρά των ιστορικών τιμών της ΟΤΣ, η οποία αφορά 5 συνολικά έτη (2006-2011), από τα οποία τα 3 πρώτα αποτελούν τα in-sample δεδομένα κατά τη διαδικασία της προσομοίωσης που παρουσιάστηκε στο Κεφάλαιο 5. Η ανάλυση της εν λόγω χρονοσειράς, αφορά στην γραφική αναπαράσταση των παρατηρήσεων με στόχο τον εντοπισμό των ποιοτικών χαρακτηριστικών της, ενώ σε δεύτερη φάση υπολογίζονται και ορισμένοι ποσοτικοί δείκτες, όπως παρουσιάστηκαν στο Κεφάλαιο 3, προκειμένου να υπάρξει μία πλήρως ολοκληρωμένη εικόνα της φύσης της χρονοσειράς. Αξίζει να αναφερθεί πως οι παρατηρήσεις είναι σε επίπεδο ώρας.

Ιδιαίτερη μνεία γίνεται στην παράγραφο 6.1.3, για την πολλαπλή εποχικότητα που χαρακτηρίζει το μέγεθος της ΟΤΣ. Συγκεκριμένα, παρουσιάζεται η επίδραση που έχει ο εποχικός παράγοντας (calendar effect) στο εν λόγω μέγεθος, ενώ αναλύονται οι δύο εμφανής συνιστώσες της εποχικότητας που διακρίνονται σε επίπεδο ημέρας αλλά και εβδομάδας.

Στην Ενότητα 6.2, παρουσιάζονται και αναλύονται κατά παρόμοια τρόπο με τη χρονοσειρά της ΟΤΣ, τέσσερις βοηθητικές χρονοσειρές, που στην ουσία αποτελούν εν δυνάμει ανεξάρτητες μεταβλητές παλινδρόμησης. Πρόκειται για τη χρονοσειρά των ιστορικών τιμών φορτίου, για τη χρονοσειρά της πρόβλεψης φορτίου, για τη χρονοσειρά ιστορικών τιμών Αργού Πετρελαίου και για τη χρονοσειρά ιστορικών τιμών Φυσικού Αερίου.

Αξίζει να αναφερθεί πως οι παρατηρήσει των δύο χρονοσειρών που αφορούν το φορτίο είναι σε επίπεδο ώρας, ενώ οι παρατηρήσει για τις χρονοσειρές του Αργού Πετρελαίου και του Φυσικού Αερίου είναι σε επίπεδο μήνα. Επίσης, είναι σημαντικό να αναφερθεί πως κατά την ποσοτική ανάλυση των παραπάνω βοηθητικών χρονοσειρών, υπολογίζεται και ο δείκτης γραμμικής αυτοσυσχέτισης, ο οποίος υποδεικνύει το κατά πόσο είναι η κάθε μία μεταβλητή κατάλληλη σαν ανεξάρτητη μεταβλητή παλινδρόμησης.

Στο Κεφάλαιο 7, γίνεται η παρουσίαση των αποτελεσμάτων όπως αυτά προέκυψαν μέσω της διαδικασίας προσομοίωσης κατά την εφαρμογή 6 διαφορετικών μοντέλων πρόβλεψης. Τα μοντέλα αυτά είναι:

1. Η μέθοδος αναφοράς Naive
2. Η μέθοδος SES
3. Η μέθοδος Holt
4. Η μέθοδος Damped
5. Η μέθοδος LRL
6. Η κλασική μέθοδος Theta

Κατά τη χρησιμοποίηση των τριών μοντέλων εξομάλυνσης (SES, Holt, Damped), αλλά και της κλασικής μεθόδου Theta, εφαρμόζεται αλγόριθμος βελτιστοποίησης (μη γραμμική/δυαδική βελτιστοποίηση) για τη βέλτιστη επιλογή των παραμέτρων εξομάλυνσης α, β, αλλά και για την παράμετρο διόρθωσης της τάσης φ. Τα αποτελέσματα, αφορούν 4 βασικές κατηγορίες, οι οποίες διαχωρίζονται με βάση τον τρόπο αντιμετώπισης της εποχικότητας στη χρονοσειρά της ΟΤΣ.

Η πρώτη κατηγορία αποτελεσμάτων παρουσιάζεται στην ενότητα 7.1, και αφορά αποτελέσματα δίχως οποιουδήποτε είδους επεξεργασία της εποχικότητας. Τα πρώτα αποτελέσματα υποδεικνύουν ικανοποιητική και παράλληλα παρόμοια συμπεριφορά της μεθόδου αναφοράς Naive, των μεθόδων εξομάλυνσης SES και Damped, αλλά και της συνδυαστικής μεθόδου Theta, ενώ ταυτόχρονα δείχνουν την ανικανότητα των μοντέλων Holt και LRL να προσαρμοστούν με επιτυχία στην υπό εξέταση χρονοσειρά.

Στην Ενότητα 7.2, παρουσιάζεται η δεύτερη ομάδα αποτελεσμάτων, η οποία αφορά στην αντιμετώπιση μονής εποχικότητας σε επίπεδο ημέρας, με την εξαγωγή 24 εποχικών δεικτών. Αυτή τη φορά τα αποτελέσματα είναι βελτιωμένα σε σχέση με εκείνα της πρώτης κατηγορίας, έχοντας πάντα ως κριτήριο το στατιστικό δείκτη sMAPE, ενώ η μέθοδος αναφοράς Naive δείχνει να χάνει έδαφος σε σχέση με τα τρία βέλτιστα μοντέλα SES, Damped και Theta. Αξίζει να αναφερθεί πως για τον υπολογισμό των εποχικών δεικτών λαμβάνονται υπόψη 6 διαφορετικές περιπτώσεις όγκου παρελθόντων παρατηρήσεων.

Η ενότητα 7.3 αφορά τα αποτελέσματα τα οποία προέρχονται από την επεξεργασία της μονής εποχικότητας, αυτή τη φορά σε επίπεδο εβδομάδας. Επομένως, εξάγονται 168 εποχικοί δείκτες, καθώς οι παρατηρήσεις βρίσκονται σε ωριαίο επίπεδο. Η συγκεκριμένη κατηγορία αποτελεσμάτων κυμαίνεται στα επίπεδα της προηγούμενης κατηγορίας, παρουσιάζοντας πολύ μικρές αποκλίσεις στην ακρίβεια των προβλέψεων σε σχέση με προηγουμένως. Σε αυτό το σημείο είναι σημαντικό να αναφερθεί πως για τον υπολογισμό των εποχικών δεικτών σε αυτή την κατηγορία, λαμβάνονται υπόψη όλα τα διαθέσιμα ιστορικά δεδομένα.

Η τελευταία ενότητα του Κεφαλαίου 7, αφορά στα αποτελέσματα τα οποία σχετίζονται με την αντιμετώπιση της διπλής εποχικότητα στη χρονοσειρά της ΟΤΣ. Ακολουθείται μία απλοϊκή διαδικασία προσδιορισμού δύο ομάδων εποχικών δεικτών που αντιστοιχίζονται στις δύο συνιστώσες της εποχικότητας (ημερήσια και εβδομαδιαία). Η συγκεκριμένη κατηγορία παρουσιάζει σημαντική βελτίωση στην ακρίβεια των παραγόμενων προβλέψεων, περιορίζοντας το δείκτη του sMAPE κοντά στα επίπεδα του 20%, ενώ τα τρία βέλτιστα αποτελέσματα προέρχονται από την εφαρμογή τον μοντέλων SES, Theta και Damped.

Έχοντας εξάγει το σύνολο των αποτελεσμάτων μέσω της διαδικασίας προσομοίωσης, στο Κεφάλαιο 8 κρίνεται απαραίτητη η ανάλυση των τριών βέλτιστων αποτελεσμάτων καθώς γίνεται προσπάθεια να ερμηνευτούν οι λόγοι για τους οποίους τα συγκεκριμένα αποτελέσματα αφορούν τα μοντέλα SES, Theta και Damped και προέρχονται όλα από την 4^{η} κατηγορία αντιμετώπισης της διπλής εποχικότητας.

Σε πρώτη φάση, στην Ενότητα 8.1 παρουσιάζεται η ποιοτική ανάλυση των τριών βέλτιστων αποτελεσμάτων, ενώ γίνεται λόγος και στην επιλογή των παραμέτρων εξομάλυνσης μέσω της διαδικασίας της μη γραμμικής/δυαδικής βελτιστοποίησης που εφαρμόζεται για τον προσδιορισμό τους. Στην Ενότητα 8.2, γίνεται μία ποσοτική προσέγγιση των βέλτιστων αποτελεσμάτων, καθώς αναλύεται και ο αντίκτυπος που έχει ο καθορισμός του ορίζοντα πρόβλεψης και της συχνότητας ανανέωσης των προβλέψεων στην ακρίβεια των προβλέψεων. Η ανάλυση των αποτελεσμάτων συμπίπτει με τη θεωρία των διαστημάτων εμπιστοσύνης, αποδεικνύοντας πως όσο μεγαλύτερος είναι ο επιθυμητός ορίζοντας πρόβλεψης τόσο μειώνεται και η ακρίβεια των παραγόμενων προβλέψεων.

Στην ενότητα 8.3, εξάγονται τα τελικά συμπεράσματα σχετικά με την καταλληλότητα των μοντέλων πρόβλεψης που χρησιμοποιήθηκαν κατά τη διαδικασία της προσομοίωσης, καθώς γίνεται λόγος και για την μεθοδολογία που αναπτύχθηκε για την αντιμετώπιση της διπλής εποχικότητας αλλά και για το βέλτιστο ορίζοντα πρόβλεψης, ώστε να παραμένουν αξιόπιστες οι παραγόμενες προβλέψεις.

Τέλος, το Κεφάλαιο 9, αφορά ορισμένες πτυχές του ζητήματος οι οποίες δεν εξετάσθηκαν στα πλαίσια της παρούσας εργασίας, όμως είναι ιδιαίτερα σημαντικές και μπορούν να αποτελέσουν μελλοντικό αντικείμενο έρευνας. Συγκεκριμένα, η Ενότητα 9.1 αναφέρεται στη χρησιμοποίηση μοντέλων γραμμικής πολλαπλής ή μη παλινδρόμησης, καθώς και στην επιλογή κατάλληλων ανεξάρτητων μεταβλητών με στόχο την όσο το δυνατό μεγαλύτερη βελτίωση της ακρίβειας των παραγόμενων προβλέψεων.

Η Ενότητα 9.2 έχει ως αντικείμενο την αντιμετώπιση της διπλής εποχικότητας, χρησιμοποιώντας πολυπλοκότερες μεθόδους απαλοιφής αυτής από τη χρονοσειρά της ΟΤΣ, και συγκεκριμένα γίνεται λόγος για τη μέθοδο απαλοιφής διπλής εποχικότητας μέσω εκθετικής εξομάλυνσης (DS exponential smoothing), η οποία αναπτύχθηκε από τον *Taylor* (2003).

Κεφάλαιο 1. Ελληνική Αγορά Ηλεκτρικής Ενέργειας

1.1 Εισαγωγή

Το έτος 1889 έφτασε ο «ηλεκτρισμός» στην Ελλάδα. Σύμφωνα με τα ιστορικά στοιχεία της ΔΕΗ Α.Ε., η «Γενική Εταιρεία Εργοληψιών» κατασκεύασε στην Αθήνα, στην οδό Αριστείδου, την πρώτη μονάδα παραγωγής ηλεκτρικής ενέργειας. Το πρώτο κτίριο που φωτίζεται είναι τα Ανάκτορα και πολύ σύντομα ο ηλεκτροφωτισμός επεκτείνεται στο σημερινό ιστορικό κέντρο της πόλης. Τον ίδιο χρόνο ηλεκτροδοτείται επίσης η Θεσσαλονίκη, η οποία ανήκει ακόμα στην Οθωμανική Αυτοκρατορία.

Η «Βελγική Εταιρία» αναλαμβάνει απ' τις τουρκικές αρχές το φωτισμό και την τροχοδρόμηση της πόλης με την κατασκευή εργοστασίου παραγωγής ηλεκτρικής ενέργειας. Δέκα χρόνια αργότερα κάνουν την εμφάνισή τους στην Ελλάδα οι πολυεθνικές εταιρίες ηλεκτρισμού. Η αμερικανική εταιρία Thomson-Houston με τη συμμετοχή της Εθνικής Τράπεζας ιδρύουν την «Ελληνική Ηλεκτρική Εταιρία» που αναλαμβάνει την ηλεκτροδότηση μεγάλων ελληνικών πόλεων. Μέχρι το 1929 θα έχουν ηλεκτροδοτηθεί 250 πόλεις με πληθυσμό άνω των 5.000 κατοίκων.

Στις πιο απόμακρες περιοχές, που ήταν οικονομικά ασύμφορο για τις μεγάλες εταιρίες να κατασκευάσουν μονάδες παραγωγής ηλεκτρικής ενέργειας, την ηλεκτροδότηση αναλαμβάνουν ιδιώτες ή δημοτικές και κοινοτικές αρχές κατασκευάζοντας μικρά εργοστάσια. Το έτος 1950 υπήρχαν στη Ελλάδα περίπου 400 εταιρίες παραγωγής ηλεκτρικής ενέργειας. Ως πρωτογενή καύσιμα χρησιμοποιούσαν το πετρέλαιο και το γαιάνθρακα, αμφότερα εισαγόμενα από το εξωτερικό.

Η κατάτμηση αυτή της παραγωγής σε πολλές μονάδες, σε συνδυασμό με τα εισαγόμενα καύσιμα, εξωθούσε την τιμή της ηλεκτρικής ενέργειας στα ύψη, φτάνοντας στο τριπλάσιο μέχρι και πενταπλάσιο των τιμών που ίσχυαν στις άλλες ευρωπαϊκές χώρες. Η ηλεκτρική ενέργεια ήταν λοιπόν ένα αγαθό πολυτελείας, αν και τις περισσότερες φορές παρεχόταν με ωράριο και οι ξαφνικές διακοπές ήταν σύνηθες φαινόμενο.

Όσον αφορά στην παραγωγή ηλεκτρικής ενέργειας, αυτή επιτυγχάνεται με την εκμετάλλευση διαφόρων πρωτογενών πηγών ενέργειας και παρουσιάζει μεγάλες διαφοροποιήσεις από χώρα σε χώρα, ανάλογα με τους διαθέσιμους εγχώριους Ενεργειακούς Πόρους, την Ενεργειακή Πολιτική της χώρας, τις γεωλογικές, γεωφυσικές και κλιματολογικές ιδιαιτερότητες αυτής.

Οι πηγές παραγωγής ενέργειας διακρίνονται στις συμβατικές που βασίζονται σε ορυκτά στερεά, υγρά ή αέρια καύσιμα, όπως το πετρέλαιο, ο άνθρακας (λιθάνθρακας και λιγνίτης), το φυσικό αέριο, στην πυρηνική ενέργεια και στις ανανεώσιμες πηγές ενέργειας που χρησιμοποιούν διαχρονικές πηγές (άνεμος, ήλιος, νερό κλπ) και δεν εξαντλούν τα περιορισμένα ενεργειακά αποθέματα (*ΡΑΕ, 2012*).

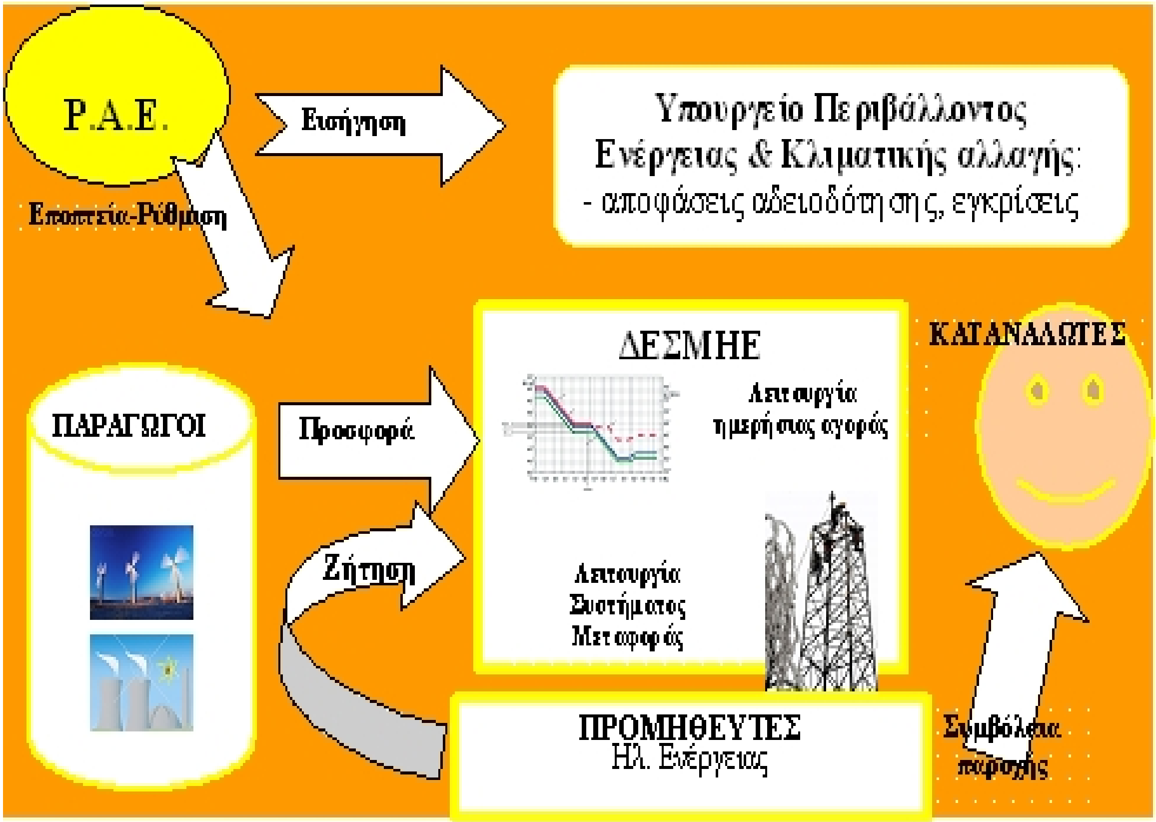
Τον Αύγουστο του 1950 ιδρύθηκε η ΔΕΗ και ως εκ τούτου, οι δραστηριότητες παραγωγής, μεταφοράς και διανομής της ηλεκτρικής ενέργειας συγκεντρώθηκαν σε ένα δημόσιο φορέα. Η ΔΕΗ αμέσως στρέφεται προς την αξιοποίηση των εγχώριων πηγών ενέργειας ενώ ξεκινά και η ενοποίηση των δικτύων μεταφοράς ηλεκτρικής ενέργειας σε ένα εθνικό διασυνδεδεμένο Σύστημα.

Τα πλούσια λιγνιτικά κοιτάσματα του ελληνικού υπεδάφους που είχαν νωρίτερα εντοπισθεί, άρχισαν να εξορύσσονται και να χρησιμοποιούνται ως καύσιμη ύλη στις λιγνιτικές μονάδες ηλεκτροπαραγωγής που δημιουργούσε η ΔΕΗ. Παράλληλα, η Επιχείρηση ξεκίνησε την αξιοποίηση της δύναμης των υδάτων με την κατασκευή υδροηλεκτρικών σταθμών στα μεγάλα ποτάμια της χώρας. Από 1.1.2001 η ΔΕΗ Α.Ε. λειτουργεί ως ανώνυμη εταιρεία ενώ από 12.12.2001 έχει εισαχθεί στα Χρηματιστήρια Αξιών Αθηνών και Λονδίνου.

Η Ρυθμιστική Αρχή Ενέργειας (ΡΑΕ) είναι υπεύθυνη για τη ρύθμιση της αγοράς ενέργειας στη χώρα μας. Στην αγορά αυτή συμμετέχουν, οι Παραγωγοί ηλεκτρικής ενέργειας που παράγουν ηλεκτρική ενέργεια από διαφορετικές πηγές, ο Διαχειριστής του Συστήματος Μεταφοράς Ηλεκτρικής Ενέργειας (ΔΕΣΜΗΕ), ο Διαχειριστής του Δικτύου Διανομής Ηλεκτρικής Ενέργειας (ΔΕΣΔΗΕ), οι Προμηθευτές ηλεκτρικής ενέργειας (όπως η ΔΕΗ Α.Ε.) και οι καταναλωτές.

Ο ΔΕΣΜΗΕ συγκροτήθηκε το 2001 και φροντίζει να υπάρχει ανά πάσα στιγμή ισορροπία παραγωγής και κατανάλωσης έτσι ώστε να καλύπτεται η ζήτηση αφού η παραγόμενη ηλεκτρική ενέργεια δεν αποθηκεύεται σε μεγάλες ποσότητες με οικονομικό τρόπο. Επιπλέον, ο ΔΕΣΜΗΕ μαζί με τον Διαχειριστή του Δικτύου Διανομής (ΔΕΣΔΗΕ), καθήκοντα του οποίου επί του παρόντος εκτελεί τμήμα της ΔΕΗ Α.Ε., είναι υπεύθυνοι να διασφαλίζουν ότι η ηλεκτρική ενέργεια μεταφέρεται με αξιοπιστία και ασφάλεια τόσο μέσω του Συστήματος (Δίκτυο υψηλής τάσης) όσο και μέσω του Δικτύου Διανομής (δίκτυο μέσης και χαμηλής τάσης) από τους ηλεκτροπαραγωγούς σταθμούς στους τελικούς καταναλωτές. Οι Προμηθευτές είναι υπεύθυνοι για την πώληση της ηλεκτρικής ενέργειας στους τελικούς καταναλωτές.

Η ενέργεια που παράγεται στους σταθμούς παραγωγής, αρχικά μεταφέρεται σε κοντινούς μετασχηματιστές που μετατρέπουν την χαμηλή τάση της ηλεκτρικής ενέργειας σε υψηλή. Με αυτό τον τρόπο, η ηλεκτρική ενέργεια μεταφέρεται με τις γραμμές μεταφοράς σε πολύ μεγάλες αποστάσεις με λιγότερες απώλειες, καθώς οι σταθμοί παραγωγής είναι συνήθως μακριά από μεγάλα αστικά κέντρα. Στο ακόλουθο σχήμα, αναπαρίσταται ο μηχανισμός λειτουργίας της Ελληνική Αγοράς Ηλεκτρικής Ενέργειας:



Σχ. 1.1.1 – Αναπαράσταση Ελληνικής Αγοράς Ηλεκτρικής Ενέργειας

Το δίκτυο μεταφοράς, οδηγεί την ηλεκτρική ενέργεια στους υποσταθμούς μέσης και χαμηλής τάσης, στους οποίους μετατρέπεται η τάση της ηλεκτρικής ενέργειας από υψηλή σε μέση και χαμηλή τάση, προκειμένου με τη βοήθεια εναέριων γραμμών να διανεμηθεί σε βιομηχανίες που χρησιμοποιούν μέση τάση και σε σπίτια που χρησιμοποιούν χαμηλή τάση.

Τα συστατικά στοιχεία των γραμμών μεταφοράς είναι τα ακόλουθα:

- Πυλώνες ή πύργοι, στους οποίους στηρίζονται οι αγωγοί των εναέριων γραμμών.
- Μονωτήρες, μέσω των οποίων αναρτώνται στους πυλώνες οι αγωγοί γραμμών.
- Αγωγοί, κυρίως από χαλκό και αλουμίνιο.

Το δίκτυο Διανομής, περιλαμβάνει:

- το δίκτυο διανομής μέσης τάσης (20kV) που μεταφέρει την ηλεκτρική ισχύ από τους υποσταθμούς μεταφοράς στους υποσταθμούς διανομής,

- το δίκτυο διανομής χαμηλής τάσης (220/380V) που μεταφέρει την ηλεκτρική ισχύ από τους υποσταθμούς διανομής στους καταναλωτές.

1.2 Περιγραφή και βασικά χαρακτηριστικά

Όταν αναφερόμαστε σε μία αγορά είναι απαραίτητο να εξετάσουμε τις συνθήκες υπό τις οποίες λειτουργεί. Στην περίπτωση της Αγοράς Ενέργειας μπορούμε να διακρίνουμε δύο βασικές κατηγορίες:

1. Απελευθερωμένη αγορά (ολιγοπώλιο).

2. Μη απελευθερωμένη αγορά (μονοπώλιο).

Πιο συγκεκριμένα, όταν κάνουμε λόγο για απελευθερωμένη αγορά σημαίνει ότι οι ποσότητες και οι τιμές καθορίζονται από τη δράση μεταξύ των παραγωγών και των καταναλωτών, και στις περισσότερες περιπτώσεις αναφερόμαστε σε ολιγοπώλιο (*Vespucci et al., 2009*).

Στην μη απελευθερωμένη ή μονοπωλιακή αγορά, το αγαθό προσφέρεται από μία μόνο επιχείρηση, δεν είναι άμεσα υποκαταστάσιμο από άλλο παρεμφερές αγαθό, και παράλληλα οι καταναλωτές είναι πολυάριθμοι. Η τιμή και η ποσότητα προσφοράς του αγαθού καθορίζονται μόνο από τις αποφάσεις που παίρνει η μονοπωλιακή επιχείρηση (*Κάπρος, 2003*).

Ήδη πολλοί Ευρωπαίοι καταναλωτές απολαμβάνουν στην καθημερινότητά τους το δικαίωμα επιλογής Προμηθευτή. Όμως στις χώρες τις Ευρώπης που δεν έχουν ενεργοποιηθεί ιδιαίτερα άλλοι Προμηθευτές ενέργειας πέραν του αρχικού, όπως στην Ελλάδα, το δικαίωμα αυτό δεν έχει ασκηθεί για τη πλειοψηφία των καταναλωτών.

Το πρώτο βήμα προς το χτίσιμο των απαραίτητων προϋποθέσεων για τη λειτουργία της απελευθερωμένης αγοράς ήταν η λήψη μέτρων σε θεσμικό επίπεδο, προκειμένου να εξαλειφθούν οι στρεβλώσεις που επικράτησαν κατά την εικοσαετία 1980-2000, όπου προτεραιότητα δόθηκε στην ασφάλεια του ενεργειακού εφοδιασμού και στη διατήρηση χαμηλών τιμολογίων.

Ωστόσο είναι γεγονός, ότι η διαδικασία απελευθέρωσης της ηλεκτρικής ενέργειας στην Ελλάδα η οποία ξεκίνησε το έτος 2001 και βρίσκεται ακόμη σε εξέλιξη προχωράει με αργά αλλά σταθερά βήματα και για το λόγο αυτό η Ελληνική Αγορά Ενέργειας μπορεί να χαρακτηριστεί ως αγορά υπό απελευθέρωση. Η δημιουργία κοινών κανόνων για την αγορά ηλεκτρισμού των χωρών της Ν.Α. Ευρώπης και οι σύγχρονες τάσεις για την ενοποίηση της Ευρωπαϊκής αγοράς ηλεκτρικής ενέργειας αποτελούν τις σημερινές προκλήσεις.

Διαχειριστής της Ελληνικής Αγοράς Ηλεκτρικής Ενέργειας είναι ο ΔΕΣΜΗΕ (Διαχειριστής Ελληνικού Συστήματος Μεταφοράς Ηλεκτρικής Ενέργειας) ο οποίος ιδρύθηκε με το Προεδρικό Διάταγμα 328/12-12-2000 ως ανώνυμη εταιρία, αντικείμενο της οποίας είναι η λειτουργία, η εκμετάλλευση, η διασφάλιση της συντήρησης και η μέριμνα για την ανάπτυξη του Συστήματος Μεταφοράς ηλεκτρικής Ενέργειας σε ολόκληρη τη χώρα, καθώς και των διασυνδέσεων του με τα άλλα δίκτυα για να διασφαλίζεται ο εφοδιασμός της χώρας με ηλεκτρική ενέργεια, κατά τρόπο επαρκή, ασφαλή, οικονομικά αποδοτικό και αξιόπιστο.

Η παρακολούθηση και η τήρηση των κανόνων (Κώδικας) που διέπουν την Ελληνική Αγορά Ενέργειας υπάγονται στις υποχρεώσεις της Ρυθμιστικής Αρχής Ενέργειας (ΡΑΕ). Πιο αναλυτικά, η ΡΑΕ συγκροτήθηκε τον Ιούλιο του 2000 και πρόκειται για μία ανεξάρτητη διοικητική αρχή, η οποία έχει ως βασικό αντικείμενο την παρακολούθηση της ανάπτυξης της Αγοράς Ενέργειας, τόσο στην Ελληνική Αγορά, αλλά και στον τρόπο με τον οποίο αυτή λειτουργεί και αναπτύσσεται σε σχέση με τις ξένες αγορές και ιδίως με εκείνες τις οποίες διασυνδέεται.

Ο Κώδικας εφαρμόστηκε προοδευτικά κατά τη διάρκεια μιας μεταβατικής περιόδου που εκτείνεται από το 2005 έως τα τέλη του 2010 και περιλαμβάνει 5 «ημέρες αναφοράς» οι οποίες χαρακτηρίζουν την πρόοδο που έγινε κατά την προαναφερθείσα περίοδο και τους νέους στόχους που τέθηκαν για το μέλλον.

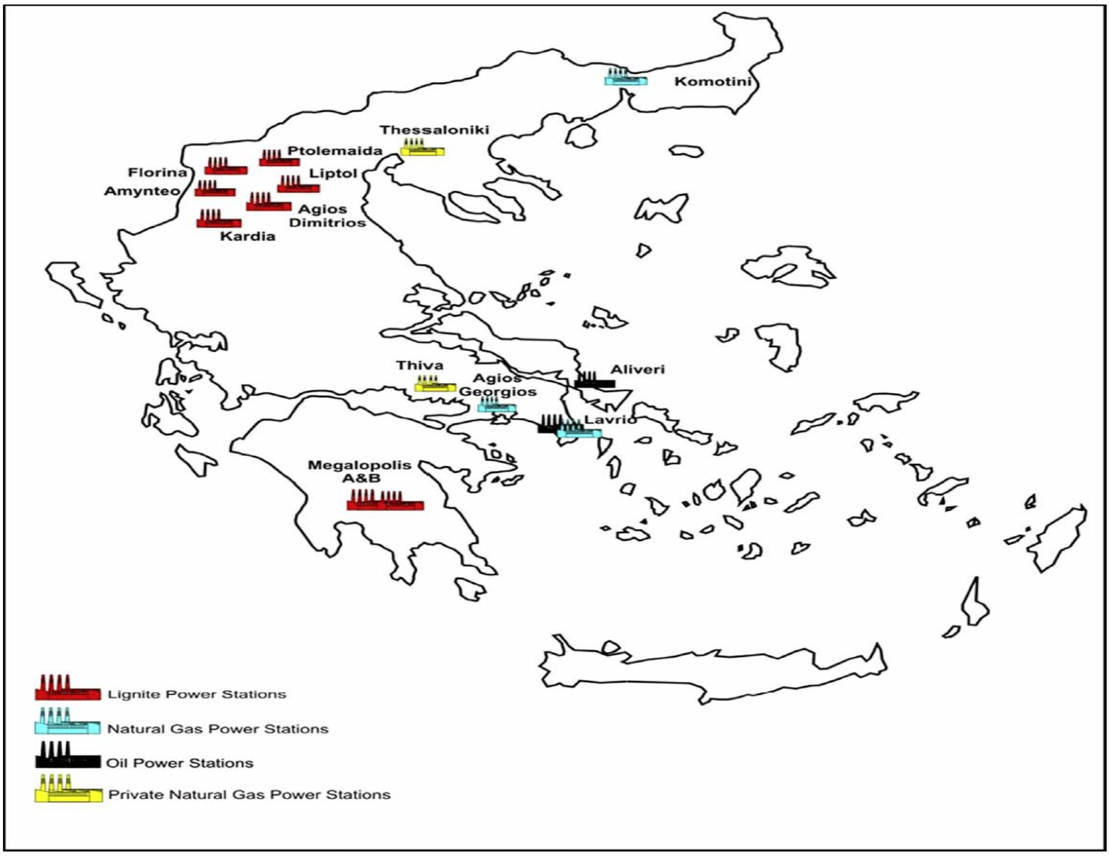
Αξίζει να αναφερθεί πως η ΡΑΕ παίζει κατά κάποιο τρόπο το ρόλο του συμβούλου του Ελληνικού Υπουργείου Ανάπτυξης το οποίο είναι αυτό που παίρνει τις τελικές αποφάσεις και καθορίζει τον τρόπο λειτουργίας της Ελληνικής Αγοράς Ενέργειας.

Ιδιαίτερα σημαντικό ρόλο στην Ελληνική Αγορά Ενέργειας έχει η ΔΕΗ Α.Ε. Δραστηριοποιείται ως Παραγωγός και είναι ο κύριος Προμηθευτής ηλεκτρικής ενέργειας, όπου σύμφωνα με τα στοιχεία του 2009, κατέχει περίπου το 89% της εγκατεστημένης ηλεκτρικής ισχύος στην Ελλάδα (12.695 MW) η οποία προέρχεται από λιγνιτικές, πετρελαϊκές, υδροηλεκτρικές μονάδες, μονάδες φυσικού αερίου καθώς και από αιολικά και ηλιακά πάρκα. Παράγοντας σχεδόν το 55% της ηλεκτρικής της παραγωγής από λιγνίτη, είναι ο 2ος μεγαλύτερος παραγωγός ηλεκτρικής ενέργειας από λιγνίτη στην Ευρωπαϊκή Ένωση.

Επιπλέον, η ΔΕΗ Α.Ε. ως Κύριος του Συστήματος μεταφοράς υποχρεούται στη συντήρηση του Συστήματος στο πλαίσιο του προγραμματισμού που εκπονεί η ίδια και υπό τις οδηγίες του ΔΕΣΜΗΕ και παράλληλα εκτελεί χρέη Διαχειριστή του Δικτύου Διανομής μέχρι την ίδρυση του ΔΕΣΔΗΕ, με σημαντικές αρμοδιότητες για τη σύνδεση των νέων χρηστών, την καθημερινή λειτουργία και συντήρηση του Δικτύου, καθώς και για την ανάπτυξή του.

Βασικό χαρακτηριστικό του Ελληνικού συστήματος είναι η γεωγραφική ανισοκατανομή μεταξύ παραγωγής και φορτίων λόγω της μεγάλης συγκέντρωσης σταθμών παραγωγής και της ύπαρξης σημαντικών διεθνών διασυνδέσεων στο Βόρειο τμήμα της χώρας, ενώ το κέντρο κατανάλωσης είναι στο Νότιο τμήμα και ιδιαίτερα στην περιοχή της Αττικής.

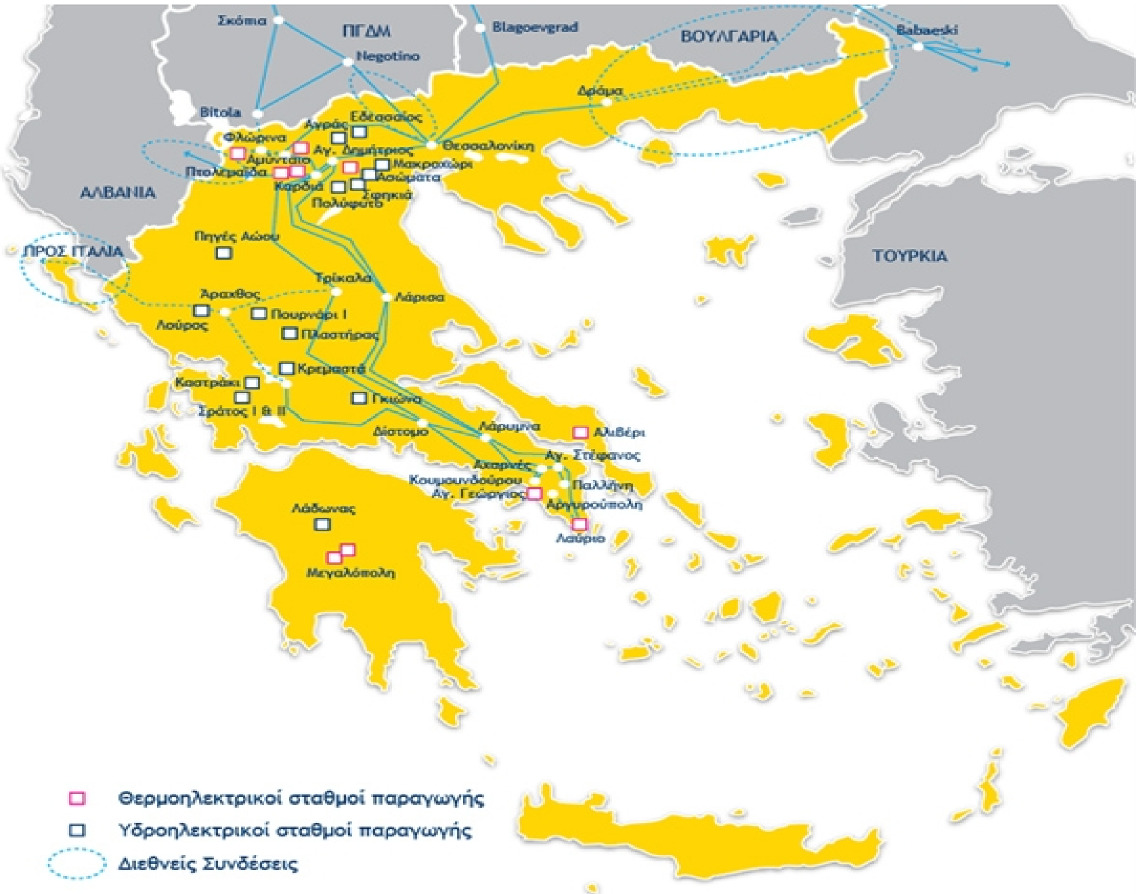
Επίσης η Πελοπόννησος δε συνδέεται στο βασικό κορμό διακίνησης ενέργειας στην ηπειρωτική χώρα (400 kV) (*ΤΕΕ, 2010*). Η γεωγραφική κατανομή των κυριότερων σταθμών παραγωγής ηλεκτρικής ενέργειας παρουσιάζονται στο ακόλουθο σχήμα(ΡΑΕ, 2012):



Σχ. 1.2.1 - Σταθμοί Παραγωγής Ηλεκτρικής Ενέργειας

Όσον αφορά τις συνδέσεις της Ελληνικής Αγοράς με το εξωτερικό, αυτές βρίσκονται στο βόρειο τμήμα της χώρας. Πιο αναλυτικά, η Ελλάδα συνδέεται στα βόρεια μέσω εναέριων γραμμών μεταφοράς (AC) με την Αλβανία, το FYROM και τη Βουλγαρία, ενώ στα δυτικά συνδέεται με την Ιταλία μέσω καλωδίου (HVDC). Από το Σεπτέμβρη του 2010 εγκαινιάστηκε νέα εναέρια γραμμή μεταφοράς με την Τουρκία. Ιδιαιτερότητες παρουσιάζουν τα Ελληνικά νησιά του Αιγαίου, τα οποία στην πλειοψηφία τους δεν είναι συνδεδεμένα με την κυρίως χώρα αλλά διαθέτουν αυτόνομους σταθμούς παραγωγής βασισμένους στο πετρέλαιο ,σε συνδυασμό με ανεμογεννήτριες και άλλων μορφών ΑΠΕ (*Andrianesis, Biskas & Liberopoulos, 2011*).

Ο χάρτης του ελληνικού διασυνδεδεμένου συστήματος μεταφοράς ηλεκτρικής ενέργειας παρουσιάζεται στο ακόλουθο σχήμα:

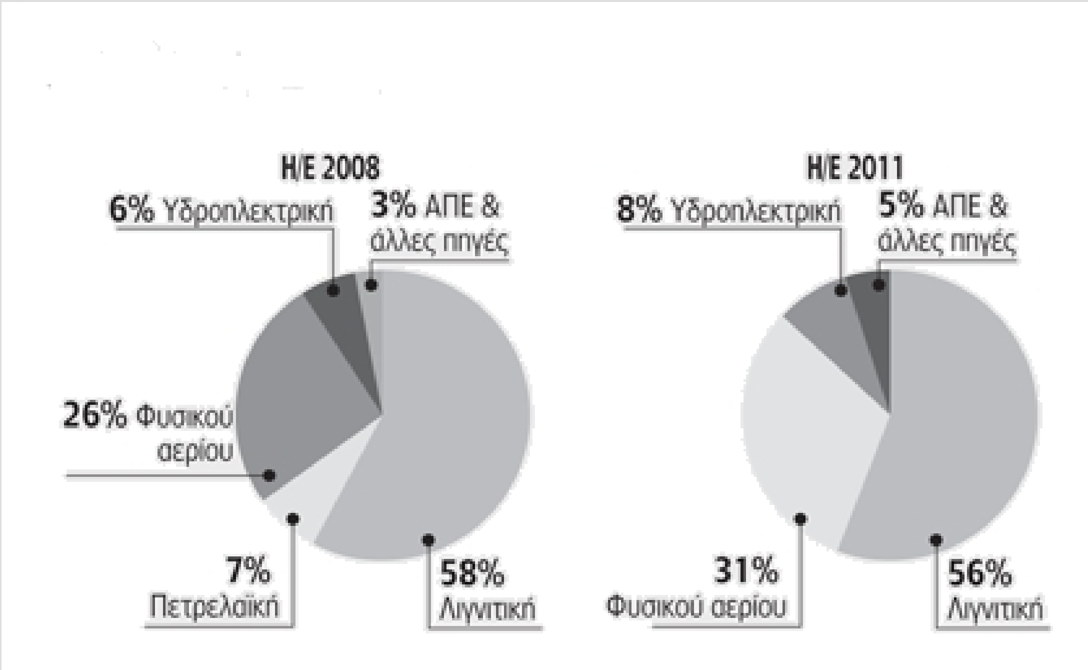


Σχ. 1.2.2 - Χάρτης Διασυνδεδεμένου Συστήματος Μεταφοράς Η.Ε.

Πηγή: http://www.investingreece.gov.gr

Το καύσιμο που χρησιμοποιείται ευρέως στην Ελληνική Αγορά Ενέργειας είναι ο λιγνίτης. Πέρα από τις μονάδες που είναι βασισμένες στο λιγνίτη, συναντάμε και μονάδες αργού πετρελαίου, μονάδες συνδυασμένου κύκλου και ανοιχτού κύκλου με φυσικό αέριο, καθώς και μονάδες ΑΠΕ, όπως μεγάλες υδροηλεκτρικές μονάδες, αιολικά πάρκα και φωτοβολταϊκές εγκαταστάσεις, μικρές υδροηλεκτρικές μονάδες, μονάδες βιομάζας και γεωθερμικές μονάδες.

Η διαμόρφωση του ενεργειακού μίγματος της χώρας (ισοζύγιο παραγωγής ηλεκτρικής ενέργειας) κατά την περίοδο 2008-2011 παρουσιάζεται στο γράφημα του ακόλουθου σχήματος:



Σχ. 1.2.3 - Ενεργειακό Μίγμα Ελληνικής Επικράτειας

Πηγή: http://www.statbank.gr

Κυρίαρχο ζητούμενο για την εύρυθμη λειτουργία της αγοράς ηλεκτρισμού στην Ελληνική επικράτεια είναι η επίτευξη των στόχων της ασφάλειας εφοδιασμού και της οικονομικής βελτιστοποίησης της αγοράς.

Αυτό επιτυγχάνεται μέσω κατάλληλου συνδυασμού μακροχρόνιων(εγκατάσταση και διαθεσιμότητα ισχύος) καθώς και βραχυχρόνιων αποφάσεων για την ορθή διαχείριση και κατανομή των πόρων στον Ημερήσιο Ενεργειακό Προγραμματισμό.

Σύμφωνα με το παραπάνω πλαίσιο (ΔΕΣΜΗΕ – Εγχειρίδιο Αγοράς, 2010) η Ελληνική Αγορά Ηλεκτρικής Ενέργειας διακρίνεται σε δύο θεμελιώδεις κατηγορίες οι οποίες είναι οι ακόλουθες:

- Η μακροχρόνια αγορά διαθεσιμότητας ισχύος (Capacity Market).
- Η βραχυχρόνια – χονδρεμπορική αγορά ενέργειας και επικουρικών υπηρεσιών, η οποία προγραμματίζεται σύμφωνα

με τον Ημερήσιο Ενεργειακό Προγραμματισμό (ΗΕΠ) (Energy and Ancillary Services Market).

1.2.1 Αγορά Μακροχρόνιας Διαθεσιμότητας Ισχύος

Η λειτουργία αυτής της αγοράς έχει ως στόχο την όσο το δυνατόν μεγαλύτερη μείωση του επιχειρηματικού κινδύνου του παραγωγού, ο οποίος λαμβάνει αμοιβή έναντι μέρους του κόστους κεφαλαίου επένδυσης, καθώς και του προμηθευτή που επιζητά την αποφυγή ιδιαίτερα υψηλών τιμών στην Ημερήσια Αγορά Ενέργειας και Επικουρικών Υπηρεσιών(βραχυχρόνια αγορά), για το λόγο ότι μειώνεται ο βραχυχρόνιος κίνδυνος του παραγωγού.

Βασικός λόγος που οδήγησε στη δημιουργία της συγκεκριμένης αγοράς είναι η ανάγκη για εξασφάλιση της επάρκειας καθώς και της ποιότητας της ηλεκτρικής ενέργειας σε μακροχρόνια βάση, αμείβοντας παράλληλα την αξιοπιστία κάθε μονάδας.

Η υλοποίηση της Αγοράς Μακροχρόνιας Διαθεσιμότητας Ισχύος επιτυγχάνεται με την έκδοση από κάθε παραγωγό που συμμετέχει σε αυτήν Αποδεικτικών Διαθεσιμότητας Ισχύος (ΑΔΙ) που αντιστοιχούν στην πραγματική διαθεσιμότητα ισχύος κάθε μονάδας του, η οποία προσδιορίζεται από το Διαχειριστή του Συστήματος. Αν για παράδειγμα διαθέτουμε μία μονάδα εγκατεστημένης ισχύος 300 MW, είναι δυνατό μετά τον προσδιορισμό της διαθεσιμότητας της από τον Διαχειριστή του Συστήματος να εκδώσει ΑΔΙ για 250 MW.

Όσον αφορά τους προμηθευτές, κάθε ένας από αυτούς συνάπτει με τους παραγωγούς Συμβάσεις Διαθεσιμότητας Ισχύος (ΣΔΙ), με οικονομικούς όρους που συμφωνούνται και από τις δύο πλευρές έτσι ώστε να καλύψουν την Υποχρέωση Επάρκειας Ισχύος που τους αναλογεί.

Η αγορά μπορεί να χαρακτηριστεί ως εξισορροπημένη (Balanced) (*ΔΕΣΜΗΕ – Εγχειρίδιο Αγοράς, 2010*) όταν ισχύει η ακόλουθη σχέση:

ΑΔΙ (MW) =ΣΔΙ (MW) + Απαιτούμενη μακροχρόνια εφεδρεία

1.2.2 Ημερήσιος Ενεργειακός Προγραμματισμός (Αγορά Ενέργειας και Επικουρικών Υπηρεσιών ή Χονδρεμπορική Αγορά)

Ο Ημερήσιος Ενεργειακός Προγραμματισμός (ΗΕΠ) αποτελεί τη χονδρεμπορική αγορά και έχει ως κυρίαρχο στόχο το βέλτιστο προγραμματισμό της λειτουργίας των θερμικών και υδροηλεκτρικών μονάδων παραγωγής του Συστήματος, όπως και των μονάδων ΑΠΕ και της διαθέσιμης ενέργειας που προέρχεται από εισαγωγές, ώστε να καλύπτεται σε ημερήσια βάση η ζήτηση ηλεκτρικής ενέργειας από καταναλωτές, η ζήτηση για εξαγωγές ενέργειας και οι απαραίτητες Επικουρικές Υπηρεσίες. Οι μονάδες παραγωγής που συμμετέχουν στον ΗΕΠ οφείλουν να προσφέρουν το σύνολο της διαθεσιμότητάς τους, τόσο σε ενέργεια όσο και σε επικουρικές υπηρεσίες. Ουσιαστικά πρόκειται για ένα μοντέλο αγοράς Υποχρεωτικής Κοινοπραξίας (Mandatory Pool).

Για να γίνει κατανοητός ο όρος «Υποχρεωτική Κοινοπραξία» απαιτείται η παράθεση του ορισμού του:

«Υποχρεωτική Κοινοπραξία» είναι το μοντέλο της Ελληνικής Αγοράς στο οποίο συναλλάσσεται το σύνολο της ηλεκτρικής ενέργειας και των συμπληρωματικών προϊόντων αυτής που θα παραχθούν, θα καταναλωθούν και θα διακινηθούν την επόμενη ημέρα στην αγορά. Όλοι οι συμμετέχοντες στην Ελληνική Αγορά υποχρεούνται να συμμετέχουν στην Υποχρεωτική Κοινοπραξία. Δεν επιτρέπονται φυσικές διμερής συναλλαγές (physical bilateral transactions) μεταξύ των συμμετεχόντων της αγοράς (*ΔΕΣΜΗΕ – Εγχειρίδιο Αγοράς έκδοση 2.0, 2010*).

Στα αυστηρά πλαίσια του παραπάνω μοντέλου κάθε διαμετακόμιση (transit) ηλεκτρικής ενέργειας από μία περιοχή εκτός Ελλάδος σε μία άλλη ,μέσω της Ελληνικής επικράτειας, μπορεί να υλοποιηθεί μόνο μέσω εισαγωγής (πώλησης) της εν λόγω ενέργειας στην Υποχρεωτική Κοινοπραξία και εξαγωγής (αγοράς) από αυτήν.

Σημαντικές επιμέρους αγορές-μηχανισμοί οι οποίες υπάγονται στον ΗΕΠ είναι οι ακόλουθες:

i. Αγορά Ενέργειας: Ικανοποιούνται οι ποσοτικές ανάγκες των καταναλωτών ηλεκτρικής ενέργειας κατά τις ώρες που αυτή είναι απαραίτητη. Προσφέρουν αλλά και αμείβονται οι εγχώριοι παραγωγοί (θερμικών – υδροηλεκτρικών σταθμών και ΑΠΕ) καθώς και οι εισαγωγείς ενέργειας ενώ αγοράζουν οι εκπρόσωποι του εγχώριου φορτίου (προμηθευτές και επιλεγμένοι πελάτες) και οι εξαγωγείς (προμηθευτές και παραγωγοί).

ii. Αγορά Επικουρικών Υπηρεσιών (Ancillary Services): Ικανοποιεί τις ανάγκες των καταναλωτών ηλεκτρικής ενέργειας για διασφάλιση της ποιότητας και αξιοπιστίας της τροφοδότησής τους. Παράλληλα προσφέρουν και αμείβονται οι εγχώριοι παραγωγοί (θερμικών – υδροηλεκτρικών σταθμών) και αγοράζουν οι εκπρόσωποι του εγχώριου φορτίου (προμηθευτές και επιλεγμένοι πελάτες) καθώς και οι εξαγωγείς (προμηθευτές και παραγωγοί).

iii. Μηχανισμός αγοράς για τη χωροθέτηση της παραγωγής κοντά στα κέντρα κατανάλωσης: Σχετίζεται με την παροχή κινήτρων χωροθέτησης των νέων μονάδων όσο τον δυνατόν πιο κοντά στην κατανάλωση . Με δεδομένη την παρούσα συγκέντρωση της ηλεκτρικής παραγωγής στο βορρά, είναι ασφαλές να πούμε ότι αναμένεται αυξημένη συμμετοχή στην ετήσια χρέωση χρήσεως συστήματος των παραγωγών του βορρά, όπως και διαφοροποιημένη αμοιβή παραγωγών που βρίσκονται στο νότιο τμήμα της χώρας ,κυρίως σε περιπτώσεις που εμφανίζεται συμφόρηση στη μεταφορά ηλεκτρικής ενέργειας από το βορρά στο νότο κατά τη διάρκεια του ημερήσιου προγραμματισμού.

1.3 ΠΕΡΙΓΡΑΦΗ ΚΑΙ ΕΠΙΛΥΣΗ ΤΟΥ ΗΕΠ

Η επίλυση του ΗΕΠ αφορά κυρίως τον προσδιορισμού του τρόπου λειτουργίας κάθε μονάδας για κάθε ώρα της επόμενης ημέρας, με απώτερο στόχο τη μεγιστοποίηση του κοινωνικού οφέλους που προκύπτει από την ικανοποίηση του ενεργειακού ισοζυγίου της αγοράς αλλά και των αναγκών των επικουρικών υπηρεσιών την επόμενη μέρα λαμβάνοντας πάντα υπόψη τους περιορισμούς του συστήματος μεταφοράς.

Ενσωματώνοντας στον ΗΕΠ τις απαραίτητες Επικουρικές Υπηρεσίες και τους τεχνικούς περιορισμούς του συστήματος, οι οποίοι πιθανότατα περιορίζουν την ποσότητα ενέργειας που είναι δυνατόν να διακινείται από το βορρά στο νότο, προκύπτει η πραγματική συνολική αξία της καταναλισκόμενης ηλεκτρικής ενέργειας, έχοντας λάβει υπόψη τα ποσοτικά (όγκος και χρόνος κατανάλωσης) αλλά και τα ποιοτικά χαρακτηριστικά της (συχνότητα, τάση , απαιτούμενη αξιοπιστία παροχής σε κάθε καταναλωτή).

Από την επίλυση του ΗΕΠ προκύπτει η **Οριακή Τιμή Συστήματος** (ΟΤΣ) η οποία αποτελεί το αντικείμενο μελέτης της παρούσας διπλωματικής εργασίας και θα αναλυθεί εκτενέστερα στη συνέχεια.

Όπως έχει αναφερθεί και προηγουμένως, ο Ημερήσιος Ενεργειακός Προγραμματισμός (ΗΕΠ) είναι η ημερήσια διαδικασία επίλυσης της αγοράς ενέργειας η οποία πραγματοποιείται την προηγούμενη ημέρα της Ημέρας Κατανομής. Ημέρα Κατανομής θεωρείται μία ημερολογιακή ημέρα σε ώρες Ελλάδος. Αποτελείται από 24 ωριαίες περιόδους κατανομής, εκτός φυσικά από τις «μικρές» και «μεγάλες» ημέρες που παρουσιάζονται το φθινόπωρο και την άνοιξη αντίστοιχα και οφείλονται στην αλλαγή της τοπικής ώρας. Επομένως η Μικρή Ημέρα Κατανομής αποτελείται από 23 περιόδους κατανομής (απουσιάζει η ώρα από τις 3:00 ως τις 4:00), και η Μεγάλη Ημέρα Κατανομής αποτελείται από 25 περιόδους Κατανομής (εμφανίζεται 2 φορές η ώρα από τις 3:00 ως τις 4:00).

1.3.1 Συμμετοχή στον ΗΕΠ

Όλοι όσοι συμμετέχουν στον ΗΕΠ αλλά και στις άλλες αγορές που διαχειρίζεται ο Διαχειριστής του Συστήματος, είναι απαραίτητο να εγγραφούν στο Μητρώο Συμμετεχόντων του Διαχειριστή. Μέσω αυτής της εγγραφής οι συμμετέχοντες συνάπτουν με το Διαχειριστή του Συστήματος «Σύμβαση Συναλλαγών Ηλεκτρικής Ενέργειας» η οποία διέπεται από τις διατάξεις του ΚΔΣ&ΣΗΕ (Κώδικας Διαχείρισης Συστήματος και Συναλλαγών Ηλεκτρικής Ενέργειας).

Η Σύμβαση Συναλλαγών Ηλεκτρικής Ενέργειας αφορά τις συναλλαγές που πραγματοποιούνται στα πλαίσια του συστήματος Συναλλαγών Ηλεκτρικής Ενέργειας, το οποίο περιλαμβάνει τον ΗΕΠ, τη Διαδικασία Κατανομής, τις Εντολές Κατανομής, την Εκκαθάριση Αποκλίσεων και τον Μηχανισμό Διασφάλισης Επαρκούς Ισχύος. Επιπλέον, οι συμμετέχοντες έχουν το δικαίωμα να προβαίνουν σε κάθε σχετική νόμιμη ενέργεια συμμετοχής στο Σύστημα Συναλλαγών Ηλεκτρικής Ενέργειας, τηρώντας πάντα τις διατάξεις του Κώδικα, και παράλληλα μπορούν να λαμβάνουν και να εξοφλούν τις πληρωμές και τις χρεώσεις που τους αναλογούν. Τα δικαιώματα και οι υποχρεώσεις του Διαχειριστή του Συστήματος ορίζονται από τις διατάξεις του Κώδικα για το Σύστημα Συναλλαγών Ηλεκτρικής Ενέργειας.

Πιο αναλυτικά, η συμμετοχή στον ΗΕΠ περιλαμβάνει τις ακόλουθες διαδικασίες (*ΔΕΣΜΗΕ – Εγχειρίδιο Αγοράς έκδοση 2.0, 2010*):

- Υποβολή μη τιμολογούμενων Δηλώσεων Φορτίου για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους Εκπροσώπους Φορτίου για τους Πελάτες που καταναλώνουν ενέργεια εντός της Ελληνικής Επικράτειας.

- Υποβολή τιμολογούμενων Δηλώσεων Φορτίου για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους Συμμετέχοντες, όταν αυτοί είναι είτε κάτοχοι Φυσικών Δικαιωμάτων Μεταφοράς (ΦΔΜ) είτε αντισυμβαλλόμενοι κατόχων ΦΔΜ, για εξαγωγές ενέργειας.

- Υποβολή τιμολογούμενων Δηλώσεων Φορτίου για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από Παραγωγούς για αντλητικές μονάδες σε λειτουργία άντλησης
  .
- Υποβολή μη τιμολογούμενων Δηλώσεων Φορτίου για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους Παραγωγούς και τους Προμηθευτές για τα βοηθητικά φορτία των μονάδων όταν αυτά δεν καλύπτονται από την παραγωγή των μονάδων.

- Υποβολή τιμολογούμενων Προσφορών Έγχυσης για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από Παραγωγούς για Κατανεμόμενες Μονάδες.

- Υποβολή τιμολογούμενων Προσφορών Έγχυσης για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από το Διαχειριστή του Συστήματος για τις Συμβεβλημένες Μονάδες. Οι τιμολογούμενες προσφορές ενέργειας για τις Συμβεβλημένες Μονάδες πρέπει να είναι σύμφωνες με τις τιμές της σχετικής σύμβασης.

- Υποβολή τιμολογούμενων Προσφορών Έγχυσης για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους συμμετέχοντες, όταν αυτοί είναι είτε κάτοχοι Φυσικών Δικαιωμάτων Μεταφοράς (ΦΔΜ) είτε αντισυμβαλλόμενοι κατόχων ΦΔΜ, για εισαγωγές ενέργειας.

- Υποβολή μη τιμολογούμενων Προσφορών Έγχυσης για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από το Διαχειριστή του Συστήματος για τις μονάδες του Άρθρου 35 του Νόμου 2773/1999[2].

- Υποβολή Δηλώσεων Διαχείρισης Υδάτινων Πόρων για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους Παραγωγούς για τις Υδροηλεκτρικές Μονάδες, συμπεριλαμβανομένων των Υδροηλεκτρικών Μονάδων Άντλησης.

- Υποβολή τιμολογούμενων Προσφορών Εφεδρειών για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από Παραγωγούς για τις Κατανεμόμενες Μονάδες.

- Υποβολή τιμολογούμενων Προσφορών Εφεδρειών για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από το Διαχειριστή τους Συστήματος για τις Συμβεβλημένες Μονάδες. Οι τιμολογούμενες προσφορές εφεδρειών για τις Συμβεβλημένες Μονάδες πρέπει να είναι σύμφωνες με τις τιμές της σχετικής σύμβασης.

- Υποβολή Δηλώσεων Ολική ή Μερικής Μη Διαθεσιμότητας για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής από τους Παραγωγούς για κάθε μονάδα.

- Υποβολή δηλώσεων Τεχνοοικονομικών Στοιχείων από τους Παραγωγούς για τις μονάδες.

1.3.2 Οι 5 Ημέρες Αναφοράς

Ακρογωνιαίο λίθο για τη διαδικασία επίλυσης του ΗΕΠ στην Ελλάδα αποτελούν όπως αναφέρθηκε και προηγουμένως οι 5 «Ημέρες Αναφοράς» οι οποίες ουσιαστικά αποτελούν τους σταθμούς για την ομαλή μετάβαση της Ελληνικής Αγοράς Ενέργειας από τη μονοπωλιακή στην ολιγοπωλιακή αγορά. Όπως αναφέρει και ο *Α.Μπακιρτζής (2009)* οι 5 «Ημέρες Αναφοράς» περιείχαν τα παρακάτω:

- <u>Έως 12 Ιανουαρίου 2006</u>
    - Περιορισμοί τεχνικών ελαχίστων των μονάδων στον ΗΕΠ.
    - Μια προσφορά ανά μονάδα για όλη την Ημέρα Κατανομής.
    - Επίλυση ΗΕΠ ανά Περίοδο Κατανομής (ώρα – ώρα).

- <u>Από 13 Ιανουαρίου 2006</u>
    - Κατάργηση περιορισμών τεχνικών ελαχίστων στον ΗΕΠ με συνέπεια την αύξηση της ΟΤΣ.
    - Μία προσφορά ανά μονάδα για όλη την Ημέρα Κατανομής.
    - Επίλυση ΗΕΠ ανά Περίοδο Κατανομής (ώρα – ώρα).

- <u>Από 1 Απριλίου 2007</u>
    - Χωρίς περιορισμούς τεχνικών ελαχίστων στον ΗΕΠ με συνέπεια την αύξηση της ΟΤΣ.
    - Ωριαίες προσφορές έγχυσης μονάδων παραγωγής.
    - Δυνατότητα υποβολής προσφορών κάτω του μεταβλητού κόστους για το 30% της διαθέσιμης ισχύος με συνέπεια τη μείωση της ΟΤΣ.
    - Επίλυση ΗΕΠ ανά Περίοδο Κατανομής (ώρα – ώρα).

- <u>Από 1 Ιανουαρίου 2009 (Τέταρτη Ημέρα Αναφοράς)</u>
    - 24ωρη βελτιστοποίηση ενέργειας και εφεδρειών.
    - Επίλυση προγράμματος ένταξης μονάδων (Τεχνικά Ελάχιστα, ρυθμοί, κλπ.).
    - Πληρωμή εφεδρειών με χρήση της «κλείδας επιμερισμού».

- <u>Από 1 Μαΐου 2009 (Πέμπτη Ημέρα Αναφοράς)</u>
    - Επίλυση ΗΕΠ όπως και κατά την Τέταρτη Ημέρα Αναφοράς.
    - Εκκαθάριση ΗΕΠ για την ενέργεια.
    - Εκκαθάριση εφεδρειών με βάση τις τιμές του ΗΕΠ και τις ποσότητες εφεδρειών που προσφέρθηκαν σε πραγματικό χρόνο.
    - Μεθοδολογία υπολογισμού ΟΤΑ (Οριακή Τιμή Αποκλίσεων).

- Εκκαθάριση αποκλίσεων για την ενέργεια.
- Ελεύθερη διαπραγμάτευση ΑΔΙ (Αποδεικτικά Διαθεσιμότητας Ισχύος).

1.3.3 ΑΝΤΙΚΕΙΜΕΝΟ ΤΟΥ ΗΕΠ

Ο Ημερήσιος Ενεργειακός Προγραμματισμός περιλαμβάνει τα ακόλουθα:

- Σχεδιασμός βέλτιστου προγράμματος ένταξης μονάδων και έγχυσης ενέργειας για την ικανοποίηση της ζήτησης ενέργειας για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής, με τέτοιο τρόπο που να επιφέρει μεγιστοποίηση του κοινωνικού πλεονάσματος(διαφορά κοινωνικού οφέλους από το συνολικό κόστος παραγωγής).

- Κατάρτιση προγραμμάτων Επικουρικών Υπηρεσιών ώστε με το ελάχιστο δυνατό κόστος να είναι δυνατό να καλυφθούν οι σχετικές ανάγκες για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής.

- Προσδιορισμός της Οριακή Τιμής Παραγωγής σε κάθε Λειτουργική Ζώνη για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής.

- Υπολογισμός της **Οριακής Τιμής του Συστήματος (ΟΤΣ)** για την ενέργεια για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής.

- Υπολογισμός των Τιμών Εφεδρείας για κάθε τιμολογούμενη Επικουρική Υπηρεσία για κάθε Περίοδο Κατανομής της Ημέρας Κατανομής.

1.4 ΟΡΙΑΚΗ ΤΙΜΗ ΣΥΣΤΗΜΑΤΟΣ (ΟΤΣ)

Η Οριακή Τιμή του Συστήματος (ΟΤΣ) είναι η τιμή στην οποία εκκαθαρίζεται η αγορά ηλεκτρικής ενέργειας και είναι η τιμή που εισπράττουν όλοι οι όσοι εγχέουν ενέργεια στο Σύστημα και πληρώνουν όλοι όσοι ζητούν ενέργεια από το Σύστημα. Συγκεκριμένα, η Οριακή Τιμή του Συστήματος διαμορφώνεται από τον συνδυασμό των προσφορών τιμών και ποσοτήτων που υποβάλλουν κάθε μέρα οι διαθέσιμες μονάδες παραγωγής ηλεκτρικής ενέργειας, και του ωριαίου φορτίου ζήτησης ηλεκτρικής ενέργειας, που διαμορφώνεται σε καθημερινή βάση από τους καταναλωτές.

Επιχειρώντας μια απλή περιγραφή του τρόπου υπολογισμού της Οριακής Τιμής του Συστήματος, σύμφωνα με τις βασικές αρχές της μικροοικονομικής θεωρίας, μπορεί να αναφερθεί ότι οι μονάδες παραγωγής κατατάσσονται αναλόγως των προσφορών τους σε αύξουσα σειρά, ξεκινώντας από την χαμηλότερη προσφερόμενη τιμή για ορισμένη ποσότητα ενέργειας και καταλήγοντας στην υψηλότερη προσφερόμενη τιμή.

Στο σημείο όπου οι προσφερόμενες ποσότητες ενέργειας εξυπηρετούν το ζητούμενο φορτίο, καθορίζεται και η Οριακή Τιμή του Συστήματος. Στην ουσία, η Οριακή τιμή του Συστήματος συμπίπτει με την προσφορά της τελευταίας μονάδας που πρέπει να λειτουργήσει για να καλυφθεί η ζήτηση.

Για λόγους προστασίας των καταναλωτών και διαμόρφωσης συνθηκών υγιούς ανταγωνισμού τίθεται διοικητικά ανώτερο όριο ως προς την προσφερόμενη τιμή, το οποίο έχει τεθεί ίσο με 150€/MWh καθώς και κατώτερο επίπεδο προσφορών, το οποίο είναι το μεταβλητό κόστος της μονάδας, ώστε στις περισσότερες περιπτώσεις οι παραγωγοί να πληρώνονται το κόστος καυσίμου τους (*ΡΑΕ, 2010*).

Η Οριακή Τιμή Συστήματος, πέρα από το γεγονός ότι συνιστά την τιμή εκκαθάρισης των συναλλαγών σε ημερήσια βάση, επιτρέπει την ανάπτυξη καμπυλών αναφοράς οι οποίες αποτελούν σημείο αναφοράς ως προς τη χρηματοοικονομική διαχείριση της αγοράς ηλεκτρικής ενέργειας και την υποκείμενη αξία για επιπλέον προϊόντα, όπως είναι τα προϊόντα αντιστάθμισης, τα προθεσμιακά προϊόντα, τα παράγωγα κ.α., που μπορούν να εξυπηρετούν συγκεκριμένες ανάγκες των συμμετεχόντων (*ΔΕΣΜΗΕ – Εγχειρίδιο Αγοράς, 2010*).

Η ΟΤΣ υπολογίζεται από τη ΔΕΣΜΗΕ με τη βοήθεια της Οριακής τιμής Παραγωγής (ΟΤΠ). Η ΟΤΠ κάθε λειτουργικής ζώνης ισούται με το σκιώδες κόστος (πολλαπλασιαστής Lagrange) του περιορισμού του ισοζυγίου ισχύος της ζώνης. Στην περίπτωση που οι διαζωνικοί περιορισμοί ροής δεν είναι δεσμευτικοί, οι δύο δυικές τιμές ταυτίζονται και η ΟΤΣ ισούται με αυτή την κοινή τιμή. Στην περίπτωση όμως που οι διαζωνικοί περιορισμοί είναι δεσμευτικοί, οι δύο δυικές τιμές μπορεί να είναι διαφορετικές και τότε η ΟΤΣ υπολογίζεται ως ο σταθμισμένος (με την παραγωγή των ζωνών) μέσος όρος των οριακών τιμών παραγωγής των ζωνών. Σε κάθε περίπτωση η ΟΤΠ κάθε λειτουργικής ζώνης καθορίζεται από μία ή περισσότερες οντότητες εντός της εκάστοτε λειτουργικής ζώνης, οι οποίες μπορεί να είναι μονάδες παραγωγής, κατανεμόμενα φορτία, εισαγωγές ή εξαγωγές.

Σε αυτό το σημείο είναι σημαντικό να αναφερθεί πως η επίλυση του ΗΕΠ από τη ΔΕΣΜΗΕ λαμβάνει χώρα από τις 12:30 έως τις 14:00 της προηγούμενης της Ημέρας Κατανομής, και η δημοσίευση των αποτελεσμάτων του ΗΕΠ, συνεπώς και της ΟΤΣ, γίνεται στις 14:00 της προηγούμενης της Ημέρας Κατανομής.

Κεφάλαιο 2. Πρόβλεψη Οριακής Τιμής Συστήματος (ΟΤΣ)

2.1 Χρησιμότητα Πρόβλεψης της ΟΤΣ

Οι δύο πιο σημαντικοί παράγοντες της Αγοράς Ηλεκτρικής Ενέργειας είναι οι ακόλουθοι:

- ΟΤΣ (οριακή τιμή συστήματος) ή αλλιώς SMP (system marginal price) ή MCP (market clearing price),
- ΟΤΦ (οριακή τιμή φορτίου) ή MCQ (market clearing quantity).

Στα πλαίσια της παρούσας διπλωματικής θα επιχειρήσουμε να ασχοληθούμε μόνο με την πρόβλεψη της ΟΤΣ.

Σε γενικές γραμμές, η όσο το δυνατόν ακριβέστερη πρόβλεψη της ΟΤΣ αποτελεί ένα αρκετά δύσκολο εγχείρημα με το οποίο έχουν ασχοληθεί αρκετοί ερευνητές τα τελευταία χρόνια. Όμως, όπως θα δούμε και στη συνέχεια, η ΟΤΣ είναι ένα μέγεθος ιδιαίτερα ευμετάβλητο και ασταθές, γεγονός που καθιστά τη διαδικασία πρόβλεψής της εξαιρετικά δύσκολη. Χαρακτηριστικό είναι το γεγονός που αναφέρουν οι *Sanjeev, Lalit και Ashwani (2009)* πως ενώ η πρόβλεψη φορτίου είναι σε ικανοποιητικά επίπεδα (MAPE<3%) η πρόβλεψη της τιμής βρίσκεται ακόμα σε αρχικό στάδιο. Στη συνέχεια παραθέτουμε την ετήσια μεταβλητότητα κάποιων χαρακτηριστικών μεγεθών της αγοράς ενέργειας:

- 30% για πετρέλαιο.
- 50% για φυσικό αέριο.
- 60% για ενέργεια (φορτίο).
- Πάνω από 200% για spot market ηλεκτρικής ενέργειας.

Όμως γιατί η πρόβλεψη της ΟΤΣ απασχολεί σε τόσο μεγάλο βαθμό τους συμμετέχοντες στην αγορά; Για να γίνει κατανοητή η χρησιμότητα – αναγκαιότητα της πρόβλεψης είναι απαραίτητο να διαχωρίσουμε την πρόβλεψη σε δύο μεγάλες κατηγορίες με βάση τον χρονικό ορίζοντα:

- την βραχυπρόθεσμη (ημέρα – εβδομάδα), και
- τη μακροπρόθεσμη (μήνας – έτος).

Ενίοτε θα μπορούσαμε να αναφέρουμε και τη μεσοπρόθεσμη (μήνας), αλλά μπορούμε στη γενικότητα να τη συμπεριλάβουμε στην μακροπρόθεσμη.

2.1.1. Βραχυπρόθεσμη Πρόβλεψη

Οι κυρίαρχοι λόγοι που είναι χρήσιμη η βραχυπρόθεσμη πρόβλεψη (ημέρα έως εβδομάδα) της ΟΤΣ είναι οι ακόλουθοι:

- Βοηθάει στη διαδικασία αγοραπωλησίας ενέργειας και στη υιοθέτηση κατάλληλης στρατηγικής από μία εταιρεία που συμμετέχει στην αγορά ενέργειας με στόχο την ελαχιστοποίηση του κόστους.

- Πραγματοποιούνται ευμενέστερες διμερής συμφωνίες.

- Προγραμματίζεται δέσμευση των παραγωγικών μονάδων με βάση τις προβλεπόμενες τιμές και προσφορά της αναμενόμενης παραγόμενης ενέργειας.

2.1.2 Μεσοπρόθεσμη Πρόβλεψη

Η μεσοπρόθεσμη πρόβλεψη παρουσιάζει κοινά χαρακτηριστικά με την μακροπρόθεσμη, όμως είναι δυνατό να αποτελέσει και προέκταση της βραχυπρόθεσμης. Αυτό συμβαίνει διότι δεν είναι σαφές πότε μπορούμε να χαρακτηρίσουμε μεσοπρόθεσμη μία πρόβλεψη, ιδιαίτερα στο πρόβλημα πρόβλεψης της ΟΤΣ, καθώς τα διαθέσιμα δεδομένα είναι συνήθως σε επίπεδο ώρας. Επομένως, ανάλογα με τον ορίζοντα της πρόβλεψης, καθορίζεται και το κατά πόσο μπορεί αυτή να θεωρηθεί μεσοπρόθεσμη.

2.1.3 Μακροπρόθεσμη Πρόβλεψη

Η μακροπρόθεσμη πρόβλεψη (μήνας έως έτος) της ΟΤΣ είναι χρήσιμη για τους ακόλουθους λόγους:

- Προγραμματισμός επενδύσεων στη δυναμικότητα ηλεκτροπαραγωγής (generation capacity).

- Ανάλυση ρίσκου αγοράς στο εγγύς η απώτερο μέλλον.

- Προγραμματισμός παραγωγής και συντήρησης.

- Σύναψη συμβολαίων για καύσιμα.

- Τιμολόγηση Παράγωγων.

Περισσότερο κρίσιμη για τους συμμετέχοντες στην αγορά ενέργειας μπορεί να χαρακτηριστεί η βραχυπρόθεσμη πρόβλεψη, καθώς είναι αυτή που μπορεί να τους επιφέρει άμεσα κέρδη ή ζημίες και να καθορίσει την οργάνωση της παραγωγής, στην περίπτωση των παραγωγών, ή αντίστοιχα την κατανάλωση ενέργειας στην περίπτωση των καταναλωτών, στο άμεσο μέλλον.

Για αυτόν το λόγο στην παρούσα διπλωματική θα ασχοληθούμε με το **"day ahead forecasting"** που αφορά τη βραχυπρόθεσμη πρόβλεψη της ΟΤΣ και είναι αυτό που επιζητούν οι συμμετέχοντες στην ηλεκτρική αγορά ενέργειας

2.2 Βασικά Χαρακτηριστικά ΟΤΣ

Το μεγαλύτερο πρόβλημα στην πρόβλεψη της ΟΤΣ είναι η μεταβλητότητα, δηλαδή το μέτρο της μεταβολής της τιμής της ηλεκτρικής ενέργειας σε ένα δεδομένο χρονικό διάστημα (μη ομογενής καμπύλη σε σύγκριση με το φορτίο) (*Deepak & Swarup, 2011*).

Υπάρχουν αρκετοί παράγοντες που επηρεάζουν ,άλλοι περισσότερο και άλλοι λιγότερο, την μεταβλητότητα της ΟΤΣ. Οι βασικότεροι από αυτούς είναι:

- Συμφόρηση κατά τη μεταφορά.
- Βασικός λόγος για τα spikes στην τιμή είναι ότι η ζήτηση και η προσφορά πρέπει να συνδέονται με μια δευτερόλεπτο προς δευτερόλεπτο σχέση, κάτι το οποίο δεν συμβαίνει.
- Μεταβλητότητα στην τιμή των καυσίμων.
- Αβεβαιότητα φορτίου.
- Διακυμάνσεις στην υδροηλεκτρική παραγωγή.
- Αβεβαιότητα παραγωγής (διακοπές).
- Συμπεριφορά των συμμετεχόντων στην αγορά ενέργειας (βασισμένη στην αναμενόμενη τιμή).
- Καθοδήγηση της αγοράς.

Εξετάζοντας τώρα τα ιδιαίτερα χαρακτηριστικά της ΟΤΣ, είναι πολύ εύκολο να διακρίνουμε το λόγο που το συγκεκριμένο μέγεθος παρουσιάζει υψηλή μεταβλητότητα. Ανάμεσα στις ιδιαιτερότητες της ΟΤΣ μπορούμε να ξεχωρίσουμε τα ακόλουθα 5 στοιχεία (*Sanjeev et al., 2009*):

- ❖ Υψηλή συχνότητα.
- ❖ Μη σταθερή μέση τιμή και διασπορά.
- ❖ Πολλαπλή εποχικότητα.

- Επηρεάζεται από το ημερολόγιο (calendar effect).
- Υψηλό επίπεδο αστάθειας και υψηλό ποσοστό ασυνήθιστων τιμών (Spikes).

Αναζητώντας του λόγους για τους οποίους η ΟΤΣ παρουσιάζει τα παραπάνω χαρακτηριστικά, μπορούμε να αναφέρουμε ότι αυτά προκύπτουν από το γεγονός ότι:

- Η ενέργεια δεν αποθηκεύεται
- Απαιτείται να υπάρχει ισορροπία ζήτησης – προσφοράς
- Η ζήτηση είναι ανελαστική μέσα σε μικρό χρονικό διάστημα
- Η αγορά λειτουργεί υπό συνθήκες ολιγοπωλίου
- Η ζήτηση και η προσφορά επηρεάζουν την ισορροπία της αγοράς

2.3 Βασικές Μεθοδολογίες Πρόβλεψης

Σύμφωνα με τους *Sanjeev, Lalit* και *Ashwani (2009)*, υπάρχουν 3 βασικές κατηγορίες μοντέλων πρόβλεψης που χρησιμοποιούνται για την παραγωγή προβλέψεων της ΟΤΣ. Οι κατηγορίες αυτές είναι οι ακόλουθες:

- Game Theory Models (Nash equilibrium, Cournot, Bertrand).
- Simulation Models (MAPS, UPLAN).
- **Time Series Models** (Stochastic models(ARIMA,AR,MA,GARCH),AI models, Regression/causal models).

Στα πλαίσια της παρούσας διπλωματικής θα εξεταστούν κυρίως τα μοντέλα χρονοσειρών (Time Series Models), τα οποία απαιτούν την έκφραση των υπό μελέτη μεγεθών με τη μορφή χρονοσειρών. Η κατηγορία αυτή μπορεί να διαχωριστεί περαιτέρω με τον ακόλουθο τρόπο (*Πετρόπουλος, Ασημακόπουλος, 2011*):

- **Μοντέλα Χρονοσειρών**: εξετάζουν τη χρονοσειρά από αριθμητική σκοπιά ως σύνολο τιμών δίνοντας μικρή έμφαση στο τι εκφράζει. Περιλαμβάνουν την αποσύνθεση (decomposition), την εξομάλυνση (smoothing) και τις αυτοπαλινδρομικές μεθόδους κινητού μέσου όρου (Autoregressive Moving Average ή ARMA).

- **Αιτιοκρατικά Μοντέλα**: εστιάζουν στις σχέσεις αιτίου-αποτελέσματος με συνέπεια να έχει ιδιαίτερη σημασία το τι εκφράζει η εκάστοτε χρονοσειρά. Περιλαμβάνουν τις γραμμικές ή μη γραμμικές παλινδρομήσεις (πολλαπλές ή μη).

Στα αιτιοκρατικά μοντέλα, και συγκεκριμένα στις μη γραμμικές παλινδρομήσεις μπορούν να συμπεριληφθούν και τα μοντέλα νευρωνικών δικτύων. Αναλυτικότερη περιγραφή των μεθόδων κάθε κατηγορίας, καθώς και της συνδυαστικής μεθόδου Theta, θα γίνει στη συνέχεια.

2.3.1 Μοντέλα Χρονοσειρών

Τα μοντέλα χρονοσειρών, αποτελούν αρκετά δημοφιλή κατηγορία για την πρόβλεψη της ΟΤΣ, κυρίως ως μοντέλα αναφοράς (benchmark models) και έχουν χρησιμοποιηθεί ευρέως σε πάρα πολλές σχετικές έρευνες, είτε μεμονωμένα είτε σε συνδυασμό μεταξύ τους (συνδυαστικές – υβριδικές μέθοδοι). Η διαδεδομένη χρήση τους οφείλεται κυρίως στην ευκολία τους ως προς τη μαθηματική τους διατύπωση σε αντίθεση με την κατηγορία των αιτιοκρατικών.

Απαραίτητη προϋπόθεση για την εφαρμογή των συγκεκριμένων μοντέλων είναι η ύπαρξη ιστορικών δεδομένων για την τιμή του μεγέθους το οποίο θέλουμε να προβλέψουμε, σε σταθερές χρονικές περιόδους. Κυρίαρχη υπόθεση είναι ότι η τιμή του μεγέθους μεταβάλλεται ακολουθώντας ένα συγκεκριμένο λανθάνον πρότυπο, χρονικά επαναλαμβανόμενο και σταθερό. Η παραγωγή των προβλέψεων επιτυγχάνεται με την αναγνώριση του ακολουθούμενου προτύπου και την προέκταση του στο μέλλον, με βάση τον χρονικό ορίζοντα που επιθυμούμε.

Αξίζει να αναφερθεί πως η σχέση που περιγράφει τέτοιου είδους μοντέλα είναι γραμμική.

2.3.2 Αιτιοκρατικά Μοντέλα

Τα αιτιοκρατικά μοντέλα θεωρούνται ως τα πιο διαδεδομένα μοντέλα για την πρόβλεψη της τιμής της ηλεκτρικής ενέργειας, καθότι έχουν τη δυνατότητα να ενσωματώνουν διάφορες παραμέτρους που επηρεάζουν είτε σε μικρό είτε σε μεγάλο βαθμό της τιμή της ΟΤΣ.

Τα μοντέλα αυτά λειτουργούν υπό την θεμελιώδη υπόθεση ότι το υπό πρόβλεψη μέγεθος (εξαρτημένη μεταβλητή) σχετίζεται μέσω κάποιας σταθερής σχέσης με κάποιες παραμέτρους (ανεξάρτητες μεταβλητές), οι οποίες και το επηρεάζουν.

2.3.3 Βασικές διαφορές

Η ουσιαστική διαφορά μεταξύ των μοντέλων χρονοσειρών και των αιτιοκρατικών μοντέλων είναι η συνάρτηση περιγραφής του κάθε συστήματος. Όπως αναφέρθηκε προηγουμένως, στην περίπτωση των μοντέλων χρονοσειρών η συνάρτηση που περιγράφει το εκάστοτε σύστημα είναι προκαθορισμένη και γραμμική, ενώ στα αιτιοκρατικά μοντέλα προσδιορίζεται από τη συσχέτιση που έχουν οι ανεξάρτητες με την εξαρτημένη μεταβλητή του συστήματος. Με αυτό τον τρόπο μπορεί να προκύψει μία συνάρτηση μη γραμμική, και επιπλέον συγκεκριμένη τιμή της εξαρτημένης μεταβλητής να προκύπτει από συγκεκριμένο συνδυασμό των ανεξάρτητων μεταβλητών.

Ως βασικό μειονέκτημα των αιτιοκρατικών μοντέλων θα μπορούσαμε να θεωρήσουμε την ανάγκη για σημαντικά περισσότερο αριθμό δεδομένων σε σύγκριση με τα μοντέλα χρονοσειρών, όπως και η ευαισθησία που μπορεί να παρουσιάζει η εξαρτημένη μεταβλητή σε πιθανές αλλαγές στον τρόπο με τον οποίο συνδέονται οι ανεξάρτητες μεταβλητές του συστήματος.

Επιπλέον, σε κάποιες περιπτώσεις είναι απαραίτητη η επιμέρους πρόβλεψη της τιμής των ανεξάρτητων μεταβλητών ώστε να μπορέσουν να χρησιμοποιηθούν στην πρόβλεψη της εξαρτημένης μεταβλητής, γεγονός το οποίο περιπλέκει τη διαδικασία της πρόβλεψης και απαιτεί περισσότερο χρόνο.

Παρά τα μειονεκτήματα που μπορεί να παρουσιάζουν τα αιτιοκρατικά μοντέλα, και αφορούν κυρίως τη διαχείριση των σαφώς περισσότερων δεδομένων που απαιτούνται για την παραγωγή προβλέψεων, χαρακτηρίζονται από ένα ιδιαίτερα σημαντικό πλεονέκτημα, που δεν είναι άλλο από τη δυνατότητα εξέτασης πολλών εναλλακτικών σεναρίων.

Αυτό σημαίνει ότι ο ενδιαφερόμενος για την πρόβλεψη ενός μεγέθους, μπορεί να συνδυάσει με διαφορετικό τρόπο τις ανεξάρτητες μεταβλητές, δίνοντάς τους κάθε φορά διαφορετική βαρύτητα, οδηγούμενος έτσι σε διαφορετικά αποτελέσματα που μπορεί να καλύπτουν κάθε πιθανή εξέλιξη της εξαρτημένης μεταβλητής συναρτήσει των παραμέτρων που την επηρεάζουν.

Όσον αφορά τα μοντέλα χρονοσειρών, η μεγάλη τους αδυναμία είναι η συσχέτιση του προβλεπόμενου μεγέθους με τους παράγοντες που το επηρεάζουν, το οποίο όπως προαναφέρθηκε αποτελεί το σπουδαιότερο προτέρημα των αιτιοκρατικών μοντέλων. Με πιο απλά λόγια τα μοντέλα χρονοσειρών μπορούν να μας ¨πουν¨ τι θα συμβεί αλλά όχι και το λόγο για τον οποίο συμβαίνει.

Παρά αυτή τους την αδυναμία, οι μέθοδοι που υπάγονται στα μοντέλα χρονοσειρών προτιμώνται για τους ακόλουθους πολύ σημαντικούς λόγους:

- Δεν είναι πάντα εφικτό να συσχετίσουμε ένα μέγεθος με ορισμένους παράγοντες που θεωρούμε ότι το επηρεάζουν, και ακόμη και αν το καταφέρουμε ίσως να μην μπορούμε να προσδιορίσουμε τον τρόπο με τον οποίο συνδέονται μεταξύ τους.

- Σε αρκετές περιπτώσεις μας ενδιαφέρει να προσδιορίσουμε απλά την τιμή ενός μεγέθους και όχι το πώς προέκυψε αυτή η τιμή.

- Το κόστος, σε κόπο και χρόνο, είναι σαφέστατα μεγαλύτερο όταν προσπαθούμε να εντοπίσουμε το λόγο μεταβολής του εξεταζόμενου μεγέθους.

2.4 Έρευνες για την πρόβλεψη της ΟΤΣ

Όπως αναφέρθηκε και προηγουμένως, μπορούμε να κατηγοριοποιήσουμε την πρόβλεψη της ΟΤΣ με βάση τον χρονικό ορίζοντα στον οποίο στοχεύουμε. Έχει εκπονηθεί πληθώρα ερευνών πάνω στο συγκεκριμένο ζήτημα και θα ήταν ασφαλές να πούμε ότι η πλειοψηφία αυτών επικεντρώνεται στη βραχυπρόθεσμη πρόβλεψη της ΟΤΣ ή αλλιώς στο **“day ahead forecasting”.** Αυτό σημαίνει πως οι περισσότεροι ερευνητές ασχολήθηκαν με την παραγωγή κυρίως ωριαίων προβλέψεων για την επόμενη ημέρα ή και για μία εβδομάδα. Σπανιότερες είναι οι αναφορές για μακροχρόνιες προβλέψεις και ελάχιστες για τις μεσοπρόθεσμες.

Βέβαια, σχετικά με τις μεσοπρόθεσμες, εξαρτάται από τον τύπο των δεδομένων. Για παράδειγμα, αν έχουμε 24ωρα δεδομένα, μία πρόβλεψη με ορίζοντα 7 ημερών (εβδομάδα) μπορεί κάλλιστα να θεωρηθεί μεσοπρόθεσμη, όπως επίσης αν έχουμε ημερήσιες τιμές και λάβουμε ως ορίζοντα 1 μήνα. Επομένως, μπορούμε να ενσωματώσουμε τη μεσοπρόθεσμη πρόβλεψη στην κατηγορία των βραχυπρόθεσμων, αρκεί να θεωρήσουμε τον κατάλληλο χρονικό ορίζοντα για τα κατάλληλα δεδομένα. Ακολουθεί ένας ενδεικτικός πίνακας (Πίνακας 1) με τα ονόματα ορισμένων ερευνητών που ασχολήθηκαν με βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη της ΟΤΣ καθώς και ποια ήταν η μέθοδος που χρησιμοποίησαν, θεωρώντας μεσοπρόθεσμες και τις προβλέψεις που έχουν ως ορίζοντα την εβδομάδα:

Ερευνητής	**Έτος**	**Μέθοδος**	**Είδος**
J. Conejo et al.	2005	Various FM	Βραχυπρόθεσμη
N. V. Karakatsani & D. W. Bunn	2008	Regression Models	Βραχυπρόθεσμη
N. Bowden, J. E. Payne	2008	ARIMA-EGARCH	Βραχυπρόθεσμη
R. Weron & A. Misiorek	2008	Parametric and not Timeseries Models	Βραχυπρόθεσμη
Whei-Min Lin, Hong-Jey Gow & Ming-Tang Tsai	2010	ERBFN	Βραχυπρόθεσμη
Jinxing Che, Jianzhou Wang	2010	SVR-ARIMA	Βραχυπρόθεσμη
A. Karsaz, H.R. Mashhadi & M.M. Mirsalehi	2010	Co-Co	Βραχυπρόθεσμη
F. Serinaldi	2011	GAMLSS	Βραχυπρόθεσμη
Deepak Singhal and K.S Swarup	2011	ANN	Βραχυπρόθεσμη
H.M.I. Pousinho, V.M.F. Mendes & J.P.S. Catalão	2012	Hybrid PSO-ANFIS	Βραχυπρόθεσμη
S. S. Torghabam	2010	Regression Models	Μεσοπρόθεσμη
M. Shafie-khah et al.	2011	WAV-ARIMA-RBFN	Μεσοπρόθεσμη
H.M.I. Pousinho, V.M.F. Mendes, J.P.S. Catalao	2011	WNF	Μεσοπρόθεσμη
G. Ham & A. Borison	2006	Proposing Methods	Μακροπρόθεσμη
S. Schlueter	2010	Various FM	Βραχυ-Μακρο

Πίνακας 2.4.1 – Κατηγοριοποίηση ερευνών με βάση το είδος πρόβλεψης

2.4.1 Μέθοδοι για Βραχυπρόθεσμη Πρόβλεψη

Όπως αναφέρθηκε και προηγουμένως η βραχυπρόθεσμη πρόβλεψη της ΟΤΣ αποτελεί το αντικείμενο της πλειοψηφίας των περισσότερων ερευνών σχετικά με το συγκεκριμένο ζήτημα. Όπως διαπιστώνουμε και από τον παραπάνω πίνακα, οι ερευνητές χρησιμοποιούν μία μεγάλη γκάμα από μεθόδους προβλέψεων, ίσως δίνοντας μεγαλύτερη βαρύτητα σε συνδυαστικές μεθόδους, όπως είναι για παράδειγμα η Wavelet-ARIMA, ή συνδυαστικές μέθοδοι παλινδρόμησης, ή ακόμα και συνδυαστικές μέθοδοι που περιέχουν Νευρωνικά Δίκτυα (RBFN, ERBFN, PSO-ANFIS). Τέτοιες μέθοδοι προτιμούνται στην πράξη ,καθώς μπορούν να απομονώσουν περισσότερα από τα ιδιαίτερα χαρακτηριστικά του μεγέθους της ΟΤΣ, όπως για παράδειγμα την πολλαπλή εποχικότητα ή τις ασυνήθιστές τιμές (spikes).

Το θέμα της πολλαπλής εποχικότητας ,απασχόλησε κατά καιρούς πάρα πολλούς ερευνητές, καθώς αποτελεί ίσως το πιο περίπλοκο χαρακτηριστικό της ΟΤΣ, μαζί με τις ξαφνικές τις μεταβολές των τιμών (spikes). Βασικές τεχνικές που ακολουθούνται για την εξάλειψη της πολλαπλής εποχικότητας είναι οι μέθοδοι εκθετικής εξομάλυνσης, όπως η μέθοδος διπλής εποχικότητας (Double Seasonal Exponential Smoothing) του Taylor (2003) που αφορά χρονοσειρές που παρουσιάζουν δύο εποχιακούς κύκλους, έναν μικρό που επαναλαμβάνεται με μεγάλη συχνότητα μέσα σε ένα μεγαλύτερο, και η μέθοδος Holt – Winters η οποία αφορά τόσο πολλαπλασιαστική όσο και αθροιστική πολλαπλή εποχικότητα (*Phillip G. Gould et al., 2008*). Θα αναφερθούμε εκτενέστερα στο θέμα της πολλαπλής εποχικότητας αλλά και στον τρόπο αντιμετώπισής του στο πρακτικό μέρος της διπλωματικής.

2.4.1.1 Μέθοδοι Παλινδρόμησης

Οι μέθοδοι παλινδρόμησης χρησιμοποιούνται ευρέως στην διαδικασία πρόβλεψης της ΟΤΣ, κυρίως λόγω των πολλών παραμέτρων που μπορούν να επηρεάσουν το συγκεκριμένο μέγεθος. Χαρακτηριστική είναι η προσπάθεια των Karakatsani και Burn (2008) να αποτυπώσουν την επίδραση διάφορων παραμέτρων, όπως οικονομικών, στρατηγικών, τεχνικών, καθώς και παραγόντων ρίσκου, σε συνδυασμό με την δυναμική αυτών των παραγόντων μέσα στο χρόνο χρησιμοποιώντας ένα μοντέλο παλινδρόμησης με χρονική παράμετρο (time–varying parameter (TVP) regression). Στην έρευνα τους λαμβάνουν ως ανεξάρτητες μεταβλητές τις ακόλουθες:

- Ζήτηση φορτίου (demand).
- Κλίση και καμπυλότητα ζήτησης (demand slope and curvature), υπολογιζόμενες από τις πρώτες και δεύτερες διαφορές της χρονοσειράς ζήτησης.
- Μεταβλητότητα ζήτησης (demand volatility) ,δηλαδή τη διακύμανση της ζήτησης λόγω πρόσκαιρων, καιρικών και καταναλωτικών προτύπων.
- Περιθώριο (Margin), ένα μέτρο που αφορά την περίσσεια δυναμικότητα ηλεκτροπαραγωγής, και ως εκ τούτου λειτουργεί ως δείκτης ανεπάρκειας.

- Lag-1 Margin, που είναι η τιμή του περιθωρίου την προηγούμενη μέρα σε συγκεκριμένη περίοδο.

- Ανεπάρκεια (scarcity), αφορά τον έντονο αντίκτυπο που έχει το περιθώριο της δυναμικότητας ηλεκτροπαραγωγής πάνω στην τιμή πάνω από ένα όριο.

- Μεταβλητότητα τιμής.

- Διαφορά (spread), που εκφράζει τη διαφορά ανάμεσα στην ανεπάρκεια και στο πλεόνασμα της ενέργειας.

- Εποχικότητα (seasonality).

- Τάση (trend).

- Ημερήσιες και εβδομαδιαίες επιδράσεις (diurnal and weekly effects).

Χρησιμοποιώντας λοιπόν τις παραπάνω μεταβλητές ως παράγοντες που επηρεάζουν την ΟΤΣ, έθεσαν ως μοντέλο αναφοράς ένα μοντέλο αυτοπαλινδρόμησης (AR) και το συνέκριναν με άλλες μεθόδους παλινδρόμησης (όπως LRL, TVP- Regression, RS-Regression, TVP-AR, RS-AR, Regression with Trend), με κριτήριο τα σφάλματα MAPE, MAE, RMSE, MaxAE και MaxAPE.

Αξίζει να σημειωθεί πως υπήρξε διαχωρισμός στην διαδικασία πρόβλεψης σχετικά με το περιθώριο (Margin), όπου στην πρώτη περίπτωση λήφθηκε υπόψη η αβεβαιότητα σχετικά με το περιθώριο, ενώ στη δεύτερη οι προβλέψεις έγιναν δίχως προκατειλημμένες προσδοκίες για το περιθώριο.

Τα αποτελέσματα της έρευνας ήταν θετικά και έδειξαν πως ένα μοντέλο μπορεί να δώσει καλύτερες προβλέψεις όταν λαμβάνουμε υπόψη μας την μεταβολή των ανεξάρτητων μεταβλητών μέσα στο χρόνο, στην περίπτωση όμως που έχουμε ως ορίζοντα την ημέρα ή ακόμα καλύτερα συγκεκριμένες περιόδους της ημέρας.

Όσον αφορά τις μεθόδους παλινδρόμησης, είναι σαφές ότι η ορθή επιλογή των παραγόντων που επηρεάζουν την ΟΤΣ είναι καθοριστική για την παραγωγή όσο το δυνατόν ακριβέστερων προβλέψεων. Μεγάλο ρόλο παίζει και ο τρόπος που συσχετίζονται τα μεγέθη μεταξύ τους και απαιτείται μεγάλη προσοχή στην αποσαφήνιση του τρόπου που μπορεί να επιδρά κάθε ανεξάρτητη μεταβλητή στην εξαρτημένη, η οποία στο πρόβλημα μας είναι η ΟΤΣ. Στη συνέχεια παρατίθεται ένας συγκεντρωτικός πίνακας με πιθανούς παράγοντες που μπορούν να ληφθούν ως ανεξάρτητες μεταβλητές σε μία πολλαπλή παλινδρόμηση, χωρισμένους σε κατηγορίες ως προς το είδος τους:

ΠΑΡΑΓΟΝΤΕΣ	**K**
Ιστορικά δεδομένα φορτίου	K1
Ρυθμός φορτίου συστήματος	K1
Εισαγωγές/Εξαγωγές	K1
Πλεόνασμα/έλλειμμα Χωρητικότητας Ηλεκτροπαραγωγής	K1
Εφεδρείες	K1
Πυρηνική Παραγωγή	K1
Θερμική Παραγωγή	K1
Υδροηλεκτρική Παραγωγή	K1
Χωρητικότητα Ηλεκτροπαραγωγής	K1
Δεσμευτικοί Περιορισμοί τους Συστήματος	K1
Όρια γραμμών μεταφοράς	K1
Ιστορικά δεδομένα Οριακής Τιμής Φορτίου	K1
Πρόβλεψη Φορτίου	K2
Πρόβλεψη εφεδρειών	K2
Θερμοκρασία	K2
Θερμοκρασία σημείου δρόσου	K2
Καιρικές συνθήκες	K2
Τιμή Καυσίμων(Πετρέλαιο, Φυσικό Αέριο κτλ.)	K2
Κενά Ηλεκτροπαραγωγής	K3
Κατάσταση γραμμών μεταφοράς	K3
Απρόβλεπτοι παράγοντες γραμμών μεταφοράς	K3
Δείκτης Συμφόρησης μεταφοράς ΗΕ	K3
Ιστορικές δεδομένα ΟΤΣ	K4
Ελαστικότητα ζήτησης	K4
Στρατηγικές προσφορών	K4
Δείκτης ύπαρξης ασυνήθιστων τιμών (spikes)	K4
Περίοδος επίλυσης	K5
Ημερολόγιο(ημέρα /μήνας /γιορτές /εποχές) (calendar effects)	K5
Αλλαγή ώρας	K5

Πίνακας 2.4.2 – Κατηγοριοποίηση Παραγόντων Παλινδρόμησης

Όπου οι κατηγορίες είναι οι ακόλουθες:

- K1 : Χαρακτηριστικά Αγοράς Ενέργειας.

- K2 : Μη στρατηγικές αβεβαιότητες.
- K3 : Στοχαστικές αβεβαιότητες.
- K4 : Δείκτες «συμπεριφοράς» της Αγοράς Ενέργειας.
- K5 : Χρονικές/Πρόσκαιρες επιδράσεις.

2.4.1.2 Νευρωνικά Δίκτυα (μη γραμμική παλινδρόμηση)

Τα Νευρωνικά Δίκτυα (Neural Networks (NN)) υπάγονται στην κατηγορία των μοντέλων Τεχνητής Νοημοσύνης (Artificial Intelligence models) και ουσιαστικά αποτελούν μία μορφή μη γραμμικής παλινδρόμησης. Όπως αναφέρουν οι *Sanjeev, Lalit* και *Ashwani (2009)*, τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks(ANN)) έχουν τη δυνατότητα να αντιλαμβάνονται την αυτοσυσχετιστική δομή μέσα σε μία χρονοσειρά ακόμη και αν η υποκείμενη σχέση που διέπει τη χρονοσειρά είναι άγνωστη ή πολύ δύσκολο να υπολογιστεί.

Από τη στιγμή που η ποσοτική πρόβλεψη βασίζεται στην εξόρυξη προτύπων παρατηρούμενη από γεγονότα του παρελθόντος και με προβολή τους στο μέλλον, τα Νευρωνικά Δίκτυα μπορούν να θεωρηθούν οι καλύτεροι υποψήφιοι για την παραγωγή προβλέψεων. Τα διαθέσιμα μοντέλα Νευρωνικών Δικτύων είναι τα ακόλουθα:

- Multilayer Feed Forward Neural Networks (FFNN).
- Radial Basis Function Network (RBFN).
- Support Vector Machine (SVM).
- Self-organizing map (SOM).
- Committee machine on NNs.
- Recurrent Neural Network (RNN).

Για τους λόγους που προαναφέρθηκαν, τα Νευρωνικά Δίκτυα χαίρουν ευρείας αποδοχής από τους ερευνητές που ασχολούνται με την πρόβλεψη της ΟΤΣ. Χαρακτηριστική είναι η έρευνα των *Deepak Singhal και K.S. Swarup (2011)*, οι οποίοι χρησιμοποίησαν έναν τριών επιπέδων πίσω-πολλαπλασιασμού Νευρωνικό Δίκτυο, με στόχο να προβλέψουν ωριαίες τιμές της ΟΤΣ με ορίζοντα 2 ημερών. Για το σκοπό αυτό ,χώρισαν τα δεδομένα τους σε 3 κατηγορίες, οι οποίες είναι:

1. Τιμή με κανονική τάση.
2. Τιμή με μικρές απότομες αλλαγές (small spikes).
3. Τιμή με μεγάλες απότομες αλλαγές (large spikes).

Σαν εισόδους στο Νευρωνικό Δίκτυο θεώρησαν:

i. Τις ιστορικές τιμές της ΟΤΣ.
ii. Χρονικούς δείκτες που σχετίζονται με τις τιμές της ΟΤΣ.
iii. Δεδομένα ζήτησης φορτίου.

Χρησιμοποιώντας ως δείκτες αξιολόγησης τα σφάλματα MAE και RMSE, τα οποία θα αναλύσουμε στη συνέχεια της διπλωματικής, κατέληξαν στο συμπέρασμα ότι το συγκεκριμένο μοντέλο ήταν ιδιαίτερα αποδοτικό για τις ημέρες με κανονική τάση, όμως παρουσιάζει μία λιγότερο καλή συμπεριφορά για τις ημέρες με απότομες αλλαγές τιμών, προτείνοντας παράλληλα ως μέτρο αντιμετώπισης αυτής της αδυναμίας του Νευρωνικού Δικτύου, την επιλογή εισόδων που εξηγούν τους λόγους εμφάνισης ασυνήθιστων τιμών

2.4.1.3 Συνδυαστικές Μέθοδοι

Η κατηγορία αυτή, αποτελεί ίσως την πιο αποτελεσματική και πιο διαδεδομένη λύση στο πρόβλημα πρόβλεψης της ΟΤΣ. Είναι πάρα πολλοί οι ερευνητές οι οποίοι συγχώνευσαν πολλά διαφορετικά μοντέλα δημιουργώντας έτσι υβριδικές μεθόδους, ώστε να έχουν τη δυνατότητα με κάθε ένα από αυτά να απομονώσουν τα ιδιαίτερα χαρακτηριστικά του συγκεκριμένου μεγέθους. Τα αποτελέσματα τους δικαιώνουν, καθώς οι συνδυαστικές μέθοδοι τείνουν να δίνουν ακριβέστερα αποτελέσματα, μειώνοντας σημαντικά τα σφάλματα σε σχέση με μεμονωμένες μεθόδους πρόβλεψης.

Σε ορισμένες από τις έρευνες που έχουν εκπονηθεί χρησιμοποιούνται μοντέλα αυτοπαλινδρομήσεων (AR, ARIMA), τα οποία μπορεί να συνδυάζονται με μοντέλα βασισμένα σε Νευρωνικά Δίκτυα (*M. Shafie-khah et* *al., 2011*) ή και με άλλες μεθόδους παλινδρόμησης (*Jinxing Che & Jianzhoy Wang, 2010*). Αντίστοιχα ισχύουν και για τις μεθόδους που έχουν ως βάση τα Νευρωνικά Δίκτυα, τα οποία μπορεί να συνδυάζονται μεταξύ τους (*A. Karsaz et al., 2010*) ή να χρησιμοποιούνται με μεθόδους βελτιστοποίησης (*H.M.I. Pousinho et al., 2012*).

2.4.2 Μέθοδοι για Μεσοπρόθεσμη Πρόβλεψη

Η Μεσοπρόθεσμη Πρόβλεψη της ΟΤΣ δεν έχει αποτελέσει μέχρι σήμερα αντικείμενο πολλών επιστημονικών ερευνών. Σημαντικό ρόλο στο γεγονός αυτό παίζει το ότι το συγκεκριμένο μέγεθος παρουσιάζει χαρακτηριστικά, όπως είναι η υψηλή μεταβλητότητα, τα οποία δυσκολεύουν σε μεγάλο βαθμό την πρόβλεψη του στο απώτερο μέλλον, καθώς και έλλειψη πολλών ιστορικών δεδομένων λόγω της βραχύβιας πορείας των ελεύθερων Αγορών Ηλεκτρικής Ενέργειας. Επιπλέον, όπως αναφέρθηκε και σε προηγούμενη ενότητα, η βραχυπρόθεσμη πρόβλεψη είναι εκείνη που ενδιαφέρει περισσότερο τους συμμετέχοντες στην Αγορά Ενέργειας, επομένως η ανάγκη για μεσοπρόθεσμες προβλέψεις είναι περισσότερο περιορισμένη αλλά όχι και δίχως χρησιμότητα.

Όπως χαρακτηριστικά αναφέρει στην έρευνά του ο *Torghaban (2010)*, η μεσοπρόθεσμη πρόβλεψη της τιμής της ηλεκτρικής ενέργειας είναι απαραίτητη για διάφορες εφαρμογές στην Ηλεκτρική Αγορά Ενέργειας, όπως η τιμολόγηση των παραγώγων, ο προγραμματισμός της συντήρησης για τις εταιρείες παραγωγής και ο καθορισμός του προϋπολογισμού και οι συμφωνίες για καύσιμα.

Στη συγκεκριμένη έρευνα επιχειρείται ο υπολογισμός αμιγώς μεσοπρόθεσμων προβλέψεων (ορίζοντας είναι το 1 έτος), γιατί όπως αναφέρθηκε και στην αρχή της ενότητας, μπορούμε να χαρακτηρίσουμε μία πρόβλεψη μεσοπρόθεσμη ανάλογα με τη συσχέτιση που υπάρχει μεταξύ του χρονικού επιπέδου των δεδομένων μας (στην περίπτωση της ΟΤΣ είναι συνήθως ανά ώρα) και στον χρονικό ορίζοντα που επιθυμούμε.

Πιο συγκεκριμένα, τα δεδομένα προέρχονται από την αγορά της Σκανδιναβικής Κοινοπραξίας (Nord Pool) και αφορούν εβδομαδιαίες μέσες τιμές της ΟΤΣ για 10 χρόνια, οι οποίες συναθροίζονται σε μέσες μηνιαίες τιμές για τις ανάγκες της έρευνας.

Ο χρονικός ορίζοντας των προβλέψεων είναι το 1 έτος (12 μήνες), και εφαρμόζονται δύο διαφορετικές μέθοδοι βασισμένες στην παλινδρόμηση, ένα υβριδικό μοντέλο (stochastic-regression model) και ένα μοντέλο αυτοπαλινδρόμησης (18-lagged Auto Regressive model), χρησιμοποιώντας τα μετασχηματισμένα δεδομένα που αναφέραμε σε συνδυασμό με άλλους παράγοντες ,όπως τα πρώτα παράγωγα της υδροηλεκτρικής παραγωγής, που επηρεάζουν την τμη της ΟΤΣ.

Ως μέτρο σύγκρισης επιλέγονται το Απόλυτο Ποσοστιαίο Σφάλμα (APE), το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE) και η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE), καθώς και η τυπική απόκλιση (STD) και διασπορά (τετραγωνική ρίζα της τυπικής απόκλισης).

Τα αποτελέσματα, θεωρούνται μάλλον ενθαρρυντικά, κυρίως όσον αφορά το υβριδικό μοντέλο του οποίου το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE) ήταν σημαντικό μικρότερο από αυτό της μεθόδου αυτοπαλινδρόμησης (9.67% έναντι 28,67%), και επιπλέον το υβριδικό μοντέλο παρήγαγε ελαφρώς καλύτερα αποτελέσματα στην περίπτωση που λήφθηκαν υπόψη τα μέσα ετήσια δεδομένα υδροηλεκτρικής παραγωγής για κάθε μήνα και όχι τα πραγματικά μηνιαία δεδομένα.

2.4.3 Μέθοδοι για Μακροπρόθεσμη Πρόβλεψη

Η συγκεκριμένη κατηγορία, παρουσιάζει παρόμοια χαρακτηριστικά με την προηγούμενη, με τη μόνη διαφορά ότι η αβεβαιότητα για την διαμόρφωση της ΟΤΣ και των παραγόντων που την επηρεάζουν είναι ακόμη μεγαλύτερη. Επιπλέον, επειδή πλέον ο ορίζοντας μας είναι μεγαλύτερος από το 1 έτος, είναι σαφές ότι τα ιστορικά δεδομένα για την ΟΤΣ είναι πλέον ανεπαρκή και τα αποτελέσματα ιδιαίτερα αναξιόπιστα.

Η σπουδαιότητα της Μακροπρόθεσμης Πρόβλεψης της ΟΤΣ είναι δεδομένη, και δεν αποτελεί απλά αντικείμενο Ακαδημαϊκής έρευνας αλλά έχει σοβαρό αντίκτυπο στην Αγορά Ενέργειας, βοηθώντας κυρίως στην κατανόηση και αξιολόγηση των επενδυτικών ρίσκων. Πολλές αποφάσεις λαμβάνονται από τους συμμετέχοντες με γνώμονα το απώτερο μέλλον, μιλώντας ακόμη και για 20 έως 50 χρόνια μπροστά.

Σε αρκετές περιπτώσεις οι Ρυθμιστικές Αρχές των Αγορών Ενέργειας ζητούν από τους συμμετέχοντες τον προγραμματισμό των πόρων τους σε ένα χρονικό ορίζοντα 20-30 ετών, και επιπλέον πολλά συμβόλαια για καύσιμα έχουν ισχύ για 20 ή και περισσότερα χρόνια (*G. Hamm & A. Borison, 2006*).

Στην παραπάνω έρευνα αναφέρονται οι 4 βασικές πηγές δεδομένων που απαιτούνται για την παραγωγή μακροπρόθεσμων προβλέψεων και είναι οι ακόλουθες:

- Ιστορικά δεδομένα τιμών Αγοράς Ενέργειας.
- Μελλοντικές τιμές Αγοράς Ενέργειας.

- Αποτελέσματα για ζήτηση/προσφορά από μοντέλα προσομοίωσης .
- Εκτιμήσεις των ειδικών, ιδίως σε σχέση με τις μελλοντικές τεχνολογίες και τους κανονισμούς.

Επιπλέον αναφέρονται δύο βασικοί λόγοι που οι προσπάθειες για παραγωγή μακροπρόθεσμων προβλέψεων δεν έχουν αποδώσει τα αναμενόμενα αποτελέσματα, οι οποίοι είναι:

- Οι ερευνητές βασίζονται σχεδόν αποκλειστικά είτε σε οικονομικά δεδομένα (ιστορικές και μελλοντικές τιμές) είτε σε τεχνικά δεδομένα (μοντέλα προσομοίωσης, κρίση των ειδικών), αντί να τα χρησιμοποιούν από κοινού ως πηγές πληροφόρησης.

- Ανεξάρτητα από το είδος των δεδομένων που χρησιμοποιούνται ,οι ερευνητές εστιάζονται τόσο πολύ στο παρελθόν και το παρών, βασίζοντας τις προβλέψεις στην προέκταση των υπαρχουσών μοντέλων, δίχως να σκέφτονται δημιουργικά και ολοκληρωμένα για το μέλλον.

Καθότι η παρούσα διπλωματική έχει ως αντικείμενο κυρίως την παραγωγή και μελέτη βραχυπρόθεσμων/μεσοπρόθεσμων προβλέψεων, η περαιτέρω μελέτη των μακροπρόθεσμων προβλέψεων δεν κρίνεται απαραίτητη.

Κεφάλαιο 3. Χαρακτηριστικά Χρονοσειρών

Απαραίτητη προϋπόθεση για την παραγωγή όσο το δυνατόν ακριβέστερων προβλέψεων, είναι η κατανόηση και ανάλυση των ιδιαίτερων χαρακτηριστικών κάθε χρονοσειράς, αλλά και η επιλογή των κατάλληλων μέτρων σύγκρισης, προκειμένου να αποφανθούμε για το κατά πόσο η πρόβλεψη που παράχθηκε είναι αποδεκτή. Βασικό εργαλείο για αυτό το σκοπό είναι τα διάφορα σφάλματα τα οποία και θα εξεταστούν στο συγκεκριμένο κεφάλαιο, ενώ θα γίνει αναφορά και στα ποιοτικά αλλά και ποσοτικά χαρακτηριστικά των χρονοσειρών.

3.1 Ποιοτικά Χαρακτηριστικά των Χρονοσειρών

Τα βασικά χαρακτηριστικά που προκύπτουν από την αποσύνθεση μιας χρονοσειράς είναι τα εξής:

1. Τάση.
2. Κυκλικότητα.
3. Εποχικότητα.
4. Ασυνέχειες - Μη κανονικές Διακυμάνσεις.

Στη συνέχεια ακολουθεί μια πιο αναλυτική περιγραφή σχετικά με τη σημασία κάθε ποιοτικού χαρακτηριστικού και πώς αυτό επηρεάζει την διαδικασία πρόβλεψης μιας χρονοσειράς.

1. Τάση (Trend)

Το μέγεθος της τάσης ορίζεται ως μία ¨μακροπρόθεσμη¨ μεταβολή του μέσου επιπέδου των τιμών μιας χρονοσειράς. Όμως θα πρέπει να εξεταστεί πότε μπορεί να θεωρηθεί μία μεταβολή ως μακροπρόθεσμη. Για να προκύψει ένα ασφαλές συμπέρασμα σχετικά με την τάση μιας χρονοσειράς, θα πρέπει να έχουμε στη διάθεση μας έναν ικανό αριθμό ιστορικών δεδομένων και παράλληλα να θεωρήσουμε κατάλληλο μήκος για την περίοδο μέσα στην οποία θα αναζητήσουμε την ύπαρξη τάσης.

2. Κυκλικότητα (Cycle)

Το συγκεκριμένο μέγεθος εκφράζει μια κυματοειδή μεταβολή που συνήθως οφείλεται σε ειδικές εξωγενείς συνθήκες και εμφανίζεται κατά περιόδους. Οι περίοδοι δεν είναι πάντοτε σταθερές και τις περισσότερες φορές έχουν μήκος μεγαλύτερο του έτους. Η κυκλικότητα εμφανίζεται κυρίως σε χρονοσειρές οικονομικών μεγεθών, όπως είναι το Ακαθάριστο Εγχώριο Προϊόν (ΑΕΠ), οι δείκτες της βιομηχανικής παραγωγής καθώς και οι τιμές των μετοχών. Οι συγκεκριμένες κυκλικές μεταβολές οφείλονται στις γενικότερες επικρατούσες οικονομικές συνθήκες και συχνά αναφέρονται με την ονομασία «επιχειρηματικός κύκλος».

3. Εποχικότητα (Seasonality)

Ως εποχικότητα ορίζεται η παρουσία μίας περιοδικής διακύμανσης, συνήθως κατανοητής και προβλέψιμης, με σταθερό και μικρότερο του έτους μήκος. Ουσιαστικά πρόκειται για αλλαγές που επαναλαμβάνονται με τον ίδιο ακριβώς τρόπο μέσα στο χρόνο. Για παράδειγμα, η κατανάλωση ενέργειας είναι αυξημένη τους καλοκαιρινούς μήνες λόγω της χρήσης κλιματιστικών. Επομένως η απομόνωση του εποχιακού παράγοντα αποτελεί τις περισσότερες φορές εύκολο εγχείρημα, με εξαίρεση ίσως τις χρονοσειρές που παρουσιάζουν πολλαπλή εποχικότητα, όπως για παράδειγμα στην περίπτωσή μας η ΟΤΣ. Η χρονοσειρά που προκύπτει απαλείφοντας τον εποχιακό παράγοντα καλείται αποεποχικοποιημένη χρονοσειρά.

4. Ασυνέχειες – Μη κανονικές Διακυμάνσεις

Ασυνέχειες καλούνται ορισμένες απομονωμένες παρατηρήσεις που παρουσιάζονται στο γράφημα μίας χρονοσειράς και ξεφεύγουν από το γενικό πρότυπο συμπεριφοράς της. Τέτοιου είδους τιμές δεν είναι δυνατό να προβλεφθούν με βάση τα ιστορικά δεδομένα και μπορούν να έχουν είτε μόνιμο είτε παροδικό χαρακτήρα. Εάν έχουν παροδικό χαρακτήρα ονομάζονται outliers ή special events ή spikes και επιδρούν στη χρονοσειρά για μικρό χρονικό διάστημα. Αυτού του είδους οι ασυνέχειες είναι δυνατό να εμφανιστούν λόγω κάποιου απρόβλεπτου εξωτερικού γεγονότος. Στην περίπτωση που έχουν μόνιμο χαρακτήρα καλούνται level-shifts, αφού εμφανίζονται ως απότομες αλλαγές στο επίπεδο της χρονοσειράς. Ένα παράδειγμα μόνιμης ασυνέχειας είναι η πτώση των πωλήσεων μιας εταιρίας λόγω της εμφάνισης μιας ανταγωνίστριας εταιρίας στην αγορά.

Οι μη κανονικές διακυμάνσεις, είναι εκείνες οι οποίες απομένουν μετά την απομόνωση όλων των λοιπών χαρακτηριστικών μιας χρονοσειράς. Οι διακυμάνσεις αυτές μπορεί να εκφράζουν είτε τον τυχαίο παράγοντα μιας στοχαστικής διαδικασίας, είτε κάποια ασυνέχεια που οφείλεται σε κάποιο ιδιαίτερο γεγονός.

3.2 Ποσοτικά Χαρακτηριστικά των Χρονοσειρών

Πρόκειται για ορισμένους βασικούς στατιστικούς δείκτες, οι οποίοι μπορούν να υπολογιστούν εύκολα για οποιαδήποτε δεδομένη χρονοσειρά με συγκεκριμένο πλήθος παρατηρήσεων, και είναι οι ακόλουθοι:

i. Μέση τιμή (Average)

Είναι ο απλός γραμμικός μέσος όρος των τιμών των παρατηρήσεων μιας χρονοσειράς και υποδεικνύει το επίπεδο γύρω από το οποίο κινούνται οι πραγματικές τιμές της χρονοσειράς. Ο υπολογισμός του συγκεκριμένου δείκτη μπορεί να γίνει μέσω υπολογιστικού φίλου με τη συνάρτηση average ή με χρήση του ακόλουθου τύπου:

$$\bar{Y} = \frac{1}{n} \bullet \sum_{i=1}^{n} Y_i$$

όπου:

- Y_i είναι η i-οστή παρατήρηση,
- n είναι το πλήθος των παρατηρήσεων.

ii. Μέγιστη και Ελάχιστη Τιμή (Maximum and Minimum)

Πρόκειται για τις ακραίες τιμές μιας χρονοσειράς και μπορούν να αποτελέσουν μία εκτίμηση της διακύμανσης των δεδομένων καθώς και της τυχαιότητάς τους. Συνήθως, όταν η εξεταζόμενη χρονοσειρά αποτελείται από μεγάλο πλήθος δεδομένων, η μέγιστη και ελάχιστη τιμή της υπολογίζονται μέσω υπολογιστικού φύλλου με τις συναρτήσεις max και min αντίστοιχα.

iii. Τυπική Απόκλιση (Standard Deviation)

Ο συγκεκριμένος στατιστικός δείκτης αναφέρεται και ως διασπορά και αντιπροσωπεύει το κατά πόσο οι παρατηρήσεις είναι διεσπαρμένες γύρω από τη μέση τιμή. Και αυτός ο δείκτης μπορεί να υπολογιστεί μέσω υπολογιστικού φύλλου, αλλά και με χρήση των παρακάτω σχέσεων για δείγμα ή πληθυσμό αντίστοιχα:

$$\sigma_{\delta\varepsilon\acute{\iota}\gamma\mu\alpha\tau o\varsigma} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}}$$

$$\sigma_{\pi\lambda\eta\theta\upsilon\sigma\mu o\acute{\upsilon}} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n}}$$

iv. Διακύμανση (Variance)

Το μέγεθος αυτό ορίζεται απλά ως το τετράγωνο της τυπικής απόκλισης και μπορεί να υπολογιστεί μέσω υπολογιστικού φύλλου, ανάλογα με το αν αναφερόμαστε σε δείγμα ή πληθυσμό.

v. Συντελεστής Γραμμικής Συσχέτισης (Linear Correlation Coefficient)

Είναι γνωστός και ως συντελεστής του Pearson και εκφράζει το κατά πόσο είναι συγκεντρωμένες οι τιμές ενός διαγράμματος διασποράς γύρω από την ευθεία παλινδρόμησης δίνοντας παράλληλα ένα μέτρο της γραμμικής συσχέτισης μεταξύ των δύο υπό εξέταση μεταβλητών. Ο εν λόγω δείκτης μπορεί να πάρει τιμές στο διάστημα [-1, 1]. Όταν ο συντελεστής παίρνει τις ακραίες τιμές ± 1 υπάρχει τέλεια γραμμική συσχέτιση μεταξύ των δύο μεταβλητών. Όσο η τιμή του συντελεστή λαμβάνει τιμές κοντά στις ακραίες, τόσο μεγαλύτερη είναι η γραμμική συσχέτιση, ενώ στην αντίθετη περίπτωση η γραμμική συσχέτιση θεωρείται ασθενής. Ο υπολογισμός του συντελεστή γραμμικής συσχέτισης μπορεί να γίνει μέσω υπολογιστικού φύλου με τη συνάρτηση correl ή με χρήση του ακόλουθου τύπου, θεωρώντας ως μία μεταβλητή το χρόνο:

$$r_{XY} = \frac{\sum_{i=1}^{n}[(X_\iota - \bar{X}) \bullet (Y_\iota - \bar{Y})]}{\sqrt{\sum_{\iota=1}^{n}(X_\iota - \bar{X})^2} \bullet \sqrt{\sum_{\iota=1}^{n}(Y_\iota - \bar{Y})^2}}$$

vi. Συντελεστής Αυτοσυσχέτισης (Autocorrelation Coefficient)

Υποδεικνύει τη συσχέτιση μεταξύ παρατηρήσεων της ίδιας μεταβλητής με χρονική υστέρηση k περιόδους. Ο συγκεκριμένος συντελεστής παίρνει τιμές στο διάστημα [0, 1], και αν λάβει τιμή κοντά στο μηδέν, θεωρείται πως η συσχέτιση των παρατηρήσεων χρονικής υστέρησης k είναι μηδενική, ενώ αν λάβει τιμή κοντά στη μονάδα, τα δεδομένα μας παρουσιάζουν υψηλή συσχέτιση. Πρόκειται για ένα πολύ χρήσιμο στατιστικό δείκτη, ο οποίος χρησιμοποιείται σε περιπτώσεις που απαιτείται αλγεβρικός εντοπισμός της εποχιακής συμπεριφοράς μιας χρονοσειράς. Μπορεί να υπολογιστεί είτε μέσω υπολογιστικού φύλλου με τη συνάρτηση acf, είτε με χρήση του παρακάτω τύπου:

$$ACF_k = \frac{\sum_{i=1+k}^{n}[(Y_i - \bar{Y}) \bullet (Y_{i-k} - \bar{Y})]}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

vii. Συντελεστής Μεταβλητότητας (Coefficient of Variation)

Πρόκειται για ένα κανονικοποιημένο μέτρο της διασποράς των παρατηρήσεων ενός δείγματος ή ενός πληθυσμού, και σε σύγκριση με την απλή τυπική απόκλιση έχει ως βασικό πλεονέκτημα ότι δεν επηρεάζεται από το επίπεδο των παρατηρήσεων. Εντούτοις δεν μπορεί να υπολογισθεί στην περίπτωση που η μέση τιμή είναι μηδενική. Ο συντελεστής μεταβλητότητας μπορεί να υπολογισθεί με χρήση του παρακάτω τύπου:

$$C_v = \frac{\sigma}{\bar{Y}} \bullet 100\%$$

viii. Μέση τιμή διαστήματος μεταξύ ζητήσεων (Intermittent Demand Interval)

Ο εν λόγω στατιστικός δείκτης αναφέρεται σε χρονοσειρές διακοπτόμενης ζήτησης και εκφράζει τη μέση τιμή των αποστάσεων διαδοχικών περιόδων με μη μηδενική τιμή. Στην περίπτωση που ο δείκτης ισούται με τη μονάδα, η χρονοσειρά είναι συνεχούς ζήτησης, ενώ σε περίπτωση που παίρνει τιμές μεγαλύτερες της μονάδας, αναφερόμαστε σε χρονοσειρά διακοπτόμενης ζήτησης. Όσο μεγαλύτερη είναι η τιμή του συγκεκριμένου δείκτη, τόσο μεγαλύτερα είναι τα μεσοδιαστήματα μεταξύ μη μηδενικών παρατηρήσεων.

3.3 Ανάλυση Ακρίβειας Προβλέψεων (Σφάλματα)

Όπως αναφέρθηκε και στην αρχή αυτού του κεφαλαίου, η ακρίβεια μιας πρόβλεψης είναι αυτή που καθορίζει το κατά πόσο αυτή είναι αποδεκτή και μας παρέχει όσο το δυνατόν ασφαλέστερα αποτελέσματα. Κυρίαρχο ρόλο στην ανάλυση της ακρίβειας των προβλέψεων παίζουν τα σφάλματα.

Ως σφάλμα ορίζεται η διαφορά μεταξύ της πραγματικής και της προβλεπόμενης τιμής για μία περίοδο. Η μαθηματική διατύπωση του παραπάνω ορισμού είναι η ακόλουθη:

$$e_i = Y_i - F_i$$

όπου:

- i είναι η εξεταζόμενη περίοδος,

- e_i είναι το σφάλμα,
- Y_i είναι η πραγματική τιμή του μεγέθους,
- F_i είναι η πρόβλεψη της τιμής του μεγέθους.

Όμως η παραπάνω απλή σχέση δεν είναι σε θέση να κρίνει την αποδοτικότητα των μεθόδων πρόβλεψης, γι αυτό και συνίσταται η χρήση πιο ολοκληρωμένων στατικών δεικτών σφάλματος, τους οποίους και παρουσιάζουμε στη συνέχεια.

3.3.1 Μέσο Σφάλμα (Mean Error)

Το μέσο σφάλμα (Mean Error), υπολογίζεται ως ο απλός προσημασμένος μέσος όρος των σφαλμάτων και εκφράζει ένα μέτρο συστηματικότητας του σφάλματος. Όταν το μέσο σφάλμα παίρνει τιμές κοντά στο μηδέν τότε τα σφάλματα θεωρούνται τυχαία και όχι συστηματικά. Στην περίπτωση που λαμβάνει θετικές τιμές, είναι σαφής η ύπαρξη απαισιοδοξίας στις προβλέψεις, ενώ όταν παίρνει αρνητικές τιμές, οι προβλέψεις θεωρούνται αισιόδοξες. Λόγο της σχέσης που έχει ο συγκεκριμένος δείκτης με την προκατάληψη, σχυνά καλείται και bias, και δίνεται από τον ακόλουθο τύπο:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)$$

3.3.2 Μέσο Απόλυτο Σφάλμα (Mean Absolute Error)

Ο συγκεκριμένος δείκτης εκφράζει ένα μέτρο της ακρίβειας της πρόβλεψης έναντι των πραγματικών τιμών διατηρώντας τις αρχικές μονάδες της χρονοσειράς. Μεγάλη τιμή του δείκτη σημαίνει μικρή ακρίβεια της μεθόδου πρόβλεψης που εφαρμόστηκε. Η μαθηματική διατύπωση αυτού του στατιστικού δείκτη είναι η εξής:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - F_i|$$

3.3.3 Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error)

Το μέσο τετραγωνικό σφάλμα είναι ένας στατιστικός δείκτης που λειτουργεί ως μέτρο της ακρίβειας της πρόβλεψης, δίνοντας όμως περισσότερο βάρος στα μεγάλα σφάλματα και λιγότερο στα μικρά (λογικό αφού ο τύπος είναι υψωμένος στο τετράγωνο). Ο συγκεκριμένος δείκτης χρησιμοποιείται κυρίως ως κριτήριο επιλογής των κατάλληλων παραμέτρων εξομάλυνσης και υπολογίζεται από τον ακόλουθο τύπο:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)^2$$

3.3.4 Ρίζα Μέσου Τετραγωνικού Σφάλματος (Root Mean Squared Error)

Έχει όμοιες ιδιότητες με τον προηγούμενο δείκτη, μόνο που είναι εκφρασμένος σε μονάδες της αρχικής χρονοσειράς. Ο τύπος που περιγράφει αυτό το στατιστικό δείκτη είναι ο παρακάτω:

$$RMSE = \sqrt{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - F_i)^2}$$

3.3.5 Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Mean Absolute Percentage Error)

Ο δείκτης του μέσου απόλυτου ποσοστιαίου σφάλματος είναι ιδιαίτερα χρήσιμος σε περιπτώσεις όπου οι πραγματικές τιμές (Y) είναι σημαντικά υψηλές. Είναι εκφρασμένος επί τις εκατό και μπορεί να πάρει τιμές ίσες ή μεγαλύτερες του μηδενός. Αξίζει να αναφερθεί πως όσο μικρότερη είναι η τιμή του συγκεκριμένου δείκτη, τόσο καλύτερα αποδίδει η μέθοδος πρόβλεψης που εξετάζουμε. Η σχέση που περιγράφει το μέσο απόλυτο ποσοστιαίο σφάλμα είναι η εξής:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{Y_i - F_i}{Y_i} \right| \bullet 100\ (\%)$$

Ο δείκτης του MAPE δεν είναι δυνατό να εφαρμοστεί σε χρονοσειρές διακοπτόμενης ζήτησης (χρονοσειρές με μεγάλο πλήθος μηδενικών), διότι καταλήγει σε απροσδιοριστία.

3.3.6 Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (Symmetric Mean Absolute Percentage Error)

Πρόκειται ουσιαστικά για μία παραλλαγή του μέσου απόλυτου ποσοστιαίου σφάλματος και υπολογίζεται σύμφωνα με τον ακόλουθο τύπο:

$$sMAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{Y_i - F_i}{\left(\frac{Y_i + F_i}{2}\right)} \right| \bullet 100 = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{2 \bullet (Y_i - F_i)}{Y_i + F_i} \right| \bullet 100\ (\%)$$

Από του δείκτες που παρουσιάστηκαν σε αυτό το κεφάλαιο, σημαντικότεροι για τον πρόβλημα πρόβλεψης της ΟΤΣ, μπορούν να θεωρηθούν ο δείκτης του Μέσου Απόλυτου Ποσοστιαίου Σφάλματος (MAPE) και ο δείκτης της Ρίζας Μέσου Τετραγωνικού Σφάλματος.

Η άποψη αυτή στηρίζεται στο γεγονός πως η πλειοψηφία των ερευνητών που ασχολήθηκαν με το συγκεκριμένο θέμα, έθεσαν ως μέτρα αξιολόγησης αυτούς τους δύο στατιστικούς δείκτες, γεγονός που μας οδηγεί να κινηθούμε και εμείς προς αυτήν την κατεύθυνση. Η αξιολόγηση των σφαλμάτων που θα χρησιμοποιηθούν στην παρούσα διπλωματική, θα αναλυθεί εκτενώς στο πρακτικό μέρος της εργασίας.

Κεφάλαιο 4. Μοντέλα Πρόβλεψης

Στα πλαίσια της διπλωματικής αυτής, θα επικεντρωθούμε στη χρήση κυρίως μοντέλων χρονοσειρών (μέθοδοι εξομάλυνσης) και αιτιοκρατικών μοντέλων (απλή ή πολλαπλή γραμμική παλινδρόμηση) για την παραγωγή βραχυπρόθεσμων προβλέψεων της ΟΤΣ. Ως μέθοδος αναφοράς (benchmark) θα χρησιμοποιηθεί η μέθοδος Naive, και επιπλέον θα επιχειρήσουμε να προσεγγίσουμε το συγκεκριμένο πρόβλημα με τη συνδυαστική μέθοδο Theta.

4.1 Μέθοδος Αναφοράς - Naive

Πρόκειται για την απλούστερη στατιστική μέθοδο, η οποία ενώ δεν παράγει ακριβείς προβλέψεις, χρησιμοποιείται από τους ερευνητές ως μέθοδος αναφοράς (benchmark), κάτι το οποίο θα εφαρμόσουμε κι εμείς σε αυτή τη διπλωματική. Η πρόβλεψη που προκύπτει από τη μέθοδο Naive για μια χρονική στιγμή t είναι ίση με την πραγματική παρατήρηση της ακριβώς προηγούμενης χρονικής περιόδου (t-1) (*Πετρόπουλος & Ασημακόπουλος, 2011*):

$$F_t = Y_{t-1}$$

Στην περίπτωση που θέλουμε να παράγουμε περισσότερες από μία σημειακές προβλέψεις (ορίζοντας m>1), η πρόβλεψη για κάθε χρονική στιγμή στο μέλλον ισούται με την τελευταία διαθέσιμη ιστορική παρατήρηση και η αντίστοιχη εξίσωση που περιγράφει τη μέθοδο είναι:

$$F_{t+m} = Y_t$$

Ο αναλυτικός αλγόριθμος που περιγράφει τη μέθοδο Naive, παρουσιάζεται στο Παράρτημα Matlab.

4.2 Μέθοδοι Εκθετικής Εξομάλυνσης

Η συγκεκριμένη κατηγορία μεθόδων πρόβλεψης θεωρείται από τις πιο διαδεδομένες στους επιχειρηματικούς κύκλους, γεγονός το οποίο οφείλεται στην μαθηματική ευκολία τους, στην ελάχιστη απαίτηση σε υπολογιστικό χρόνο και στην ανάγκη για σχετικά μικρό αριθμό παρατηρήσεων ούτως ώστε να παράγουν προβλέψεις. Οι μέθοδοι εξομάλυνσης θεωρούνται καταλληλότερες για βραχυπρόθεσμες και μεσοπρόθεσμες προβλέψεις μεγάλου όγκου χρονοσειρών και παρουσιάζουν καλύτερη απόδοση σε δεδομένα με στασιμότητα ή μικρό ρυθμό ανάπτυξης. Εμπειρικές μελέτες έχουν δείξει ότι οι μέθοδοι εκθετικής εξομάλυνσης αποδίδουν καλύτερα όσον αφορά τα ποσοστά ακρίβειας σε σύγκριση με πιο περίπλοκες μεθόδους πρόβλεψης, κάτι το οποίο οφείλεται στο ότι δεν επηρεάζονται από τις ιδιομορφίες των προτύπων των δεδομένων ή από περιστασιακά εμφανιζόμενες ακραίες τιμές (spikes) που εμφανίζονται συχνά σε επιχειρησιακά δεδομένα. Στην παρούσα διπλωματική θα χρησιμοποιηθούν 3 από τις πιο βασικές μεθόδους εξομάλυνσης οι οποίες είναι οι ακόλουθες:

- Απλή Εκθετική Εξομάλυνση σταθερού επιπέδου (Simple Exponential Smoothing).
- Εκθετική Εξομάλυνση Γραμμικής Τάσης (Holt Exponential Smoothing).
- Εκθετική Εξομάλυνση μη γραμμικής ή φθίνουσας τάσης (Damped Exponential Smoothing).

4.2.1 Απλή Εκθετική Εξομάλυνση σταθερού επιπέδου (SES)

Το συγκεκριμένο μοντέλο σταθερού επιπέδου, το οποίο είναι γνωστό και ως απλή εκθετική εξομάλυνση (Simple Exponential Smoothing), περιγράφεται από τις ακόλουθες εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + \alpha \bullet e_t$$

$$F_{t+1} = S_t$$

όπου:

- e_t δηλώνει το σφάλμα, δηλαδή την απόκλιση της πραγματικής τιμές από την πρόβλεψη,
- S δηλώνει το επίπεδο,
- F δηλώνει την πρόβλεψη,
- t δηλώνει τη χρονική περίοδο,
- α είναι ο συντελεστής εξομάλυνσης της μεθόδου, ο οποίος λαμβάνει τιμές στο διάστημα [0,1].

Από τις παραπάνω σχέσεις που περιγράφουν την απλή εκθετική εξομάλυνση, καθίσταται σαφές ότι είναι απαραίτητο να οριστεί ένα **αρχικό επίπεδο** (S_0)

ώστε να μπορέσει να ξεκινήσει η διαδικασία υπολογισμού του μοντέλου πρόβλεψης αλλά και των ζητούμενων σημειακών προβλέψεων δοθέντος κάποιου χρονικού ορίζοντα. Υπάρχουν 4 τρόποι για να υπολογίσουμε το αρχικό επίπεδο, η επιλογή των οποίων εξαρτάται από τα ποιοτικά χαρακτηριστικά της χρονοσειράς που εξετάζουμε. Ως σταθερό επίπεδο μπορούμε να θεωρήσουμε ένα από τα ακόλουθα:

1. Το μέσο όρο όλων των παρατηρήσεων.
2. Το μέσο όρο των ν πρώτων παρατηρήσεων.
3. Την πρώτη παρατήρηση.
4. Το σταθερό επίπεδο από το μοντέλο απλής γραμμικής παλινδρόμησης (LRL).

Η ορθή επιλογή του σταθερού επιπέδου είναι σημαντική, διότι αφ’ ενός επηρεάζει την επιλογή του συντελεστή εξομάλυνσης, αφ’ ετέρου αν το σταθερό επίπεδο που θα επιλέξουμε δεν είναι αντιπροσωπευτικό των δεδομένων μας, θα προκαλέσει αρνητική επίδραση στην παραγωγή των προβλέψεων. Αξίζει να σημειωθεί πως στο αρχικό επίπεδο αποτελεί ουσιαστικά και την αρχική πρόβλεψη (F_1).

Ιδιαίτερα σημαντικό ρόλο στο συγκεκριμένο μοντέλο πρόβλεψης παίζει ο συντελεστής εξομάλυνσης (α). Η επιλογή του βέλτιστου συντελεστή εξομάλυνσης καθορίζεται από δύο αλληλοεξαρτώμενους παράγοντες, το ποσοστό του θορύβου στη χρονοσειρά και τη σταθερότητα του μέσου όρου της χρονοσειράς. Όσον αφορά τον πρώτο παράγοντα, όσο περισσότερος είναι ο θόρυβος που εμφανίζεται στα δεδομένα τόσο μικρότερη πρέπει να είναι η τιμή του συντελεστή εξομάλυνσης, προκειμένου να αποφευχθεί υπερβολική αντίδραση στο θόρυβο.

Σχετικά με το δεύτερο παράγοντα, αν ο μέσος όρος παρουσιάζει μεταβολές, ο συντελεστής εξομάλυνσης πρέπει να λάβει μεγάλες τιμές ώστε να μπορούν οι προβλέψεις να ακολουθούν τις μεταβολές των δεδομένων. Στην περίπτωση που ο μέσος όρος είναι σχετικά σταθερός, λαμβάνεται μικρότερη τιμή του α.

Η βελτιστοποίηση της απλής εκθετικής εξομάλυνσης βασίζεται στην εύρεση του βέλτιστου συντελεστή εξομάλυνσης. Η συνηθέστερη διαδικασία που ακολουθείται είναι η γραμμική αναζήτηση η οποία ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα (MSE), θέτοντας από την αρχή την ακρίβεια με την οποία θα γίνει η αναζήτηση καθώς και τα επιθυμητά όρια μέσα στα οποία θα κινηθεί ο συντελεστής α.

Πέρα της γραμμικής αναζήτησης, μπορεί να χρησιμοποιηθεί και η μη γραμμική ή δυαδική αναζήτηση για τη βελτιστοποίηση της μεθόδου. Πρόκειται για έναν αναδρομικό αλγόριθμο αναζήτησης ο οποίος χρησιμοποιεί την τεχνική διαίρει και βασίλευε. Η συγκεκριμένη μέθοδος εφαρμόζεται μόνο σε ταξινομημένους πίνακες, σε αντίθεση με τη γραμμική αναζήτηση, γεγονός που μειώνει σημαντικά την πολυπλοκότητα των υπολογισμών. Το ζητούμενο στη δυαδική αναζήτηση είναι να εντοπιστεί μία συγκεκριμένη τιμή που ονομάζεται κλειδί, στην περίπτωση φυσικά που αυτό είναι δυνατό, δηλαδή αυτή η τιμή υπάρχει μέσα στον πίνακα που εξετάζουμε.

Χρησιμοποιώντας την δυαδική αναζήτηση, ζητούμενο είναι ξανά η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE). Αρχικά η παράμετρος α εκφράζεται μέσω ενός πίνακα διαστάσεων 1Χ2. Ως ιδανικό α, θεωρείται το πρώτο στοιχείο αυτού του πίνακα, και ως ιδανικό MSE, το αντίστοιχο που προκύπτει από τη συγκεκριμένη τιμή του α.

Ο αλγόριθμος ξεκινάει, υπολογίζοντας για κάθε τιμή του α το αντίστοιχο MSE, και εν συνεχεία το συγκρίνει με το ιδανικό MSE, το οποίο είναι και η τιμή κλειδί. Σε περίπτωση που προκύψει μικρότερο MSE, τότε αυτό παίρνει τη θέση του κλειδιού, και κρατείται το α από το οποίο προέκυψε αυτό το κλειδί. Στη συνέχεια, οι τιμές του αρχικού πίνακα μεταβάλλονται κατά τη διαφορά του πρώτου στοιχείου διαιρούμενο με τον αριθμό 2.

Συγκεκριμένα, πλέον το πρώτο στοιχείο του πίνακα, δηλαδή το α(1), ισούται με την προηγούμενη τιμή του α(1) μείον τη διαφορά που περιγράφηκε, ενώ το νέο στοιχείο α(2) ισούται με το παλιό συν τη διαφορά. Επιπλέον, η νέα διαφορά γίνεται μισή της προηγούμενης, και ο αλγόριθμος συνεχίζει από την αρχή υπολογίζοντας και συγκρίνοντας ξανά τα MSE, σύμφωνα με τις νέες τιμές που προέκυψαν για την παράμετρο α.

Όπως αναφέρθηκε, η δυαδική αναζήτηση παρουσιάζει μικρότερη υπολογιστική πολυπλοκότητα από τη γραμμική, γεγονός το οποίο την καθιστά πολύ πιο γρήγορη. Για το λόγο αυτό, και λόγω του μεγάλου όγκου των δεδομένων που χρησιμοποιούνται, η μέθοδος της μη γραμμικής βελτιστοποίησης προτιμάται στην παρούσα διπλωματική και ο αλγόριθμος αυτής αλλά και συνολικά της μεθόδου παρουσιάζεται στο Παράρτημα Matlab.

Αξίζει να αναφερθεί πως όταν ο συντελεστής εξομάλυνσης λάβει την ακραία τιμή 1, τότε το μοντέλο ταυτίζεται με την μέθοδο Naive, ενώ αν είναι α=0 ,κάθε πρόβλεψη είναι ίση με το αρχικό επίπεδο.

Επιπλέον, αν ζητείται ο υπολογισμός περισσότερων από μίας σημειακής πρόβλεψης, τότε όλες οι επόμενες θέτονται ίσες με την τελευταία υπολογισμένη πρόβλεψη (flat forecast) καθώς το μοντέλο θεωρεί την ύπαρξη ενός σταθερού επιπέδου (*Πετρόπουλος & Ασημακόπουλος, 2011*).

4.2.2 Μοντέλο Γραμμικής Τάσης (Holt Exponential Smoothing)

Το μοντέλο αυτό αποτελεί μία επέκταση του μοντέλου της απλής εκθετικής εξομάλυνσης, η οποία έχει τη επιπλέον δυνατότητα να διαχειριστεί τη συνιστώσα της τάσης, μέγεθος το οποίο παρατηρείται συχνά στην πλειοψηφία των επιχειρησιακών δεδομένων. Η μέθοδος πήρε το όνομα της το οποίο διατηρεί μέχρι και σήμερα από τον Holt, ο οποίος πρότεινε τη συγκεκριμένη επέκταση το 1957. Η μέθοδος Holt περιγράφεται από τις ακόλουθες εξισώσεις:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + \alpha \bullet e_t$$

$$T_t = T_{t-1} + \beta \bullet e_t$$

$$F_{t+m} = S_t + m \bullet T_t$$

όπου:

- e_t δηλώνει το σφάλμα, δηλαδή την απόκλιση της πραγματικής τιμής από την πρόβλεψη,
- S δηλώνει το επίπεδο,
- F δηλώνει την πρόβλεψη,
- t δηλώνει τη χρονική περίοδο,
- T δηλώνει την τάση,
- m δηλώνει το χρονικό ορίζοντα της πρόβλεψης,
- α είναι ο συντελεστής εξομάλυνσης της μεθόδου για το επίπεδο, ο οποίος λαμβάνει τιμές στο διάστημα [0,1],
- β είναι ο συντελεστής εξομάλυνσης της μεθόδου για την τάση, ο οποίος λαμβάνει τιμές στο διάστημα [0,1].

Όπως και στην απλή εκθετική εξομάλυνση, έτσι και στην μέθοδο Holt απαιτείται ο προσδιορισμός ενός **αρχικού επιπέδου** (S_0) αλλά και μίας **αρχικής τάσης** (T_0), έτσι ώστε να είναι δυνατή η εκκίνηση της διαδικασίας υπολογισμού του μοντέλου πρόβλεψης αλλά και η παραγωγή των ζητούμενων σημειακών προβλέψεων δοθέντος κάποιου χρονικού ορίζοντα.

Το αρχικό επίπεδο προσδιορίζεται όπως και στην μέθοδο SES που αναφέρθηκε στην ενότητα 4.2.1. Η αρχική τάση υπολογίζεται ως:

1. Η διαφορά δεύτερης και πρώτης παρατήρησης (Y_2 – Y_1).

2. Η διαφορά ν-οστής και πρώτης παρατήρησης διαιρεμένης με ν-1 (πχ. (Y_8 – Y_1)/7).

3. Η σταθερά της κλίσης από το μοντέλο της απλής γραμμικής παλινδρόμησης (LRL).

Και στην μέθοδο Holt, είναι σημαντική η σωστή επιλογή του αρχικού επιπέδου αλλά και της αρχικής τάσης, για την ορθή λειτουργία του μοντέλου πρόβλεψης και την παραγωγή ακριβέστερων σημειακών προβλέψεων. Η επιλογή των δύο αυτών μεταβλητών γίνεται με γνώμονα τα ποιοτικά χαρακτηριστικά των διαθέσιμων ιστορικών δεδομένων για τη χρονοσειρά που εξετάζουμε. Στην περίπτωση μας επιλέγουμε την μέθοδο της απλής γραμμικής παλινδρόμησης λόγω του όγκου των δεδομένων.

Η μέθοδος Holt, όπως και η μέθοδος SES, επιδέχεται βελτιστοποίησης, η οποία βασίζεται στην επιλογή του βέλτιστου συνδυασμού των παραμέτρων α και β. Η γραμμική αναζήτηση αποτελεί και πάλι την πιο διαδεδομένη λύση στο πρόβλημά μας, όμως για τη βελτιστοποίηση της συγκεκριμένης μεθόδου είναι ασφαλέστερη η επιλογή της μη γραμμικής (δυαδικής) αναζήτησης με κριτήριο και πάλι τη μείωση του μέσου τετραγωνικού σφάλματος (MSE), καθώς διαπιστώθηκε με τη βοήθεια πειραμάτων ότι παρουσιάζει καλύτερα αποτελέσματα ενώ παράλληλα μειώνει την πολυπλοκότητα της διαδικασίας υπολογισμού των παραμέτρων α και β. Ο αλγόριθμος για τη μη γραμμική βελτιστοποίηση αλλά και για την μέθοδο Holt, παρουσιάζονται αναλυτικά στο Παράρτημα Matlab.

Σε αυτό το σημείο είναι σημαντικό να αναφερθεί πως στις περισσότερες περιπτώσεις η βέλτιστη τιμή του συντελεστή β για την τάση είναι μικρότερη από την τιμή του συντελεστή α για το επίπεδο, λόγω του ότι η τιμή της τάσης για κάθε περίοδο είναι σημαντικά μικρότερη από την τιμή του αντίστοιχου επιπέδου. Στην περίπτωση που προσθέσουμε το ίδιο ποσοστό σφάλματος και στο επίπεδο αλλά και στην τάση, υπάρχει το ενδεχόμενο να προκύψουν προβλέψεις που θα χαρακτηρίζονται από μεγάλη αστάθεια.

4.2.3 Μοντέλο μη Γραμμικής Τάσης

Το μοντέλο μη γραμμικής τάσης αποτελεί ουσιαστικά μία τροποποίηση του μοντέλου γραμμικής τάσης. Η ανάγκη τροποποίησης του μοντέλου Holt έγκειται στο γεγονός ότι παρουσιάζει μία τάση για υπερεκτίμηση, δηλαδή οι προβλέψεις που παράγει χαρακτηρίζονται από θετική προκατάληψη, ιδιαίτερα σε περίπτωση μεσοπρόθεσμων ή μακροπρόθεσμων προβλέψεων. Η τροποποίηση αυτή επιτυγχάνεται με τη προσθήκη μίας επιπλέον παραμέτρου η οποία ελέγχει το ρυθμό αύξησης των τιμών. Η παράμετρος αυτή καλείται παράμετρος διόρθωσης της τάσης (trend-modification parameter) και συμβολίζεται φ. Οι εξισώσεις που περιγράφουν το συγκεκριμένο μοντέλο με τον τρόπο που αυτό παρουσιάστηκε από τους Gardner και McKenzie το 1985 είναι οι ακόλουθες:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + \varphi \bullet T_{t-1} + \alpha \bullet e_t$$

$$T_t = \varphi \bullet T_{t-1} + \beta \bullet e_t$$

$$F_{t+m} = S_t + \sum_{i=1}^{m} \varphi^i \bullet T_t$$

όπου:

- e_t δηλώνει το σφάλμα, δηλαδή την απόκλιση της πραγματικής τιμής από την πρόβλεψη,
- S δηλώνει το επίπεδο,
- F δηλώνει την πρόβλεψη,
- t δηλώνει τη χρονική περίοδο,
- T δηλώνει την τάση,
- m δηλώνει το χρονικό ορίζοντα της πρόβλεψης,
- α είναι ο συντελεστής εξομάλυνσης της μεθόδου για το επίπεδο, ο οποίος λαμβάνει τιμές στο διάστημα [0,1],
- β είναι ο συντελεστής εξομάλυνσης της μεθόδου για την τάση, ο οποίος λαμβάνει τιμές στο διάστημα [0,1],
- φ είναι η παράμετρος διόρθωσης της τάσης.

Από τις παραπάνω σχέσεις είναι σαφές ότι το μοντέλο είναι σχεδόν πανομοιότυπο με αυτό της γραμμικής τάσης με εξαίρεση την παρουσία της παραμέτρου φ, όπου συμβάλει στην παραγωγή προβλέψεων με μη γραμμικό υπολογισμό της τάσης μέσω του συντελεστή m, σε αντίθεση με πριν (Holt) όπου ο υπολογισμός ήταν γραμμικός.

Σε αντίθεση με τις παραμέτρους α και β, η παράμετρος φ μπορεί να λάβει οποιαδήποτε μεγαλύτερη του μηδενός τιμή, όμως είναι αναγκαία η επιβολή άνω και κάτω ορίων για ορισμένες χαρακτηριστικές τιμές. Επομένως, μπορούμε να χαρακτηρίσουμε το μοντέλο μη γραμμικής τάσης ανάλογα με την τιμή που παίρνει κάθε φορά η παράμετρος φ ,ως εξής (Πετρόπουλος & Ασημακόπουλος, 2011):

- Αν φ=0, τότε προκύπτει το μοντέλο της απλής εκθετικής εξομάλυνσης (SES), διότι η τάση δεν έχει πλέον καμία επίδραση στον καθορισμό των στατιστικών σημειακών προβλέψεων.

- Αν 0<φ<1, τότε προκύπτει το **μοντέλο φθίνουσας τάσης** (damped exponential smoothing), το οποίο θεωρείται καταλληλότερο όταν μας ενδιαφέρει η παραγωγή μεσοπρόθεσμων προβλέψεων, καθώς χαρακτηρίζεται από έλλειψη τάσης για υπεραισιοδοξία.

- Αν φ=1, τότε προκύπτει το μοντέλο της γραμμικής τάσης, αφού στην τελευταία εξίσωση η οποία παράγει και την πρόβλεψη, τη θέση του αθροίσματος παίρνει το γινόμενο της μεταβλητής χρονικού ορίζοντα m και τα προηγούμενης τάσης T_t.

- Αν φ>1, τότε προκύπτει το μοντέλο της εκθετικής τάσης, το οποίο χαρακτηρίζεται από μεγάλη θετική προκατάληψη και χρησιμοποιείται μόνο σε περιπτώσεις όπου το ζητούμενο είναι η πρόβλεψη τιμής στην αρχή του κύκλου ζωής ενός προϊόντος ή μιας υπηρεσίας.

Το μεγάλο μειονέκτημα του συγκεκριμένου μοντέλου, είναι η παραγωγή ιδιαίτερα προκατειλημμένων προβλέψεων, κάτι το οποίο ενδέχεται να προκύψει από λανθασμένη επιλογή του συντελεστή φ. Γι αυτό το λόγο, συνίσταται ο περιορισμούς της παραμέτρου φ στο διάστημα [0,1].

Αξίζει να σημειωθεί πως η ακρίβεια των προβλέψεων που παράγονται από το μοντέλο μη γραμμικής τάσης είναι σημαντικά μεγαλύτερη σε σχέση με αυτές του μοντέλου γραμμικής τάσης. Γενικά, όσο μεγαλύτερος είναι ο απαιτούμενος ορίζοντας της πρόβλεψης, τόσο μεγαλύτερο πλεονέκτημα παρουσιάζει το συγκεκριμένο μοντέλο έναντι των υπόλοιπων μοντέλων εκθετικής εξομάλυνσης.

Όσον αφορά τη βέλτιστη επιλογή των παραμέτρων α, β και φ, αυτή επιτυγχάνεται με μη γραμμική (δυαδική) αναζήτηση, με κριτήριο πάντα την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE), ενώ η επιλογή αρχικού επιπέδου και αρχικής τάσης πραγματοποιείται με εφαρμογή της απλής γραμμικής παλινδρόμησης, όπως ακριβώς και στην μέθοδο Holt.

4.3 Μέθοδοι Παλινδρόμησης

Η κατηγορία αυτή των μεθόδων πρόβλεψης έχει ως κυρίαρχο ζητούμενο της εύρεση συσχετίσεων μεταξύ μιας εξαρτημένης και μίας ή και περισσότερων ανεξάρτητων μεταβλητών, προσφέροντας τη δυνατότητα να κατανοήσουμε τον τρόπο με τον οποίο μεταβάλλεται η εξαρτημένη μεταβλητή σε μία πιθανή αλλαγή κάποιας ανεξάρτητης μεταβλητές, με όλες τις υπόλοιπες σταθερές.

Στην περίπτωση αυτή, η πρόβλεψη της ανεξάρτητης μεταβλητής είναι μια μαθηματική συνάρτηση των ανεξάρτητων μεταβλητών, η οποία ονομάζεται εξίσωση της παλινδρόμησης. Στην ενότητα αυτή θα αναλύσουμε δύο συγκεκριμένες μεθόδους παλινδρόμησης, την απλή γραμμική παλινδρόμηση και την πολλαπλή γραμμική παλινδρόμηση.

4.3.1 Απλή Γραμμική Παλινδρόμηση

Στη μέθοδο της απλής γραμμικής παλινδρόμησης (Linear regression line (LRL)), θεωρούμε ότι υπάρχει γραμμική σχέση ανάμεσα στην μεταβλητή την οποία θέλουμε να προβλέψουμε (εξαρτημένη) και σε κάποια άλλη μεταβλητή (ανεξάρτητη) η οποία συνήθως είναι ο χρόνος στην περίπτωση των χρονοσειρών.

Λόγω της υπόθεσης ότι η σχέση μεταξύ των δύο μεταβλητών είναι γραμμική, η συγκεκριμένη μέθοδος έχει πολύ περιορισμένο πεδίο εφαρμογής. Στόχος της απλής γραμμικής παλινδρόμησης είναι να εκφράσει τη σχέση που αναφέραμε ανάμεσα στις μεταβλητές X και Y με την εξίσωση μίας ευθείας γραμμής, δηλαδή μια σχέση της μορφής:

$$\hat{Y}_i = a + b \bullet X_i$$

όπου :

- a δηλώνει την τεταγμένη του σημείου τομής της ευθείας με τον άξονα της εξαρτημένης μεταβλητής, δηλαδή όταν θεωρηθεί μηδενική τιμή για την ανεξάρτητη μεταβλητή,

- b δηλώνει την κλίση της ευθείας, δηλαδή το ρυθμό μεταβολής του Y ανά μοναδιαία αύξηση του X.

Με δεδομένες τις πραγματικές τιμές Y της εξαρτημένης μεταβλητής, ο υπολογισμός των συντελεστών a και b γίνεται με γνώμονα την ελαχιστοποίηση των τετραγώνων των διαφορών των πραγματικών τιμών Y από τις τιμές που παράγει η εξίσωση παλινδρόμησης $\hat{Y}$ και περιγράφεται από την ακόλουθη σχέση:

$$(a, b) \mid \min \left[\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2\right]$$

Η λογική της παραπάνω μεθόδου βασίζεται στο ότι πρέπει να ελαχιστοποιηθεί η απόσταση των πραγματικών παρατηρήσεων Y από τη βέλτιστη γραμμή παλινδρόμησης. Η μέθοδος αυτή καλείται μέθοδος ελαχίστων τετραγώνων και με βάση αυτή οι σχέσεις που δίνουν τις παραμέτρους a και b είναι οι εξής:

$$b = \frac{\frac{\sum_{i=1}^{n} X_i \bullet Y_i}{n} - \overline{X} \bullet \overline{Y}}{\frac{\sum_{i=1}^{n} X_i^2}{n} - \overline{X}^2} \quad \text{ή ισοδύναμα} \quad b = \frac{\sum_{i=1}^{n} [(X_i - \overline{X}) \bullet (Y_i - \overline{Y})]}{\sum_{i=1}^{n} (X_i - \overline{X})^2}$$

$$a = \overline{Y} - b \bullet \overline{X}$$

όπου:

- $\overline{X}$ και $\overline{Y}$ είναι οι μέσες τιμές των διανυσμάτων X και Y,
- n είναι ο αριθμός των παρατηρήσεων (ιστορικών δεδομένων του υπό πρόβλεψη μεγέθους) με βάση τις οποίες υπολογίζεται η ευθεία της παλινδρόμησης.

4.3.2 Πολλαπλή Γραμμική Παλινδρόμηση

Η πολλαπλή γραμμική παλινδρόμηση αποτελεί στην ουσία μία γενίκευση της απλής γραμμικής παλινδρόμησης, και χρησιμοποιείται στην περίπτωση που θέλουμε να εξετάσουμε την επίδραση περισσότερων της μίας ανεξάρτητων μεταβλητών στην εξαρτημένη μεταβλητή. Η γενική σχέση που εκφράζει τη συγκεκριμένη μέθοδο είναι η εξής:

$$Y = b_0 + b_1 \bullet X_1 + b_2 \bullet X_2 + \ldots + b_k \bullet X_k + e$$

όπου:

- Y εκφράζει την εξαρτημένη μεταβλητή,
- X_1, X_2 έως X_k εκφράζουν τις ανεξάρτητες μεταβλητές,
- Οι συντελεστές b_0, b_1 έως b_k είναι σταθερές παράμετροι,

- e εκφράζει τον τυχαίο παράγοντα, ο οποίος θεωρείται κανονικά κατανεμημένος γύρω από το μηδέν.

Η παραπάνω εξίσωση είναι γραμμική ως προς τους συντελεστές. Για να επιτύχουμε τη γραμμικότητα της σχέσης θεωρούμε ότι κάθε συντελεστής b είναι υψωμένος σε εκθέτη που ισούται με τη μονάδα, και μπορούμε να υπολογίσουμε τις τιμές των συντελεστών με την χρήση της μεθόδου των ελαχίστων τετραγώνων.

Στην περίπτωση της πολλαπλής παλινδρόμησης, το γράφημα της συνάρτησης που συνδέει την μεταβλητή Y με τις διάφορες μεταβλητές X, δεν είναι πλέον μία ευθεία γραμμή όπως στην περίπτωση της απλής γραμμικής παλινδρόμησης και επομένως είναι δύσκολο να περιγραφεί. Κυρίαρχο ζητούμενο στη διαδικασία της πολλαπλής παλινδρόμησης είναι ο προσδιορισμός των συντελεστών b αλλά και της διακύμανσης του τυχαίου παράγοντα, έχοντας πάντα ένα συγκεκριμένο σύνολο δεδομένων στο οποίο εφαρμόζουμε τη μέθοδο των ελαχίστων τετραγώνων.

Πιο αναλυτικά, ο υπολογισμός των συντελεστών b της πολλαπλής παλινδρόμησης, για παράδειγμα στην περίπτωση που έχουμε δύο ανεξάρτητες μεταβλητές, βασίζεται στην εφαρμογή της μεθόδου των ελαχίστων τετραγώνων ώστε να υπολογιστεί το ελάχιστο άθροισμα των τετραγώνων των σφαλμάτων, ως εξής:

$$(b_0, b_1, b_2) \mid \min \left[\sum_{i=1}^{n} e_i^2\right] ,$$

όπου:

$$e_i = Y_i - \hat{Y}_i$$

Όμως ισχύει η ακόλουθη σχέση για τα σφάλματα:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_{1,i} - b_2 X_{2,i})^2$$

Άρα, για να υπολογίσουμε τους άγνωστους συντελεστές b_0, b_1, b_2, οι οποίοι ελαχιστοποιούν την παραπάνω σχέση, πρέπει να υπολογίσουμε τις μερικές παραγώγους αυτής για κάθε έναν από τους συντελεστές, στη συνέχεια να θέσουμε τις προκύπτουσες παραγώγους ίσες με το μηδέν και τέλος να λύσουμε ένα γραμμικό σύστημα τριών αγνώστων.

Η διαδικασία που περιγράφηκε, η οποία για λόγους κατανόησης αναφέρεται μόνο σε δύο ανεξάρτητες μεταβλητές, μπορεί να γενικευθεί σε οποιοδήποτε μοντέλο πολλαπλής γραμμικής παλινδρόμησης, με περισσότερες των δύο ανεξάρτητες μεταβλητές.

Οι αλγόριθμοι που χρησιμοποιούνται για την παραγωγή προβλέψεων με τις μεθόδους της απλής και πολλαπλής γραμμικής παλινδρόμησης, παρουσιάζονται αναλυτικά στο Παράρτημα Matlab.

4.4 Συνδυαστική μέθοδος Theta

Η μέθοδος Theta, η οποία προτάθηκε για πρώτη φορά από τους Ασημακόπουλο και Νικολόπουλο το 2000, είναι μια μονοδιάστατη μέθοδος πρόβλεψης, η οποία βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσα από την παράμετρο θ (Theta) που εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων (*Πετρόπουλος & Ασημακόπουλος, 2011*). Μέσω της παραπάνω διαδικασίας δημιουργείται μια νέα χρονοσειρά η οποία διατηρεί τη μέση τιμή και κλίση (παλινδρόμησης) της αρχικής χρονοσειράς αλλά δεν διατηρεί τις τοπικές καμπυλότητες και τη διακύμανση.

Οι παραγόμενες από αυτή τη διαδικασία χρονοσειρές καλούνται γραμμές Theta (Theta lines), και έχουν ως βασικό τους προτέρημα την καλύτερη προσέγγιση της μακροπρόθεσμης συμπεριφοράς, συγκεκριμένα της τάσης, των δεδομένων, αλλά και την ανάδειξη των βραχυπρόθεσμων χαρακτηριστικών της χρονοσειράς, εξαρτώμενες πάντα από την τιμή της παραμέτρου θ, που λαμβάνει τιμές μεγαλύτερες ή μικρότερες της μονάδας.

Αξίζει να αναφερθεί πως όσο μικρότερη είναι η τιμή του συντελεστή θ, τόσο μεγαλύτερος είναι ο βαθμός μείωσης των καμπυλοτήτων. Για παράδειγμα, στην ειδική περίπτωση όπου είναι $\theta=0$ η χρονοσειρά ταυτίζεται με την ευθεία της απλής γραμμικής παλινδρόμησης, στην περίπτωση που είναι $\theta=-1$ παίρνουμε τη συμμετρική της αρχικής χρονοσειράς ως προς την ευθεία της γραμμικής παλινδρόμησης, ενώ αν είναι $\theta>1$ τότε η χρονοσειρά που προκύπτει είναι διογκωμένη σε σύγκριση με την αρχική.

4.4.1 Διαδικασία Υπολογισμού Γραμμών Theta και τελικής πρόβλεψης

Η μέθοδος Theta διαχωρίζει την αρχική χρονοσειρά σε δύο ή και περισσότερες γραμμές Theta, οι οποίες προεκτείνονται στο μέλλον η καθεμία ξεχωριστά, και ανασυνθέτονται για την παραγωγή της τελικής πρόβλεψης. Ουσιαστικά η συγκεκριμένη συνδυαστική μέθοδος, μπορεί να θεωρηθεί ως μία διαφορετική μορφή αποσύνθεσης, ή επέκταση της θεωρίας των γραμμικών συνδυασμών παραγόμενων προβλέψεων.

Η προέκταση των επιμέρους γραμμών Theta μπορεί να υλοποιηθεί με οποιαδήποτε από τις διαθέσιμες μεθόδους πρόβλεψης που αναφέραμε μέχρι τώρα. Επιπλέον μπορεί να εφαρμοστεί διαφορετικός συνδυασμός γραμμών

Theta για διαφορετικό ορίζοντα πρόβλεψης. Στην απλή περίπτωση όπου η αρχική χρονοσειρά διαχωρίζεται σε 2 γραμμές Theta με την παράμετρο θ να παίρνει τις τιμές 0 και 2 αντίστοιχα, η παραγόμενη πρόβλεψη δίνεται από την ακόλουθη σχέση που αποτελεί απλό συνδυασμό των δύο παραγόμενων προβλέψεων από τις γραμμές Theta:

$$\mathrm{Y_t} = \frac{1}{2} \bullet (\mathrm{Y_t^{\theta=0}} + \mathrm{Y_t^{\theta=2}})$$

όπου:

- $\mathrm{Y_t^{\theta=0}}$ είναι η τιμή για την περίοδο t της γραμμής Theta με παράμετρο θ=0,

- $\mathrm{Y_t^{\theta=2}}$ είναι η τιμή για την περίοδο t της γραμμής Theta με παράμετρο θ=2.

Η γραμμή Theta με παράμετρο θ=0 (Theta(0)) είναι στην ουσία η απλή γραμμική παλινδρόμηση της αρχικής χρονοσειράς περιγράφοντάς την μέσω μιας γραμμικής τάσης, ενώ η γραμμή Theta με παράμετρο θ=2 (Theta(2)) έχει διπλάσιες τοπικές καμπυλότητες σε σχέση με την αρχική, δίνοντας μεγαλύτερη βαρύτητα στη βραχυπρόθεσμη συμπεριφορά. Η γραμμή Theta (2) συνήθως προεκτείνεται χρησιμοποιώντας τη μέθοδο της απλής εκθετικής εξομάλυνσης (SES), η οποία περιγράφηκε αναλυτικά στην ενότητα 4.2.1

Ο συνδυασμός των προβλέψεων των δύο γραμμών Theta που αναφέρθηκαν, μας εξασφαλίζει την αξιοποίηση τόσο της μακροπρόθεσμης τάσης (Theta line (0)), όσο και της βραχυπρόθεσμης πληροφορίας (Theta line (2)), παράγοντας προβλέψεις οι οποίες πετυχαίνουν καλύτερη προσέγγιση του σωστού επιπέδου, εξασφαλίζοντας παράλληλα μία συντηρητική αλλά σταθερή συνέχιση της μακροπρόθεσμης τάσης.

Ο αλγόριθμος που χρησιμοποιείται για τον υπολογισμό προβλέψεων με τη μέθοδο Theta, παρουσιάζεται αναλυτικά στο Παράρτημα Matlab.

Κεφάλαιο 5. Διαδικασία Προσομοίωσης

Στο σύγχρονο επιχειρηματικό κόσμο οι επικρατούσες συνθήκες χαρακτηρίζονται από ρευστότητα και υψηλή μεταβλητότητα. Οι επιχειρήσεις οι οποίες δεν είναι σε θέση να ανταποκριθούν γρήγορα και αποτελεσματικά στις συνεχόμενες αλλαγές της αγοράς, δε μπορούν να θεωρηθούν ανταγωνιστικές. Σε αυτό ακριβώς το πρόβλημα μπορεί να δώσει λύση η μέθοδος του rolling forecasting (κυλιόμενη πρόβλεψη). Χρησιμοποιώντας τις κυλιόμενες προβλέψεις μία επιχείρηση μπορεί να παρακολουθήσει καλύτερα την εξέλιξη των γεγονότων και να επαναπροσδιορίσει τη στρατηγική της για το εγγύς ή απώτερο μέλλον.

Όσον αφορά το πρόβλημα της πρόβλεψης της Οριακής τιμής Συστήματος (ΟΤΣ), οι διάφοροι ερευνητές που κατά καιρούς έχουν ασχοληθεί με αυτό, έχουν χρησιμοποιήσει ποικίλες διαδικασίες πρόβλεψης ανάμεσα στις οποίες είναι και το rolling forecasting. Στο παρών κεφάλαιο παρουσιάζεται αναλυτικά η συγκεκριμένη μέθοδος καθώς και η γενικότερη διαδικασία που θα ακολουθηθεί για την παραγωγή προβλέψεων στα πλαίσια αυτής της Διπλωματικής.

5.1 Rolling Forecasts

Η διαδικασία του rolling forecasting βασίζεται σε μία βασική αρχή, σύμφωνα με την οποία επανεξετάζονται οι υποθέσεις της αρχικής πρόβλεψης και παράγεται μία νέα πρόβλεψη η οποία βασίζεται στις νέες υποθέσεις που λήφθηκαν πλέον υπόψη. Υπό αυτή την ευρεία έννοια, η πλειοψηφία των οργανισμών χρησιμοποιεί πλέον την μέθοδο του rolling forecasting για την παραγωγή ασφαλέστερων προβλέψεων.

Στην ουσία πρόκειται για μία δυναμική μέθοδο προσομοίωσης, αφού τα δεδομένα μας συνεχώς ανανεώνονται. Αυτό έχει ως αποτέλεσμα την παραγωγή προβλέψεων οι οποίες λαμβάνουν υπόψη τα πιο πρόσφατα διαθέσιμα δεδομένα, επομένως μπορούν να ακολουθήσουν με μεγαλύτερη ακρίβεια τη συμπεριφορά της υπό εξέταση χρονοσειράς.

Για την καλύτερη κατανόηση του τρόπου λειτουργίας της μεθόδου είναι απαραίτητη η παρουσίαση ενός παραδείγματος. Έστω ότι θέλουμε να προβλέψουμε τη ζήτηση ενός προϊόντος για τους επόμενους 12 μήνες. Εφαρμόζοντας τη διαδικασία του rolling forecasting ο αριθμός των περιόδων που προβλέπουμε, δηλαδή ο ορίζοντας της πρόβλεψης, παραμένει σταθερός και στη συγκεκριμένη περίπτωση θα ισούται με12.

Όταν η πραγματική ζήτηση για μία περίοδο γίνεται διαθέσιμη, τότε παύει να θεωρείται πρόβλεψη και ενσωματώνεται στα πραγματικά δεδομένα. Άρα παύει να είναι μέρος των προβλέψεων και έτσι μία νέα περίοδος προστίθεται στο τέλος των προβλέψεων, δηλαδή γίνεται κύλιση των προβλέψεων μπροστά κατά μία περίοδο με συνέπεια ο ορίζοντας της πρόβλεψης να παραμένει πάντα σταθερός. Η διαδικασία του rolling forecasting για το παραπάνω παράδειγμα παρουσιάζεται στο ακόλουθο σχήμα:

Σχ. 5.1.1 - Rolling Forecasts

5.2 Σχεδιασμός Προσομοίωσης

Έχοντας επιλέξει ως διαδικασία προσομοίωσης τη μέθοδο του rolling forecasting, είναι απαραίτητη η κατάλληλη προσαρμογή της συγκεκριμένης διαδικασίας πάνω στο πρόβλημα πρόβλεψης της ΟΤΣ, που αποτελεί και το αντικείμενο αυτής της Διπλωματικής. Ο σχεδιασμός της προσομοίωσης έχει ως κυρίαρχο γνώμονα τα ιδιαίτερα χαρακτηριστικά που παρουσιάζει το μέγεθος της ΟΤΣ, και αποτελείται από 5 βασικά στάδια τα οποία παρουσιάζουν υψηλή εξάρτηση μεταξύ τους και είναι τα εξής:

i. Συγκέντρωση και τροποποίηση δεδομένων

ii. Γραφική αναπαράσταση δεδομένων

iii. Προσδιορισμός ποιοτικών χαρακτηριστικών των χρονοσειρών

iv. Εφαρμογή rolling forecasting

v. Αξιολόγηση των προβλέψεων

Στη συνέχεια της παρούσας ενότητας, παρουσιάζονται αναλυτικά τα 5 αυτά στάδια, και προσδιορίζεται η λειτουργία του καθενός μέσα στην όλη διαδικασία της προσομοίωσης.

5.2.1 Συγκέντρωση και τροποποίηση δεδομένων

Τα απαραίτητα δεδομένα συγκεντρώνονται αλλά και τροποποιούνται σε κατάλληλη μορφή χρονοσειρών με τη βοήθεια χρήσης υπολογιστικού φύλλου και μακροεντολών (Microsoft Excel & Visual Basic), προκειμένου να είναι δυνατή η χρησιμοποίηση τους για την παραγωγή των επιθυμητών προβλέψεων. Αξίζει να αναφερθεί πως τα αρχικά δεδομένα έχουν τη μορφή δισδιάστατου πίνακα, γεγονός που δε διευκολύνει την όλη διαδικασία πρόβλεψης και επομένως κρίθηκε απαραίτητη η τροποποίησή τους.

Η συγκεκριμένη τροποποίηση επιτυγχάνεται μέσω μακροεντολής η οποία μετασχηματίζει τον αρχικό δισδιάστατο πίνακα σε μία ενιαία χρονοσειρά, τοποθετώντας σε σειρά τα ωριαία δεδομένα κάθε ημέρας, σχηματίζοντας με αυτό τον τρόπο μία χρονοσειρά 43848 ωριαίων παρατηρήσεων.

5.2.2 Γραφική αναπαράσταση δεδομένων

Σε αυτό το στάδιο, τα δεδομένα που συλλέχθηκαν αναπαρίστανται γραφικά μέσω του προγράμματος Matlab, ούτως ώστε να γίνει μία πρώτη επαφή με τη μορφή των χρονοσειρών που θα χρησιμοποιηθούν αλλά και για τον εντοπισμό ορισμένων ποιοτικών χαρακτηριστικών που μπορεί να τις χαρακτηρίζουν, τα οποία δεν είναι εύκολο να προσδιοριστούν από τις αριθμητικές τιμές. Σε αυτό το στάδιο τα δεδομένα δεν υπόκεινται σε καμία μορφή επεξεργασίας, καθώς ζητούμενο αυτού του σταδίου είναι η «γνωριμία» με τη φυσική μορφή των χρονοσειρών. Εκτενέστερη αναφορά στη μορφή των δεδομένων θα γίνει στο επόμενο κεφάλαιο.

5.2.3 Προσδιορισμός ποιοτικών χαρακτηριστικών χρονοσειρών

Πρόκειται για τη συνέχεια του 2^{ου} σταδίου, όπου εντοπίζονται και απομονώνονται τα ποιοτικά χαρακτηριστικά των διαθέσιμων δεδομένων, όπως είναι η τάση, ο κύκλος, η εποχικότητα κτλ. Σχετικά με την εποχικότητα, αξίζει να αναφέρουμε πως το μέγεθος της ΟΤΣ χαρακτηρίζεται από πολλαπλή εποχικότητα, την οποία θα εξετάσουμε διεξοδικά στο Κεφάλαιο 6.

5.2.4 Εφαρμογή του rolling forecasting

Πρόκειται για το αμιγώς πρακτικό κομμάτι της προσομοίωσης, όπου πλέον τα δεδομένα είναι έτοιμα προς χρήση. Σε αυτό το στάδιο θα εφαρμοστούν και οι μέθοδοι πρόβλεψης που παρουσιάστηκαν εκτενώς στο Κεφάλαιο 4. Το συγκεκριμένο στάδιο αποτελείται από 4 μέρη τα οποία παρουσιάζονται αναλυτικά στη συνέχεια της παρούσας ενότητας.

5.2.4.1 Προεπεξεργασία (Pre – Process)

Οι χρονοσειρές οι οποίες θα χρησιμοποιηθούν υφίστανται μία περαιτέρω επεξεργασία με στόχο την διασφάλιση ασφαλέστερων προβλέψεων. Στα πλαίσια αυτής της προεπεξεργασίας απαλείφονται οι όποιες μηδενικές τιμές καθώς γίνεται και κατάλληλη κύλιση της χρονοσειράς της ΟΤΣ προκειμένου η πρώτη διαθέσιμη τιμή να αναφέρεται στις 14:00 η ώρα, όπου και γίνονται διαθέσιμες οι τιμές για την επόμενη μέρα, σύμφωνα με τη διαδικασία υπολογισμού του ΗΕΠ που ακολουθεί ο ΔΕΣΜΗΕ και περιγράφηκε εκτενώς στην ενότητα 1.3. Συγκεκριμένα, όσον αφορά τα μηδενικά, κάθε μηδενική τιμή αντικαθίσταται με το μέσο όρο της προηγούμενης και επόμενης μη μηδενικής τιμής (γραμμική παρεμβολή).

Επίσης, σε αυτό το στάδιο καθορίζεται και ο επιθυμητός ορίζοντας πρόβλεψης (horizon) αλλά και το βήμα ανανέωσης των δεδομένων (step) κατά τη διάρκεια του rolling. Φυσικά, ο καθορισμός του ορίζοντα των προβλέψεων βασίζεται πάντα στη λογική του *day ahead forecasting*, το οποίο αποτελεί και το βασικό άξονα γύρω από τον οποίο γίνεται η παραγωγή των προβλέψεων.

Επιπλέον ορίζεται το μέγεθος των in-sample δεδομένων που θα χρησιμοποιηθούν για το στήσιμο του εκάστοτε μοντέλου πρόβλεψης που θα χρησιμοποιηθεί, και ισούται με το 60% του συνόλου των διαθέσιμων παρατηρήσεων. Αξίζει να αναφερθεί πως τα παραπάνω πραγματοποιούνται πριν την επαναληπτική διαδικασία του rolling forecasting.

Αφότου ξεκινήσει η επαναληπτική διαδικασία του rolling, πραγματοποιείται η αποεποχικοποίηση της επεξεργασμένης χρονοσειράς που προέκυψε από τα προηγούμενα βήματα, μέσω της κλασικής μεθόδου αποσύνθεσης με κριτήριο τις παρατηρήσεις των τελευταίων x μηνών. Η επιλογή της παραμέτρου x δεν είναι τυχαία, αλλά βασίζεται στην ανάγκη για ομαλή μετάβαση από τους χειμερινούς στους καλοκαιρινούς μήνες του έτους. Οι δείκτες εποχικότητας που προκύπτουν από την παραπάνω διαδικασία εκφράζουν την εποχικότητα σε επίπεδο ώρας μέσα στην ημέρα. Χρησιμοποιώντας αυτούς τους δείκτες, υπολογίζεται η αποεποχικοποιημένη χρονοσειρά, η οποία και θα χρησιμοποιηθεί για την παραγωγή προβλέψεων.

Η συγκεκριμένη μεθοδολογία αφορά μόνο την περίπτωση όπου αντιμετωπίζεται η μονή εποχικότητα σε επίπεδο ημέρας. Υπάρχουν τρείς ακόμη διαφορετικές περιπτώσεις αντιμετώπισης της εποχικότητας, συμπεριλαμβανομένης και της διπλής εποχικότητας, οι οποίες παρουσιάζονται αναλυτικά στο Κεφάλαιο 7. Σε αυτό το σημείο ολοκληρώνεται η διαδικασία της προεπεξεργασίας, η οποία φυσικά επαναλαμβάνεται, όσον αφορά την αποεποχικοποίηση (προσδιορισμός εποχικών δεικτών), σε κάθε βήμα του rolling forecasting.

5.2.4.2 Παραγωγή Προβλέψεων (Forecasting)

Σε αυτό το στάδιο παράγονται προβλέψεις σύμφωνα με τη λογική του rolling forecasting, για σταθερό ορίζοντα, στα πλαίσια πάντα των βραχυπρόθεσμων προβλέψεων και του day ahead forecasting, το οποίο είναι αυτό που κυρίως επιζητούν και οι συμμετέχοντες στην Αγορά Ενέργειας. Η αποεποχικοποιημένη χρονοσειρά που προέκυψε από το στάδιο της προεπεξεργασίας χρησιμοποιείται ως είσοδος για τις διάφορες μεθόδους πρόβλεψης οι οποίες εφαρμόζονται σε αυτό το στάδιο και περιγράφηκαν αναλυτικά στο Κεφάλαιο 4.

Στο σημείο αυτό, αξίζει να αναφερθεί πως η διαδικασία της προσομοίωσης διαφέρει στην περίπτωση που παράγονται προβλέψεις με χρήση των αιτιοκρατικών μοντέλων σε σχέση με των μοντέλων χρονοσειρών, καθώς επίσης εξαρτάται και από τον τρόπο με τον οποίο αντιμετωπίζεται η εποχικότητα. Και οι δύο περιπτώσεις παρουσιάζονται αναλυτικά στο Παράρτημα Matlab. Η χρονοσειρά που προκύπτει και περιλαμβάνει τις προβλέψεις που παράχθηκαν με τον τρόπο που περιγράφηκε παραπάνω, υποβάλλεται σε μετεπεξεργασία.

5.2.4.3 Μετεπεξεργασία (Post – Process)

Αφού έχουν παραχθεί οι προβλέψεις σύμφωνα πάντα με τον επιθυμητό ορίζοντα, τα αποτελέσματα περνούν από το στάδιο της μετεπεξεργασίας όπου γίνεται η επανεποχικοποίησή τους, αλλά και τίθεται ως ανώτατο όριο τιμών τα 150 €/MWh σύμφωνα με τους κανονισμούς της ΡΑΕ. Συγκεκριμένα, η χρονοσειρά των προβλέψεων που παρήχθησαν επανεποχικοποιείται χρησιμοποιώντας τους δείκτες εποχικότητας που υπολογίστηκαν στο στάδιο της προεπεξεργασίας, με διαφορετικό τρόπο για κάθε μία από τις 4 περιπτώσεις αντιμετώπισης της εποχικότητας, και προκύπτει η τελική χρονοσειρά που περιλαμβάνει τις επανεποχικοποιημένες προβλέψεις.

Είναι σημαντικό να τονιστεί το γεγονός ότι γίνεται λόγος για πολλαπλασιαστικό μοντέλο εποχικότητας, επομένως η επανεποχικοποίηση πραγματοποιείται με πολλαπλασιασμό των δεικτών εποχικότητας με τις αντίστοιχες παρατηρήσεις, δίνοντας μεγάλη προσοχή στον απόλυτο συγχρονισμό των δεικτών με τις παρατηρήσεις, προκειμένου να αποφευχθούν λανθασμένα αποτελέσματα.

Στο σημείο αυτό αξίζει να αναφερθεί πως όλα τα παραπάνω, από το σημείο της αποεποχικοποίησης της επεξεργασμένης χρονοσειράς μέχρι και την επανεποχικοποίηση της χρονοσειράς των προβλέψεων, πραγματοποιούνται μέσα στα πλαίσια της επαναληπτικής διαδικασίας του rolling forecasting, η οποία ανανεώνεται σύμφωνα με το βήμα που έχει καθοριστεί στην αρχή της διαδικασίας και η διάρκειά της εξαρτάται από τον ορίζοντα των προβλέψεων.

5.2.4.4 Υπολογισμός Σφαλμάτων (Evaluation)

Για την αξιολόγηση των παραγόμενων προβλέψεων απαιτείται ο υπολογισμός ορισμένων στατιστικών δεικτών, όπως είναι τα σφάλματα. Για το λόγο αυτό, μετά το πέρας της επαναληπτικής διαδικασίας του rolling forecasting, υπολογίζεται η διαφορά μεταξύ της χρονοσειράς των πραγματικών δεδομένων (δίχως αποεποχικοποίηση) και αυτής των επανεποχικοποιημένων προβλέψεων.

Αξίζει να αναφερθεί πως οι παραπάνω χρονοσειρές ανανεώνονται σε κάθε βήμα της διαδικασίας του rolling forecasting, και επομένως ο υπολογισμός τους γίνεται στα πλαίσια της όλης επαναληπτικής διαδικασίας.

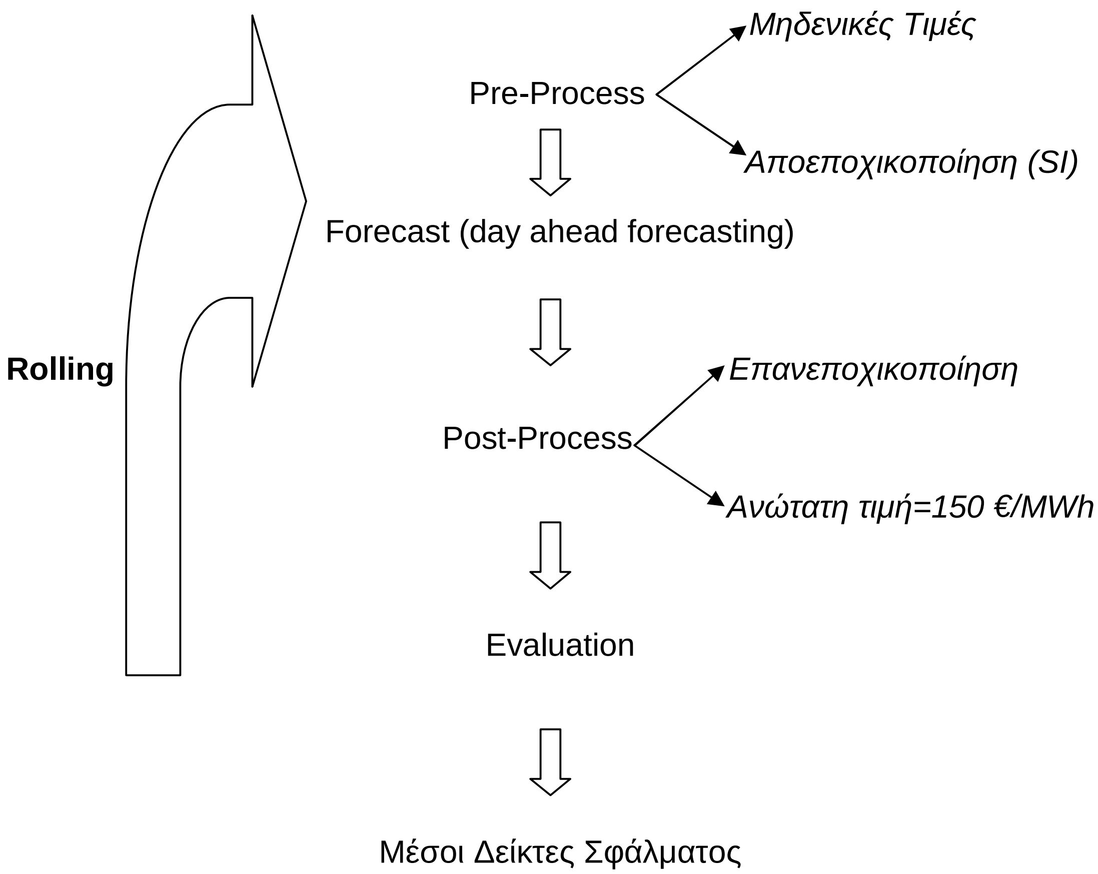
Η εν λόγω διαφορά εκφράζεται με τη μορφή ενός δισδιάστατου πίνακα, όπου έχει ως γραμμές το σύνολο των ημερών που προβλέπονται και ως στήλες τις 24 ώρες που περιλαμβάνει η κάθε ημέρα, αλλά με ορίζοντα μίας εβδομάδας.

Αυτό σημαίνει ότι, έχοντας ως ορίζοντα πρόβλεψης την μία εβδομάδα, οι 24 ώρες της πρώτης ημέρας που προβλέπεται απέχουν 7 ημέρες μακριά από την τελευταία διαθέσιμη ημέρα, και όσο προχωράει το rolling και ανανεώνονται τα διαθέσιμα δεδομένα αυτή η ημέρα απέχει πλέον 6 ημέρες, έπειτα 5 ημέρες κοκ. Έχοντας πλέον υπολογίσει το συγκεκριμένο πίνακα, υπάρχει η δυνατότητα της αξιολόγηση των προβλέψεων που ακολουθεί στο τελευταίο στάδιο της προσομοίωσης.

5.2.5 Αξιολόγηση των προβλέψεων

Πρόκειται για το τελευταίο στάδιο της προσομοίωσης, κατά το οποίο υπολογίζονται οι τιμές των σφαλμάτων APE και sMAPE, με τη βοήθεια του πίνακα που υπολογίστηκε στο προηγούμενο μέρος της προσομοίωσης. Έχοντας υπολογίσει τους συγκεκριμένους δείκτες, είναι εύκολο να υπολογιστούν και οι μέσες τιμές αυτών. Αυτές οι μέσες τιμές θα είναι και το τελικό κριτήριο αξιολόγησης των προβλέψεων που παράχθηκαν και θα χρησιμοποιηθούν για την περαιτέρω ανάλυση των αποτελεσμάτων.

Η διαδικασία της προσομοίωσης που περιγράφηκε μέσα από το παρών Κεφάλαιο, παρουσιάζεται συνοπτικά στο ακόλουθο συγκεντρωτικό διάγραμμα:



Σχ. 5.2.1 - Προσομοίωση Rolling Forecasting

Αναλυτική παρουσίαση ολόκληρης της διαδικασίας προσομοίωσης όπως περιγράφηκε στην παρούσα ενότητα, είναι διαθέσιμη στο Παράρτημα Matlab.

Κεφάλαιο 6. Επεξεργασία Δεδομένων

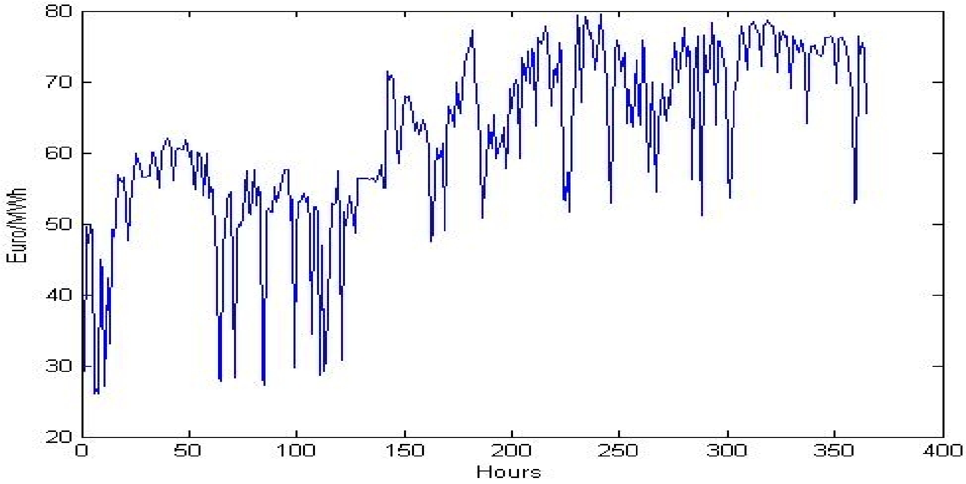
Στο συγκεκριμένο κεφάλαιο θα γίνει η παρουσίαση των δεδομένων που θα χρησιμοποιηθούν κατά τη διαδικασία του rolling forecasting η οποία περιγράφηκε εκτενώς στο προηγούμενο κεφάλαιο. Πέραν της απλής παρουσίασης αυτών των δεδομένων, κρίνεται απαραίτητη και η επεξεργασία τους η οποία αφορά κυρίως στα ποσοτικά και ποιοτικά χαρακτηριστικά των χρονοσειρών που είναι διαθέσιμες. Ο κύριος όγκος των δεδομένων προέρχεται από τον διαδικτυακό ιστότοπο του ΔΕΣΜΗΕ, και περιλαμβάνει τις ιστορικές τιμές της ΟΤΣ, καθώς και της ζήτησης και πρόβλεψης φορτίου, αλλά και ορισμένα δεδομένα σχετικά με τις εισαγωγές/εξαγωγές ενέργειας. Επιπρόσθετα, συλλέχθηκαν δεδομένα που αφορούν τις τιμές των καυσίμων που χρησιμοποιούνται στην Παραγωγή Ηλεκτρικής Ενέργειας στην Ελλάδα, με στόχο τη χρησιμοποίησή τους ως ανεξάρτητες μεταβλητές στη μέθοδο της πολλαπλής ή μη, γραμμικής παλινδρόμησης. Στη συνέχεια αυτού του κεφαλαίου ακολουθεί η ανάλυση και περιγραφή της κάθε χρονοσειράς των δεδομένων που περιγράφηκαν.

6.1 Ιστορικές Τιμές ΟΤΣ

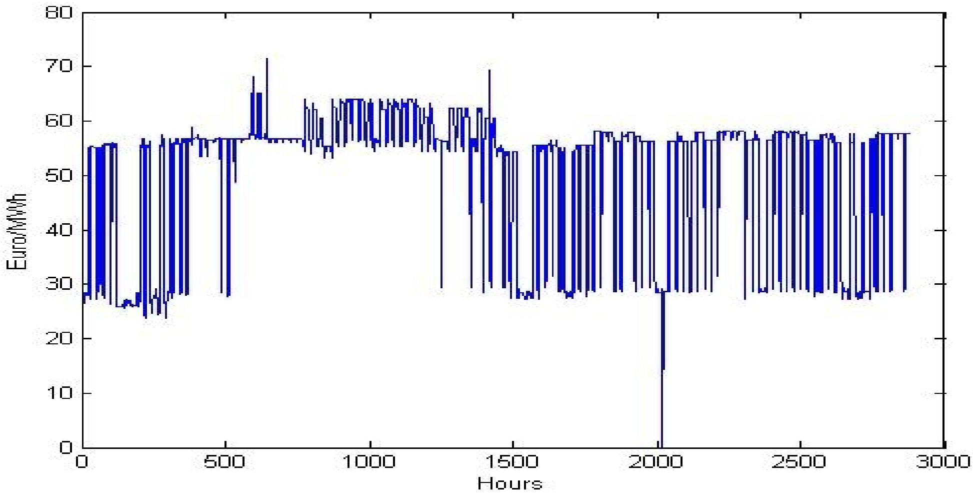
Οι ιστορικές τιμές της Οριακή Τιμής Συστήματος (ΟΤΣ) αποτελούν το βασικό άξονα για την παραγωγή προβλέψεων στα πλαίσια αυτής της Διπλωματικής. Πρόκειται για ωριαία δεδομένα που εκφράζουν την τιμή της ηλεκτρικής ενέργειας σε €/MWh συναρτήσει του χρόνου. Τα δεδομένα αυτά αφορούν 5 έτη, από την 1/1/2006 έως την 1/1/2011. Όπως αναφέρθηκε και στο Κεφάλαιο 5, το 60% του όγκου αυτών των παρατηρήσεων, αποτελεί τα in-sample δεδομένα μας. Επομένως τα 3 πρώτα έτη χρησιμοποιούνται για την «εκπαίδευση» του εκάστοτε μοντέλου πρόβλεψης που χρησιμοποιείται, ενώ υπολογίζεται η ακρίβεια των προβλέψεων για τα 2 εναπομείναντα έτη, συγκρίνοντας τις παραγόμενες προβλέψεις με τις πραγματικές διαθέσιμες ιστορικές τιμές.

6.1.1 Γραφική Αναπαράσταση ΟΤΣ

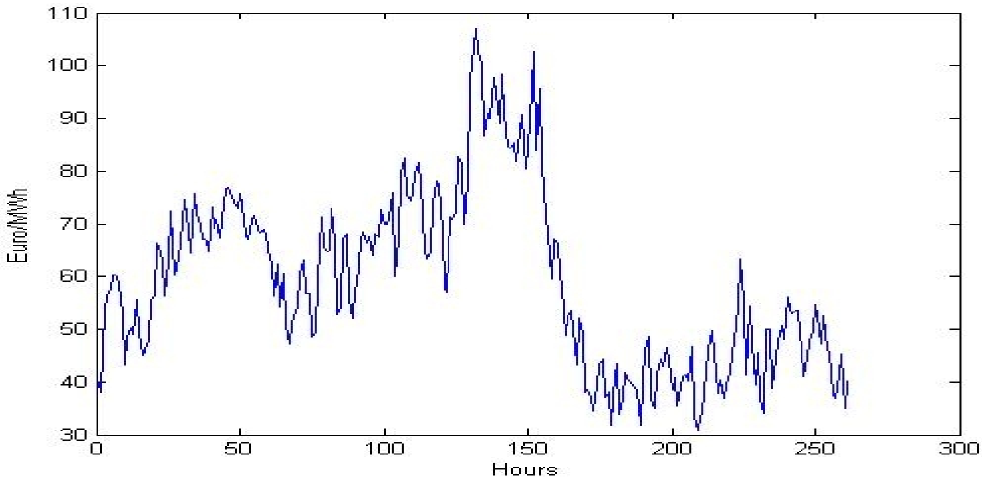
Το μέγεθος της ΟΤΣ, όπως έχει αναφερθεί εκτενώς και στην Ενότητα 2.2, παρουσιάζει μερικά ιδιαίτερα χαρακτηριστικά, όπως είναι η υψηλή τυχαιότητα (volatility), οι πολλές ασυνήθιστες τιμές (spikes) αλλά και η πολλαπλή εποχικότητα (multiple seasonality). Προκειμένου να εντοπιστούν και να απομονωθούν αυτά τα ποιοτικά χαρακτηριστικά, σε πρώτη φάση κρίθηκε απαραίτητη η απλή γραφική αναπαράσταση της χρονοσειράς της ΟΤΣ. Στα διαγράμματα που ακολουθούν, παρουσιάζεται μέρος των διαθέσιμων παρατηρήσεων της ΟΤΣ, σε επίπεδο ώρας, ημέρας αλλά και εβδομάδας:



Σχ. 6.1.1 - Ημερήσια δεδομένα ΟΤΣ για το έτος 2006



Σχ. 6.1.2 - Ωριαία δεδομένα ΟΤΣ για τους 4 πρώτους μήνες του έτους 2006

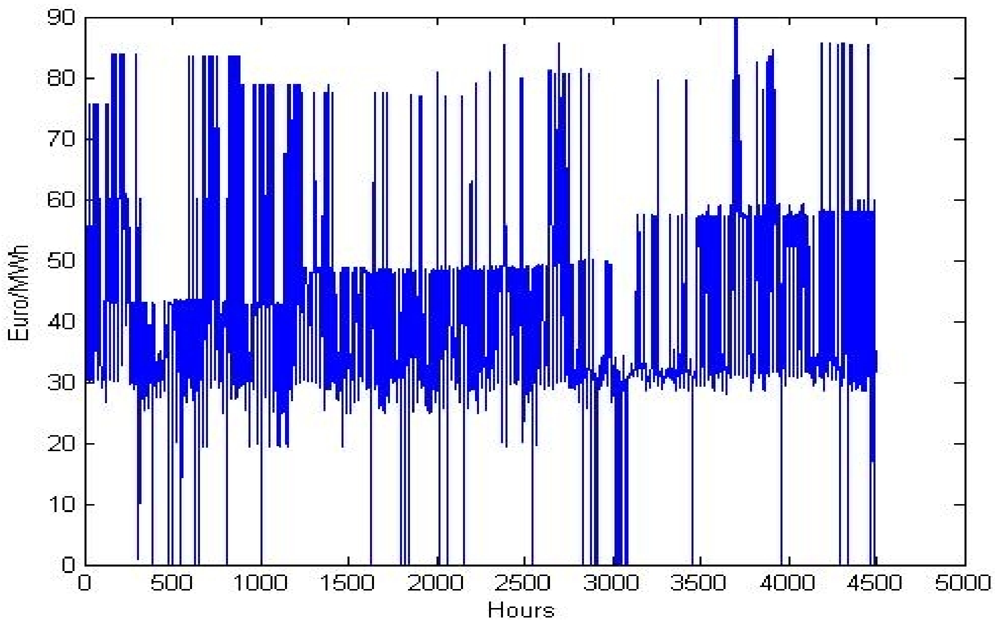


Σχ. 6.1.3 - Εβδομαδιαία δεδομένα ΟΤΣ για τα έτη 2006-2011

Παρατηρώντας το Σχ. 6.1.1, διακρίνουμε την ύπαρξη τάσης, καθώς και την ύπαρξη εποχικότητας στο επίπεδο της ημέρας. Εντούτοις, στο Σχ. 6.1.2 είναι εμφανές πως πλέον δεν υπάρχει τάση, ενώ η εποχικότητα είναι μεν υπαρκτή αλλά πλέον είναι σε ωριαίο επίπεδο.

Από αυτά τα δύο σχήματα, είναι πλέον δυνατό να εξαχθεί ένα πρώτο συμπέρασμα για την πολλαπλή εποχικότητα που χαρακτηρίζει το μέγεθος της ΟΤΣ, και αναλύεται σε 2 συνιστώσες, την εποχικότητα μέσα στη μέρα και την εποχικότητα στην εβδομάδα. Η δεύτερη συνιστώσα είναι ακόμη πιο εμφανής στο Σχ. 6.1.3, όπου τα δεδομένα έχουν αναχθεί στο επίπεδο της εβδομάδας.

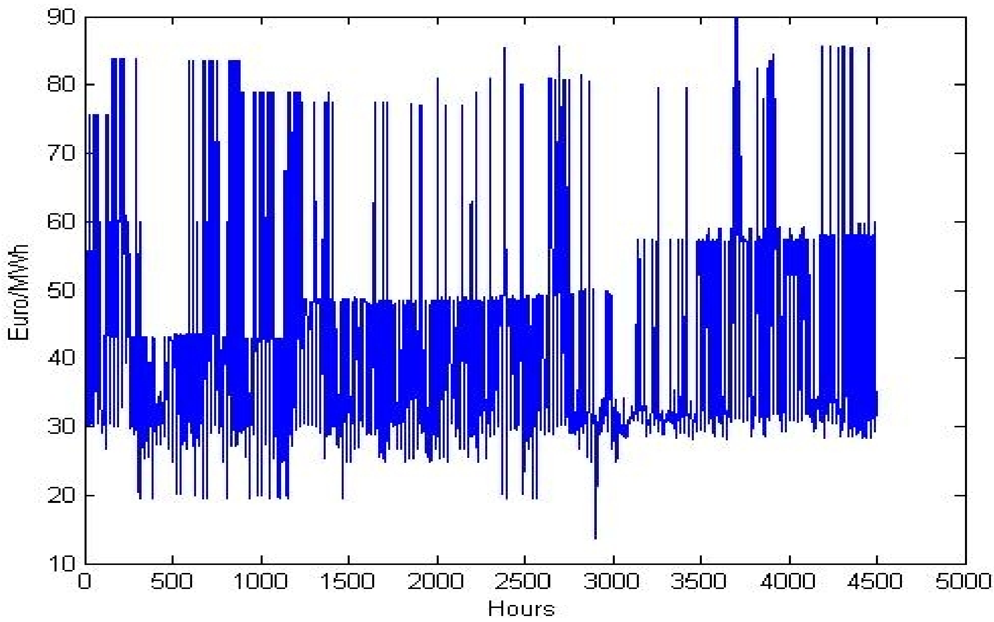
Πλέον είναι ξεκάθαρη η απουσία τάσης, καθώς οι παρατηρήσεις κατανέμονται γύρω από ένα επίπεδο τιμών. Επίσης, στο Σχ. 6.1.2, διακρίνουμε μία μηδενική τιμή. Στη συνέχεια της χρονοσειράς το φαινόμενο αυτό γίνεται όλο και συχνότερο, καθώς πέρα από τις αλλαγές τις ώρας μέσα σε ένα έτος, υπάρχουν και αρκετές μη διαθέσιμες τιμές (missing values), που πιθανότατα οφείλονται σε λάθος του Συστήματος. Στο ακόλουθο σχήμα παρουσιάζεται ένα τμήμα της χρονοσειράς της ΟΤΣ, στο οποίο είναι ιδιαίτερα έντονο το παραπάνω φαινόμενο:



Σχ. 6.1.4 – Τμήμα ΟΤΣ (επίπεδο ώρας) με Ανεξήγητες Μηδενικές Τιμές

Όπως είναι ξεκάθαρο από το παραπάνω σχήμα, δεν είναι λογική η εμφάνιση τέτοιου όγκου μηδενικών τιμών, επομένως είναι προφανές ότι πρόκειται για σφάλμα του Συστήματος, ή οφείλεται σε τυχόν μη προγραμματισμένες διακοπές.

Προκειμένου να αντιμετωπιστούν οι τόσες μηδενικές τιμές, χρησιμοποιείται για την απαλοιφή τους η μέθοδος της γραμμικής παρεμβολής, καθώς η ύπαρξη αυτών των μηδενικών παρατηρήσεων, θα καθιστούσε προβληματική την προσπάθεια παραγωγής ικανοποιητικών προβλέψεων. Επιπλέον τίθεται ένα κατώτατο όριο τιμών, για την εξάλειψη των ασυνήθιστα χαμηλών τιμών που μπορεί να υπάρχουν στις παρατηρήσεις. Στο ακόλουθο σχήμα φαίνεται το ίδιο τμήμα με το Σχ. 6.1.4 της χρονοσειράς της ΟΤΣ, έπειτα από την απαλοιφή των μηδενικών τιμών και την επιβολή κατωφλίου ίσου με 5 €/MWh:



Σχ. 6.1.5 – Τμήμα ΟΤΣ (επίπεδο ώρας) έπειτα από απαλοιφή των μηδενικών τιμών

6.1.2 Ποσοτικά Χαρακτηριστικά

Έχοντας μία πρώτη εικόνα, προερχόμενη από την γραφική αναπαράσταση της χρονοσειράς της ΟΤΣ, κρίνεται απαραίτητος ο προσδιορισμός ορισμένων ποσοτικών αλλά και ποιοτικών χαρακτηριστικών. Όσον αφορά την πρώτη κατηγορία, πρόκειται κυρίως για κάποιους βασικούς στατιστικούς δείκτες, έχοντας ως στόχο την περαιτέρω ανάλυση και κατανόηση του εν λόγω μεγέθους.

Στον πίνακα που ακολουθεί, παρουσιάζονται ορισμένα χρήσιμα στατιστικά μεγέθη, τα οποία προέρχονται από τα διαθέσιμα ιστορικά δεδομένα της ΟΤΣ, μετά από απαλοιφή των μηδενικών τιμών σύμφωνα με τη μεθοδολογία που αναφέρθηκε προηγουμένως. Οι ακόλουθοι στατιστικοί δείκτες έχουν υπολογιστεί με τη βοήθεια του προγράμματος Matlab:

Ποσοτικά Χαρακτηριστικά ΟΤΣ
ΜΕΣΗ ΤΙΜΗ (mean) : 59,3735 €/MWh
ΜΕΓΙΣΤΗ ΤΙΜΗ (max) : 138 €/MWh
ΕΛΑΧΙΣΤΗ ΤΙΜΗ (min) : 13,7410 €/MWh
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (standard deviation) : 22,5746
ΔΙΑΚΥΜΑΝΣΗ (variance): 509,6132

Πίνακας 6.1.1 - Ποσοτικά Χαρακτηριστικά ΟΤΣ

Υπολογίζοντας τους παραπάνω δείκτες, παρατηρούμε ότι προφανώς και τηρείται το ανώτατο όριο τιμών που επιβάλει η ΡΑΕ (150 €/MWh), καθώς επίσης και μία περισσότερο λογική κατώτατη τιμή. Στην περίπτωση που η χρονοσειρά δεν είχε υποστεί επεξεργασία για την απαλοιφή των μηδενικών τιμών, μοιραία η κατώτατη τιμή θα ήταν ίση με το μηδέν, γεγονός το οποίο απέχει πολύ από την πραγματικότητα.

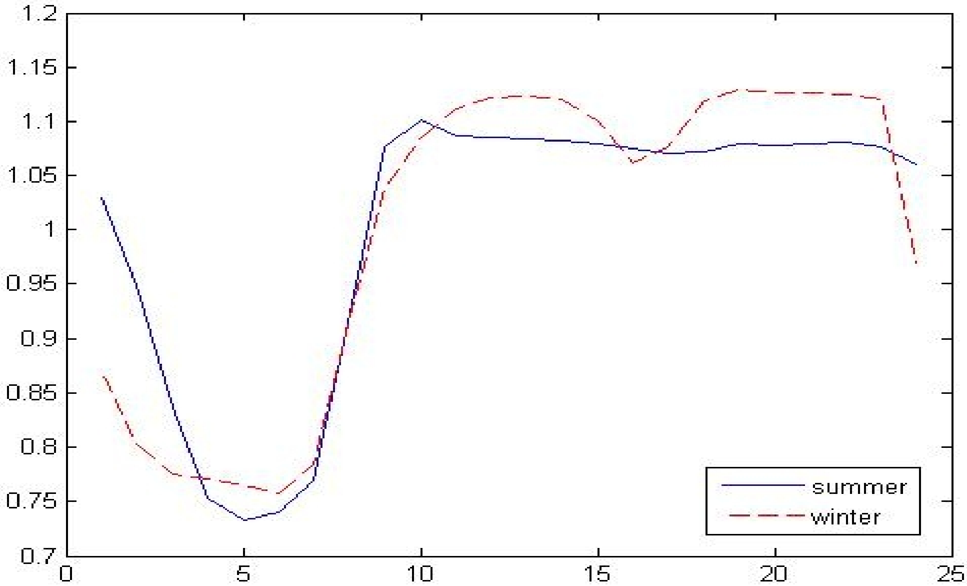
Η επεξεργασία αυτή, επηρεάζει ελάχιστα τον δείκτη της τυπικής απόκλισης, καθώς συμπεριλαμβανομένων των μηδενικών τιμών, η τυπική απόκλιση λαμβάνει την τιμή 22,831, τιμή που απέχει ελάχιστα από αυτή που αναγράφεται στον παραπάνω πίνακα.

Εντούτοις, ο δείκτης της διακύμανσης παρουσιάζεται ελαφρώς διαφορετικός, καθώς ισούται με το τετράγωνο της τυπικής απόκλισης, και οποιαδήποτε διαφορά έχει μεγαλύτερο αντίκτυπο στην τιμή του. Ενδεικτικά αναφέρεται ότι χωρίς την απαλοιφή των μηδενικών τιμών, η διακύμανση λαμβάνει την τιμή 521,3514, δηλαδή μία διαφορά της τάξης των 11,7382 μονάδων.

6.1.3 Ποιοτικά Χαρακτηριστικά – Πολλαπλή Εποχικότητα

Όσον αφορά τα ποιοτικά χαρακτηριστικά, ήδη αναφέρθηκε προηγουμένως η απουσία τάσης στη χρονοσειρά της ΟΤΣ, καθώς οι παρατηρήσεις κινούνται γύρω από ένα σταθερό επίπεδο. Σε αντίθεση με την τάση, είναι έντονη η παρουσία του ποιοτικού χαρακτηριστικού της εποχικότητας. Πιο αναλυτικά, γίνεται λόγος για πολλαπλή εποχικότητα, η οποία αναφέρεται σε επίπεδο ώρας μέσα σε μία ημέρα και σε επίπεδο ημέρας μέσα σε μία εβδομάδα.

Υπάρχουν δηλαδή δύο εποχικότητες, όπου η μία εμπεριέχει την άλλη. Αξίζει να αναφερθεί, πως το μέγεθος της ΟΤΣ επηρεάζεται αισθητά από τις εποχές (calendar effect), με συνέπεια να είναι δυνατή η διάκριση δύο διαφορετικών εποχικών δεικτών, ενός για τους καλοκαιρινούς μήνες και ενός για τους χειμερινούς. Προκειμένου να γίνει περισσότερο κατανοητή αυτή η εξάρτηση από το ημερολόγιο, στο διάγραμμα που ακολουθεί, παρουσιάζονται οι δύο διαφορετικοί δείκτες για το επίπεδο της ώρας μέσα σε μία ημέρα:

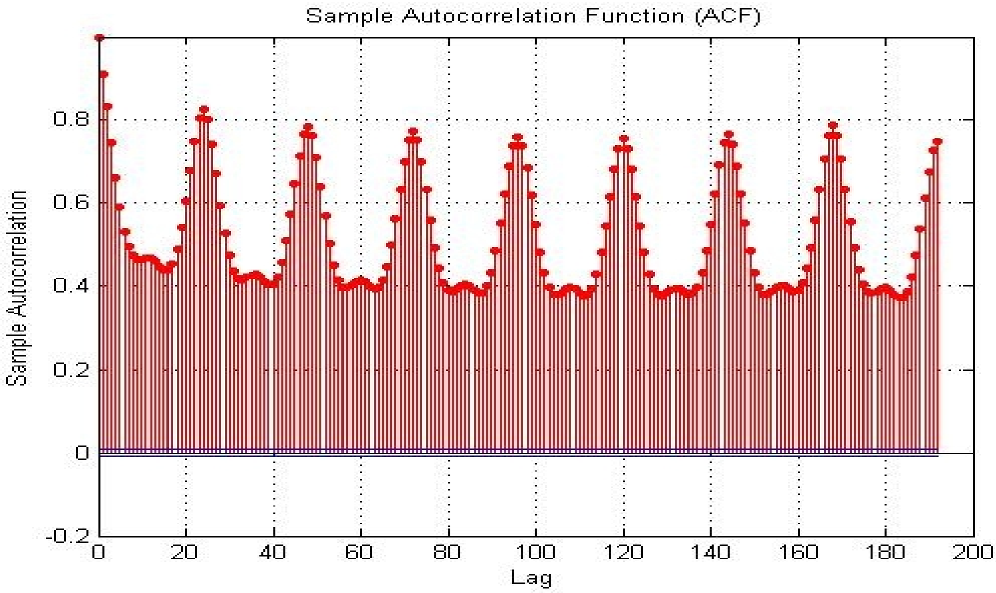


Σχ. 6.1.6 – Καλοκαιρινοί/Χειμερινοί Δείκτες εποχικότητας

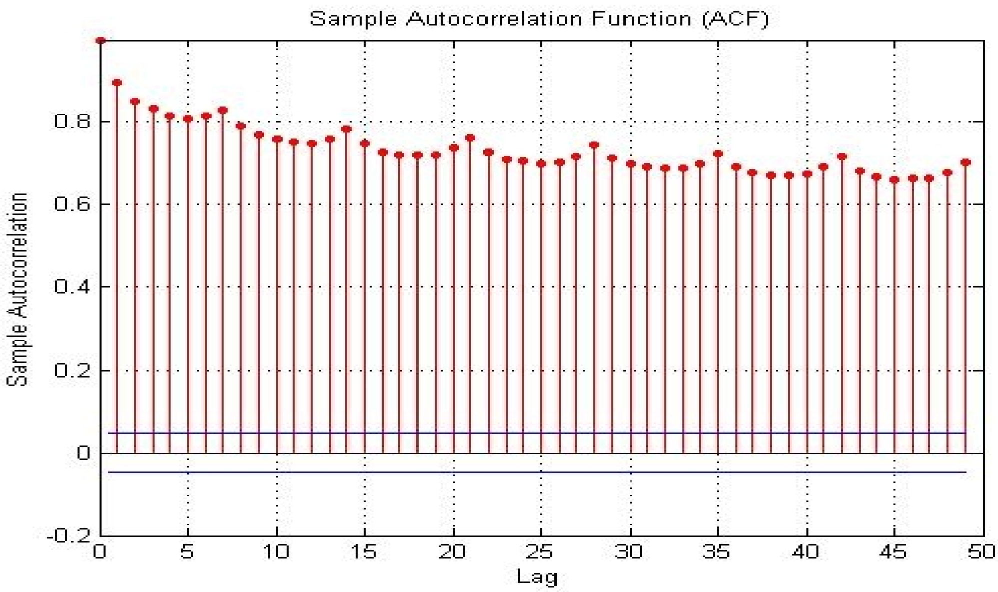
Παρατηρώντας το παραπάνω διάγραμμα, είναι εμφανής η διαφορά που υπάρχει μεταξύ καλοκαιρινής και χειμερινής περιόδου. Ο καλοκαιρινός δείκτης εποχικότητας παρουσιάζει μία πιο ομαλή συμπεριφορά από τις πρωινές έως τις βραδινές ώρες, σε αντίθεση με το χειμερινό δείκτη, ο οποίος χαρακτηρίζεται από περισσότερες μεταβολές, αλλά και υψηλότερες τιμές. Αυτό οφείλεται πιθανότατα στις αλλαγές των δραστηριοτήτων των ανθρώπων κατά τη διάρκεια του καλοκαιριού σε σχέση με τη χειμερινή περίοδο.

Ένα ακόμη χρήσιμο εργαλείο, το οποίο μπορεί να βοηθήσει στον εντοπισμό αλλά και την ανάλυση της εποχικότητας στη χρονοσειρά της ΟΤΣ, είναι η αυτοσυσχέτιση, η οποία περιγράφηκε αναλυτικά στην Παράγραφο 3.2. Μέσω της αυτοσυσχέτισης είναι δυνατό να αποτυπωθεί το εποχικό μοτίβο που ακολουθεί μία χρονοσειρά ως προς το επιθυμητό επίπεδο. Στην περίπτωση της ΟΤΣ, μας αφορά η εποχικότητα σε επίπεδο ημέρας αλλά και εβδομάδας.

Στα διαγράμματα που ακολουθούν, παρουσιάζεται το δείγμα από τη διαδικασία αυτοσυσχέτισης για τα δύο επίπεδα που αναφέρθηκαν προηγουμένως:



Σχ. 6.1.7 – Δείγμα Αυτοσυσχέτισης με καθυστέρηση 8 ημερών



Σχ. 6.1.8 – Δείγμα Αυτοσυσχέτισης με καθυστέρηση 7 εβδομάδων

Στο Σχ. 6.1.7, παρουσιάζεται το δείγμα αυτοσυσχέτισης της χρονοσειράς της ΟΤΣ σε επίπεδο ώρας, με καθυστέρηση 8 ημερών, προκειμένου να γίνει αντιληπτό το εποχικό μοτίβο που ακολουθούν οι παρατηρήσεις μέσα σε μία ημέρα κατά τη διάρκεια μιας εβδομάδας. Είναι εμφανές πως τα δεδομένα χαρακτηρίζονται από ισχυρή εποχικότητα σε επίπεδο ώρας, καθώς παρατηρείται επανάληψη του προτύπου των 24 ωρών μέσα στην εβδομάδα.

Έχοντας αναγάγει τα δεδομένα στο επίπεδο της ημέρας, στο Σχ. 6.1.8 παρουσιάζεται το δείγμα αυτοσυσχέτισης της ΟΤΣ με καθυστέρηση 7 εβδομάδων. Με τον τρόπο αυτό απομονώνεται το πρότυπο της εποχικότητας που παρουσιάζεται μέσα σε μία εβδομάδα στο επίπεδο της ημέρας. Όπως και προηγουμένως, έτσι και εδώ γίνεται εύκολα αντιληπτή η ισχυρή ύπαρξη ενός προτύπου εποχικότητας, καθώς κάθε ημέρα της εβδομάδας επαναλαμβάνεται με παρόμοιο τρόπο με το πέρας των εβδομάδων.

6.2 Ανεξάρτητες Μεταβλητές Παλινδρόμησης

Έχοντας εξετάσει και αναλύσει διεξοδικά τη χρονοσειρά της ΟΤΣ, κρίνεται πλέον απαραίτητο να γίνει αναφορά και στα δευτερεύοντα δεδομένα, τα οποία χρησιμοποιούνται για την εφαρμογή της απλής ή και πολλαπλής γραμμικής παλινδρόμησης. Τα δεδομένα αυτά είναι τα εξής:

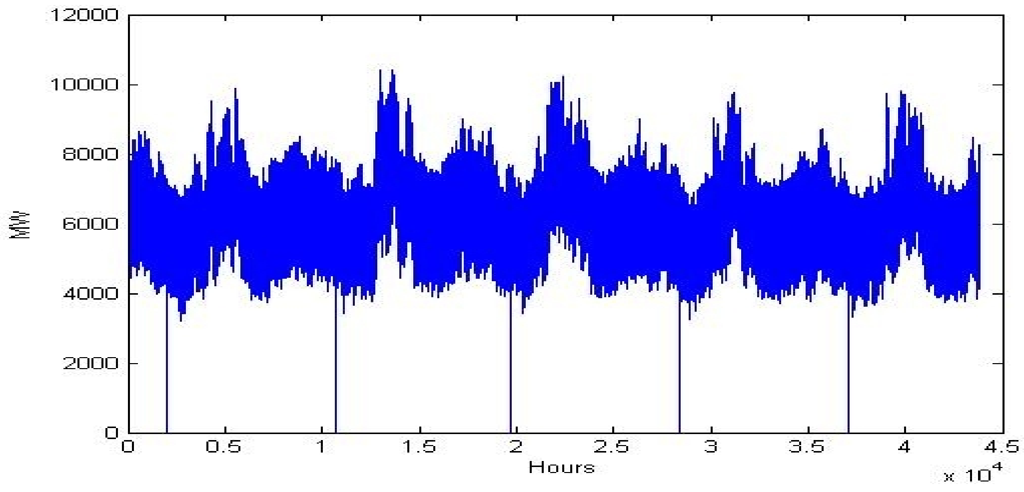
- Ιστορικές Τιμές Φορτίου (actual load).
- Πρόβλεψη Φορτίου (load forecast).
- Μηνιαίες Ιστορικές Τιμές Πετρελαίου (crude oil).
- Μηνιαίες Ιστορικές Τιμές Φυσικού Αερίου (natural gas).

Η επιλογή των τεσσάρων παραπάνω χρονοσειρών δεν είναι τυχαία, αλλά βασίζεται στη μεγάλη επίδραση που αυτές έχουν στη χρονοσειρά της ΟΤΣ και θεωρούνται αρκούντως ικανοποιητικές ως ανεξάρτητες μεταβλητές στη διαδικασία της απλής ή πολλαπλής γραμμικής παλινδρόμησης.

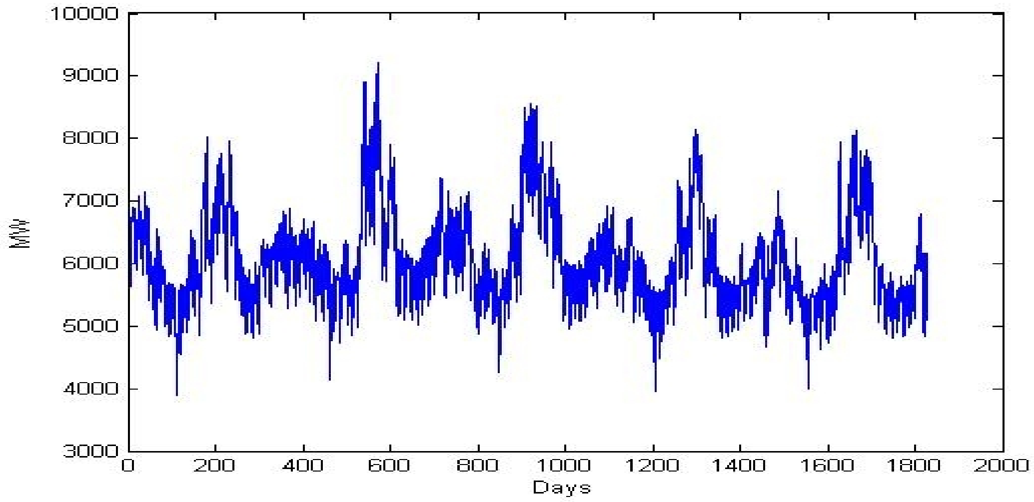
6.2.1 Ιστορικές Τιμές Φορτίου

Πρόκειται για τις πραγματικές τιμές του φορτίου (MW) όπως αυτές καταγράφηκαν από το ΔΕΣΜΗΕ και αφορούν την περίοδο από την 1/1/2006 έως την 1/1/2011, δηλαδή τα εν λόγω δεδομένα έχουν τον ίδιο όγκο με αυτά της ΟΤΣ. Και σε αυτή την περίπτωση παρατηρούνται ορισμένες μηδενικές τιμές, οι οποίες σε καμία περίπτωση δεν φτάνουν σε αριθμό αυτές της ΟΤΣ.

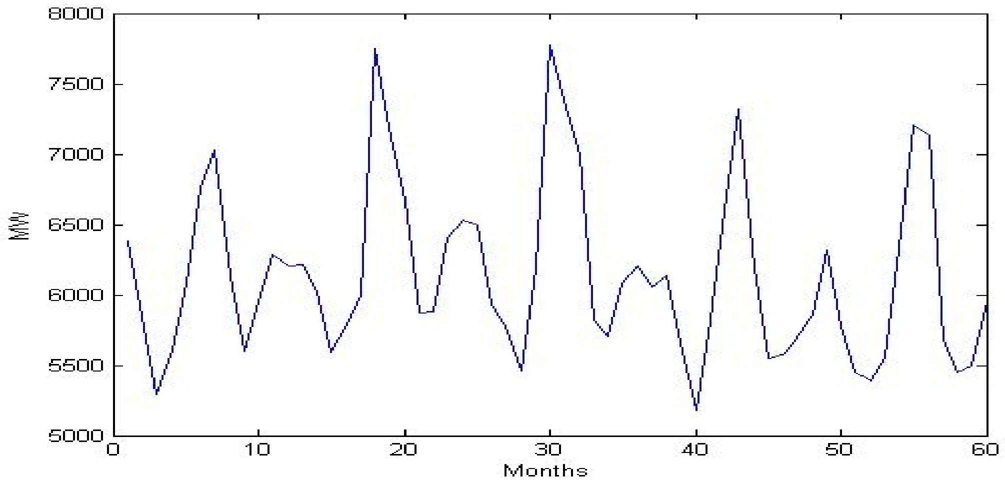
Επομένως οι μηδενικές τιμές που παρατηρούνται στα συγκεκριμένα δεδομένα οφείλονται αποκλειστικά στην αλλαγή της ώρας. Στα παρακάτω γραφήματα παρουσιάζονται και τα 5 χρόνια των ιστορικών τιμών φορτίου δίχως κάποιου είδους επεξεργασία αναφερόμενα σε επίπεδο ώρας, ημέρας και μήνα:



Σχ. 6.2.1 – Ιστορικές Τιμές Φορτίου σε επίπεδο ώρας για τα έτη 2006-2011



Σχ. 6.2.2 – Ιστορικές Τιμές Φορτίου σε επίπεδο ημέρας για τα έτη 2006-2011

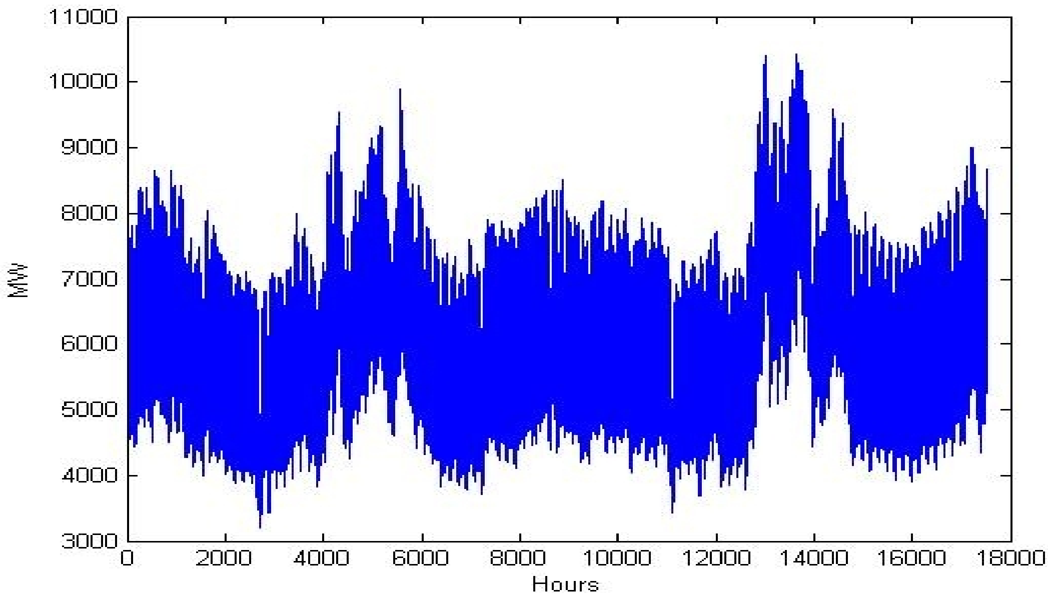


Σχ. 6.2.3- Ιστορικές Τιμές Φορτίου σε επίπεδο μήνα για τα έτη 2006-2011

Όπως αναφέρθηκε και προηγουμένως, στο Σχ. 6.2.1 παρατηρούνται 5 συνολικά μηδενικές τιμές, μία για κάθε χρόνο, και συγκεκριμένα αφορούν την 31η Μαρτίου όπου υπάρχει μία επιπλέον ώρα. Επίσης, σύμφωνα με το εν λόγω σχήμα είναι προφανής η απουσία τάσης στη συγκεκριμένη χρονοσειρά, όχι όμως και η απουσία εποχικότητας, καθώς με μία πρώτη ματιά είναι εμφανές πως τα δεδομένα ακολουθούν κάποιο συγκεκριμένο πρότυπο.

Αυτό το πρότυπο τείνει να γίνει ακόμη πιο ξεκάθαρο στο Σχ. 6.2.2, όπου οι παρατηρήσεις είναι σε επίπεδο ημέρας και ακόμη περισσότερο στο Σχ. 6.2.3 όπου πλέον είναι ξεκάθαρη η επανάληψη ενός ετήσιου προτύπου.

Στο σημείο αυτό αξίζει να αναφερθεί πως η μόνη επεξεργασία που κρίνεται απαραίτητη είναι η απαλοιφή των μηδενικών. Εξάλλου, η χρονοσειρά των ιστορικών τιμών φορτίου αποτελεί μία εκ των ανεξάρτητων μεταβλητών που χρησιμοποιούνται για την παραγωγή προβλέψεων μέσω της μεθόδου της γραμμικής παλινδρόμησης. Ενδεικτικά, παρουσιάζονται στο ακόλουθο διάγραμμα τα 2 πρώτα έτη των ιστορικών τιμών φορτίου σε επίπεδο ώρας, αφότου εφαρμόζεται η απαλοιφή μηδενικών στην αρχική χρονοσειρά με τον ίδιο τρόπο που εφαρμόστηκε και στα δεδομένα της ΟΤΣ:



Σχ. 6.2.4 – Ιστορικές Τιμές Φορτίου έπειτα από απαλοιφή μηδενικών τιμών για τα έτη 2006-2008

Επιπρόσθετα, για λόγους πληρότητας, κρίνεται απαραίτητη η αναφορά στα ποσοτικά χαρακτηριστικά της χρονοσειράς των ιστορικών τιμών φορτίου, όπως είναι η μέση τιμή, η διακύμανση, η μέγιστη και ελάχιστη τιμή καθώς και η τυπική απόκλιση. Για το λόγο αυτό, στον πίνακα που ακολουθεί, παρουσιάζονται αυτοί οι 5 βασικοί στατιστικοί δείκτες που αφορούν τη συγκεκριμένη χρονοσειρά, οι οποίοι υπολογίζονται με τη βοήθεια του προγράμματος Matlab, έπειτα από απαλοιφή των μηδενικών τιμών, καθώς και ο Συντελεστής Γραμμικής Συσχέτισης με τη χρονοσειρά της ΟΤΣ:

Ποσοτικά Χαρακτηριστικά Ιστορικών Τιμών Φορτίου
ΜΕΣΗ ΤΙΜΗ (mean) : 6144,2 MW
ΜΕΓΙΣΤΗ ΤΙΜΗ (max) : 10421 MW
ΕΛΑΧΙΣΤΗ ΤΙΜΗ (min) : 3201 MW
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (standard deviation) : 1246,3
ΔΙΑΚΥΜΑΝΣΗ (variance): 1553263,69
ΣΥΝΤΕΛΕΣΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ (Linear Correlation Coefficient): 0,5849

Πίνακας 6.2.1 - Ποσοτικά Χαρακτηριστικά Πραγματικού Φορτίου

Συγκρίνοντας τον παραπάνω πίνακα με τον Πίνακα 6.1.1 που αναφέρεται στα ποσοτικά χαρακτηριστικά της ΟΤΣ, γίνεται εμφανές ότι η χρονοσειρά των ιστορικών τιμών φορτίου παρουσιάζει μικρότερη μεταβλητότητα και πολύ λιγότερες ασυνήθιστες τιμές. Αυτό γίνεται αντιληπτό με τη βοήθεια της μέση τιμής, η οποία είναι σχεδόν ίση με το μέσο όρο της ελάχιστης και μέγιστης τιμής της χρονοσειράς, σε πλήρη αντίθεση με αυτή της ΟΤΣ. Επομένως, αποδεικνύεται και μέσω των στατιστικών δεικτών η ιδιαίτερα ευμετάβλητη μορφή της χρονοσειράς της ΟΤΣ σε σχέση με τα υπόλοιπα δεδομένα που χρησιμοποιούνται.

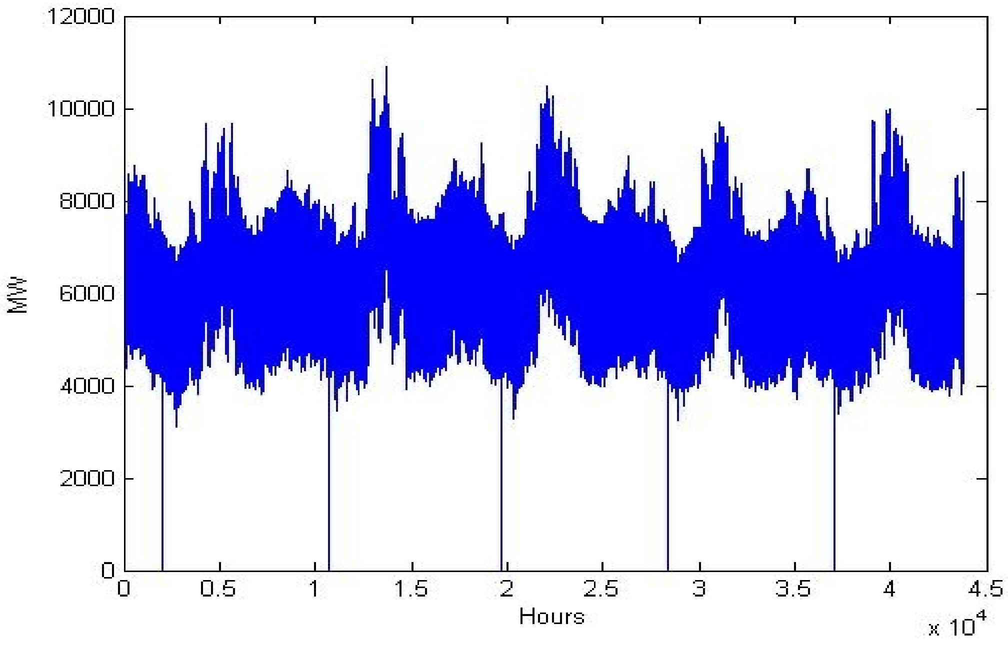
Παρατηρώντας το Συντελεστή Γραμμικής Συσχέτισης, επιβεβαιώνεται ο ισχυρισμός πως η χρονοσειρά των ιστορικών τιμών φορτίου παρουσιάζει ικανοποιητική γραμμική συσχέτιση με εκείνη της ΟΤΣ. Συγκεκριμένα, ο εν λόγω δείκτης βρίσκεται στο διάστημα [0,5, 0,7], γεγονός που πιστοποιεί την παρουσία μέσης γραμμικής συσχέτισης, σύμφωνα με τη θεωρία που παρουσιάστηκε στην Ενότητα 3.2, δικαιολογώντας την επιλογή των συγκεκριμένων δεδομένων ως ανεξάρτητη μεταβλητή στην μέθοδο της απλής ή πολλαπλής, γραμμικής παλινδρόμησης.

6.2.2 Πρόβλεψη Φορτίου

Πρόκειται για τις προβλεπόμενες τιμές του φορτίου, τις οποίες προβλέπει ο ίδιος ο ΔΕΣΜΗΕ, σε καθημερινή βάση για κάθε ώρα της ημέρας. Η συγκεκριμένη χρονοσειρά κατέχει ένα ξεχωριστό ρόλο στην διαδικασία της πολλαπλής γραμμικής παλινδρόμησης. Πιο συγκεκριμένα, χρησιμοποιείται αποκλειστικά στο τμήμα της παραγωγής των προβλέψεων κατά τη διαδικασία του rolling forecasting, και όχι στην εκπαίδευση του μοντέλου που αφορά τα in-sample δεδομένα, σε αντίθεση με τις ιστορικές τιμές φορτίου που σχετίζονται μόνο με το κομμάτι της εκπαίδευσης του μοντέλου της πολλαπλής γραμμικής παλινδρόμησης.

Η χρονοσειρά της πρόβλεψης φορτίου αφορά 5 συνολικά έτη, και συγκεκριμένα αναφέρεται στην περίοδο από την 1/1/2006 έως την 1/1/2011. Τα δεδομένα είναι σε επίπεδο ώρας και η μονάδα μέτρησης του φορτίου είναι το MW. Όπως και στις ιστορικές τιμές φορτίου, έτσι κι εδώ, δεν παρατηρείται σημαντικός όγκος μηδενικών τιμών, σε σύγκριση πάντα με τη χρονοσειρά της ΟΤΣ.

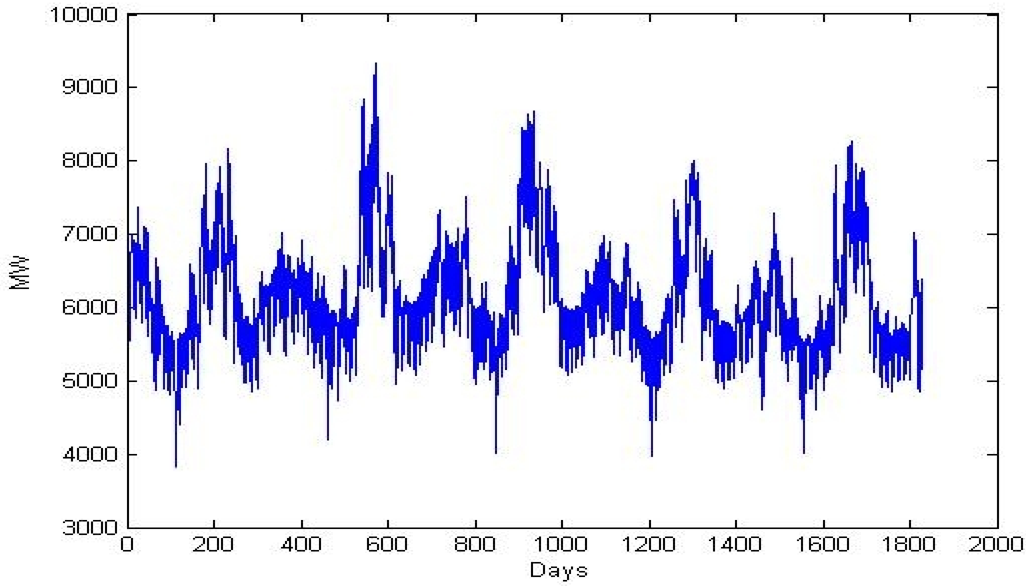
Συγκεκριμένα, παρατηρούνται μόλις 5 μηδενικές τιμές, οι οποίες οφείλονται αποκλειστικά στην αλλαγή της ώρας, ακριβώς όπως και στην περίπτωση των πραγματικών τιμών φορτίου, γεγονός απόλυτα φυσιολογικό καθώς πρόκειται για παρόμοιες χρονοσειρές. Στο γράφημα που ακολουθεί, παρουσιάζονται οι προβλεπόμενες τιμές του φορτίου, όπως τις προέβλεψε ο ΔΕΣΜΗΕ, χωρίς κανενός είδους επεξεργασία, για τα έτη 2006-2011:



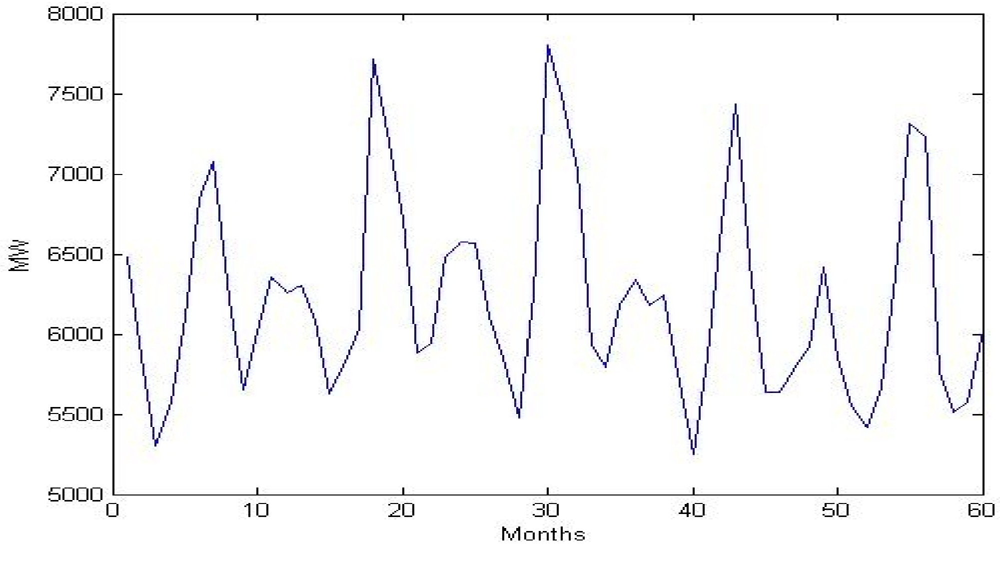
Σχ. 6.2.5 – Πρόβλεψη Φορτίου σε επίπεδο ώρας για τα έτη 2006-2011

Από το παραπάνω σχήμα, είναι εμφανέστατη η ομοιότητα των συγκεκριμένων δεδομένων με αυτά της χρονοσειράς των πραγματικών τιμών φορτίου (Σχ. 6.2.1). Και σε αυτήν την περίπτωση, δεν παρατηρείται σαφής παρουσία τάσης, όμως είναι ευδιάκριτη η παρουσία εποχικότητας, δηλαδή οι παρατηρήσεις ακολουθούν ένα συγκεκριμένο πρότυπο το οποίο δείχνει να επαναλαμβάνεται σε ετήσιο επίπεδο.

Ανάγοντας τα δεδομένα σε επίπεδο ημέρας, είναι πιο εύκολο πλέον να δικαιολογηθεί η παρουσία του προτύπου που αυτά ακολουθούν. Προχωρώντας δε και σε επίπεδο μήνα, είναι πλέον ξεκάθαρη η παρουσία εποχικότητας μέσα στο έτος. Στα δύο γραφήματα που ακολουθούν, παρουσιάζεται η χρονοσειρά της πρόβλεψης φορτίου σε επίπεδο ημέρας και μήνα για τα έτη 2006-2011, χωρίς να έχουν υποστεί κάποια περαιτέρω επεξεργασία:

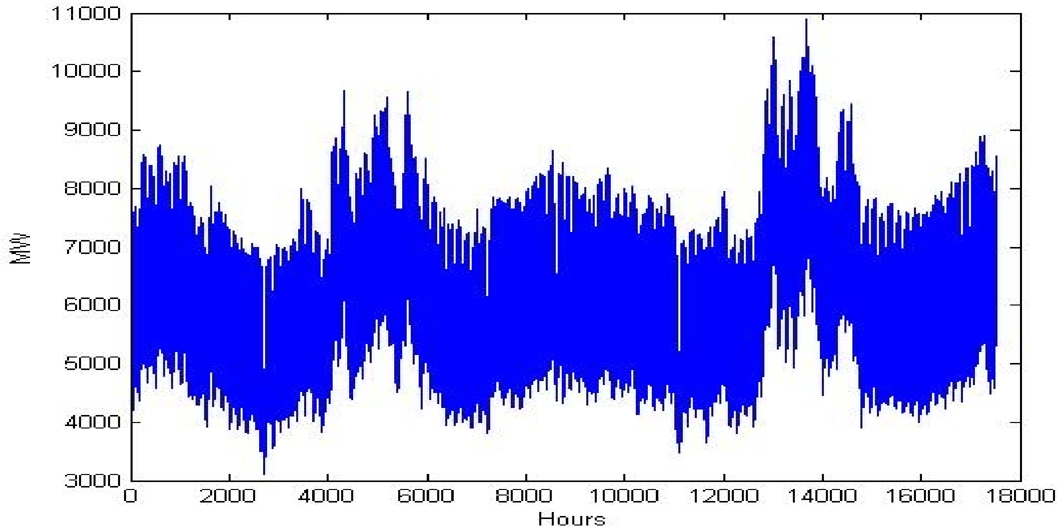


Σχ. 6.2.6 – Πρόβλεψη Φορτίου σε επίπεδο ημέρας για τα έτη 2006-2011



Σχ. 6.2.7 – Πρόβλεψη Φορτίου σε επίπεδο μήνα για τα έτη 2006-2011

Σε αυτό το σημείο, έχοντας εντοπίσει και αποσαφηνίσει τα ποιοτικά χαρακτηριστικά της χρονοσειράς της πρόβλεψης φορτίου, δεν θεωρείται απαραίτητη η όποια επεξεργασία αυτών των δεδομένων, πέρα της απαλοιφής των μηδενικών τιμών, σύμφωνα πάντα με τη μέθοδο της γραμμικής παρεμβολής που ακολουθήθηκε και στις δύο προηγούμενες χρονοσειρές (ΟΤΣ και Ιστορικές Τιμές Φορτίου). Τα 2 πρώτα έτη της χρονοσειράς που προκύπτει έπειτα από την εν λόγω επεξεργασία παρουσιάζονται στο ακόλουθο διάγραμμα σε επίπεδο ώρας:



Σχ. 6.2.8 – Πρόβλεψη Φορτίου έπειτα από απαλοιφή των μηδενικών τιμών για τα έτη 2006-2008

Όσον αφορά τα ποσοτικά χαρακτηριστικά της χρονοσειράς τα οποία είναι όμοια με αυτά που υπολογίστηκαν και στην Ενότητα 6.2.1, αυτά υπολογίζονται με τη βοήθεια του προγράμματος Matlab, και παρουσιάζονται στον ακόλουθο πίνακα έπειτα από απαλοιφή των μηδενικών τιμών:

Ποσοτικά Χαρακτηριστικά Πρόβλεψης Φορτίου
ΜΕΣΗ ΤΙΜΗ (mean) : 6210,8 MW
ΜΕΓΙΣΤΗ ΤΙΜΗ (max) : 10900 MW
ΕΛΑΧΙΣΤΗ ΤΙΜΗ (min) : 3100 MW
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (standard deviation) : 1253,5
ΔΙΑΚΥΜΑΝΣΗ (variance): 1571262,25
ΣΥΝΤΕΛΕΣΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ (Linear Correlation Coefficient): 0,5849

Πίνακας 6.2.2. Ποσοτικά Χαρακτηριστικά Πρόβλεψης Φορτίου

Όπως ήταν αναμενόμενο, οι στατιστικοί δείκτες που παρουσιάζονται στον Πίνακα 3, δεν διαφέρουν σημαντικά από αυτούς του Πίνακα 2. Αυτό είναι φυσικό, καθώς οι δύο αυτές χρονοσειρές, των ιστορικών τιμών φορτίου και της πρόβλεψης φορτίου, είναι πανομοιότυπες και παρουσιάζουν παρόμοια ποιοτικά και ποσοτικά χαρακτηριστικά.

Χαρακτηριστικό είναι το γεγονός πως ο Συντελεστής Γραμμικής Συσχέτισης, λαμβάνει ίδια τιμή με αυτή του Πίνακα 2, δηλαδή η χρονοσειρά των προβλέψεων φορτίου παρουσιάζει και αυτή μέση γραμμική συσχέτιση με τη χρονοσειρά της ΟΤΣ. Επομένως, αυτή είναι μία επιπλέον ένδειξη η οποία φανερώνει την υψηλή ομοιότητα των δύο αυτών χρονοσειρών, αλλά και τον παρόμοιο αντίκτυπο που έχουν στο μέγεθος της ΟΤΣ.

6.2.3 Ιστορικές Τιμές Αργού Πετρελαίου

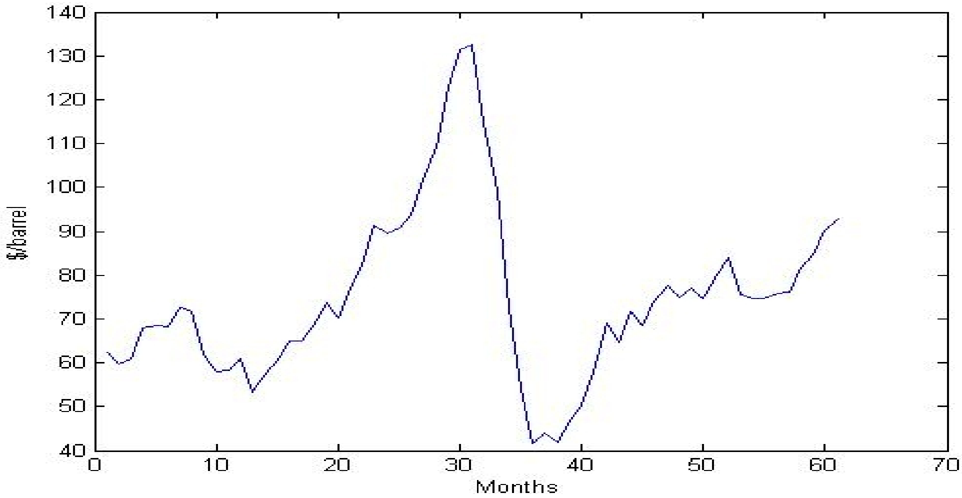
Ένα πολύ σημαντικό στοιχείο το οποίο καθορίζει εν πολλοίς το μέγεθος της ΟΤΣ είναι η τιμή των καυσίμων που χρησιμοποιούνται στην παραγωγή ηλεκτρικής ενέργειας. Όπως αναφέρθηκε και στην Ενότητα 1.1, το Ενεργειακό Μίγμα της Ελλάδας συμπεριλαμβάνει το αργό πετρέλαιο σε ένα ποσοστό της τάξης του 7% κατά το έτος 2008.

Όμως, παρατηρείται μια προοδευτική μείωση της συμμετοχής του αργού πετρελαίου στο Ενεργειακό Μίγμα της χώρας, με αποκορύφωμα το έτος 2011 κατά το οποίο παρατηρείται μηδενική συμμετοχή της πετρελαϊκής παραγωγής στην Αγορά Ενέργειας. Αυτό οφείλεται και στην ανάγκη για «καθαρότερες» μορφές ενέργειας, αλλά και στην συνεχώς αυξανόμενη τιμή του αργού πετρελαίου το οποίο προμηθεύεται η Ελλάδα κυρίως από τις ΗΠΑ.

Παρόλα αυτά, η χρήση της τιμής του πετρελαίου ως ανεξάρτητη μεταβλητή στη διαδικασία της γραμμικής παλινδρόμησης, παρουσιάζει ενδιαφέρον, καθώς τα διαθέσιμα δεδομένα δεν αφορούν ουσιαστικά το έτος 2011, αλλά μέχρι και το πέρας του 2010, όπου η πετρελαϊκή παραγωγή κατείχε ακόμη κάποιο μικρό ποσοστό στο ενεργειακό μίγμα.

Επιπλέον, κατά τα πρώτα έτη στα οποία αναφέρονται τα δεδομένα (2006 και μετά), η συμμετοχή του αργού πετρελαίου στην παραγωγή ηλεκτρικής ενέργειας είναι σαφέστατα μείζονος σημασίας.

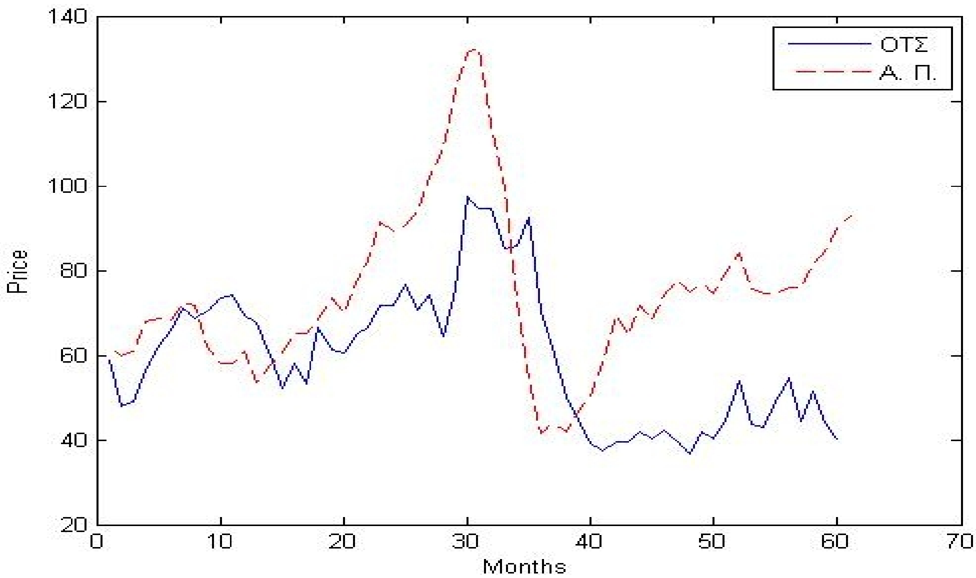
Όσον αφορά τη χρονοσειρά των τιμών του αργού πετρελαίου, πρόκειται για μηνιαία δεδομένα που αφορούν την περίοδο από τον Ιανουάριο του 2006 έως τον Ιανουάριο του 2011. Η τιμή του πετρελαίου είναι σε δολάρια ανά βαρέλι ($ per barrel). Στο παρακάτω γράφημα, παρουσιάζονται οι εν λόγω παρατηρήσεις δίχως κάποιου είδους επεξεργασία:



Σχ. 6.2.9 – Μηνιαίες τιμές Αργού Πετρελαίου για τα έτη 2006-2011

Πηγή: http://www.indexmundi.com

Σύμφωνα με το παραπάνω διάγραμμα, είναι εμφανής η ανοδική πορεία που ακολουθεί η τιμή του αργού πετρελαίου μέσα στη 5ετία 2006-2011, με αποκορύφωμα τα έτη 2007-2008, όπου η τιμή του βαρελιού έφτασε περίπου τα 130$. Γενικά, είναι αισθητή η παρουσία ανοδικής τάσης στη χρονοσειρά, σε αντίθεση με την παρουσία εποχικότητας, η οποία φαίνεται να απουσιάζει. Ιδιαίτερο ενδιαφέρον παρουσιάζει μία πρώτη γραφική συσχέτιση της χρονοσειράς του αργού πετρελαίου με τη χρονοσειρά της ΟΤΣ, έχοντας αναγάγει τα δεδομένα σε επίπεδο μήνα. Οι δύο αυτές χρονοσειρές παρουσιάζονται στο ακόλουθο κοινό διάγραμμα, όπου για λόγους απλότητας θεωρείται μία κοινή μονάδα μέτρησης τιμής και όχι οι πραγματικές:



Σχ. 6.2.10 – Χρονοσειρές ΟΤΣ και Α.Π. σε επίπεδο μήνα για τα έτη 2006-2011

Όπως είναι εμφανές στο παραπάνω διάγραμμα, παρατηρείται ομοιότητα στη μορφή των δύο χρονοσειρών, και ιδιαίτερα στο σημείο όπου η τιμή του αργού πετρελαίου λαμβάνει πολύ υψηλές τιμές (2007-2008), παρατηρείται ανάλογη συμπεριφορά και από τη χρονοσειρά της ΟΤΣ.

Πέραν όμως αυτού του συγκεκριμένου σημείου, δεν θα ήταν υπερβολή να ειπωθεί ότι οι δύο χρονοσειρές ακολουθούν γενικώς παρόμοιες πορείες μέσα στο χρονικό διάστημα το οποίο εξετάζεται. Επομένως, μπορεί να θεωρηθεί δεδομένη η υψηλή εξάρτηση που παρουσιάζει το μέγεθος της ΟΤΣ από την τιμή του αργού πετρελαίου.

Στο σημείο αυτό, αξίζει να γίνει αναφορά και στα ποσοτικά χαρακτηριστικά της χρονοσειράς των τιμών του αργού πετρελαίου. Υπολογίζονται όπως και προηγουμένως με τη βοήθεια του προγράμματος Matlab οι ίδιοι έξι βασικοί στατιστικοί δείκτες, οι οποίοι παρουσιάζονται στον παρακάτω πίνακα:

Ποσοτικά Χαρακτηριστικά Τιμής Αργού Πετρελαίου
ΜΕΣΗ ΤΙΜΗ (mean) : 74.9438 $/barrel
ΜΕΓΙΣΤΗ ΤΙΜΗ (max) : 135.55 $/barrel
ΕΛΑΧΙΣΤΗ ΤΙΜΗ (min) : 41,53 $/barrel
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (standard deviation) : 19,6599
ΔΙΑΚΥΜΑΝΣΗ (variance): 386,5119
ΣΥΝΤΕΛΕΣΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ (Linear Correlation Coefficient): 0,4279

Πίνακας 6.2.3 -. Ποσοτικά Χαρακτηριστικά Τιμής Αργού Πετρελαίου

Αξίζει να σημειωθεί το γεγονός ότι η χρονοσειρά των τιμών του αργού πετρελαίου δεν χρειάζεται να υποστεί κάποιου είδους επεξεργασία, όπως για παράδειγμα είναι η απαλοιφή μηδενικών τιμών, καθώς τα δεδομένα δεν παρουσιάζουν κάποια ανωμαλία η οποία χρήζει περαιτέρω επεξεργασίας.

Επιπλέον, είναι σημαντικό να γίνει ειδική αναφορά στο Συντελεστή Γραμμικής Συσχέτισης, ο οποίος λαμβάνει την τιμή 0,4279, τιμή η οποία είναι αρκετά χαμηλότερη από αυτές που υπολογίστηκαν για τις ιστορικές τιμές αλλά και την πρόβλεψη φορτίου.

Σε αυτή την περίπτωση, σύμφωνα πάντα με τη θεωρία που αναπτύχθηκε για το συγκεκριμένο στατιστικό δείκτη στην Ενότητα 3.2, υπάρχει ασθενής γραμμική συσχέτιση μεταξύ της χρονοσειράς του αργού πετρελαίου και αυτής της ΟΤΣ. Αυτό ήταν αναμενόμενο, καθώς όπως αναφέρθηκε και στην αρχή της παρούσας ενότητας, η συμμετοχή του αργού πετρελαίου στην ενεργειακή παραγωγή ακολούθησε πτωτική πορεία κατά τη διάρκεια του χρονικού διαστήματος που εξετάζεται.

6.2.4 Ιστορικές Τιμές Φυσικού Αερίου

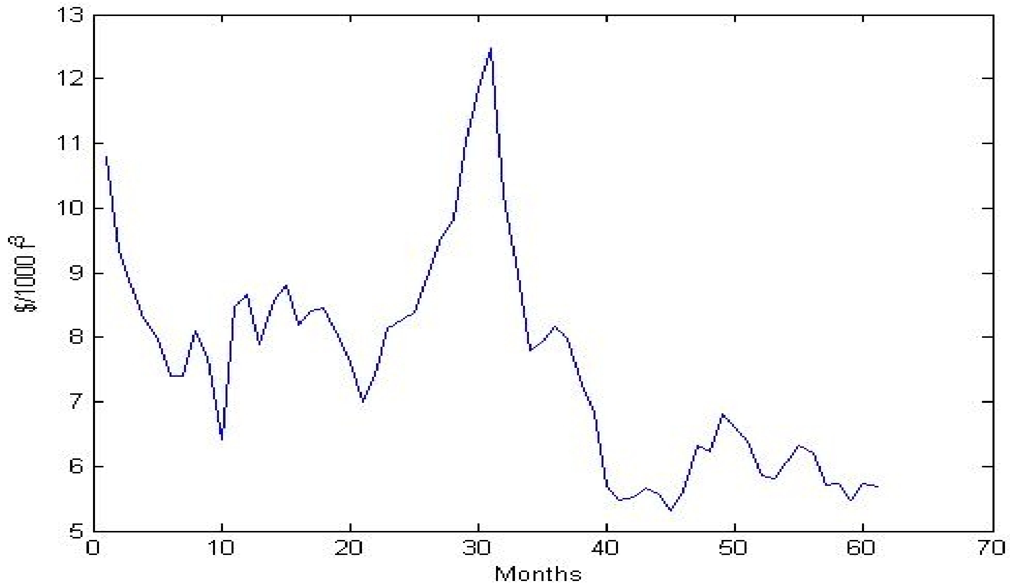
Το Φυσικό Αέριο (natural gas) είναι ένα μίγμα υδρογονανθράκων, το οποίο εξάγεται από υπόγειες κοιλότητες και λόγω των ιδιοτήτων του συγκαταλέγεται στην κατηγορία των οικολογικών καυσίμων. Αυτό οφείλεται κυρίως στο γεγονός ότι η καύση του έχει λιγότερο επιβλαβείς συνέπειες για το περιβάλλον, όπως για παράδειγμα παράγει μικρότερες ποσότητες διοξειδίου του άνθρακα για κάθε μονάδα παραγόμενης ενέργειας. Η προμήθεια του συγκεκριμένου καυσίμου στην Ελλάδα γίνεται κυρίως από τη Ρωσία και την Αλγερία.

Ο ρόλος του Φυσικού Αερίου στην παραγωγή ηλεκτρικής ενέργειας είναι αδιαμφισβήτητος. Κατέχει σημαντικό ποσοστό στον ενεργειακό μίγμα της Ελλάδας, το οποίο συνεχώς αυξάνεται, ιδιαίτερα τα τελευταία χρόνια. Χαρακτηριστικό είναι το γεγονός πως ενώ το 2008 το Φυσικό Αέριο κατείχε το 26% περίπου του ενεργειακού μίγματος, μετά το πέρας τριών χρόνων, δηλαδή το έτος 2011, το ποσοστό αυτό σκαρφάλωσε στο 31% (Ενότητα 1.1), εκμεταλλευόμενο και την απομάκρυνση του αργού πετρελαίου από την παραγωγή ηλεκτρικής ενέργειας.

Λόγω των παραπάνω, η τιμή του Φυσικού Αερίου είναι βέβαιο ότι επηρεάζει το μέγεθος της ΟΤΣ και μάλιστα σε αρκετά υψηλό βαθμό. Σε αντίθεση με το αργό πετρέλαιο, η χρήση του Φυσικού Αερίου στην παραγωγή ενέργειας αυξάνεται συνεχώς, με συνέπεια να θεωρείται ένας εκ το βασικότερων παραγόντων καθορισμού του ενεργειακού μίγματος της χώρας.

Επομένως, παρουσιάζει ιδιαίτερο ενδιαφέρον η επιλογή του ως ανεξάρτητη μεταβλητή στην μέθοδο της πολλαπλής ή μη, γραμμικής παλινδρόμησης, αλλά και η μελέτη της επίδρασης που έχει στην διαμόρφωση των τιμών των παρατηρήσεων της χρονοσειράς της ΟΤΣ.

Όσον αφορά τα δεδομένα, πρόκειται για μηνιαίες παρατηρήσεις της τιμής του Φ.Α., οι οποίες αναφέρονται στην περίοδο από τον Ιανουάριο του 2006 έως τον Ιανουάριο του 2011. Η τιμή του Φ.Α. μετριέται σε δολάρια ανά 1000 κυβικά πόδια ($ per 1000 cubic feet). Στο γράφημα που ακολουθεί, παρουσιάζεται οι εν λόγω παρατηρήσεις δίχως κάποιου είδους επεξεργασία:



Σχ. 6.2.11 – Μηνιαίες τιμές Φυσικού Αερίου για τα έτη 2006-2011

Πηγή: http://www.economagic.com

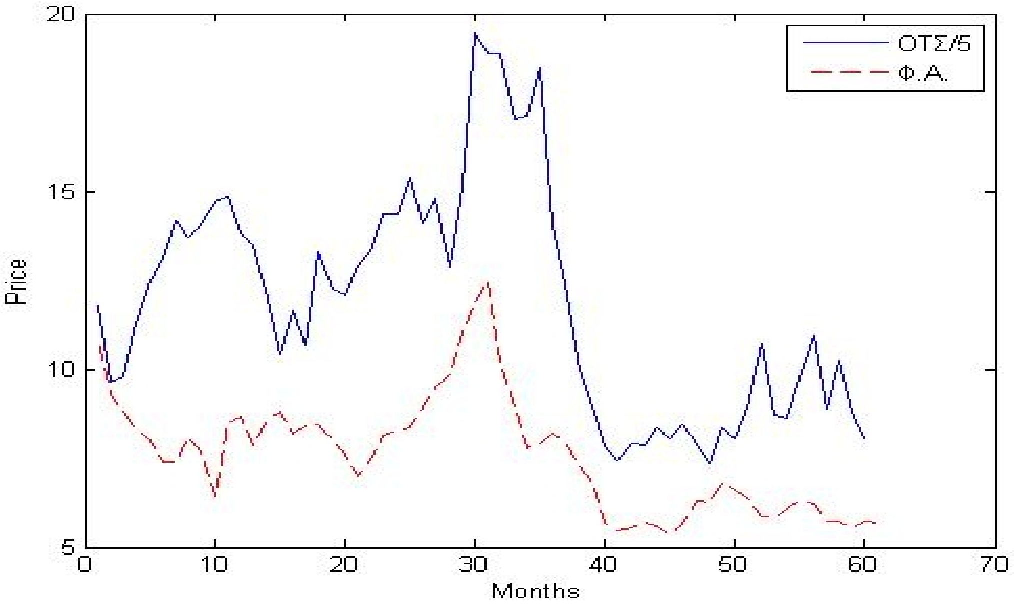
Όπως είναι εμφανές από το παραπάνω διάγραμμα, παρατηρείται και εδώ μία απότομη αύξηση της τιμής του Φ.Α. κατά το έτος 2008, όπως συνέβη και με την τιμή του αργού πετρελαίου. Η γενική εικόνα αυτής της χρονοσειράς, παρουσιάζει μη εμφανή πτωτική τάση, όχι όμως και κάποιου είδους εποχικότητα, καθώς οι παρατηρήσεις δεν φαίνεται να ακολουθούν κάποιο συγκεκριμένο πρότυπο μέσα στο χρόνο.

Σε αυτό το σημείο κρίνεται απαραίτητο να εξετασθούν ταυτόχρονα, όσον αφορά τις γραφικές τους αναπαραστάσεις, οι χρονοσειρές της ΟΤΣ αλλά και του Φ.Α., προκειμένου να εξαχθεί ένα πρώτο συμπέρασμα σχετικά με το κατά πόσο επηρεάζει η πορεία της τιμής του Φ.Α αυτήν της ΟΤΣ.

Φυσικά, τον τελικό λόγο για το κατά πόσο μπορεί να επηρεάζει η συγκεκριμένη ανεξάρτητη μεταβλητή την διαμόρφωση των τιμών της χρονοσειράς της ΟΤΣ, τον έχουν οι στατιστικοί δείκτες, οι οποίοι θα υπολογιστούν στην συνέχεια της παρούσας ενότητας.

Για να είναι περισσότερη δόκιμη η συσχέτιση των δύο χρονοσειρών, η χρονοσειρά της ΟΤΣ διαιρείται με τον αριθμό 5, προκειμένου η κλίμακα των τιμών να είναι στα ίδια επίπεδα με αυτή των τιμών του Φ.Α., καθιστώντας έτσι περισσότερο εύκολη την σύγκριση μεταξύ των δύο εξεταζόμενων χρονοσειρών.

Στο παρακάτω διάγραμμα παρουσιάζονται οι δύο αυτές χρονοσειρές, με την ΟΤΣ να έχει αναχθεί σε επίπεδο μήνα, αλλά και θεωρώντας κοινή μονάδα μέτρησης της τιμής για λόγους ευκολίας σύγκρισης:



Σχ. 6.2.12 – Χρονοσειρές ΟΤΣ και Φ.Α. σε επίπεδο μήνα για τα έτη 2006-2011

Το συμπέρασμα που εξάγεται από το παραπάνω γράφημα είναι ότι σαφώς και υπάρχει υψηλή συσχέτιση της πορείας των τιμών του Φ.Α. με την πορεία της ΟΤΣ. Μάλιστα, η χρονοσειρά των τιμών του Φ.Α. φαίνεται να ακολουθεί καλύτερα εκείνη της ΟΤΣ, από ότι η χρονοσειρά των τιμών του αργού πετρελαίου. Αυτό φυσικά οφείλεται στο γεγονός που αναφέρθηκε στην αρχή της συγκεκριμένης ενότητας, δηλαδή στην πολύ μεγάλη συμμετοχή που έχει το Φ.Α. στο ελληνικό ενεργειακό μίγμα.

Στο σημείο αυτό, είναι σημαντικό να αναφερθεί πως η συγκεκριμένη χρονοσειρά δεν χρήζει περεταίρω επεξεργασίας, καθώς δεν παρουσιάζει κάποιου είδους ανωμαλία, όπως για παράδειγμα μηδενικές τιμές.

Επιπλέον, όπως έχει ήδη αναφερθεί, κρίνεται απαραίτητη η αναφορά σε ορισμένους βασικούς στατιστικού δείκτες, όμοιους με εκείνους που αναφέρθηκαν μέχρι τώρα για τις υπόλοιπες ανεξάρτητες μεταβλητές παλινδρόμησης, που χαρακτηρίζουν τη χρονοσειρά των τιμών Φ.Α. δίχως αυτή να έχει υποστεί οποιουδήποτε είδους επεξεργασία, προκειμένου να δικαιολογηθεί μέσω αυτών η καταλληλότητα της παρούσας χρονοσειράς ως ανεξάρτητη μεταβλητή παλινδρόμησης. Οι εν λόγω ποσοτικοί δείκτες παρουσιάζονται στον ακόλουθο πίνακα:

Ποσοτικά Χαρακτηριστικά Τιμής Αργού Πετρελαίου
ΜΕΣΗ ΤΙΜΗ (mean) : 7.5920 $/1000 f^3
ΜΕΓΙΣΤΗ ΤΙΜΗ (max) : 12.4800 $/1000 f^3
ΕΛΑΧΙΣΤΗ ΤΙΜΗ (min) : 5.3200 $/1000 f^3
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ (standard deviation) : 1.6628
ΔΙΑΚΥΜΑΝΣΗ (variance): 2.7651
ΣΥΝΤΕΛΕΣΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ (Linear Correlation Coefficient): 0,7450

Πίνακας 6.2.4 - Ποσοτικά Χαρακτηριστικά Τιμής Φυσικού Αερίου

Παρατηρώντας τον παραπάνω πίνακα, εξάγεται ένα πολύ σημαντικό συμπέρασμα σχετικά με το κατά πόσο επηρεάζονται οι τιμές της ΟΤΣ από την τιμή του Φυσικού Αερίου. Σε αντίθεση με τις υπόλοιπες ανεξάρτητες μεταβλητές που παρουσιάστηκαν μέχρι τώρα στο παρών κεφάλαιο, η τιμή του φυσικού αερίου παρουσιάζει ιδιαίτερα υψηλή τιμή όσον αφορά το Συντελεστή Γραμμικής Συσχέτισης. Συγκεκριμένα ο εν λόγω στατιστικός δείκτης λαμβάνει την τιμή 0,745, τιμή η οποία βρίσκεται στο διάστημα [0,7, 0,8], και επομένως παρουσιάζει ισχυρή γραμμική συσχέτιση, αν ληφθούν υπόψη τα όσα λέχθηκαν για τις πιθανές τιμές του συγκεκριμένου δείκτη στα πλαίσια της Ενότητας 3.2.

Σε αυτό το σημείο, συνοψίζοντας τα όσα αναφέρθηκαν για τις 4 ανεξάρτητες μεταβλητές που παρουσιάστηκαν στην παρούσα ενότητα, εξάγεται το συμπέρασμα πως η επιλογή των συγκεκριμένων δεδομένων κρίνεται ορθή, κυρίως μέσω του υπολογισμού αλλά και της ερμηνείας του Συντελεστή Γραμμικής Συσχέτισης, δικαιώνοντας και παράλληλα αιτιολογώντας τις αρχικές υποθέσεις και παραδοχές που έγιναν προκειμένου να στηριχθεί η επιλογή τους ως ανεξάρτητες μεταβλητές παλινδρόμησης ανάμεσα σε τόσους άλλους παράγοντες που επηρεάζουν το μέγεθος της ΟΤΣ.

Κεφάλαιο 7. Αποτελέσματα Προσομοίωσης

Στο παρών Κεφάλαιο θα παρουσιαστούν αλλά και θα αναλυθούν όλα τα αποτελέσματα που προέκυψαν από τη διαδικασία της προσομοίωσης που περιγράφηκε στο Κεφάλαιο 5. Στην διαδικασία της προσομοίωσης χρησιμοποιούνται τα μοντέλα χρονοσειρών, και συγκεκριμένα η μέθοδος Naive, οι μέθοδοι εξομάλυνσης όπως είναι οι SES, Holt και Damped, η συνδυαστική μέθοδος Theta, αλλά και μέθοδοι παλινδρόμησης (LRL), όπως αυτές περιγράφηκαν στο Κεφάλαιο 4, δεχόμενες σαν είσοδο τη χρονοσειρά της Οριακής Τιμής Συστήματος (ΟΤΣ).

Ως κριτήριο αξιολόγησης των αποτελεσμάτων που προκύπτουν εφαρμόζοντας τις παραπάνω μεθόδους πρόβλεψης, θεωρείται η ελαχιστοποίηση της μέσης τιμής του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE), το οποίο παρουσιάστηκε αναλυτικά στην Ενότητα 3.2.6.

Αρχικά, είναι απαραίτητος ο διαχωρισμός των αποτελεσμάτων της προσομοίωσης σε τέσσερις κατηγορίες, οι οποίες διαφοροποιούνται με βάση τον ποιοτικό δείκτη της εποχικότητας και τον τρόπο με τον οποίο προσεγγίζεται κάθε φορά στα πλαίσια της διαδικασίας του rolling forecasting. Οι κατηγορίες αυτές είναι οι ακόλουθες:

i. Κανενός είδους επεξεργασία εποχικότητας (no seasonality index – no SI).

ii. Μονή εποχικότητα (single seasonality) σε επίπεδο ώρας μέσα στην ημέρα (SI=24).

iii. Μονή εποχικότητα (single seasonality) σε επίπεδο ώρας μέσα στην εβδομάδα (SI=168).

iv. Διπλή εποχικότητα (double seasonality) σε επίπεδο ώρας μέσα στην ημέρα και στην εβδομάδα (SI_1=24, SI_2=168).

Η επίδραση της εποχικότητας στη χρονοσειρά της ΟΤΣ, όπως έχει αναφερθεί και στο Κεφάλαιο 6, είναι ιδιαίτερη σημαντική και επηρεάζει εν πολλοίς τα αποτελέσματα που παράγονται από την προσομοίωση. Επιπλέον, σε κάθε μία από τις τέσσερις παραπάνω περιπτώσεις, το μέγεθος της εποχικότητας επιδρά με διαφορετικό τρόπο σε κάθε μέθοδο πρόβλεψης, μεταβάλλοντας άλλες φορές περισσότερο και άλλες λιγότερο τις τιμές των αποτελεσμάτων.

Στη συνέχεια ακολουθεί παρουσίαση καθώς και αναλυτική περιγραφή των αποτελεσμάτων που προέκυψαν από τη διαδικασία της προσομοίωσης για τις διάφορες μεθόδους πρόβλεψης που χρησιμοποιήθηκαν, και για τις 4 κατηγορίες που αφορούν τον τρόπο αντιμετώπισης της εποχικότητας.

7.1 Αποτελέσματα δίχως απαλοιφή της εποχικότητας

Πρόκειται για την 1^{η} ομάδα αποτελεσμάτων, και ίσως την απλούστερη από θέμα πολυπλοκότητας. Η προεπεξεργασία των δεδομένων, και συγκεκριμένα της ΟΤΣ, δεν περιλαμβάνει απαλοιφή της εποχικότητας, παρά μόνο την κύλιση των τιμών, ώστε η πρώτη διαθέσιμη παρατήρηση να είναι στις 14:00 η ώρα, καθώς και την απαλοιφή των μηδενικών τιμών με τη μεθοδολογία που αναπτύχθηκε στα Κεφάλαια 5 και 6. Επίσης, κατά το στάδιο της μετεπεξεργασίας, η χρονοσειρά των προβλέψεων που προκύπτει απλώς περιορίζεται από το ανώτατο όριο τιμών που τίθεται από τη ΡΑΕ (150€/MWh).

Αξίζει να σημειωθεί, πως στο σημείο εκείνο που επιλέγεται ως in-sample δείγμα, το 60% των διαθέσιμων παρατηρήσεων της χρονοσειράς της ΟΤΣ, παρουσιάζεται η ανάγκη της διόρθωσης αυτού του μεγέθους, το οποίο συμβολίζεται με N, κυρίως για τις ακόλουθες περιπτώσεις όπου γίνεται επεξεργασία της εποχικότητας.

Αυτό οφείλεται στο γεγονός ότι ολισθαίνοντας τις παρατηρήσεις ώστε η πρώτη από αυτές να αναφέρεται στις 14:00 η ώρα, όπου σύμφωνα με το ΔΕΣΜΗΕ γίνεται η επίλυση του ΗΕΠ και η δημοσιοποίηση των αποτελεσμάτων του για το επόμενο 24ωρο, παρουσιάζεται ολίσθηση και στους εποχιακούς δείκτες, με αποτέλεσμα να μην υπάρχει συγχρονισμός μεταξύ παρατηρήσεων και δεικτών.

Αξίζει να σημειωθεί πως στη συγκεκριμένη κατηγορία, όπου δεν πραγματοποιείται κάποιου είδους επεξεργασία επί της εποχικότητας, θα ήταν δυνατό να μη διορθωθεί το N, και να παραμείνει ως έχει, αλλά για λόγους ομοιογένειας μεταξύ των αποτελεσμάτων των τεσσάρων κατηγοριών, αποφασίστηκε να εφαρμοσθεί η διόρθωση σε κάθε μία από αυτές.

Όπως αναφέρθηκε στην αρχή του παρόντος κεφαλαίου, κριτήριο αξιολόγησης των αποτελεσμάτων των μεθόδων πρόβλεψης που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης, είναι ο ποσοτικός στατιστικός δείκτης του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE), καθώς είναι εκείνος ο δείκτης που μπορεί να αποτυπώσει καλύτερα την ακρίβεια των διάφορων μεθόδων πρόβλεψης, εξασφαλίζοντας χρησιμότερα συμπεράσματα που βοηθούν στην επιλογή του βέλτιστου μοντέλου.

Στον συγκεντρωτικό πίνακα που ακολουθεί, παρουσιάζονται οι μέσες τιμές (%) του sMAPE για τις 6 μεθόδους πρόβλεψης (Naive, LRL, Theta, SES, Holt, Damped), όπως αυτές προέκυψαν από τη διαδικασία του rolling forecasting:

	ΜΕΘΟΔΟΣ	Naive	LRL	Theta	SES	Holt	Damped
No SI	***sMAPE* (%)**	31,0848	42,5993	31,0893	31,0904	44,6469	31,3484

Πίνακας 7.1.1. - Μέσοι Δείκτες sMAPE χωρίς αποεποχικοποίηση

Παρατηρώντας τον παραπάνω πίνακα, είναι δυνατό να εξαχθούν αρκετά συμπεράσματα, όπως για παράδειγμα το κατά πόσο καλά δείχνει να ανταποκρίνεται κάθε μέθοδος στη χρονοσειρά της ΟΤΣ. Σε πρώτη φάση, είναι προφανές πως η μέθοδος της Απλής Γραμμικής Παλινδρόμησης, δεν ακολουθεί ικανοποιητικά τα δεδομένα, και παρουσιάζει sMAPE της τάξης του 42,6%, ποσοστό ιδιαίτερα υψηλό το οποίο δε μπορεί σε καμία περίπτωση να αποτελέσει αξιόλογο αποτέλεσμα. Αυτό οφείλεται στο γεγονός ότι η ευθεία της απλής γραμμικής παλινδρόμησης, δεν είναι δυνατό να ανταποκριθεί γρήγορα στις ταχείες και απότομες μεταβολές της ΟΤΣ.

Ακόμη πιο απογοητευτικό είναι το αποτέλεσμα που προκύπτει για το sMAPE χρησιμοποιώντας το μοντέλο γραμμικής τάσης Holt. Για τη βέλτιστη επιλογή των παραμέτρων α και β του μοντέλου Holt, χρησιμοποιείται η μη γραμμική βελτιστοποίηση (δυαδική), όπως περιγράφηκε στο Κεφάλαιο 4. Στην περίπτωση αυτή, το σφάλμα αγγίζει το 45%, γεγονός το οποίο καθιστά τη μέθοδο Holt, αναξιόπιστη για την παραγωγή προβλέψεων όσον αφορά τη χρονοσειρά της ΟΤΣ.

Αυτό το απογοητευτικό αποτέλεσμα, είναι πιθανό να οφείλεται στην επιλογή των παραμέτρων α και β, οι οποίες προκύπτουν από τη δυαδική βελτιστοποίηση της μεθόδου. Συγκεκριμένα, η παράμετρος α λαμβάνει πάρα πολύ υψηλή τιμή (0,9661), ενώ η παράμετρος β λαμβάνει τιμή σημαντικά μικρότερη και μάλιστα κοντά στο μηδέν (0,0103).

Το γεγονός ότι η παράμετρος β παίρνει τόσο μικρή τιμή, οφείλεται στο ότι η χρονοσειρά της ΟΤΣ συμπεριφέρεται ως στάσιμη χρονοσειρά (παραμένει σταθερή ως προς το μέσο), και επομένως ορθώς το β λαμβάνει τόσο μικρή τιμή από τη στιγμή που δεν παρουσιάζεται εμφανής τάση στα δεδομένα. Όσον αφορά την παράμετρο εξομάλυνσης α, είναι πιθανό μικρότερη τιμή αυτής να έδινε και καλύτερο αποτέλεσμα, όμως η περαιτέρω ανάλυση αυτού του εγχειρήματος είναι έξω από τα πλαίσια της παρούσας Διπλωματικής.

Αξιοσημείωτο είναι το γεγονός πως οι τέσσερις εναπομείναντες μέθοδοι (Naive, Theta, SES, Damped), παρουσιάζουν παραπλήσια αποτελέσματα. Συγκεκριμένα, ο δείκτης του sMAPE κινείται κοντά στην τιμή 31%, με το μοντέλο της Damped να παρουσιάζει ελαφρώς χειρότερο δείκτη από τα υπόλοιπα τρία μοντέλα (31,3484). Μεταξύ των τριών καλύτερων μοντέλων, η Naive είναι εκείνη που κόβει το νήμα πρώτη, αφού παρουσιάζει sMAPE της τάξης του 31,0848%, με δεύτερη καλύτερη Theta και sMAPE 31.0893%, και τέλος τρίτη τη SES με σφάλμα 31,0904%.

Το γεγονός, ότι έστω και οριακά το μοντέλο Naive παρουσιάζει το καλύτερο αποτέλεσμα, δεν αποτελεί έκπληξη. Ήδη έχει αναφερθεί σε προηγούμενο κεφάλαιο, ότι η μέθοδος Naive μπορεί να ανταποκριθεί ιδιαίτερα ικανοποιητικά σε περιπτώσεις όπου η υπό πρόβλεψη χρονοσειρά παρουσιάζει υψηλή αστάθεια και μεταβλητότητα, και επομένως είναι πολύ δύσκολο να την ακολουθήσει ακριβώς κάποιο μοντέλο πρόβλεψης.

Σχετικά με τα μοντέλα Theta και SES, παρατηρείται σχεδόν όμοια συμπεριφορά, γεγονός απόλυτα λογικό, αφού η γραμμή Theta (2) είναι στην ουσία το μοντέλο SES, το οποίο επιδέχεται μικρής βελτίωσης χάρη στη γραμμή Theta(0), η οποία δεν είναι άλλη από την ευθεία απλής γραμμικής παλινδρόμησης (LRL). Η βελτίωση αυτή οφείλεται στο ότι η γραμμή Theta(0), συλλαμβάνει την μακροπρόθεσμη τάση που υπάρχει στη χρονοσειρά, όμως ακριβώς επειδή η χρονοσειρά της ΟΤΣ δεν παρουσιάζει τάση, δεν υπάρχει και ουσιαστική βελτίωση στο δείκτη του sMAPE, σε σύγκριση πάντα με το μοντέλο SES.

Σε αυτό το σημείο αξίζει να σημειωθεί, πως η επιλογή των παραμέτρων εξομάλυνσης στα μοντέλα των SES και Damped, αλλά και στη γραμμή Theta(2) που εμπεριέχεται στο μοντέλο Theta, γίνεται μέσω της μη γραμμικής (δυαδικής) βελτιστοποίησης, ακριβώς όπως και στην περίπτωση της μεθόδου Holt που αναφέρθηκε προηγουμένως.

Η επιλογή της δυαδικής αναζήτησης έναντι της γραμμικής γίνεται για λόγους ταχύτητας και μικρότερης πολυπλοκότητας κατά τη διάρκεια της εφαρμογής της διαδικασίας της προσομοίωσης με τις συγκεκριμένες μεθόδους, μέσω του προγράμματος Matlab, το οποίο χρησιμοποιείται για την παραγωγή των αποτελεσμάτων.

7.2 Αποτελέσματα με μονή εποχικότητα (SI=24)

Πρόκειται για μια πρώτη προσπάθεια απαλοιφής της εποχικότητας από τη χρονοσειρά της ΟΤΣ, στοχεύοντας στην παραγωγή πιο ικανοποιητικών αποτελεσμάτων σε σύγκριση με αυτών της 1ης κατηγορίας, στα πλαίσια της οποίας δεν επιχειρήθηκε καμία ενέργεια απαλοιφής της εποχικότητας.

Η όλη διαδικασία βασίζεται στην υπόθεση πως υπάρχει μόνο εποχικότητα σε επίπεδο ώρας μέσα στην ημέρα, και επομένως είναι απαραίτητη 24 εποχικοί δείκτες (seasonality indexes - SI), αδιαφορώντας για την εμφάνιση εποχικότητας σε εβδομαδιαία βάση. Η παραπάνω υπόθεση, σημαίνει παράλληλα ότι το ημερήσιο πρότυπο το οποίο θεωρείται ότι υπάρχει, επαναλαμβάνεται αυτούσιο για κάθε ημέρα μέσα στην εβδομάδα, δίχως να λαμβάνει υπόψη τις υπαρκτές διαφορές που παρουσιάζονται ανάμεσα στα πρότυπα που ακολουθούν οι καθημερινές σε σχέση με τα Σαββατοκύριακα.

Το στάδιο της προεπεξεργασίας εμπλουτίζεται με τη διαδικασία απαλοιφής της εποχικότητας και τον προσδιορισμό των 24 εποχικών δεικτών που είναι απαραίτητοι για την παραγωγή μιας αποεποχικοποιημένης χρονοσειράς της ΟΤΣ, η οποία αποτελεί πλέον την είσοδο για τα διάφορα μοντέλα πρόβλεψης που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης.

Επιπλέον, κατά το στάδιο της μετεπεξεργασίας, η χρονοσειρά που προκύπτει ως έξοδος από το εκάστοτε μοντέλο πρόβλεψης που εφαρμόζεται κάθε φορά, υπόκειται σε επανεποχικοποίηση, χρησιμοποιώντας τους 24 εποχικούς δείκτες που υπολογίστηκαν κατά την προεπεξεργασία.

Αξίζει να αναφερθεί πως η επανεποχικοποίηση βασίζεται στον πολλαπλασιασμό των δεικτών εποχικότητας με τις αντίστοιχα δεδομένα, επομένως πρόκειται για πολλαπλασιαστικό μοντέλο εποχικότητας.

Σε αυτή την κατηγορία, παίζει μεγάλο ρόλο και η διόρθωση του μεγέθους N, το οποίο, όπως αναφέρθηκε και στην αρχή της προηγούμενης ενότητας, εκφράζει το μέγεθος των in-sample δεδομένων. Η διόρθωση του N θεωρείται μείζονος σημασίας προκειμένου οι εποχικοί δείκτες να είναι συγχρονισμένοι με τα δεδομένα και να μην παρουσιάζεται το φαινόμενο της ολίσθησης, δηλαδή για παράδειγμα ο εποχικός δείκτης που αναφέρεται στις 20:00 η ώρα να πολλαπλασιάζεται με την τιμή που αντιστοιχεί στις 15:00 η ώρα, με αποτέλεσμα η επανεποχικοποιημένη χρονοσειρά που παράγεται να είναι λανθασμένη.

Από τη στιγμή που εξετάζεται η εποχικότητα, δεν μπορεί να παραλειφθεί η πιθανή επίδραση που μπορεί να έχει σε αυτήν η αλλαγή των εποχών μέσα στο ημερολογιακό έτος (calendar effect). Για το λόγο αυτό, είναι απαραίτητη η επιλογή κατάλληλου αριθμού διαθέσιμων παρατηρήσεων για τον υπολογισμό των εποχικών δεικτών, έτσι ώστε να μπορούν αυτοί να συμπεριλάβουν την πληροφορία που οφείλεται στις ημερολογιακές αλλαγές.

Οι 24 εποχικοί δείκτες υπολογίζονται για τις περιπτώσεις όπου λαμβάνονται υπόψη 1,2,3,6,12 μήνες διαθέσιμων παρατηρήσεων, αλλά και στην περίπτωση που λαμβάνεται το σύνολο των παρατηρήσεων. Η διαφορά που προκύπτει από την κάθε περίπτωση, έγκειται στο γεγονός ότι με τη χρησιμοποίηση περισσότερων ή λιγότερων μηνών για την εξαγωγή των δεικτών εποχικότητας, ενσωματώνεται κάθε φορά διαφορετικός όγκος πληροφορίας από τις προηγούμενες παρατηρήσεις.

Στον πίνακα που ακολουθεί, παρουσιάζεται ο μέσος στατιστικός δείκτης του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE), για τις 6 μεθόδους που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης, όπως αυτός προκύπτει μετά τον υπολογισμό των 24 εποχικών δεικτών (SI=24):

		ΜΕΘΟΔΟΣ	Naive	LRL	Theta	SES	Holt	Damped
SI=24	sMAPE (%)	*1 Μήνας*	26,1606	38,2304	26,0253	26,0219	48,1635	26,0414
		2 Μήνες	26,2117	38,1508	26,0724	26,069	46,6392	26,0842
		3 Μήνες	26,3579	38,154	26,228	26,2193	46,9235	26,223
		6 Μήνες	26,5694	38,1082	26,4075	26,403	47,9758	26,4121
		12 Μήνες	26,433	37,9152	26,2541	26,2493	47,9319	26,2602
		Όλα	25,6652	37,3632	25,351	25,3464	49,3313	25,3696

Πίνακας 7.2.1. – Μέσοι Δείκτες sMAPE με μονή εποχικότητα σε επίπεδο ώρας (SI=24)

Με μία πρώτη ματιά στον Πίνακα 7.2.1, είναι εμφανές ότι οι μέσες τιμές του δείκτη sMAPE έχουν μειωθεί αισθητά στο σύνολο τους, σε σχέση με αυτές που παρουσιάστηκαν στον Πίνακα 7.1.1 και αφορούσαν την παραγωγή προβλέψεων δίχως απαλοιφή της εποχικότητας στη χρονοσειρά της ΟΤΣ. Συγκεκριμένα, για τις μεθόδους Naive, LRL, Theta, SES και Damped, το μεγαλύτερο σε τιμή σφάλμα που υπολογίστηκε για κάθε μία, είναι σημαντικά μικρότερο ακόμη και από το καλύτερο σφάλμα που προέκυψε χωρίς την επεξεργασία της εποχικότητας στην προηγούμενη ενότητα.

Επομένως, υπάρχει σαφής ένδειξη ότι σε αυτή την περίπτωση, απαλείφοντας δηλαδή την εποχικότητα σε επίπεδο ώρας μέσα στην ημέρα και υπολογίζοντας 24 εποχικούς δείκτες, τα αποτελέσματα παρουσιάζουν ιδιαίτερη βελτίωση και ο δείκτης του sMAPE μειώνεται ακόμη και κατά 6 ποσοστιαίες μονάδες, για κάθε μοντέλο πρόβλεψης, εκτός από το μοντέλο Holt.

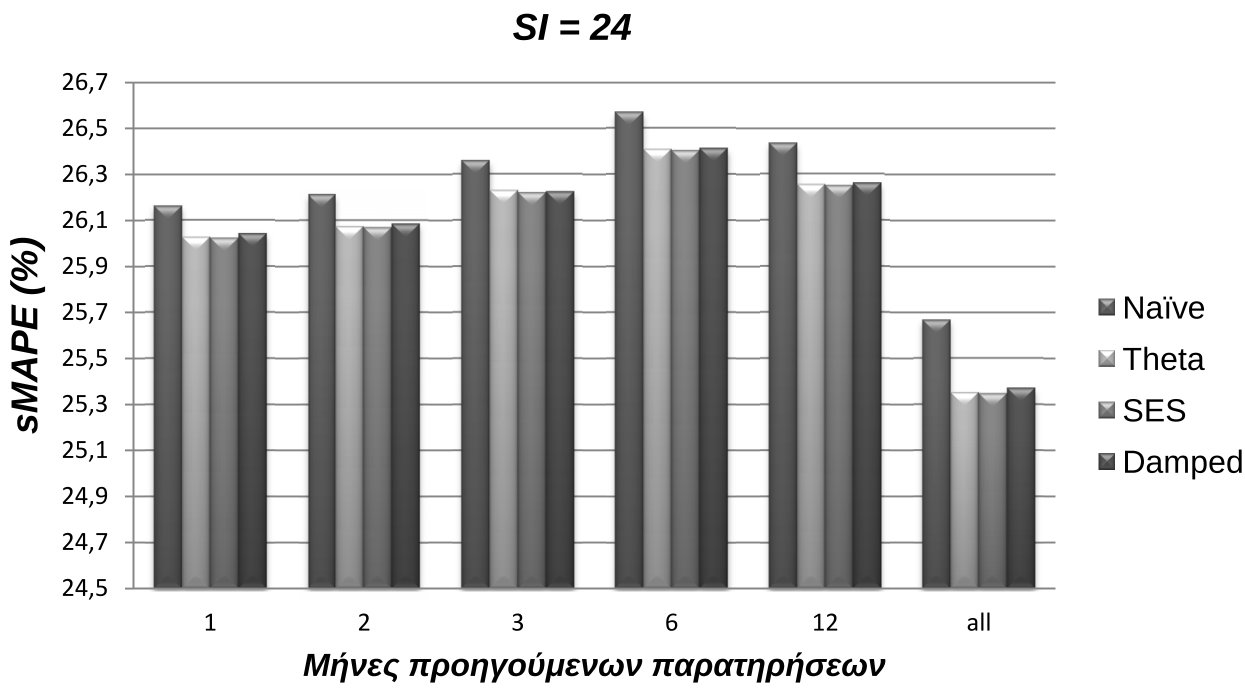
Το μοντέλο Holt αποτελεί τη μόνη παραφωνία σε αυτή τη μαζική βελτίωση του στατιστικού δείκτη sMAPE, γεγονός το οποίο δείχνει ότι πιθανότατα δε μπορεί να προσαρμοστεί στη χρονοσειρά της ΟΤΣ, παράγοντας ιδιαιτέρως αναξιόπιστες προβλέψεις.

Μάλιστα, ο δείκτης του sMAPE, όχι μόνο δεν βελτιώθηκε, αλλά αυξήθηκε, με τη μεγαλύτερη αύξηση να παρουσιάζεται στην περίπτωση που οι δείκτες εποχικότητας υπολογίστηκαν λαμβάνοντας υπόψη όλα τα διαθέσιμα δεδομένα, φτάνοντας το 49,3313%, δηλαδή πρόκειται για μία ποσοστιαία μεταβολή της τάξης του 4,7% σε σχέση με το αποτέλεσμα του Πίνακα 7.1.1. Αξίζει να σημειωθεί, πως η καλύτερη επίδοση που καταγράφεται για το συγκεκριμένο μοντέλο είναι το 46,6392%, ποσοστό μεγαλύτερο κατά 2 ποσοστιαίες μονάδες από εκείνο του Πίνακα 7.1.1

Σε αντίθεση με το μοντέλο Holt, τα υπόλοιπα πέντε μοντέλα παρουσιάζουν σημαντική βελτίωση, και μάλιστα καταγράφουν την καλύτερη τιμή για το δείκτη sMAPE στην περίπτωση που το μοντέλο Holt παρουσιάζει το μεγαλύτερο σφάλμα, δηλαδή στον υπολογισμό των εποχικών δεικτών με βάση όλες τις διαθέσιμες παρατηρήσεις.

Αξιοσημείωτο είναι το γεγονός, ότι η δεύτερη καλύτερη επίδοση για τις 5 μεθόδους (Naive, LRL, Theta, SES, Damped) καταγράφεται στο σημείο όπου χρησιμοποιείται μόνο 1 μήνας διαθέσιμων παρατηρήσεων για την εξαγωγή των δεικτών εποχικότητας. Στη συνέχεια, όσο αυξάνεται ο όγκος των διαθέσιμων παρατηρήσεων, συγκεκριμένα μέχρι και τους 6 μήνες (με εξαίρεση μόνο το μοντέλο LRL το οποίο δείχνει να μειώνει το sMAPE στους 6 μήνες), ο δείκτης του sMAPE αυξάνεται σταδιακά, για να μειωθεί και πάλι στην περίπτωση που χρησιμοποιούνται 12 μήνες παρατηρήσεων.

Στο συγκεντρωτικό διάγραμμα που ακολουθεί, παρουσιάζεται η πορεία που διαγράφει ο μέσος δείκτης του sMAPE για τις διάφορες τιμές των προηγούμενων παρατηρήσεων που λαμβάνονται υπόψη, για τις 4 βέλτιστες μεθόδους της παρούσας κατηγορίας, δηλαδή τις μεθόδους Naive, Theta, SES και Damped:



Σχ. 7.2.1 – Μέσοι Δείκτες sMAPE για τα μοντέλα Naive, Theta, SES, Damped (SI=24)

Παρατηρώντας το παραπάνω διάγραμμα, γίνεται αντιληπτή η αισθητή βελτίωση του μέσου δείκτη sMAPE, κατά την εφαρμογή των μοντέλων SES, Theta, Damped, σε σχέση με το μοντέλο αναφοράς Naive. Το γεγονός αυτό, υποδεικνύει πως η αύξηση της υπολογιστικής πολυπλοκότητας των τριών μοντέλων σε σχέση με το μοντέλο αναφοράς, αντισταθμίζεται μέσω της ικανοποιητικής βελτίωσης που επέρχεται στην ακρίβεια των παραγόμενων προβλέψεων

Όσον αφορά τις καλύτερες επιδόσεις, αυτές σίγουρα δεν αφορούν το μοντέλο LRL, το οποίο αν και παρουσιάζει βελτίωση στο δείκτη σφάλματος, παραμένει αρκετά αναξιόπιστο για την παραγωγή προβλέψεων, έχοντας ως μικρότερο δείκτη sMAPE ένα ποσοστό της τάξης του 37,3632%.

Επομένως, όπως και στην προηγούμενη περίπτωση όπου δεν λήφθηκε υπόψη η παρουσία εποχικότητας, έτσι και εδώ τα μοντέλα Naive, Theta, SES και Damped, παρουσιάζουν τα καλύτερα αποτελέσματα, σημαντικά βελτιωμένα σε σύγκριση με αυτών της περίπτωσης i (no SI).

Τα 4 αυτά μοντέλα, παρουσιάζουν παρόμοια αποτελέσματα, και για τις 6 διαφορετικές περιπτώσεις όγκου των διαθέσιμων δεδομένων για την εξαγωγή των εποχικών δεικτών, με τη μόνη διαφορά ότι πλέον το μοντέλο SES καταλαμβάνει την πρώτη θέση, με τη Theta να απέχει ελάχιστα και στις 6 περιπτώσεις.

Αυτό οφείλεται στο ότι το μοντέλο SES δε λαμβάνει υπόψη την ύπαρξη τάσης, που εκ των πραγμάτων δεν έχει έντονη παρουσία στη χρονοσειρά της ΟΤΣ αλλά και απαλείφεται σε μεγάλο βαθμό μέσω της διαδικασίας αποεποχικοποίησης που εφαρμόζεται, σε αντίθεση με το μοντέλο Theta, όπου η γραμμή Theta(0) ενσωματώνει όποια πιθανή πληροφορία τάσης μέσα στις παραγόμενες προβλέψεις. Φυσικά οι διαφορές είναι ελάχιστες, αλλά παραμένουν εμφανείς και χρήζουν εξήγησης.

Ένα πολύ σημαντικό γεγονός που προκύπτει από τα αποτελέσματα αυτής της κατηγορίας, είναι ότι η μέθοδος Naive, ενώ παρήγαγε τα καλύτερα αποτελέσματα στην περίπτωση i, εδώ βρίσκεται στην 4η θέση, πίσω ακόμη και από τη μέθοδο Damped.

Επομένως, εξάγεται το συμπέρασμα πως η επιτυχία της μεθόδου Naive προηγουμένως, οφείλεται στη μη επεξεργασία της εποχικότητας στη χρονοσειρά της ΟΤΣ, καθώς οι υπόλοιπες 3 μέθοδοι οι οποίες παρήγαγαν εξίσου καλά αποτελέσματα συμπεριλαμβανομένης της εποχικότητας, με την εισαγωγή των εποχικών δεικτών βελτιώθηκαν σε μεγαλύτερο βαθμό από ότι η Naive.

7.3 Αποτελέσματα με μονή εποχικότητα (SI=168)

Ένας δεύτερος τρόπος απομόνωσης της εποχικότητας στη χρονοσειρά της ΟΤΣ, είναι η υπόθεση πως υπάρχει μονή εποχικότητα σε επίπεδο ώρας αλλά μέσα σε μία εβδομάδα, και όχι μέσα στην ημέρα όπως υποτέθηκε στην προηγούμενη ενότητα. Αυτό πρακτικά σημαίνει ότι πλέον είναι απαραίτητοι 168 εποχικοί δείκτες και όχι μόνο 24, καθώς εξετάζεται η εποχικότητα σε επίπεδο εβδομάδας, όπου περιλαμβάνονται επτά 24ωρα.

Στόχος της συγκεκριμένης προσπάθειας είναι να επιτευχθούν ακόμη χαμηλότεροι δείκτες σφάλματος από αυτούς της περίπτωσης ii, εξετάζοντας παράλληλα και την αντίδραση των διάφορων μοντέλων πρόβλεψης που χρησιμοποιούνται για την παραγωγή αποτελεσμάτων μέσω της διαδικασίας της προσομοίωσης.

Με αυτή την υπόθεση, γίνεται στην ουσία μία πρώτη προσπάθεια να αντιμετωπιστεί η δεύτερη συνιστώσα της πολλαπλής εποχικότητας, εκείνη που αναφέρεται σε εβδομαδιαίο επίπεδο, που όπως πολλάκις έχει αναφερθεί στα πλαίσια αυτής της Διπλωματικής, χαρακτηρίζει τη χρονοσειρά της ΟΤΣ.

Εξάγοντας 168 εποχικούς δείκτες, είναι δυνατό να αντιμετωπιστεί το πρότυπο που υπάρχει μέσα στην εβδομάδα, καθώς είναι σαφές ότι παρατηρείται σημαντική διαφορά στο πρότυπο που ακολουθείται τις καθημερινές από αυτό του Σαββατοκύριακου. Επομένως, διαφορετικές μέρες μπορούν να εμφανίζουν διαφορετικές μορφές καμπύλης ΟΤΣ, κάτι το οποίο μπορεί να αποτυπωθεί με τη συγκεκριμένη μεθοδολογία

Στον πίνακα που ακολουθεί, παρουσιάζονται οι μέσες τιμές του στατιστικού δείκτη του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE), όπως αυτές υπολογίστηκαν για κάθε μία από τις 6 μεθόδους πρόβλεψης που χρησιμοποιούνται κατά τη διαδικασία του rolling forecasting:

		ΜΕΘΟΔΟΣ	Naive	LRL	Theta	SES	Holt	Damped
SI=168	**sMAPE**	***Όλα***	25,7923	37,4809	25,2775	25,2736	46,4451	25,2803

Πίνακας 7.3.1. – Μέσοι Δείκτες sMAPE με μονή εποχικότητα σε επίπεδο ώρας (SI=168)

Αξίζει να αναφερθεί το γεγονός ότι για τον υπολογισμό των 168 εποχικών δεικτών, λαμβάνονται υπόψη όλες οι διαθέσιμες παρατηρήσεις, καθώς με αυτό τον τρόπο αποτρέπεται η πιθανότητα εμφάνισης ολίσθησης μεταξύ των δεικτών εποχικότητας και των παρατηρήσεων.

Με μία πρώτη ματιά, είμαι εμφανής η αντίδραση του μοντέλου Naive απέναντι στην εφαρμογή 168 εποχικών δεικτών. Σε σύγκριση με την βέλτιστη επίδοση για την περίπτωση ii, ο δείκτης sMAPE για το μοντέλο Naive είναι μεγαλύτερος εδώ κατά 0,13 ποσοστιαίες μονάδες.

Όπως παρατηρήθηκε και στην προηγούμενη ενότητα, αυτό οφείλεται στο γεγονός ότι το μοντέλο Naive, ενώ βελτίωσε σημαντικά τις επιδόσεις του από τη στιγμή που αντιμετωπίστηκε η εποχικότητα, σε σύγκριση πάντα με την περίπτωση i, δεν παρουσίασε τόσο μεγάλη βελτίωση όσο οι υπόλοιπες πέντε μέθοδοι, και ειδικά οι τρείς καλύτερες (Theta, SES, Damped).

Το μοντέλο LRL, παρουσιάζει παρόμοια συμπεριφορά με αυτό της Naive, δηλαδή εμφανίζεται ελαφρώς χειρότερο από τη βέλτιστη επίδοσή του για την περίπτωση ii, με μία διαφορά στο δείκτη sMAPE της τάξης των 0,1177 ποσοστιαίων μονάδων. Αυτό πιθανότατα οφείλεται στο ότι η εισαγωγή 168 εποχικών δεικτών εξασθένισε ακόμη περισσότερο την παρουσία τάσης στη χρονοσειρά της ΟΤΣ, με αποτέλεσμα το μοντέλο LRL να ανταποκριθεί σε αυτό το γεγονός οδηγούμενο σε μεγαλύτερο σφάλμα.

Το μοντέλο Holt, παραμένει και σε αυτή την περίπτωση η χειρότερη επιλογή, παράγοντας αποτελέσματα ελαφρώς βελτιωμένα σε σχέση με την προηγούμενη περίπτωση (SI=24), τα οποία όμως παραμένουν χειρότερα από αυτά της περίπτωσης i (no SI). Σε αυτό το σημείο, θεωρείται ασφαλές να ειπωθεί πως τα μέχρι στιγμής αποτελέσματα δείχνουν ότι η συγκεκριμένη μέθοδος παραμένει σταθερά στην τελευταία θέση από τη σκοπιά των επιδόσεων, και με μεγάλη διαφορά ακόμη και από το μοντέλο LRL, γεγονός που υποδεικνύει πως δεν μπορεί να ανταποκριθεί στις ιδιομορφίες της χρονοσειράς της ΟΤΣ, ανεξαρτήτως οποιασδήποτε διαδικασίας προεπεξεργασίας υποβάλλεται κάθε φορά η υπό πρόβλεψη χρονοσειρά.

Όσον αφορά τα τρία εναπομείναντα μοντέλα, δηλαδή τα μοντέλα Theta, SES, Damped, αυτά παρουσιάζουν όμοια συμπεριφορά, κάτι το οποίο ήταν εμφανές και από τις δύο προηγούμενες περιπτώσεις που εξετάσθηκαν (SI=24, no SI). Τα τρία μοντέλα εμφανίζουν πολύ μικρή βελτίωση αναφορικά με το δείκτη sMAPE, σε σύγκριση με την περίπτωση ii, διατηρώντας την ίδια ακριβώς σειρά από σκοπιά απόδοσης. Καλύτερη μέθοδος εμφανίζεται να είναι η SES, παρουσιάζοντας μια ποσοστιαία μεταβολή της τάξης του 0,0728% σε σύγκριση με την περίπτωση Ii, και 18,71% σε σύγκριση με την περίπτωση i.

Αντίστοιχη συμπεριφορά παρατηρείται και από το μοντέλο Theta, το οποίο παρουσιάζει μία ποσοστιαία μεταβολή της τάξης του 0,0735% σε σύγκριση με την περίπτωση ii, και 18,69% σε σύγκριση με την περίπτωση i.

Τέλος, το τρίτο καλύτερο μοντέλο (Damped), χαρακτηρίζεται από ποσοστιαία μεταβολή 0,0893% σε σύγκριση με την περίπτωση ii, και 19,36% σε σχέση με την περίπτωση I, μεταβολές μεγαλύτερες σε σύγκριση με τα άλλα δύο μοντέλα, γεγονός που υποδεικνύει καλύτερη ανταπόκριση του μοντέλου Damped στην εισαγωγή 168 εποχικών δεικτών έναντι 24.

7.4 Αποτελέσματα με διπλή εποχικότητα (double SI)

Πρόκειται για την $4^{η}$ και τελευταία κατηγορία αποτελεσμάτων. Το κυρίαρχο στοιχείο σε αυτή την περίπτωση είναι η αναγνώριση της παρουσίας διπλής εποχικότητας, και η παρουσίαση κατάλληλου τρόπου εξάλειψής της από τη χρονοσειρά της ΟΤΣ. Όπως και στις προηγούμενες κατηγορίες, έτσι και εδώ, ζητούμενο είναι η όσο το δυνατόν μεγαλύτερη βελτίωση της απόδοσης των μοντέλων πρόβλεψης κατά τη διαδικασία της προσομοίωσης, με κριτήριο πάντα το στατιστικό δείκτη του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE).

Αρχικά, η υπό εξέταση χρονοσειρά, υπόκειται στην γνωστή πλέον διαδικασία προεπεξεργασίας (κύλιση παρατηρήσεων, απαλοιφή μηδενικών τιμών). Στη συνέχεια ακολουθεί ο υπολογισμός 24 εποχικών δεικτών σε επίπεδο ώρας που αφορούν το ημερήσιο πρότυπο της εποχικότητας, το οποίο αποτελεί την πρώτη συνιστώσα της διπλής εποχικότητας στη χρονοσειρά της ΟΤΣ. Ουσιαστικά πρόκειται για την περίπτωση ii (SI=24), κατά την οποία αντιμετωπίστηκε η ημερήσια εποχικότητα με τον ίδιο ακριβώς τρόπο.

Το σημείο που διαφοροποιείται η συγκεκριμένη κατηγορία από την κατηγορία ii, είναι εκείνο όπου μετά τον υπολογισμό των 24 εποχικών δεικτών που αφορούν το ημερήσιο πρότυπο, το οποίο θεωρείται ότι επαναλαμβάνεται αυτούσιο για κάθε ημέρα, ακολουθεί ο υπολογισμός 168 εποχικών δεικτών σε επίπεδο ώρας, οι οποίοι αφορούν το εβδομαδιαίο πρότυπο, το οποίο αποτελεί τη δεύτερη συνιστώσα της διπλής εποχικότητας.

Επομένως, η βασική υπόθεση αυτής της κατηγορίας αποτελεσμάτων είναι ότι αναγνωρίζεται η ύπαρξη διπλής εποχικότητας στη χρονοσειρά της ΟΤΣ, διαχωρίζεται σε δύο συνιστώσες, δηλαδή ημερήσιο και εβδομαδιαίο πρότυπο, και υπολογίζονται για την κάθε συνιστώσα οι αντίστοιχοι εποχικοί δείκτες, 24 για το ημερήσιο πρότυπο και 168 για το εβδομαδιαίο.

Είναι πολύ σημαντικό η αποεποχικοποίηση να γίνει με τη σειρά που αναφέρθηκε κατά το στάδιο της προεπεξεργασίας, δηλαδή πρώτα να υπολογιστούν οι 24 εποχικοί δείκτες της πρώτης εποχικής συνιστώσας, και στη συνέχεια, σύμφωνα πλέον με την ενδιάμεση αποεποχικοποιημένη χρονοσειρά που προέκυψε, να υπολογιστούν οι 168 εποχικοί δείκτες της δεύτερη συνιστώσας.

Ο λόγος για τον οποίο πρέπει να τηρείται αυτή η σειρά είναι ότι, στην περίπτωση που εξάγονταν πρώτα οι 168 εποχικοί δείκτες που αφορούν το εβδομαδιαίο πρότυπο, θα υπερκάλυπταν την πληροφορία που υπάρχει μέσα στο ημερήσιο πρότυπο. Με απλά λόγια, το εβδομαδιαίο πρότυπο θα απορροφούσε την πληροφορία του ημερήσιου προτύπου, όπως συνέβη και στην περίπτωση iii (SI=168), καθώς οι εποχικοί δείκτες είναι σε ωριαίο επίπεδο.

Επομένως, για να αποφευχθεί η απώλεια της πληροφορίας του ημερήσιου προτύπου, προηγείται ο υπολογισμός των 24 εποχικών δεικτών (SI=24), και αφού έχει εξασφαλιστεί η απαλοιφή της πρώτης εποχικής συνιστώσας, ακολουθεί ο προσδιορισμός των 168 εποχικών δεικτών, απαλείφοντας έτσι και το εβδομαδιαίο πρότυπο εποχικότητας. Αξίζει να αναφερθεί, πως στην περίπτωση που το εβδομαδιαίο πρότυπο αντιμετωπίζονταν με 7 εποχικούς δείκτες, έναν για κάθε ημέρα της εβδομάδας, τότε το ενδεδειγμένο θα ήταν να γίνει πρώτη η απαλοιφή του εβδομαδιαίου προτύπου και ύστερα του ημερήσιου.

Βασική προϋπόθεση σε αυτή την περίπτωση θα ήταν οι παρατηρήσεις να είναι σε επίπεδο ημέρας και όχι ώρας για την εξαγωγή των 7 πρώτων δεικτών, και στη συνέχεια να ανάγονταν και πάλι σε επίπεδο ώρας για την εξαγωγή των 24 δεικτών ημερήσιας εποχικότητας. Η συγκεκριμένη διαδικασία είναι πιθανό να εκφύλιζε τηη αρχική φύση των παρατηρήσεων, λόγος για τον οποίο αυτό το εγχείρημα παραμένει έξω από τα όρια της παρούσας Διπλωματικής.

Σχετικά με το στάδιο της μετεπεξεργασίας, δεν παίζει ρόλο η σειρά με την οποία γίνεται η επανεποχικοποίηση για κάθε μία από τις δύο συνιστώσες, καθώς πρόκειται για πολλαπλασιαστική εποχικότητα, και επομένως η σειρά με την οποία θα γίνει ο πολλαπλασιασμός της κάθε ομάδας εποχικών δεικτών με τις παρατηρήσεις δεν προκαλεί κάποιου είδους μεταβολή στο τελικό αποτέλεσμα. Είναι σημαντικό να αναφερθεί πως πρόκειται για έναν απλοϊκό τρόπο αντιμετώπισης της διπλής εποχικότητας. Υπάρχουν πιο εξειδικευμένες μέθοδοι, αυξημένης πολυπλοκότητας, οι οποίες δεν εντάσσονται στα πλαίσια της παρούσας Διπλωματικής.

Στον πίνακα που ακολουθεί, παρουσιάζονται οι μέσες τιμές του στατιστικού δείκτη sMAPE, για τα 6 μοντέλα που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης, για τους διάφορους συνδυασμούς των δεικτών ημερήσιας εποχικότητας με τους δείκτες εβδομαδιαίας εποχικότητας (1, 2, 3, 6, 12 και όλοι οι μήνες):

		ΜΕΘΟΔΟΣ	**Naive**	**LRL**	**Theta**	**SES**	**Holt**	**Damped**
Double SI	**sMAPE**	*1 Μήνας*	24,3993	36,6691	23,8551	23,8509	45,6236	23,8582
		2 Μήνες	24,3999	36,6692	23,8564	23,8521	45,6236	23,8591
		3 Μήνες	24,3999	36,6692	23,8563	23,8521	45,6244	23,8591
		6 Μήνες	24,4008	36,6696	23,8572	23,8529	45,6293	23,8598
		12 Μήνες	24,4026	36,6695	23,8589	23,8547	45,6357	23,8615
		Όλα	24,403	36,6688	23,8602	23,8559	45,6383	23,862

Πίνακας 7.4.1. - Μέσοι Δείκτες sMAPE με διπλή εποχικότητα σε επίπεδο ώρας (SI=24, SI2=168)

Όπως φαίνεται και στον Πίνακα 7.4.1., για την εξαγωγή των 24 δεικτών ημερήσιας εποχικότητας, χρησιμοποιήθηκε όπως και στην περίπτωση ii, διαφορετικός όγκος διαθέσιμων παρατηρήσεων, με στόχο την εύρεση του ιδανικότερου συνδυασμού μεταξύ της πρώτης και της δεύτερης εποχικής συνιστώσας της διπλής εποχικότητας. Επιπλέον, για τον υπολογισμό των 168 δεικτών εβδομαδιαίας εποχικότητας, λήφθηκαν υπόψη όλα τα διαθέσιμα δεδομένα, ακριβώς όπως και στην περίπτωση iii (SI=168).

Το εγχείρημα της αντιμετώπισης της διπλής εποχικότητας με τη μεθοδολογία που αναπτύχθηκε παραπάνω, δείχνει σε πρώτη φάση να βελτιώνει αισθητά το δείκτη του sMAPE, για όλα τα μοντέλα πρόβλεψης που χρησιμοποιήθηκαν στην προσομοίωση του rolling forecasting. Η μόνη παραφωνία σε αυτή τη μαζική βελτίωση είναι και πάλι το μοντέλο Holt, που ενώ βελτιώθηκαν οι επιδόσεις του σε σχέση με τις περιπτώσεις ii (SI=24) και iii (SI=168), παράγει αποτελέσματα με αρκετά υψηλούς δείκτες σφάλματος.

Αξίζει να σημειωθεί, πως ακόμη και σε αυτή την περίπτωση όπου αντιμετωπίστηκε η διπλή εποχικότητα, το καλύτερο αποτέλεσμα για τη μέθοδο Holt, το οποίο είναι ένα ποσοστό σφάλματος της τάξης του 45,6236%, παραμένει χειρότερο από το αποτέλεσμα που παρουσιάστηκε στον Πίνακα 7.1.1 και αφορούσε την μη απαλοιφή της εποχικότητας από τη χρονοσειρά της ΟΤΣ, γεγονός που υποδεικνύει πως το μοντέλο Holt δεν βελτιώνεται σε καμία από τις τρείς περιπτώσεις αντιμετώπισης της εποχικότητας, διπλής ή μη, παρουσιάζοντας τα καλύτερα αποτελέσματα εκεί όπου δεν υπήρξε καμία απαλοιφή της εποχικότητας, δηλαδή στην περίπτωση I (no SI).

Αναφορικά με το μοντέλο LRL, αυτό παρουσιάζει τις καλύτερες επιδόσεις του, σε σύγκριση και με τις 3 προηγούμενες κατηγορίες αποτελεσμάτων. Όμως, ο δείκτης σφάλματος sMAPE διατηρείται σε αρκετά υψηλά επίπεδα ακόμη και σε αυτή την περίπτωση.

Το γεγονός αυτό δεν αναιρεί την μεγάλη πρόοδο που επέδειξε το μοντέλο LRL καθώς υπήρξε μετάβαση από την 1η στην 4η κατηγορία αποτελεσμάτων, φτάνοντας τελικά σε μία ποσοστιαία μεταβολή της τάξης του 16,17%, συγκρίνοντας την βέλτιστη τιμή του Πίνακα 7.4.1. με αυτή του Πίνακα 7.1.1.

Επιπλέον, παρατηρείται πως η μέθοδος LRL, παρουσιάζει το βέλτιστο αποτέλεσμά της όταν για τον υπολογισμό των 24 δεικτών ημερήσιας εποχικότητας λαμβάνονται υπόψη όλες οι προηγούμενες παρατηρήσεις, ακριβώς όπως συνέβη και στην περίπτωση ii (SI=24), όπου οι καλύτερες τιμές του δείκτη sMAPE παρατηρούνταν στην πλειοψηφία τους στο ίδιο σημείο.

Αντίθετα, οι μέθοδοι Naive, Theta, SES, Damped ανταποκρίνονται με διαφορετικό τρόπο στην απαλοιφή της διπλής εποχικότητας, παρουσιάζοντας τις βέλτιστες τιμές τους για το συνδυασμό όπου λαμβάνεται υπόψη μόνο 1 μήνας παρελθόντων δεδομένων για την εξαγωγή των 24 ημερήσιων εποχικών δεικτών.

Σχετικά με το μοντέλο αναφοράς Naive, παρουσιάζεται σημαντική βελτίωση στο δείκτη του sMAPE, με τη βέλτιστη τιμή να είναι ακόμη και 1,393 ποσοστιαίες μονάδες χαμηλότερη από εκείνη της περίπτωσης iii, και να βελτιώνεται κατά 21,5% σε σχέση με το αποτέλεσμα που παρουσιάζεται στον Πίνακα 7.1.1.

Επομένως, το μοντέλο Naive, ακολουθεί μια πορεία ανάλογη με το μοντέλο LRL, εντούτοις δείχνει να βελτιώνει σε μεγαλύτερο βαθμό τις επιδόσεις του και να ανταποκρίνεται καλύτερα στις διαφορετικές προσεγγίσεις που πραγματοποιήθηκαν, μεταβαίνοντας από την 1η στην 4η κατηγορία επεξεργασίας της εποχικότητας στη χρονοσειρά της ΟΤΣ.

Αναφορικά με τις υπόλοιπες μεθόδους πρόβλεψης, παρατηρείται πως και σε αυτήν την περίπτωση καταλαμβάνουν τις τρείς πρώτες θέσεις από σκοπιά επιδόσεων, παρουσιάζοντας τους χαμηλότερους δείκτες sMAPE. Η μέθοδος SES, διατηρεί τα πρωτεία, πετυχαίνοντας την καλύτερη επίδοση της στο συνδυασμό όπου λαμβάνεται υπόψη 1 μήνας προηγούμενων παρατηρήσεων για την εξαγωγή των πρώτων 24 δεικτών εποχικότητας, όπως συμβαίνει και με τα μοντέλα Theta και Damped.

Η τιμή αυτή είναι ένα σφάλμα της τάξης του 23,8509%, χαμηλότερο κατά 1,4227 ποσοστιαίες μονάδες από το καλύτερο αποτέλεσμα της περίπτωσης iii (SI=168), φτάνοντας ακόμη και σε μία ποσοστιαία μεταβολή της τάξης του 23,29% σε σχέση με την αρχική τιμή του sMAPE που παρουσιάστηκε στον Πίνακα 7.1.1.

Η ποσοστιαία αυτή μεταβολή, είναι μεγαλύτερη ακόμη και από εκείνη που παρατηρείται για το μοντέλο Naive, υποδεικνύοντας το μεγάλο αντίκτυπο που είχε η αντιμετώπιση της διπλής εποχικότητας πάνω στο συγκεκριμένο μοντέλο πρόβλεψης.

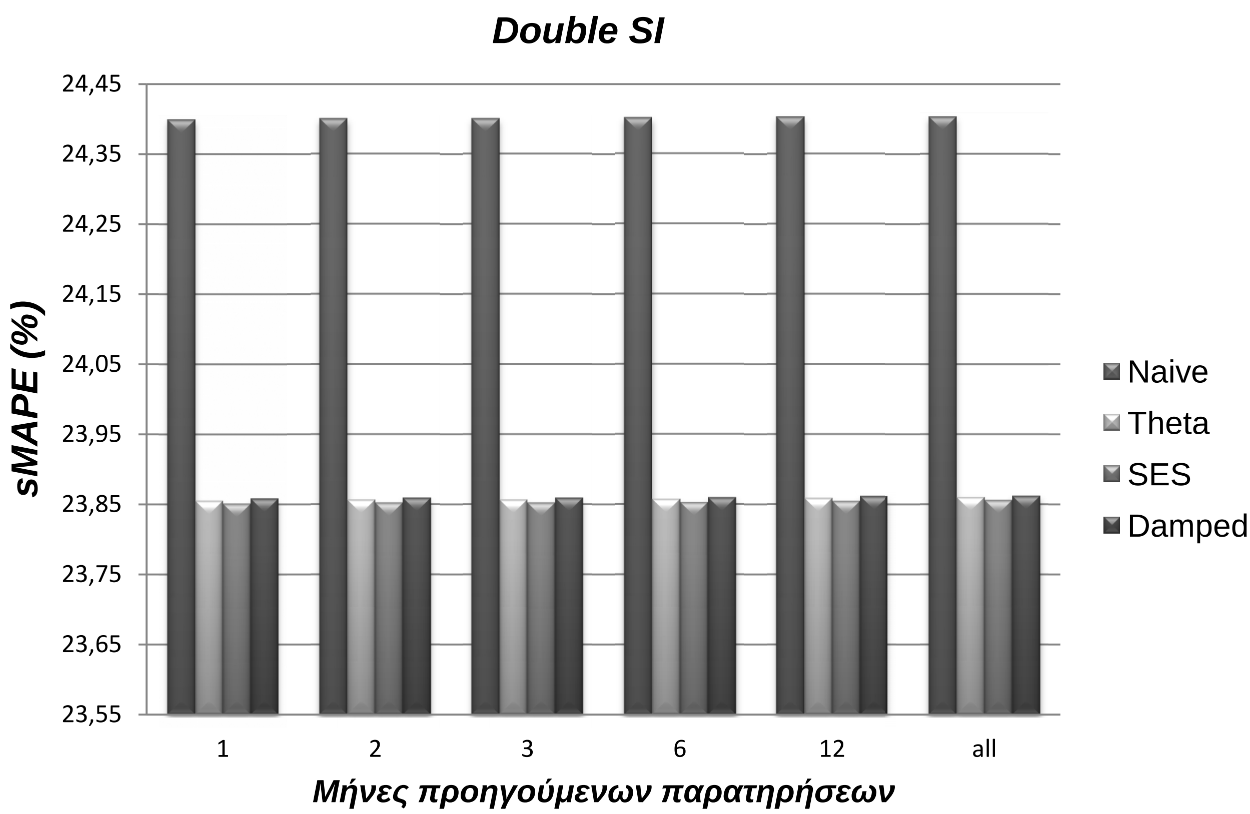
Η μέθοδος Theta, διαγράφει εξίσου σημαντική βελτίωση, ακολουθώντας ανάλογη πορεία με εκείνη της μεθόδου SES. Συγκεκριμένα, ο βέλτιστος δείκτης sMAPE σε αυτή την κατηγορία ισούται με 23,8551%, γεγονός που καθιστά το μοντέλο Theta το δεύτερο πιο αξιόπιστο ανάμεσα στα 6 μοντέλα που εξετάσθηκαν. Καταγράφεται μία διαφορά της τάξης των 1,4224 ποσοστιαίων μονάδων σε σχέση με την περίπτωση iii, νούμερο σχεδόν ταυτόσημο με εκείνο της μεθόδου SES, και μία συνολική ποσοστιαία μεταβολή της τάξης του 23,27%, ποσοστό ελαφρώς μικρότερο από εκείνο της μεθόδου SES.

Επομένως, είναι ασφαλές πλέον να ειπωθεί, πως η μέθοδος Theta, ανταποκρίνεται με πανομοιότυπο τρόπο σε σχέση με τη μέθοδο SES, και στις

4 διαφορετικές κατηγορίες αντιμετώπισης της εποχικότητας, διαγράφοντας σημαντική βελτίωση.

Τέλος, το μοντέλο Damped, κινείται πολύ κοντά στα 2 προαναφερθέντα μοντέλα, καταγράφοντας εξίσου μεγάλη βελτίωση αναφορικά με την ακρίβεια των προβλέψεων που παράγει. Η συνολική ποσοστιαία μεταβολή που παρουσιάζει το συγκεκριμένο μοντέλο συγκρίνοντας τη βέλτιστη τιμή του Πίνακα 7.4.1 με την τιμή του Πίνακα 7.1.1, φτάνει και το 23,90%. Το γεγονός αυτό, καθιστά το μοντέλο Damped, ως το πιο βελτιωμένο μοντέλο κατά τη διάρκεια της μετάβασης από την $1^{η}$ στην $4^{η}$ κατηγορία αντιμετώπισης της εποχικότητας στη χρονοσειρά της ΟΤΣ.

Στο συγκεντρωτικό διάγραμμα που ακολουθεί, παρουσιάζεται η πορεία που διαγράφει ο μέσος δείκτης του sMAPE, για τα 4 βέλτιστα μοντέλα, για τις διάφορες τιμές του όγκου των προηγούμενων παρατηρήσεων που λαμβάνονται υπόψη για τον υπολογισμό των 24 δεικτών ημερήσιας εποχικότητας, για την περίπτωση αντιμετώπισης της διπλής εποχικότητας με τη μεθοδολογία που αναπτύχθηκε στη παρούσα ενότητα:



Σχ. 7.4.1 – Μέσοι Δείκτες sMAPE για τα μοντέλα Naive, Theta, SES, Damped (Double SI)

Συγκρίνοντας το παραπάνω διάγραμμα με το διάγραμμα 7.2.1, γίνεται αντιληπτή η πολύ μεγάλη επίδραση που είχε η αντιμετώπιση της διπλής εποχικότητας στην ακρίβεια των παραγόμενων προβλέψεων. Επιπλέον, τα τρία βέλτιστα μοντέλα (SES, Theta, Damped), δείχνουν σαφέστατα μεγαλύτερη βελτίωση σε σύγκριση με το μοντέλο αναφοράς, ενώ στην περίπτωση αντιμετώπισης της μονής εποχικότητας η διαφορά του μοντέλου Naive με τα υπόλοιπα κυμαίνονταν σε πολύ χαμηλότερα επίπεδα.

Επομένως, όπως αναφέρθηκε και στην Ενότητα 7.2, η αύξηση της πολυπλοκότητας, αντισταθμίζεται με ακόμη μεγαλύτερη επιτυχία, και σε αυτή την περίπτωση, μέσω της σημαντικής βελτίωσης της μέσης τιμής του στατιστικού δείκτη sMAPE, για όλες τις περιπτώσεις που λαμβάνεται υπόψη διαφορετικός συνδυασμός για την εξαγωγή των δύο ομάδων δεικτών εποχικότητας (SI_1, SI_2).

Κεφάλαιο 8. Ανάλυση Βέλτιστων Αποτελεσμάτων Προσομοίωσης

Στο παρών κεφάλαιο, παρουσιάζονται και αναλύονται οι 3 βέλτιστες μέσες τιμές του στατιστικού δείκτη sMAPE, όπως αυτές προέκυψαν από τη διαδικασία της προσομοίωσης, σύμφωνα με τη μεθοδολογία που αναπτύχθηκε για τις 5 κατηγορίες αποτελεσμάτων, χρησιμοποιώντας τα 6 διαφορετικά μοντέλα πρόβλεψης που παρουσιάστηκαν στο Κεφάλαιο 7. Επιπλέον, γίνεται προσπάθεια επεξήγησης των παραμέτρων που καθόρισαν την εμφάνιση των συγκεκριμένων αποτελεσμάτων, καθώς και λεπτομερής γραφική ανάλυση της ακρίβειας των τριών καλύτερων μοντέλων πρόβλεψης.

8.1. Ποιοτική Ανάλυση Ακρίβειας Αποτελεσμάτων

Έχοντας πλέον υπολογίσει τους μέσους δείκτες σφάλματος sMAPE, για κάθε μία από τις 4 κατηγορίες αντιμετώπισης της εποχικότητας και για κάθε μοντέλο πρόβλεψης που χρησιμοποιείται κατά τη διαδικασία προσομοίωσης, είναι δυνατό να εντοπιστούν οι τρείς βέλτιστες επιδόσεις, ανάμεσα στο σύνολο των αποτελεσμάτων.

Όπως αναφέρθηκε και στην προηγούμενη Ενότητα, η περίπτωση αντιμετώπισης της διπλής εποχικότητας, προσδιορίζοντας 2 διαφορετικές ομάδες εποχικών δεικτών (24 δείκτες ημερήσιας εποχικότητας και 168 δείκτες εβδομαδιαίας εποχικότητας), βελτίωσε κατά πολύ τις επιδόσεις των 5 μοντέλων πρόβλεψης που χρησιμοποιήθηκαν κατά τη διαδικασία της προσομοίωσης, με εξαίρεση πάντα το μοντέλο Holt το οποίο παρουσίασε την καλύτερη επίδοσή του στην περίπτωση μη απαλοιφής της εποχικότητας.

Επομένως, οι 3 βέλτιστες τιμές του στατιστικού δείκτη sMAPE, παρατηρούνται στον Πίνακα 7.4.1, καθιστώντας με αυτό τον τρόπο τη συγκεκριμένη κατηγορία (double SI), τη βέλτιστη επιλογή μεταξύ των τεσσάρων κατηγορ

ιών αντιμετώπισης της εποχικότητας που εξετάστηκαν στο παρών Κεφάλαιο. Το γεγονός αυτό, επιβεβαιώνει την σημαντική επιρροή που έχει η συνιστώσα της διπλής εποχικότητας στη χρονοσειρά της ΟΤΣ, και παρά τον απλοϊκό τρόπο αντιμετώπισής της στα πλαίσια αυτής της Διπλωματικής, παράχθηκαν αποτελέσματα εμφανώς βελτιωμένα σε σχέση με κάθε άλλη κατηγορία αποτελεσμάτων.

Αξίζει να σημειωθεί πως και τα τρία καλύτερα αποτελέσματα, παρατηρούνται για το συνδυασμό εκείνο στον οποίο λαμβάνονται υπόψη προηγούμενα δεδομένα ενός μήνα για την εξαγωγή των 24 εποχικών δεικτών ημερήσιας εποχικότητας.

Στον ακόλουθο πίνακα, παρουσιάζονται τα 3 καλύτερα αποτελέσματα του δείκτη sMAPE, όπως αυτά υπολογίστηκαν σύμφωνα με τη μεθοδολογία που αναπτύχθηκε για την αντιμετώπιση της διπλής εποχικότητας, καθώς και τα μοντέλα πρόβλεψης τα οποία παρήγαγαν τα συγκεκριμένα βέλτιστα αποτελέσματα:

Double SI	
ΜΟΝΤΕΛΟ	***sMAPE (%)***
SES	**23,8509**
Theta	**23,8551**
Damped	**23,8582**

Πίνακας 8.1.1. - Βέλτιστοι Μέσοι Δείκτες sMAPE με διπλή εποχικότητα σε επίπεδο ώρας (SI=24, SI2=168)

Παρατηρώντας τον παραπάνω πίνακα, καθίσταται σαφές το γεγονός ότι δεν υπάρχει σημαντική διαφορά στην απόδοση των τριών μοντέλων SES, Theta και Damped. Ο στατιστικός δείκτης sMAPE κινείται στα ίδια επίπεδα και για τις τρείς προαναφερθείσες μεθόδους, με το μοντέλο SES να παρουσιάζει το βέλτιστο δείκτη, με μικρή διαφορά από τις άλλες δύο μεθόδους.

Αυτό που ενδιαφέρει στην παρούσα φάση, είναι ο λόγος για τον οποίο το μοντέλο SES, παρουσιάζεται ως το βέλτιστο, έστω και με πολύ μικρή διαφορά, για την παραγωγή προβλέψεων της ΟΤΣ. Συγκρίνοντας το συγκεκριμένο μοντέλο με το μοντέλο Theta, είναι ασφαλές να ειπωθεί πως η καλύτερη προσαρμογή του πρώτου έναντι του δεύτερου έγκειται στο γεγονός ότι το μοντέλο SES, δε λαμβάνει υπόψη την παρουσία τάσης στη χρονοσειρά της ΟΤΣ.

Αυτό λειτουργεί ευεργετικά στην παραγωγή των προβλέψεων, καθώς όπως έχει αναφερθεί πολλάκις μέχρι τώρα, η χρονοσειρά της ΟΤΣ δεν παρουσιάζει τάση (στάσιμη ως προς το μέσο), και επομένως η μη αντιμετώπισή της από το μοντέλο πρόβλεψης τείνει να βοηθά στην καλύτερη προσαρμογή του μοντέλου πάνω στα δεδομένα της υπό πρόβλεψης χρονοσειράς.

Αντίθετα, το μοντέλο Theta, και συγκεκριμένα η γραμμή Theta(0), ενσωματώνει στις παραγόμενες προβλέψεις την όποια πιθανή πληροφορία προέρχεται από την παρουσία μακροπρόθεσμης τάσης, αφού στην ουσία πρόκειται για εφαρμογή της ευθείας απλής γραμμικής παλινδρόμησης (LRL) πάνω στη χρονοσειρά της ΟΤΣ. Το εγχείρημα αυτό αποδεικνύεται πως δεν παράγει ικανοποιητικά αποτελέσματα, γεγονός το οποίο επιβεβαιώνεται παρατηρώντας τις ιδιαίτερα μέτριες επιδόσεις που παρουσιάζει το μοντέλο LRL (Πίνακες 7.1.1 - 7.4.1).

Επομένως, το εγχείρημα του μοντέλου Theta, να αποτυπώσει την βραχυπρόθεσμη πληροφορία μέσω της γραμμής Theta(2), η οποία δεν είναι άλλη από το μοντέλο SES, σε συνδυασμό με την παρουσία μακροπρόθεσμης τάσης μέσω της γραμμής Theta(0), αποδεικνύεται μη ικανοποιητικό, σε σύγκριση πάντα με το μοντέλο SES, καθώς η απουσία μακροπρόθεσμης τάσης στη χρονοσειρά της ΟΤΣ ακυρώνει σε μεγάλο βαθμό τη λειτουργία της γραμμής Theta(0), περιορίζοντας τη δυνατότητα του συγκεκριμένου μοντέλου να προσαρμοστεί όσο το δυνατόν καλύτερα στην υπό εξέταση χρονοσειρά.

Όσον αφορά τη σύγκριση μεταξύ των δύο μοντέλων εξομάλυνσης (SES, Damped), παρατηρείται και σε αυτή την περίπτωση πολύ μικρή διαφορά στο δείκτη του sMAPE. Από καθαρά θεωρητική σκοπιά, το μοντέλο Damped θεωρείται ότι ανταποκρίνεται καλύτερα σε σχέση με τα άλλα δύο μοντέλα εκθετικής εξομάλυνσης, αλλά μόνο σε περιπτώσεις όπου ορίζεται μεγάλος ορίζοντας πρόβλεψης. Αντίθετα, στην περίπτωση που εξετάζεται, ο ορίζοντας πρόβλεψης διατηρείται σε χαμηλά επίπεδα (24 ώρες), καθώς ενδιαφέρει περισσότερο η παραγωγή βραχυπρόθεσμων προβλέψεων στα πλαίσια της λογικής του *day ahead forecasting.*

Επομένως, με δεδομένο ότι το μοντέλο Damped ανταποκρίνεται καλύτερα σε περιπτώσεις όπου ζητείται η παραγωγή μεσοπρόθεσμων προβλέψεων, σύμφωνα με τη θεωρία που παρουσιάστηκε στην Ενότητα 4.2.3, θεωρείται λογική η απόκλιση στην απόδοση που παρουσιάζει το εν λόγω μοντέλο σε σχέση με τα μοντέλα SES και Theta, εντούτοις η απόκλιση αυτή διατηρείται σε πολύ χαμηλά επίπεδα, γεγονός που υποδεικνύει πως το μοντέλο Damped μπορεί να παράγει αξιόλογες προβλέψεις, κάτι το οποίο οφείλεται στην βέλτιστη επιλογή των παραμέτρων α, β, και φ.

Η κατάλληλη επιλογή των συγκεκριμένων παραμέτρων μέσω της διαδικασίας της δυαδικής αναζήτησης, εξασφαλίζει την όσο το δυνατό καλύτερη προσαρμογή του μοντέλου στα δεδομένα της χρονοσειράς της ΟΤΣ, με αποτέλεσμα το μοντέλο Damped να καταλαμβάνει την 3η καλύτερη θέση, από σκοπιάς ακρίβειας των παραγόμενων προβλέψεων, ανάμεσα στα 6 μοντέλα που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης.

Βεβαίως, το ίδιο ισχύει και για τα 2 καλύτερα μοντέλα (SES, Theta), για τα οποία χρησιμοποιείται επίσης η μη γραμμική βελτιστοποίηση για την παράμετρο α του μοντέλου SES, προκειμένου να επιλέγεται κάθε φορά η βέλτιστη τιμή της συγκεκριμένης μεταβλητής εξομάλυνσης.

Στον παρακάτω πίνακα, παρουσιάζονται οι τιμές που λαμβάνουν οι παράμετροι εξομάλυνσης α, β, καθώς και η τιμή της παραμέτρου διόρθωσης της τάσης φ, κατά την τελευταία επανάληψη της διαδικασία του rolling, για τα μοντέλα Damped και SES, για την περίπτωση όπου παρουσίασαν το βέλτιστο δείκτη sMAPE στην κατηγορία επεξεργασίας της διπλής εποχικότητας:

DOUBLE SI			
ΜΟΝΤΕΛΟ	***Τιμή Παραμέτρου***		
	α	***β***	***φ***
SES	0,6773	-	-
Damped	0,6567	0,6497	0,0103

Πίνακας 8.1.2. - Επιλεγμένες τιμές παραμέτρων για διπλή εποχικότητα σε επίπεδο ώρας (SI=24 – 1 μήνας, SI2=168)

Η τιμή που λαμβάνει η παράμετρος α κατά τη διαδικασία υπολογισμού της γραμμής Theta(2), δεν εμφανίζεται στον παραπάνω πίνακα, διότι αποτελεί εσωτερική διεργασία της μεθόδου Theta και δεν κρίνεται σκόπιμη η καταγραφή της συγκεκριμένης τιμής, εντούτοις τονίζεται πως χρησιμοποιείται η διαδικασία της μη γραμμικής βελτιστοποίησης και σε αυτή την περίπτωση.

Παρατηρώντας τον Πίνακα 7.5.2, είναι εμφανές ότι τηρούνται ορισμένες βασικές θεωρητικές προϋποθέσεις, όπως είναι για παράδειγμα το γεγονός ότι η τιμή της παραμέτρου β είναι χαμηλότερη από εκείνη που λαμβάνει η παράμετρος α. Επιπλέον, είναι σημαντικό το γεγονός πως η παράμετρος α, λαμβάνει παρόμοια τιμή και για τα δύο μοντέλα, κάτι το οποίο ενδεχομένως προκαλεί την παραγωγή προβλέψεων με τόσο μικρή διαφορά ακρίβειας μεταξύ των βέλτιστων μοντέλων πρόβλεψης, όπως παρουσιάστηκαν στον Πίνακα 7.5.1.

Εντούτοις, προβληματίζει το γεγονός πως η παράμετρος β, αν και χαμηλότερη από την α, παρουσιάζει αρκετά υψηλή τιμή, κάτι το οποίο δεν θεωρείται ιδανικό, καθώς η χρονοσειρά της ΟΤΣ δεν παρουσιάζει τάση, και επομένως μια χαμηλή τιμή της παραμέτρου β πιθανό να βελτίωνε αισθητά την απόδοση του μοντέλου Damped.

Αναφορικά με τη τιμή που επιλέγεται για την παράμετρο φ, θεωρείται περισσότερο από αποδεκτή, καθώς επιλέγεται μέσα στο διάστημα [0, 1], το οποίο είναι λογικό καθώς πρόκειται για το μοντέλο φθίνουσας τάσης (Damped), και επιπλέον λαμβάνει τιμή κοντά στο 0, γεγονός το οποίο αποτυπώνει την απουσία τάσης στη χρονοσειρά της ΟΤΣ. Σημειώνεται, πως όσο πιο εμφανής είναι η παρουσία τάσης σε μία χρονοσειρά, τόσο πιο κοντά στη μονάδα είναι η τιμή της παραμέτρου φ. Στην περίπτωση που εξετάζεται, η απουσία τάσης είναι σαφής, και επομένως ορθώς επιλέγεται πολύ χαμηλή τιμή για τη μεταβλητή φ.

Ένα ακόμη γεγονός άξιο σχολιασμού, είναι εκείνο της επιλογής του κατάλληλου συνδυασμού των δύο ομάδων εποχικών δεικτών που υπολογίζονται στα πλαίσια της περίπτωσης αντιμετώπισης της διπλής εποχικότητας. Είναι σημαντικό να τονιστεί πως και οι τρείς βέλτιστες τιμές του δείκτη sMAPE, παρατηρούνται για την περίπτωση όπου οι 24 δείκτες ημερήσιας εποχικότητας υπολογίζονται με βάση μόνο 1 μήνα ιστορικών παρατηρήσεων.

Ένας πιθανός λόγος για το γεγονός αυτό, είναι ότι οι παρατηρήσεις βρίσκονται σε ωριαίο επίπεδο. Επομένως, έχοντας σα δείγμα 720 παρατηρήσεις (30 ημέρες επί 24 ώρες), είναι δυνατό να αποτυπωθεί στους εποχικούς δείκτες ένας αρκετά ικανός όγκος πληροφορίας σχετικά με το πρότυπο που ακολουθεί η χρονοσειρά.

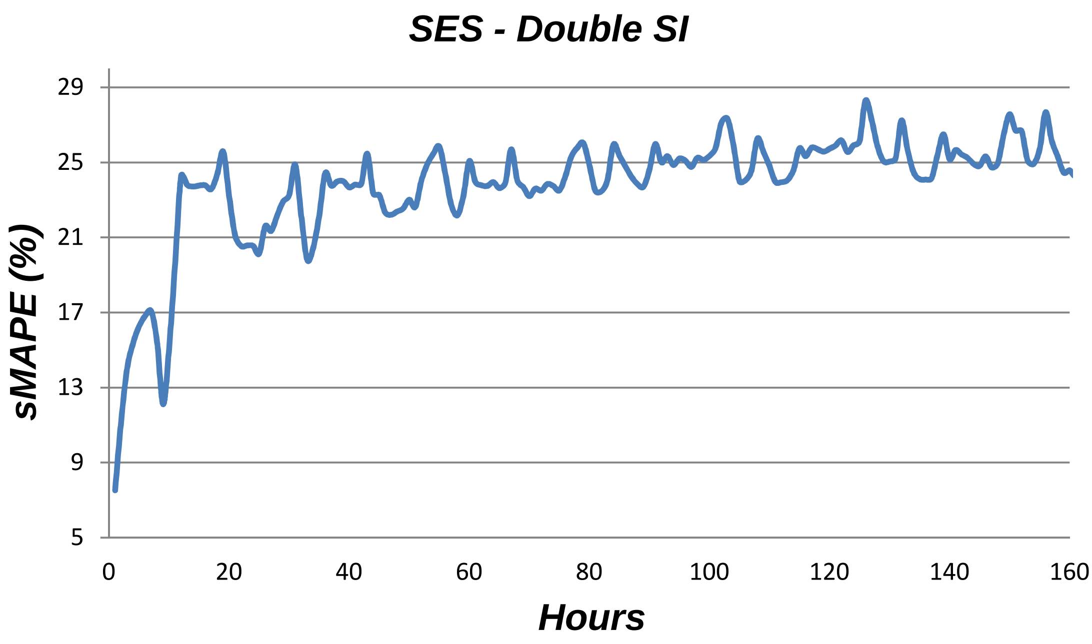
Επιπρόσθετα, η ανάδειξη του συγκεκριμένου συνδυασμού ως του βέλτιστου σε σύγκριση με τους υπόλοιπους που εξετάζονται, σημαίνει πως 1 μήνας παρατηρήσεων, θεωρείται ιδανικός για να καταγράψει τη σταδιακή αλλαγή μεταξύ των εποχών του έτους, και να αποτυπώσει καλύτερα την επίδραση που αυτή η αλλαγή μπορεί να έχει στη χρονοσειρά της ΟΤΣ.

Επίσης, η επιλογή όλων των προηγούμενων παρατηρήσεων στην εξαγωγή των 168 δεικτών εβδομαδιαίας εποχικότητας, θεωρείται ως η πιο ασφαλής, αποφεύγοντας έτσι οποιοδήποτε ενδεχόμενο εμφάνισης αποσυγχρονισμού των εποχικών δεικτών και των παρατηρήσεων με τις οποίες αντιστοιχίζονται.

8.2 Ποσοτική Ανάλυση Ακρίβειας Αποτελεσμάτων

Σε αυτό το σημείο, θεωρείται πολύ χρήσιμη η γραφική αναπαράσταση, καθώς και η ανάλυση της ακρίβειας των τριών βέλτιστων μοντέλων (SES, Theta, Damped), όπως αυτή προέκυψε από τη διαδικασία της προσομοίωσης, για την κατηγορία αποτελεσμάτων που σχετίζεται με την αντιμετώπιση της διπλής εποχικότητας, αλλά και το βέλτιστο συνδυασμό υπολογισμού των δύο ομάδων εποχικών δεικτών που χρησιμοποιήθηκαν κατά τη διαδικασία της κλασικής αποσύνθεσης.

Σε πρώτη φάση, εξετάζεται η απόδοση του «νικητήριου» μοντέλου SES, σε σχέση πάντα με τη μέση τιμή του στατιστικού δείκτη sMAPE που παρουσιάστηκε στον Πίνακα 7.4.1. Στο διάγραμμα που ακολουθεί, παρουσιάζονται οι τιμές του δείκτη sMAPE, όπως αυτές προκύπτουν για κάθε ώρα, με ορίζοντα μίας εβδομάδας:

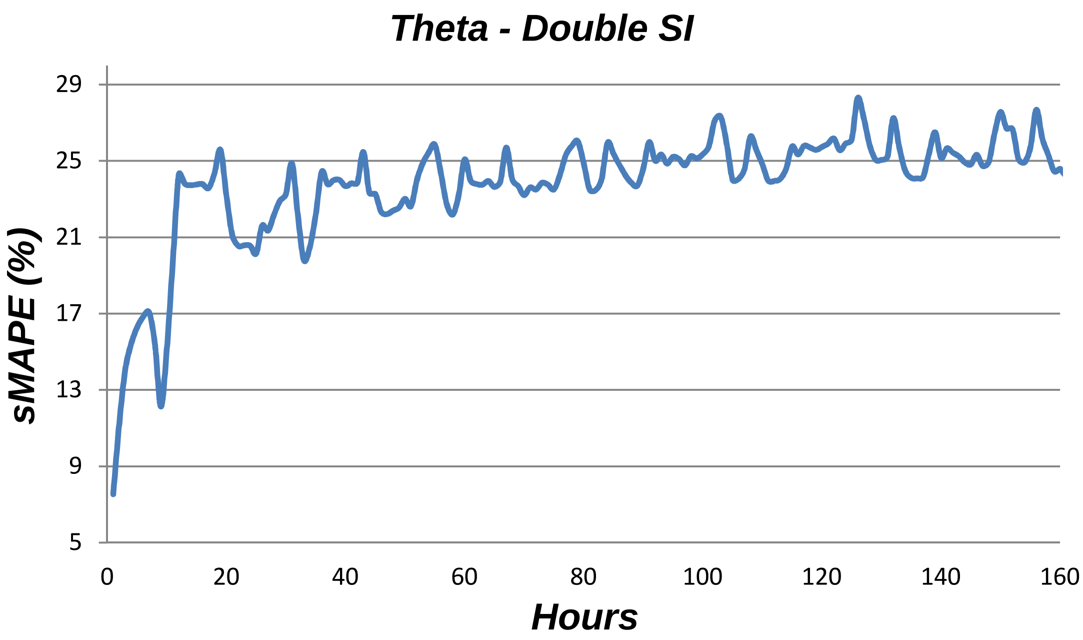


Σχ. 8.2.1 – Δείκτες sMAPE σε επίπεδο ώρας για ορίζοντα εβδομάδας, για το μοντέλο SES (double SI)

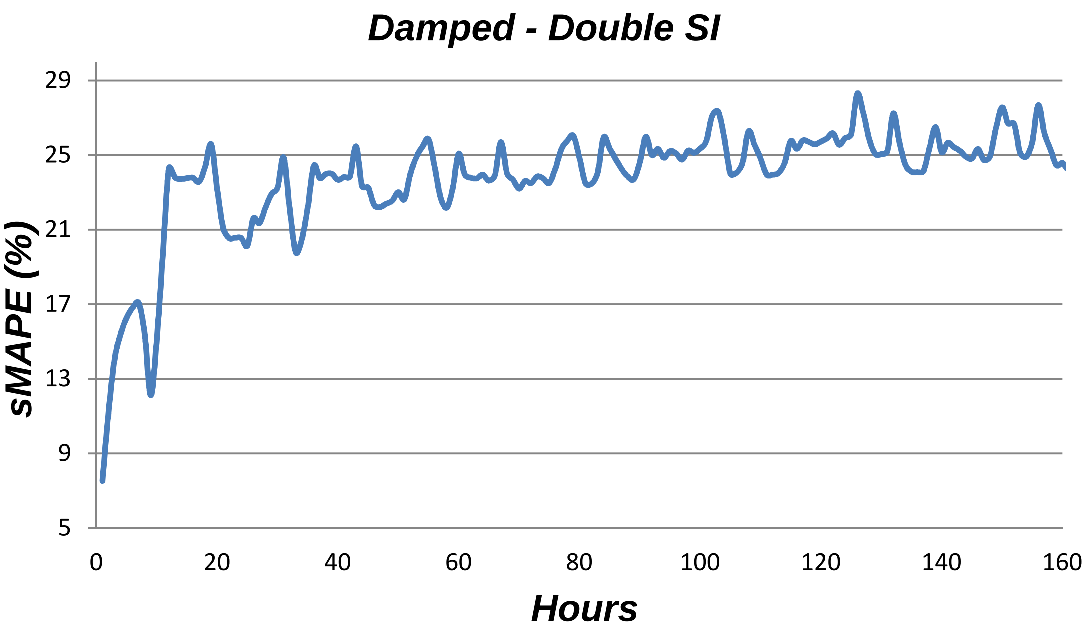
Παρατηρώντας το παραπάνω σχήμα, γίνεται εμφανής η παρουσία ενός εποχικού προτύπου, το οποίο οφείλεται στην αδυναμία πλήρους εξάλειψης της εποχικότητας από τη χρονοσειρά της ΟΤΣ, καθώς όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, χρησιμοποιείται ένας αρκετά απλοϊκός τρόπος για την αντιμετώπιση αυτής.

Όπως είναι αναμενόμενο, όσο προχωρούν οι ώρες αυξάνεται και ο δείκτης του sMAPE, με αποτέλεσμα να είναι εμφανής η παρουσία τάσης στο συγκεκριμένο διάγραμμα. Εντούτοις, παρατηρείται μία προοδευτική σταθεροποίηση των τιμών του σφάλματος γύρω από την τιμή 25%, γεγονός που υποδεικνύει την ικανοποιητική προσαρμογή του μοντέλου πρόβλεψης στην χρονοσειρά της ΟΤΣ.

Ενδιαφέρον παρουσιάζει η σύγκριση του σχήματος 8.2.1, με τα αντίστοιχα σχήματα που προκύπτουν για τα άλλα δύο μοντέλα, τα μοντέλα Theta και Damped. Στα ακόλουθα σχήματα, παρουσιάζονται οι δείκτες sMAPE για κάθε ώρα με ορίζοντα εβδομάδας, για τα μοντέλα Theta και Damped αντίστοιχα:



Σχ. 8.2.2 – Δείκτες sMAPE σε επίπεδο ώρας για ορίζοντα εβδομάδας, για το μοντέλο Theta (double SI)



Σχ. 8.2.3 – Δείκτες sMAPE σε επίπεδο ώρας για ορίζοντα εβδομάδας, για το μοντέλο Damped (double SI)

Παρατηρώντας και τα 3 σχήματα (8.2.1 - 8.2.3), καθίσταται εμφανές ότι οι διαφορές είναι ελάχιστες, και δε θα ήταν υπερβολή να ειπωθεί ότι είναι σχεδόν αδύνατο να εντοπιστεί κάποια κραυγαλέα διαφορά στην ακρίβεια των τριών μοντέλων πρόβλεψης που εξετάζονται. Αυτό ήταν αναμενόμενο, καθώς οι βέλτιστες μέσες τιμές του δείκτη sMAPE που παρουσιάστηκαν στην Ενότητα 8.1, απέχουν ελάχιστα μεταξύ τους, γεγονός που προμήνυε παρόμοια γραφική συμπεριφορά των ωριαίων σφαλμάτων των τριών διαφορετικών μοντέλων πρόβλεψης. Επιπλέον, και στα 3 διαγράμματα εμφανίζεται ένα κοινό εποχικό πρότυπο, το οποίο όπως προαναφέρθηκε οφείλεται στην ατελή εξάλειψη της εποχικότητας από τη χρονοσειρά της ΟΤΣ.

Έχοντας μία ολοκληρωμένη εικόνα της μορφής των ωριαίων σφαλμάτων, για τα τρία βέλτιστα μοντέλα πρόβλεψης, θεωρείται πλέον απαραίτητο να εξεταστεί αλλά και αναλυθεί η επίδραση του ορίζοντα πρόβλεψης στην επίδοση του βέλτιστου μοντέλου SES, έχοντας πάντα ως κριτήριο τη βελτίωση του στατιστικού δείκτη sMAPE.

Η ανάλυση αυτή αφορά κυρίως στη συμπεριφορά των δεικτών σφάλματος sMAPE, δηλαδή την πορεία που διαγράφουν οι τιμές των σφαλμάτων κατά τη διάρκεια του ορίζοντα που έχει τεθεί κατά τη διαδικασία της προσομοίωσης. Για το λόγο αυτό, η ορθή επιλογή του ορίζοντα πρόβλεψης είναι μείζονος σημασίας και καθορίζει εν πολλοίς την ακρίβεια των παραγόμενων προβλέψεων.

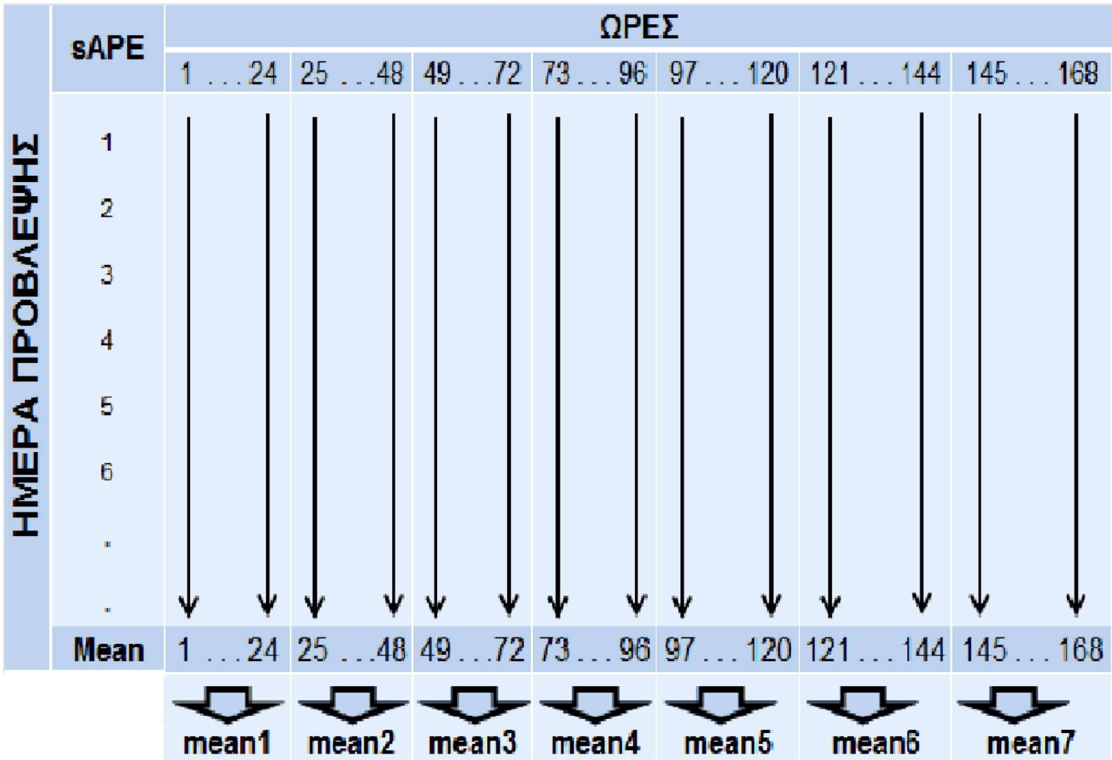
Ο ορίζοντας πρόβλεψης δεν έχει μόνο μία έννοια. Αποτελεί παράλληλα και την περίοδο της διαδικασίας του rolling, καθώς επηρεάζει και τη συχνότητα ανανέωσης των προβλέψεων. Ουσιαστικά πρόκειται για μία τρισυπόστατη παράμετρο, η οποία παίζει ίσως το σπουδαιότερο ρόλο στην προσπάθεια βελτίωσης της ακρίβειας των παραγόμενων προβλέψεων. Επομένως, η σχέση που περιγράφει τον τρόπο με τον οποίο σχετίζονται οι τρείς διαφορετικές έννοιες του ορίζοντα πρόβλεψης είναι η εξής:

$$\textit{horizon} = \textit{Περίοδος rolling} = \frac{1}{Συχνότητα\ Ανανέωσης\ Προβλέψεων}$$

Από την παραπάνω σχέση καθίσταται σαφές, ότι όσο αυξάνεται ο ορίζοντας τόσο μειώνεται η συχνότητα ανανέωσης των προβλέψεων, και επομένως επηρεάζεται η ακρίβεια των προβλέψεων. Αυτό που ενδιαφέρει είναι ο τρόπος με τον οποίο επιδρά η αυξομείωση της συχνότητας ανανέωσης των προβλέψεων, και επομένως του ορίζοντα πρόβλεψης, στην ακρίβεια των προβλέψεων.

Για να γίνει κατανοητή αυτή η επίδραση, κρίνεται απαραίτητο να εξετασθεί η ακρίβεια των προβλέψεων σε ημερήσιο επίπεδο, και όχι σε ωριαίο, όπως παρουσιάστηκε στα σχήματα 8.2.1 έως 8.2.3. Αυτό επιτυγχάνεται με τον υπολογισμό των μέσων ημερήσιων δεικτών sMAPE, δηλαδή υπολογίζεται η μέση τιμή των δεικτών sAPE κάθε ώρας για κάθε ημέρα πρόβλεψης (sMAPE).

Στη συνέχεια, έχοντας τις μέσες τιμές των σφαλμάτων για κάθε ώρα για μία εβδομάδα, υπολογίζεται ο μέσος όρος για κάθε 24 ώρες, δηλαδή για κάθε ημέρα από τις 7 ημέρες της εβδομάδας (mean(sMAPE)), καθώς αυτός είναι και ο αρχικά καθορισμένος ορίζοντας πρόβλεψης κατά τη διαδικασία προσομοίωσης. Τελικά, προκύπτουν 7 μέσες τιμές του δείκτη sMAPE, μία για κάθε ημέρα της εβδομάδας. Για λόγους κατανόησης, η μεθοδολογία που περιγράφηκε, παρουσιάζεται στο ακόλουθο σχήμα:



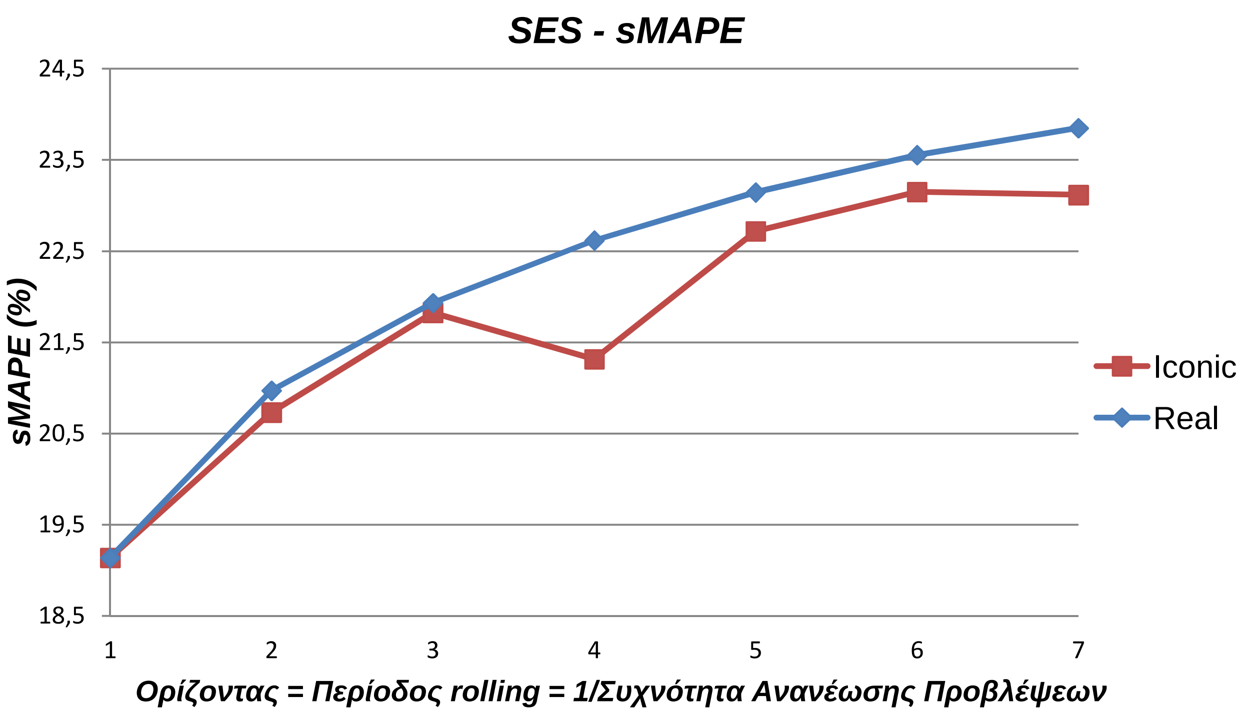
Σχ. 8.2.4 – Μεθοδολογία υπολογισμού μέσων τιμών sMAPE

Έχοντας, υπολογίσει τα ημερήσια μέσα σφάλματα sMAPE, είναι δυνατή η παραγωγή ενός γραφήματος το οποίο δύναται να παρουσιάσει την επίδραση της συχνότητας ανανέωσης των προβλέψεων στην ακρίβεια των προβλέψεων. Το διάγραμμα αυτό, απαρτίζεται από δύο διαφορετικές γραφικές παραστάσεις.

Η πρώτη, αφορά το δείκτη του μέσου sMAPE που παρουσιάζεται καθώς ο ορίζοντας πρόβλεψης αυξάνεται από τη μία στις επτά ημέρες. Επίσης, η συγκεκριμένη γραφική παράσταση έχει την ιδιομορφία ότι οι νέες παραγόμενες προβλέψεις αντικαθιστούν τις προηγούμενες, με αποτέλεσμα οι προβλέψεις που υπολογίστηκαν για ορίζοντα ίσο με 1 ημέρα, και αφορούν, να επικαλύπτονται από εκείνες που παράχθηκαν για ορίζοντα 2 ημερών κοκ. Επομένως, η πρώτη γραφική παράσταση δεν παρουσιάζει τις πραγματικά παρατηρούμενες τιμές του μέσου δείκτη sMAPE.

Αντίθετα, η δεύτερη γραφική παράσταση αφορά τις πραγματικά παρατηρούμενες τιμές του δείκτη του μέσου sMAPE. Στην περίπτωση αυτή, οι νέες προβλέψεις δεν επικαλύπτουν τις προηγούμενες, κάτι το οποίο επιτυγχάνεται μέσω της μεταβολής της περιόδου του rolling, δηλαδή οι προβλέψεις παράγονται κάθε 2η μέρα, ή κάθε 3η ημέρα κοκ.

Στο διάγραμμα που ακολουθεί, παρουσιάζονται τα δύο παραπάνω γραφήματα, που αφορούν την τιμή του μέσου δείκτη sMAPE συναρτήσει του ορίζοντα πρόβλεψης (ή περίοδος rolling, ή 1/συχνότητα ανανέωσης προβλέψεων):

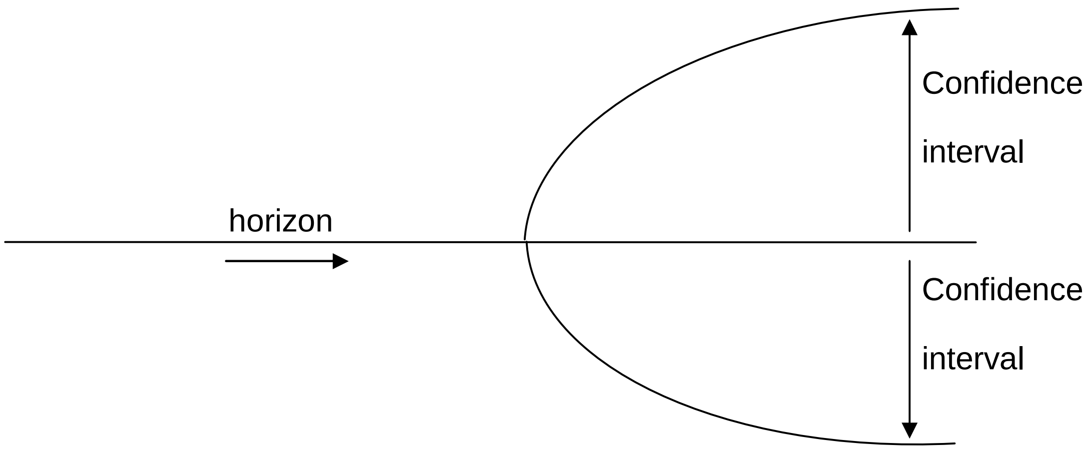


Σχ. 8.2.5 – Μέσοι δείκτες sMAPE συναρτήσει του ορίζοντα πρόβλεψης γα το μοντέλο SES (double SI)

Παρατηρώντας το παραπάνω σχήμα, γίνεται εμφανές πως όσο αυξάνεται ο ορίζοντας πρόβλεψης, και επομένως μειώνεται η συχνότητα ανανέωσης των προβλέψεων, τόσο μεγαλύτερη είναι και η μέση τιμή του δείκτη sMAPE, και επομένως τόσο χειρότερη είναι η ακρίβεια των παραγόμενων προβλέψεων.

Αυτό σημαίνει πως, μπορεί να επιλεγεί μεγάλος ορίζοντας πρόβλεψης, όπως για παράδειγμα οι 7 ημέρες (1 εβδομάδα), αλλά το τίμημα είναι η αύξηση του δείκτη sMAPE κατά περίπου 5 ποσοστιαίες μονάδες σε σχέση με την περίπτωση όπου επιλέγεται ορίζοντας μίας ημέρας. Αξίζει να τονισθεί, πως η έντονη διαφορά που παρατηρείται μεταξύ των δύο γραφημάτων, στην περίπτωση που ο ορίζοντας είναι 4 ημέρες, μπορεί να αποδοθεί σε καθαρά τυχαίο γεγονός.

Η μορφή των δύο γραφικών παραστάσεων του σχήματος 8.2.5 δεν είναι τυχαία, αλλά βασίζεται στη θεωρία των διαστημάτων εμπιστοσύνης των παραγόμενων προβλέψεων. Η θεωρία αυτή αναφέρει πως όσο αυξάνεται ο ορίζοντας πρόβλεψης, τόσο αυξάνεται το εύρος των διαστημάτων εμπιστοσύνης, αλλά μέχρι ενός σημείου, από το οποίο παρατηρείται τάση σύγκλισης. Στο ακόλουθο σχήμα, παρουσιάζεται η θεωρητική μορφή του εύρους των διαστημάτων εμπιστοσύνης, προκειμένου να γίνει αντιληπτή η ομοιομορφία σε σύγκριση με τα γραφήματα του σχήματος 8.2.5:

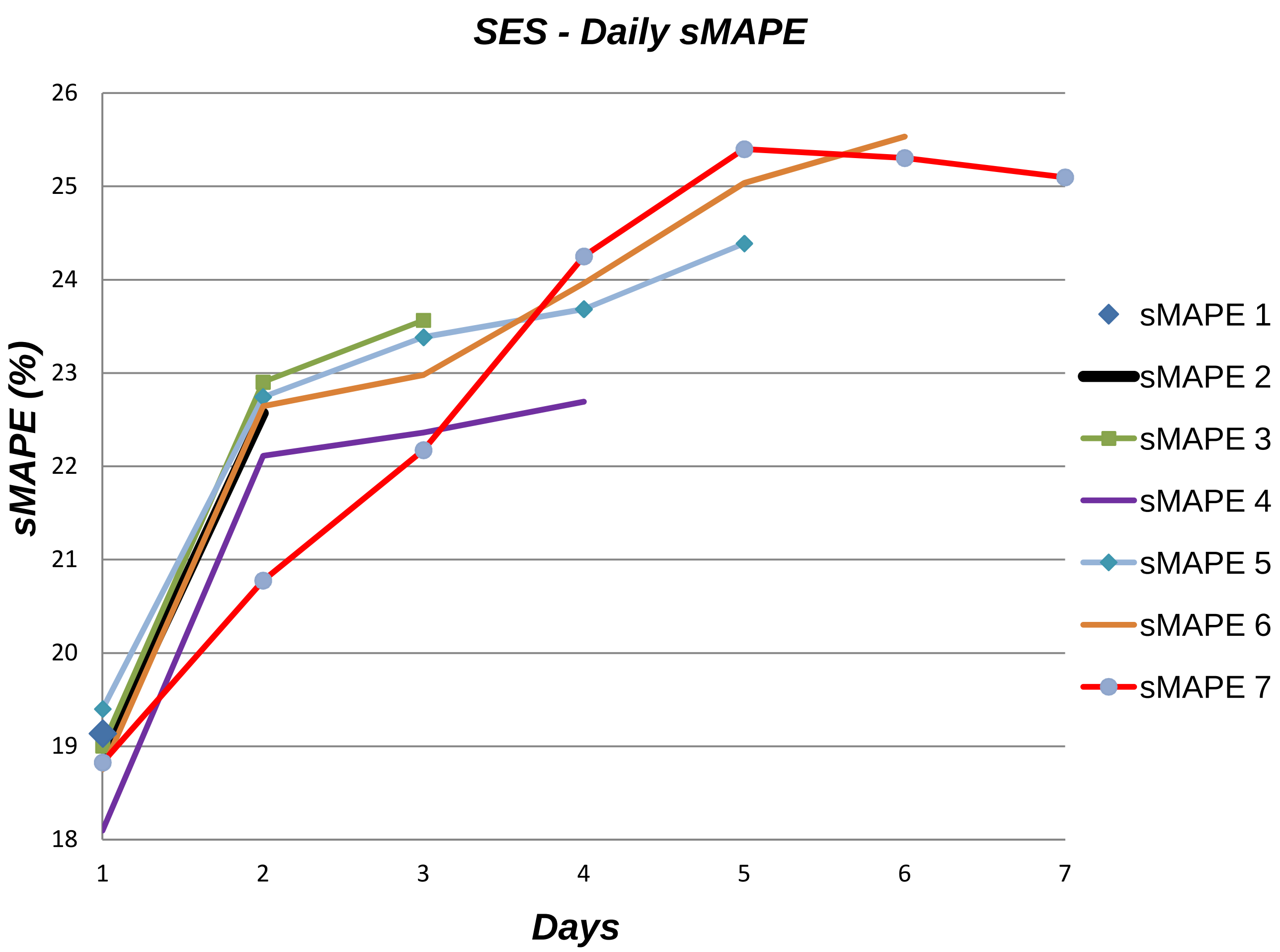


Σχ. 8.2.6 – Εύρος διαστημάτων εμπιστοσύνης συναρτήσει του ορίζοντα πρόβλεψης

Στο σημείο αυτό, παρουσιάζει μεγάλο ενδιαφέρον η περαιτέρω ανάλυση του σχήματος 8.2.5, μέσω της παραγωγής ενός δεύτερου διαγράμματος στο οποίο θα αποτυπώνεται ο τρόπος με τον οποίο κατανέμεται η κάθε μία τιμή των 7 δεικτών σφάλματος που παρουσιάστηκαν, για κάθε ημέρα πρόβλεψης.

Στην ουσία, πρόκειται για ένα γράφημα το οποίο παρουσιάζει τους μέσους ημερήσιους δείκτες σφάλματος, όπως αυτοί διαμορφώνονται για κάθε διαφορετική περίπτωση του ορίζοντα πρόβλεψης (1-7 ημέρες).

Στο συγκεντρωτικό διάγραμμα που ακολουθεί, παρουσιάζονται τα γραφήματα που αφορούν τους πραγματικά παρατηρούμενους ημερήσιους μέσους δείκτες sMAPE, που αντιστοιχούν στις 7 διαφορετικές περιπτώσεις που λαμβάνονται υπόψη για τον ορίζοντα πρόβλεψης:



Σχ. 8.2.7 – Μέσοι ημερήσιοι δείκτες sMAPE για τις 7 διαφορετικές τιμές του ορίζοντα πρόβλεψης γα το μοντέλο SES (double SI)

Παρατηρώντας το παραπάνω σχήμα, είναι εμφανής η διαφορά που επέρχεται στο δείκτη των μέσων ημερήσιων σφαλμάτων sMAPE, για κάθε διαφορετική τιμή του ορίζοντα πρόβλεψης. Ουσιαστικά, η πληροφορία που εξάγεται από το σχήμα 8.2.7, αφορά τον τρόπο που μεταβάλλεται η ακρίβεια των προβλέψεων, για τις περιπτώσεις όπου οι προβλέψεις ανανεώνονται ανά 1, 2, 3, 4, 5, 6 και 7 ημέρες αντίστοιχα.

Αξίζει να αναφερθεί πως η κάθε μία από τις 7 διαφορετικές γραφικές παραστάσεις, αναπαριστά τιμές του ημερήσιου δείκτη sMAPE ανάλογα με τον ορίζοντα στον οποίο αντιστοιχεί. Για παράδειγμα, όταν ο ορίζοντας είναι 1 ημέρα, έχει νόημα να αναπαρασταθεί μόνο το σφάλμα για την πρώτη ημέρα, καθώς πάντα προβλέπεται μία ημέρα μπροστά. Όταν είναι 2 ημέρες, αναπαρίστανται τα σφάλματα για 2 ημέρες, διότι κάθε φορά προβλέπονται 2 ημέρες μπροστά κοκ.

Επίσης, είναι εμφανής η ομοιότητα της γραφικής παράσταση που αντιστοιχεί στον ορίζοντα 7 ημερών, αλλά και των υπολοίπων, με τη θεωρητική μορφή του εύρους των διαστημάτων εμπιστοσύνης, που παρουσιάστηκε στο σχήμα 8.2.6, σύμφωνα με το οποίο όσο αυξάνεται ο ορίζοντας πρόβλεψης τόσο αυξάνονται και οι τιμές των σφαλμάτων, γεγονός απόλυτα φυσιολογικό καθώς υπάρχει άμεση σχέση του παραπάνω σχήματος με το σχήμα 8.2.5.

Συγκεκριμένα, το σχήμα 8.2.7, είναι ουσιαστικά η μεγέθυνση του σχήματος 8.2.5, καθώς η μέση τιμή των σημείων που αποτελούν κάθε γράφημα για κάθε μία από τις 7 περιπτώσεις του ορίζοντα πρόβλεψης, ισούται με την αντίστοιχη τιμή του μέσου δείκτη sMAPE, στο σχήμα 8.2.5. Επίσης, αξίζει να τονιστεί, πως στην περίπτωση που ο ορίζοντας πρόβλεψης ισούται με 7 ημέρες (sMAPE7, Σχ. 8.2.7), η ανανέωση των προβλέψεων πραγματοποιείται πάντα την ίδια μέρα, γεγονός που δεν επηρεάζει αρνητικά την ακρίβεια των προβλέψεων, σε περίπτωση φυσικά που έχει πραγματοποιηθεί η κατάλληλη επεξεργασία για την απαλοιφή της εποχικότητας.

Σε κάθε άλλη περίπτωση (sMAPE1-sMAPE6, Σχ. 8.2.7), η ανανέωση των προβλέψεων γίνεται σε διαφορετική μέρα κάθε φορά, έως ότου ολοκληρωθεί ο κύκλος και να επιστρέψει η διαδικασία στην ίδια μέρα στην οποία είχε ξεκινήσει. Ο κύκλος αυτός, παρουσιάζει μεγαλύτερη διάρκεια όσο μεγαλύτερη είναι και η επιλογή του ορίζοντα πρόβλεψης.

8.3 Συμπεράσματα

Με δεδομένη την ολοκλήρωση της διαδικασίας της προσομοίωσης, καθώς και την εκτενή ανάλυση του συνόλου των αποτελεσμάτων για κάθε μία από τις 4 κατηγορίες αντιμετώπισης της εποχικότητας, είναι δυνατή η εξαγωγή των τελικών συμπερασμάτων αναφορικά με τη βέλτιστη επιλογή όλων εκείνων των παραμέτρων οι οποίες καθορίζουν την ακρίβεια των παραγόμενων προβλέψεων.

8.3.1 Βέλτιστη Κατηγορία Αντιμετώπισης Εποχικότητας

Για την αντιμετώπισης της εποχικότητας στη χρονοσειρά της ΟΤΣ, αναπτύχθηκαν και εφαρμόστηκαν 4 διαφορετικές μεθοδολογίες προσέγγισης του συγκεκριμένου προβλήματος. Ουσιαστικά, οι τρείς από αυτές τις μεθοδολογίες αναφέρονται σε απαλοιφή της εποχικότητας, καθώς η 1^{η} κατηγορία αφορά την παραγωγή προβλέψεων δίχως οποιουδήποτε είδους επεξεργασία της εποχικότητας.

Σύμφωνα με τις παραπάνω κατηγορίες, παράχθηκαν τέσσερις αντίστοιχες ομάδες αποτελεσμάτων, λαμβάνοντας ως κριτήριο τη μέση τιμή του δείκτη του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος (sMAPE). Σύμφωνα με τα παραχθέντα αποτελέσματα, η 4^{η} κατηγορία, η οποία αναφέρεται στην αντιμετώπιση της διπλής εποχικότητας, παρουσίασε τα βέλτιστα αποτελέσματα σε σύγκριση με τις υπόλοιπες, γεγονός το οποίο αποδίδεται στην έντονη επίδραση της διπλής εποχικότητας στη χρονοσειρά της ΟΤΣ, γεγονός το οποίο δε λαμβάνεται υπόψη από τις άλλες κατηγορίες προσέγγισης της εποχικότητας.

Η μεθοδολογία που αναπτύχθηκε βασίζεται στην παρουσία δύο συνιστωσών εποχικότητας, και στην εξαγωγή δύο ομάδων εποχικών δεικτών για την αντιμετώπισή τους. Η πρώτη ομάδα, αφορά το ημερήσιο εποχικό πρότυπο (Σχ. 6.1.7), και αποτελείται από 24 εποχικούς δείκτες, οι οποίοι αντιστοιχίζονται σε κάθε ώρα της ημέρας.

Καθοριστικός παράγοντας για τον υπολογισμό των συγκεκριμένων δεικτών, είναι η κατάλληλη επιλογή του όγκου των παρελθόντων παρατηρήσεων, σύμφωνα με τις οποίες εξάγονται οι 24 εποχικοί δείκτες, ενσωματώνοντας με αυτό τον τρόπο την απαραίτητη πληροφορία του ημερήσιου προτύπου εποχικότητας.

Η δεύτερη ομάδα, αφορά το εβδομαδιαίο εποχικό πρότυπο (Σχ. 6.1.8), το οποίο είναι ιδιαίτερα εμφανές στη χρονοσειρά της ΟΤΣ. Για την αντιμετώπιση αυτής της συνιστώσας εποχικότητας, εξάγονται 168 δείκτες εβδομαδιαίας εποχικότητας, καθώς τα δεδομένα είναι σε επίπεδο ώρας.

Στην περίπτωση αυτή, παίζει και πάλι ρόλο ο όγκος των προηγούμενων παρατηρήσεων που λαμβάνονται υπόψη για την εξαγωγή των 168 εποχικών δεικτών, όμως για την αποφυγή ολίσθησης μεταξύ των δεικτών και των παρατηρήσεων στις οποίες αντιστοιχούν, επιλέγεται η εξαγωγή των εβδομαδιαίων εποχικών δεικτών με βάση όλα τα παρελθόντα δεδομένα.

Έχοντας υπολογίσει τις δύο ομάδες εποχικών δεικτών, ήταν απαραίτητο να ευρεθεί ο βέλτιστος συνδυασμός μεταξύ αυτών, συναρτήσει του όγκου των παρελθόντων παρατηρήσεων. Τα βέλτιστα αποτελέσματα, σύμφωνα με τον Πίνακα 7.4.1, παρουσιάστηκαν για το συνδυασμό όπου λήφθηκε υπόψη 1 μήνας παρελθόντων παρατηρήσεων για την εξαγωγή των 24 δεικτών ημερήσιας εποχικότητας. Τα γεγονός αυτό οφείλεται στην ωριαία φύση των δεδομένων της ΟΤΣ, με αποτέλεσμα ο 1 μήνας, δηλαδή 720 παρατηρήσεις, να αποδεικνύονται επαρκείς για την αποτύπωση της πληροφορίας του ημερήσιου προτύπου εποχικότητας.

Σε αυτό το σημείο, πρέπει να τονιστεί για ακόμη μία φορά, πως η μεθοδολογία που ακολουθείται για την αντιμετώπιση της διπλής εποχικότητας, θεωρείται ιδιαίτερα απλοϊκή. Υπάρχουν ποικίλες μεθοδολογίες προσέγγισης της διπλής εποχικότητας, όπως για παράδειγμα η μέθοδος Taylor για διπλή εποχικότητα ή η μέθοδος Holt-Winters, για τις οποίες θα γίνει αναφορά στο Κεφάλαιο 9.

8.3.2 Βέλτιστο Μοντέλο Πρόβλεψης

Μετά την επιλογή της βέλτιστης κατηγορίας αντιμετώπισης της εποχικότητας, σειρά έχει η επιλογή του βέλτιστου μοντέλου πρόβλεψης με κριτήριο τη μέση τιμή του στατιστικού δείκτη sMAPE. Κατά τη διαδικασία της προσομοίωσης , χρησιμοποιούνται 6 μοντέλα, συμπεριλαμβανομένου και του μοντέλου αναφοράς Naive. Τα υπόλοιπα 5 μοντέλα είναι τα ακόλουθα:

- Μοντέλο Απλής Γραμμικής Παλινδρόμησης (LRL)
- Μοντέλο Theta (κλασική μέθοδος Theta)
- Μοντέλα Εκθετικής Εξομάλυνσης (SES, Holt, Damped)

Το μοντέλο αναφοράς Naive, παρουσίασε ικανοποιητικά αποτελέσματα, συγκριτικά με την απλότητά του, γεγονός το οποίο αποδίδεται στην ιδιαίτερη φύση της χρονοσειράς της ΟΤΣ. Γενικώς, το μοντέλο Naive, τείνει να προσαρμόζεται καλύτερα σε χρονοσειρές που παρουσιάζουν υψηλή μεταβλητότητα, κάτι το οποίο επιβεβαιώθηκε και μέσω των αποτελεσμάτων και για τις 4 κατηγορίες χειρισμού της εποχικότητας, καθώς το συγκεκριμένο μοντέλο διατηρείται ψηλά από σκοπιά επιδόσεων σε κάθε μία από αυτές.

Εντελώς αντίθετη είναι η συμπεριφορά του μοντέλου εκθετικής εξομάλυνσης Holt. Το μοντέλο Holt, έχει σαν βασικό χαρακτηριστικό του την αναγνώριση τάσης που είναι δυνατό να χαρακτηρίζει μία χρονοσειρά. Εντούτοις, η χρονοσειρά της ΟΤΣ, χαρακτηρίζεται από απουσία τάσης (στάσιμη προς το μέσο), γεγονός το οποίο καθιστά το μοντέλο Holt ανίκανο να προσαρμοστεί στην υπό εξέταση χρονοσειρά. Η αδυναμία αυτή του συγκεκριμένου μοντέλου, αποτυπώνεται ξεκάθαρα στις επιδόσεις του, και για τις 4 κατηγορίες αντιμετώπισης εποχικότητας, όπου παρουσιάζει πολύ μεγάλη απόκλιση σε σχέση με την ακρίβεια των υπόλοιπων μοντέλων πρόβλεψης.

Αναφορικά με το μοντέλο απλής γραμμικής παλινδρόμησης, αξίζει να σημειωθεί πως παρουσιάζει καλύτερες επιδόσεις από εκείνες του μοντέλου Holt, αλλά παραμένει αρκετά μακριά σε σύγκριση με τα υπόλοιπα μοντέλα και το μοντέλο αναφοράς Naive. Το γεγονός αυτό δεν προκαλεί εντύπωση καθώς το συγκεκριμένο μοντέλο πρόβλεψης, ακολουθεί ένα μονοδιάστατο τρόπο προσαρμογής στη χρονοσειρά, αδυνατώντας να συλλάβει τις όποιες ιδιαιτερότητες αυτή παρουσιάζει. Εντούτοις, η απόδοση του μοντέλου LRL, είναι δυνατό να κριθεί ικανοποιητική σε σχέση με τις αρχικές προσδοκίες για το εν λόγω μοντέλο, γεγονός το οποίο πιθανό να οφείλεται στη στασιμότητα της χρονοσειράς.

Τα τρία βέλτιστα μοντέλα, είναι τα δύο εναπομείναντα μοντέλα εκθετικής εξομάλυνσης και η συνδυαστική μέθοδος Theta. Οι διαφορές μεταξύ των τριών νικητήριων μοντέλων είναι ελάχιστες, όσον αφορά την τιμή του μέσου δείκτη sMAPE. Με μικρή διαφορά, η πρώτη θέση ανήκει στο μοντέλο SES, η δεύτερη στο μοντέλο Theta, και τρίτο έρχεται το μοντέλο Damped.

Οι μικρές αλλά υπαρκτές διαφορές μεταξύ των τριών μοντέλων, καθορίζονται εν πολλοίς από την απουσία τάσης στη χρονοσειρά της ΟΤΣ. Το μοντέλο SES, δε λαμβάνει υπόψη του την παρουσία τάσης σε μία χρονοσειρά, γεγονός που υποδεικνύει το πλεονέκτημα που έχει έναντι των μοντέλων Theta και Damped, τα οποία αμφότερα βασίζουν τη λειτουργία τους στην παρουσία τάσης.

Ιδιαίτερη μνεία γίνεται για το μοντέλο Theta, το οποίο δεν είχε εφαρμοστεί ποτέ μέχρι τώρα για την παραγωγή προβλέψεων για τη χρονοσειρά της ΟΤΣ. Τα αποτελέσματα είναι μάλλον ενθαρρυντικά για τη χρήση του συγκεκριμένου μοντέλου πρόβλεψης σε τέτοιου είδους χρονοσειρές, καθώς δείχνει να προσαρμόζεται σε μεγάλο βαθμό στα ιδιαίτερα χαρακτηριστικά της χρονοσειρά της ΟΤΣ, λόγος για τον οποίο καταλαμβάνει την 2^{η} καλύτερη θέση μετά το μοντέλο SES, και μάλιστα με πολύ μικρή διαφορά. Επομένως, η πρώτη απόπειρα χρήσης της κλασικής μεθόδου Theta, για την παραγωγή προβλέψεων στη χρονοσειρά της ΟΤΣ, θεωρείται αρκούντως ικανοποιητική.

Η βασική διαφορά του μοντέλου Theta με το μοντέλο SES, είναι η γραμμή Theta(0), η οποία δεν είναι άλλη από την απλή γραμμική παλινδρόμηση, ενώ η γραμμή Theta(2) είναι το μοντέλο SES. Επομένως, ο συνυπολογισμός της γραμμής Theta(0), η οποία αφορά την παρουσία μακροπρόθεσμης τάσης στη χρονοσειρά, επηρεάζει αρνητικά την ακρίβεια των προβλέψεων του μοντέλου Theta, γεγονός που εξηγεί την μικρή αλλά σημαντική διαφορά από το πρωτοπόρο μοντέλο SES.

Το μοντέλο Damped, παρουσιάζει την 3^{η} καλύτερη επίδοση από πλευράς ακρίβειας, με ελάχιστη διαφορά από τα δύο καλύτερα μοντέλα. Η διαφορά αυτή, οφείλεται ξανά στην απουσία τάσης από τη χρονοσειρά της ΟΤΣ, γεγονός το οποίο αντιμετωπίζεται με την κατάλληλη τιμή της παραμέτρου εξομάλυνσης β, μέσω της διαδικασίας της μη γραμμικής βελτιστοποίησης. Εντούτοις, η τιμή της συγκεκριμένης παραμέτρου λαμβάνει σημαντικά υψηλές τιμές, ακόμη και μέσω της διαδικασίας βελτιστοποίησης (Πίνακας 8.1.2). Το γεγονός αυτό είναι και ο λόγος για τον οποίο το μοντέλο Damped αδυνατεί να βελτιώσει την απόδοσή του και να πλησιάσει το νικητήριο μοντέλο SES.

Τελικά, το συμπέρασμα που προκύπτει για την επιλογή του βέλτιστου μοντέλου πρόβλεψης, βασίζεται σε μεγάλο βαθμό από την απουσία τάσης στη χρονοσειρά της ΟΤΣ. Το γεγονός αυτό, οδηγεί τα αυξημένης πολυπλοκότητας μοντέλα σε αδυναμία να αξιοποιήσουν το μέγιστο των δυνατοτήτων τους, ενώ αντίθετα δημιουργεί ευνοϊκές συνθήκες για την εφαρμογή απλούστερων μεθόδων, όπως είναι το μοντέλο SES ή ακόμη και το μοντέλο Naive, μετατρέποντας σε πλεονέκτημα την αδυναμία τους να αναλύσουν κάθε ποιοτικό χαρακτηριστικό μία χρονοσειράς.

8.3.3 Βέλτιστος Ορίζοντας Πρόβλεψης

Η επιλογή του ορίζοντα πρόβλεψης παίζει πολύ σημαντικό ρόλο στην ακρίβεια των παραγόμενων προβλέψεων. Η σημασία του ορίζοντα πρόβλεψης αναλύθηκε διεξοδικά στην προηγούμενη ενότητα (Ενότητα 8.2), και σύμφωνα με αυτή την ανάλυση είναι δυνατό να εξαχθούν ορισμένα χρήσιμα συμπεράσματα σχετικά με τη συγκεκριμένη παράμετρο.

Όπως έχει ήδη αναφερθεί, ο ορίζοντας πρόβλεψης αποτελεί παράλληλα και την περίοδο της επαναληπτικής διαδικασίας του rolling, ενώ ταυτόχρονα επηρεάζει με αντιστρόφως ανάλογο τρόπο και τη συχνότητα ανανέωσης των προβλέψεων. Το συμπέρασμα που εξάγεται από τη μελέτη και ανάλυση της επίδρασης του ορίζοντα πρόβλεψης στην ακρίβεια των προβλέψεων, είναι ότι όσο μεγαλύτερος είναι ο ορίζοντας που επιλέγεται, τόσο μικρότερη είναι η συχνότητα ανανέωσης των προβλέψεων, και επομένως τόσο μεγαλύτερα είναι και τα σφάλματα που παρατηρούνται.

Εντούτοις, στα πλαίσια της παρούσας διπλωματικής, το ζητούμενο είναι η παραγωγή προβλέψεων, με ορίζοντα μίας εβδομάδας. Επομένως, προκειμένου να επιτευχθεί ο συγκεκριμένος στόχος, είναι αναγκαίο να πληρωθεί το τίμημα της απώλειας ακρίβειας, προκειμένου να παραχθεί μεγαλύτερος όγκος προβλέψεων.

Φυσικά, η ακρίβεια μπορεί να αυξηθεί, στην περίπτωση που επιλεχθεί μικρός ορίζοντας πρόβλεψης, όμως είναι ξεκάθαρη η πρόθεση του παρόντος εγχειρήματος πρόβλεψης της χρονοσειράς της ΟΤΣ, να παράγει προβλέψεις που είναι μεν μέσα στα πλαίσια της λογικής του day ahead forecasting, αλλά προσπαθούν να εξετάσουν το όλο ζήτημα και από μία περισσότερο μεσοπρόθεσμη σκοπιά, καθώς είναι δυνατό να θεωρηθεί μεσοπρόθεσμη η πρόβλεψη μίας εβδομάδας από τη στιγμή που οι παρατηρήσεις είναι σε ωριαίο επίπεδο, κάτι το οποίο δεν επιχειρείται ευρέως στη διαθέσιμη μέχρι σήμερα βιβλιογραφία.

Κεφάλαιο 9. Προεκτάσεις

Στο παρών Κεφάλαιο, δίνεται η ευκαιρία να παρουσιαστούν ορισμένα εγχειρήματα τα οποία διατηρήθηκαν εκτός του φάσματος εξέτασης της συγκεκριμένης Διπλωματικής, αλλά η σημασία τους θεωρείται δεδομένη και κρίνεται απαραίτητο να αποτελέσουν αντικείμενα προς μελέτη σε μελλοντικές προσπάθειες που αφορούν την παραγωγή προβλέψεων για την ΟΤΣ.

Ένα από αυτά τα εγχειρήματα είναι η χρησιμοποίηση μοντέλων παλινδρόμησης (γραμμικής ή μη) κατά τη διαδικασία της προσομοίωσης για την παραγωγή προβλέψεων. Επίσης, ένας καθοριστικός παράγοντας για την βελτίωση της ακρίβειας των προβλέψεων είναι η απαλοιφή της διπλής εποχικότητας, η οποία χαρακτηρίζει τη χρονοσειρά της ΟΤΣ. Για το λόγο αυτό, προτείνονται ορισμένες μέθοδοι αυξημένης πολυπλοκότητας σε σχέση με εκείνη που αναπτύχθηκε στα πλαίσια αυτής της Διπλωματικής, οι οποίες είναι πιθανό να συμβάλλουν σημαντικά στην αύξηση της ακρίβειας τον παραγόμενων προβλέψεων.

9.1 Γραμμική Παλινδρόμηση

Οι μέθοδοι παλινδρόμησης, γραμμικής ή μη, χρησιμοποιούνται ευρέως στην προσπάθεια πρόβλεψης της ΟΤΣ, και κατέχουν σημαντικό κομμάτι στη βιβλιογραφία για το συγκεκριμένο ζήτημα. Στα πλαίσια της παρούσας διπλωματικής, δεν υπήρξε η δυνατότητα να παρουσιαστεί εκτενώς η συγκεκριμένη κατηγορία μεθόδων πρόβλεψης, εντούτοις υλοποιήθηκαν ορισμένες χρήσιμες μεθοδολογίες, οι οποίες είναι δυνατό να αποτελέσουν μελλοντικό αντικείμενο μελέτης.

Σε πρώτο στάδιο, επιχειρήθηκε η απαραίτητη τροποποίηση της διαδικασίας προσομοίωσης, με τον τρόπο που αυτή παρουσιάστηκε στο Κεφάλαιο 5, βασιζόμενη πάντοτε στη λογική του rolling forecasting, ώστε να είναι δόκιμη η εφαρμογή της πολλαπλής ή μη, γραμμικής παλινδρόμησης. Αρχικά, θεωρήθηκε απαραίτητος ο προσδιορισμός των κατάλληλων ανεξάρτητων μεταβλητών παλινδρόμησης, καθώς το ρόλο της εξαρτημένης μεταβλητής κατέχει η χρονοσειρά της ΟΤΣ.

Η επιλογή των κατάλληλων ανεξάρτητων μεταβλητών θεωρείται δύσκολο εγχείρημα, καθώς είναι μεγάλος ο όγκος των παραμέτρων οι οποίες επηρεάζουν το μέγεθος της ΟΤΣ, και ενδεχομένως στον ίδιο βαθμό. Μερικές από αυτές τις παραμέτρους παρουσιάστηκαν στο Κεφάλαιο 4, καθώς έγινε και μία προσπάθεια κατηγοριοποίησης τους. Εν τέλη, επιλέχθηκαν 4 διαφορετικές παράμετροι, οι οποίες παρουσιάστηκαν αναλυτικά στο Κεφάλαιο 6 και είναι οι εξής:

- Ιστορικές τιμές ζήτησης Φορτίου
- Πρόβλεψη Φορτίου
- Ιστορικές Τιμές Αργού Πετρελαίου
- Ιστορικές Τιμές Φυσικού Αερίου

Είναι σημαντικό να αναφερθεί, πως οι παρατηρήσεις των ιστορικών τιμών αργού πετρελαίου και φυσικού αερίου, βρίσκονται σε μηνιαίο επίπεδο. Για το λόγο αυτό, θεωρείται απαραίτητο να διαχωριστεί η μεθοδολογία της παλινδρόμησης σε 2 κατηγορίες. Στην πρώτη κατηγορία, οι χρονοσειρές που χρησιμοποιούνται κατά τη διαδικασία της προσομοίωσης, βρίσκονται σε ωριαίο επίπεδο, ενώ στην δεύτερη σε μηνιαίο.

Πέρα από αυτή την ειδοποιό διαφορά, η μεθοδολογία που ακολουθείται και στις 2 κατηγορίες είναι πανομοιότυπη. Προτού ξεκινήσει η επαναληπτική διαδικασία του rolling, καθορίζεται το μέγεθος των in-sample δεδομένων, που ισούται με το 60% του συνόλου των παρατηρήσεων της χρονοσειράς της ΟΤΣ, όπως ακριβώς και στη διαδικασία προσομοίωσης που περιγράφηκε στο Κεφάλαιο 5.

Στη συνέχεια, αφού καθοριστεί ο ορίζοντας πρόβλεψης αλλά και το βήμα ανανέωσης του rolling, ανάλογα σε ποια κατηγορία αναφερόμαστε (μηνιαία ή ωριαία δεδομένα), ξεκινά η επαναληπτική διαδικασία. Αξίζει να τονιστεί, για την κατηγορία των ωριαίων δεδομένων, πως οι χρονοσειρές της ΟΤΣ, της Ζήτησης Φορτίου αλλά και της Πρόβλεψης Φορτίου, περνούν από το στάδιο της προεπεξεργασίας, το οποίο περιλαμβάνει την απαλοιφή των μηδενικών, σύμφωνα με τη γνωστή μεθοδολογία που αναπτύχθηκε στα προηγούμενα κεφάλαια.

Το επόμενο βήμα, είναι ο υπολογισμός του διανύσματος των συντελεστών παλινδρόμησης, το οποίο υπολογίζεται μέσω τυποποιημένης συνάρτησης του προγράμματος Matlab, χρησιμοποιώντας τις ανεξάρτητες και εξαρτημένες μεταβλητές. Σχετικά με τις ανεξάρτητες μεταβλητές, πρέπει να τονιστεί πως χρησιμοποιείται κάθε φορά μόνο μία εκ των δύο χρονοσειρών που αφορούν τη ζήτηση φορτίου, είτε οι πραγματικές τιμές είτε οι προβλέψεις.

Στη συνέχεια, ανανεώνονται τα δεδομένα του διανύσματος ανεξάρτητων μεταβλητών ανάλογα με τον ορίζοντα πρόβλεψης, και η τελική πρόβλεψης παράγεται μέσω του πολλαπλασιασμού του διανύσματος συντελεστών με εκείνο των ανανεωμένων ανεξάρτητων μεταβλητών.

Για την αξιολόγηση των προβλέψεων, κατά τη διάρκεια της επαναληπτικής διαδικασίας, αποθηκεύονται οι πραγματικές τιμές των παρατηρήσεων της χρονοσειράς της ΟΤΣ, καθώς και οι παραγόμενες προβλέψεις, ακριβώς όπως και στην διαδικασία προσομοίωσης που χρησιμοποιήθηκε για την παραγωγή των αποτελεσμάτων που παρουσιάστηκαν στο Κεφάλαιο 7. Με αυτή την ενέργεια κλείνει ο κύκλος της επαναληπτικής διαδικασίας του rolling.

Στο σημείο αυτό, είναι απαραίτητο να τονιστεί πως η όλη μεθοδολογία που αναπτύχθηκε παραπάνω, θεωρείται απλή προσέγγιση της διαδικασίας της πολλαπλής γραμμικής παλινδρόμησης και παρουσιάζει ικανό περιθώριο βελτίωσης.

Παρά το γεγονός ότι η παραγωγή προβλέψεων μέσω της διαδικασίας της πολλαπλής γραμμικής παλινδρόμησης, διατηρήθηκε εκτός του ορίων μελέτης της παρούσας Διπλωματικής, το συγκεκριμένο ζήτημα παρουσιάζει έντονο ενδιαφέρον και θεωρείται ιδιαίτερα χρήσιμο να διατυπωθούν ορισμένες ενέργειες που είναι δυνατό να εφαρμοστούν σε μελλοντικές έρευνες με κατεύθυνση το συγκεκριμένο πρόβλημα. Οι ενέργειες αυτές αφορούν τη βελτιστοποίηση της μεθοδολογίας που παρουσιάστηκε στην παρούσα ενότητα και είναι οι εξής:

i. Αντιμετώπιση εποχικότητας στη χρονοσειρά της ΟΤΣ, αλλά και στις χρονοσειρές που επιλέγονται ως ανεξάρτητες μεταβλητές.

ii. Βέλτιστη επιλογή ανεξάρτητων μεταβλητών, μέσω της μελέτης βασικών στατιστικών δεικτών.

iii. Βέλτιστη επιλογή επιπέδου δεδομένων (ώρα-ημέρα-μήνας), για την παραγωγή ακριβέστερων αποτελεσμάτων.

iv. Βέλτιστη επιλογή ορίζοντα πρόβλεψης.

v. Εφαρμογή απλής έναντι πολλαπλής γραμμικής παλινδρόμησης

Από τις παραπάνω ενέργειες που παρουσιάστηκαν, η αντιμετώπιση της εποχικότητας αναμένεται να έχει το μεγαλύτερο αντίκτυπο στην βελτίωση της ακρίβειας των αποτελεσμάτων. Η υπόθεση αυτή μπορεί κάλλιστα να βασιστεί στα συμπεράσματα που εξάχθηκαν μέσω ανάλυσης των αποτελεσμάτων των μοντέλων χρονοσειρών που χρησιμοποιήθηκαν κατά τη διαδικασία της προσομοίωσης και παρουσιάστηκαν στην Ενότητα 8.3.

Επίσης, σημαντικό ρόλο αναμένεται να παίξει και η βέλτιστη επιλογή του ορίζοντα πρόβλεψης, καθώς όπως αποδείχθηκε στα προηγούμενα κεφάλαια, πρόκειται για μία εκ των σημαντικότερων παραμέτρων που επηρεάζουν την ακρίβεια των παραγόμενων προβλέψεων. Επιπλέον, μεγάλο ενδιαφέρον παρουσιάζει η περίπτωση της απλής γραμμικής παλινδρόμησης λαμβάνοντας υπόψη μόνο μία ανεξάρτητη μεταβλητή κάθε φορά, καθώς με τον τρόπο αυτό είναι δυνατό να εξαχθούν χρήσιμα συμπεράσματα για την επίδραση που μπορεί να έχει κάθε ανεξάρτητη μεταβλητή ξεχωριστά στη χρονοσειρά της ΟΤΣ. Η απλή γραμμική παλινδρόμηση, χρησιμοποιήθηκε για την παραγωγή αποτελεσμάτων στο Κεφάλαιο 7, εντούτοις λήφθηκε ως ανεξάρτητη μεταβλητή ο χρόνος και όχι κάποια από τις μεταβλητές που παρουσιάστηκαν σε αυτή την ενότητα.

Τέλος, η βέλτιστη επιλογή του επιπέδου των δεδομένων αλλά και των παραμέτρων εκείνων οι οποίες επηρεάζουν σε μεγαλύτερο βαθμό το μέγεθος της ΟΤΣ στην Ελληνική Αγορά Ενέργειας, όπως για παράδειγμα είναι η τιμή του λιγνίτη, είναι πιθανό να βελτιώσουν κατά πολύ την απόδοση της μεθόδου της πολλαπλής γραμμικής παλινδρόμησης.

Η μεθοδολογία της διαδικασίας προσομοίωσης που περιγράφηκε για την εφαρμογή της πολλαπλής γραμμικής παλινδρόμησης, και για τις 2 διαφορετικές κατηγορίες επιπέδου δεδομένων, παρουσιάζεται στο Παράρτημα Matlab.

9.2 Διπλή Εποχικότητα

Σε αρκετές περιπτώσεις, οι χρονοσειρές ενδέχεται να χαρακτηρίζονται από πολλαπλά εποχικά πρότυπα, διαφορετικού μήκους το καθένα. Μια τέτοια περίπτωση είναι η χρονοσειρά της ΟΤΣ, στην οποία είναι ευδιάκριτη η συνύπαρξη δύο εποχικών προτύπων, ενός εβδομαδιαίου και ενός ημερήσιου.

Προκειμένου να αντιμετωπιστούν οι συγκεκριμένες συνιστώσες εποχικότητας, εφαρμόστηκε μία απλοϊκή μέθοδος, χαμηλής υπολογιστικής πολυπλοκότητας, η οποία βασίστηκε στην απαλοιφή κάθε συνιστώσας εποχικότητας μέσω του προσδιορισμού 24 εποχικών δεικτών για την περίπτωση του ημερήσιου προτύπου, και 168 δεικτών για την περίπτωση του εβδομαδιαίου προτύπου, αμφότεροι αναφερόμενοι σε ωριαίο επίπεδο δεδομένων.

Η συγκεκριμένη προσέγγιση της διπλής εποχικότητας, δεν είναι τυχαία, αλλά βασίζεται στη λογική της μεθόδου αντιμετώπισης διπλής εποχικότητας μέσω εκθετικής εξομάλυνσης (DS exponential smoothing), η οποία αναπτύχθηκε από τον *Taylor* (2003). Ο βασικός άξονας γύρω από τον οποίο κινείται η συγκεκριμένη μέθοδος, είναι η υπόθεση πως ένας κύκλος εποχικότητα είναι φωλιασμένος μέσα σε έναν άλλο μεγαλύτερο κύκλο. Επομένως, η απλή μεθοδολογία που παρουσιάστηκε στο Κεφάλαιο 7, για την προσέγγιση της διπλής εποχικότητας, βασίστηκε σε αυτή τη θεμελιώδη παραδοχή.

Στα πλαίσια της παρούσας διπλωματικής, αποφεύχθηκε η μελέτη ιδιαίτερα πολύπλοκων μεθόδων για την αντιμετώπιση της διπλής εποχικότητας, καθώς ένα τέτοιο εγχείρημα βρισκόταν έξω από τη σφαίρα ενδιαφέροντος της συγκεκριμένης έρευνας. Εντούτοις, η χρησιμοποίηση μεθόδων όπως του Taylor, είναι πιθανό να οδηγούσε σε σημαντική βελτίωση της ακρίβειας των παραγόμενων προβλέψεων, και θεωρείται σημαντικό εγχείρημα προς μελλοντική έρευνα.

Η μέθοδος διπλής εποχικότητας του Taylor, βασίζεται στην παραδοχή ύπαρξης δύο εποχικών προτύπων, όπου το ένα πρότυπο περιέχεται μέσα στο άλλο. Κοινώς, αναφέρεται σε ένα ημερήσιο πρότυπο, το οποίο επαναλαμβάνεται στα πλαίσια ενός μεγαλύτερου κύκλου δηλαδή ενός εβδομαδιαίου προτύπου.

Η συγκεκριμένη παραδοχή παρουσιάζει μία σημαντική αδυναμία. Το ημερήσιο πρότυπο θεωρείται ότι επαναλαμβάνεται για κάθε ημέρα της εβδομάδας, αγνοώντας με αυτό τον τρόπο τις υπαρκτές διαφορές που παρατηρούνται ανάμεσα στις καθημερινές και στα Σαββατοκύριακα.

Στην ουσία δεν πρόκειται για παρθενογένεση, αλλά για τροποποίηση της κλασικής μεθόδου Holt-Winters εκθετικής εξομάλυνσης, έτσι ώστε αυτή να μπορεί να επεξεργαστεί ταυτόχρονα, περισσότερα του ενός εποχικά πρότυπα. Η κλασική μέθοδος Holt-Winters, αποτελεί έναν από τους πιο διαδεδομένους τρόπους προσέγγισης της εποχικότητας, κατά τη διαδικασία πρόβλεψης χρονοσειρών.

Εντούτοις, όπως αναφέρθηκε, μπορεί να συλλάβει μόνο ένα εποχικό πρότυπο, σε αντίθεση για παράδειγμα με το πολλαπλασιαστικό μοντέλο ARIMA διπλής εποχικότητας, το οποίο αποτελεί παράλληλα τον κύριο ανταγωνιστή της μεθόδου Holt-Winters στην αντιμετώπιση της εποχικότητας, και επιπλέον μπορεί να μετατραπεί εύκολα για να επεξεργάζεται ακόμη και τρία πρότυπα εποχικότητας ταυτόχρονα. Φυσικά αυτό δεν ενδιαφέρει στην περίπτωση όπου ο ορίζοντας των προβλέψεων καθορίζεται σύμφωνα με τη λογική του day ahead forecasting, καθώς το ετήσιο πρότυπο δεν επιδρά σημαντικά σε τέτοιου είδους περιπτώσεις.

Η προσέγγιση της μεθόδου Holt-Winters που παρουσιάζεται από τον *Taylor* (2003), αφορά την τροποποίηση της συγκεκριμένης πολλαπλασιαστικής μεθόδου εκθετικής εξομάλυνσης, ώστε να μπορεί να αντιμετωπίσει περισσότερα του ενός εποχικά πρότυπα, εγχείρημα το οποίο δεν είχε εξεταστεί μέχρι τότε στη σχετική βιβλιογραφία. Συγκεκριμένα, η μετατροπή αυτή αφορά στην ικανότητα της μεθόδου να μπορεί να αντιμετωπίσει δύο εποχικά πρότυπα, ένα ημερήσιο και ένα εβδομαδιαίο, εισάγοντας δύο ομάδες εποχικών δεικτών, μία για κάθε συνιστώσα εποχικότητας. Η συγκεκριμένη μεθοδολογία, είναι δυνατό να επεκταθεί για την αντιμετώπιση τριών ή και περισσότερων εποχικών προτύπων, εισάγοντας έναν επιπλέον εποχικό δείκτη και μία αντίστοιχη εξίσωση εξομάλυνσης, για κάθε επιπλέον πρότυπο εποχικότητας.

Τα αποτελέσματα του συγκεκριμένου εγχειρήματος, θεωρούνται θετικά, καθώς η τροποποιημένη μέθοδος Holt-Winters για διπλή εποχικότητα, παρουσίασε σημαντικά βελτιωμένα αποτελέσματα σε σύγκριση με την κλασική μέθοδο Holt-Winters, η οποία εφαρμόστηκε για ημερήσιο και για εβδομαδιαίο πρότυπο ξεχωριστά. Μάλιστα, η εφαρμογή της κλασικής μεθόδου για ημερήσια εποχικότητα, παρουσίασε μακράν τα χειρότερα αποτελέσματα. Εντούτοις, ακόμη και η τροποποιημένη μέθοδος του Taylor δεν μπόρεσε να σταθεί ικανοποιητικά απέναντι στην μέθοδο διπλής εποχικότητας ARIMA, η οποία αποτελεί την κύρια εναλλακτική λύση αντιμετώπισης της εποχικότητας έναντι στις μεθόδους εξομάλυνσης.

Προκειμένου να μειωθεί το χάσμα μεταξύ της απόδοσης της τροποποιημένης μεθόδου Holt-Winters και της μεθόδου ARIMA, ο Taylor επιχείρησε να ρυθμίσει την αυτοσυσχέτιση (AR) των υπολοίπων των σφαλμάτων στη μέθοδο Holt-Winters, καθώς μελετώντας τα σφάλματα για 1 βήμα μπροστά, παρατηρήθηκε σημαντική πρώτης τάξης αυτοσυσχέτιση, και η αντιμετώπισή της είναι πιθανό να οδηγούσε σε βελτιωμένα αποτελέσματα. Η συγκεκριμένη ρύθμιση, δεν κρίθηκε απαραίτητη για το μοντέλο ARIMA, καθώς εκείνο δεν παρουσίαζε αυτοσυσχέτιση στα υπόλοιπα των σφαλμάτων.

Με αυτή την τελική τροποποίηση, η μέθοδος διπλής εποχικότητας Holt-Winters, παρουσίασε σημαντική βελτίωση, όσον αφορά το στατιστικό δείκτη MAPE, ο οποίος τέθηκε ως κριτήριο αξιολόγησης των παραγόμενων προβλέψεων, και επιπλέον ξεπέρασε ακόμη και τη μέθοδο ARIMA, ενώ και το ίδιο συνέβη και με την κλασική μέθοδο Holt-Winters για εποχικότητα εβδομάδας.

Εντούτοις, ενώ το γεγονός ότι η τροποποιημένη μέθοδος Holt-Winters παρουσιάζει καλύτερα αποτελέσματα σε σύγκριση με την κλασική θεωρείται απόλυτα φυσιολογικό, καθώς η κλασική μέθοδος αδυνατεί να συλλάβει περισσότερα του ενός εποχικά πρότυπα, ο λόγος για τον οποίο ξεπερνάει τη μέθοδο ARIMA παραμένει λιγότερο ευδιάκριτος.

Αυτό είναι πιθανό να οφείλεται στο γεγονός ότι η μέθοδος ARIMA λειτουργεί καλύτερα σε χρονοσειρές όπου παρουσιάζεται βραχυπρόθεσμη συσχέτιση, ενώ παρουσιάζουν αδυναμία προσαρμογής σε χρονοσειρές με τάση και εποχικές διακυμάνσεις. Αν συνυπολογιστεί και η φύση των δεδομένων που χρησιμοποιούνται για την παραγωγή προβλέψεων, που στην προκειμένη περίπτωση είναι ημιωριαίες τιμές ζήτησης φορτίου, τότε μπορεί να εξηγηθεί η επικράτηση της μεθόδου διπλής εποχικότητας Holt-Winters με αυτοσυσχέτιση υπολοίπων (DS Holt-Winters with AR(1) adjustment).

Τελικά, έχοντας πλέον μία σφαιρική εικόνα της μεθόδου που παρουσίασε ο Taylor για την προσέγγιση της διπλής εποχικότητας, είναι δυνατό να ειπωθεί πως θα μπορούσε να εφαρμοστεί σε ανάλογες μελλοντικές έρευνες σχετικές με την πρόβλεψη της ΟΤΣ στην Ελληνική Αγορά Ενέργειας. Η υπόθεση αυτή ενισχύεται από το γεγονός πως η μεθοδολογία που εφαρμόστηκε στην παρούσα Διπλωματική για την αντιμετώπιση της διπλής εποχικότητας, η οποία αποτελεί μία απλοϊκή προσέγγιση της μεθόδου που παρουσίασε ο Taylor, παρουσίασε τα βέλτιστα αποτελέσματα σχετικά με την ακρίβεια των προβλέψεων και το στατιστικό δείκτη sMAPE.

Βιβλιογραφία

- Andrianesis P., Biskas P. & Liberopoulos G. (2011) 'An overview of Greece's wholesale electricity market with emphasis on ancillary services', *Electric Power Systems Research*, Vol. 81, pp. 1631-1642.

- Che Jinxing & Wang Jianzhou (2010) 'Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average modeling', *International Journal of Energy Conversion and Management*, Vol. 51, pp. 1911-1917.

- Deepak S. & Swarup K.S. (2011) 'Electricity price forecasting using artificial neural networks', *International Journal of Electrical Power and Energy Systems*, Vol. 33, pp. 550-555.

- Gould P.G., Koehler A.B., Ord J.K., Snyder R.D., Hyndman R.J. & Farshid Vahid-Araghi (2008) 'Forecasting time series with multiple seasonal patterns', *European Journal of Operational Research*, Vol. 191, pp. 207-222.

- Hamm G. & Borison A. (2006) 'Forecasting Long-Run Electricity Prices', *The electricity Journal*, Vol. 17, Issue 7, pp. 47-57

- Karakatsani N.V. & Burn D.W. (2008) 'Forecasting electricity prices: The impact of fundamentals and time-varying coefficients', *International Journal of Forecasting*, Vol. 24, pp. 764-785.

- Karsaz A., Rajabi Mashhadi H. & Mirsalehi M.M. (2010) 'Market clearing price and load forecasting using cooperative co-evolutionary approach', *International Journal of Electrical Power and Energy Systems*, Vol. 32, pp. 408-415.

- Pousinho H.M.I., Mendes V.M.F. & Catalão J.P.S. (2012) 'Short-term electricity prices forecasting in a competitive market by a hybrid PSO-ANFIS approach', *International Journal of Electrical Power and Energy Systems*, Vol. 39, pp. 29-35.

- Sanjeev Kumar Aggarwal, Lalit Mohan Saini & Ashwani Kumar (2009) 'Electricity price forecasting in deregulated markets: A review and evaluation', *International Journal of Electrical Power and Energy Systems*, Vol. 31, pp. 13-22.

- Shafie khah M., Parsa Moghaddam M. & Sheikh-El-Eslami M.K. (2011) 'Price forecasting of day-ahead electricity markets using a hybrid forecast method', *International Journal of Energy Conversion and Management*, Vol. 52, pp. 2165-2169.

- Taylor J.W. (2003) 'Short-term electricity demand forecasting using double seasonal exponential smoothing', *Journal of the Operational Research Society*, Vol. 54, pp. 799-805.

- Torghaban S.S. (2010) 'Medium-Term Electricity Market Price Forecasting: A Data-driven Approach', Thesis for the Degree of Master of Science, Chalmers University of Technology, Goteborg, Sweden.

- Vespucci M.T., Allevi E., Gnudi A. & Innorta M. (2010) 'Cournot equilibria in oligopolistic electricity market', *IMA Journal of Management Mathematics*, Vol. 21, pp. 183-193.

- ΜΕΤΑΦΟΡΑ ΚΑΙ ΔΙΑΝΟΜΗ ΗΛΕΚΤΡΙΚΗΣ ΕΝΕΡΓΕΙΑΣ–ΗΛΕΚΤΡΙΚΑ ΔΙΚΤΥΑ, Συνέδριο 8 – 10 Μαρτίου 2010, Εισήγηση Μόνιμης Επιτροπής Ενέργειας ΤΕΕ, Τεχνικό Επιμελητήριο Ελλάδας, Αθήνα.

- Κάπρος Π. (2003), Σημειώσεις Οικονομικής Ανάλυσης – Μέρος Γ', Ε.Μ.Π., Αθήνα.

- Πετρόπουλος Φ. & Ασημακόπουλος Β. (2011), *Επιχειρησιακές Προβλέψεις*, Αθήνα: Εκδόσεις Συμμετρία.

- Configuration of System's Marginal Price. Obtained through the Internet: http://www.rae.gr [accessed 19/3/2012].

- Egcheiridio_Agoras_version_2.0_2010.09.06. Obtained through the Internet: http://www.desmie.gr/ [accessed 19/3/2012].

Για τη συγγραφή της παρούσας Διπλωματικής, πληροφορίες αντλήθηκαν και από τους παρακάτω διαδικτυακούς τόπους:

- http://www.investingreece.gov.gr
- http://www.statbank.gr
- http://www.indexmundi.com
- http://www.economagic.com
- http://www.cashfocus.com

ΠΑΡΑΡΤΗΜΑ

Matlab

1.Δείκτες Σφάλματος

1.1 Μέσο Τετραγωνικό Σφάλμα (MSE)

```
function [ error ] = mse( x,y )

    error = mean((x-y).^2);

end
```

1.2 Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE)

```
function [ error ] = mape( x , y )

error = mean(abs((x-y)./x))*100;

end
```

1.3 Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (sMAPE)

```
function [ error ] = smape( x , y )

error = mean(abs(2*(x-y)./(x+y)))*100;

end
```

2. Μοντέλο Αναφοράς Naive

```
function [ y ] = naive(x, hor )

N = length(x);
y(1) = x(1);
for t = 2:N
    y(t) = x(t-1);
end

y ((length(x) +1) : (length(x) + hor)) = x(t);

end
```

3. Μοντέλα Εκθετικής Εξομάλυνσης

3.1 Μοντέλο SES

```
function [ y ] = ses( x, a, S0, hor)

y(1) = S0;
S = S0+a*(x(1)-y(1));

for t = 2:length(x)
    y(t) = S;
    e = x(t)-y(t);
    S = S+a*e;
end

y(length(x)+1:length(x)+ hor) = S;

end
```

3.2 Μοντέλο SES (Γραμμική Βελτιστοποίηση)

```
function [ y best_a ] = seslinear( x, rangea, S0, hor )

y = ses(x,rangea(1),S0,hor);
best_mse = mse(x,y(1:length(x)));
best_a = rangea(1);

for I = rangea(2:end)
   y = ses(x,i,S0,hor);
   current_mse = mse(x,y(1:length(x)));
   if current_mse < best_mse
      best_mse = current_mse;
      best_a = i;
   end
end

y = ses(x,best_a,S0,hor);

end
```

3.3 Μοντέλο SES (Μη Γραμμική Βελτιστοποίηση)

```
function [ y best_a ] = sesnonlinear( x, S0, M, hor )

N = length(x);
a= [0.33 0.667];
dV = a(1)/2;
y = ses(x,a(1),S0,hor);
best_mse = mse(x,y(1:N));
best_a =a (1);

for i = 1:M
   for j = 1:2
         y = ses(x,a(j),S0,hor);
         current_mse = mse(x,y(1:N));
         if current_mse < best_mse
            best_mse = current_mse;
            best_a = a(j);
         end
   end

   a(1) = best_a-dV;
   a(2) = best_a+dV;
   dV = dV/2;
end
y = ses(x,best_a,S0,hor);
end
```

3.4 Μοντέλο Holt

```
function [ y ] = holt( x, a, b, S0, T0, hor )

N = length(x);
y(1) = S0 + T0;
e = x(1)-y(1);
S = S0 + T0 +a*e;
T = T0 + b*e;

for t = 2:N
    y(t) = S + T;
    e = x(t)- y(t);
    S = S+T+a*e;
    T = T+b*e;
end

for m = 1:hor
    y(N+m) = S+m*T;
end

end
```

3.5 Μοντέλο Holt (Γραμμική Βελτιστοποίηση)

```
function [ f best_a best_b] = holtlinear( x, rangea, rangeb, S0, T0, hor )

N = length(x);
f = holt(x,rangea(1),rangeb(1),S0,T0,hor);
best_mse = mse(x,f(1:N));
best_a = rangea(1);
best_b = rangeb(1);

for i = rangea(1:end)
    for j = rangeb(1:end)
        f = holt(x,i,j,S0,T0,hor);
        current_mse = mse(x,f(1:N));
        if current_mse < best_mse
            best_mse = current_mse;
            best_a = i;
            best_b = j;
        end
    end
end

f = holt(x,best_a,best_b,S0,T0,hor);
a = best_a;
b = best_b;

end
```

3.6 Μοντέλο Holt (Μη Γραμμική Βελτιστοποίηση)

```
function [ y best_a best_b ] = holtnonlinear( x, S0, T0, M, hor )

N = length(x);
a = [0.33 0.667];
b = [0.33 0.667];
dV = a(1)/2;

y = holt(x,a(1),b(1),S0,T0,hor);
best_mse = mse(x,y(1:N));
best_a = a(1);
best_b = b(1);

for i = 1:M
    for j = 1:2
        for k = 1:2
            y = holt(x,a(j),b(k),S0,T0,hor);
            current_mse = mse(x,y(1:N));
            if current_mse < best_mse
                best_mse = current_mse;
                best_a = a(j);
                best_b = b(k);
            end
        end
    end

    a(1) = best_a-dV;
    a(2) = best_a+dV;
    b(1) = best_b-dV;
    b(2) = best_b+dV;
    dV = dV/2;
end

y = holt(x,best_a,best_b,S0,T0,hor);

end
```

3.7 Μοντέλο Damped

```
function [ y ] = damped( x, a, b, f, S0, T0, hor )

N = length(x);
y(1) = S0 + f*T0;
e = x(1)-y(1);
S = S0 + f*T0 +a*e;
T = f*T0 + b*e;

for t = 2:N
    y(t) = S + f*T;
    e = x(t)- y(t);
    S = S+f*T+a*e;
    T = f*T+b*e;
end
y(N+1) = S+f*T;

for m = 2:hor
    y(N+m) = y(N+m-1)+(f^m)*T;
end

end
```

3.8 Μοντέλο Damped (Γραμμική Βελτιστοποίηση)

```
function [ y best_a best_b best_f ] = dampedlinear( x, rangea, rangeb, rangef,
S0, T0, hor )

N = length(x);
Y = damped( x, rangea(1), rangeb(1), rangef(1), S0, T0, hor );

best_mse = mse(x,y(1:N));
best_a = rangea(1);
best_b = rangeb(1);
best_f = rangef(1);

for i = rangea(1:end)
    for j = rangeb(1:end)
        for k = rangef(1:end)
            y = damped(x,i,j,k,S0,T0,hor);
            current_mse = mse(x,y(1:N));
            if current_mse < best_mse
            best_mse = current_mse;
            best_a = i;
            best_b = j;
            best_f = k;
            end
        end
    end
end

y = damped(x,best_a,best_b,best_f,S0,T0,hor);
a = best_a;
b = best_b;
f = best_f;

end
```

3.9 Μοντέλο Damped (Μη Γραμμική Βελτιστοποίηση)

```
function [ y best_a best_b best_f ] = dampednonlinear( x, S0, T0, M, hor )

N = length(x);
a = [0.33 0.667];
b = [0.33 0.667];
f = [0.33 0.667];
dV = a(1)/2;

y = damped(x,a(1),b(1),f(1),S0,T0,hor);
best_mse = mse(x,y(1:N));
best_a = a(1);
best_b = b(1);
best_f = f(1);

for i = 1:M
    for j = 1:2
        for k = 1:2
            for l = 1:2
                y = damped(x,a(j),b(k),f(l),S0,T0,hor);
                current_mse = mse(x,y(1:N));
                if current_mse < best_mse
                    best_mse = current_mse;
                    best_a = a(j);
                    best_b = b(k);
                    best_f = f(l);
                end
            end
        end
    end
    a(1) = best_a-dV;
    a(2) = best_a+dV;
    b(1) = best_b-dV;
    b(2) = best_b+dV;
    f(1) = best_f-dV;
    f(2) = best_f+dV;
    dV = dV/2;
end

y = damped(x,best_a,best_b,best_f,S0,T0,hor);

end
```

4.1 Γραμμική Παλινδρόμηση (LR)

```
function [ S0 T0 ] = lr( x )

t = 1:length(x);
p = polyfit(t,x,1);
T0 = p(1);
S0 = p(2);

end
```

4.2 Απλή Γραμμική Παλινδρόμηση (LRL)

```
function [ y ] = lrl( x,hor )

   [S0 T0] = lr(x);
   t = 1:length(x)+hor;
   y = S0+T0*t;

end
```

5. Μοντέλο Theta (κλασική μέθοδος Theta)

```
function [ y ] = theta( x , hor )

N = length(x);

f0 = rl(x,hor);
e = x-f0(1:N);

y2 = f0(1:N)+2*e;
S0 = lr(y2);
f2 = sesnonlinear(y2,S0,6,hor);
y = 0.5*(f0+f2);

end
```

6. Βοηθητικές Συναρτήσεις

6.1 Συνάθροιση (Aggregate)

```
function [ y, offset ] = aggregate( x, L )
%Προκαλεί χρονική συνάθροιση των δεδομένων

                    --------------- Είσοδοι ---------------
%x:       διάνυσμα αρχικής χρονοσειράς
%L:       επίπεδο συνάθροισης

                    --------------- Έξοδοι ---------------
%y:      διάνυσμα συναθροισμένης χρονοσειράς
%offset: πλήθος παρατηρήσεων της αρχικής συμπεριφοράς που δεν
%        θα συμμετάσχουν στο μοντέλο ADIDA

    N=length(x);
    %αποκοπή πρώτων παρατηρήσεων που πλεονάζουν
    offset = mod(N,L);
    x = x(1+offset:end);
    %υπολογισμός αθροισμάτων
    y = filter(ones(1,L),1,x);
    %κρατάμε κάθε L-στό άθροισμα
    y = downsample(y,L, L-1);

end
```

6.2 Διάσπαση (Disaggregate)

```
function [ y ] = disaggregate( x, L, w, offset )
%Προκαλεί χρονική συνάθροιση των δεδομένων

                              --------------- Είσοδοι ---------------
%x:      διάνυσμα συναθροισμένης χρονοσειράς
%L:      επίπεδο συνάθροισης
%w:      διάνυσμα βαρών διάσπασης
%offset: πλήθος αχρησιμοποίητων αρχικών παρατηρήσεων

                              --------------- Έξοδοι ---------------
%y:      διάνυσμα διασπασμένης χρονοσειράς

   x = upsample(x,L);
   y = filter(w,1,x);
   y = [ones(1,offset)*y(1) y];

end
```

6.3 Απαλοιφή Μηδενικών Τιμών

```
function [ in ] = fixzero( in, v )

k = find(in<=v);
i = 1;

while i <= length(k)
   x1 = k(i)-1;
   y1 = in(x1);

   j = 1;
   if i+j <= length(k)
      while k(i+j)-k(i) == j
         j = j+1;
         if i+j > length(k)
            break;
         end
      end
      j = j-1;
   else
         j = 0;
   end

   x2 = k(i+j)+1;
   y2 = in(x2);
   a = (y2-y1)/(x2-x1);
   b = y2-a*x2;
   for l = k(i):k(i+j)
      in(l) = a*l+b;
   end
   i = i+j+1;
end

end
```

6.4 Προσδιορισμός Εποχικών Δεικτών

```
function [SI] = getSI( x, seas)

   if mod(seas,2) == 0
      offset = ceil(seas/2);
      d = filter([0.5 ones(1,seas-1) 0.5]/seas,1,x);
      d = d(1+seas:end);
   else
      offset = floor(seas/2);
      d = filter(ones(1,seas)/seas,1,x);
      d = d(seas:end);
   end

   x = x(1+offset:end-offset);
   d = x./d;
   d = [zeros(1,offset) d'];
   d = buffer(d,seas);
   N = sum(d~=0,2);
   SI = sum(d,2);
   if min(N) >= 4
      SI = SI-max(d,[],2)-min(d,[],2);
      N = N-2;
   end
   SI = SI./N;
   SI = SI/sum(SI)*seas;
end
```

6.5 Αποεποχικοποίηση Χρονοσειράς

```
function [ y ] = deseasonalize( x, SI, offset )
%Αφαιρεί την εποχικότητα από μία χρονοσειρά

                              --------------- Είσοδοι ---------------
%x:      διάνυσμα εποχικοποιημένης χρονοσειράς
%SI:     διάνυσμα δεικτών εποχικότητας
%offset: μετατόπιση πρώτου δείκτη σε σχέση
          με την πρώτη περίοδο της χρονοσειράς

                              --------------- Έξοδοι ---------------
%y: διάνυσμα αποεποχικοποιημένης χρονοσειράς

   seasonality = length(SI);
   y = x;

   for i = 1:length(x)
      y(i) = x(i)/SI(mod(i-1+offset,seasonality)+1);
   end

end
```

6.6 Επανεποχικοποίηση Χρονοσειράς

```
function [ y ] = reseasonalize(  x, SI, offset )
%Εφαρμόζει εποχικότητα σε χρονοσειρά

                              --------------- Είσοδοι ---------------
%x:      διάνυσμα αποεποχικοποιημένης χρονοσειράς
%SI:     διάνυσμα δεικτών εποχικότητας
%offset: μετατόπιση πρώτου δείκτη σε σχέση
          με την πρώτη περίοδο της χρονοσειράς

                              --------------- Έξοδοι ---------------
%y: διάνυσμα επανεποχικοποιημένης χρονοσειράς

   seasonality = length(SI);
   y = x;

   for i = 1:length(x)
      y(i) = x(i)*SI(mod(i-1+offset,seasonality)+1);
   end

end
```

7. Διαδικασία Προσομοίωσης (Rolling)

7.1.1 Προσομοίωση για μοντέλα χρονοσειρών – Δίχως Απαλοιφή Εποχικότητας (no SI)

```
clear
load('ots.mat')                 %φόρτωση χρονοσειράς ΟΤΣ

                   --------------- ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ---------------

ots = fixzero(ots,5);           %απαλοιφή μηδενικών τιμών
ots = ots(14:end);              %πρώτη διαθέσιμη τιμή στις 14:00 η ώρα

L = length(ots);                %μήκος χρονοσειράς για rolling

N = round(0.6*L);               %μέγεθος in-sample δεδομένων
N = N-mod(N,24);                %αποκατάσταση ολίσθησης των δεικτών
                                 εποχικότητας

C=7;                            %αριθμός 24-ωρων για ορίζοντα πρόβλεψης
hor=C*24;                       %ορίζοντας πρόβλεψης για C-24ωρα
step=24;                        %βήμα ανανέωσης δεδομένων rolling

j = 1;
for i= N:step:L-hor
  i/L*100
  z = ots(1:i)';
                   --------------- ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ ---------------

  %y = naive(z,hor);             %μέθοδος Naive

  %y = lrl(z,hor);               %μέθοδος LRL

  %y = theta(z,hor);             %μέθοδος Theta

  [S0 T0]=lr(z);                %υπολογισμός αρχικού επιπέδου και αρχικής τάσης


  [y a] = sesnonlinear(z,S0,6,hor);         %μέθοδοι εξομάλυνσης
  %[y a] = seslinear(z,0.1:0.01:0.9,S0,hor);
  %[y] = ses(z,0.1,S0,hor);


  %[y a b] = holtnonlinear(z,S0,T0,6,hor);
  %[y] = holt(z,0.98,0.0001,S0,T0,hor);

  %[y a b f] = dampednonlinear(z,S0,T0,6,hor);
  %[y] = damped(z,0.99,0.001,0.25,S0,T0,hor);
```

```
--------------- ΜΕΤΕΠΕΞΕΡΓΑΣΙΑ ---------------

    y(find(y>150)) = 150;

    Y(j,1:hor)= ots (i+(1:hor));    %πίνακας με πραγματικές τιμές δεδομένων
    F(j,1:hor) = y(i+(1:hor));     %πίνακας με επανεποχικοποιημένες προβλέψεις

    j = j+1;
end

--------------- ΣΦΑΛΜΑΤΑ ---------------

e = Y-F; %υπολογισμός σφάλματος σε πίνακα: ημέρες_πρόβλεψης*168_ώρες

sAPE = (abs(2*e./(Y+F)))*100; %sAPE

sMAPE = mean(sAPE(1:1:end,:),1);
dsMAPE = aggregate(sMAPE,24)/24;

for k = 1:7                                  %εικονικά σφάλματα
  p = mean(sAPE(1:1:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

for k = 1:7                                  %πραγματικά σφάλματα
  p = mean(sAPE(1:k:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

mean(mean(sAPE,1),2)                         %μέση τιμή δείκτη sMAPE
```

7.1.2 Προσομοίωση για μοντέλα χρονοσειρών – Μονή Εποχικότητα (SI=24)

```
clear
load('ots.mat')                  %φόρτωση χρονοσειράς ΟΤΣ

               --------------- ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ---------------

ots = fixzero(ots,5);            %απαλοιφή μηδενικών τιμών
ots = ots(14:end);               %πρώτη διαθέσιμη τιμή στις 14:00 η ώρα

L = length(ots);                 %μήκος χρονοσειράς για rolling

N = round(0.6*L);                %μέγεθος in-sample δεδομένων
N = N-mod(N,24);                 %αποκατάσταση ολίσθησης των δεικτών
                                 εποχικότητας

C = 7;                           %αριθμός 24-ωρων για ορίζοντα πρόβλεψης
hor = C*24;                      %ορίζοντας πρόβλεψης για C-24ωρα
step = 24;                       %βήμα ανανέωσης δεδομένων rolling

j = 1;
for i = N:step:L-hor
   i/L*100

   %z = ots(1:i)';
   x = ots(1:i)';
              --------------- ΜΟΝΗ ΕΠΟΧΙΚΟΤΗΤΑ ---------------

   delay = i-1*7*24;             %καθυστέρηση με δεδομένα με t (t=1:7) μήνες
   %delay = 1;                   %καθυστέρηση με όλα τα δεδομένα

   SI = getSI(x(delay:i)',24);      %ωριαίοι δείκτες εποχικότητας με SI=24

   z = deseasonalize(x,SI,0);       %αποεποχικοποιημένη χρονοσειρά

              --------------- ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ ---------------

   %y = naive(z,hor);             %μέθοδος Naive

   %y = lrl(z,hor);               %μέθοδος LRL

   %y = theta(z,hor);             %μέθοδος Theta

   [S0 T0]=lr(z);                 %υπολογισμός αρχικού επιπέδου και αρχικής
                                   τάσης

   [y a] = sesnonlinear(z,S0,6,hor);        %μέθοδοι εξομάλυνσης
   %[y a] = seslinear(z,0.1:0.01:0.9,S0,hor);
   %[y] = ses(z,0.1,S0,hor);
```

```
    %[y a b] = holtnonlinear(z,S0,T0,6,hor);
    %[y] = holt(z,0.98,0.0001,S0,T0,hor);

    %[y a b f] = dampednonlinear(z,S0,T0,6,hor);
    %[y] = damped(z,0.99,0.001,0.25,S0,T0,hor);

                    --------------- ΜΕΤΕΠΕΞΕΡΓΑΣΙΑ ---------------

    y(i+(1:hor)) = reseasonalize(y(i+(1:hor)),SI,0);    %επανεποχικοποίηση
                                                         προβλέψεων για ημέρα

    y(find(y>150)) = 150;

    Y(j,1:hor) = ots(i+(1:hor));    %πίνακας με πραγματικές τιμές δεδομένων
    F(j,1:hor) = y(i+(1:hor));     %πίνακας με επανεποχικοποιημένες προβλέψεις

    j = j+1;
end

                    --------------- ΣΦΑΛΜΑΤΑ ---------------

e = Y-F; %υπολογισμός σφάλματος σε πίνακα: ημέρες_πρόβλεψης*168_ώρες

sAPE = (abs(2*e./(Y+F)))*100; %sAPE

sMAPE = mean(sAPE(1:1:end,:),1);
dsMAPE = aggregate(sMAPE,24)/24;

for k = 1:7                                  %εικονικά σφάλματα
  p = mean(sAPE(1:1:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

for k = 1:7                                  %πραγματικά σφάλματα
  p = mean(sAPE(1:k:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

mean(mean(sAPE,1),2)                         %μέση τιμή δείκτη sMAPE
```

7.1.3 Προσομοίωση για μοντέλα χρονοσειρών – Μονή Εποχικότητα (SI=168)

```
clear
load('ots.mat')                %φόρτωση χρονοσειράς ΟΤΣ

                --------------- ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ---------------

ots = fixzero(ots,5);           %απαλοιφή μηδενικών τιμών
ots = ots(14:end);              %πρώτη διαθέσιμη τιμή στις 14:00 η ώρα

L = length(ots);                %μήκος χρονοσειράς για rolling

N = round(0.6*L);               %μέγεθος in-sample δεδομένων
N = N-mod(N,24);                %αποκατάσταση ολίσθησης των δεικτών
                                 εποχικότητας

C = 7;                          %αριθμός 24-ωρων για ορίζοντα πρόβλεψης
hor = C*24;                     %ορίζοντας πρόβλεψης για C-24ωρα
step = 24;                      %βήμα ανανέωσης δεδομένων rolling

j = 1;
for i = N:step:L-hor
   i/L*100

   %z = ots(1:i)';
   x = ots(1:i)';
               --------------- ΜΟΝΗ ΕΠΟΧΙΚΟΤΗΤΑ ---------------

   delay = i-1*7*24;            %καθυστέρηση με δεδομένα με t (t=1:7) μήνες
   %delay = 1;                  %καθυστέρηση με όλα τα δεδομένα

   SI = getSI(x(delay:i)',168);    %εβδομαδιαίοι δείκτες εποχικότητας με SI=168

   z = deseasonalize(x,SI,0);      %αποεποχικοποιημένη χρονοσειρά

               --------------- ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ ---------------

   %y = naive(z,hor);           %μέθοδος Naive

   %y = lrl(z,hor);             %μέθοδος LRL

   %y = theta(z,hor);           %μέθοδος Theta

   [S0 T0]=lr(z);               %υπολογισμός αρχικού επιπέδου και αρχικής
                                  τάσης

   [y a] = sesnonlinear(z,S0,6,hor);         %μέθοδοι εξομάλυνσης
   %[y a] = seslinear(z,0.1:0.01:0.9,S0,hor);
   %[y] = ses(z,0.1,S0,hor);
```

```
    %[y a b] = holtnonlinear(z,S0,T0,6,hor);
    %[y] = holt(z,0.98,0.0001,S0,T0,hor);

    %[y a b f] = dampednonlinear(z,S0,T0,6,hor);
    %[y] = damped(z,0.99,0.001,0.25,S0,T0,hor);

                    --------------- ΜΕΤΕΠΕΞΕΡΓΑΣΙΑ ---------------

    y(i+(1:hor)) = reseasonalize(y(i+(1:hor)),SI,0);    %επανεποχικοποίηση
                                                          προβλέψεων για εβδομάδα

    y(find(y>150)) = 150;

    Y(j,1:hor) = ots(i+(1:hor));    %πίνακας με πραγματικές τιμές δεδομένων
    F(j,1:hor) = y(i+(1:hor));     %πίνακας με επανεποχικοποιημένες προβλέψεις

    j = j+1;
end

                    --------------- ΣΦΑΛΜΑΤΑ ---------------

e = Y-F; %υπολογισμός σφάλματος σε πίνακα: ημέρες_πρόβλεψης*168_ώρες

sAPE = (abs(2*e./(Y+F)))*100; %sAPE

sMAPE = mean(sAPE(1:1:end,:),1);
dsMAPE = aggregate(sMAPE,24)/24;

for k = 1:7                                   %εικονικά σφάλματα
  p = mean(sAPE(1:1:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

for k = 1:7                                   %πραγματικά σφάλματα
  p = mean(sAPE(1:k:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

mean(mean(sAPE,1),2)                          %μέση τιμή δείκτη sMAPE
```

7.1.4 Προσομοίωση για μοντέλα χρονοσειρών – Διπλή Εποχικότητα (SI_1=24, SI_2=168)

```
clear
load('ots.mat')                    %φόρτωση χρονοσειράς ΟΤΣ

                    --------------- ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ---------------

ots = fixzero(ots,5);      %απαλοιφή μηδενικών τιμών
ots = ots(14:end);         %πρώτη διαθέσιμη τιμή στις 14:00 η ώρα

L = length(ots);           %μήκος χρονοσειράς για rolling

N = round(0.6*L);          %μέγεθος in-sample δεδομένων
N = N-mod(N,24);           %αποκατάσταση ολίσθησης των δεικτών εποχικότητας

C = 7;                     %αριθμός 24-ωρων για ορίζοντα πρόβλεψης
hor = C*24;                %ορίζοντας πρόβλεψης για C-24ωρα
step = 24;                 %βήμα ανανέωσης δεδομένων rolling

j = 1;
for i = N:step:L-hor
   i/L*100

   %z = ots(1:i)';
   x = ots(1:i)';

                    --------------- ΔΙΠΛΗ ΕΠΟΧΙΚΟΤΗΤΑ ---------------

   delay = i-1*t*24;              %καθυστέρηση με δεδομένα με t (t=1:7) μήνες
   %delay = 1;                    %καθυστέρηση με όλα τα δεδομένα

   SI = getSI(x(delay:i)',24);       %ωριαίοι δείκτες εποχικότητας με SI=24

   z = deseasonalize(x,SI,0);     %ενδιάμεση αποεποχικοποιημένη χρονοσειρά

   SI2 = getSI(z',168);              %εβδομαδιαίοι δείκτες εποχικότητας
   z = deseasonalize(z,SI2,0);    %τελική αποεποχικοποιημένη χρονοσειρά

                    --------------- ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ ---------------

   %y = naive(z,hor);             %μέθοδος Naive

   %y = lrl(z,hor);               %μέθοδος LRL

   %y = theta(z,hor);             %μέθοδος Theta

   [S0 T0]=lr(z);              %υπολογισμός αρχικού επιπέδου και αρχικής τάσης
```

```
    [y a] = sesnonlinear(z,S0,6,hor);          %μέθοδοι εξομάλυνσης
    %[y a] = seslinear(z,0.1:0.01:0.9,S0,hor);
    %[y] = ses(z,0.1,S0,hor);

    %[y a b] = holtnonlinear(z,S0,T0,6,hor);
    %[y] = holt(z,0.98,0.0001,S0,T0,hor);

    %[y a b f] = dampednonlinear(z,S0,T0,6,hor);
    %[y] = damped(z,0.99,0.001,0.25,S0,T0,hor);

                    --------------- ΜΕΤΕΠΕΞΕΡΓΑΣΙΑ ---------------

     y(i+(1:hor)) = reseasonalize(y(i+(1:hor)),SI,0);        %επανεποχικοποίηση
                                                              προβλέψεων για ημέρα
    y(i+(1:hor)) = reseasonalize(y(i+(1:hor)),SI2,mod(i,168));
                                                            %επανεποχικοποίηση
                                                            προβλέψεων για βδομάδα
    y(find(y>150)) = 150;          %ανώτατο όριο τιμής τα 150€/MWh

    Y(j,1:hor) = ots(i+(1:hor));    %πίνακας με πραγματικές τιμές δεδομένων
    F(j,1:hor) = y(i+(1:hor));     %πίνακας με επανεποχικοποιημένες προβλέψεις

    j = j+1;
end
                    --------------- ΣΦΑΛΜΑΤΑ ---------------

e = Y-F; %υπολογισμός σφάλματος σε πίνακα: ημέρες_πρόβλεψης*168_ώρες

sAPE = (abs(2*e./(Y+F)))*100; %sAPE

sMAPE = mean(sAPE(1:1:end,:),1);
dsMAPE = aggregate(sMAPE,24)/24;      % μέσος ημερήσιος δείκτης sMAPE

for k = 1:7                                 %εικονικά σφάλματα
  p = mean(sAPE(1:1:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

for k = 1:7                                 %πραγματικά σφάλματα
  p = mean(sAPE(1:k:end,:),1);
  horsMAPE(k) = mean(p(1:k*24));
end

mean(mean(sAPE,1),2)                         %μέση τιμή δείκτη sMAPE
```

7.2 Προσομοίωση για πολλαπλή Γραμμική Παλινδρόμηση

```
clear
load('ng_month.mat')
load('oil_month.mat')
load('pqdata.mat')

ots = fixzero(ots,5);                    %απαλοιφή μηδενικών τιμών

%demand = fixzero(demand,5);

demandf = fixzero(demandf,5);

                  --------------- μετατροπή σε μηνιαίες τιμές ---------------

ots = aggregate(ots,24*30);
%demand = aggregate(demand,24*30);
demandf = aggregate(demandf,24*30);

                  --------------- μετατροπή σε ωριαίες τιμές ---------------

%oil = disaggregate(oil,24*30,ones(24,1)/24,0);
%ng_month = disaggregate(ng_month,24*30,ones(24,1)/24,0);

L = length(ots);
N = round(0.6*L);

         --------------- ανεξάρτητες μεταβλητές παλινδρόμησης ----------------

%k1 = demand(1:L);
k2 = demandf(1:L);
k3 = oil(1:L);
k4 = ng_month(1:L);

%hor = 7*24;              %ορίζοντας πρόβλεψης για ωριαία δεδομένα
hor = 12;                 %ορίζοντας πρόβλεψης για μηνιαία δεδομένα
%step = 24;               %βήμα ανανέωσης για ωριαία δεδομένα
step = 12;                %βήμα ανανέωσης για μηνιαία δεδομένα
j = 1;

for i = N:step:L-hor

      ---------------- υπολογισμός του διανύσματος συντελεστών ----------------

   l = ots(1:i);
   k = [k2(1:i) k3(1:i) k4(1:i) ones(i,1)];
   w = regress(l,k);

   %z1 = k1(1:i);
   f1 = k2(1:i+hor);
```

```
    f2 = k3(1:i+hor);
    f3 = k4(1:i+hor);

    f = [f1 f2 f3 ones(i+hor,1)];    %νέο διάνυσμα των ανεξάρτητων μεταβλητών

    y = w'*f';        % παραγωγή πρόβλεψης της ΟΤΣ με το ανανεωμένο διάνυσμα

    y(find(y<0)) = 0;

    Y(j,1:hor) = ots(i+(1:hor));    %πίνακας με πραγματικές τιμές δεδομένων
    F(j,1:hor) = y(i+(1:hor));     %πίνακας με επανεποχικοποιημένες προβλέψεις

    j = j+1;
end

e = Y-F; %υπολογισμός σφάλματος σε πίνακα: ημέρες_πρόβλεψης*168_ώρες

sAPE = (abs(2*e./(Y+F)))*100; %sAPE

mean(mean(sAPE,1),2)            %μέση τιμή sMAPE
```