



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Τομέας Ηλεκτρικών Βιομηχανικών Διατάξεων και
Συστημάτων Αποφάσεων

**Βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου ηλεκτρικής
ενέργειας και εξέταση της επίδρασης των ειδικών ημερών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ειρήνη-Ελισάβετ Θεοδώρου

Επιβλέπων: Βασίλειος Ασημακόπουλος
Καθηγητής Ε.Μ.Π.

Υπεύθυνος: Φώτιος Πετρόπουλος
Διδάκτωρ Ε.Μ.Π.

Αθήνα, Οκτώβριος 2012



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Τομέας Ηλεκτρικών Βιομηχανικών Διατάξεων και
Συστημάτων Αποφάσεων

**Βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου ηλεκτρικής
ενέργειας και εξέταση της επίδρασης των ειδικών ημερών**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ειρήνη-Ελισάβετ Θεοδώρου

Επιβλέπων: Βασίλειος Ασημακόπουλος
Καθηγητής Ε.Μ.Π.

Υπεύθυνος: Φώτιος Πετρόπουλος
Διδάκτωρ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την ...^η Οκτωβρίου 2012.

.....
Β. Ασημακόπουλος,
Καθηγητής Ε.Μ.Π

.....
Ι. Ψαρράς,
Καθηγητής Ε.Μ.Π.

.....
Δ. Ασκούνης
Αναπλ. Καθηγητής Ε.Μ.Π

Αθήνα, Οκτώβριος 2012

.....
ΕΙΡΗΝΗ Ε. ΘΕΟΔΩΡΟΥ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών
Ε.Μ.Π

Copyright © ΕΙΡΗΝΗ Ε. ΘΕΟΔΩΡΟΥ, 2012
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Ευχαριστίες

Η διπλωματική αυτή εργασία εντάσσεται στα πλαίσια της ερευνητικής δραστηριότητας της Μονάδας Προβλέψεων και Στρατηγικής που υπάγεται στον Τομέα Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων. Προσωπικά, θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα Καθηγητή μου, κύριο Βασίλειο Ασημακόπουλο, ο οποίος μου έδωσε την ευκαιρία να ασχοληθώ με τον πραγματικά ενδιαφέροντα τομέα των προβλέψεων, όπως και στον Καθηγητή, κύριο Ιωάννη Ψαρρά και στον αναπληρωτή Καθηγητή, κύριο Δημήτριο Ασκούνη, για τη συμμετοχή τους στην επιτροπή εξέτασης της διπλωματικής εργασίας. Επίσης, θα ήθελα να ευχαριστήσω θερμά το διδάκτορα Φώτη Πετρόπουλο, ο οποίος ήταν παρών καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας, με υποστήριξε και με καθοδήγησε με συνέπεια και υπευθυνότητα. Ακόμη, ευχαριστώ τα υπόλοιπα μέλη της Μονάδας Προβλέψεων και Στρατηγικής, για την υποστήριξη και την παροχή βοήθειας, όποτε τη χρειαζόμουν. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, χάρη στις θυσίες και στην αδιάκοπη συμπαράσταση της οποίας έχω τη δυνατότητα να επιτυγχάνω τους στόχους μου, καθώς και τους φίλους μου, για την ηθική υποστήριξη που μου προσέφεραν.

Οκτώβριος 2012,
Ειρήνη Ε. Θεοδώρου

Περίληψη

Η πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας είναι ιδιαίτερα σημαντική τις τελευταίες δεκαετίες, με την έναρξη της απελευθέρωσης των αγορών ηλεκτρισμού. Συγκεκριμένα, η βραχυπρόθεσμη πρόβλεψη φορτίου ηλεκτρισμού είναι απαραίτητη για τη σχεδίαση των συστημάτων ενέργειας και την εκτίμηση της ασφάλειάς τους, ενώ σημαντικός είναι ο ρόλος της και στη διαμόρφωση της τιμής ηλεκτρικής ενέργειας. Για τους ανωτέρω λόγους, η παρούσα μελέτη δίνει έμφαση στην αξιολόγηση μεθόδων πρόβλεψης ζήτησης φορτίου ηλεκτρισμού (με χρονικό ορίζοντα μία ημέρα) και στην ενσωμάτωση παραγόντων που επηρεάζουν τη ζήτηση αυτή.

Αρχικά, παρουσιάζονται τα γενικά χαρακτηριστικά των χρονοσειρών ζήτησης φορτίου και αναλύεται η διαδικασία αποσύνθεσής στα μεμονωμένα ποιοτικά χαρακτηριστικά τους, με έμφαση στην εποχικότητα των δεδομένων. Δίνεται στη συνέχεια μια επισκόπηση της ποικιλίας των μεθόδων που χρησιμοποιούνται για βραχυπρόθεσμη πρόβλεψη φορτίου ζήτησης. Ακόμη, παρουσιάζονται οι μέθοδοι που αξιολογούνται στην παρούσα εργασία, βασικές μέθοδοι πρόβλεψης, τροποποιημένες ώστε να ενσωματώνουν μία εποχικότητα (Naive, SES, Damped, Holt και Theta), όπως επίσης και η προσαρμογή του Taylor στην εκθετική μέθοδο εξομάλυνσης των Holt-Winters για διπλή εποχικότητα. Επιπλέον, αναλύονται τα σφάλματα μέτρησης της ακρίβειας των μεθόδων και συγκρίνεται η απόδοσή τους επί διαφορετικού τύπου χρονοσειρών.

Έπειτα από το πρώτο, θεωρητικό τμήμα της μελέτης, αναλύεται η χρονοσειρά με τα πραγματικά δεδομένα ζήτησης φορτίου ηλεκτρισμού. Που μελετάται. Στην παρούσα εργασία εξετάζεται η απόδοση μιας μεθοδολογίας πρόβλεψης, αποτελούμενης από δύο στάδια. Στο πρώτο στάδιο, γίνεται βραχυπρόθεσμη προέκταση των δεδομένων από τις τεχνικές που αναφέρθηκαν παραπάνω, με εφαρμογή της κλασσικής αποσύνθεσης όπου είναι κρίνεται απαραίτητη. Εφόσον η πρόβλεψη ζήτησης φορτίου απαιτεί γρήγορη αντίδραση σε εξωτερικές μεταβολές της αγοράς, η αποτελεσματικότητα της κυλιόμενης πρόβλεψης γίνεται έντονα αισθητή και η διαδικασία της (εφαρμόζεται σε κάθε βήμα της μεθοδολογίας) περιγράφεται λεπτομερώς. Τα αποτελέσματα όλων των μεθόδων που αναφέρθηκαν παρουσιάζονται στη συνέχεια και είναι αρκετά ικανοποιητικά, όπως αναμενόταν.

Το δεύτερο στάδιο της μεθοδολογίας περιλαμβάνει την ενσωμάτωση των ειδικών ημερών (αργίες, απεργίες) στις προβλέψεις του πρώτου σταδίου. Ιδιαίτερη προσοχή δίνεται στους διαφορετικούς συνδυασμούς ειδικών ημερών και οι προβλέψεις των πιο αποδοτικών εξ' αυτών παρουσιάζονται στη συνέχεια και συγκρίνονται με τα αρχικά αποτελέσματα του πρώτου σταδίου. Με συγκεντρωτική παρουσίαση των αποτελεσμάτων, αποδεικνύονται ανώτερες οι απλούστερες μέθοδοι προέκτασης των δεδομένων και υπογραμμίζεται η ανάγκη για προσαρμογές επί των μεθόδων, βάσει των ειδικών ημερών. Τέλος, αναφέρεται η δυνατότητα περαιτέρω μείωσης του σφάλματος με ενσωμάτωση των καιρικών συνθηκών στις μεθόδους που εξετάζονται στην παρούσα μελέτη.

Λέξεις κλειδιά: Πρόβλεψη ζήτησης ηλεκτρισμού, αποσύνθεση, διπλή εποχικότητα, κυλιόμενη πρόβλεψη, επίδραση ειδικών ημερών.

Abstract

Electricity load demand forecasting has been of great importance over the past decades, since the deregulation of electricity markets. Specifically, short-term load forecasting is required for scheduling power systems and assessing their security, and it also plays a great role in electricity price fixing. Therefore, this study is focused on the evaluation of methods for electricity load demand forecasting (24-hours ahead forecasts) and the incorporation of factors that influence this demand.

At first, the characteristics of load demand time series are specified and the process of decomposition into its individual components is analyzed, with emphasis on the seasonality of the data. A general review of the variety of methods used for short-term electricity load demand is given. Moreover, the methods evaluated in this paper are presented, basic forecasting methods modified to incorporate one seasonality (Naive, SES, Damped, Holt and Theta), as well as the adaptation of Holt-Winters exponential smoothing method for double seasonality. Furthermore, the errors for accuracy measurement are analyzed along with their performance for different types of time series.

After the first, theoretical part of the study, the time series with empirical data of electricity load demand is analyzed. A two-step forecasting methodology is investigated on this study. Firstly, short-term extrapolation of the data is performed by the techniques mentioned above, with classical decomposition applied when necessary. Furthermore, since load demand forecasting requires quick response to external market changes, the effectiveness of rolling forecasting is keenly felt and its procedure (implemented on every step of the methodology) is described in detail. The results of all methods mentioned are presented and -as expected- they are pretty satisfying.

The second step of the methodology includes the incorporation of special days (bank holidays, strikes) on the produced forecasts. Specific attention is given on different combinations of special days and the forecasts of the most efficient of them are presented and compared with existing results of the first step. Overall, simpler extrapolation techniques are proven superior than more complex methods and the need for adjustments based on special days is underlined. Finally, potential for further error reduction by incorporating weather's impact on the methods examined on this study is mentioned.

Key words: Electricity demand forecasting, decomposition, double seasonality, rolling forecasting, special days' impact.

Πίνακας περιεχομένων

Περίληψη	- 7 -
Abstract.....	- 9 -
ΚΕΦΑΛΑΙΟ 1: Ευρεία περίληψη	- 21 -
1.1 Πρόβλεψη Ζήτησης Φορτίου	- 21 -
1.2 Χρονοσειρές.....	- 22 -
1.3 Μέθοδοι Πρόβλεψης.....	- 24 -
1.4 Μεθοδολογία	- 25 -
1.5 Αποτελέσματα Προβλέψεων.....	- 27 -
1.6 Επίδραση Ειδικών Ημερών.....	- 29 -
1.7 Συμπεράσματα και Προοπτικές.....	- 31 -
ΚΕΦΑΛΑΙΟ 2: Εισαγωγή.....	- 33 -
2.1 Σημασία της πρόβλεψης φορτίου	- 33 -
2.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου ηλεκτρισμού.....	- 34 -
2.3 Το αντικείμενο της διπλωματικής	- 36 -
ΚΕΦΑΛΑΙΟ 3: Χρονοσειρές.....	- 39 -
3.1 Γενικά χαρακτηριστικά χρονοσειρών	- 39 -
3.2 Κλασική μέθοδος αποσύνθεσης.....	- 41 -
3.3 Επίδραση ειδικών γεγονότων επί των χρονοσειρών.....	- 43 -
3.4 Στατιστική ανάλυση χρονοσειρών.....	- 44 -
3.5 Χαρακτηριστικά χρονοσειρών ζήτησης φορτίου.....	- 46 -
ΚΕΦΑΛΑΙΟ 4: Μέθοδοι πρόβλεψης	- 49 -
4.1 Γενικά	- 49 -
4.2 Κατηγορίες πρόβλεψης.....	- 50 -
4.3 Στατιστικές μέθοδοι πρόβλεψης	- 50 -
4.3.1 Απλοϊκή μέθοδος Naive.....	- 50 -
4.3.2 Μέθοδοι εκθετικής εξομάλυνσης.....	- 51 -
4.3.2.1. Απλή εκθετική εξομάλυνση σταθερού επιπέδου (Seasonal Exponential Smoothing).....	- 51 -
4.3.2.2 Εκθετική εξομάλυνση γραμμικής τάσης (Holt Exponential Smoothing)	- 52 -
4.3.2.3 Εκθετική εξομάλυνση μη γραμμικής τάσης.....	- 53 -
4.3.3 Μοντέλα παλινδρόμησης.....	- 54 -
4.3.3.1 Απλή γραμμική παλινδρόμηση	- 55 -
4.3.3.2 Πολλαπλή γραμμική παλινδρόμηση	- 56 -

4.3.4 Μέθοδος Theta	- 56 -
4.3.4 Μοντέλα ARIMA.....	- 57 -
4.3.5 Νευρωνικά δίκτυα.....	- 57 -
4.4 Εφαρμογή εποχικότητας στα χρησιμοποιούμενα μοντέλα	- 58 -
4.5 Η προσαρμοσμένη μέθοδος Holt-Winters για την ενσωμάτωση διπλής εποχικότητας (από τον J. W. Taylor).....	- 59 -
4.6 Αξιολόγηση της ακρίβειας των προβλέψεων.....	- 61 -
4.7 Μοντέλα βραχυπρόθεσμης πρόβλεψης	- 63 -
ΚΕΦΑΛΑΙΟ 5: Μεθοδολογία	- 67 -
5.1 Δεδομένα.....	- 67 -
5.2 Διαδικασία παραγωγής προβλέψεων	- 69 -
5.3 Εφαρμογή κυλιόμενης πρόβλεψης (rolling forecasting)	- 71 -
5.4 Αξιολόγηση	- 74 -
ΚΕΦΑΛΑΙΟ 6: Αποτελέσματα	- 77 -
6.1 Παρουσίαση αποτελεσμάτων	- 77 -
6.2 Ανάλυση αποτελεσμάτων	- 81 -
ΚΕΦΑΛΑΙΟ 7: ΕΠΙΔΡΑΣΗ ΕΙΔΙΚΩΝ ΗΜΕΡΩΝ	- 91 -
7.1 Επιλογή ειδικών ημερών.....	- 91 -
7.2 Εισαγωγή ειδικών ημερών στα υπάρχοντα μοντέλα πρόβλεψης.....	- 93 -
7.2.1 Ενσωμάτωση μόνο των αργιών	- 95 -
7.2.2 Ενσωμάτωση μόνο των απεργιών	- 95 -
7.2.3 Ενσωμάτωση συνδυασμού αργιών και απεργιών.....	- 96 -
7.3 Αποτελέσματα νέων προβλέψεων	- 96 -
7.3.1 Αποτελέσματα με επίδραση επίσημων αργιών	- 96 -
7.3.1.1 Επίσημες αργίες	- 96 -
7.3.1.2 Επίσημες και μη αργίες	- 101 -
7.3.2 Αποτελέσματα με επίδραση των απεργιών (Μέσος όρος όλων των απεργιών των ετών 2009 και 2010).....	- 105 -
7.3.3 Αποτελέσματα με επίδραση αργιών και απεργιών	- 110 -
7.4 Ανάλυση αποτελεσμάτων	- 114 -
ΚΕΦΑΛΑΙΟ 8: ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΡΟΕΚΤΑΣΕΙΣ	- 121 -
8.1 Συνολικά αποτελέσματα εργασίας	- 121 -
8.2 Προεκτάσεις	- 124 -
Βιβλιογραφία	- 127 -

Ευρετήριο Πινάκων

Πίνακας 1. Πλήθος διαθέσιμων δεδομένων.....	- 69 -
Πίνακας 2. Σφάλματα MAPE και sMAPE μεθόδων (α).....	- 77 -
Πίνακας 3. Σφάλματα MAPE και sMAPE μεθόδων (β).....	- 78 -
Πίνακας 4. Σφάλματα MdAPE και sMdAPE μεθόδων (α).....	- 78 -
Πίνακας 5. Σφάλματα MdAPE και sMdAPE μεθόδων (β).....	- 79 -
Πίνακας 6. Τα μικρότερα σφάλματα όλων των μεθόδων.....	- 83 -
Πίνακας 7. Ελάχιστο και μέγιστο σφάλμα MAPE των μεθόδων και η διακύμανσή του (ΜΟ των ωριαίων σφαλμάτων για όλο το 2011).....	- 88 -
Πίνακας 8. Ποσοστό πλήθους σφαλμάτων (MAPE, sMAPE) μικρότερο του 3%.....	- 88 -
Πίνακας 9. Επίσημες αργίες για τα έτη 2009, 2010, 2011	- 92 -
Πίνακας 10. Μη επίσημες αργίες για τα έτη 2009, 2010, 2011	- 92 -
Πίνακας 11. Απεργίες για τα έτη 2009, 2010, 2011.....	- 93 -
Πίνακας 12. Σφάλματα MAPE και sMAPE με επίδραση των επίσημων αργιών (α)	- 97 -
Πίνακας 13. Σφάλματα MAPE και sMAPE με επίδραση των επίσημων αργιών (β)	- 97 -
Πίνακας 14. Σφάλματα MdAPE και sMdAPE με επίδραση των επίσημων αργιών (α)	- 99 -
Πίνακας 15. Σφάλματα MdAPE και sMdAPE με επίδραση των επίσημων αργιών (β)	- 100 -
Πίνακας 16. Σφάλματα MAPE και sMAPE με επίδραση όλων των αργιών, επίσημων και μη (α)	- 101 -
Πίνακας 17. Σφάλματα MAPE και sMAPE με επίδραση όλων των αργιών, επίσημων και μη (β)	- 102 -
Πίνακας 18. Σφάλματα MdAPE και sMdAPE με επίδραση όλων των αργιών, επίσημων και μη (α)	- 103 -
Πίνακας 19. Σφάλματα MdAPE και sMdAPE με επίδραση όλων των αργιών, επίσημων και μη (β)	- 104 -
Πίνακας 20. Σφάλματα MAPE και sMAPE με επίδραση του ΜΟ	- 106 -
Πίνακας 21. Σφάλματα MAPE και sMAPE με επίδραση του ΜΟ	- 106 -
Πίνακας 22. Σφάλματα MdAPE και sMdAPE επίδραση του ΜΟ όλων των απεργιών (α)	- 108 -
Πίνακας 23. Σφάλματα MdAPE και sMdAPE επίδραση του ΜΟ όλων των απεργιών (β)	- 108 -
Πίνακας 24. Σφάλματα MAPE και sMAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (α).....	- 110 -
Πίνακας 25. Σφάλματα MAPE και sMAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (β)	- 110 -

Πίνακας 26. Σφάλματα MdAPE και sMdAPE με την επίδραση των επίσημων αργιών και του MO όλων των απεργιών (α)	- 112 -
Πίνακας 27. Σφάλματα MdAPE και sMdAPE με την επίδραση των επίσημων αργιών και του MO όλων των απεργιών (α)	- 112 -
Πίνακας 28α. Σφάλματα MAPE για τη μέθοδο Naive πριν και μετά την επίδραση των ειδικών ημερών	- 114 -
Πίνακας 28β. Σφάλματα MAPE για τη μέθοδο Naive πριν και μετά την επίδραση των ειδικών ημερών	- 115 -
Πίνακας 29α. Ποσοστά ωριαίων προβλέψεων με σφάλμα (MAPE και sMAPE) μικρότερο του 3%, πριν και μετά από την ενσωμάτωση της επίδρασης των ειδικών ημερών.....	- 116 -
Πίνακας 29β. Ποσοστά ωριαίων προβλέψεων με σφάλμα (MAPE και sMAPE) μικρότερο του 3%, πριν και μετά από την ενσωμάτωση της επίδρασης των ειδικών ημερών.....	- 117 -
Πίνακας 30α. Βέλτιστα σφάλματα μεθόδων για όλες τις προβλέψεις	- 118 -
Πίνακας 30β. Βέλτιστα σφάλματα μεθόδων για όλες τις προβλέψεις	- 119 -
Πίνακας 31. Βέλτιστες εκτελέσεις της μεθόδου Naive	- 122 -
Πίνακας 32. Βέλτιστες εκτελέσεις της μεθόδου SES.....	- 122 -
Πίνακας 33. Βέλτιστες εκτελέσεις της μεθόδου Holt	- 122 -
Πίνακας 34. Βέλτιστες εκτελέσεις της μεθόδου Theta	- 122 -
Πίνακας 35. Βέλτιστες εκτελέσεις της μεθόδου Damped.....	- 123 -
Πίνακας 36. Βέλτιστες εκτελέσεις της προσαρμοσμένης μεθόδου του Taylor.....	- 123 -

Ευρετήριο Εικόνων

Εικόνα 1. Ημερήσια ζήτηση φορτίου για την 21 Ιανουαρίου (σε kWh)	- 67 -
Εικόνα 2. Ημερήσια ζήτηση φορτίου για την 21 Ιουλίου (σε kWh).....	- 68 -
Εικόνα 3. Εβδομαδιαία ζήτηση φορτίου για δύο εβδομάδες του Ιανουαρίου 2009 (σε kWh).....	- 68 -
Εικόνα 4. Εβδομαδιαία ζήτηση φορτίου για δύο εβδομάδες του Ιουλίου 2009 (σε kWh)	- 69 -
Εικόνα 5. Παράδειγμα μεθόδου rolling forecasting.....	- 72 -
Εικόνα 6. Ολοκληρωμένη η διαδικασία προβλέψεων.....	- 76 -
Εικόνα 7. Σφάλμα MAPE για ημερήσια εποχικότητα.....	- 79 -
Εικόνα 8. Σφάλμα MAPE για εβδομαδιαία εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα -	80 -
Εικόνα 9. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με εφαρμογή rolling για διαφορετικά χρονικά διαστήματα	- 80 -
Εικόνα 10. Σφάλμα MdAPE για ημερήσια εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα..	80 -
Εικόνα 11. Σφάλμα MdAPE για εβδομαδιαία εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα.....	- 81 -
Εικόνα 12. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με εφαρμογή rolling για διαφορετικά χρονικά διαστήματα	- 81 -
Εικόνα 13. Ζήτηση φορτίου ηλεκτρισμού για το διάστημα 1-14/4/2011 (πραγματική και βέλτιστη πρόβλεψη, Naive)-	85 -
Εικόνα 14. Ζήτηση φορτίου ηλεκτρισμού και για τους 8 μήνες του 2011 (πραγματική χρονοσειράς και βέλτιστη πρόβλεψη με Naive)	- 86 -
Εικόνα 15. Ζήτηση φορτίου ηλεκτρισμού και για τους 8 μήνες του 2011 (πραγματική και βέλτιστη προβλεπόμενη μέσω Damped- χρονοσειρά)	- 86 -
Εικόνα 16. Μέσο ωριαίο σφάλμα MAPE για τις βέλτιστες προβλεπόμενες χρονοσειρές όλων των μεθόδων.....	- 87 -
Εικόνα 17. Πραγματική ζήτηση φορτίου για τη χρονική περίοδο.....	- 93 -
Εικόνα 18. Σφάλμα MAPE με επίδραση των επίσημων αργιών.....	- 98 -
Εικόνα 19. Σφάλμα MAPE με επίδραση των επίσημων αργιών και εβδομαδιαία εποχικότητα	- 98 -
Εικόνα 20. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών.....	- 99 -

Εικόνα 21. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και ημερήσια εποχικότητα.....	- 100 -
Εικόνα 22. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και εβδομαδιαία εποχικότητα	- 100 -
Εικόνα 23. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών.....	- 101 -
Εικόνα 24. Σφάλμα MAPE με επίδραση επίσημων και μη αργιών και ημερήσια εποχικότητα	- 102 -
Εικόνα 25. Σφάλμα MAPE με επίδραση επίσημων και μη αργιών και εβδομαδιαία εποχικότητα	- 102 -
Εικόνα 26. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση επίσημων και μη αργιών	- 103 -
Εικόνα 27. Σφάλμα MdAPE με επίδραση επίσημων και μη αργιών και ημερήσια εποχικότητα	- 104 -
Εικόνα 28. Σφάλμα MdAPE με επίδραση επίσημων και μη αργιών και εβδομαδιαία εποχικότητα	- 105 -
Εικόνα 29. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση επίσημων και μη αργιών	- 105 -
Εικόνα 30. Σφάλμα MAPE επίδραση του MO όλων των απεργιών και ημερήσια εποχικότητα.....	- 107 -
Εικόνα 31. Σφάλμα MAPE με επίδραση του MO όλων των απεργιών και εβδομαδιαία εποχικότητα	- 107 -
Εικόνα 32. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση του MO όλων των απεργιών	- 107 -
Εικόνα 33. Σφάλμα MdAPE με επίδραση του MO όλων των απεργιών και ημερήσια εποχικότητα.....	- 109 -
Εικόνα 34. Σφάλμα MdAPE με επίδραση του MO όλων των απεργιών και εβδομαδιαία εποχικότητα	- 109 -
Εικόνα 35. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση του MO όλων των απεργιών	- 109 -
Εικόνα 36. Σφάλμα MAPE με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (και ημερήσια εποχικότητα).....	- 111 -
Εικόνα 37. Σφάλμα MAPE με επίδραση των επίσημων αργιών & όλων των απεργιών (και εβδομαδιαία εποχικότητα)	- 111 -
Εικόνα 38. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών και του MO όλων των απεργιών	- 111 -
Εικόνα 39. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (και ημερήσια εποχικότητα).....	- 113 -
Εικόνα 40. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (και εβδομαδιαία εποχικότητα).....	- 113 -

Εικόνα 41. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών- 113 -

ΚΕΦΑΛΑΙΟ 1: Ευρεία περίληψη

1.1 Πρόβλεψη Ζήτησης Φορτίου

Η πρόβλεψη της ζήτησης φορτίου ηλεκτρικής ενέργειας είναι ένας έντονα ενεργός κλάδος της επιστήμης των προβλέψεων, στον οποίο οι ερευνητές έχουν επικεντρωθεί μόλις στο πρόσφατο παρελθόν. Βασική αιτία για την επίδειξη μεγάλης προσοχής στο συγκεκριμένο κλάδο, είναι η απελευθέρωση των αγορών ηλεκτρικής ενέργειας. Με την κίνηση αυτή, καταλύθηκαν τα μονοπώλια που προϋπήρχαν, παρουσιάστηκαν νέες εταιρείες παραγωγής ενέργειας στις επιμέρους αγορές ηλεκτρικής ενέργειας και όλες οι εταιρείες πλέον λειτουργούν υπό κλίμα έντονου ανταγωνισμού. Είναι λοιπόν σημαντική για τις εταιρείες η κατοχή της προβλεπόμενης ζήτησης του φορτίου ηλεκτρισμού, ώστε να μπορεί να ανταπεξέρχεται στις απαιτήσεις της αγοράς. Στην Ελλάδα, η Δημόσια Εταιρεία Ηλεκτρισμού, ΔΕΗ, έχασε το προνόμιο του αποκλειστικού παραγωγού ηλεκτρικής ενέργειας της χώρας και έκτοτε αρκετές ιδιωτικές εταιρείες έχουν εμφανιστεί στην παροχή της Υψηλής και της Μέσης Τάσης.

Εκτός από την απελευθέρωση της αγοράς, υπάρχουν πολλοί ακόμη σημαντικοί λόγοι για τους οποίους μια εταιρεία παραγωγής ηλεκτρικής ενέργειας χρειάζεται την πρόβλεψη των τιμών του ζητούμενου φορτίου. Οι λόγοι αυτοί φαίνονται από τις εφαρμογές της πρόβλεψης του φορτίου, που αναφέρονται αμέσως μετά. Συγκεκριμένα, η βασικότερη εφαρμογή της πρόβλεψης του ζητούμενου φορτίου παρατηρείται κατά τον οικονομικό προγραμματισμό του συστήματος παραγωγής ενέργειας. Αυτό σημαίνει ότι η πληροφορία της πρόβλεψης αξιοποιείται για το σχεδιασμό της ένταξης μονάδων παραγωγής στο δίκτυο με τον καλύτερο δυνατό τρόπο, όπως και για προσοδοφόρες συμφωνίες ανταλλαγής ενέργειας με γειτονικά δίκτυα. Ακόμη, η πρόβλεψη ζητούμενου φορτίου συμβάλλει στην ασφάλεια του συστήματος παραγωγής ενέργειας (ΣΗΕ), καθώς προειδοποιεί τους χειριστές του συστήματος για ενδεχόμενες έντονες διακυμάνσεις της ζήτησης και διαταραχές, ώστε να προβούν σε ανάλογες ενέργειες για την αποφυγή ανεπιθύμητων καταστάσεων. Μερικές από τις ενέργειες των χειριστών είναι η αγορά ενέργειας από γειτονικά δίκτυα και η ένταξη μονάδων αιχμής (σε περίπτωση που προβλεφθεί έντονη αύξηση της ζήτησης φορτίου. Αυτή η κατάσταση οδηγεί σε καλύτερη διαχείριση του ΣΗΕ και αύξηση της αξιοπιστίας του συστήματος. Τέλος, η κατοχή προβλέψεων ακριβείας σχετικά με το ζητούμενο φορτίο ενέργειας, συμβάλλει στη μείωση του κόστους παραγωγής και λειτουργίας ενός ΣΗΕ

Η κατηγοριοποίηση των προβλέψεων ζήτησης φορτίου ανάλογα με το χρονικό ορίζοντά τους, είναι σημαντική για την επιλογή της σωστής μεθόδου που θα χρησιμοποιήσει κανείς. Συγκεκριμένα, οι κατηγορίες βάσει χρονικού ορίζοντα πρόβλεψης είναι : πολύ βραχυπρόθεσμη, βραχυπρόθεσμη, μεσοπρόθεσμη και μακροπρόθεσμη. Στην παρούσα μελέτη το ενδιαφέρον μας περιορίζεται στη βραχυπρόθεσμη πρόβλεψη, που χρησιμεύει στην ενεργειακή διαχείριση των ΣΗΕ.

Βασικό βήμα πριν την ανάλυση και αξιολόγηση οποιαδήποτε χρονοσειράς ζήτησης ηλεκτρικού φορτίου αποτελεί ο εντοπισμός των παραγόντων που επηρεάζουν τη συγκεκριμένη μορφή ζήτησης και ο τρόπος και το μέγεθος της επιρροής αυτής. Στην παρούσα μελέτη μας ενδιαφέρει η βραχυπρόθεσμη πρόβλεψη ωριαίων τιμών ζήτησης φορτίου, για χρονικό ορίζοντα μίας ημέρας. Οι παράγοντες που καταγράψαμε κατά την έρευνά μας είναι κυρίως εποχικής, μετεωρολογικής και οικονομικής φύσης. Υπάρχουν βέβαια και τυχαίοι παράγοντες, τυχαία γεγονότα δηλαδή, όπως εκλογές ή μεγάλης διάρκειας και συμμετοχής απεργίες.

Στους εποχικούς παράγοντες συγκαταλέγονται οι εναλλαγές των εποχών μέσα στο ημερολογιακό έτος, οι διάφορες εποχικότητες που παρουσιάζουν τα δεδομένα μιας χρονοσειράς και οι ειδικές ημέρες που παρατηρούνται και μεταβάλλονται ανάλογα με την ευρύτερη περιοχή (συνήθως χώρα) στην οποία διενεργείται η μελέτη. Ένα παράδειγμα εναλλαγής εποχιακών μεταβλητών μέσα στο ημερολογιακό έτος, είναι η διάρκεια της ημέρας (με αύξηση της διάρκειας όσο πλησιάζει το καλοκαίρι και μείωση αντίστοιχα το χειμώνα), που οδηγεί σε σταδιακή αλλαγή του επιπέδου ζήτησης φορτίου. Έντονη είναι η επίδραση της ημερήσιας και εβδομαδιαίας εποχικότητας μιας χρονοσειράς ζήτησης ηλεκτρικής ενέργειας, καθώς η καμπύλη της χρονοσειράς παρουσιάζει πάντα ένα επαναλαμβανόμενο μοτίβο τόσο μέσα στην ημέρα, όσο και σε επίπεδο εβδομάδας, που υποδεικνύει ότι η ημέρα της εβδομάδας παίζει σημαντικό ρόλο στη διαμόρφωση της ζήτησης. Τέλος, οι διάφορες ειδικές ημέρες, όπως αργίες (πχ. Χριστούγεννα και Πάσχα) και απεργίες, οδηγούν σε απότομη μεταβολή της ζήτησης ενέργειας από το καταναλωτικό κοινό.

Όσον αφορά τους μετεωρολογικούς παράγοντες, εμφανής είναι η συμβολή τους στη διαμόρφωση της ζήτησης του φορτίου ηλεκτρικής ενέργειας. Σημαντικότερος παράγοντας της κατηγορίας αυτής θεωρείται η θερμοκρασία, ενώ άλλοι παράγοντες είναι η υγρασία, η ηλιοφάνεια και οι άνεμοι που παρατηρούνται στην περιοχή που μελετάται. Ένα παράδειγμα επιρροής της μεταβολής της ζήτησης από μετεωρολογικό παράγοντα είναι η αύξηση της ζήτησης φορτίου λόγω έντονης χρήσης κλιματιστικών κατά την παρουσίαση έντονου καύσωνα το καλοκαίρι.

Ως οικονομικούς παράγοντες αντιλαμβανόμαστε την κατάσταση της οικονομίας μιας περιοχής, καθώς και την ύπαρξη ή μη ρυθμού ανάπτυξης. Οι παράγοντες αυτοί επιδρούν στη διαμόρφωση της ζήτησης φορτίου ηλεκτρικής ενέργειας μέσω αλλαγής του βιοτικού επιπέδου των καταναλωτών. Η μεταβολή του πλήθους των καταναλωτών και της σύνθεσής του, όπως και η εξέλιξη των τεχνολογικών επιτευγμάτων οδηγούν σε αντίστοιχη μεταβολή του ζητούμενου φορτίου ενέργειας.

Στα πλαίσια της παρούσας μελέτης, επιλέγεται να εξεταστεί η επίδραση των εποχικών παραγόντων σε ορισμένα μοντέλα πρόβλεψης, τα οποία με τη σειρά τους επιλέγονται ανάμεσα σε πλήθος μοντέλων που χρησιμοποιούνται στη βραχυπρόθεσμη πρόβλεψη ζητούμενου φορτίου ενέργειας. Η διαδικασία πρόβλεψης αναλύεται στη συνέχεια.

1.2 Χρονοσειρές

Σημαντικό μέρος της διαδικασίας παραγωγής προβλέψεων αποτελεί η απόκτηση και η ανάλυση των χρονοσειρών προς μελέτη. Οι χρονοσειρές αποτελούν σύνολα διαδοχικών παρατηρήσεων της τιμής ενός μεγέθους (φυσικού ή και όχι). Μέσω της παρατήρησης των χρονοσειρών, εξάγονται συμπεράσματα για την πορεία του μεγέθους στο χρόνο και αναδεικνύονται τα ποιοτικά χαρακτηριστικά των χρονοσειρών, η τάση, η εποχικότητα, ο κύκλος (αλλιώς κυκλικότητα) και η τυχαιότητα. Ακόμη, εντοπίζονται οι ενδεχόμενες ακραίες τιμές των χρονοσειρών, όπως εκλιπούσες ή μηδενικές τιμές είτε τιμές που χρήζουν ημερολογιακής προσαρμογής. Διευκολύνεται επομένως το έργο του χειριστή, ώστε να επέμβει και να διορθώσει τις ακραίες τιμές.

Πριν προβεί κανείς σε εφαρμογή μοντέλων πρόβλεψης επί μιας χρονοσειράς, ή ακόμη και μόνο για την καλύτερη αντίληψη της συμπεριφοράς του μεγέθους που περιγράφει η χρονοσειρά αυτή, σημαντικό βήμα είναι η απομόνωση μίας ή και περισσότερων από τις συνιστώσες της. Η απομόνωση των συνιστωσών επιτυγχάνεται με εφαρμογή μεθόδων

αποσύνθεσης. Να σημειώσουμε ότι οι πιο απλές διατυπώσεις αποσύνθεσης είναι η προσθετική και η πολλαπλασιαστική, με τη δεύτερη περίπτωση να αποτελεί το πιο συχνό μοντέλο. Απομονώνοντας από τη χρονοσειρά τις συνιστώσες της εποχικότητας και της τυχαιότητας, προκύπτει η σειρά τάσης-κύκλου, στην οποία εφαρμόζονται τα μοντέλα πρόβλεψης, τα αποτελέσματα των οποίων εποχικοποιούνται ξανά στη συνέχεια για τη λήψη των τελικών τιμών.

Οι δύο βασικές μέθοδοι παραγωγής της σειράς τάσης-κύκλου είναι η εξομάλυνση και η αποσύνθεση. Στο κεφάλαιο 3 παρουσιάζεται η μέθοδος εξομάλυνσης με χρήση κινητών μέσων όρων και η μέθοδος της κλασσικής αποσύνθεσης. Σημειώνουμε ότι στην εργασία μας χρησιμοποιείται η μέθοδος της κλασσικής αποσύνθεσης, με Κεντρικό Κινητό Μέσο Όρο, όπως θα αναλυθεί στο κεφάλαιο 4 (παράγραφος 4.4), λόγω του άρτιου μήκους της εποχικότητας των δεδομένων [με ωριαίες τιμές, η ημερήσια εποχικότητα έχει μήκος $s_1 = 24$ ενώ η εβδομαδιαία έχει αντίστοιχο μήκος $s_2 = 168$]. Κατά την εξομάλυνση, υπολογίζονται οι δείκτες εποχικότητας των δεδομένων που χρησιμοποιούνται, παράγεται η αποεποχικοποιημένη χρονοσειρά με διαίρεση των αρχικών δεδομένων με τους δείκτες αυτούς και, έπειτα από την εφαρμογή των μοντέλων, πολλαπλασιάζονται οι ωριαίες προβλέψεις με τους αντίστοιχους δείκτες εποχικότητας, ώστε το τελικό αποτέλεσμα να ανταποκρίνεται στην πραγματικότητα με μεγαλύτερη ακρίβεια.

Ένα ακόμη χαρακτηριστικό της χρονοσειράς ενός μεγέθους είναι η ύπαρξη διακυμάνσεων των τιμών των δεδομένων της, λόγω ειδικών γεγονότων που συμβαίνουν κατά το χρονικό διάστημα που γίνεται η καταγραφή των δεδομένων αυτών. Τα ειδικά γεγονότα έχουν παρατηρούνται στο άμεσο, όπως και στο ευρύτερο περιβάλλον του μελετώμενου μεγέθους και η επίδρασή τους αποτυπώνεται στη χρονοσειρά με την αλλαγή του επιπέδου της προς τη θετική ή την αρνητική κατεύθυνση, καθώς και με την εμφάνιση περιόδων ιδιαίτερα ασυνήθιστων τιμών. Ένα παράδειγμα είναι η διεξαγωγή Ολυμπιακών Αγώνων σε μία χώρα, οπότε και επηρεάζεται έντονα η ζήτηση ενέργειας της ευρύτερης αυτής περιοχής (προς τη θετική φυσικά κατεύθυνση εδώ). Είναι επομένως σημαντικό να αφαιρούνται οι επιδράσεις των ειδικών γεγονότων από τα in-sample δεδομένα με εξομάλυνση των τιμών, ώστε να οδηγούμαστε σε περισσότερο ακριβείς προβλέψεις.

Όπως ειπώθηκε, πριν από την παραγωγή προβλέψεων, σημαντική θεωρείται η επισκόπηση της χρονοσειράς ενός μεγέθους. Αυτό επιτυγχάνεται με στατιστική ανάλυση της χρονοσειράς. Συγκεκριμένα, εξετάζονται οι τιμές βασικών στατιστικών δεικτών, όπως η μέση τιμή, οι ακραίες τιμές, η τυπική απόκλιση και η διακύμανση. Έχοντας καλύτερη εικόνα της χρονοσειράς, είναι δυνατή και η πιο σωστή επιλογή μοντέλου πρόβλεψης.

Τέλος, συνοψίζουμε τα χαρακτηριστικά των χρονοσειρών ζήτησης φορτίου ηλεκτρικής ενέργειας:

- Η έντονη διακύμανση των τιμών των χρονοσειρών, με τη μέση τιμή τους να μην είναι σταθερή.
- Η ύπαρξη έντονης εποχικότητας και μάλιστα όχι μόνο ενός είδους, καθώς όλες οι χρονοσειρές ζήτησης φορτίου (στις οποίες μπορεί να εφαρμοστεί βραχυπρόθεσμη πρόβλεψη) παρουσιάζουν τουλάχιστον ημερήσια και εβδομαδιαία εποχικότητα.
- Η σύνδεση της εμφάνισης ειδικών ημερών με μεταβολές του ζητούμενου φορτίου.
- Η φανερή επίδραση των καιρικών συνθηκών, με βασικότερη τη θερμοκρασία, στη διαμόρφωση του ζητούμενου φορτίου.

Η μελέτη της βραχυπρόθεσμης πρόβλεψης στην παρούσα εργασία χωρίζεται σε δύο μέρη, με το πρώτο να περιλαμβάνει τις εποχικότητες της χρονοσειράς, ενώ το δεύτερο ενσωματώνει και τις ειδικές ημέρες που καταγράφονται. Στην παράγραφο 1.4 περιγράφεται η μεθοδολογία που ακολουθείται για την παραγωγή προβλέψεων για το πρώτο μέρος της μελέτης, ενώ η διαδικασία του δεύτερου μέρους παρουσιάζεται στην παράγραφο 1.6.

1.3 Μέθοδοι Πρόβλεψης

Αναφέραμε στην παράγραφο 1.2 ότι οι χρονοσειρές είναι πολύ σημαντικές για την εξαγωγή συμπερασμάτων σχετικά με την εξέλιξη των μεγεθών που περιγράφουν. Στην επιστήμη των προβλέψεων ωστόσο, το βασικότερο κομμάτι κατά την επεξεργασία μιας χρονοσειράς είναι η εύρεση της καλύτερης δυνατής προσέγγισης της πορείας της χρονοσειράς στο μέλλον. Η ανάγκη μας αυτή ικανοποιείται με την ανάπτυξη μοντέλων πρόβλεψης, τα οποία εφαρμόζονται επί των διαθέσιμων δεδομένων της χρονοσειράς και προβλέπουν τις τιμές των δεδομένων για έναν ορισμένο χρονικό ορίζοντα. Τρεις είναι οι κατηγορίες πρόβλεψης, η στατιστική, η κριτική και η πρόβλεψη στόχου.

Όπως αντιλαμβάνεται εύκολα κανείς, η κριτική πρόβλεψη στηρίζεται σε προσωπική κρίση του χειριστή, με βάση προηγούμενη εμπειρία ή γνώση του και είναι διαδεδομένη η χρήση της στις επιχειρήσεις. Η πρόβλεψη στόχου είναι ουσιαστικά η επιθυμητή πορεία της χρονοσειράς και όχι κάποια παραγόμενη τιμή. Τέλος, η στατιστική πρόβλεψη είναι που μας ενδιαφέρει περισσότερο, καθώς περιλαμβάνει τα μοντέλα πρόβλεψης που έχουν αναπτυχθεί και παράγει πιο ακριβείς προβλέψεις. Τα μοντέλα αυτά χωρίζονται σε δύο κατηγορίες, τα μοντέλα χρονοσειρών και τα αιτιοκρατικά μοντέλα. Στην πρώτη κατηγορία γίνεται η υπόθεση της μεταβολής του μεγέθους σύμφωνα με ένα καθορισμένο πρότυπο, που επαναλαμβάνεται στο χρόνο. Μοντέλα αυτής της κατηγορίας είναι οι αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (ARMA). Η δεύτερη κατηγορία υποθέτει την ύπαρξη συγκεκριμένης σταθερής σχέσης μεταξύ του προβλεπόμενου μεγέθους και των παραμέτρων που το επηρεάζουν και περιλαμβάνει τα μοντέλα παλινδρόμησης.

Στη συνέχεια περιγράφονται συνοπτικά οι στατιστικές μέθοδοι πρόβλεψης που θα αναλυθούν στην παρούσα εργασία, η απλοϊκή μέθοδος (Naive), οι μέθοδοι εκθετικής εξομάλυνσης SES (σταθερού επιπέδου), Holt (γραμμικής τάσης) και Damped (φθίνουσας τάσης), η κλασική μέθοδος Theta και μέθοδος διπλής εποχικότητας που αναπτύχθηκε από τον Taylor [Taylor, 2003b].

Η απλοϊκή μέθοδος (Naive) αποτελεί την πιο απλή στατιστική μέθοδο, της οποίας η πρόβλεψη προκύπτει ως η πραγματική παρατήρηση της ακριβώς προηγούμενης χρονικής περιόδου. Έχει καθιερωθεί η χρήση των αποτελεσμάτων της ως σημείο αναφοράς για τις υπόλοιπες μεθόδους.

Οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα δημοφιλείς από την αρχή κιόλας της εμφάνισής τους. Δίνουν μεγαλύτερη βαρύτητα στα πιο πρόσφατα δεδομένα της χρονοσειράς, καθώς θεωρούν ότι αυτά παρέχουν περισσότερη πληροφορία συγκριτικά με τα παλαιότερα δεδομένα της. Το μοντέλο σταθερού επιπέδου (SES) υποθέτει την απουσία τάσης από τα δεδομένα και παράγει προβλέψεις με οριζόντια προέκταση των δεδομένων. Το μοντέλο γραμμικής τάσης (Holt) αποτελεί προέκταση του μοντέλου SES, στο οποίο πλέον γίνεται διαχείριση της συνιστώσας τάσης που παρατηρείται στα δεδομένα πολλών χρονοσειρών. Η παραγωγή των προβλέψεων γίνεται με απλή προέκταση μιας ευθείας

γραμμής. Τέλος, το μοντέλο φθίνουσας τάσης (Damped) μειώνει σταδιακά την επίδραση της τάσης στα αποτελέσματα πρόβλεψης και ελέγχει το ρυθμό αύξησης των τιμών τους.

Η μέθοδος Theta αποτελεί μια μονοδιάστατη μέθοδο βασιζόμενη στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσω της παραμέτρου θ , που εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων. Οι παραγόμενες γραμμές, γραμμές Theta, προεκτείνονται ξεχωριστά και στη συνέχεια γίνεται συνδυασμός τους με βάρη (ίσα για την κλασική μέθοδο) για την παραγωγή της τελικής πρόβλεψης. Ανάλογα την τιμή της παραμέτρου θ , δίνεται μεγαλύτερη έμφαση στη μακροπρόθεσμη (μικρό θ) είτε στη βραχυπρόθεσμη (μεγάλο θ) συμπεριφορά των δεδομένων.

Τέλος, η μέθοδος διπλής εποχικότητας του Taylor αποτελεί προσαρμογή της μεθόδου των Holt-Winters, οι οποίοι είχαν ήδη αναπτύξει το μοντέλο εκθετικής εξομάλυνσης ώστε να ενσωματώνει την επίδραση μίας εποχικότητας. Στην εκδοχή του μοντέλου αυτού, χρησιμοποιείται το πρότυπο πολλαπλασιαστικής εποχικότητας με προσθετική τάση. Μέσω δύο νέων μεταβλητών εισάγεται η επιρροή των δύο ξεχωριστών εποχικότητων, ενώ γίνεται χρήση και ενός ακόμη όρου για προσαρμογή του σφάλματος μέσω μοντέλου AR (Autoregressive) και την επίτευξη καλύτερων αποτελεσμάτων.

Επιπλέον, υπάρχουν οι μέθοδοι γραμμικής παλινδρόμησης, αλλά και αρκετές ακόμη μέθοδοι πιο σύνθετες από αυτές που αναφέρθηκαν προηγουμένως και που χρησιμοποιούνται στη βραχυπρόθεσμη πρόβλεψη φορτίου ζητούμενης ηλεκτρικής ενέργειας. Μερικά από αυτά τα μοντέλα είναι τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου (ARIMA), τα τεχνητά νευρωνικά δίκτυα και διάφορες προσαρμογές των μεθόδων παλινδρόμησης.

Όπως έχουμε ήδη σημειώσει, μεγάλη σημασία έχει κατά την πρόβλεψη ζήτησης φορτίου η ενσωμάτωση των εποχικότητων της χρονοσειράς στα μοντέλα πρόβλεψης. Εκτός από τη μέθοδο του Taylor για διπλή εποχικότητα, στις υπόλοιπες μεθόδους εφαρμόζεται απλή διαδικασία αποσύνθεσης με χρήση κινητών κεντρικών μέσων όρων, ξεχωριστά μάλιστα για κάθε εποχικότητα (ημερήσια αρχικά με ΚΚΜΟ(24) και εβδομαδιαία σε δεύτερο στάδιο, με ΚΚΜΟ(168)).

Η αξιολόγηση των μεθόδων που εφαρμόζουμε στην παρούσα εργασία γίνεται με χρήση στατιστικών δεικτών σφάλματος και συγκεκριμένα επιλέγονται τα ποσοστιαία σφάλματα MAPE, sMAPE, MdAPE και sMdAPE, η ανάλυση των οποίων γίνεται στο κεφάλαιο 4 (παράγραφος 4.6). Γίνεται υπολογισμός των out-of-sample σφαλμάτων τους, όταν τα πραγματικά δεδομένα της ζήτησης είναι πλέον διαθέσιμα.

1.4 Μεθοδολογία

Βασικό τμήμα της παραγωγής προβλέψεων είναι η επιλογή και η ανάλυση των δεδομένων του μεγέθους που μελετάται. Στην εργασία μας, έχουμε διαθέσιμα τα ωριαία δεδομένα ζήτησης φορτίου (σε kWh) από μια ιδιωτική εταιρεία παραγωγής ηλεκτρικής ενέργειας, που δραστηριοποιείται στην Ελλάδα. Το χρονικό διάστημα για το οποίο έχει γίνει η καταγραφή των δεδομένων είναι από την 1/1/2009 ως και την 31/8/2011. Με γραφική απεικόνιση των δεδομένων, παρατηρείται η συμπεριφορά της χρονοσειράς ζήτησης και εντοπίζονται τα μοτίβα που επαναλαμβάνονται, καθώς και οι ασυνήθιστες τιμές των τιμών των δεδομένων.

Συγκεκριμένα, εντοπίζεται πολύ εύκολα η ημερήσια εποχικότητα των δεδομένων, καθώς ακολουθείται η ίδια συμπεριφορά ανά εικοσιτέσσερις παρατηρήσεις. Το επίπεδο των τιμών ενδέχεται να μεταβάλλεται από τη μία ημέρα στην άλλη, αλλά η εμφάνιση των μέγιστων και των ελάχιστων τιμών τους και η εναλλαγή από ανοδική σε καθοδική πορεία ή και στασιμότητα γίνονται με την ίδια σειρά και για τα αντίστοιχα χρονικά διαστήματα μέσα στην ημέρα. Ακόμη, εμφανής είναι και η εβδομαδιαία εποχικότητα, με αυξημένο επίπεδο ζήτησης φορτίου τις καθημερινές (Δευτέρα έως και Παρασκευή), ενώ παρατηρούνται μικρότερες τιμές το Σαββατοκύριακο. Αυτός ο διαχωρισμός είναι γενικός, αφού και στις καθημερινές εντοπίζεται τάση στα δεδομένα και επιπλέον η ζήτηση τα Σάββατα διαφέρει από την αντίστοιχη τις Κυριακές.

Από το πλήθος των διαθέσιμων παρατηρήσεων ζητούμενου φορτίου επιλέγεται να χρησιμοποιηθεί το 25% αυτών για την αξιολόγηση των μεθόδων πρόβλεψης που θα εφαρμοστούν στην πορεία. Επομένως, οι διαδικασίες που ακολουθούνται στην παρούσα εργασία περιλαμβάνουν την παραγωγή πρόβλεψης ζήτησης ηλεκτρικού φορτίου για τη χρονική περίοδο από 1/1/2011 ως και την 31/8/2011. Επί των δεδομένων αυτών υπολογίζονται τα ποσοστιαία σφάλματα που αναφέραμε στην παράγραφο 1.3 και κρίνεται η απόδοση των μοντέλων και των συνδυασμών που εξετάζονται.

Η διαδικασία παραγωγής προβλέψεων στο πρώτο μέρος της εργασίας μπορεί να περιγραφεί εν συντομία ως εξής:

- Εφαρμόζεται κλασική αποσύνθεση στα διαθέσιμα δεδομένα σε δύο στάδια, στο πρώτο με χρήση ΚΚΜΟ(24) και στο δεύτερο με ΚΚΜΟ(168), για την ημερήσια και την εβδομαδιαία εποχικότητα αντίστοιχα.
- Στην αποεποχικοποιημένη σειρά εφαρμόζεται το επιλεγμένο μοντέλο πρόβλεψης και καταγράφονται σε κάθε εκτέλεση οι προβλεπόμενες τιμές.
- Εποχικοποιούνται ξανά οι προβλεπόμενες τιμές για να ληφθεί το τελικό αποτέλεσμα, το οποίο είναι πιο κοντά στην πραγματικότητα.
- Επί της συνολικής προβλεπόμενης σειράς ζήτησης (για τη χρονική περίοδο 1/1/2011-31/8/2011) υπολογίζονται τα σφάλματα, με βάση τα οποία αξιολογούνται τα διαφορετικά μοντέλα πρόβλεψης.

Σημειώνουμε ότι το πρώτο βήμα δεν ακολουθείται όταν εξετάζεται η μέθοδος του Taylor, καθώς είναι η μοναδική μέθοδος που μελετάμε για διπλή εποχικότητα. Το ίδιο ισχύει και για το τρίτο βήμα.

Πολύ σημαντικό χαρακτηριστικό της διαδικασίας πρόβλεψης είναι η εφαρμογή της μεθόδου κύλισης των δεδομένων. Ο όγκος των δεδομένων που έχουμε στη διάθεσή μας είναι αρκετά μεγάλος και θεωρούμε ότι η χρήση όλων των παρελθοντικών δεδομένων για την παραγωγή των προβλέψεων μίας ημέρας ενδέχεται να εμποδίζει την αποδοτική προέκταση της χρονοσειράς αντί να διευκολύνει τη διαδικασία (για την πρώτη πρόβλεψη, της 1/1/2011, χρησιμοποιούνται με τη νοσοτροπία αυτή δεδομένα χρονικής περιόδου 2 ετών). Αποφασίζουμε λοιπόν, να εφαρμόσουμε κυλιόμενη πρόβλεψη σε κάθε βήμα της εργασίας μας, ώστε να χρησιμοποιούνται για την πρόβλεψη κάθε ημερήσιας ζήτησης (24-τιμών) τα πιο πρόσφατα πραγματικά δεδομένα ζήτησης, που περιέχουν περισσότερο σημαντική πληροφορία.

Για την εφαρμογή της κυλιόμενης πρόβλεψης επιλέγεται αρχικά ένα χρονικό διάστημα, το οποίο παραμένει σταθερό καθ' όλη τη διαδικασία της πρόβλεψης. Για κάθε τρέχον εικοσιτετράωρο, την ζήτηση του οποίου θέλουμε να προβλέψουμε, χρησιμοποιούνται τα προηγούμενα (διαθέσιμα πλέον) πραγματικά δεδομένα, πλήθους ίσου με αυτό που αντιστοιχεί στο επιλεγμένο χρονικό διάστημα. Επομένως, με την εναλλαγή των τρεχόντων εικοσιτετράωρων ανανεώνονται και τα δεδομένα που χρησιμοποιούνται για τον υπολογισμό του in-sample σφάλματος και των συντελεστών των μεθόδων και η παραγωγή προβλέψεων μπορεί να ενσωματώνει πιο αποτελεσματικά τις μεταβολές του περιβάλλοντος (και οι οποίες επηρεάζουν τη ζήτηση φορτίου ηλεκτρικής ενέργειας). Είναι λοιπόν πολύ σημαντική η επιλογή του χρονικού διαστήματος εφαρμογής τα μεθόδου, ώστε να χρησιμοποιηθεί ο βέλτιστος αριθμός παρατηρήσεων για την παραγωγή προβλέψεων. Αναλυτικά παρουσιάζεται η μεθοδολογία του rolling forecasting στο κεφάλαιο 5 (παράγραφος 5.3).

Όπως αναφέρθηκε προηγουμένως, η αξιολόγηση των μεθόδων που χρησιμοποιούνται γίνεται με χρήση στατιστικών δεικτών σφάλματος και συγκεκριμένα από την κατηγορία των ποσοστιαίων σφαλμάτων. Αυτά είναι τα εξής: μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE), συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (sMAPE), η διάμεσος του απόλυτου ποσοστιαίου σφάλματος (MdAPE) και η διάμεσος του συμμετρικού απόλυτου ποσοστιαίου σφάλματος (sMdAPE). Η επιλογή των σφαλμάτων είναι και αυτή σημαντική, αφού κάθε σφάλμα αποδίδει καλύτερα σε διαφορετικής φύσης χρονοσειρές. Καθώς οι τιμές των παρατηρήσεων του ζητούμενου φορτίου είναι πολύ υψηλές, η χρήση του MAPE δεν παρουσιάζει μειονεκτήματα, εκτός από την έλλειψη συμμετρίας, για την οποία και υπολογίζεται το sMAPE. Ακόμη, γίνεται χρήση των διαμέσων των σφαλμάτων, MdAPE και sMdAPE, επειδή είναι πιο σταθερά απέναντι σε ανωμαλίες των τιμών της χρονοσειράς. Ο υπολογισμός των σφαλμάτων γίνεται επί όλων των προβλεπόμενων τιμών για κάθε εκτέλεση κάθε μεθόδου (για τους μήνες Ιανουάριο ως και Αύγουστο του 2011).

1.5 Αποτελέσματα Προβλέψεων

Έπειτα από την ανάλυση των δεδομένων και τον καθορισμό της μεθοδολογίας που ακολουθείται για την παραγωγή των προβλέψεων, υπολογίζονται τα σφάλματα και παρουσιάζονται συγκεντρωτικά, ώστε να ληφθούν τα απαραίτητα συμπεράσματα. Συγκεκριμένα, παρουσιάζονται στο κεφάλαιο 6 και τα τέσσερα ποσοστιαία σφάλματα που αναφέρθηκαν προηγουμένως σε πίνακες (πίνακες 2 έως και 5). Για καλύτερη αντίληψη του επιπέδου των σφαλμάτων και ειδικά για την πιο αποτελεσματική σύγκριση μεταξύ των μεθόδων, γίνεται γραφική απεικόνιση των σφαλμάτων MAPE (εικόνες 7, 8 και 11) και MdAPE (εικόνες 9, 10 και 12). Επειδή τα αντίστοιχα συμμετρικά σφάλματα ακολουθούν την ίδια πορεία, δεν κρίνεται αναγκαία η απεικόνιση και αυτών γραφικά.

Τα αποτελέσματα παρουσιάζονται για τα διαφορετικά χρονικά διαστήματα εφαρμογής του rolling forecasting και για τα δύο στάδια αποεποχικοποίησης (για τις απλές μεθόδους). Σημαντική διαφορά επιπέδου τιμών σημειώνεται ανάμεσα στα σφάλματα MAPE και sMAPE από τη μία πλευρά και στα σφάλματα MdAPE και sMdAPE από την άλλη. Ακόμη, είναι εμφανής η διαφορά της ακρίβειας των απλών μεθόδων για τις δύο ξεχωριστές εποχικότητες των δεδομένων. Το επίπεδο των σφαλμάτων είναι πολύ μικρότερο στην περίπτωση της ενσωμάτωσης της εβδομαδιαίας εποχικότητας, ενώ για ημερήσια εποχικότητα τα σφάλματα είναι υψηλά.

Βασικό συμπέρασμα είναι η πολύ καλή απόδοση των απλών μεθόδων πρόβλεψης. Μάλιστα, βέλτιστα αποτελέσματα παράγει η απλούστερη όλων των μεθόδων, Naive. Αυτό

το συμπέρασμα προκύπτει από ανάλυση των σφαλμάτων στο δεύτερο στάδιο, με τη συμπερίληψη της επίδρασης της εβδομαδιαίας εποχικότητας, γιατί στο πρώτο στάδιο (με την ημερήσια εποχικότητα) η Naive είναι η δεύτερη χειρότερη μέθοδος μετά τη Holt (συγκρίνοντας τις απλές μεθόδους). Τα απλά μοντέλα των μεθόδων SES, Damped και Theta παρουσιάζουν παρόμοια αποτελέσματα, ενώ η μέθοδος γραμμικής τάσης Holt χαρακτηρίζεται από μεγαλύτερα σφάλματα. Αναμέναμε ότι η αισιοδοξία που διακρίνει τη μέθοδο Holt θα οδηγούσε σε λίγο μεγαλύτερα σφάλματα από ό, τι οι υπόλοιπες μέθοδοι εξομάλυνσης. Ωστόσο, η κατάταξη της μεθόδου του Taylor για διπλή εποχικότητα ως τελευταία, για σύγκριση με πολύ απλές μεθόδους, είναι ένα γεγονός που προκαλεί εντύπωση.

Συγκεκριμένα για τη μέθοδο διπλής εποχικότητας, μπορούμε να αναφέρουμε ότι παρουσιάζει μεγάλη διακύμανση στα σφάλματα των επαναλήψεων της με διαφορετικό χρονικό διάστημα κύλισης των δεδομένων. Επίσης, λαμβάνοντας υπόψη τη μη ικανοποιητική συμπεριφορά των απλών μοντέλων στο στάδιο ενσωμάτωσης της ημερήσιας εποχικότητας και γνωρίζοντας ότι στην παραγωγή κάθε προβλεπόμενης τιμής η μέθοδος του Taylor συμπεριλαμβάνει και τις δύο εποχικότητες, μπορεί να εξηγηθεί η χαμηλή απόδοσή της.

Εντύπωση προκαλεί η παραγωγή των ίδιων ακριβώς προβλέψεων σε κάθε εφαρμογή (για όλα τα διαφορετικά χρονικά διαστήματα) της κυλιόμενης μεθόδου στο στάδιο της ημερήσιας εποχικότητας. Ακόμη, να σημειώσουμε ότι τα βέλτιστα σφάλματα δεν εντοπίζονται για το ίδιο χρονικό διάστημα εφαρμογής rolling forecasting σε όλες τις μεθόδους πρόβλεψης. Τέλος, η αξιολόγηση με βάση το MAPE οδηγεί ορισμένες φορές σε διαφορετικά συμπεράσματα από την αντίστοιχη αξιολόγηση με τη διάμεσο MdAPE, όπως το χρονικό διάστημα επίτευξης των βέλτιστων σφαλμάτων τους σε μερικές από αυτές.

Για να δούμε την πορεία των σφαλμάτων των μεθόδων κατά τη διάρκεια του εικοσιτετραώρου, γίνεται παρουσίαση του σφάλματος MAPE σε εικοσιτετράωρη βάση για τις καλύτερες περιπτώσεις όλων των μεθόδων (εικόνα 16). Παρατηρείται μεγάλη αύξηση του σφάλματος για όλες τις μεθόδους από την έναρξη του πρωινού ωραρίου εργασίας, με τις μέγιστες τιμές του να σημειώνονται στο χρονικό διάστημα 15:00 – 17:00 το απόγευμα, ενώ στη συνέχεια ακολουθεί πτωτική πορεία. Και στην παρουσίαση αυτή, είναι φανερό η μεγαλύτερη ακρίβεια της απλοϊκής μεθόδου και η μη ικανοποιητική απόδοση της μεθόδου διπλής εποχικότητας (ως η μέθοδος με τις υψηλότερες τιμές σφάλματος).

Τέλος, γίνεται παρουσίαση του ποσοστού των ωριαίων παρατηρήσεων που προβλέψαμε με τα επιλεγμένα μοντέλα, στις οποίες το σφάλμα MAPE προέκυψε μικρότερο του 3% (πίνακας 8). Και σε αυτό τον πίνακα φαίνεται ότι η απλοϊκή μέθοδος (Naive) είναι η πλέον αποδοτική από όσες μελετήσαμε, με σημαντικά μεγαλύτερο επίπεδο ποσοστού (καθώς ο υπολογισμός γίνεται σε κάθε μέθοδο για όλες τις εκτελέσεις της, για όλα τα χρονικά διαστήματα κύλισης και για τα δύο στάδια εποχικότητας δηλαδή). Παρατηρείται εδώ ότι η μέθοδος διπλής εποχικότητας έχει μεγάλο ποσοστό σφαλμάτων μικρότερων του 3% και η μικρότερη τιμή του είναι αρκετά μεγαλύτερη από τα κάτω όρια των ποσοστών των υπόλοιπων απλών μεθόδων (εκτός από τη Naive).

Με βάση λοιπόν τα αποτελέσματα του πρώτου μέρους της μελέτης μας, επιβεβαιώνεται ότι είναι πολύ σημαντική η καλή προσαρμογή των μοντέλων πρόβλεψης που διατίθενται, για την ενσωμάτωση των εποχικότητων των δεδομένων ζήτησης φορτίου. Ακόμη, αποδεικνύεται ότι οι απλές μέθοδοι πρόβλεψης της εργασίας αυτής, με την ενσωμάτωση της εβδομαδιαίας εποχικότητας, παράγουν καλύτερα αποτελέσματα από τη μέθοδο διπλής εποχικότητας.

1.6 Επίδραση Ειδικών Ημερών

Στο δεύτερο μέρος της παρούσας μελέτης εξετάζεται η ενσωμάτωση ενός ακόμη εποχικού παράγοντα που επηρεάζει τη ζήτηση φορτίου ηλεκτρικής ενέργειας, τις ειδικές ημέρες. Η παρατήρηση της γραφικής απεικόνισης των δεδομένων, αναδεικνύει την ύπαρξη έντονων διακυμάνσεων των παρατηρήσεων της χρονοσειράς ζήτησης. Πρόκειται για ημέρες ειδικών αργιών της Ελλάδας, όπως και ημέρες στις οποίες πραγματοποιούνται πανελλαδικές πανεργατικές απεργίες. Η επίδραση των ειδικών ημερών στα δεδομένα ζήτησης, κυμαίνεται από λίγες ώρες ως και μερικές ημέρες πριν και μετά από αυτές τις ημέρες.

Αρχικά καταγράφονται οι ημέρες αργιών για ένα ημερολογιακό έτος, τόσο αυτές που αναφέρονται ως επίσημες αργίες (εθνικές εορτές και θρησκευτικές εορτές), όσο και ορισμένες ακόμη που θεωρούμε εμείς, σύμφωνα και με τις μεταβολές των δεδομένων της χρονοσειράς (πίνακες 9 και 10). Επίσης, καταγράφονται οι ημέρες στις οποίες πραγματοποιήθηκαν πανελλαδικές πανεργατικές απεργίες στη διάρκεια του χρονικού διαστήματος, για το οποίο διαθέτουμε δεδομένα (πίνακας 11). Τα δεδομένα αυτά παρέχονται από τα αρχεία της Γενικής Συνομοσπονδίας Εργατών Ελλάδας (ΓΣΕΕ).

Στη συνέχεια, καθορίζεται ο τρόπος ενσωμάτωσης της επίδρασης των ειδικών ημερών που καταγράφηκαν στην παραγωγή των προβλέψεων. Εξετάζονται τρεις διαφορετικοί τρόποι, βασισμένοι στην ποσοστιαία μεταβολή ζήτησης που υπολογίζεται για τα ωριαία δεδομένα των ειδικών ημερών συγκριτικά με τις αντίστοιχες τιμές της προηγούμενης ημέρας και της προηγούμενης εβδομάδας (πλέον αναφερόμαστε σε δεδομένα της πραγματικής χρονοσειράς ζήτησης). Οι ποσοστιαίες αυτές μεταβολές υπολογίζονται για τις ειδικές ημέρες των δύο πρώτων ετών, ενώ η εφαρμογή τους γίνεται επί των ειδικών ημερών του 2011. Να σημειώσουμε εδώ ότι δε γίνεται πρόβλεψη ολόκληρης της χρονοσειράς για την περίοδο του 2011 που εξετάζουμε, αλλά εφαρμόζονται τα ποσοστά που αναφέραμε επί των παρατηρήσεων της πραγματικής χρονοσειράς και μόνο για τις ειδικές ημέρες που έχουμε καταγράψει. Με τους τρεις τρόπους, που θα περιγράψουμε αμέσως μετά, υπολογίζονται οι τελικές προβλεπόμενες τιμές της ζήτησης φορτίου για τις ειδικές ημέρες και στη συνέχεια αντικαθιστώνται οι τιμές των ειδικών ημερών των χρονοσειρών που είχαμε παράγει στο πρώτο μέρος της εργασίας με αυτές τις νέες τιμές και προκύπτει η τελική πρόβλεψη της χρονοσειράς ζήτησης για το διάστημα 1/1/2011-31/8/2011. Σύμφωνα με τον πρώτο τρόπο, η προβλεπόμενη ωριαία τιμή ζήτησης σε ειδικές ημέρες ισούται με την αντίστοιχη πραγματική ωριαία τιμή ζήτησης της προηγούμενης ημέρας, στην οποία έχει εφαρμοστεί το ποσοστό μεταβολής συγκριτικά με την προηγούμενη ημέρα, ο δεύτερος με την αντίστοιχη σχέση για δεδομένα και ποσοστό βάσει της προηγούμενης εβδομάδας, ενώ ο τρίτος τρόπος προκύπτει ως ο μέσος όρος των δύο πρώτων.

Εκτός από τη μελέτη διαφορετικών τρόπων παραγωγής των τελικών προβλέψεων, μελετάμε και διαφορετικούς συνδυασμούς των ειδικών ημερών, για να εντοπίσουμε το βέλτιστο συνδυασμό όλων. Οι συνδυασμοί αυτοί χωρίζονται σε τρεις κατηγορίες: ενσωμάτωση μόνο των αργιών, ενσωμάτωση μόνο των απεργιών και ενσωμάτωση συνδυασμού αυτών. Η πρώτη κατηγορία έχει δύο περιπτώσεις, όπου στην πρώτη λαμβάνονται υπόψη τα ποσοστά (που αναφέραμε προηγουμένως) των επίσημων αργιών και στη δεύτερη τα ποσοστά όλων των αργιών (και των δύο πινάκων 9 και 10). Τα ποσοστά αυτά υπολογίζονται και στις δύο περιπτώσεις για τις αργίες των πρώτων δύο ετών και στη συνέχεια υπολογίζεται ο μέσος όρος τους (για κάθε αργία από τα δύο έτη), ο οποίος και χρησιμοποιείται εν τέλει

Η δεύτερη κατηγορία έχει τρεις περιπτώσεις. Αρχικά, υπολογίζονται τα ποσοστά όλων των απεργιών των δύο ετών. Στη συνέχεια, στην πρώτη περίπτωση χρησιμοποιείται τελικά ο

μέσος όρος τους, στη δεύτερη περίπτωση υπολογίζεται ο μέσος όρος των ποσοστών για ίδιες ημέρες (πχ. πραγματοποιήθηκαν τα πρώτα δύο έτη τρεις απεργίες ημέρα Τρίτη και υπολογίζεται το τελικό ποσοστό ως ο μέσος όρος των τριών ποσοστών αυτών των ημερών) και στην τρίτη περίπτωση δεν μας ενδιαφέρει κανένας μέσος όρος ποσοστών, αλλά χρησιμοποιούμε τις ποσοστιαίες μεταβολές για αντίστοιχες ημέρες (παραπλήσιες τιμές εβδομαδιαία και σε ίδιους μήνες).

Τέλος, στην τρίτη κατηγορία γίνεται μελέτη όλων των δυνατών συνδυασμών των δύο παραπάνω κατηγοριών, στη εργασία ωστόσο παρουσιάζονται τα αποτελέσματα μόνο του καλύτερου συνδυασμού, δηλαδή της πρώτης περίπτωσης των αργιών με την πρώτη περίπτωση των απεργιών. Όμοια και για την δεύτερη κατηγορία, γίνεται παρουσίαση των σφαλμάτων μόνο της πρώτης μεθόδου με το μέσο όρο των ποσοστιαίων μεταβολών όλων των απεργιών.

Αναλυτικά η περιγραφή της ενσωμάτωσης των ειδικών ημερών, όσο και τα αποτελέσματα των προβλέψεων που προκύπτουν με την παραπάνω μεθοδολογία, παρουσιάζονται στο κεφάλαιο 7. Τα αποτελέσματα αυτά επιβεβαιώνουν την καλύτερη απόδοση που εμφανίζουν οι μέθοδοι μιας εποχικότητας συγκριτικά με τη μέθοδο του Taylor. Ακόμη, η απλοϊκή μέθοδος παραμένει και πάλι η καλύτερη όλων (σύμφωνα με τα σφάλματα που βασίζονται στην επίδραση της εβδομαδιαίας εποχικότητας, γιατί τα σφάλματά της για ημερήσια εποχικότητα συγκαταλέγονται και πάλι στα δύο χειρότερα, μαζί με αυτά της μεθόδου Holt). Αρχικά να σημειωθεί ότι ο καλύτερος τρόπος ενσωμάτωσης των ποσοστιαίων μεταβολών των ειδικών ημερών είναι ο ίδιος για όλες τις μεθόδους σε κάθε συνδυασμό που μελετάται. Συγκεκριμένα, για τις περιπτώσεις ενσωμάτωσης των επίσημων αργιών, των επίσημων και μη αργιών μαζί και τον συνδυασμό των αργιών και απεργιών, οι μέθοδοι είναι πιο αποδοτικές όταν χρησιμοποιείται η ποσοστιαία μεταβολή της προηγούμενης ημέρας και στην περίπτωση της ενσωμάτωσης του μέσου όρου όλων των απεργιών καλύτερες προβλέψεις παράγονται με τη χρήση του μέσου όρου των ποσοστιαίων μεταβολών προηγούμενης ημέρας και εβδομάδας μαζί.

Η μείωση των σφαλμάτων στο σύνολο των δεδομένων της χρονικής περιόδου 1/1/2011-31/8/2011 είναι μικρή. Ωστόσο παρουσιάζεται για όλες τις εκτελέσεις των προβλέψεων, με κάθε δυνατό συνδυασμό και, έχοντας υπόψη ότι πρόκειται για μεταβολή που προκύπτει από την αλλαγή των προβλέψεων μόλις σε 5 έως 15 ημερών από τις 243 της περιόδου αυτής, συμπεραίνουμε πως δεν είναι αδιάφορη η επίδραση των ειδικών ημερών στη ζήτηση του φορτίου ηλεκτρικής ενέργειας. Αυτό το συμπέρασμα ενισχύεται από τα αποτελέσματα του πίνακα 28 (παράγραφος 7.4), στον οποίο παρουσιάζονται ενδεικτικά τα καλύτερα σφάλματα μόνο για τις ειδικές ημέρες, για όλους τους συνδυασμούς προβλέψεων που υλοποιήθηκαν με τη μέθοδο Naive και η ποσοστιαία μείωση που παρατηρείται από τα σφάλματα της αρχικής πρόβλεψης (κεφάλαιο 5) επί των ειδικών ημερών στα νέα σφάλματα με την εισαγωγή των ποσοστιαίων μεταβολών.

Ακόμη, παρουσιάζονται και πάλι (πίνακας 29) τα ποσοστά των ωριαίων προβλέψεων με σφάλμα (MAPE και sMAPE) συγκεντρωτικά για όλες τις περιπτώσεις παραγωγής προβλέψεων. Φαίνεται εδώ ότι τα μεγαλύτερα ποσοστά παρατηρούνται για την ενσωμάτωση της ποσοστιαίας μεταβολής όλων των αργιών.

Τέλος, στον πίνακα 30 παρουσιάζονται τα καλύτερα σφάλματα των μοντέλων πρόβλεψης για κάθε συνδυασμό πρόβλεψης που υλοποιήθηκε. Η παρουσίαση γίνεται από τον καλύτερο προς το χειρότερο συνδυασμό και όμοια είναι ταξινομημένες οι μέθοδοι σε κάθε συνδυασμό. Προκύπτει ότι το καλύτερο επίπεδο σφαλμάτων (μικρότερα σφάλματα)

προκύπτει για την ενσωμάτωση της ποσοστιαίας μεταβολής των επίσημων αργιών, όπως και για την ενσωμάτωση της ποσοστιαίας μεταβολής όλων των αργιών. Η διαφορά του επιπέδου σφαλμάτων για τις δύο περιπτώσεις είναι πολύ μικρή, με το MdAPE να υποδεικνύει μικρή βελτίωση για την περίπτωση όλων των αργιών. Ενώ η καλύτερη απόδοση ανάμεσα σε όλες τις εκτελέσεις, σημειώνεται για τη μέθοδο Naive και για την περίπτωση της ενσωμάτωσης της ποσοστιαίας μεταβολής όλων των αργιών.

1.7 Συμπεράσματα και Προοπτικές

Η βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας είναι πολύ σημαντική υπόθεση και απαιτείται η περαιτέρω προσαρμογή επί των μοντέλων που χρησιμοποιούνται, ώστε να επιτευχθεί όσο το δυνατόν πιο ακριβής πρόβλεψή της. Με τη μελέτη που γίνεται στα πλαίσια της εν λόγω διπλωματικής εργασίας, αποδεικνύεται η πολύ καλή απόδοση των απλών μεθόδων πρόβλεψης, με την ενσωμάτωση της εβδομαδιαίας εποχικότητας στα μοντέλα αυτά και την εφαρμογή της κύλισης των δεδομένων, ώστε να χρησιμοποιείται σε κάθε πρόβλεψη η πιο σημαντική πληροφορία των παρελθοντικών δεδομένων της πραγματικής χρονοσειράς. Ακόμη, επιβεβαιώνεται η σημασία των ειδικών ημερών στη διαμόρφωση της ζήτησης του φορτίου και συνίσταται η ενσωμάτωσή τους στην παραγωγή προβλέψεων βραχυπρόθεσμης ζήτησης φορτίου.

Σαφώς όμως, υπάρχουν αρκετά σημεία τα οποία μπορούν να δεχτούν βελτίωση. Αυτά συνοπτικά είναι τα παρακάτω:

- ✓ Βελτίωση της ενσωμάτωσης των εποχικότητων των δεδομένων ζήτησης φορτίου. Μελέτη συνδυασμών αυτών με άνισα βάρη, μέχρι να εντοπιστεί ο βέλτιστος.
- ✓ Ενσωμάτωση μετεωρολογικών παραγόντων και κυρίως της θερμοκρασίας στα μοντέλα πρόβλεψης. Εξέταση της ενσωμάτωσής της με διάφορους συνδυασμούς (ακραίες τιμές, συνδυασμός παρελθοντικών τιμών θερμοκρασίας, κλπ).
- ✓ Προσπάθεια ενσωμάτωσης των οικονομικών παραγόντων στη διαδικασία πρόβλεψης (ιδιαίτερα πολύπλοκη υπόθεση, ίσως με εξελιγμένα μοντέλα πολλών εισόδων και εφόσον πρώτα έχει επιτευχθεί αποτελεσματική ενσωμάτωση της θερμοκρασίας).
- ✓ Εξέταση των προσαρμογών που μπορεί να δεχτεί το μοντέλο διπλής εποχικότητας του Taylor, ενδεχομένως με διαφορετικά βάρη επί των συντελεστών των δύο διαφορετικών εποχικότητων.

ΚΕΦΑΛΑΙΟ 2: Εισαγωγή

2.1 Σημασία της πρόβλεψης φορτίου

Η πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας είναι ένας τομέας της επιστήμης των προβλέψεων, που προϋπάρχει εδώ και μερικές δεκαετίες, όμως έντονη ενασχόληση με αυτόν παρατηρείται μόλις τα τελευταία χρόνια. Η πρόβλεψη ζήτησης φορτίου είναι ιδιαίτερης σημασίας για τις εταιρείες παραγωγής ηλεκτρικής ενέργειας, καθώς έχει διάφορες εφαρμογές με βασικότερη τον οικονομικό προγραμματισμό του συστήματος παραγωγής. Κατέχοντας την πρόβλεψη ζήτησης ηλεκτρικού φορτίου για την επόμενη ημέρα έως και εβδομάδα, μπορεί η εταιρεία να διαχειριστεί τα υπάρχοντα συστήματα ηλεκτρικής ενέργειας (ΣΗΕ) και να προγραμματίσει τη βέλτιστη ένταξη μονάδων στο δίκτυο και την ανταλλαγή ηλεκτρικής ενέργειας με γειτονικά δίκτυα. Έχοντας υπόψη ότι βασικό χαρακτηριστικό της ηλεκτρικής ενέργειας είναι η δυσκολία αποθήκευσής της, καταλαβαίνει κανείς πόσο σημαντική είναι η διατήρηση της παραγωγής και της ζήτησης ηλεκτρικής ενέργειας στο ίδιο επίπεδο.

Η γνώση της προβλεπόμενης ζήτησης φορτίου διευκολύνει ακόμη το έργο των χειριστών του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ). Με βάση την πρόβλεψη αυτή αποφασίζουν ποιες διαδικασίες θα ακολουθήσουν για τη διόρθωση της παραγωγής, από την ένταξη μονάδων, την αγορά επιπλέον ενέργειας από γειτονικά δίκτυα, τη ρύθμιση της τάσης των ζυγών για την επόμενη ημέρα -ως και εβδομάδα, ώστε να υπάρχει η απαιτούμενη εφεδρεία για την οικονομική λειτουργία και την εξασφάλιση της κάλυψης της ζήτησης, σε περιπτώσεις σφαλμάτων και ανωμαλιών του συστήματος. Σε διαφορετική περίπτωση, η εταιρεία παραγωγής ενέργειας αντιμετωπίζει ιδιαίτερα δαπανηρές επιλογές δράσης. Συγκεκριμένα, αν υποτιμηθεί αρκετά η μελλοντική ζήτηση φορτίου, η εταιρεία θα προβεί στην ενσωμάτωση μονάδων αιχμής στο δίκτυο για την κάλυψη της επιπλέον ζήτησης, εκτός του ότι κινδυνεύει η αξιοπιστία του συστήματος και ενδέχεται να παρατηρηθούν διακοπές ενέργειας. Ενώ σε περίπτωση έντονης υπερτίμησης της ζήτησης, θα έχει παραχθεί φορτίο χωρίς λόγο και αγοραστικό κοινό και έτσι οδηγείται και πάλι σε αύξηση του λειτουργικού κόστους. Επίσης, η κατοχή προβλέψεων ακριβείας για το συνδυασμό ζήτησης και τιμής του φορτίου δίνουν στις εταιρείες τη δύναμη να αναπτύξουν προσοδοφόρα πολιτική του προσφερόμενου φορτίου τους απέναντι στις ανταγωνίστριες εταιρείες και στον καταναλωτή, να εξυπηρετούν μεγαλύτερο μέρος της αγοράς και να μεγιστοποιήσουν τα κέρδη τους.

Όπως αναφέρθηκε και παραπάνω, ο επιστημονικός κόσμος έχει παρουσιάσει ιδιαίτερο ενδιαφέρον για την πρόβλεψη ζήτησης φορτίου μόλις τα τελευταία χρόνια. Παλαιότερα δεν υπήρχε λόγος για έντονη ενασχόληση με τον τομέα αυτό των προβλέψεων, καθώς συγκεκριμένοι οργανισμοί αναλάμβαναν την παραγωγή ηλεκτρικής ενέργειας σε κάθε χώρα. Στην αρχή της δεκαετίας του '90 όμως ξεκίνησε η απελευθέρωση των αγορών ενέργειας σε πολλές χώρες, με αποτέλεσμα την κατάλυση των ισχυρών μονοπωλίων και την ύπαρξη πλέον ανταγωνιστικών αγορών με πολλές εταιρείες παραγωγής ενέργειας. Αυτή η νέα κατάσταση οδήγησε και στην έντονη ανάγκη για σωστή και όσο το δυνατόν πιο ακριβή πρόβλεψη της ζήτησης φορτίου ηλεκτρικής ενέργειας, για τον αποτελεσματικό προγραμματισμό και έλεγχο των ΣΗΕ (όπως προαναφέραμε), όπως και της τιμής της ενέργειας, ώστε να αποφασίζει η εκάστοτε εταιρεία την επόμενη κίνησή της για να επιβιώσει ή ακόμη και να ηγηθεί του ανταγωνισμού της αγοράς. Επιπλέον, η κατανάλωση ενέργειας με το πέρασμα του χρόνου αυξήθηκε σημαντικά, εφόσον στα νοικοκυριά όπως και στις υπηρεσίες ενσωματώθηκαν ηλεκτρικές συσκευές που δεν υπήρχαν στο παρελθόν (ηλεκτρονικός υπολογιστής, κλιματιστικό, κλπ) και η χρήση τους εξαρτάται από την χρονική

περίοδο που βρισκόμαστε μέσα στη διάρκεια του ημερολογιακού έτους και άλλους παράγοντες. Στην Ελλάδα, η απελευθέρωση της αγοράς ηλεκτρικής ενέργειας οδήγησε στην κατάργηση του αποκλειστικού προνομίου που κατείχε η Δημόσια Εταιρεία Ηλεκτρισμού (ΔΕΗ) το Δεκέμβριο του 1999. Πλέον από το Φεβρουάριο του 2001 οι πελάτες που συνδέονται στην Υψηλή και τη Μέση Τάση έχουν το δικαίωμα επιλογής του προμηθευτή τους. Έτσι αποκτά σημασία και στη χώρα μας η πρόβλεψη ζήτησης φορτίου ηλεκτρικής ενέργειας και παρατηρείται ενασχόληση πολλών Ελλήνων ερευνητών με το αντικείμενο.

Στα πλαίσια αυτά κινείται η μελέτη της εν προκειμένω διπλωματικής εργασίας, η οποία βασίζεται στα στοιχεία ζήτησης ενέργειας μιας ιδιωτικής εταιρείας παροχής ηλεκτρισμού στην ελληνική αγορά

Συγκεκριμένα υπάρχουν τέσσερις κατηγορίες πρόβλεψης φορτίου, οι οποίες κατηγοριοποιούνται σύμφωνα με τον χρονικό ορίζοντά τους και είναι οι εξής:

- Η μακροπρόθεσμη πρόβλεψη φορτίου έχει ορίζοντα από ένα ως και αρκετά έτη, χρησιμεύει στο σχεδιασμό και τη συντήρηση των συστημάτων παραγωγής και διανομής ενέργειας. Είναι ακόμη βασικό κριτήριο, στο οποίο στηρίζονται οι εταιρείες για αποφάσεις σχετικά με μεγάλες επενδύσεις αγοράς εγκαταστάσεων και νέων μονάδων παραγωγής.

- Η μεσοπρόθεσμη πρόβλεψη φορτίου, με ορίζοντα από ένα μήνα ως και τρία έτη, χρειάζεται κατά κύριο λόγο στη ρύθμιση προγραμμάτων συντήρησης των διαθέσιμων ΣΗΕ, καθώς και στο χειρισμό και το σχεδιασμό τους.

- Η βραχυπρόθεσμη πρόβλεψη, με ορίζοντα από μερικές ώρες ως και μία εβδομάδα, απαιτείται για τον έλεγχο και τον προγραμματισμό των συστημάτων ηλεκτρικής ενέργειας, για την εκτίμηση της ασφάλειάς τους (για παράδειγμα καθορίζει τη σειρά ένταξης των μονάδων παραγωγής του ΣΗΕ). Η βραχυπρόθεσμη πρόβλεψη είναι επίσης απαραίτητη από εταιρείες μετάδοσης ενέργειας, στην περίπτωση λειτουργίας self-dispatching αγορών. Σε ορισμένες τέτοιες αγορές χρησιμεύει στην εξισορρόπηση της παροχής με τη ζήτηση, έπειτα από την παράπλευρη συναλλαγή ανάμεσα στους παραγωγούς και τους προμηθευτές της ενέργειας, ώστε να εξασφαλίζεται η αξιόπιστη λειτουργία των συστημάτων ηλεκτρικής ενέργειας(ΣΗΕ).

- Η πολύ βραχυπρόθεσμη πρόβλεψη, με ορίζοντα από λίγα λεπτά ως και μία ώρα, χρησιμεύει για την κατανομή φορτίου στις μονάδες του ΣΗΕ κατά το βέλτιστο και πλέον οικονομικό τρόπο. Συμβάλλει επομένως στην κάλυψη των αναγκών του συστήματος των γεννητριών και με την άμεση παρακολούθηση της συχνότητας με το φορτίο, εξασφαλίζεται καλύτερα η ασφάλεια του συστήματος. Ακόμη, μπορεί να συμβάλει σημαντικά στη ρύθμιση της τιμής της ηλεκτρικής ενέργειας.

2.2 Παράγοντες που επηρεάζουν τη ζήτηση φορτίου ηλεκτρισμού

Στη μελέτη αυτή μας ενδιαφέρει η επίτευξη καλής προσέγγισης της ωριαίας ζήτησης φορτίου ηλεκτρισμού, με ορίζοντα μία ημέρα. Έχει επομένως σημασία να καταγράψουμε τους παράγοντες που επηρεάζουν την πορεία της ζήτησης φορτίου και να κρίνουμε ποιους θεωρούμε απαραίτητους να συμπεριληφθούν στις μεθόδους που θα εξετάσουμε στην πορεία. Παραθέτουμε λοιπόν τους σημαντικότερους παράγοντες που εντοπίσαμε και είναι εποχικοί, καιρικοί και οικονομικοί.

Ως εποχικούς παράγοντες θεωρούμε τις εναλλαγές των εποχών κατά τη διάρκεια του έτους, τις εποχικότητες που παρουσιάζουν τα δεδομένα και τις ειδικές ημέρες που μεταβάλλονται ανάλογα με τη χώρα και τις συνθήκες που επικρατούν σε αυτή. Πιο συγκεκριμένα, μέσα σε ένα ημερολογιακό έτος υπάρχει η διαδοχή των τεσσάρων εποχών (χειμώνας, άνοιξη, καλοκαίρι και φθινόπωρο). Η εναλλαγή των εποχών προκαλεί μεταβολή και στην ένταση των ανθρώπινων δραστηριοτήτων που απαιτούν ηλεκτρική ενέργεια, όπως για παράδειγμα στην Ελλάδα το καλοκαίρι αυξάνεται η ζήτηση ενέργειας λόγω της ύπαρξης μεγάλου πλήθους τουριστών, των οποίων οι ανάγκες πρέπει να ικανοποιηθούν. Παρατηρείται έτσι αλλαγή στο επίπεδο ζήτησης της ηλεκτρικής ενέργειας, η οποία παρουσιάζει διαφορετική μέση τιμή και τιμή αιχμής (πολύ σημαντικές για τη σωστή λειτουργία των ΣΗΕ) σε κάθε εποχή. Έχοντας μεγάλο πλήθος δεδομένων μπορεί να κατανοήσει κανείς την εποχικότητα που παρουσιάζουν τα δεδομένα ζήτησης φορτίου σε ετήσια βάση. Ακόμη, η ζήτηση ηλεκτρικού φορτίου παρουσιάζει έντονη εποχικότητα, τόσο ημερήσια όσο και εβδομαδιαία. Αυτή έχει ως αιτία τη διαφορετικότητα των ημερών της εβδομάδας, όπου έχουμε τις εργάσιμες ημέρες (καθημερινές) από Δευτέρα έως και Παρασκευή και το Σάββατο και την Κυριακή. Η πρώτη κατηγορία ημερών χαρακτηρίζεται από έντονη ενασχόληση όλων των εργαζομένων, το Σάββατο θεωρείται ως ημι-εργάσιμη ημέρα, ενώ η Κυριακή είναι θεωρητικά ημέρα ανάπαυσης για όλους. Αντίστοιχα διαμορφώνεται και η ζήτηση του φορτίου μέσα στην εβδομάδα. Επίσης, σε επίπεδο ημέρας, παρατηρείται αυξομείωση της ζήτησης του φορτίου, καθώς από τις 6:00 το πρωί ως και τις 22:00 το βράδυ υπάρχει έντονη δραστηριότητα του ανθρώπου και με ένα διάλειμμα το απόγευμα, τα επίπεδα ζήτησης είναι αυξημένα (λόγω λειτουργίας όλων των υπηρεσιών, εργοστασίων, κλπ την ημέρα και λόγω λοιπών δραστηριοτήτων το απόγευμα και βράδυ, όπως τηλεόραση και ηλεκτρική κουζίνα). Τέλος, κάθε χώρα έχει τις δικές της ειδικές ημέρες, ημέρες εθνικών και μεγάλων θρησκευτικών εορτών, ιδιαίτερες αργίες και τις εκάστοτε απεργίες. Αυτές οι ημέρες προκαλούν ανωμαλίες στην πορεία της χρονοσειράς ζήτησης φορτίου είτε αυξάνοντας είτε μειώνοντας πολύ το επίπεδο (όπως Πρωτομαγιά).

Οι καιρικοί παράγοντες είναι άμεσα συνδεδεμένοι με την εναλλαγή των εποχών λόγω των μεταβολών της θερμοκρασίας. Για παράδειγμα, κατά τη διάρκεια του χειμώνα με τη μείωση της θερμοκρασίας είναι πιο έντονη η χρήση μέσων θέρμανσης, ενώ κατά τους καλοκαιρινούς μήνες αυξάνεται πολύ η θερμοκρασία και με αυτή η λειτουργία κλιματιστικών και άλλων μέσων ψύξης. Επιπλέον, ασυνήθιστες για την κάθε εποχή τιμές της θερμοκρασίας προκαλούν μεγάλες αποκλίσεις στη ζήτηση φορτίου, οι οποίες μάλιστα είναι έντονες και προκαλούν κινδύνους στα ΣΗΕ. Σε συνδυασμό με τη θερμοκρασία υπάρχει επιρροή και από το επίπεδο υγρασίας, τις βροχοπτώσεις, τη συννεφιά, την ύπαρξη έντονων ανέμων (ως και τυφώνων), το επίπεδο της ατμοσφαιρικής πίεσης και άλλους παρόμοιους παράγοντες.

Ως οικονομικούς παράγοντες αντιλαμβανόμαστε την οικονομική κατάσταση μιας χώρας (εφόσον αναφερόμαστε σε μελέτη ζήτησης φορτίου σε εθνικό επίπεδο) και το ρυθμό θετικής ή αρνητικής ανάπτυξης που την χαρακτηρίζει. Η σημασία τους είναι προφανής αφού με τη βελτίωση της οικονομικής κατάστασης μιας χώρας μεταβάλλεται και το βιοτικό επίπεδο των κατοίκων της, αυξάνεται η κατανάλωση σε αγαθά και υπηρεσίες, αναπτύσσονται νέες τεχνολογίες στον πρωτογενή και δευτερογενή τομέα για την κάλυψη των μεγαλύτερων πλέον αναγκών και υπάρχει γενικότερη εφορία. Προκαλείται με τον τρόπο αυτό αύξηση και της κατανάλωσης ηλεκτρικής ενέργειας, ώστε μεσοπρόθεσμα παρατηρείται άνοδος του επιπέδου των τιμών ζήτησης που μελετάμε. Σε κάποιο σημείο επέρχεται κορεσμός από την υπερκατανάλωση ενέργειας και οι καταναλωτές στρέφονται σε συσκευές με μικρότερη κατανάλωση ενέργειας στο πνεύμα ενός πιο οικολογικού τρόπου ζωής, με σκοπό τη μείωση του επιπέδου της ζητούμενης ενέργειας. Σε αντίθετη πορεία από μια χώρα

ή μια περιοχή που βιώνει ανάπτυξη, όταν παρατηρείται αρνητική πορεία (ύφεση) σε αυτή και το βιοτικό επίπεδο των καταναλωτών υποβαθμίζεται, ο κόσμος απομακρύνεται από τα αστικά κέντρα, και υιοθετεί πιο εγκρατή νοοτροπία, η οποία αντικατοπτρίζεται και στη μείωση της ζήτησης ηλεκτρικής ενέργειας.

Υπάρχουν πολλοί ακόμη παράγοντες, εκτός από όσους αναφέρθηκαν, που δεν ανήκουν σε κάποια συγκεκριμένη κατηγορία. Πρόκειται για γεγονότα τυχαία ή απρόσμενα, όπως οι εκλογές ή ένας πόλεμος.

Από τους παράγοντες που καταγράφηκαν, ορισμένοι είναι πολύ δύσκολο να συνυπολογιστούν στα μοντέλα πρόβλεψης της ζήτησης φορτίου, όπως οι οικονομικοί παράγοντες (ο ρυθμός ανάπτυξης και η οικονομική κατάσταση της χώρας). Στα πλαίσια της εργασίας αυτής λαμβάνονται υπόψη οι εποχικοί παράγοντες. Με παρατήρηση της χρονοσειράς των διαθέσιμων δεδομένων ζήτησης φορτίου ηλεκτρισμού αντιλαμβάνεται κανείς ότι η εποχικότητα, ειδικά σε ημερήσια και εβδομαδιαία βάση, είναι το πιο έντονο φαινόμενο της ζήτησης φορτίου. Επομένως αποτελεί και το κύριο μέλημά μας, ενώ η ενσωμάτωση του καιρού δε θα μας απασχολήσει προς το παρόν. Ο καιρός είναι όπως αναφέραμε σημαντικός παράγοντας στη διαμόρφωση της ζήτησης φορτίου, ωστόσο στη βραχυπρόθεσμη πρόβλεψη και μάλιστα με διασύνδεση με το δίκτυο, θεωρείται ιδιαίτερα πολύπλοκη η διαδικασία ενσωμάτωσης της επίδρασής του. Όπως δε γνωρίζουμε τις μελλοντικές τιμές για τη ζήτηση φορτίου και παράγουμε προβλέψεις τους, έτσι και με τον καιρό χρησιμοποιούνται οι προβλεπόμενες τιμές της θερμοκρασίας (η θερμοκρασία είναι η βασικότερη μεταβλητή των καιρικών συνθηκών που επηρεάζει τη ζήτηση φορτίου). Αυτό το χαρακτηριστικό των τιμών θερμοκρασίας, υγρασίας και λοιπών παραγόντων (όποιων χαρακτηριστικών του καιρού αποφασίζεται να χρησιμοποιηθούν στην παραγωγή πρόβλεψης φορτίου ηλεκτρισμού) προσθέτει επιπλέον αβεβαιότητα στις τελικές προβλέψεις του ζητούμενου φορτίου.

Ακόμη, η πρόβλεψη των καιρικών συνθηκών δε συνηθίζεται να γίνεται σε ωριαία βάση, αλλά για συγκεκριμένα χρονικά διαστήματα της ημέρας. Για παράδειγμα, το www.meteo.gr παράγει τέσσερις προβλέψεις για κάθε ημέρα, (για τις ώρες 9:00 και 15:00 κατά τη διάρκεια της ημέρας και 21:00 και 3:00 κατά τη διάρκεια της νύχτας). Χρειάζεται συνεπώς ειδική μεταχείριση των δεδομένων των καιρικών συνθηκών για την αποδοτική ενσωμάτωσή τους στις μεθόδους πρόβλεψης. Πρέπει να τονίσουμε ακόμη, ότι σε κάθε περιοχή της χώρας επικρατούν διαφορετικές καιρικές συνθήκες, οι οποίες επηρεάζουν ανάλογα τη ζήτηση των καταναλωτών κατά τόπους. Η ξεχωριστή πρόβλεψη για τις διάφορες περιοχές της χώρας και ο συνδυασμός τους για την τελική προβλεπόμενη τιμή είναι χρονοβόρα και απαιτητική διαδικασία, η οποία δεν είναι πρακτική όταν στόχος είναι η πρόβλεψη των επόμενων εικοσιτεσσάρων ωριαίων τιμών του ζητούμενου ηλεκτρικού φορτίου. Τέλος, μπορεί να παρατηρήσει κανείς ότι η αλλαγή του καιρού δεν είναι τόσο απότομη όσο ένα απρόσμενο τυχαίο γεγονός. Έτσι, με την ομαλή μεταβολή των καιρικών συνθηκών, οι μέθοδοι μίας μεταβλητής (της εποχικότητας στην περίπτωσή μας) θεωρούνται πολύ ικανοποιητικές, αφού η μεταβολή του καιρού αποτυπώνεται από μόνη της στη χρονοσειρά ζήτησης.

2.3 Το αντικείμενο της διπλωματικής

Συνοψίζοντας τα ανωτέρω (παράγραφοι 2.1 και 2.2), έχουμε ότι οι συνθήκες της αγοράς ηλεκτρικής ενέργειας έχουν μεταβληθεί έντονα τα τελευταία χρόνια, οδηγώντας τις εταιρείες παραγωγής της σε μεγάλο μεταξύ τους ανταγωνισμό. Δημιουργείται λοιπόν στις εταιρείες η ανάγκη για γνώση, εκ των προτέρων, της μελλοντικής συμπεριφοράς των

καταναλωτών της αγοράς στην οποία κινούνται. Εκτός από τον ανταγωνισμό, σημαντική είναι η γνώση αυτή και για το σωστό σχεδιασμό και την εύρυθμη λειτουργία των συστημάτων παραγωγής ηλεκτρικής ενέργειας. Εδώ συμβάλλει η επιστήμη των προβλέψεων, η οποία αναπτύσσει εργαλεία και μεθοδολογίες πρόβλεψης της ζήτησης του φορτίου, ώστε να εξυπηρετούνται οι ανάγκες των παραγωγών ηλεκτρικής ενέργειας.

Οι μέθοδοι πρόβλεψης ζήτησης φορτίου ηλεκτρικής ενέργειας διακρίνονται, ανάλογα με τον ορίζοντα της πρόβλεψης ζήτησης, σε πολύ βραχυπρόθεσμες, βραχυπρόθεσμες, μεσοπρόθεσμες και μακροπρόθεσμες προβλέψεις. Κάθε κατηγορία πρόβλεψης δίνει έμφαση σε ορισμένα χαρακτηριστικά των χρονοσειρών ζήτησης και εξυπηρετεί για το λόγο αυτό, διαφορετικές ανάγκες των εταιρειών, όπως παρουσιάστηκε στην παράγραφο 2.1. Στην παρούσα μελέτη, στόχος είναι η εξέταση μεθόδων για τη βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου. Η αξιολόγηση των μεθόδων γίνεται πάνω σε δεδομένα που μας διατίθενται από μία εταιρεία παραγωγής ηλεκτρικής ενέργειας στην Ελλάδα. Αναλυτική περιγραφή των δεδομένων γίνεται στο κεφάλαιο 5 (παράγραφος 5.1), εδώ να αναφερθεί μόνο ότι η διαθέσιμη χρονοσειρά αποτελείται από καταγεγραμμένες τιμές ωριαίας ζήτησης φορτίου ηλεκτρικής ενέργειας, για χρονικό διάστημα λίγο μεγαλύτερο των δυόμισι ετών. Οι μέθοδοι επομένως που επιλέγονται, εξετάζονται στη βραχυπρόθεσμη πρόβλεψη ωριαίας ζήτησης φορτίου ηλεκτρικής ενέργειας και με χρονικό ορίζοντα πρόβλεψης μία ημέρα (εικοσιτέσσερις παρατηρήσεις).

Είναι αναμενόμενη η ύπαρξη μεγάλου πλήθους μεθόδων βραχυπρόθεσμης πρόβλεψης της ζήτησης ηλεκτρικού φορτίου, όπως και η ποικιλία των προσαρμογών επί των μοντέλων που ήδη χρησιμοποιούνται. Η επιλογή των μεθόδων πρόβλεψης στη συγκεκριμένη εργασία γίνεται με βάση το γνωστικό μας πεδίο, την παρατήρηση της τάσης των ερευνών στον τομέα αυτό των προβλέψεων και στους στόχους που έχουμε θέσει από την αρχή, όπως περιγράφεται στη συνέχεια.

Υπάρχουν ποικίλες σύνθετες μέθοδοι, με δυνατότητα συμπερίληψης πολλών μεταβλητών ως εισόδους των μοντέλων για την εξαγωγή προβλέψεων της (μίας) εξαρτημένης μεταβλητής, που είναι η ζήτηση φορτίου. Αυτές οι μέθοδοι λαμβάνουν υπόψη την επίδραση των καιρικών συνθηκών, των εποχικών παραγόντων και όλων γενικά των παραμέτρων που μπορούν να επηρεάσουν τη ζήτηση του φορτίου και μπορούν να εισαχθούν στα μοντέλα πρόβλεψης. Ακόμη, γίνεται χρήση συνδυαστικών μεθόδων με βάρη επί των επιμέρους μεθόδων, στις οποίες συνηθίζεται η χρήση των πιο απλών μεθόδων πρόβλεψης μαζί με περισσότερο πολύπλοκα μοντέλα. Έπειτα από μελέτες των Hyndman και Taylor πάνω σε μεθόδους εκθετικής, παρατηρείται μεγάλο ενδιαφέρον στα μοντέλα αυτής της κατηγορίας, λόγω της απλότητάς τους και της μικρής τους απαίτησης σε δεδομένα εισόδου, ενώ παράλληλα η παραγωγή προβλέψεων ολοκληρώνεται σε σύντομο χρονικό διάστημα και έχει αρκετά ικανοποιητικά αποτελέσματα. Οι μέθοδοι βραχυπρόθεσμης πρόβλεψης, που ερευνώνται κατά κόρον, αναλύονται στο κεφάλαιο 4 (παράγραφος 4.7) και για το λόγο αυτό δεν γίνεται εκτενέστερη αναφορά στο σημείο αυτό.

Στόχος μας από την αρχή της κατάστρωσης της παρούσας εργασίας, είναι η αξιολόγηση των όποιων μεθόδων πρόβλεψης επιλεγούν και στη συνέχεια, μέσω των δικών επεμβάσεων μας να οδηγηθούμε στο καλύτερο δυνατό επίπεδο σφαλμάτων για όλες τις μεθόδους (στα πλαίσια πάντα της εργασίας). Παράλληλα, μας ενδιαφέρει ο εντοπισμός της μεθόδου η οποία να παράγει τα πλέον ικανοποιητικά αποτελέσματα. Με βάση λοιπόν τα όσα αναφέραμε, αποφασίζουμε να μην εξετάσουμε πολύπλοκα μοντέλα, αλλά να επικεντρωθούμε στις απλές μεθόδους προβλέψεων και να ασχοληθούμε και με προσαρμοσμένα μοντέλα εκθετικής εξομάλυνσης. Συγκεκριμένα, από την κατηγορία των

μεθόδων εκθετικής εξομάλυνσης επιλέγεται η προσαρμογή του Taylor στο ήδη υπάρχον μοντέλο των Holt-Winters (το αρχικό για ενσωμάτωση μίας εποχικότητας) για διπλή εποχικότητα [Taylor, 2003b], [Taylor, 2006]. Από τις απλές μεθόδους πρόβλεψης επιλέγονται οι εξής: η απλοϊκή μέθοδος Naive, οι μέθοδοι εκθετικής εξομάλυνσης SES, Damped, Holt και το κλασσικό μοντέλο της Theta. Οι μέθοδοι αναφέρονται όλες αναλυτικά στο κεφάλαιο 4, οι απλές μέθοδοι στην παράγραφο 4.3 και η προσαρμογή του Taylor στην παράγραφο 4.5.

Έχοντας επιλέξει τις μεθόδους που θέλουμε να εξετάσουμε, συνεχίζουμε με την εφαρμογή τους στα δεδομένα της διαθέσιμης χρονοσειράς. Έπειτα από ανάλυση των δεδομένων της αρχικής χρονοσειράς, όπως παρουσιάζεται στο κεφάλαιο 5 (παράγραφος 5.1), έχουν εντοπιστεί οι εποχικότητες της χρονοσειράς και εφαρμόζεται μέθοδος αποεποχικοποίησης των δεδομένων της. Πλέον, τα δεδομένα είναι έτοιμα για την παραγωγή προβλέψεων μέσω των μοντέλων που τους επιβάλλονται. Οι τελικές προβλέψεις σε ημερήσια βάση (24 προβλεπόμενες τιμές) προκύπτουν από την επαναφορά της εποχικότητας των δεδομένων στις προβλέψεις των μεθόδων, με χρήση των αντίστοιχων δεικτών εποχικότητας για κάθε ωριαία τιμή. (Οι δείκτες αυτοί υπολογίζονται και καταγράφονται κατά τη διαδικασία της αποεποχικοποίησης, όπως παρουσιάζεται αναλυτικά στο κεφάλαιο 4 (παράγραφος 4.4).)

Η ανάγκη για γρήγορη αντίδραση στις αλλαγές του περιβάλλοντος, οδηγεί στην ανάπτυξη ανάλογων μεθόδων. Στην παρούσα εργασία εφαρμόζεται η μεθοδολογία της κυλιόμενης παραγωγής προβλέψεων ή αλλιώς rolling forecasting. Αναλυτικότερα για αυτή τη μέθοδο θα αναφερθούμε στο κεφάλαιο 5 (παράγραφος 5.2). Αντικείμενο λοιπόν της μελέτης μας είναι η βέλτιστη εφαρμογή της μεθόδου κύλισης επί των δεδομένων της συγκεκριμένης χρονοσειράς. Μέσω επαναλήψεων με διαφορετικές τιμές των παραμέτρων της μεθόδου κύλισης, προσδιορίζονται εκείνες στις οποίες πετυχαίνεται το καλύτερο επίπεδο σφαλμάτων, για την εφαρμογή κάθε μεθόδου.

Στην παράγραφο 2.1 σημειώθηκε ότι οι ειδικές ημέρες αποτελούν τμήμα των εποχικών παραγόντων που επηρεάζουν τη ζήτηση φορτίου ηλεκτρικής ενέργειας και όπως είναι εμφανές, αντικείμενο της μελέτης μας είναι η αξιοποίηση των εποχικών παραγόντων στα μοντέλα πρόβλεψης. Επομένως, σημαντικό μέρος της διπλωματικής εργασίας είναι η ενσωμάτωση των ειδικών ημερών στα επιλεγμένα μοντέλα πρόβλεψης. Ο τρόπος εξέτασης της επίδρασης των ειδικών ημερών, καθώς και της ένταξής τους στη διαδικασία των προβλέψεων, παρουσιάζεται αναλυτικά στο κεφάλαιο 7.

Αυτό είναι, εν συντομία, το αντικείμενο της παρούσας διπλωματικής εργασίας. Τα καθολικά αποτελέσματα και τα συμπεράσματα που προκύπτουν παρουσιάζονται στο τελευταίο κεφάλαιο, 8, μαζί με προτάσεις για προέκταση της μελέτης μας.

ΚΕΦΑΛΑΙΟ 3: Χρονοσειρές

3.1 Γενικά χαρακτηριστικά χρονοσειρών

Οι χρονοσειρές είναι σύνολα διαδοχικών παρατηρήσεων της τιμής κάποιου φυσικού ή άλλου μεγέθους. Θα μπορούσε να πει κανείς ότι οι χρονοσειρές αντιπροσωπεύουν ντετερμινιστικές διαδικασίες, αφού παρουσιάζουν τη χρονική εξέλιξη ενός φυσικού ή άλλου μεγέθους όπως αναφέραμε ήδη. Όμως στην πράξη υπεισέρχεται ο παράγοντας της τυχαιότητας στα περισσότερα μεγέθη και για το λόγο αυτό οι χρονοσειρές αντιπροσωπεύουν εν τέλει στοχαστικές διαδικασίες.

Η επιστήμη των προβλέψεων αξιοποιεί τις εκάστοτε χρονοσειρές που μελετά για την εξαγωγή συμπερασμάτων σχετικά με την μελλοντική εξέλιξη των μεγεθών που περιγράφουν. Ο έλεγχος της γραφικής αναπαράστασης μιας χρονοσειράς στο πεδίο του χρόνου αποτελεί βάση της μελέτης της χρονοσειράς, καθώς δίνει μια πρώτη εικόνα της εξέλιξης του μεγέθους και αναδεικνύει τα ποιοτικά χαρακτηριστικά της χρονοσειράς: την τάση, την εποχικότητα, την κυκλικότητα και την τυχαιότητα. Διευκολύνεται έτσι ο χειριστής για την επιλογή της κατάλληλης μεθοδολογίας ανάλογα με τη χρονοσειρά που εξετάζεται. Ακόμη, ισχύει ότι με την παρατήρηση της γραφικής απεικόνισης μιας χρονοσειράς μπορεί να εντοπιστούν ακραίες τιμές (εκλιπούσες τιμές ή τιμές που απαιτούν ημερολογιακές προσαρμογές), οπότε ο χειριστής επεξεργάζεται περαιτέρω τα δεδομένα, για να διορθωθούν οι παρατηρούμενες ακραίες τιμές.

Συγκεκριμένα, η γραφική αναπαράσταση των δεδομένων της χρονοσειράς και η στατιστική ανάλυση με χρήση κατάλληλων δεικτών διευκολύνει στον εντοπισμό των ποιοτικών χαρακτηριστικών αυτής, καθώς επίσης και των διάφορων διαζόντων τιμών της. Είναι ωστόσο συχνή η ανάγκη για περαιτέρω επεξεργασία των δεδομένων, σε περιπτώσεις ύπαρξης κενών ή διακοπτόμενων τιμών, όπως και σε περιπτώσεις που χρειάζονται ημερολογιακές προσαρμογές επί των αρχικών δεδομένων, πριν εφαρμοστούν σε αυτά οι εκάστοτε μέθοδοι πρόβλεψης για την αποφυγή προβλημάτων και την επίτευξη καλύτερων αποτελεσμάτων.

Συγκεκριμένα, η ύπαρξη κενών τιμών οφείλεται σε μη καταγραφή της τιμής του φυσικού ή άλλου παρατηρούμενου μεγέθους λόγω αστοχίας του συστήματος ή σε λάθος χειρισμό του χρήστη και η αντιμετώπιση της ύπαρξής τους γίνεται με διαδικασίες εκτίμησης των όποιων κενών τιμών. Διαδικασίες εκτίμησης των τιμών αυτών είναι η εύρεσή τους από άλλες πηγές ή ο άμεσος ορισμός τους είτε η χρήση μέσων όρων από τις γύρω παρατηρήσεις με διάφορους συνδυασμούς. Οι μηδενικές τιμές οφείλονται με τη σειρά τους είτε σε καταγραφή τους ως μηδενικές τιμές από το σύστημα ή πρόκειται για πραγματικά μηδενικές τιμές. Οπότε στην πρώτη περίπτωση αντιμετωπίζονται όπως και οι κενές τιμές, ενώ στη δεύτερη περίπτωση πρόκειται για χρονοσειρές διακοπτόμενης ζήτησης, στις οποίες εφαρμόζονται ειδικές μέθοδοι πρόβλεψης για την αντιμετώπιση των μηδενικών τιμών τους. Τέλος, εφαρμόζεται συγκεκριμένη μεθοδολογία για την εξομάλυνση της διακύμανσης χρονοσειρών με περιόδους ορισμένου αριθμού ημερών. Η μεθοδολογία αυτή αποτελεί ημερολογιακή προσαρμογή επί των χρονοσειρών.

Σχετικά με τις τέσσερις συνιστώσες μιας χρονοσειράς, την τάση, τον κύκλο, την εποχικότητα και την τυχαιότητα, ακολουθεί τώρα η ανάλυσή τους.

❖ Η τάση (Trend, T) ορίζεται ως μια μακροπρόθεσμη μεταβολή του μέσου επιπέδου των τιμών της χρονοσειράς. Μπορεί να είναι αύξουσα, φθίνουσα είτε μηδενική τάση, ανάλογα με την πορεία των στοιχείων της χρονοσειράς σε μακροπρόθεσμο ορίζοντα. Η τάση κατηγοριοποιείται ακόμη ως γραμμική και μη γραμμική (παραδείγματα μη γραμμικής τάσης: εκθετική, λογαριθμική). Εφόσον αναφερόμαστε σε μακροπρόθεσμη μεταβολή, είναι σημαντικό να διευκρινίζεται ο χρονικός ορίζοντας της μεταβολής που θεωρείται ως μακροπρόθεσμος, καθώς είναι δυνατόν το ίδιο τμήμα της χρονοσειράς να θεωρείται ότι παρουσιάζει τάση για σχετικά μικρή περίοδο, ενώ για μεγαλύτερη χρονική περίοδο θα παρατηρούσαμε κυκλική διακύμανση των δεδομένων.

❖ Η κυκλικότητα (Cycle, C) ορίζεται ως μια κυματοειδής μεταβολή οφειλόμενη σε ειδικές εξωγενείς συνθήκες, η οποία παρουσιάζεται σε περιόδους όχι απαραίτητα σταθερής διάρκειας και με μήκος κατά κανόνα μεγαλύτερο του έτους. Οι εξωγενείς συνθήκες είναι κυρίως οι οικονομικές συνθήκες των διαφόρων περιόδων, στις οποίες παρουσιάζονται διαδοχικές άνοδοι και υφέσεις και ονομάζονται επιχειρηματικοί κύκλοι. Είναι φανερό λοιπόν, ότι κυκλικότητα εμφανίζουν οι χρονοσειρές των περισσότερων οικονομικών μεγεθών, όπως του Ακαθάριστου Εθνικού Προϊόντος (ΑΕΠ) και του ρυθμού των επιτοκίων.

❖ Η εποχικότητα (Seasonality, S) αποτελεί μια περιοδική διακύμανση με σταθερό μήκος, μικρότερο του έτους. Είναι εμφανής η διαφορά της εποχικότητας από την κυκλικότητα στο μήκος της περιόδου (μικρότερη έναντι μεγαλύτερης του έτους), όπως και στη σταθερότητα της περιόδου (σταθερή έναντι όχι απαραίτητα σταθερής). Οι εποχικές διακυμάνσεις είναι επαναλαμβανόμενες κατά τον ίδιο τρόπο με το χρόνο και έτσι εύκολα αντιληπτές και προβλέψιμες από τον παρατηρητή, με αποτέλεσμα να θεωρούνται κανονικές. Είναι εύκολη η μέτρηση και απομόνωση της επίδρασής τους ώστε να προκύψουν τα δεδομένα αποεποχικοποιημένα πλέον.

❖ Η τυχαιότητα (Randomness, R) αναφέρεται σε τυχαίες και μη προβλέψιμες διακυμάνσεις των δεδομένων μιας χρονοσειράς. Είναι η διακύμανση που παραμένει στη χρονοσειρά έπειτα από την απομόνωση της τάσης, της κυκλικότητας και της τυχαιότητας. Οι διακυμάνσεις τυχαιότητας ενδέχεται να είναι ασυνέχειες με παροδικό ή μόνιμο χαρακτήρα (ασυνήθιστες τιμές, outliers, ή αλλαγή επιπέδου, level shift, αντίστοιχα).

Οι τέσσερις βασικές συνιστώσες των χρονοσειρών μπορούν να απομονωθούν από αυτές με μεθόδους αποσύνθεσης. Με απομόνωση της εποχικότητας και της τυχαιότητας προκύπτει η σειρά τάσης-κύκλου κάθε αρχικής χρονοσειράς, στην οποία εφαρμόζονται εν συνεχεία οι επιλεγμένες μέθοδοι πρόβλεψης και οι τιμές που προκύπτουν από την εκάστοτε πρόβλεψη εποχικοποιούνται ξανά, ώστε να ληφθεί το τελικό αποτέλεσμα. Με τη διαδικασία αυτή επιτυγχάνεται αύξηση της αξιοπιστίας των αποτελεσμάτων της πρόβλεψης και δίνεται στο χρήστη η δυνατότητα να εισάγει και τη δική του κρίση, βάσει της εμπειρίας και της διαίσθησής του.

Οι δύο απλούστερες διατυπώσεις της αποσύνθεσης είναι η προσθετική και η πολλαπλασιαστική, με τη μορφή που παρουσιάζεται στις επόμενες δύο σχέσεις:

$$Y_t = S_t \cdot T_t \cdot C_t \cdot R_t, \text{ (πολλαπλασιαστικό μοντέλο)}$$

$$Y_t = S_t + T_t + C_t + R_t, \text{ (προσθετικό μοντέλο)}$$

Ο υπολογισμός της σειράς τάσης-κύκλου γίνεται με διάφορες μεθόδους (εξομάλυνση ή αποσύνθεση), με την απλούστερη αυτών να είναι η εξομάλυνση με χρήση κινητών μέσων όρων. Αμέσως μετά αναλύεται η εξομάλυνση με χρήση των κινητών μέσων όρων, ενώ στην επόμενη παράγραφο (παράγραφος 3.2) αναλύεται η μέθοδος της κλασσικής αποσύνθεσης.

Η μέθοδος των κινητών μέσων όρων δεν παρέχει πληροφορίες σχετικά με την εποχικότητα και την τυχαιότητα της χρονοσειράς τις οποίες απομακρύνει από αυτή για την παραγωγή της σειράς τάσης-κύκλου. Πιο λεπτομερώς, υπάρχουν τέσσερα είδη κινητών μέσων όρων για την παραγωγή της εξομαλυμένης σειράς. Σημαντική είναι η επιλογή του μήκους n πριν τον υπολογισμό των κινητών μέσων όρων. Με την αύξηση του μήκους n επιτυγχάνεται και μεγαλύτερη εξομάλυνση των δεδομένων, ωστόσο πολύ μεγάλη τιμή του μήκους οδηγεί στον υπολογισμό της μέσης τιμής των δεδομένων της χρονοσειράς και δεν διακρίνονται η τάση και ο κύκλος που μας ενδιαφέρουν, ενώ επιπλέον αυξάνεται πολύ ο αριθμός των κενών τιμών στην αρχή και στο τέλος της παραγόμενης σειράς. Να σημειωθεί ότι στην περίπτωση έντονης εμφάνισης εποχικότητας προτιμάται η χρήση μήκους ίσου (ή μεγαλύτερου) με το μήκος της περιόδου της εποχικότητας.

- Ο απλός κινητός μέσος όρος (ΚΜΟ) πρόκειται για τον απλό μέσο όρο n τιμών (όπου n : περιττός για τη διατήρηση συμμετρίας στους υπολογισμούς) της αρχικής χρονοσειράς γύρω από την εκάστοτε παρατήρηση. Υπολογίζοντας τον απλό κινητό μέσο όρο για όλες τις παρατηρήσεις της χρονοσειράς προκύπτει η επιθυμητή σειρά τάσης-κύκλου.

- Ο σταθμισμένος κινητός μέσος όρος (ΣΚΜΟ) αποτελεί παραλλαγή του απλού ΚΜΟ με συμμετοχή των γειτονικών παρατηρήσεων στον υπολογισμό της σειράς τάσης-κύκλου με άνισα βάρη. Το άθροισμα από τα βάρη πρέπει να ισούται με τη μονάδα, δίνεται προσοχή ώστε να είναι συμμετρικά ως προς την εκάστοτε (τρέχουσα) παρατήρηση, ενώ η επιλογή τους γίνεται με στόχο να δοθεί μεγαλύτερη βαρύτητα σε παρατηρήσεις πλησίον της εκάστοτε παρατήρησης και μικρότερη στις πιο μακρινές παρατηρήσεις. Η χρήση του σταθμισμένου κινητού μέσου όρου οδηγεί σε περισσότερο εξομαλυμένη σειρά τάσης-κύκλου.

- Στο διπλό κινητό μέσο όρο (ΔΚΜΟ) με διπλή εφαρμογή του απλού ΚΜΟ, με ίσα ή και άνισα μήκη, οδηγούμαστε σε διπλή εξομάλυνση, αλλά και στην ύπαρξη περισσότερων κενών τιμών -συγκριτικά πάντα με τον απλό κινητό μέσο όρο (ΚΜΟ). Ακόμη, η εφαρμογή του ΔΚΜΟ ισοδυναμεί με την εφαρμογή του ΣΚΜΟ με ορισμένα βάρη (εξαρτώμενα από την κάθε περίπτωση).

- Ο κεντρικός κινητός μέσος όρος (ΚΚΜΟ) αποτελεί συνδυασμό του απλού και του διπλού κινητού μέσου όρου και χρησιμεύει για τον υπολογισμό κινητού μέσου όρου άρτιου μήκους. Η εφαρμογή του κεντρικού κινητού μέσου όρου ισοδυναμεί με ένα διπλό κινητό μέσο όρο $2 \times n$ (όπου n το μήκος του απλού ΚΜΟ), καθώς και με ένα σταθμισμένο κινητό μέσο όρο με ορισμένα βάρη. Η χρήση του ΚΚΜΟ είναι πλέον συνήθης επί χρονοσειρών με παρατηρήσεις μηνιαίες, τριμηνιαίες, εξαμηνιαίες, κλπ για την εξομάλυνση της όποια εποχικότητας ενδέχεται να παρουσιάζουν (με δώδεκα, τέσσερις και δύο παρατηρήσεις ανά έτος αντίστοιχα).

3.2 Κλασσική μέθοδος αποσύνθεσης

Όσον αφορά στη μέθοδο της αποσύνθεσης, η πιο κοινή είναι η κλασσική μέθοδος αποσύνθεσης, η οποία εφαρμόζεται τόσο σε χρονοσειρά με πολλαπλασιαστική όσο και με

προσθετική διατύπωση της αποσύνθεσης (με την πολλαπλασιαστική να αποτελεί τον πιο συνηθισμένο τρόπο διατύπωσης). Υπάρχει η δυνατότητα παραγωγής προβλέψεων μέσω της αποσύνθεσης, όπου έπειτα από ανάλυση της αρχικής χρονοσειράς στις συνιστώσες της ακολουθεί η σύνθεσή τους ξανά για τη λήψη της τελικής τιμής πρόβλεψης. Η βασική συνιστώσα που αφαιρείται από τα δεδομένα είναι η εποχικότητα, η οποία έπειτα από την παραγωγή της πρόβλεψης ενσωματώνεται και πάλι στην τελική χρονοσειρά. Στη συνέχεια αναλύονται τα βήματα που ακολουθούνται κατά την κλασική αποσύνθεση σε χρονοσειρά πολλαπλασιαστικής σχέσης των συνιστωσών της. Ο τρόπος με τον οποίο γίνεται η αποεποχικοποίηση στην παρούσα εργασία και η συμπερίληψη της εποχικότητας στο τέλος περιγράφονται στη συνέχεια (παράγραφος 4.3).

Η μεθοδολογία της κλασικής αποσύνθεσης αποτελείται από πέντε βασικά βήματα:

1° βήμα:

Αρχικά, επιλέγεται ένα από τα τέσσερα είδη κινητού μέσου όρου που σημειώθηκαν στην παράγραφο 3.1 με βάση το μήκος της εποχικότητας των δεδομένων. Ο υπολογιζόμενος κινητός μέσος όρος είναι απαλλαγμένος από εποχικότητα, ενώ ενδέχεται να περιέχει μόνο μικρή τυχαιότητα, δίνεται έτσι μια καλή εικόνα της πορείας των δεδομένων της χρονοσειράς σχετικά με την τάση και την κυκλικότητα. Μπορεί να θεωρηθεί επομένως ότι ισχύει το εξής:

$$KMO(n)_t = T_t \times C_t$$

Στην παραπάνω σχέση συμβολίζονται με n το μήκος του κινητού μέσου όρου, με T η συνιστώσα της τάσης και με C η συνιστώσα του κύκλου της χρονοσειράς. Η χρονοσειρά ζήτησης φορτίου παρουσιάζει ημερήσια και εβδομαδιαία εποχικότητα και, όπως θα δούμε στο κεφάλαιο 5, τα δεδομένα που διατίθενται για την παρούσα εργασία είναι σε ωριαία βάση. Επομένως, γνωρίζοντας ότι για την περίπτωση που τα δεδομένα έχουν άρτιο μήκος εποχικότητας χρησιμοποιείται ο κεντρικός κινητός μέσος όρος (ΚΚΜΟ) και έχοντας στην περίπτωση αυτή μήκη $s_1 = 24$ και $s_2 = 168$ αντίστοιχα για ημερήσια και εβδομαδιαία εποχικότητα, συμπεραίνει κανείς ότι θα χρησιμοποιηθεί ο επόμενος τύπος με ΚΚΜΟ:

$$KKMO(n)_t = T_t \times C_t$$

2° βήμα:

Έπειτα διαιρούνται τα πραγματικά δεδομένα με τους κεντρικούς μέσους όρους που υπολογίστηκαν στο 1° βήμα, όπως φαίνεται στην επόμενη σχέση. Λαμβάνονται με τον τρόπο αυτό οι λόγοι εποχικότητας, που χαρακτηρίζονται ακόμη από την εποχικότητα και την τυχαιότητα της χρονοσειράς.

$$\frac{Y_t}{KKMO(n)_t} = \frac{T_t \times C_t \times S_t \times R_t}{T_t \times C_t} = S_t \times R_t$$

3° βήμα:

Στο βήμα αυτό, στόχος είναι η απαλοιφή της τυχαιότητας από τους λόγους εποχικότητας που μόλις υπολογίστηκαν. Για να συμβεί αυτό, υπολογίζονται οι μέσοι όροι των αντίστοιχων λόγων εποχικότητας, οι οποίοι αποτελούν τους δείκτες εποχικότητας της χρονοσειράς (της αρχικής χρονοσειράς). Ενδέχεται βέβαια να χρειάζεται ακόμη μια επεξεργασία αυτών των μέσων όρων ώστε το άθροισμά τους να είναι ίσο με το μήκος της εποχικότητας. Η διαδικασία αυτή ονομάζεται κανονικοποίηση των δεικτών εποχικότητας της χρονοσειράς και γίνεται με υπολογισμό των αντίστοιχων μέσων όρων των λόγων εποχικότητας, χωρίς τις ακραίες τιμές (τη μέγιστη και την ελάχιστη τιμή τους). Στην παρούσα εργασία, ακολουθείται η διαδικασία της κανονικοποίησης των δεικτών εποχικότητας. Συγκεκριμένα, έπειτα από τον

υπολογισμό των λόγων εποχικότητας, αγνοείται η μέγιστη και η ελάχιστη τιμή τους και υπολογίζεται ο μέσος όρος των υπόλοιπων. Οι νέοι μέσοι όροι ονομάζονται μη κανονικοποιημένοι δείκτες εποχικότητας και το πλήθος τους ισούται με το μήκος της εποχικότητας. Μετά, υπολογίζεται ο συντελεστής κανονικοποίησης, ως ο λόγος του μήκους της εποχικότητας προς το άθροισμα όλων των μη κανονικοποιημένων δεικτών εποχικότητας. Με πολλαπλασιασμό των μη κανονικοποιημένων δεικτών επί τον συντελεστή κανονικοποίησης προκύπτουν εν τέλει οι κανονικοποιημένοι δείκτες εποχικότητας.

4^ο βήμα:

Στο βήμα αυτό, διαιρούνται τα δεδομένα της πραγματικής χρονοσειράς με τους αντίστοιχους κανονικοποιημένους δείκτες εποχικότητας που υπολογίστηκαν στο 3^ο βήμα. Παράγεται έτσι η αποεποχικοποιημένη σειρά, η οποία περιέχει πλέον τις υπόλοιπες τρεις συνιστώσες της, τάση, κύκλο και τυχαιότητα, έχοντας απαλλαγεί από την εποχικότητα.

$$\frac{Y_t}{S_t} = \frac{T_t \times C_t \times S_t \times R_t}{S_t} = T_t \times C_t \times R_t$$

Στην παρούσα εργασία, προχωρούμε μέχρι και το 4^ο βήμα της κλασσικής μεθόδου αποσύνθεσης, όπου επιτυγχάνεται η αποεποχικοποίηση των επιλεγμένων δεδομένων και στη συνέχεια ακολουθεί η εφαρμογή των μεθόδων πρόβλεψης. Εδώ συνεχίζουμε λίγο ακόμη, περιγράφοντας το 5^ο βήμα της διαδικασίας.

5^ο βήμα:

Ενδιαφέρον πλέον έχει η απαλοιφή και της τυχαιότητας από την ήδη αποεποχικοποιημένη σειρά. Υπολογίζεται, για το σκοπό αυτό, ο κινητός μέσος όρος μήκους $n = 3$ ή $n = 6$ παρατηρήσεων την αποεποχικοποιημένης σειράς, ώστε να παραχθεί μια ικανοποιητική και ακριβής σειρά τάσης-κύκλου. Για βέλτιστα αποτελέσματα συνίσταται η χρήση διπλού κινητού μέσου όρου, ΔΚΜΟ(3x3). Στην περίπτωση που ζητείται η συνιστώσα της τυχαιότητας, υπολογίζεται με βάση τη σχέση:

$$\frac{T_t \times C_t \times R_t}{\text{ΚΚΜΟ}(3 \times 3)_t} = \frac{T_t \times C_t \times R_t}{T_t \times C_t} = R_t$$

Αυτά είναι τα πέντε κύρια βήματα της μεθόδου της κλασσικής αποσύνθεσης. Μπορεί κανείς να συνεχίσει και στον υπολογισμό της συνιστώσας της τάσης έπειτα από το 5^ο βήμα και ακόμη, με πολλαπλασιαστική σύνθεση των συνιστωσών τάσης, εποχικότητας και κύκλου (η τυχαιότητα δεν είναι προβλέψιμη συνιστώσα) να παράγει προβλέψεις. Σημειώνεται ότι η πρόβλεψη γίνεται μέσω πολλαπλασιαστικής σύνθεσης των συνιστωσών, επειδή από την αρχή της παραγράφου αναφερόμαστε σε χρονοσειρά με πολλαπλασιαστική σχέση μεταξύ των συνιστωσών της.

3.3 Επίδραση ειδικών γεγονότων επί των χρονοσειρών

Κατά τη γραφική απεικόνιση των χρονοσειρών παρατηρούνται συχνά ορισμένες διακυμάνσεις των τιμών των δεδομένων, πιο έντονες από τις συνήθειες. Αυτές οι διακυμάνσεις οφείλονται σε ειδικά γεγονότα και ενέργειες, που προκαλούν αλλαγή του επιπέδου των τιμών προς τα πάνω ή προς τα κάτω ή και καταγραφή περιόδων με ιδιαίτερα χαμηλές ή υψηλές παρατηρήσεις. Τα αίτια πρόκλησής τους βρίσκονται στο άμεσο αλλά και στο ευρύτερο περιβάλλον του μεγέθους υπό μελέτη. Για παράδειγμα, στην περίπτωση της

ζήτησης ηλεκτρικής ενέργειας κατά το διάστημα των Αλκυονίδων ημερών παρατηρείται μείωση της ζητούμενης κατανάλωσης, λόγω αύξησης της θερμοκρασίας του περιβάλλοντος κατά τη διάρκεια του κρύου κατά τα άλλα χειμώνα. Προξενείται, έτσι, πτώση του επιπέδου της χρονοσειράς για το διάστημα που διαρκεί το φαινόμενο αυτό και αναφερόμαστε σε ένα ειδικό γεγονός. Η επίδραση ωστόσο του ειδικού γεγονότος υπεισέρχεται και στην πρόβλεψη των επόμενων τιμών της χρονοσειράς (δεν αναφερόμαστε μόνο στην απλή παρατήρηση της χρονοσειράς, αλλά και στην προέκτασή της εδώ). Μειώνεται επομένως η ακρίβεια των παραγόμενων προβλέψεων και απαιτείται η αφαίρεση του αντίκτυπου του γεγονότος από τις τιμές των δεδομένων της συγκεκριμένης περιόδου, ώστε η πρόβλεψη των επόμενων δεδομένων να γίνεται με βάση τις εξομαλυμένες τιμές τους και να υπάρχει μεγαλύτερη ακρίβεια στο επίπεδο που κινούνται οι παρατηρήσεις της χρονοσειράς.

Είναι προφανής η ανάγκη για τον εντοπισμό των ειδικών γεγονότων, η καταγραφή του αντίκτυπου αυτών επί των δεδομένων του μεγέθους που περιγράφει η εκάστοτε χρονοσειρά και η εξομάλυνση των δεδομένων της. Με τον τρόπο αυτό μπορεί ο χρήστης να αντιληφθεί καλύτερα την πορεία της σειράς δεδομένων και να επιτύχει ακριβέστερες προβλέψεις, συνδυάζοντας τα αποτελέσματα των στατιστικών μεθόδων πρόβλεψης (επί της εξομαλυμένης σειράς) και τη δική του κρίση βάσει της παρατήρησης παρόμοιων ή ίδιας φύσης γεγονότων σε παλαιότερες χρονικές περιόδους. Οι μέθοδοι εντοπισμού ειδικών γεγονότων και ενεργειών αναλύουν κάθε παρατήρηση ξεχωριστά για να καταλήξουν αν είναι ασυνήθιστη ή όχι και στη συνέχεια ομαδοποιούν τις γειτονικές ασυνήθιστες παρατηρήσεις, εντοπίζοντας τα ειδικά γεγονότα. Έπειτα, ο χειριστής με την κρίση του επιβεβαιώνει τα ειδικά γεγονότα και τα κατηγοριοποιεί. Ακολουθεί ο υπολογισμός της εξομαλυμένης σειράς δεδομένων, επί της οποίας πλέον μπορούν να εφαρμοστούν οι διάφορες στατιστικές μέθοδοι πρόβλεψης. Και στο τέλος ο χειριστής δύναται να συμπεριλάβει τον αντίκτυπο των ειδικών γεγονότων στα αποτελέσματα των στατιστικών μεθόδων (σε περίπτωση που αναμένεται αντίστοιχο φαινόμενο στο χρονικό διάστημα της πρόβλεψης) για να ληφθεί η επίσημη τιμή των προβλεπόμενων παρατηρήσεων.

3.4 Στατιστική ανάλυση χρονοσειρών

Αναφέρθηκε προηγουμένως ότι η αποκόμιση πληροφοριών για τις χρονοσειρές που μελετώνται γίνεται μέσω παρατήρησης της γραφικής απεικόνισης αυτών είτε με τη χρήση στατιστικής ανάλυσης, μέσω ορισμένων δεικτών. Στη καταγράφονται οι δείκτες αυτοί που βοηθούν στην απόκτηση καλύτερης κρίσης της πορείας των δεδομένων και στην πιο σωστή επιλογή μεθόδων επέκτασής τους. Οι βασικοί στατιστικοί δείκτες που υπολογίζονται πάνω στα n τον αριθμό δεδομένα της μελετώμενης χρονοσειράς είναι οι εξής:

- Μέση τιμή, $\bar{Y} = \frac{1}{n} \cdot \sum_{i=1}^n Y_i$

Πρόκειται για τον απλό γραμμικό μέσο όρο των τιμών των δεδομένων της χρονοσειράς και δείχνει το επίπεδο γύρω από το οποίο μεταβάλλονται.

- Μέγιστη και ελάχιστη τιμή

Οι ακραίες τιμές της χρονοσειράς βοηθούν στην αντίληψη της διακύμανσης των δεδομένων της, καθώς και της τυχαιότητας αυτών.

- Τυπική απόκλιση, $\sigma_\alpha = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}}$ και $\sigma_\beta = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}$

Ο δείκτης αυτός εκφράζει πόσο μεγάλη είναι η διασπορά των τιμών των δεδομένων γύρω από τη μέση τιμή, \bar{Y} . Με σ_α συμβολίζεται η τυπική απόκλιση που αναφέρεται σε ένα δείγμα, ενώ με σ_β συμβολίζεται η τυπική απόκλιση πληθυσμού.

- Διακύμανση, σ^2

Δείχνει το βαθμό συγκέντρωσης των παρατηρήσεων γύρω από τη μέση τιμή (όμοια με την τυπική απόκλιση, αλλά με τα μεγέθη υψωμένα στο τετράγωνο).

- Συνδιακύμανση, $\text{Cov}(X, Y) = \frac{1}{n} \cdot \sum_{i=1}^n [(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]$

Ο δείκτης αυτός εμφανίζει τη σχέση ανάμεσα σε δύο μεταβλητές, συγκρίνοντας τη φορά μεταβολής τους. Δείχνει δηλαδή αν μεταβάλλονται προς την ίδια (θετική συνδιακύμανση) ή προς την αντίθετη κατεύθυνση (αρνητική συνδιακύμανση) και τη συσχέτιση ή μη των συνόλων των δύο μεταβλητών (όταν δεν υπάρχει συσχέτιση, πρόκειται για μηδενική συνδιακύμανση). Στην περίπτωση της μελέτης χρονοσειράς γίνεται υπολογισμός της συνδιακύμανσης των παρατηρήσεων της συγκριτικά με τον αύξοντα αριθμό της χρονικής περιόδου (off).

- Συντελεστής γραμμικής συσχέτισης, $r_{XY} = \frac{\sum_{i=1}^n [(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}$

Είναι το μέτρο της συγκέντρωσης των σημείων ενός διαγράμματος διασποράς γύρω από την ευθεία παλινδρόμησης, που δηλώνει τη συσχέτιση μεταξύ δύο συνόλων δεδομένων (κατεύθυνση και βαθμό συσχέτισης).

- ✓ Για $0 < r < 1$ οι δύο μεταβλητές X και Y είναι θετικά γραμμικά συσχετισμένες.
- ✓ Για $-1 < r < 0$ οι δύο μεταβλητές είναι αρνητικά γραμμικά συσχετισμένες.
- ✓ Για $r = \pm 1$ υπάρχει τέλεια θετική γραμμική συσχέτιση, οπότε όλα τα σημεία βρίσκονται επί της ευθείας παλινδρόμησης.
- ✓ Για $r = 0$ δεν υπάρχει καθόλου συσχέτιση μεταξύ των δύο μεταβλητών.

Να σημειώσουμε ότι όσο η τιμή του δείκτη r τείνει στις ακρινές τιμές του διαστήματος [-1, 1] στο οποίο κινείται, τόσο πιο ισχυρή είναι η συσχέτιση μεταξύ των δύο μεταβλητών.

- Συντελεστής αυτοσυσχέτισης, $\text{ACF}_k = \frac{\sum_{i=1}^n [(Y_i - \bar{Y}) \cdot (Y_{i-k} - \bar{Y})]}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$

Ο δείκτης αυτός εκφράζει τη συσχέτιση μεταξύ παρατηρήσεων της ίδιας πλέον μεταβλητής, οι οποίες έχουν χρονική διαφορά k περιόδους. Η τιμή του κινείται στο διάστημα [0, 1] και όταν είναι κοντά στο άνω όριο της φανερώνει μεγάλη συσχέτιση, ενώ το αντίθετο συμβαίνει για τιμή κοντά στο μηδέν. Η σημασία του δείκτη αυτού φαίνεται στην περίπτωση που μας ενδιαφέρει ο εντοπισμός της εποχικής συμπεριφοράς μια χρονοσειράς (με αλγεβρικό τρόπο).

- Συντελεστής μεταβλητότητας, $c_v = \frac{\sigma}{\bar{y}} \cdot 100\%$

Πρόκειται για κανονικοποιημένο μέτρο της διασποράς των παρατηρήσεων δείγματος ή πληθυσμού και πλεονεκτεί έναντι της απλής τυπικής απόκλισης (σ) λόγω της μη συμπερίληψης της επίδρασης του επιπέδου των παρατηρήσεων. Δε γίνεται όμως να υπολογιστεί σε περίπτωση μηδενικής μέσης τιμής.

- Μέση τιμή μεταξύ ζητήσεων

Στην περίπτωση διακοπτόμενης ζήτησης αποτελεί σημαντικό μέτρο της ασυνέχειας των ζητήσεων και η τιμή του δηλώνει κάθε πόσες, κατά μέσο όρο, περιόδους εμφανίζεται μηδενική τιμή στη χρονοσειρά. Όσο μεγαλύτερες τιμές (μεγαλύτερες της μονάδας για την οποία η ζήτηση είναι συνεχής) λαμβάνει, τόσο μεγαλύτερα είναι τα διαστήματα της χρονοσειράς χωρίς να παρατηρηθεί μη μηδενική τιμή.

Μία ακόμη κατηγορία της στατιστικής ανάλυσης εκτός από τη βασική, που αναλύθηκε μόλις, είναι ο ρυθμός ανάπτυξης. Πρόκειται για ένα μέτρο της πορείας των δεδομένων μιας χρονοσειράς, είτε αύξουσα ή φθίνουσα, για ορισμένο χρονικό διάστημα. Η τιμή του παρουσιάζεται σε ποσοστιαία μορφή και προκύπτει ως η σύγκριση των δεδομένων του τελευταίου έτους με τα δεδομένα της υπόλοιπης χρονικής περιόδου που διατίθεται.

Οι παραπάνω δείκτες υπολογίζονται επί της χρονοσειράς που μελετάται χωρίς να χρειαστεί προέκτασή της. Με την εφαρμογή μεθόδων πρόβλεψης επί της χρονοσειράς, εφαρμόζεται στατιστική ανάλυση της ακρίβειας των προβλεπόμενων τιμών και οι αντίστοιχοι δείκτες αναλύονται στο επόμενο κεφάλαιο (παράγραφος 4.6).

3.5 Χαρακτηριστικά χρονοσειρών ζήτησης φορτίου

Οι χρονοσειρές ζήτησης ηλεκτρικού φορτίου στις περισσότερες αγορές ενέργειας παρουσιάζουν τα ακόλουθα χαρακτηριστικά. Η μέση τιμή των χρονοσειρών δεν είναι σταθερή και παρατηρείται έντονη διακύμανση στις τιμές των δεδομένων τους. Σε ποικίλες χρονοσειρές ζήτησης φορτίου ηλεκτρικής ενέργειας είναι εμφανές ένα κοινό χαρακτηριστικό, η ύπαρξη εποχικότητας στα δεδομένα και μάλιστα όχι μόνο ενός είδους. Η ημερήσια και η εβδομαδιαία εποχικότητα έχουν έντονη παρουσία στις χρονοσειρές αυτές. Επιπλέον, οι καιρικές συνθήκες, με κυρίαρχη τη θερμοκρασία, επιδρούν ιδιαίτερα στη ζήτηση φορτίου ηλεκτρισμού και επηρεάζουν μη γραμμικά τις χρονοσειρές. Φυσικά υπάρχουν και τυχαίες διακυμάνσεις, που παρουσιάζονται ως ανωμαλίες πάνω στη γραφική απεικόνιση των δεδομένων ζήτησης και προκαλούνται από τυχαίους και ασυνήθιστους παράγοντες. Αναφέρθηκε και προηγουμένως ότι στα πλαίσια της εργασίας αυτής δε θα υπάρξει αναφορά στην επίδραση των καιρικών και των οικονομικών συνθηκών, αλλά σημασία θα δοθεί στις ειδικές ημέρες (αργίες, απεργίες, κλπ).

Η εμφανής ημερήσια και εβδομαδιαία εποχικότητα των καμπυλών ζήτησης φορτίου είναι εύκολα παρατηρήσιμη, ειδικά σε χρονοσειρές με ωριαία ως και λεπτό-προς-λεπτό δεδομένα. Σαφώς οι χρονοσειρές ζήτησης φορτίου χαρακτηρίζονται και από ετήσια εποχικότητα, η οποία όμως δεν είναι μεγάλης σημασίας όταν γίνεται βραχυπρόθεσμη πρόβλεψη ζήτησης ηλεκτρικού φορτίου. Επομένως δε θα ληφθεί υπόψη στη συγκεκριμένη εργασία η ετήσια εποχικότητα της χρονοσειράς, αλλά μόνο η ημερήσια και η εβδομαδιαία

εποχικότητα είναι αυτές που θα μελετηθούν. Ο τρόπος χειρισμού των εποχικοτήτων της χρονοσειράς που μελετάται θα αναλυθεί σε επόμενη παράγραφο.

ΚΕΦΑΛΑΙΟ 4: Μέθοδοι πρόβλεψης

4.1 Γενικά

Οι χρονοσειρές χρησιμεύουν για την παρατήρηση της εξέλιξης των μεγεθών που περιγράφουν. Εκτός από την μελέτη των καταγεγραμμένων δεδομένων, πολύ σημαντική είναι συνεπώς η γνώση -έστω και κατά προσέγγιση- της μελλοντικής πορείας της ακολουθίας των παρατηρήσεων. Αυτή η ανάγκη καλύπτεται από την παραγωγή προβλέψεων με τη χρήση μοντέλων πρόβλεψης, με το κάθε μοντέλο να ακολουθεί μια διαφορετική τεχνική και έτσι έχουμε την παραγωγή ποικίλων μοντέλων και αποτελεσμάτων πρόβλεψης.

Στην παράγραφο αυτή θα αναφερθούμε στα δύο μοντέλα που υπάγονται στο είδος της στατιστικής πρόβλεψης, το μοντέλο χρονοσειρών και το αιτιοκρατικό μοντέλο [Πετρόπουλος, Ασημακόπουλος, 2011]. Υπάρχουν και ακόμη δύο είδη πρόβλεψης, η κριτική πρόβλεψη και η πρόβλεψη στόχου, που θα παρουσιαστούν πιο αναλυτικά στην επόμενη παράγραφο.

- Το μοντέλο χρονοσειρών είναι το πλέον γνωστό είδος στατιστικής (ή αλλιώς ποσοτικής) πρόβλεψης, η εφαρμογή του οποίου προϋποθέτει την ύπαρξη ιστορικών δεδομένων σε προηγούμενες, σταθερές χρονικές περιόδους. Υποθέτουμε ότι η τιμή του μεγέθους μεταβάλλεται με ένα καθορισμένο, σταθερό πρότυπο, που επαναλαμβάνεται στο χρόνο και η πρόβλεψη πραγματοποιείται με την αναγνώριση του προτύπου αυτού και την προέκτασή του. Το μοντέλο αυτό περιγράφει τις λεγόμενες μεθόδους των χρονοσειρών, μερικές εκ των οποίων είναι η αποσύνθεση, η εξομάλυνση και οι αυτοπαλινδρομικές μέθοδοι κινητού μέσου όρου (Autoregressive Moving Average ή αλλιώς ARMA). Οι μέθοδοι αυτοί θα αναλυθούν στην παράγραφο 4.2 Η χρήση του μοντέλου χρονοσειρών παράγει ικανοποιητικά αποτελέσματα στις περιπτώσεις που η εξέλιξη της χρονοσειράς γίνεται με βάση τις τωρινές συνθήκες (εφόσον οι προβλέψεις στηρίζονται στις ιστορικές παρατηρήσεις). Το μειονέκτημα του μοντέλου χρονοσειρών είναι η μη συσχέτιση του προβλεπόμενου μεγέθους με τους παράγοντες που το επηρεάζουν και έτσι παράγονται προβλέψεις χωρίς όμως να αιτιολογείται η πορεία της προβλεπόμενης χρονοσειράς.

- Το αιτιοκρατικό μοντέλο βασίζεται στην υπόθεση της ύπαρξης μιας σταθερής σχέσης ανάμεσα στο μέγεθος προς πρόβλεψη, ουσιαστικά μια εξαρτημένη μεταβλητή, και σε ορισμένες ανεξάρτητες παραμέτρους που το επηρεάζουν. Προσδιορίζονται οι τιμές των ανεξάρτητων μεταβλητών (μέσω εκτιμήσεων, προβλέψεων, μετρήσεων, κλπ) και στη συνέχεια προκύπτει η πρόβλεψη της εξαρτημένης (και ζητούμενης) μεταβλητής. Στην περίπτωση αυτού του μοντέλου η συνάρτηση που περιγράφει το σύστημα προκύπτει από μία ορισμένη διαδικασία, σε αντίθεση με το μοντέλο χρονοσειρών στο οποίο η συνάρτηση προσδιορίζεται από το πρότυπο που έχει οριστεί από την υπόθεση. Μειονέκτημα του αιτιοκρατικού μοντέλου είναι η ανάγκη για αρκετά περισσότερα δεδομένα, λόγω των διαφορετικών ανεξάρτητων μεταβλητών και μάλιστα πολλές φορές για μεγάλα χρονικά διαστήματα πριν το τρέχον. Ακόμη, οι παράγοντες που επηρεάζουν όλες τις ανεξάρτητες μεταβλητές επιδρούν και στο μοντέλο αυτό. Έχει όμως το θετικό ότι μπορεί ο χειριστής να λάβει προβλέψεις για την πορεία μιας χρονοσειράς με πολλούς συνδυασμούς των μεταβλητών εισόδου.

4.2 Κατηγορίες πρόβλεψης

Υπάρχουν τρία είδη πρόβλεψης, στα οποία βασίζεται η τελική πρόβλεψη της επιχείρησης σε κάθε περίπτωση [Πετρόπουλος, Ασημακόπουλος, 2011], η στατιστική πρόβλεψη, η κριτική πρόβλεψη και η πρόβλεψη στόχου.

❖ Η στατιστική πρόβλεψη αναφέρεται στην εφαρμογή των στατιστικών μοντέλων χρονοσειρών ή των αιτιοκρατικών μοντέλων πάνω σε μια σειρά δεδομένων, ώστε να παραχθεί με συστηματικό τρόπο η πρόβλεψη της πορείας της ακολουθίας. Σημαντικό στοιχείο της στατιστικής πρόβλεψης είναι η ακρίβειά της συγκριτικά με τα προηγούμενα δύο είδη πρόβλεψης, όταν συνδυαστεί με κατάλληλα διαστήματα εμπιστοσύνης. Διευκολύνει τη δουλειά των διευθυνόντων, καθώς οι μέθοδοί της μπορούν να εφαρμοστούν εύκολα και παράγουν αποτελέσματα σε σύντομο χρονικό διάστημα, χωρίς να απαιτούνται ειδικές γνώσεις.

❖ Η κριτική πρόβλεψη μπορεί να βασίζεται στη γνώση, την κρίση και την πείρα ενός ατόμου ή μιας ομάδας ατόμων χωρίς την απαίτηση δεδομένων όπως γίνεται στη στατιστική πρόβλεψη. Ακόμη, μπορεί να λαμβάνει υπόψη ειδικά γεγονότα, να αντισταθμίζει τις ανεπάρκειες και να εξισορροπεί τις ανομοιογένειες των καταγεγραμμένων δεδομένων. Καθίσταται έτσι κατάλληλη σε περιπτώσεις που περισσότερο από τεχνοοικονομικούς παράγοντες έχουμε ηθικούς και όταν είναι επιθυμητός ο έλεγχος του μεγέθους του οποίου η ζήτηση προβλέπεται, ενώ το μειονέκτημά της είναι η προκατάληψη των ατόμων στη κρίση των οποίων στηρίζονται τα αποτελέσματά της. Για το λόγο αυτό κρίνεται σκόπιμος ο συνδυασμός των προβλέψεων από στατιστικές και κριτικές μεθόδους για την επίτευξη καλύτερων αποτελεσμάτων. Μερικές από τις βασικότερες μεθόδους της είναι η απλή κρίση, η μέθοδος Delphi και οι δομημένες αναλογίες.

❖ Τέλος, η πρόβλεψη στόχου είναι ουσιαστικά η επιθυμητή πορεία της χρονοσειράς που θέλει η επιχείρηση να έχει και όχι τόσο η πραγματική πρόβλεψή της. Με ανάλυση των ιστορικών δεδομένων του μεγέθους που μελετάται υπολογίζεται ο ρυθμός ανάπτυξής του για το παρελθόν διάστημα. Στη συνέχεια, η εφαρμογή των στατιστικών προβλέψεων δίνει μια εκτίμηση της πορείας της χρονοσειράς και πλέον μένει ο χειριστής των δεδομένων να τροποποιήσει το μελλοντικό ρυθμό ανάπτυξης, σύμφωνα με την επιθυμητή μελλοντική πορεία της επιχείρησης. Η τελική πρόβλεψη είναι συνεπώς εδώ η πρόβλεψη του διευθυντή ανάπτυξης (χειριστή) και περιέχει αρκετή αισιοδοξία και μεροληψία, καθώς καθορίζει τους στόχους της επιχείρησης.

Οι πιο βασικές στατιστικές μέθοδοι πρόβλεψης αναλύονται στη συνέχεια:

4.3 Στατιστικές μέθοδοι πρόβλεψης

4.3.1 Απλοϊκή μέθοδος Naive

Η απλοϊκή μέθοδος, Naive, αποτελεί την πιο απλή στατιστική μέθοδο. Έχει καθιερωθεί οι προβλέψεις που παράγει να χρησιμοποιούνται ως σημείο αναφοράς (benchmark) για τις υπόλοιπες περισσότερο πολύπλοκες μεθόδους, καθώς η ίδια δεν έχει ιδιαίτερα ακριβείς προβλέψεις στις περισσότερες των περιπτώσεων. Η προβλεπόμενη τιμή του υπό μελέτη μεγέθους για μια χρονική στιγμή προκύπτει στη Naive ως η πραγματική παρατήρηση της ακριβώς προηγούμενης χρονικής περιόδου [Πετρόπουλος και Ασημακόπουλος, 2011].

Ο τύπος αυτής της μεθόδου πρόβλεψης είναι ο εξής:

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

4.3.2 Μέθοδοι εκθετικής εξομάλυνσης

Οι επόμενες τρεις μέθοδοι στατιστικής πρόβλεψης υπάγονται στην κατηγορία της εκθετικής εξομάλυνσης. Η εκθετική εξομάλυνση ως μέθοδος πρόβλεψης παράγει τα αποτελέσματά της έπειτα από εξομάλυνση των δεδομένων για την απομόνωση των πραγματικών προτύπων της χρονοσειράς από τις διάφορες τυχαίες διακυμάνσεις. Θεωρείται, στις μεθόδους εκθετικής εξομάλυνσης ότι όσο πιο πρόσφατα τα δεδομένα τόσο περισσότερη η πληροφορία τους. Έτσι δίνεται μεγαλύτερη βαρύτητα στα πρόσφατα δεδομένα, η οποία (βαρύτητα) φθίνει εκθετικά όσο παλαιότερα είναι τα δεδομένα της χρονοσειράς.

Σχεδόν από την αρχή της εμφάνισής τους, οι συγκεκριμένες μέθοδοι αναδείχτηκαν ιδιαίτερα και αποτελούν τις δημοφιλέστερες μεθόδους για τους επιχειρηματίες λόγω της ευκολίας στη χρήση τους, της παραγωγής αποτελεσμάτων σε μικρό χρονικό διάστημα και της μικρής απαίτησης σε παρατηρήσεις για την παραγωγή των προβλέψεων. Αποδεικνύεται ότι παράγουν πιο ακριβείς προβλέψεις στην περίπτωση δεδομένων με στασιμότητα ή έστω μικρό ρυθμό ανάπτυξης. Ακόμη τα αποτελέσματα των μεθόδων εκθετικής εξομάλυνσης παρουσιάζουν ικανοποιητική ακρίβεια συγκριτικά με τις περισσότερο πολύπλοκες μεθόδους πρόβλεψης. Αυτό συμβαίνει επειδή οι ιδιομορφίες των προτύπων των δεδομένων ή οι όποιες διακυμάνσεις παρουσιάζονται εξομαλύνονται πριν την παραγωγή πρόβλεψης και έτσι οι μέθοδοι δεν επηρεάζονται από αυτές.

Οι πλέον καθιερωμένες μέθοδοι αυτής της κατηγορίας είναι η απλή εκθετική εξομάλυνση, η εκθετική εξομάλυνση γραμμικής τάσης και η εκθετική εξομάλυνση μη γραμμικής ή φθίνουσας τάσης.

4.3.2.1. Απλή εκθετική εξομάλυνση σταθερού επιπέδου (*Seasonal Exponential Smoothing*)

Η απλή εκθετική εξομάλυνση, SES, όπως καλείται το μοντέλο σταθερού επιπέδου, έχει ως βασική της υπόθεση την απουσία τάσης από τα δεδομένα και η εκάστοτε πρόβλεψη προκύπτει προεκτείνοντας μια οριζόντια ευθεία γραμμή. Για την παραγωγή πρόβλεψης με τη μέθοδο αυτή, χρειάζεται πρώτα να οριστεί ένα αρχικό επίπεδο της χρονοσειράς, (S_0). Κάθε νέα τιμή του προβλεπόμενου μεγέθους ισούται με το αμέσως προηγούμενο επίπεδο που υπολογίζεται. Ο υπολογισμός του επιπέδου χρονοσειράς γίνεται σε κάθε χρονική περίοδο και εξομαλύνεται από το σφάλμα της διαδικασίας προσαρμογής του μοντέλου, e_t . Η διαδικασία αυτή επαναλαμβάνεται για όλα τα καταγεγραμμένα δεδομένα της χρονοσειράς και οι προβλέψεις του μοντέλου ισούνται με την τελευταία υπολογισμένη τιμή του επιπέδου της χρονοσειράς, εφόσον θεωρείται η ύπαρξη σταθερού επιπέδου στο μοντέλο αυτό. Έτσι οι προβλέψεις είναι όλες επίπεδες και η απλή εκθετική εξομάλυνση θεωρείται χρήσιμη μέθοδος για προβλέψεις ενός βήματος ή όταν οι χρονοσειρές έχουν αυξημένο θόρυβο ή τυχαιότητα [Πετρόπουλος και Ασημακόπουλος, 2011].

Οι εξισώσεις που περιγράφουν το συγκεκριμένο μοντέλο πρόβλεψης είναι οι εξής:

$$\begin{aligned}e_t &= Y_t - F_t \\S_t &= S_{t-1} + \alpha \cdot e_t \\F_{t+1} &= S_t\end{aligned}$$

Ή αλλιώς, έχουμε την παρακάτω εξίσωση με ενσωμάτωση των τριών προηγούμενων, την οποία χρησιμοποιήσαμε και στη δική μας μοντελοποίηση:

$$F_t = \alpha \cdot Y_{t-1} + (1 - \alpha) \cdot F_{t-1}$$

Με F_t συμβολίζεται η προβλεπόμενη τιμή για τη χρονική στιγμή t . Και με α , ο συντελεστής εξομάλυνσης του επιπέδου της μεθόδου και λαμβάνει τιμές στο διάστημα $[0,1]$.

Το αρχικό επίπεδο της χρονοσειράς, S_0 , επιλέγεται ως ένα από τα παρακάτω:

- Ο μέσος όρος όλων των παρατηρήσεων
- Ο μέσος όρος ορισμένου πλήθους από τις αρχικές παρατηρήσεις
- Η πρώτη παρατήρηση
- Το σταθερό επίπεδο από το μοντέλο απλής γραμμικής παλινδρόμησης.

Η δεύτερη και η τελευταία επιλογή προτιμώνται, ως πιο αντιπροσωπευτικές του δείγματος της χρονοσειράς. Η επιλογή του αρχικού επιπέδου βασίζεται στα ποιοτικά χαρακτηριστικά της χρονοσειράς και επηρεάζει τόσο τις προβλεπόμενες τιμές όσο και την τιμή του συντελεστή εξομάλυνσης, α .

Είναι σημαντικό όπως αντιλαμβάνεται κανείς να καθοριστεί επιτυχώς και η τιμή του συντελεστή εξομάλυνσης, ώστε να παραχθούν οι καλύτερες δυνατές προβλέψεις. Η πλέον συνηθισμένη μέθοδος για τον υπολογισμό της βέλτιστης τιμής του συντελεστή α είναι η γραμμική αναζήτηση της τιμής του, ώστε να ελαχιστοποιείται το μέσο τετραγωνικό σφάλμα (Mean Square Error, MSE). Να αναφερθεί ότι ο συντελεστής α επηρεάζεται από το ποσοστό θορύβου της χρονοσειράς. Όσο μεγαλύτερο το ποσοστό του θορύβου τόσο μικρότερη τιμή του συντελεστή α επιθυμούμε και αντίστροφα. Ακόμη, σημασία έχει και η σταθερότητα του μέσου όρου της χρονοσειράς καθώς, αν η σταθερότητά του είναι μικρή, είναι επιθυμητές μεγάλες τιμές για το συντελεστή α . Αξίζει να σημειωθεί ότι για τιμές του α κοντά στη μονάδα, αποδίδεται μεγαλύτερη βαρύτητα στα πιο πρόσφατα δεδομένα με ακραία την περίπτωση όπου $\alpha = 1$ και η μέθοδος αποτελεί ουσιαστικά την απλοϊκή μέθοδο πρόβλεψης (Naive). Από την άλλη πλευρά, στην περίπτωση που $\alpha = 0$, οι προβλεπόμενες τιμές ισούνται όλες με το αρχικό επίπεδο της χρονοσειράς, S_0 .

4.3.2.2 Εκθετική εξομάλυνση γραμμικής τάσης (Holt Exponential Smoothing)

Το μοντέλο γραμμικής τάσης αποτελεί προέκταση του μοντέλου της απλής εκθετικής εξομάλυνσης, στο οποίο μπορεί πλέον να γίνει διαχείριση της συνιστώσας της τάσης που παρατηρείται στα δεδομένα πολλών χρονοσειρών.

Οι εξισώσεις που περιγράφουν το μοντέλο αυτό παρατίθενται αμέσως μετά [Taylor, 2003a]:

$$\begin{aligned} S_t &= \alpha \cdot X_t + (1 - \alpha) \cdot (S_{t-1} + T_{t-1}) \\ T_t &= \gamma \cdot (S_t - S_{t-1}) + (1 - \gamma) \cdot T_{t-1} \\ F_{t+m} &= S_t + m \cdot T_t \end{aligned}$$

Με e συμβολίζεται το σφάλμα ως η απόκλιση της πραγματικής τιμής από την προβλεπόμενη, με X_t η πραγματική τιμή της χρονοσειράς (την οποία έχουμε πλέον διαθέσιμη), με S το επίπεδο της χρονοσειράς, με T η τάση της χρονοσειράς, με F η προβλεπόμενη τιμή, με α ο συντελεστής εξομάλυνσης του επιπέδου και με γ ο συντελεστής

εξομάλυνσης της τάσης της χρονοσειράς. Τέλος, με t συμβολίζεται η χρονική περίοδος και με m ο χρονικός ορίζοντας της πρόβλεψης.

Η τιμή του αρχικού επιπέδου της χρονοσειράς, S_0 , υπολογίζεται όπως και στη απλή μέθοδο εκθετικής εξομάλυνσης, ενώ ως τιμή της αρχικής τάσης, T_0 , λαμβάνεται ένα από τα παρακάτω [Πετρόπουλος και Ασημακόπουλος, 2011]:

- Η διαφορά της δεύτερης και της πρώτης παρατήρησης, $(Y_2 - Y_1)$
- Η διαφορά μιας τυχαίας παρατήρησης (n) της χρονοσειράς με την προηγούμενη ακριβώς παρατήρησή της ($n-1$), διαιρεμένης με την απόσταση των παρατηρήσεων, δηλαδή με $n-1$, $(Y_n - Y_{n-1}) / (n-1)$
- Η σταθερά της κλίσης από το μοντέλο απλής γραμμικής παλινδρόμησης.

Και στη μέθοδο αυτή η επιλογή των αρχικών τιμών των S και T είναι σημαντικό κομμάτι της διαδικασίας που ακολουθείται, εφόσον επηρεάζουν το μοντέλο και τις παραγόμενες προβλέψεις και γίνεται με βάση τα ποιοτικά χαρακτηριστικά της χρονοσειράς.

Οι συντελεστές α και γ λαμβάνουν τιμές στο διάστημα $[0,1]$ και η εύρεση της βέλτιστης τιμής τους στηρίζεται, όπως και στην προηγούμενη μέθοδο, στην ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Είναι σύνηθες η βέλτιστη τιμή του συντελεστή εξομάλυνσης τάσης, γ , να είναι μικρότερη από την αντίστοιχη βέλτιστη τιμή του συντελεστή εξομάλυνσης επιπέδου, α . Αυτό συμβαίνει επειδή η τιμή της τάσης κάθε περιόδου είναι συνήθως πολύ μικρότερη από την τιμή του επιπέδου της χρονοσειράς.

Πλέον οι προβλέψεις δεν είναι σταθερές όπως στην απλή εκθετική εξομάλυνση, αλλά εμφανίζουν τάση, όπως φαίνεται και από την τελευταία εξίσωση, της προβλεπόμενης τιμής. Έχει παρατηρηθεί ότι το συγκεκριμένο μοντέλο πρόβλεψης παράγει αρκετές φορές ιδιαίτερα αισιόδοξα αποτελέσματα σε σχέση με τις πραγματικές τιμές που καταγράφονται, ιδίως για μεσοπρόθεσμες και μακροπρόθεσμες προβλέψεις.

4.3.2.3 Εκθετική εξομάλυνση μη γραμμικής τάσης

Για τις περιπτώσεις στις οποίες το μοντέλο εξομάλυνσης για γραμμική τάση αποτυγχάνει να αποτυπώσει σωστά τη μελλοντική πορεία των χρονοσειρών (εξαιτίας της υπερεκτίμησης που το χαρακτηρίζει), αναπτύχθηκε το μοντέλο μη γραμμικής τάσης από τους Gardner και McKenzie το 1985, το οποίο προκαλεί βαθμιαία μείωση του μεγέθους αύξησης των τιμών της χρονοσειράς σε κάθε χρονική περίοδο.

Οι εξισώσεις που περιγράφουν το μοντέλο εκθετικής εξομάλυνσης μη γραμμικής τάσης είναι οι εξής [Taylor, 2003a]:

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot (S_{t-1} + \varphi \cdot T_{t-1})$$

$$T_t = \gamma \cdot (S_t - S_{t-1}) + (1 - \gamma) \cdot \varphi \cdot T_{t-1}$$

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

Οι περισσότεροι συμβολισμοί είναι γνωστοί ήδη από τις προηγούμενες μεθόδους εκθετικής εξομάλυνσης. Συγκεκριμένα, με e συμβολίζεται η απόκλιση της πραγματικής τιμής από την προβλεπόμενη, με X_t η πραγματική τιμή της χρονοσειράς την τρέχουσα περίοδο t , με S το επίπεδο της χρονοσειράς, με T η τάση της χρονοσειράς, με F η προβλεπόμενη τιμή, με α ο συντελεστής εξομάλυνσης του επιπέδου και με γ ο συντελεστής εξομάλυνσης της τάσης της

χρονοσειράς. Και, όπως ήδη σημειώθηκε, με t συμβολίζεται η χρονική περίοδος, ενώ με m ο χρονικός ορίζοντας της πρόβλεψης.

Υπεισέρχεται όμως πλέον μία παράμετρος που ελέγχει το ρυθμό αύξησης των τιμών για την αφαίρεση της θετικής προκατάληψης (που παρουσιάζει το μοντέλο γραμμικής τάσης). Η παράμετρος αυτή ονομάζεται παράμετρος διόρθωσης της τάσης, ϕ , η οποία μπορεί να λάβει τιμές θετικές (ή και τη μηδενική τιμή) χωρίς άνω περιορισμό, σε αντίθεση με τους συντελεστές α και γ . Είναι, ωστόσο, σημαντική η κατηγοριοποίηση της μορφής του μοντέλου, ανάλογα με τα διαστήματα τιμών της παραμέτρου ϕ :

- ✓ Για $\phi = 0$ προκύπτει το μοντέλο απλής εκθετικής εξομάλυνσης, καθώς η τάση T δεν έχει κανένα ρόλο στον υπολογισμό της πρόβλεψης.
- ✓ Για $0 < \phi < 1$ προκύπτει το μοντέλο φθίνουσας τάσης (Damped Exponential Smoothing), το οποίο μειώνει σταδιακά την επίδραση της τάσης στο αποτέλεσμα της πρόβλεψης και ελέγχει με τον τρόπο αυτό το ρυθμό αύξησης των τιμών τους.
- ✓ Για $\phi = 1$ προκύπτει το μοντέλο γραμμικής τάσης όπως φαίνεται από την τελευταία εξίσωση.
- ✓ Για $\phi > 1$ προκύπτει το μοντέλο εκθετικής τάσης, το οποίο είναι υπεραισιόδοξο και οδηγεί σε υψηλές τιμές πρόβλεψης. Έτσι συνίσταται η χρήση του μόνο για πρόβλεψη ζήτησης ενός προϊόντος στα αρχικά στάδια εμφάνισής του στην αγορά.

Για την αποφυγή της παραγωγής προκατειλημμένων προβλέψεων λόγω λανθασμένης βελτιστοποίησης του συντελεστή ϕ και για επίτευξη καλύτερων αποτελεσμάτων, περιορίζεται το διάστημα τιμών της παραμέτρου ϕ στο διάστημα $[0,1]$. Η επιλογή των αρχικών τιμών του επιπέδου και της τάσης της χρονοσειράς συνίσταται να γίνεται με εφαρμογή της γραμμικής παλινδρόμησης και ο υπολογισμός του βέλτιστου συνδυασμού των συντελεστών α , γ και ϕ με τη διαδικασία ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος, όπως και στις προηγούμενες μεθόδους.

Έχει παρατηρηθεί ότι το μοντέλο μη γραμμικής τάσης παράγει πιο ακριβείς προβλέψεις σε σχέση με το μοντέλο γραμμικής τάσης, που είναι όμως πιο συχνά χρησιμοποιούμενο. Μπορεί να πει κανείς ότι το μοντέλο μη γραμμικής τάσης δίνει ικανοποιητικά αποτελέσματα σε περιπτώσεις μεγάλου πλήθους χρονοσειρών, οπότε και απαιτείται πλήρως αυτοματοποιημένο σύστημα στην παραγωγή προβλέψεων και μάλιστα όσο μεγαλύτερος ο χρονικός ορίζοντας πρόβλεψης, τόσο καλύτερο είναι το μοντέλο αυτό (συγκριτικά με τις υπόλοιπες μεθόδους εκθετικής εξομάλυνσης). Τέλος, ικανοποιητικά αποτελέσματα δίνει και σε περιπτώσεις αδυναμίας επιλογής ορισμένου μοντέλου πρόβλεψης.

4.3.3 Μοντέλα παλινδρόμησης

Τα μοντέλα παλινδρόμησης (Regression Models) είναι μία ακόμη μέθοδος στατιστικής πρόβλεψης και διακρίνονται σε μοντέλα απλής ή πολλαπλής γραμμικής παλινδρόμησης [Πετρόπουλος και Ασημακόπουλος, 2011], ανάλογα με το πλήθος των ανεξάρτητων μεταβλητών τους (αν περιλαμβάνουν μία ή περισσότερες ανεξάρτητες μεταβλητές αντίστοιχα). Βασική χρήση των μοντέλων παλινδρόμησης είναι η ανάλυση των σχέσεων ανάμεσα στην εξαρτημένη μεταβλητή (τη μεταβλητή πρόβλεψης) και τις ανεξάρτητες μεταβλητές, παρότι χρησιμοποιούνται πολύ και στις προβλέψεις. Συντελεί δηλαδή στην κατανόηση της αλλαγής της εξαρτημένης μεταβλητής με την οποιαδήποτε αλλαγή μίας από των ανεξάρτητων μεταβλητών και με τις υπόλοιπες να παραμένουν σταθερές. Ακόμη, με ορισμένες τιμές για όλες τις ανεξάρτητες μεταβλητές, εκτιμάται η μέση τιμή της αναμενόμενης εξαρτημένης μεταβλητής μέσω της εξίσωσης παλινδρόμησης. Ενδιαφέρουσα

είναι και η διακύμανση των τιμών που λαμβάνει η εξαρτημένη μεταβλητή γύρω από την εξίσωση παλινδρόμησης, η οποία μπορεί να περιγραφεί μέσω πιθανοτικής κατανομής. Η κύρια μέθοδος αυτής της κατηγορίας είναι η απλή γραμμική παλινδρόμηση.

4.3.3.1 Απλή γραμμική παλινδρόμηση

Η απλή γραμμική παλινδρόμηση στηρίζεται στην υπόθεση ύπαρξης ορισμένης σχέσης ανάμεσα σε μια εξαρτημένη μεταβλητή (τη μεταβλητή πρόβλεψης) και σε μια ανεξάρτητη μεταβλητή (τον χρόνο στην περίπτωση χρονοσειράς). Ακόμη υποτίθεται ότι η σχέση των δύο μεταβλητών είναι γραμμική, παρότι σε πολλές περιπτώσεις η υπόθεση αυτή δεν ανταποκρίνεται στην πραγματικότητα. Η παρακάτω εξίσωση περιγράφει τη σχέση αυτή:

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

Με a συμβολίζεται η τεταγμένη του σημείου τομής της ευθείας με τον άξονα των εξαρτημένων μεταβλητών, με b η κλίση της ευθείας, με X η ανεξάρτητη μεταβλητή -στην περίπτωση πρόβλεψης χρονοσειρών είθισται να είναι ο χρόνος- και με \hat{Y} η εξαρτημένη μεταβλητή, δηλαδή η τιμή πρόβλεψης. Οι τιμές των συντελεστών a και b υπολογίζονται με βάση την αρχή των ελαχίστων τετραγώνων, επιλέγονται δηλαδή οι συντελεστές που ελαχιστοποιούν το άθροισμα των τετραγώνων των διαφορών των πραγματικών τιμών από τις προβλεπόμενες σε κάθε χρονική περίοδο, όπως φαίνεται και στην επόμενη σχέση:

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

Με βάση τη λογική της ελαχιστοποίησης της απόστασης των πραγματικών παρατηρήσεων Y από τη βέλτιστη γραμμή παλινδρόμησης προκύπτουν οι εξισώσεις υπολογισμού των συντελεστών a και b , όπως παρουσιάζονται στη συνέχεια:

$$b = \frac{\sum_{i=1}^n [(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad \text{και} \quad a = \bar{Y} - b \cdot \bar{X}$$

Με \bar{X} και \bar{Y} συμβολίζονται οι μέσες τιμές των μεταβλητών X και Y αντίστοιχα, με n ο αριθμός των παρατηρήσεων, βάση των οποίων υπολογίζεται η ευθεία παλινδρόμησης και με \hat{Y} συμβολίζεται η εκτίμηση της εξαρτημένης μεταβλητής μέσω της ευθείας παλινδρόμησης.

Η πρόβλεψη με χρήση τη μέθοδο της απλής γραμμικής παλινδρόμησης δίνει μια καλή εικόνα της μέσης και της μακροπρόθεσμης συμπεριφοράς του υπό μελέτη μεγέθους.

Σε περίπτωση που η σχέση ανάμεσα σε δύο μεταβλητές, την εξαρτημένη και την ανεξάρτητη, δεν είναι γραμμική, μπορεί και πάλι να εφαρμοστεί η μέθοδος της απλής γραμμικής παλινδρόμησης, αφού πρώτα γίνει μετασχηματισμός της σχέσης των δύο μεταβλητών σε γραμμική. Ακόμη εφόσον η απλή γραμμική παλινδρόμηση στηρίζεται στην ύπαρξη σχέσης ανάμεσα στις δύο μεταβλητές, μας ενδιαφέρει ο έλεγχος του βαθμού συσχέτισής τους. Αυτό το πετυχαίνουμε με τη χρήση του συντελεστή γραμμικής συσχέτισης των δύο μεταβλητών, ο οποίος λαμβάνει τιμές στο διάστημα $[-1, 1]$. Για μηδενική τιμή δεν υπάρχει καμία σχέση ανάμεσα στις δύο μεταβλητές. Για τιμή ± 1 υπάρχει απόλυτη συσχέτιση ανάμεσα στις δύο μεταβλητές και γενικά για θετικές τιμές του δείκτη οι μεταβλητές αυξάνονται ή μειώνονται συγχρόνως, ενώ για αρνητικές τιμές του δείκτη συσχέτισης, η αυξομειώσεις τους είναι αντίθετης κατεύθυνσης. (Σημείωση: Ο συντελεστής γραμμικής συσχέτισης έχει αναλυθεί στην παράγραφο 3.3.)

4.3.3.2 Πολλαπλή γραμμική παλινδρόμηση

Η πολλαπλή γραμμική παλινδρόμηση εφαρμόζεται για τη συμπερίληψη περισσότερων από μία ανεξάρτητων μεταβλητών στις περιπτώσεις που αυτό απαιτείται. Ζητείται η πρόβλεψη μιας εξαρτημένης μεταβλητής βάσει των πολλών ανεξάρτητων μεταβλητών. Η γενική μορφή της πολλαπλής παλινδρόμησης δίνεται από την εξής σχέση:

$$Y = b_0 + b_1 \cdot X_{1,i} + b_2 \cdot X_{2,i} + \dots + b_k \cdot X_{k,i} + e_i$$

Με Y συμβολίζεται η εξαρτημένη μεταβλητή, με X οι ανεξάρτητες μεταβλητές, με b_i οι συντελεστές (σταθερές παράμετροι) και με e ο τυχαίος παράγοντας που θεωρείται κανονικά κατανομημένος γύρω από το μηδέν. Η συνάρτηση της πολλαπλής γραμμικής παλινδρόμησης είναι γραμμική ως προς τους συντελεστές, καθώς οι εκθέτες όλων ισούνται με τη μονάδα και οι τιμές των συντελεστών μπορούν να υπολογιστούν με τη μέθοδο των ελαχίστων τετραγώνων.

Στην περίπτωση τα πολλαπλής γραμμικής παλινδρόμησης είναι δύσκολος ο προσδιορισμός της μορφής της συνάρτησης περιγραφής της σχέσης μεταξύ των δύο μεταβλητών και δίνεται έμφαση στον υπολογισμό των συντελεστών b_i και της διακύμανσης του τυχαίου παράγοντα e , έχοντας ορισμένο σύνολο δεδομένων εφαρμογής της μεθόδου των ελαχίστων τετραγώνων. Συγκεκριμένα για τον υπολογισμό των συντελεστών b_i υπολογίζονται για κάθε συντελεστή οι μερικές παράγωγοι του αθροίσματος του τετραγώνου των σφαλμάτων $e_i = Y_i - \hat{Y}_i$, με \hat{Y}_i την εκτίμηση της τιμής της εξαρτημένης μεταβλητής Y . Στη συνέχεια τίθενται οι παράγωγοι αυτές ίσες με το μηδέν και λύνεται ένα γραμμικό σύστημα εξισώσεων με αριθμό αγνώστων ίσο με τον αριθμό των συντελεστών b_i .

4.3.4 Μέθοδος Theta

Η μέθοδος πρόβλεψης Theta [Assimakopoulos and Nikolopoulos, 2000] αποτελεί μια μονοδιάστατη μέθοδο, που βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσω της παραμέτρου θ , η οποία και εφαρμόζεται πολλαπλασιαστικά στις διαφορές δεύτερης τάξης των δεδομένων, όπως φαίνεται στην παρακάτω σχέση:

$$Y_t^\theta = \theta \cdot Y_t'', \text{ με το } Y_t'' \text{ να ισούται με } Y_t'' = Y_t - 2 \cdot Y_{t-1} + Y_{t-2}$$

Η ιδιαίτερη προσέγγιση της μεθόδου έγκειται στο γεγονός ότι διαχωρίζει τα αποεποχικοποιημένα δεδομένα σε γραμμές ξεχωριστές, με έμφαση στη βραχυπρόθεσμη και στη μακροπρόθεσμη τάση τους, ανάλογα με την τιμή της παραμέτρου θ . Συγκεκριμένα, η μείωση της τιμής της παραμέτρου θ επιτυγχάνει αύξηση του βαθμού μείωσης των καμπυλοτήτων, μέχρι που για μηδενική τιμή της παραμέτρου η μέθοδος παράγει γραμμή που ισοδυναμεί με την ευθεία των ελαχίστων τετραγώνων (LRL). Αντίθετα, για τιμές του θ μεγαλύτερες της μονάδας ενισχύονται οι τοπικές καμπυλότητες της χρονοσειράς, αυξάνεται ο βαθμός διόγκωσής της και έτσι προκαλείται τόνωση και της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς. Η παράμετρος θ λαμβάνει και αρνητικές τιμές, όπως για παράδειγμα $\theta = -1$, η οποία αντιστοιχεί στη συμμετρική χρονοσειρά της αρχικής ως προς την ευθεία γραμμικής παλινδρόμησης.

Έπειτα από την αποσύνθεση της αρχικής χρονοσειράς σε δύο ή περισσότερες γραμμές Theta, προεκτείνεται η καθεμία ξεχωριστά στο μέλλον με οποιαδήποτε στατιστική μέθοδο πρόβλεψης έχει ήδη αναφερθεί και με συνδυασμό των διαφορετικών παραγόμενων προβλέψεων προκύπτει η τελική πρόβλεψη. Να αναφερθεί ότι κάθε γραμμή Theta διατηρεί

τη μέση τιμή, αλλά και την κλίση της παλινδρόμησης της αρχικής χρονοσειράς, ενώ δέχονται μεταβολή οι καμπυλότητες και η διακύμανσή της.

Η απλοποιημένη εξίσωση που περιγράφει την κάθε γραμμή Theta είναι η εξής:

$$Y_t^\theta = \theta \cdot Y_t + (1 - \theta) \cdot LRL_t$$

Ενώ το κλασικό μοντέλο Theta, που θα χρησιμοποιηθεί και στην παρούσα εργασία, συνδυάζει τις γραμμές Theta με $\theta=0$ και $\theta=2$ με ίσα βάρη, ώστε να παραχθεί η ζητούμενη πρόβλεψη. Αυτή είναι και η απλούστερη περίπτωση της συγκεκριμένης μεθόδου, στην οποία η γραμμή με $\theta = 0$ παρέχει πληροφορία για τη μακροπρόθεσμη τάση της χρονοσειράς και η γραμμή με $\theta = 2$ παρέχει πληροφορία για τη βραχυπρόθεσμη τάση της αντίστοιχα. Και η εξίσωση από την οποία προκύπτει η τελική πρόβλεψη της μεθόδου είναι η ακόλουθη:

$$Y_t = \frac{1}{2} \cdot (Y_t^{\theta=1+\alpha} + Y_t^{\theta=1-\alpha})$$

Η μέθοδος αυτή στόχο είχε την αύξηση του βαθμού αξιοποίησης της χρήσιμης πληροφορίας που βρίσκεται στα δεδομένα, πριν εφαρμοστεί σε αυτά κάποιο μοντέλο επέκτασής τους στο μέλλον. Μέσω των διαδικασιών της μεθόδου αυτής, γίνεται πιο αποδοτικός ο γραμμικός συνδυασμός των προβλέψεων των συνιστωσών. Καταλήγοντας, συμπεραίνεται ότι η μέθοδος Theta οδηγεί σε καλύτερα αποτελέσματα όσον αφορά μακροπρόθεσμες προβλέψεις της εξέλιξης των δεδομένων ή ενίσχυση της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς ανάλογα με την τιμή της παραμέτρου θ , όπως προαναφέρθηκε.

4.3.4 Μοντέλα ARIMA

Τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητών μέσων όρων (Auto-Regressive Integrated Moving Average, ARIMA) είναι στοχαστικά μαθηματικά μοντέλα, τα οποία δίνουν μια καλή εικόνα της διαχρονικής εξέλιξης της χρονοσειράς, παράγοντας ικανοποιητικές προβλέψεις των μελλοντικών τιμών του υπό μελέτη μεγέθους. Η διαχρονική περιγραφή ενός φυσικού μεγέθους είναι συχνά πολύ δύσκολη μέσω ντετερμινιστικού μοντέλου και ακόμη η εξάρτηση του φυσικού μεγέθους από μη ντετερμινιστικούς παράγοντες οδηγούν στην περιγραφή της διαχρονικής εξέλιξης του μεγέθους μέσω στοχαστικού μοντέλου. Οι Box και Jenkins [Box and Jenkins, 1970] μελέτησαν σε βάθος τα μοντέλα ARIMA και πρότειναν μια ομάδα αλγεβρικών μοντέλων πρόβλεψης, οι προβλέψεις των οποίων στηρίζονται αποκλειστικά στις παρελθοντικές τιμές και το μοτίβο συμπεριφοράς της χρονοσειράς.

4.3.5 Νευρωνικά δίκτυα

Τα νευρωνικά δίκτυα (Neural Networks) έχουν ως βασική ιδέα το φιλτράρισμα των ανεξάρτητων μεταβλητών (είσοδοι του μοντέλου), μέσω ενός ή περισσότερων κρυφών επιπέδων αποτελούμενων από κρυφούς κόμβους, πριν την παραγωγή των προβλέψεων (έξοδοι του μοντέλου). Παρέχουν διευκόλυνση στη μελέτη χρονοσειρών, των οποίων οι μεταβλητές εισόδου και εξόδου δεν έχουν γραμμική σχέση μεταξύ τους.

Να αναφερθεί ότι στη συγκεκριμένη εργασία μελετώνται τα εξής στατιστικά μοντέλα:

- Η απλοϊκή μέθοδος, Naive
- Η απλή εκθετική εξομάλυνση, Simple Exponential Smoothing
- Η εκθετική εξομάλυνση για γραμμική τάση, Holt Exponential Smoothing

- Η εκθετική εξομάλυνση μη γραμμικής τάσης (φθίνουσας για την ακρίβεια),
Damped Exponential Smoothing
- Η κλασσική μέθοδος Theta

4.4 Εφαρμογή εποχικότητας στα χρησιμοποιούμενα μοντέλα

Όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, οι καμπύλες ζήτησης φορτίου ηλεκτρικής ενέργειας παρουσιάζουν έντονη ημερήσια και εβδομαδιαία εποχικότητα. Η επίδραση αυτού του χαρακτηριστικού τους μας είναι το επίκεντρο του ενδιαφέροντος στην παρούσα εργασία και κρίνεται απαραίτητο να ληφθεί υπόψη στα μοντέλα πρόβλεψης που εφαρμόζονται επί της διαθέσιμης χρονοσειράς. Συγκεκριμένα, τα δεδομένα της χρονοσειράς που μελετάται, υφίστανται στο πρώτο στάδιο αποεποχικοποίηση πριν εφαρμοστούν επί αυτών οι διάφορες μέθοδοι πρόβλεψης που αναλύθηκαν στην παράγραφο 4.3.

Η διαδικασία της αποεποχικοποίησης έχει ως εξής:

- Αρχικά υπολογίζεται ο κεντρικός κινητός μέσος όρος (ΚΚΜΟ) των δεδομένων που χρησιμοποιούνται με μήκος ίσο με την περίοδο της εποχικότητας που εμφανίζεται, δηλαδή ημέρα ή εβδομάδα. Στην περίπτωση που τα δεδομένα είναι καταγεγραμμένα ανά ημίωρο, υπολογίζονται δηλαδή οι κεντρικοί κινητοί μέσοι όροι με μήκος $n_1 = 48$ και με μήκος $n_2 = 336$, εφόσον σε μία ημέρα αντιστοιχούν σαράντα οκτώ παρατηρήσεις και σε μία εβδομάδα τριακόσιες τριάντα έξι παρατηρήσεις της χρονοσειράς. Αντίστοιχα, για ωριαία καταγραφή των δεδομένων, γίνεται υπολογισμός κεντρικών κινητών μέσων όρων με μήκη $n_1 = 24$ και $n_2 = 168$, όπως αντιλαμβάνεται εύκολα κανείς. Οι υπολογισμένοι κεντρικοί κινητοί μέσοι όροι δίνουν μια καλή εικόνα της χρονοσειράς ως προς την τάση και την κυκλικότητά της.
- Στη συνέχεια διαιρούνται τα δεδομένα της αρχικής χρονοσειράς με τις αντίστοιχες τιμές των ΚΚΜΟ και προκύπτουν οι λόγοι εποχικότητας, οι οποίοι χαρακτηρίζονται όμως από τυχαιότητα και για το λόγο αυτό υπολογίζονται οι μέσες τιμές των αντίστοιχων λόγων εποχικότητας. Για παράδειγμα, σε περίπτωση ΚΚΜΟ μήκους 12, υπολογίζονται δώδεκα μέσες τιμές των λόγων εποχικότητας. Αυτές οι μέσες τιμές αποτελούν τους δείκτες εποχικότητας και χρειάζεται συνήθως να γίνει και κανονικοποίησή τους για να ληφθούν οι τελικοί κανονικοποιημένοι δείκτες εποχικότητας.
- Τέλος, η διαίρεση των αρχικών δεδομένων με τους κανονικοποιημένους δείκτες εποχικότητας οδηγεί στη λήψη των αποεποχικοποιημένων δεδομένων της σειράς.

Έπειτα από την ολοκλήρωση της διαδικασίας αποεποχικοποίησης των χρησιμοποιούμενων δεδομένων της χρονοσειράς και ενώ έχει γίνει καταγραφή των δεικτών εποχικότητας, τα δεδομένα είναι έτοιμα για την εφαρμογή της εκάστοτε στατιστικής μεθόδου επί των δεδομένων αυτών και την πρόβλεψη της ζήτησης ηλεκτρικού φορτίου από ένα έως και περισσότερα βήματα στο μέλλον. Η βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου γίνεται με ορίζοντα από μερικές ώρες ως και μία εβδομάδα και οι παραγόμενες τιμές πολλαπλασιάζονται με τους αντίστοιχους κανονικοποιημένους δείκτες εποχικότητας για τον υπολογισμό της τελικής πρόβλεψης.

4.5 Η προσαρμοσμένη μέθοδος Holt-Winters για την ενσωμάτωση διπλής εποχικότητας (από τον J. W. Taylor)

Μελετώντας τη βραχυπρόθεσμη πρόβλεψη ζήτησης ηλεκτρικού φορτίου ο Taylor παρατήρησε την ανάγκη υποστήριξης περισσότερων από μιας εποχικότητας, η οποία αποτυπώνεται σε διάφορα πολυσύνθετα μοντέλα πρόβλεψης, όπως το μοντέλο ARIMA πολλαπλασιαστικής εποχικότητας. Ωστόσο, μέθοδοι όπως αυτή απαιτούν αρκετά δεδομένα ως εισόδους, αποδεικνύεται χρονοβόρα η παραγωγή αποτελεσμάτων και η πολυπλοκότητά τους δεν οδηγεί πάντοτε σε αντιπροσωπευτικές προβλέψεις. Σε αντίθεση με τα πολύπλοκα μοντέλα πρόβλεψης, οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα διαδεδομένες λόγω της ακρίβειας που παρουσιάζουν στα αποτελέσματά τους και της ευκολίας χρήσης τους για την αυτοματοποιημένη παραγωγή προβλέψεων. Έτσι ο Taylor προσαρμοσε τη μέθοδο Holt-Winters exponential smoothing [Taylor, 2003a] και [Taylor et al., 2006], ώστε να υποστηρίξει πλέον και τις δύο εποχικότητες των χρονοσειρών ζήτησης φορτίου, ημερήσια και εβδομαδιαία (είναι ο πρώτος που εισήγαγε περισσότερα από ένα μοτίβα εποχικότητας σε μοντέλο εκθετικής εξομάλυνσης). Να σημειωθεί ότι οι Holt-Winters είχαν ήδη προσαρμόσει το βασικό στατιστικό μοντέλο γραμμικής εκθετικής εξομάλυνσης, ώστε να ενσωματώνει την επίδραση μίας εποχικότητας.

Οι επόμενες εξισώσεις περιγράφουν το βασικό μοντέλο εκθετικής εξομάλυνσης των Holt-Winters πριν τη μετατροπή από τον Taylor:

Standard Holt-Winters method:

$$\text{Level} \quad : S_t = a \cdot (X_t / I_{t-s}) + (1-a) \cdot (S_{t-1} + T_{t-1})$$

$$\text{Trend} \quad : T_t = \gamma \cdot (S_t - S_{t-1}) + (1-\gamma) \cdot T_{t-1}$$

$$\text{Seasonality: } I_t = \delta \cdot (X_t / S_t) + (1-\delta) \cdot I_{t-s}$$

$$\text{Forecast} \quad : F_t(k) = (S_t + k \cdot T_t) \cdot I_{t-s+k}$$

Υπάρχουν δύο πρότυπα εποχικότητας, το πολλαπλασιαστικό και το αθροιστικό, με το πρώτο να είναι κατάλληλο σε περίπτωση που μια αύξηση του μέσου επιπέδου τιμών της χρονοσειράς προκαλεί ένταση της εποχικής συμπεριφοράς, ενώ το δεύτερο σε περίπτωση που η επίδραση της εποχικότητας δεν επηρεάζεται από το τρέχον μέσο επίπεδο. Στην εκδοχή του μοντέλου Holt-Winters χρησιμοποιείται το πρότυπο πολλαπλασιαστικής εποχικότητας, με προσθετική τάση και ο υπολογισμός της κλίσης γίνεται με εξομάλυνση των διαφορών του επιπέδου S από τη μία χρονική περίοδο ως την επόμενη. Με a , γ , και δ συμβολίζονται οι συντελεστές εξομάλυνσης του επιπέδου, της κλίσης και της εποχικής περιόδου αντίστοιχα, ενώ με s η περίοδος της εποχικότητας της χρονοσειράς (περίπτωση μίας εποχικότητας). Με S , T και I συμβολίζονται το επίπεδο, η κλίση (τάση) και η εποχικότητα της χρονοσειράς αντίστοιχα. Ο δείκτης εποχικότητας, I , υπολογίζεται με εξομάλυνση του λόγου της παρατήρησης X προς το επίπεδο S για κάθε χρονική περίοδο t .

Στη συνέχεια παρατίθενται οι εξισώσεις που περιγράφουν το τροποποιημένο μοντέλο των Holt-Winters από τον Taylor για διπλής εποχικότητας χρονοσειρά. Η αρχική μορφή του μοντέλου που παρουσίασε ο Taylor έχει μία διαφορά από αυτή που φαίνεται παρακάτω. Για την ακρίβεια, υπολείπεται ο όρος $\phi^k \cdot (X_t - ((S_{t-1} + T_{t-1}) \cdot D_{t-s1} \cdot W_{t-s2}))$ από την εξίσωση πρόβλεψης. Σύμφωνα με τη συγκεκριμένη μελέτη του Taylor, για την επίτευξη ικανοποιητικών τελικών προβλέψεων απαιτείται μία ακόμη προσαρμογή στο αρχικό του μοντέλο, με τη χρήση ενός μοντέλου AR (Autoregressive model). Η προσαρμογή αυτή αποτυπώνεται στην εξίσωση πρόβλεψης με την προσθήκη ενός ακόμη όρου, της μορφής $\lambda^k \cdot e_t$ (ο δεύτερος όρος που φαίνεται στην εξίσωση). Επομένως, ο δεύτερος όρος της τελευταίας εξίσωσης αποτελεί την προσαρμογή για πρώτης τάξεως αυτοσυσχέτιση.

Holt-Winters method for double seasonality:

Level : $S_t = a \cdot (X_t / (D_{t-s_1} W_{t-s_2})) + (1-a) \cdot (S_{t-1} + T_{t-1})$

Trend : $T_t = \gamma \cdot (S_t - S_{t-1}) + (1-\gamma) \cdot T_{t-1}$

Seasonality 1: $D_t = \delta \cdot (X_t / (S_t \cdot W_{t-s_2})) + (1-\delta) \cdot D_{t-s_1}$

Seasonality 2: $W_t = \omega \cdot (X_t / (S_t \cdot D_{t-s_1})) + (1-\omega) \cdot W_{t-s_2}$

Forecast : $F_t(k) = (S_t + k \cdot T_t) \cdot D_{t-s_1+k} \cdot W_{t-s_2+k}$
 $+ \phi^k \cdot (X_t - ((S_{t-1} + T_{t-1}) \cdot D_{t-s_1} \cdot W_{t-s_2}))$

Η ημερήσια εποχικότητα συμβολίζεται με D_t και η εβδομαδιαία με W_t . Με X συμβολίζονται τα δεδομένα της αρχικής χρονοσειράς, με a , γ , δ , ω οι παράμετροι εξομάλυνσης, με ϕ συμβολίζεται η παράμετρος αυτοσυσχέτισης, ενώ με s_1 και s_2 οι χρονικές περίοδοι των δύο εποχικότητων. Η πρόβλεψη βήματος k από την τρέχουσα χρονική περίοδο, t , συμβολίζεται με $F_t(k)$. Ο δείκτης ημερήσιας εποχικότητας (με περίοδο s_1), D_t , υπολογίζεται με εξομάλυνση του λόγου της πραγματικής παρατήρησης, X_t , προς το γινόμενο της τιμής του επιπέδου της χρονοσειράς, S_t , επί το δείκτη εβδομαδιαίας εποχικότητας (με περίοδο s_2), W_{t-s_2} , για κάθε χρονική περίοδο, t και ο δείκτης εβδομαδιαίας εποχικότητας, W_t , υπολογίζεται αντίστοιχα με εξομάλυνση του λόγου της πραγματικής παρατήρησης, X_t , προς το γινόμενο του επιπέδου, S_t , επί το δείκτη ημερήσιας εποχικότητας, D_{t-s_1} .

Αρχικοποίηση παραμέτρων και δεικτών:

- Η αρχική τιμή της τάσης, T_0 : υπολογίζεται ως ο μέσος όρος: 1) του $1/s_2$ επί της διαφοράς ανάμεσα στις πρώτες και τις δεύτερες s_2 (τον αριθμό) παρατηρήσεις της χρονοσειράς και 2) του μέσου όρου των πρώτων διαφορών για τις πρώτες s_2 παρατηρήσεις.

- Η αρχική τιμή του επιπέδου, S_0 : υπολογίζεται ως ο μέσος όρος των πρώτων $2 \cdot s_2$ παρατηρήσεων μείον $s_2/2$ φορές την αρχική τιμή της τάσης, T_0 .

- Δείκτης ημερήσιας εποχικότητας, D_{t_0} : ο μέσος όρος των λόγων εποχικότητας (τιμή δεδομένων αρχικής χρονοσειράς διαιρεμένη με τον ΚΚΜΟ βήματος s_1) από τις αντίστοιχες περιόδους της πρώτης εβδομάδας της χρονοσειράς

- Δείκτης εβδομαδιαίας εποχικότητας, W_{t_0} : ο μέσος όρος των λόγων εποχικότητας (τιμή δεδομένων αρχικής χρονοσειράς διαιρεμένη με τον ΚΚΜΟ βήματος s_2) από τις αντίστοιχες περιόδους των πρώτων δύο εβδομάδων της χρονοσειράς, διαιρεμένων με τους αντίστοιχους αρχικούς δείκτες ημερήσιας εποχικότητας.

Οι βέλτιστες τιμές των παραμέτρων a , γ , δ , ω και ϕ προσδιορίζονται με την κλασική διαδικασία της ελαχιστοποίησης του αθροίσματος των τετραγώνων των σφαλμάτων κάθε επόμενου βήματος. Στην παρούσα εργασία ο υπολογισμός των βέλτιστων παραμέτρων εξομάλυνσης γίνεται σε κάθε βήμα πρόβλεψης και όχι μία φορά, στην αρχή, για όλες τις εφαρμογές του μοντέλου, ώστε να επιτυγχάνονται καλύτερα αποτελέσματα.

Η εξέταση της μεθόδου εκθετικής εξομάλυνσης των Holt-Winters για διπλή εποχικότητα γίνεται μέσω του προγράμματος RStudio. Σύμφωνα λοιπόν με το πρόγραμμα αυτό, επιβάλλονται ορισμένες τροποποιήσεις στο μοντέλο του Taylor με την προσαρμογή για την αυτοσυσχέτιση του σφάλματος βάση των παρατηρήσεων του βιβλίου [Hyndman et al., 2008].

Όπως έχει παρουσιαστεί στη δημοσίευση «Short-term electricity demand forecasting using double seasonal exponential smoothing» του James T. Taylor, η βελτιωμένη μέθοδος Holt-

Winters αποδεικνύεται καλύτερη από το μοντέλο ARIMA για διπλή εποχικότητα με το οποίο συγκρίνεται. Αιτία για αυτό το αποτέλεσμα πιθανολογείται -από τους Chatfield και Yar- η έντονη ύπαρξη τάσης και διαφόρων εποχικοτήτων στη χρονοσειρά, βασικό χαρακτηριστικό της ζήτησης φορτίου ηλεκτρικής ενέργειας. Ενώ η μοντελοποίηση ARIMA είναι πιο ακριβής στην περίπτωση που η χρονοσειρά παρουσιάζει έντονη βραχυπρόθεσμη συσχέτιση.

4.6 Αξιολόγηση της ακρίβειας των προβλέψεων

Το εργαλείο για την αξιολόγηση της ακρίβειας των μεθόδων που χρησιμοποιούνται για την παραγωγή προβλέψεων είναι οι στατιστικοί δείκτες ακρίβειας προβλέψεων. Αρχικά, ως σφάλμα ορίζεται η διαφορά μεταξύ της πραγματικής και της προβλεπόμενης τιμής για μία περίοδο:

$$e_t = Y_t - F_t$$

Όταν γίνεται πρόβλεψη των τιμών ενός μεγέθους, μπορεί να γίνει υπολογισμός δύο ειδών σφαλμάτων. Πριν καταγραφούν οι πραγματικές τιμές του μεγέθους για την περίοδο στην οποία αναφέρεται η πρόβλεψη, μπορεί να υπολογιστεί μόνο το σφάλμα του μοντέλου πρόβλεψης (in-sample error). Ενώ έπειτα από ένα χρονικό διάστημα, όταν τα πραγματικά δεδομένα είναι πλέον διαθέσιμα, μπορεί να υπολογιστεί τόσο το σφάλμα του μοντέλου πρόβλεψης όσο και το πραγματικό σφάλμα (out-of-sample error). Το σφάλμα του μοντέλου πρόβλεψης χρησιμεύει στην επιλογή των παραμέτρων των μοντέλων πρόβλεψης, όπως για παράδειγμα αναφέρθηκε στις στατιστικές μεθόδους ότι οι βέλτιστες τιμές των παραμέτρων α και β επιλέγονται με διαδικασία ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος (MSE). Το δε πραγματικό σφάλμα παρουσιάζει την ακρίβεια της μεθόδου πρόβλεψης στην τελική της μορφή (εφόσον έχουν δηλαδή υπολογιστεί οι παράμετροι και έχουν γίνει οι τελικές προβλέψεις) και υποδεικνύει την ανάγκη ή μη τροποποίησης του μοντέλου για βελτίωση των σφαλμάτων.

Στη συνέχεια αναλύονται οι στατιστικοί δείκτες σφάλματος [Πετρόπουλος και Ασημακόπουλος, 2011]:

- Το μέσο σφάλμα (Mean Error) υπολογίζεται με χρήση της παρακάτω σχέσης και εκφράζει ένα μέτρο της συστηματικότητας του σφάλματος. Συγκεκριμένα, όσο μικρότερη τιμή λαμβάνει (κοντά στο μηδέν), τόσο περισσότερο χαρακτηρίζονται τα σφάλματα από τυχαιότητα. Και όσο απομακρύνεται η τιμή του σφάλματος από το μηδέν, τότε αναπτύσσεται μια τάση για απαισιοδοξία των προβλέψεων για θετικές τιμές του σφάλματος και μια τάση για αισιοδοξία για αρνητικές τιμές αντίστοιχα.

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

- Το μέσο απόλυτο σφάλμα (Mean Absolute Error) δείχνει την ακρίβεια της μεθόδου, χωρίς να δείχνει την κατεύθυνση της πρόβλεψης. Όσο μεγαλύτερη η τιμή του δείκτη αυτού, τόσο μικρότερη η ακρίβεια του μοντέλου πρόβλεψης, όπως φαίνεται και από τον τύπο υπολογισμού του. Δεν είναι πάντα κατάλληλο για χρονοσειρές διαφορετικού ύψους και ίδιας διασποράς, καθώς δεν είναι δυνατή η λήψη συγκεκριμένων πληροφοριών.

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

- Το μέσο τετραγωνικό σφάλμα (Mean Squared Error) είναι ένα μέτρο σύγκρισης της ακρίβειας πρόβλεψης των διαφόρων μοντέλων που χρησιμοποιούνται. Όπως αναφέρθηκε ήδη, χρησιμοποιείται για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης. Ιδιαίτερη βαρύτητα δίνεται στα μεγάλα σφάλματα, λόγω του τετραγωνισμού των σφαλμάτων, σε αντίθεση με τα μικρά σφάλματα που δεν επηρεάζουν τόσο έντονα το σφάλμα. Ο τύπος υπολογισμού του είναι ο εξής:

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

- Επειδή το μέσο τετραγωνικό σφάλμα εκφράζεται στις μονάδες τις αρχικής χρονοσειράς υψωμένες στο τετράγωνο, υπολογίζεται η τετραγωνική ρίζα αυτού, ώστε να είναι το σφάλμα εκφρασμένο στις μονάδες της αρχικής χρονοσειράς. Οι ιδιότητες της ρίζας του μέσου τετραγωνικού σφάλματος (Root Mean Squared Error) παραμένουν ως και του μέσου τετραγωνικού σφάλματος.

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

Ένα μειονέκτημα των προαναφερθέντων δεικτών ακρίβειας μιας μεθόδου πρόβλεψης είναι ότι εκφράζονται στις μονάδες της χρονοσειράς που μελετάται. Το χαρακτηριστικό τους αυτό δυσχεραίνει τη σύγκριση των διαφορετικών μεθόδων πρόβλεψης και ακόμη περισσότερο σε περιπτώσεις πολύ υψηλών πραγματικών τιμών. Χρησιμοποιούνται τότε σφάλματα σε ποσοστιαία μορφή, όπως παρουσιάζονται αμέσως μετά:

- Το μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error) λαμβάνει μη αρνητικές τιμές όπως φαίνεται και από τον παρακάτω τύπο. Είναι από τα δύο κύρια χρησιμοποιούμενα σφάλματα για την εκτίμηση της ακρίβειας των μοντέλων πρόβλεψης. Σε περιπτώσεις μεγάλων θετικών τιμών των δεδομένων της χρονοσειράς προτιμάται, λόγω της απλότητάς του. Ακόμη, στην αξιολόγηση της ακρίβειας μιας μεθόδου πρόβλεψης εφαρμοζόμενης σε περισσότερες από μία χρονοσειρές και μάλιστα διαφορετικού επιπέδου τιμών, χρησιμοποιείται το σφάλμα αυτό. Παρουσιάζει ωστόσο αδυναμία εφαρμογής σε χρονοσειρές διακοπτόμενης ζήτησης, αφού καταλήγει σε απροσδιοριστία.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - F_i|}{|Y_i|} \cdot 100\%$$

-Το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error) χρησιμοποιείται μαζί με το μέσο απόλυτο ποσοστιαίο σφάλμα κατά κύριο λόγο στην ανάλυση της ακρίβειας προβλέψεων. Σε αντίθεση όμως με το MAPE, το SMAPE δεν παρουσιάζει απροσδιοριστία κατά την εφαρμογή του σε χρονοσειρές διακοπτόμενης ζήτησης. Αποφεύγει τα μεγάλα σφάλματα σε περίπτωση έντονων διαφορών ανάμεσα στην απόλυτη ποσοστιαία τιμή σφάλματος όταν η πραγματική τιμή είναι μεγαλύτερη της πρόβλεψης και το ανάποδο. Το μειονέκτημα αυτού του δείκτη είναι ότι δεν

αντιμετωπίζονται οι αισιόδοξες προβλέψεις με αντίστοιχο τρόπο όπως οι απαισιόδοξες και επομένως δεν είναι τόσο συμμετρικός όσο παρουσιάζεται.

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

- Η διάμεσος του απόλυτου ποσοστιαίου σφάλματος (Median Absolute Percentage Error) είναι ένα ακόμη από τα ποσοστιαία σφάλματα που χρησιμοποιούνται. Προτιμάται έναντι του MAPE, γιατί είναι πιο σταθερό σφάλμα απέναντι σε ανωμαλίες των τιμών της χρονοσειράς. Ένα μειονέκτημα των σφαλμάτων MAPE και MdAPE είναι ότι δεν ορίζονται για την περίπτωση μηδενικής τιμής για οποιαδήποτε χρονική περίοδο που μελετάται για την περίπτωση τιμών κοντά στο μηδέν). Στην παρούσα εργασία ωστόσο δεν υφίστανται μηδενικές τιμές δεδομένων. Τέλος, αναφέρουμε ότι κατά την ύπαρξη πραγματικών τιμών κοντά στο μηδέν οι τιμές του MAPE είναι σημαντικά μεγαλύτερες από τις αντίστοιχες τιμές του σφάλματος MdAPE, λόγω της ιδιαίτερα ασύμμετρης κατανομής των τιμών του.

$$MdAPE = \text{median}(APE) = \text{median} \left(\sum_{i=1}^n \frac{|Y_i - F_i|}{|Y_i|} \cdot 100\% \right)$$

- Η διάμεσος του συμμετρικού απόλυτου ποσοστιαίου σφάλματος (Symmetric Median Absolute Percentage Error) χρησιμοποιείται για την ελάττωση των προβλημάτων που προκύπτουν από πολύ μικρές τιμές των δεδομένων της χρονοσειράς. Το συμμετρικό MdAPE όπως και το συμμετρικό MAPE δεν είναι τόσο συμμετρικό όσο υποδεικνύει η ονομασία του. Και αυτό δίνει μεγαλύτερη βαρύτητα στα θετικά σφάλματα σε αντίθεση με τα αρνητικά, σε μικρότερο βαθμό όμως από το MdAPE.

$$sMdAPE = \text{median} \left(\sum_{i=1}^n \frac{|Y_i - F_i|}{\left| \frac{Y_i + F_i}{2} \right|} \cdot 100\% \right)$$

- Υπάρχει ακόμη και η κατηγορία των σχετικών σφαλμάτων, που υπολογίζονται συγκριτικά με τη μέθοδο που επιλέγεται ως βασική. Αυτή η διαδικασία διευκολύνει στην καλύτερη αντίληψη της πορείας των προβλέψεων σε σύγκριση με τις γύρω τους. Μάλιστα, εφαρμόστηκε και το μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute Scaled Error), το οποίο έχει σαν βάση την απλοϊκή μέθοδο Naive και συγκρίνει την πορεία των υπόλοιπων μελετώμενων μεθόδων με αυτή. Έτσι, μπορεί να κατατάξει κανείς τις μεθόδους που χρησιμοποιούνται σύμφωνα με τα αποτελέσματα του σφάλματος αυτού.

4.7 Μοντέλα βραχυπρόθεσμης πρόβλεψης

Λόγω του μεγάλου ενδιαφέροντος που παρουσιάζει η πρόβλεψη της ζήτησης φορτίου ηλεκτρικής ενέργειας και με την αναγνώριση της έντονης εμφάνισης διπλής εποχικότητας, ημερήσια και εβδομαδιαία, στα δεδομένα των χρονοσειρών της ζήτησης έχουν αναπτυχθεί τα τελευταία χρόνια αρκετές μέθοδοι και συνδυασμοί αυτών, με συνυπολογισμό της επίδρασης της εποχικότητας. Στη συνέχεια παρατίθενται ορισμένες από τις πιο αντιπροσωπευτικές μεθόδους, οι οποίες στοχεύουν στη βραχυπρόθεσμη πρόβλεψη ηλεκτρικού φορτίου, που είναι και το αντικείμενο μελέτης της παρούσας εργασίας.

Χαρακτηριστικό είναι το εύρος χρήσης των μεθόδων εκθετικής εξομάλυνσης, καθώς παρατηρήθηκε ευρωστία των μοντέλων τους, καλύτερη διαίσθηση και ακρίβεια, σε συνδυασμό με την απλότητά τους και την πολύ μικρή απαίτηση σε δεδομένα. Έτσι, η εκθετική εξομάλυνση για γραμμική τάση (Holt Exponential Smoothing) και το μοντέλο μη γραμμικής τάσης με βαθμιαία μείωση του επιπέδου των τιμών (Damped trend) δέχονται μετατροπές για την ενσωμάτωση της εποχικότητας των χρονοσειρών ζήτησης. Αξίζει να αναφερθεί ότι για μεγάλο διάστημα Το 2003 ο Taylor [Taylor, 2003a] μελέτησε μια μέθοδο εκθετικής εξομάλυνσης, με πολλαπλασιαστική τάση, την οποία πρότεινε ο Pegels [Pegels, 1969] αναφερόμενος στο ότι οι περισσότερες πραγματικές χρονοσειρές παρουσιάζουν πολλαπλασιαστική και όχι προσθετική τάση. Εισηγήθηκε ωστόσο στη μέθοδο του Pegels μία παράμετρο για μη γραμμική εξομάλυνση. Η μέθοδος αυτή αποδείχτηκε καλύτερη από τη μέθοδο damped Holt, που ήταν ως τότε δημοφιλής. Στη συνέχεια, προσάρμοσε τη μέθοδο Holt-Winters μίας εποχικότητας για την ενσωμάτωση και των δύο εποχικότητων της χρονοσειράς ζήτησης φορτίου, τόσο της ημερήσιας όσο και της εβδομαδιαίας [Taylor, 2003b]. Η νέα αυτή μέθοδος αναδείχθηκε ιδιαίτερα δημοφιλής στον τομέα της πρόβλεψης ζήτησης ηλεκτρικού φορτίου, καθώς ανήκει στις απλές μεθόδους μιας μεταβλητής και συνυπολογίζει τις δύο βασικές εποχικότητες των χρονοσειρών της κατηγορίας.

Η ποικιλία μεθόδων μιας μεταβλητής για τη βραχυπρόθεσμη πρόβλεψη φορτίου είναι πολύ μεγάλη. Εκτός από την εκθετική εξομάλυνση στην οποία ανήκει και η μέθοδος του Taylor, με ενσωμάτωση δύο εποχικότητων, προσοχή δόθηκε και στη μοντελοποίηση ARIMA για διπλή εποχικότητα, στα τεχνητά νευρωνικά δίκτυα και σε μεθόδους παλινδρόμησης με διάφορους συνδυασμούς αυτών.

- ❖ Για τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου, ARIMA, γνωστή είναι η εφαρμογή των Box και Jenkins [Box and Jenkins, 1970], όπως και τα ARIMA transfer function models [Cho et al., 1995]. Επειδή παρουσιάζουν πολύ καλή συμπεριφορά συγκριτικά με τις υπόλοιπες κλασσικές μεθόδους και έχουν κατανοητή μορφή, χρησιμοποιούνται ως σημείο αναφοράς για την αξιολόγηση εναλλακτικών μεθόδων στην κατηγορία των μοντέλων μιας μεταβλητής και για σύγκριση με νέα μοντέλα νευρωνικών δικτύων.
- ❖ Τα τεχνητά νευρωνικά δίκτυα επιτρέπουν τον πειραματισμό με μοντέλα χαμηλού θεωρητικού υποβάθρου, αλλά πλούσια σε δεδομένα, τα οποία μπορούν να αναγνωρίσουν τις πολύπλοκες μη γραμμικές σχέσεις των δεδομένων της χρονοσειράς και να συμπεράνουν τη μελλοντική συμπεριφορά τους. Τα τεχνητά νευρωνικά δίκτυα παράγουν καλύτερα αποτελέσματα με το πέρασμα του χρόνου, καθώς «μαθαίνουν» μέσω παραδείγματος. Έτσι έχουν τη δυνατότητα να αναπτύξουν καλύτερη συμπεριφορά (και άρα ακρίβεια προβλέψεων) από τα μοντέλα που στηρίζονται στη γραμμική παλινδρόμηση και τη μέθοδο ελαχίστων τετραγώνων. Ακόμη, εμφανίζουν μεγάλη ευαισθησία σε μεταβολές του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ) και χρειάζεται η μετατροπή τους με την αλλαγή συστήματος (πχ. ανάλογα τη χώρα στην οποία βρίσκεται το ΣΗΕ) [Charytoniuk et al., 1998]. Έχουν αναπτυχθεί μέθοδοι δύο μόλις νευρωνικών δικτύων, που με διάδοση προς τα πίσω παράγουν γραμμικό συνδυασμό και σε αυτόν εφαρμόζεται η μέθοδος των ελαχίστων τετραγώνων με μεγαλύτερη έμφαση στα πιο σφάλματα των πιο πρόσφατων παρατηρήσεων [Khotanzad et al., 1998].
- ❖ Σχετικά με τις μεθόδους παλινδρόμησης, υπάρχει πλήθος προσαρμογών τους, όπως με εφαρμογή διάφορων μετασχηματισμών επί αυτών. Για παράδειγμα παλινδρομική μέθοδος με εισόδους τη μέγιστη θερμοκρασία της ημέρας, τη μέση θερμοκρασία

των δύο τελευταίων ημερών και τη σχετική υγρασία της ημέρας σε μορφή πολυωνυμικών συναρτήσεων [Haidda, 1994]] ή ένα συνεχές αυτοπαλινδρομικό μοντέλο με μετασχηματισμό τύπου Hilbert [Khotanzad, 1995] για χρονοσειρές στοχαστικές και χωρίς στασιμότητα. Υπάρχουν βέβαια και μη παραμετρικές μέθοδοι, ντετερμινιστικές προσεγγίσεις [Dehdashti et al., 1982] και πλήθος υβριδικών κλασικότερων μεθόδων [Papalexopoulos and Hesterrberg, 1990]. Ακόμη η παλινδρόμηση με principal component analysis, PCA, για μείωση της διάστασης των συνόλων πολυμεταβλητών δεδομένων (όπου οι μεταβλητές παρουσιάζουν μεγάλη συσχέτιση) σε ένα μικρότερο σύνολο μεταβλητών) [Guo et al., 2004].

- ❖ Τελευταία αναλύθηκε και η εκθετική εξομάλυνση με ενδοημερήσιο κύκλο για τις διαφορετικές ημέρες της εβδομάδας. Θεωρεί δηλαδή ολόκληρες ημέρες ως ταυτόσημες και τις αντιμετωπίζει αντίστοιχα (τις διαφορετικές ημέρες της εβδομάδας που επαναλαμβάνονται στη χρονοσειρά). Έπειτα από αναφορές για πιθανές περαιτέρω προσαρμογές [Hyndman et al., 2008], ο Taylor [Taylor, 2010] πρόβη στη βελτίωση της μεθόδου αυτής, θεωρώντας τμήματα των ξεχωριστών ημερών ως ταυτόσημα και όχι ολόκληρες τις ημέρες.

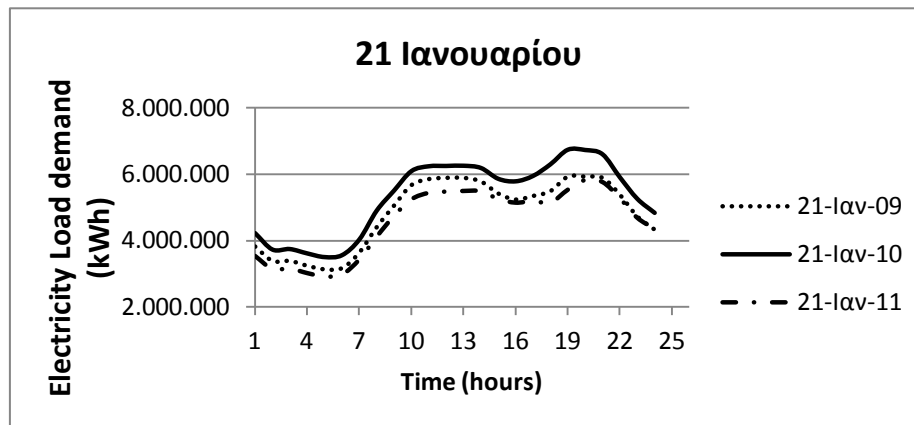
Αυτές είναι μερικές από τις μεθόδους πρόβλεψης που έχουν χρησιμοποιηθεί κατά κόρον τα τελευταία έτη και συνεχίζουν να εξελίσσονται για επίτευξη όσο το δυνατόν μεγαλύτερης ακρίβειας. Στην εργασία αυτή, όπως αναφέρθηκε και παραπάνω, θα μελετηθούν οι κλασικές στατιστικές μέθοδοι: η απλοϊκή Naive, οι μέθοδοι εκθετικής εξομάλυνσης SES, Holt και Damped , η κλασική μέθοδος Theta και η προσαρμογή του Taylor στη μέθοδο εκθετικής εξομάλυνσης Holt-Winters μιας εποχικότητας, για να ενσωματώνονται πλέον και οι δύο εποχικότητες της σειράς ζήτησης ηλεκτρικού φορτίου. Ο χειρισμός των απλών μεθόδων, όσο και της προσαρμογής του Taylor παρουσιάζεται στο κεφάλαιο 5.

ΚΕΦΑΛΑΙΟ 5: Μεθοδολογία

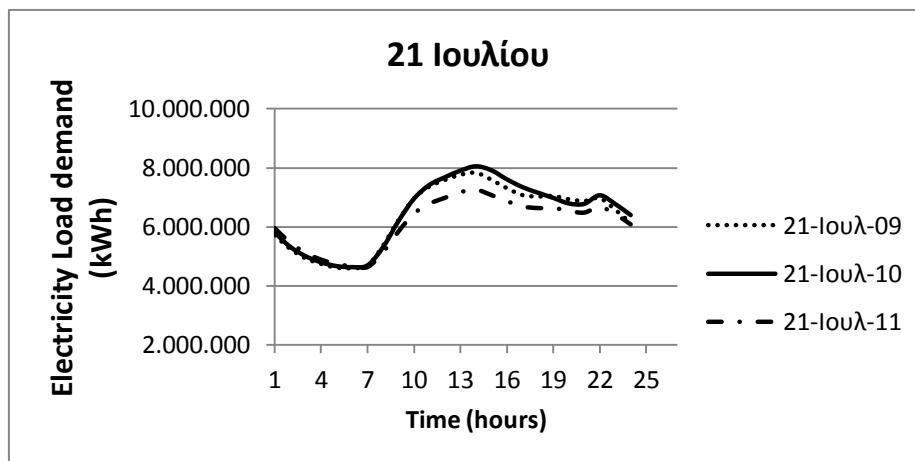
5.1 Δεδομένα

Η διαδικασία πρόβλεψης έχει ως πρώτο στάδιο τον καθορισμό του προβλήματος, δηλαδή την αναγνώριση της ανάγκης που υπάρχει και ποιος είναι ο αποδέκτης των αποτελεσμάτων της πρόβλεψης. Στη συγκεκριμένη εργασία στόχος είναι, για τις ανάγκες μιας ιδιωτικής εταιρείας παραγωγής ηλεκτρικής ενέργειας, η μελέτη μοντέλων πρόβλεψης των 24 ωριαίων τιμών ζήτησης για κάθε επόμενη ημέρα και η ανάδειξη των καλύτερων από αυτά. Το δεύτερο σημαντικό κομμάτι της διαδικασίας πρόβλεψης είναι η συλλογή και η ανάλυση των δεδομένων. Αυτό το στάδιο βοηθάει το χειριστή ώστε να λάβει καταρχάς υλικό για τη μελέτη του και ακόμη για να αποκτήσει μια σωστή και αντιπροσωπευτική εικόνα των στοιχείων με τα οποία ασχολείται. Έτσι θα μπορεί να έχει μια πιο καλή αντίληψη της κατάστασης των δεδομένων, αλλά θα μπορεί και να αναπτύξει πιο σωστή κριτική σκέψη για την εκάστοτε περίπτωση. Στην παρούσα εργασία χρησιμοποιούνται ωριαίες παρατηρήσεις ζήτησης φορτίου από μια ελληνική εταιρεία παραγωγής ηλεκτρισμού. Πρόκειται για δεδομένα που αντιστοιχούν στη Χαμηλή Τάση (εκφρασμένα σε kWh) για το χρονικό διάστημα τριάντα δύο μηνών, από την 1 Ιανουαρίου του 2009 ως και την 31 Αυγούστου 2011. Ακριβέστερα για τη χρονοσειρά που εξετάζουμε, το πλήθος των δεδομένων για το χρονικό διάστημα από την 1/1/2009 έως και την 31/8/2011 ισούται με 23352 ωριαίες παρατηρήσεις. Εκφρασμένο σε ημέρες το ίδιο πλήθος στοιχείων αντιστοιχεί σε 973 ημέρες.

Στις εικόνες 1 και 2 παρουσιάζονται τα δεδομένα ζήτησης φορτίου για τις ημέρες 21 Ιανουαρίου και 21 Ιουλίου αντίστοιχα και για τα τρία έτη στα οποία διαθέτουμε παρατηρήσεις από την εταιρεία, ώστε να έχουμε μια καλύτερη εικόνα του επιπέδου των τιμών ζήτησης σε δύο χαρακτηριστικούς μήνες διαφορετικών εποχών, χειμώνας-καλοκαίρι. Είναι φανερή η έντονη αύξηση του επιπέδου των τιμών για το έτος 2010 και στη συνέχεια η ακόμη μεγαλύτερη μείωση του επιπέδου για το έτος 2011.

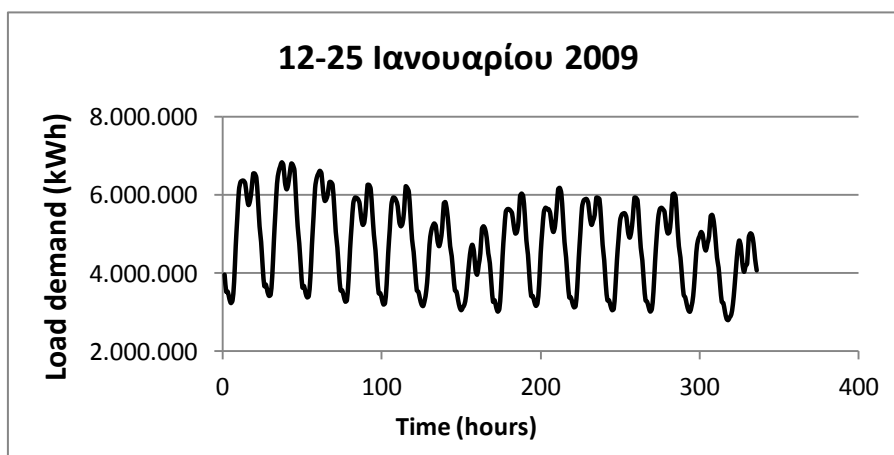


Εικόνα 1. Ημερήσια ζήτηση φορτίου για την 21 Ιανουαρίου (σε kWh)

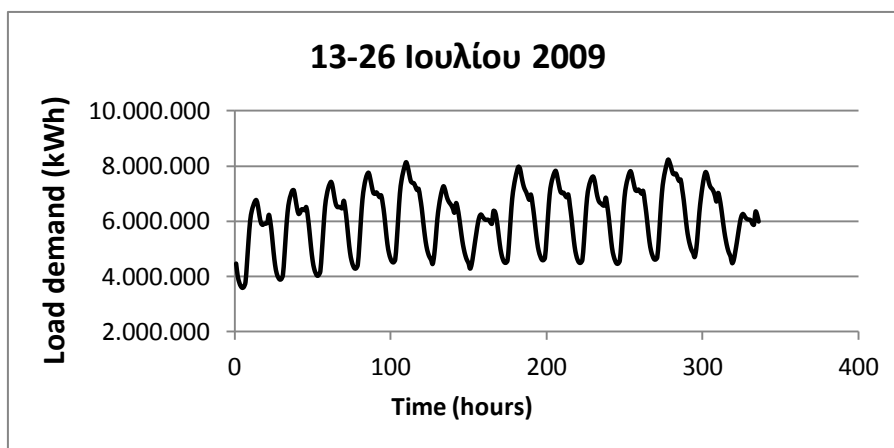


Εικόνα 2. Ημερήσια ζήτηση φορτίου για την 21 Ιουλίου (σε kWh)

Στις εικόνες 3 και 4 παρουσιάζεται ενδεικτικά η ζήτηση φορτίου ηλεκτρισμού για δύο εβδομάδες τους μήνες Ιανουάριο και Ιούλιο και συγκεκριμένα για τις εβδομάδες 12/01/2009 ως και 25/01/2009 στην εικόνα 3 και για τις εβδομάδες 13/07/2009 έως και 26/07/2009 στην εικόνα 4 (από Δευτέρα σε επόμενη Κυριακή). Στις εικόνες αυτές είναι πλέον εμφανής η εποχικότητα της χρονοσειράς σε επίπεδο ημέρας και εβδομάδας. Παρατίθενται δεδομένα για τις δύο αυτές εβδομάδες, καθώς αποτελούν δύο αντιπροσωπευτικές εβδομάδες του έτους, η πρώτη στο μεσαίο μήνα του χειμώνα και η δεύτερη κατά το μεσαίο μήνα του καλοκαιριού. Φαίνεται χαρακτηριστικά, εκτός από τη διπλή εποχικότητα και το επίπεδο των τιμών ζήτησης στις διαφορετικές αυτές χρονικές περιόδους. Αξιοπρόσεκτη είναι η πτωτική τάση των τιμών στην εικόνα 3 για τις καθημερινές, σε αντίθεση με την αύξηση της ζήτησης όσο πλησιάζει η Παρασκευή στην εικόνα 4.



Εικόνα 3. Εβδομαδιαία ζήτηση φορτίου για δύο εβδομάδες του Ιανουαρίου 2009 (σε kWh)



Εικόνα 4. Εβδομαδιαία ζήτηση φορτίου για δύο εβδομάδες του Ιουλίου 2009 (σε kWh)

Από τα δεδομένα που υπάρχουν στη διάθεσή μας, θα χρησιμοποιηθούν ορισμένα για τον υπολογισμό του σφάλματος του μοντέλου πρόβλεψης (in-sample error) και τα υπόλοιπα για την αξιολόγηση του εκάστοτε εφαρμοζόμενου μοντέλου από αυτά που αναφέρθηκαν στο τέλος της παραγράφου 4.6, με τον υπολογισμό του πραγματικού σφάλματος (out-of-sample error). Κρίθηκε σκόπιμο η αξιολόγηση των μεθόδων που θα εξεταστούν να γίνει με χρήση των δεδομένων του έτους 2011, δηλαδή με τις 243 εκ των 973 ημερών από την 1/1/2011 έως και την 31/1/2011. Ποσοστιαία, ο αριθμός των παρατηρήσεων για το out-of-sample error είναι περίπου στο 75% του συνολικού πλήθους τους. Εφαρμόζεται ωστόσο στην εργασία αυτή η μέθοδος της κύλισης των δεδομένων κατά τη διαδικασία της πρόβλεψης. Αυτό έχει σαν αποτέλεσμα να μη χρησιμοποιούνται όλα τα παρελθοντικά δεδομένα, δηλαδή το 75% που αναφέρουμε, αλλά επιλέγεται ένα μόλις τμήμα των παρατηρήσεων της πραγματικής χρονοσειράς, το οποίο ορίζεται με βάση τη χρονική περίοδο κύλισης δεδομένων που αποφασίζουμε να εξετάσουμε. Τη μέθοδο αυτή θα τη σχολιάσουμε καλύτερα στην παράγραφο 5.2.

	Χρονικό διάστημα	Πλήθος σε ημέρες	Πλήθος ωριαίων δεδομένων	Ποσοστό επί του συνόλου
Δεδομένα υπολογισμού in-sample error	1-1-2009 έως 31-12-2010	730 ημέρες	17520 στοιχεία	75%
Δεδομένα υπολογισμού out-of-sample error	1-1-2011 έως 31-8-2011	243 ημέρες	5832 στοιχεία	25%

Πίνακας 1. Πλήθος διαθέσιμων δεδομένων.

5.2 Διαδικασία παραγωγής προβλέψεων

Να σημειωθεί αρχικά ότι η αποθήκευση των δεδομένων ως χρονοσειρά, η αρχική επεξεργασία τους, αλλά και η επιβολή των απλών μεθόδων πρόβλεψης επί αυτών γίνονται

με τη χρήση υπολογιστικού φύλλου και των ενσωματωμένων σε αυτό μακροεντολών (Microsoft Excel και Visual Basic on Excel).

Στο πρώτο σκέλος της μελέτης αυτής γίνεται εφαρμογή της απλοϊκής μεθόδου Naive, των μεθόδων εκθετικής εξομάλυνσης SES, Damped και Holt, του κλασσικού μοντέλου της μεθόδου Theta και της προσαρμογής του Taylor επί της βασικής μεθόδου των Holt-Winters (Holt-Winters method with double seasonality). Όπως αναφέραμε η μοντελοποίηση των μεθόδων αυτών γίνεται με χρήση των μακροεντολών του Excel σε Visual Basic. Ενώ τα αποτελέσματα της μεθόδου του Taylor ελέγχονται με επανεκτέλεση των προβλέψεων στο πρόγραμμα του RStudio. Στη συνέχεια περιγράφονται τα βήματα που ακολουθούνται για την παραγωγή των προβλέψεων.

Διαδικασία παραγωγής προβλέψεων:

- Εφαρμογή κλασσικής αποσύνθεσης για αποεποχικοποίηση των δεδομένων για κάθε τρέχουσα ημέρα που μελετάμε. Σαφώς, η εφαρμογή της γίνεται μόνο στις περιπτώσεις που κρίνεται αναγκαίο, συγκεκριμένα για τις μεθόδους που αναφέρθηκαν εκτός από την προσαρμογή του Taylor αφού, όπως σημειώθηκε ήδη στην παράγραφο 4.5, το μοντέλο από τον ορισμό του συμπεριλαμβάνει και τις δύο διαφορετικές εποχικότητες.
- Εφαρμογή των μοντέλων που μας ενδιαφέρουν, το καθένα ξεχωριστά και παραγωγή της προέκτασης των χρονοσειρών.
- Εποχικοποίηση ξανά των παρατηρήσεων που παράχθηκαν, ώστε να είναι πιο ρεαλιστικές και ακριβείς οι προβλέψεις.
- Καταχώρηση των προβλέψεων για τη χρονική περίοδο 1/1/2011 έως και 31/8/2011 και υπολογισμός των σφαλμάτων αυτών, με έμφαση στα MAPE, sMAPE, MdAPE και sMdAPE.

Η διαδικασία της αποεποχικοποίησης σε γενικό πλαίσιο έχει αναλυθεί στην παράγραφο 4.4. Τώρα θα παρατηρήσουμε τις μετατροπές που γίνονται στα πλαίσια της εργασίας αυτής επί της διαδικασίας της αποεποχικοποίησης των δεδομένων. Στην περίπτωση μας, τα διαθέσιμα δεδομένα δίνονται σε ωριαία βάση με αποτέλεσμα οι περίοδοι των εποχικότητων να ισούνται με:

- $s_1 = 24$ (ημερήσια: 24 ώρες/ημέρα) και
- $s_2 = 168$ (εβδομαδιαία: 24ώρες/ημέρα x 7 ημέρες).

Για κάθε μία από τις μεθόδους στις οποίες χρειάζεται να εφαρμοστεί αποεποχικοποίηση επί των δεδομένων της, ακολουθούν με τη σειρά δύο στάδια. Στο πρώτο, γίνεται μελέτη της ημερήσιας εποχικότητας και στο δεύτερο της εβδομαδιαίας εποχικότητας της χρονοσειράς. Η αποεποχικοποίηση του πρώτου και του δεύτερου σταδίου γίνεται με χρήση του κεντρικού κινητού μέσου όρου ΚΚΜΟ(24) και ΚΚΜΟ(168) αντίστοιχα. Από αυτούς τους ΚΚΜΟ και με την ήδη γνωστή διαδικασία (από την παράγραφο 4.4) υπολογίζονται στα δύο ξεχωριστά στάδια οι κανονικοποιημένοι δείκτες ημερήσιας και εβδομαδιαίας εποχικότητας και εφαρμόζονται επί των δεδομένων της αρχικής χρονοσειράς, για τη λήψη της αποεποχικοποιημένης χρονοσειράς. Αυτή η μεθοδολογία ακολουθείται επειδή, όπως γνωρίζουμε, τα συνήθη χρησιμοποιούμενα μοντέλα μπορούν να συμπεριλάβουν μόλις μέχρι μία εποχικότητα, ως μοντέλα μιας ανεξάρτητης μεταβλητής.

Στη συνέχεια, σε κάθε στάδιο επιβάλλεται η υπό μελέτη μέθοδος πρόβλεψης και επί των νέων προβλέψεων εφαρμόζονται ξανά οι δείκτες εποχικότητας (με πολλαπλασιασμό των τιμών των παραγόμενων προβλέψεων με τους κανονικοποιημένους δείκτες), ώστε να προκύψει η τελική πρόβλεψη του επόμενου εικοσιτετραώρου (για την ακρίβεια εικοσιτέσσερις ωριαίες τιμές). Με τον τρόπο αυτό, στο πρώτο σκέλος της μελέτης μας λαμβάνουμε τις προβλέψεις με ενσωματωμένη την επίδραση της ημερήσιας εποχικότητας για τις τιμές ζήτησης φορτίου ηλεκτρισμού όλων των απλών μεθόδων που εξετάζουμε για το χρονικό διάστημα 1/1/2011 έως και 31/8/2011 και αντίστοιχα στο δεύτερο σκέλος λαμβάνουμε τις προβλέψεις με ενσωματωμένη την επίδραση της εβδομαδιαίας εποχικότητας. Η αποεποχικοποίηση ωστόσο στην εργασία αυτή δεν εφαρμόζεται μόνο μία φορά στην αρχή της διαδικασίας για κάθε μέθοδο πρόβλεψης, αλλά σε κάθε τρέχουσα ημέρα προς μελέτη. Αυτό είναι δυνατό με τη χρήση της μεθόδου κύλισης, η οποία αναλύεται στην παράγραφο 5.3 και στόχο έχει την καλύτερη προετοιμασία της χρονοσειράς πριν από κάθε πρόβλεψη.

Στη μέθοδο του Taylor δεν εφαρμόζεται αποεποχικοποίηση, αλλά παράγονται οι προβλέψεις της με τη βοήθεια του μοντέλου της που ενσωματώνει την επίδραση και των δύο εποχικότητων της χρονοσειράς. Κάτι ακόμη που είναι σημαντικό στη διαδικασία πρόβλεψης που ακολουθείται, είναι ότι η παραγωγή πρόβλεψης έπεται από την αποεποχικοποίηση δε γίνεται μόνο μία φορά για κάθε μέθοδο. Αντίθετα, σε κάθε ένα από τα δύο στάδια εφαρμόζονται οι μέθοδοι πρόβλεψης για αρκετές περιπτώσεις, όπως αναλύονται στην επόμενη παράγραφο (παράγραφος 5.3). Το ίδιο συμβαίνει και για τη μέθοδο του Taylor, η οποία δεν υφίσταται τη διαδικασία αποεποχικοποίησης και το πλήθος των εκτελέσεων αυτής της μεθόδου θα σημειωθεί στην επόμενη παράγραφο μαζί με των υπολοίπων απλούστερων μεθόδων μιας εποχικότητας.

5.3 Εφαρμογή κυλιόμενης πρόβλεψης (rolling forecasting)

Ένα σημαντικό χαρακτηριστικό που εισάγουμε σε όλα τα μοντέλα πρόβλεψης που εξετάζονται στην εν προκειμένω εργασία είναι η κύλιση των δεδομένων ή αλλιώς rolling forecasting.

Όπως αντιλαμβάνεται κανείς, οι διαθέσιμες παρατηρήσεις τις οποίες θεωρούμε ως φανερές αντιστοιχούν σε μια περίοδο δύο ετών (2009 και 2010), αρκετά μεγάλο χρονικό διάστημα με εναλλαγές εποχών και ποικιλία διακυμάνσεων. Σκεφτόμαστε ότι για τον υπολογισμό των τιμών ζήτησης για το έτος 2011 ενδέχεται να μη χρειάζονται οι πληροφορίες που διαθέτουν όλα αυτά τα δεδομένα, αντιθέτως ίσως και να δυσχεραίνουν το έργο των τεχνικών πρόβλεψης. Επειδή η εργασία αυτή έχει ενδιαφέρον για τις εταιρείες παραγωγής ενέργειας, δίνεται προσοχή στην όσο το δυνατόν καλύτερη εξυπηρέτηση των συμφερόντων τους. Επομένως, στα πλαίσια των πολύ ανταγωνιστικών αγορών ηλεκτρικής ενέργειας και στο συνεχώς μεταβαλλόμενο οικονομικό περιβάλλον, μας ενδιαφέρει οι εταιρείες παραγωγής να μπορούν να ανταποκρίνονται γρήγορα και αποτελεσματικά (για τις ίδιες) στις επικρατούσες συνθήκες. Στόχος λοιπόν είναι η μεγαλύτερη ακρίβεια των προβλέψεων, ώστε οι εταιρείες παραγωγής ηλεκτρικής ενέργειας να λαμβάνουν καλύτερες αποφάσεις σχετικά με τη μελλοντική τους παραγωγή, τις μετατροπές των συστημάτων ενέργειας (ΣΗΕ) τους και τη συμπεριφορά τους στην αγορά. Για την επίτευξη αυτού του στόχου εφαρμόζεται η μέθοδος της κυλιόμενης πρόβλεψης σε όλα τα μοντέλα που χρησιμοποιούνται στην εργασία. Έτσι, μόνο η πιο πρόσφατες παρατηρήσεις της χρονοσειράς λαμβάνονται υπόψη κατά την αποεποχικοποίηση των τρεχόντων δεδομένων. Εκτός όμως από την αποεποχικοποίηση, και η επιλογή των βέλτιστων συντελεστών κάθε μοντέλου πρόβλεψης,

καθώς και οι ίδιες οι παραγόμενες προβλέψεις προκύπτουν με βάση τη μέθοδο της κυλιόμενης πρόβλεψης. Παρακάτω εξηγείται καλύτερα η διαδικασία που ακολουθείται.

Με τη μέθοδο της στατικής πρόβλεψης (static forecasting) χρησιμοποιούνται τα δεδομένα των περιόδων που έχουν επιλεγεί μέχρι να μηδενιστούν (μειώνεται δηλαδή η χρησιμοποιούμενη χρονική περίοδος) και στη συνέχεια επιλέγονται νέα δεδομένα ίδιας χρονικής περιόδου. Η κυλιόμενη πρόβλεψη από την άλλη πλευρά αξιοποιεί παρατηρήσεις ίδιας χρονικής περιόδου, η οποία παραμένει σταθερή. Ορίζεται αρχικά το χρονικό διάστημα κύλισης των πραγματικών δεδομένων. Δηλαδή, για κάθε τρέχον εικοσιτετράωρο, η διαδικασία πρόβλεψης εφαρμόζεται θεωρώντας ως παρελθοντικά στοιχεία όσα προηγούνται κατά αυτό το χρονικό διάστημα. Η κύλιση γίνεται επί των δεδομένων έτσι ώστε να υπεισέρχονται οι 24 νέες πραγματικές ωριαίες τιμές αμέσως μόλις γίνουν διαθέσιμες και να αφαιρούνται οι 24 παλαιότερες πραγματικές τιμές της χρονοσειράς, εξασφαλίζοντας ότι το χρονικό διάστημα των δεδομένων που χρησιμοποιούνται για την κάθε ημερήσια πρόβλεψη παραμένει σταθερό σε όλη τη διαδικασία (κατά την εκτέλεση δηλαδή της μεθόδου για πρόβλεψη όλων των τιμών ζήτησης του 2011). Στην περίπτωση μας ζητείται η πρόβλεψη ζήτησης φορτίου σε εικοσιτετράωρη βάση. Έτσι λοιπόν, με το πέρας κάθε εικοσιτετράωρου, γίνονται διαθέσιμα τα πραγματικά δεδομένα για τα οποία έχει παραχθεί η πιο πρόσφατη πρόβλεψη και πλέον θεωρούνται καταγεγραμμένες παρατηρήσεις για τον υπολογισμό του in-sample error και κατ' επέκταση των συντελεστών εξομάλυνσης της εφαρμοζόμενης μεθόδου.

Το ακόλουθο παράδειγμα βοηθάει στην καλύτερη κατανόηση του τρόπου λειτουργίας της μεθόδου κύλισης. Έστω ότι στόχος μας είναι η πρόβλεψη ζήτησης ενός μεγέθους για χρονικό ορίζοντα δώδεκα μηνών (ενός έτους). Στην περίπτωση αυτή ο αριθμός των περιόδων εφαρμογής της μεθόδου ισούται σταθερά με 12 μήνες, με το πέρασμα κάθε μήνα η πραγματική του ζήτηση είναι πλέον διαθέσιμη και γίνεται ενσωμάτωσή της στα πραγματικά δεδομένα. Έτσι, παύει να θεωρείται τμήμα των προβλεπόμενων δεδομένων και εισέρχεται νέα περίοδος δεδομένων στο τέλος των προβλέψεων (κύλιση των προβλέψεων κατά ένα μήνα), με αποτέλεσμα να διατηρείται ο ορίζοντας της πρόβλεψης σταθερός (ίσος με 12 μήνες στο παράδειγμά μας). Η διαδικασία αυτή παρουσιάζεται και στην εικόνα 5.

Original 12 periods forecast												Rolling forecast (always 12 periods)		
July	August	September	October	November	December	January	February	March	April	May	June	July	August	September
1	2	3	4	5	6	7	8	9	10	11	12			
	1	2	3	4	5	6	7	8	9	10	11	12		
		1	2	3	4	5	6	7	8	9	10	11	12	
			1	2	3	4	5	6	7	8	9	10	11	12

Εικόνα 5. Παράδειγμα μεθόδου rolling forecasting

Σύμφωνα με τα παραπάνω, ολοκληρωμένη η διαδικασία πρόβλεψης στο πρώτο μέρος της παρούσας μελέτης μπορεί να περιγραφεί από τα παρακάτω βήματα (για μία ημέρα, δηλαδή εικοσιτέσσερις (24) προβλέψεις):

- ✓ Επιλέγονται και αντιγράφονται σε ένα υπολογιστικό φύλο του Microsoft Excel οι παρατηρήσεις της πραγματικής χρονοσειράς για διάστημα ίσο με το επιλεγμένο χρονικό διάστημα εφαρμογής κύλισης δεδομένων πριν από την τρέχουσα ημερομηνία.
- ✓ Με βάση αυτές τις παρατηρήσεις υπολογίζεται ο κινητός κεντρικός μέσος όρος (ΚΚΜΟ) μήκους ίσου με 24 ή 168, ανάλογα το στάδιο στο οποίο βρισκόμαστε, οι κανονικοποιημένοι δείκτες εποχικότητας και τέλος η αποεποχικοποιημένη χρονοσειρά του τρέχοντος βήματος.
- ✓ Στη συνέχεια εφαρμόζεται το μοντέλο της μεθόδου που εξετάζεται επί των αποεποχικοποιημένων δεδομένων. Οι βέλτιστοι συντελεστές εξομάλυνσης και οι υπόλοιποι παράμετροι του μοντέλου υπολογίζονται επίσης βάσει των δεδομένων αυτών. Παράγονται οι προβλέψεις για το εικοσιτετράωρο που μελετάται (εικοσιτέσσερις ωριαίες τιμές ζήτησης φορτίου). Αυτές δε διαθέτουν εποχικότητα, η οποία προστίθεται μετά, με πολλαπλασιασμό επί των αντίστοιχων για κάθε ωριαία τιμή δεικτών και λαμβάνεται το τελικό αποτέλεσμα. Έχουμε το πρότυπο της πολλαπλασιαστικής εποχικότητας στην περίπτωση μας, όπως συνηθίζεται άλλωστε.

Αυτή η διαδικασία επαναλαμβάνεται για όλες τις μεθόδους πρόβλεψης μιας εποχικότητας της παρούσας εργασίας, για τα δύο στάδια με ημερήσια και εβδομαδιαία εποχικότητα και (σε κάθε στάδιο) για τις διαφορετικές χρονικές περιόδους που επιλέγεται να εξεταστούν. Στην περίπτωση της μεθόδου του Taylor παραλείπεται το τμήμα της αποεποχικοποίησης αλλά και η επανεποχικοποίηση των προβλέψεων στο τέλος. Η ακρίβεια των μοντέλων πρόβλεψης ελέγχεται όπως είπαμε για το έτος 2011 για την περίοδο των 8 πρώτων μηνών, επομένως η εφαρμογή της διαδικασίας που μόλις αναφέρθηκε επαναλαμβάνεται 243 φορές, όσες και οι ημέρες της περιόδου. Ακόμη, το μήκος του χρονικού διαστήματος των δεδομένων κύλισης μπορεί να είναι σταθερό κατά τη διάρκεια όλης της διαδικασίας για μία μέθοδο, ωστόσο μεταβάλλεται στις διαφορετικές εκτελέσεις της μεθόδου.

Στην περίπτωση της συγκεκριμένης χρονοσειράς, επιλέγονται χρονικά διαστήματα μίας ως και δεκαπέντε εβδομάδων (περίπου τρεισήμισι μηνών). Θεωρήθηκε από την αρχή ότι εφόσον θα εφαρμοστεί κύλιση των δεδομένων, χρήσιμο είναι να εξεταστούν ακόμη και οι περιπτώσεις πολύ μικρού k όπως μίας ή δύο εβδομάδων, έως και περίπου ενός τριμήνου ή τετραμήνου. Η μέγιστη τιμή του δείκτη k επιλέχθηκε οριστικά κατά την εκτέλεση των πειραμάτων, οπότε και παρατηρήθηκε καλύτερα η πρόοδος των σφαλμάτων κατά τη μεταβολή της χρονικής περιόδου. Όπως φαίνεται και στην παράγραφο 6.1, με την αύξηση του δείκτη k κατά τις αρχικές επαναλήψεις των μοντέλων, προκαλείται μείωση των σφαλμάτων, μέχρι μιας τιμής (όχι πάντα η ίδια για όλες τις μεθόδους) στην οποία επιτυγχάνεται το ελάχιστο σφάλμα και έπειτα η τιμή του αρχίζει και πάλι ανοδική πορεία. Αποφασίστηκε να γίνουν οι επαναλήψεις σε κάθε μέθοδο για διαφορετικές χρονικές περιόδους έως και τη μέγιστη τιμή του δείκτη k , από την οποία και έπειτα επιβεβαιώνεται η αύξηση των σφαλμάτων.

Πιο ειδικά, για τις απλές μεθόδους μιας εποχικότητας γίνεται επανάληψη των μοντέλων με χρήση rolling forecasting για μήκη διαστήματος από μία (1) έως και δεκαπέντε (15) εβδομάδες. Το μήκος του διαστήματος συμβολίζεται με k και λαμβάνει τις συγκεκριμένες τιμές: $k=1,2,3,4,5,6,7,8,9,10,15$. Στο πρώτο στάδιο (της ημερήσιας εποχικότητας)

χρησιμοποιούνται όλα τα μήκη από $k=1$ ως και $k=15$. Στο δεύτερο στάδιο με την εβδομαδιαία εποχικότητα δε, μελετώνται οι περιπτώσεις για $k=2$ έως και $k=15$. Αυτό συμβαίνει επειδή για τον υπολογισμό των δεικτών εβδομαδιαίας εποχικότητας απαιτείται αρχικά ο ΚΚΜΟ(168), επομένως χρειάζονται περισσότερα από μιας εβδομάδας δεδομένα για να καταφέρουμε να παράγουμε προβλέψεις. Και όπως φαίνεται στην παράγραφο 6.1, η περίπτωση για $k=15$ εβδομάδες είναι περιττή, καθώς τα αποτελέσματα παύουν να δέχονται βελτίωση ήδη από τις δέκα εβδομάδες. Η μέθοδος που ανέπτυξε ο Taylor μελετάται για μεγαλύτερο εύρος χρονικών διαστημάτων κύλισης δεδομένων, από 3 έως και 55 εβδομάδες. Το κάτω όριο υπάρχει λόγω της ανάγκης των πρώτων δύο εβδομάδων για την αρχικοποίηση των δεικτών εποχικότητας D_{t0} , W_{t0} . Ενώ το αρκετά μεγάλο πλήθος μηκών που εξετάζονται, υφίσταται λόγω της αυξομείωσης των σφαλμάτων της μεθόδου με την αύξηση της τιμής του μήκους k . Θέλαμε επομένως να είμαστε σίγουροι ότι έχουμε καταγράψει σωστά τη μικρότερη δυνατή τιμή σφάλματος από τις προβλέψεις της μεθόδου.

Στο δεύτερο βήμα της διαδικασίας παραγωγής προβλέψεων σημειώνεται σε κάθε εκτέλεσή της υπολογίζονται και οι νέοι συντελεστές εξομάλυνσης οπότε και γίνεται αρχικοποίηση του επιπέδου, της τάσης και των δεικτών εποχικότητας, ανάλογα με το μοντέλο πρόβλεψης που εφαρμόζεται. Να σημειωθεί εδώ ότι για την αρχικοποίηση του επιπέδου και της τάσης των χρονοσειρών αξιοποιούνται όλα τα δεδομένα της εκάστοτε χρονικής περιόδου και χρησιμοποιείται σε όλες τις περιπτώσεις ο ίδιος τύπος για το καθένα. Συγκεκριμένα, η αρχική τιμή S_0 υπολογίζεται ως το σταθερό επίπεδο του μοντέλου απλής γραμμικής παλινδρόμησης και η τιμή T_0 ως η σταθερά της κλίσης του, επί όλων των δεδομένων (δηλαδή για περίοδο από μία έως και δεκαπέντε εβδομάδες).

5.4 Αξιολόγηση

Έπειτα από την επανάληψη της διαδικασίας παραγωγής πρόβλεψης για κάθε εξεταζόμενο μοντέλο, για τα δύο στάδια με τις εποχικότητες, για όλα τα επιλεγμένα μήκη εφαρμογής της κύλισης δεδομένων και για όλες τις ημέρες του 2011 που μας ενδιαφέρουν, ακολουθεί ο υπολογισμός σφαλμάτων, ανάμεσα στις πραγματικές παρατηρήσεις (της αρχικής χρονοσειράς με εποχικότητα) και στις παρατηρήσεις που υπολογίστηκαν από το κάθε μοντέλο πρόβλεψης (τις τελικές, με ενσωμάτωση της εποχικότητας). Στο βήμα αυτό, για την αξιολόγηση της ακρίβειας των μεθόδων που εφαρμόζονται, υπολογίζονται τα σφάλματα MAPE (μέσο απόλυτο ποσοστιαίο σφάλμα) και sMAPE (συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα), καθώς και τα MdAPE (διάμεσος του απόλυτου ποσοστιαίου σφάλματος) και sMdAPE (διάμεσος του συμμετρικού απόλυτου ποσοστιαίου σφάλματος), που αναφέρθηκαν ήδη στην παράγραφο 4.6.

Μεγάλο ζήτημα είναι η επιλογή του τύπου σφάλματος στο οποίο στηρίζεται η αξιολόγηση των μεθόδων και η λήψη των αποφάσεων από το χειριστή-εταιρεία. Όπως είναι επόμενο έχει γίνει και η αντίστοιχη μελέτη και ανάλυση των σφαλμάτων, ανάλογα με το είδος των χρονοσειρών, κλπ [Makridakis, 1993] και [Hyndman and Koehler, 2006]. Στην παράγραφο 4.6 σημειώθηκαν μαζί με τους τύπους των σφαλμάτων αξιολόγησης των μεθόδων και ορισμένα θετικά και αρνητικά χαρακτηριστικά τους. Στη συνέχεια αναφέρουμε πιο συγκεκριμένα τα χαρακτηριστικά αυτά για τα σφάλματα που χρησιμοποιούνται στην εργασία αυτή και η παρουσίαση των οποίων γίνεται στο επόμενο κεφάλαιο.

Το σφάλμα MAPE αντιμετωπίζει ευρεία χρήση παρότι έχει δεχτεί αρκετές κριτικές. Η απουσία συμμετρίας του και η αύξηση της τιμής του για τιμές δεδομένων κοντά στο μηδέν είναι ορισμένα από τα βασικά μειονεκτήματά του και για αυτό αναπτύχθηκε και το sMAPE

και προωθήθηκε η χρήση του. Ο Μακριδάκης [Makridakis, 1993] πρότεινε τη χρήση του sMAPE και την εισήγαγε στους διαγωνισμούς M2 και M3. Ωστόσο, σημειώθηκε [Koeehler 2001] ότι το sMAPE δεν αντιμετωπίζει τόσο συμμετρικά τα θετικά και τα αρνητικά σφάλματα, ευνοώντας μάλιστα και αυτό τα θετικά σφάλματα (με αποτέλεσμα την αύξηση της τιμής του). Στην εργασία αυτή, καθώς οι τιμές των δεδομένων της χρονοσειράς που διαθέτουμε είναι πολύ υψηλές (όχι κοντά στο μηδέν) και δεν υπάρχει διακοπτόμενη ζήτηση φορτίου ηλεκτρικής ενέργειας, ο δείκτης MAPE θεωρείται ότι μπορεί να δώσει πολύ ικανοποιητική εικόνα των αποτελεσμάτων κάθε μεθόδου πρόβλεψης. Η χρήση του MdAPE γίνεται επειδή είναι πιο σταθερό σφάλμα απέναντι σε ανωμαλίες των τιμών της χρονοσειράς, συγκριτικά με το MAPE. Ακόμη είναι εύκολη η ερμηνεία του, δεν είναι εξαρτώμενο από κλίμακα, είναι αξιόπιστο και έγκυρο. Πρόκειται για τη διάμεσο των σφαλμάτων APE, δηλαδή το ημίθροισμα των δύο μεσαίων σφαλμάτων APE, όταν αυτά τοποθετηθούν κατά αύξουσα σειρά. Αντίστοιχα έχουμε και το sMdAPE, σε ακολουθία του σφάλματος sMAPE.

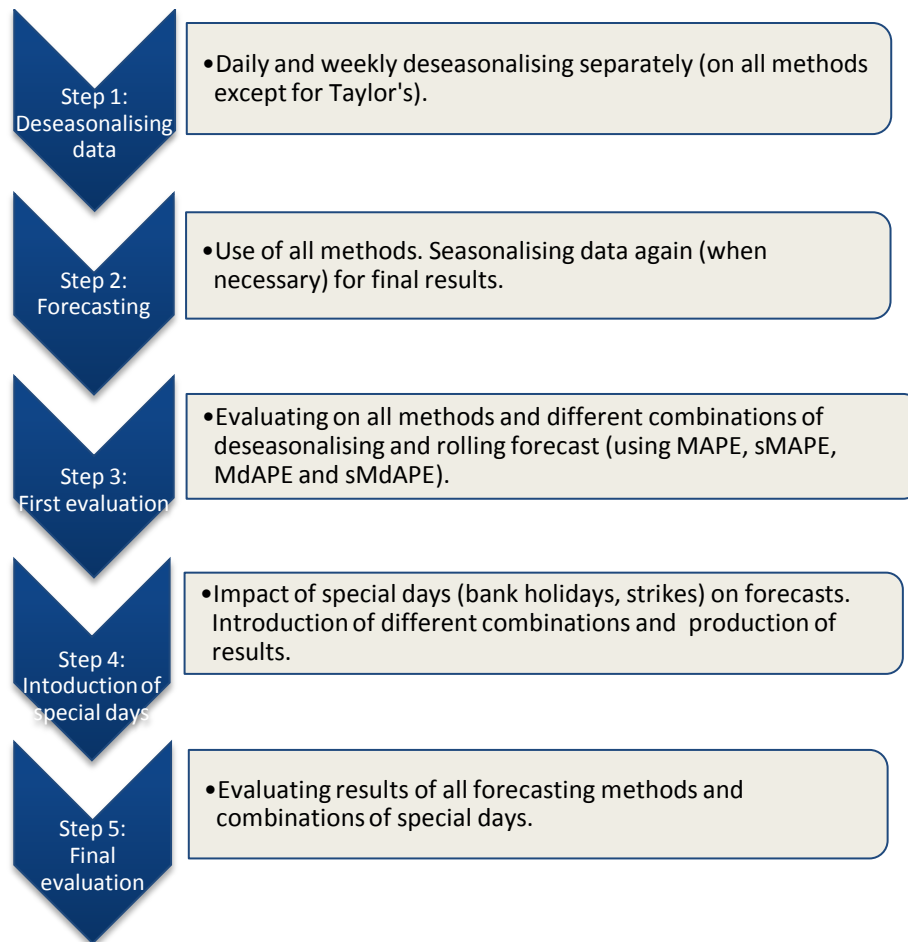
Ο υπολογισμός των σφαλμάτων MAPE και sMAPE γίνεται επί ολόκληρης της παραγόμενης χρονοσειράς για τους 8 μήνες του 2011. Ουσιαστικά αυτός ο μέσος όρος προκύπτει από σφάλματα 5832 πραγματικών παρατηρήσεων-προβλέψεων. Και για κάθε παραγόμενη χρονοσειρά υπολογίζονται και τα αντίστοιχα MdAPE και sMdAPE. Εκτός από τον υπολογισμό των ποσοστιαίων σφαλμάτων σε όλη τη χρονοσειρά για το 2011, γίνεται υπολογισμός τους και για κάθε ώρα της ημέρας. Υπολογίζονται δηλαδή 24 τιμές για κάθε σφάλμα, όσες και οι ώρες της ημέρας, ως ο μέσος όρος των σφαλμάτων των αντίστοιχων ωρών και για τους οκτώ μήνες. Κάτι ακόμη που μας ενδιαφέρει επίσης να εξετάσουμε το ποσοστό των σφαλμάτων της προβλεπόμενης χρονοσειράς, τα οποία έχουν τιμή μικρότερη του 3%.

Έπειτα από την παραγωγή των προβλέψεων με βάση τις διαδικασίες που περιγράφηκαν στο κεφάλαιο 5 (παράγραφοι 5.2 και 5.3), αναμένουμε να δούμε τη συμπεριφορά των μεθόδων που επιλέξαμε να μελετήσουμε. Ιδιαίτερο ενδιαφέρον προκαλεί η σύγκριση των πιο απλών στατιστικών μεθόδων συγκριτικά με τη «βελτιωμένη» μέθοδο εκθετικής εξομάλυνσης Holt-Winters που εισήγαγε ο Taylor, καθώς είναι η μόνη που ενσωματώνει και τις δύο εποχικότητες στο μοντέλο της. Αξιοπρόσεκτη είναι και η σύγκριση ανάμεσα στα δύο στάδια υλοποίησης κάθε στατιστικής μεθόδου ξεχωριστά. Με το πρώτο στάδιο να λαμβάνει υπόψη την επίδραση της ημερήσιας εποχικότητας στα δεδομένα, ενώ στο δεύτερο στάδιο συμπεριλαμβάνεται η συμπεριφορά της χρονοσειράς για μία ολόκληρη εβδομάδα, επομένως αναμένουμε καλύτερες προβλέψεις. Και μέσα στο ίδιο στάδιο ωστόσο, δίνεται προσοχή στα αποτελέσματα, καθώς αναμένεται η εύρεση της εκτέλεσης στην οποία κάθε μέθοδος πετυχαίνει μεγαλύτερη ακρίβεια στις προβλέψεις της (σε ποια χρονική περίοδο εφαρμογής της κυλιόμενης πρόβλεψης δηλαδή). Εκτός από την εκτέλεση για την κάθε μέθοδο ξεχωριστά, σημασία έχει και η σύγκριση των διαφορετικών εκτελέσεων μεταξύ τους. Ενσωματώνοντας την επίδραση της εποχικότητας στα μοντέλα που χρησιμοποιούνται, όπως και τη μέθοδο της κύλισης των δεδομένων, αναμένουμε αρκετά ικανοποιητικά αποτελέσματα.

Με το πέρας της διαδικασίας αυτής για όλες τις μεθόδους πρόβλεψης, προχωρούμε στην ενσωμάτωση της επίδρασης των ειδικών ημερών (αργίες και απεργίες) όπως σημειώθηκε από την αρχή της εργασίας. Έχοντας παράγει και σε αυτό το στάδιο όλες τις ζητούμενες προβλέψεις, υπολογίζονται στη συνέχεια τα νέα σφάλματα επί των χρονοσειρών. Εκτός από την απεικόνιση των σφαλμάτων που αναφέρθηκαν παραπάνω, παρουσιάζεται πλέον και η ποσοστιαία μεταβολή των σφαλμάτων για κάθε μέθοδο από την περίπτωση της πρώτης πρόβλεψης (πρώτο μέρος της μελέτης) στην περίπτωση των ειδικών ημερών (δεύτερο

μέρος). Αναμένεται, όπως είναι φυσικό, η μεταβολή αυτή να είναι αρνητική, ενώ θεωρείται ότι επιτυγχάνεται μεγαλύτερη ακρίβεια στις προβλέψεις και συνεπώς μικρότερα σφάλματα σε σχέση με το πρώτο μέρος παραγωγής πρόβλεψης ζήτησης φορτίου.

Να σημειωθεί εδώ ότι τόσο η διαδικασία ενσωμάτωσης της επίδρασης των ειδικών ημερών, όσο και τα αποτελέσματα της προσπάθειας αυτής παρουσιάζονται όλα μαζί στο κεφάλαιο 7. Ενώ στην επόμενη εικόνα (εικόνα 6) εμφανίζεται η ολοκληρωμένη διαδικασία παραγωγής προβλέψεων που ακολουθείται στα πλαίσια της παρούσας μελέτης. Δεν αναφέρεται στο σχήμα, αλλά καθ' όλη τη διαδικασία γίνεται εφαρμογή του rolling forecasting (όπως περιγράφηκε ανωτέρω).



Εικόνα 6. Ολοκληρωμένη η διαδικασία προβλέψεων

ΚΕΦΑΛΑΙΟ 6: Αποτελέσματα

6.1 Παρουσίαση αποτελεσμάτων

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα που προέκυψαν από την παραγωγή προβλέψεων ζήτησης φορτίου ηλεκτρικής ενέργειας για το 2011. Οι μέθοδοι που εξετάστηκαν είναι η απλοϊκή μέθοδος Naive, οι μέθοδοι εκθετικής εξομάλυνσης SES, Holt και Damped, η κλασική μέθοδος Theta και η προσαρμογή της μεθόδου Holt-Winters από τον Taylor για ενσωμάτωση διπλής εποχικότητας. Εκτός της τελευταίας μεθόδου, οι υπόλοιπες υφίστανται αποεποχικοποίηση των δεδομένων τους και για όλες τις μεθόδους, για ολόκληρη τη διαδικασία πρόβλεψης, εφαρμόζεται κύλιση επί των δεδομένων της χρονοσειράς.

Χρησιμοποιούνται περισσότερα του ενός κριτήρια αξιολόγησης των μεθόδων, η ελαχιστοποίηση του MAPE, του sMAPE, του MdAPE και του sMdAPE. Με έμφαση στα MAPE και MdAPE, καθώς η χρονοσειρά είναι τέτοιας μορφής που το MAPE μπορεί να εφαρμοστεί και να δώσει πολύ ικανοποιητικά αποτελέσματα (δεν είναι διακοπτόμενης ζήτησης και οι τιμές των δεδομένων είναι πολύ μεγαλύτερες του μηδενός, όπως τονίστηκε και στην παράγραφο 5.3). Έπειτα από την παρουσίαση των σφαλμάτων για όλες τις μεθόδους, παρουσιάζεται και η κατάταξη αυτών βάσει των σφαλμάτων.

Στη συνέχεια παρατίθενται τμήμα των σφαλμάτων MAPE και sMAPE των μεθόδων που χρησιμοποιήθηκαν για την παραγωγή της πρόβλεψης ζήτησης φορτίου στους πίνακες 2 και 3. Ο υπολογισμός των σφαλμάτων έγινε επί όλων των προβλεπόμενων τιμών ζήτησης για το 2011, για τις δύο εποχικότητες και για όλες τις εβδομάδες εφαρμογής rolling forecast. Τα αποτελέσματα για όλες τις εβδομάδες κυλιόμενης πρόβλεψης βρίσκονται στο Appendix (πίνακες A1 έως A4).

		Naive (deseasonal.&rolling)		SES (deseasonal.&rolling)		Holt (deseasonal.&rolling)	
		MAPE(%)	sMAPE(%)	MAPE(%)	MAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	6,624	6,651	6,060	5,866	10,633	10,278
	k=2	6,624	6,651	6,017	5,852	9,346	9,071
	k=3	6,624	6,651	6,052	5,890	8,794	8,471
	k=4	6,624	6,651	6,082	5,916	8,296	8,133
	k=5	6,624	6,651	6,133	5,963	7,667	7,615
	k=6	6,624	6,651	6,211	6,033	7,335	7,252
e=168	k=4	3,893	3,870	3,822	3,736	4,753	4,583
	k=5	3,664	3,642	3,696	3,604	4,202	4,087
	k=6	3,533	3,511	3,774	3,674	4,125	3,980
	k=7	3,512	3,487	3,914	3,808	3,926	3,819
	k=8	3,474	3,450	4,033	3,919	4,044	3,928
	k=9	3,461	3,437	4,157	4,035	4,167	4,043
	k=10	3,473	3,449	4,281	4,152	4,289	4,158

Πίνακας 2. Σφάλματα MAPE και sMAPE μεθόδων (α)

		Damped (deseasonal.&rolling)		Theta (deseasonal.&rolling)		Taylor (deseasonal.&rolling)	
		MAPE(%)	MAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	6,257	6,257	10,278	6,540	6,397		
	6,143	6,143	9,071	6,308	6,192		
	6,150	6,150	8,471	6,319	6,209		
	6,157	6,157	8,133	6,352	6,243		
	6,188	6,188	7,615	6,401	6,292		
	6,257	6,257	7,252	6,470	6,360		
e=168	3,898	3,898	4,583	3,825	3,768	4,716	4,654
	3,763	3,763	4,087	3,712	3,652	4,863	4,888
	3,831	3,831	3,980	3,768	3,704	5,254	5,332
	3,952	3,952	3,819	3,903	3,837	4,877	4,795
	4,063	4,063	3,928	4,007	3,939	4,603	4,589
	4,187	4,187	4,043	4,120	4,048	5,243	5,175
	4,291	4,291	4,158	4,243	4,170	5,711	5,922

Πίνακας 3. Σφάλματα MAPE και sMAPE μεθόδων (β)

Να σημειωθεί ότι με e συμβολίζεται το μήκος του κεντρικού κινητού μέσου όρου (ΚΚΜΟ) που υπολογίζεται στις στατιστικές μεθόδους, ώστε να ενσωματωθεί η κάθε εποχικότητα χωριστά, και έτσι είναι $e=14$ (ημερήσια) ή $e=168$ (εβδομαδιαία). Και με k συμβολίζεται το χρονικό διάστημα εφαρμογής της κύλισης στα δεδομένα, σε εβδομάδες.

Αντίστοιχα, στους πίνακες 4 και 5 παρουσιάζονται τα σφάλματα MdAPE και sMdAPE για τις περισσότερες εκτελέσεις των μεθόδων που μελετώνται, με εφαρμογή αποεποχικοποίησης και της μεθόδου κυλιόμενης πρόβλεψης. Και πάλι, συνολικά τα αποτελέσματα παρουσιάζονται στο Appendix.

		Naive (deseasonal.&rolling)		SES (deseasonal.&rolling)		Holt (deseasonal.&rolling)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,780	3,772	3,842	3,883	6,501	6,624
	k=2	3,780	3,772	4,090	4,121	5,980	6,048
	k=3	3,780	3,772	4,123	4,167	5,537	5,622
	k=4	3,780	3,772	4,092	4,123	5,396	5,442
	k=5	3,780	3,772	4,089	4,134	5,149	5,196
	k=6	3,780	3,772	4,111	4,147	4,960	5,001
e=168	k=4	2,462	2,461	2,360	2,364	2,624	2,632
	k=5	2,315	2,313	2,317	2,320	2,503	2,505
	k=6	2,202	2,203	2,394	2,396	2,468	2,474
	k=7	2,137	2,140	2,527	2,531	2,542	2,546
	k=8	2,137	2,136	2,636	2,634	2,660	2,658
	k=9	2,103	2,107	2,708	2,714	2,709	2,720
	k=10	2,122	2,121	2,822	2,819	2,840	2,839

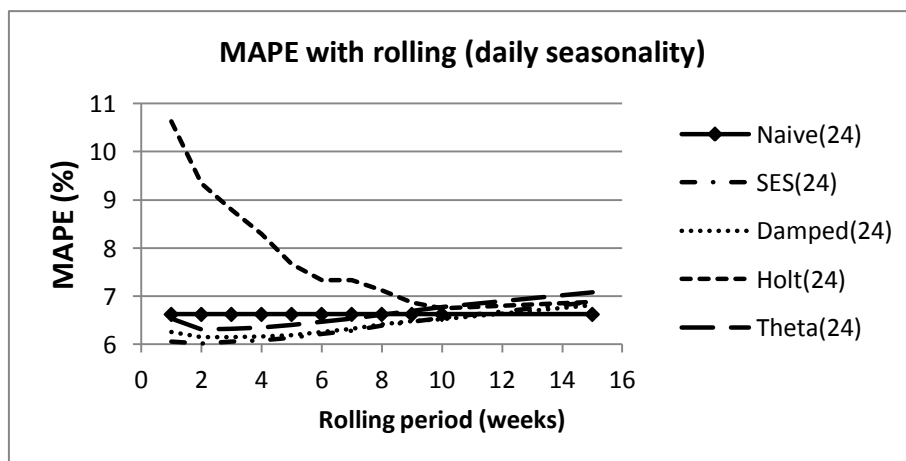
Πίνακας 4. Σφάλματα MdAPE και sMdAPE μεθόδων (α)

		Damped (deseasonal.&rolling)		Theta (deseasonal.&rolling)		Taylor (deseasonal.&rolling)	
		MdAPE(%)	MdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE (%)	sMdAPE (%)
e=24	k=1	3,765	6,624	4,194	4,240		
	k=2	4,110	6,048	4,356	4,388		
	k=3	4,156	5,622	4,407	4,419		
	k=4	4,029	5,442	4,464	4,462		
	k=5	3,998	5,196	4,492	4,539		
	k=6	3,976	5,001	4,532	4,583		
e=168	k=4	2,404	2,632	2,386	2,403	2,620	2,624
	k=5	2,321	2,505	2,377	2,393	2,791	2,786
	k=6	2,391	2,474	2,385	2,386	3,063	3,061
	k=7	2,562	2,546	2,533	2,543	2,923	2,917
	k=8	2,667	2,658	2,645	2,647	2,759	2,759
	k=9	2,726	2,720	2,719	2,721	3,070	3,082
	k=10	2,827	2,839	2,839	2,845	3,402	3,406

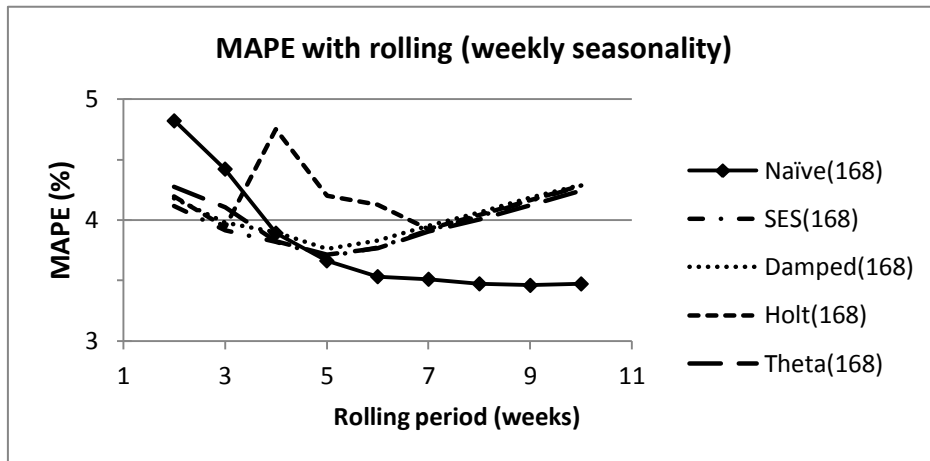
Πίνακας 5. Σφάλματα MdAPE και sMdAPE μεθόδων (β)

Ενώ στις επόμενες εικόνες, 7 ως και 12, εμφανίζονται διαγράμματα όλων των σφαλμάτων MAPE και MdAPE, για καλύτερη αντίληψη της πορείας τους με την εφαρμογή διαφορετικών μεθόδων και χρονικών διαστημάτων υπολογισμού in-sample error. Οι εικόνες 7 και 8 δείχνουν τα σφάλματα MAPE της απλοϊκής μεθόδου Naive, των μοντέλων εκθετικής εξομάλυνσης SES, Damped, Holt και της κλασσικής μεθόδου Theta για την εκτέλεση της μεθόδου με ημερήσια και με εβδομαδιαία εποχικότητα αντίστοιχα και για όλα τα χρονικά διαστήματα εφαρμογής rolling forecast (από μία έως και δεκαπέντε εβδομάδες για ημερήσια εποχικότητα και από δύο έως και δέκα εβδομάδες για εβδομαδιαία εποχικότητα). Με την ίδια λογική παρουσιάζονται και τα σφάλματα MdAPE στις εικόνες 10 και 11 για τις ίδιες μεθόδους.

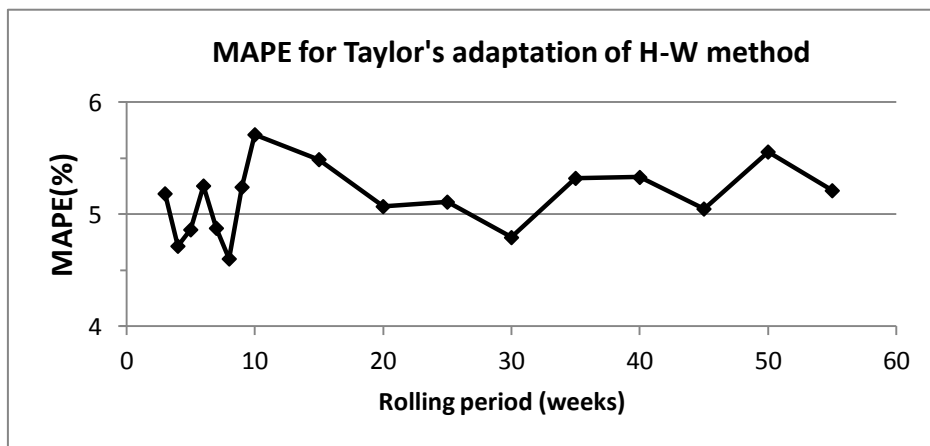
Τέλος, στην εικόνα 9 παρουσιάζεται το σφάλμα MAPE για το μοντέλο που ανέπτυξε ο Taylor και στην εικόνα 12 το αντίστοιχο σφάλμα MdAPE. Η εν λόγω μέθοδος παρουσιάζεται ξεχωριστά από τις προηγούμενες, επειδή κρίθηκε ιδιαίτερου ενδιαφέροντος λόγω ταυτόχρονης της συμπερίληψης και των δύο εποχικότητων στο μοντέλο της.



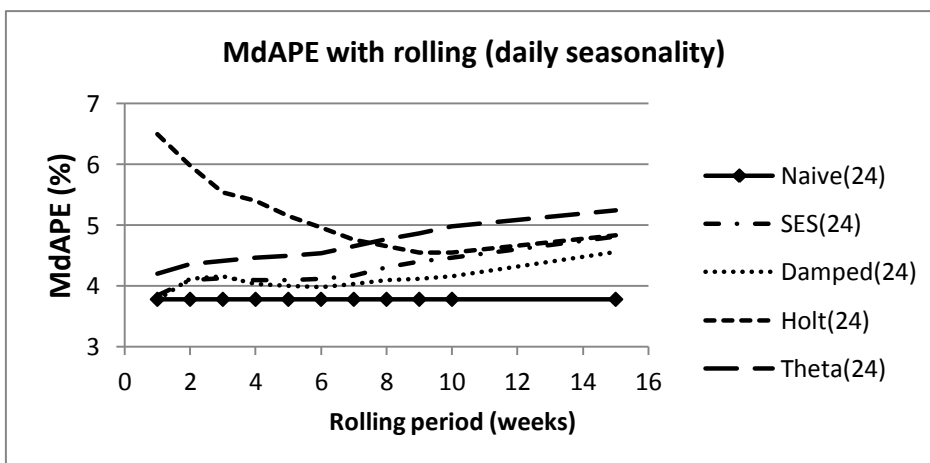
Εικόνα 7. Σφάλμα MAPE για ημερήσια εποχικότητα



Εικόνα 8. Σφάλμα MAPE για εβδομαδιαία εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα

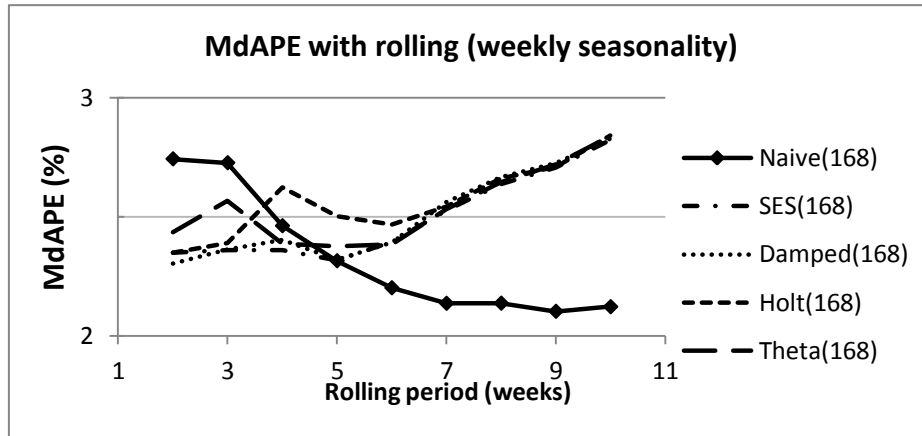


Εικόνα 9. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με εφαρμογή rolling για διαφορετικά χρονικά διαστήματα

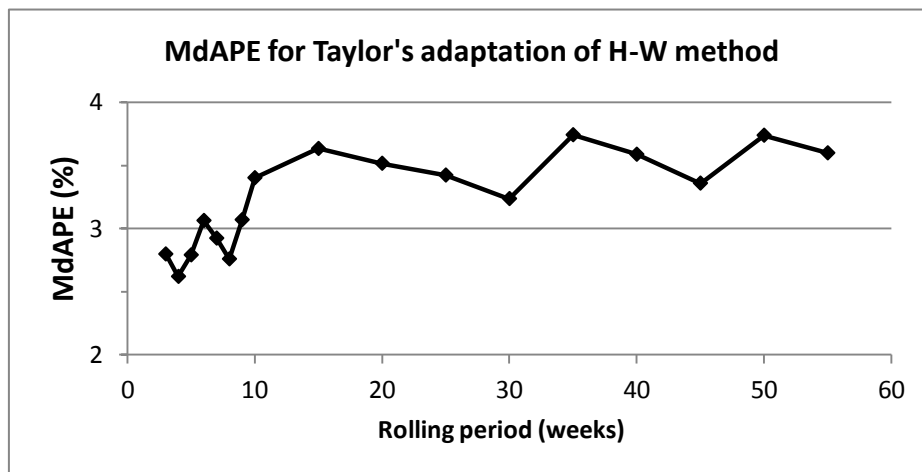


Εικόνα 10. Σφάλμα MdAPE για ημερήσια εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα

Σημειώνουμε ότι ενώ στους πίνακες αναφέρονται και οι τέσσερις κατηγορίες σφαλμάτων, MAPE, sMAPE, MdAPE και sMdAPE, στα διαγράμματα απεικονίζονται οι δύο εξ αυτών μόνο. Αυτό συμβαίνει επειδή, η πορεία των σημείων στα διαγράμματα για το σφάλμα sMAPE είναι ίδια με εκείνη του MAPE και αντίστοιχα του sMdAPE για το MdAPE.



Εικόνα 11. Σφάλμα MdAPE για εβδομαδιαία εποχικότητα και εφαρμογή rolling για διαφορετικά χρονικά διαστήματα



Εικόνα 12. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με εφαρμογή rolling για διαφορετικά χρονικά διαστήματα

6.2 Ανάλυση αποτελεσμάτων

Η παρουσίαση των σφαλμάτων κάθε μεθόδου στην προηγούμενη παράγραφο διευκολύνει την εξαγωγή συμπερασμάτων για τις διαδικασίες πρόβλεψης που ακολουθήθηκαν. Για όλες τις μεθόδους, εκτός της προσαρμογής του Taylor, έντονη είναι η μεγάλη διαφορά του επιπέδου των σφαλμάτων ανάμεσα στην παραγωγή πρόβλεψης με συμπερίληψη μόνο της ημερήσιας και μόνο της εβδομαδιαίας εποχικότητας. Στην πρώτη περίπτωση με την ημερήσια εποχικότητα παρατηρούνται πολύ υψηλότερα σφάλματα, τα οποία με μικρότερες ή μεγαλύτερες αποκλίσεις κυμαίνονται περίπου στην τιμή του 6 % για το MAPE. Στη δεύτερη

περίπτωση, της εβδομαδιαίας εποχικότητας, τα σφάλματα λαμβάνουν τιμές στο διάστημα [3% , 5%], κυμαινόμενες γύρω από το 4%. Αντίστοιχα για το MdAPE για τη συμπερίληψη της ημερήσιας εποχικότητας οι τιμές του σφάλματος κυμαίνονται στο 4%, ενώ για την περίπτωση της εβδομαδιαίας εποχικότητας οι αντίστοιχες τιμές του κινούνται γύρω από το 2.5%. Είναι εμφανής λοιπόν και η διαφορά επιπέδου τιμών ανάμεσα στα δύο αυτά σφάλματα, όπως αναμενόταν και από τη θεωρία. Αυτό συμβαίνει, επειδή το MAPE δεν είναι τόσο ανεπηρέαστο από τα μεγαλύτερα σφάλματα, όσο το MdAPE.

Συγκρίνοντας μεταξύ τους τα σφάλματα MAPE των μεθόδων που παρουσιάζονται στις εικόνες 7 και 8 συμπεραίνουμε τα ακόλουθα. Η μέθοδος Holt δεν είναι ακριβής σε σχέση με τις υπόλοιπες απλές μεθόδους μιας εποχικότητας, τα σφάλματα των πρώτων επαναλήψεων της πρόβλεψής της (για τα μικρά δηλαδή χρονικά διαστήματα εφαρμογής της μεθόδου κύλισης των δεδομένων) είναι μεγάλα και μάλιστα με διαφορά από των υπόλοιπων μοντέλων. Όσο αυξάνεται η χρονική περίοδος εφαρμογής του rolling forecasting ωστόσο, βελτιώνονται και τα σφάλματα της μεθόδου Holt, παραμένει όμως και πάλι η μέθοδος με τα χειρότερα αποτελέσματα (δεν αναφερόμαστε ακόμη στην προσαρμογή του Taylor). Οι υπόλοιπες μέθοδοι έχουν παραπλήσιες τιμές σφαλμάτων για όλες τις επαναλήψεις τους (για τα διαφορετικά k) στο πρώτο στάδιο της ενσωμάτωσης της ημερήσιας εποχικότητας και η διαφορά της μέγιστης από την ελάχιστη τιμή τους δεν ξεπερνάει τη μία ποσοστιαία μονάδα. Για το δεύτερο στάδιο πρόβλεψης, με συμπερίληψη της εβδομαδιαίας εποχικότητας, παρατηρείται ότι οι μέθοδοι ακολουθούν την ίδια πορεία, εκτός πλέον από την απλοϊκή Naive και τη Holt. Η τελευταία εμφανίζει έως και μία ποσοστιαία μονάδα αύξηση σφάλματος κατά την επανάληψη με $k=4$, διατηρεί διαφορά για τις μεσαίες επαναλήψεις, ενώ ακολουθεί το επίπεδο τιμών το υπόλοιπο διάστημα. Η Naive από την πλευρά της παρουσιάζει αρχικά αρκετά μεγαλύτερες τιμές σφαλμάτων (για μικρά k) και στη συνέχεια παρατηρείται έντονη μείωση αυτών, οπότε και μετά την επανάληψη για $k=5$ τα σφάλματά της είναι μόνιμα μικρότερα από τα αντίστοιχα των υπόλοιπων μεθόδων.

Παρόμοια εικόνα εμφανίζουν και τα σφάλματα MdAPE στις εικόνες 10 και 11, με τις παρακάτω βέβαια διαφορές. Η απλοϊκή μέθοδος στην περίπτωση αυτή φαίνεται να έχει το καλύτερο σφάλμα και σε επίπεδο ημερήσιας εποχικότητας, σε αντίθεση με προηγούμενους που οι μέθοδοι SES, Damped και Theta είχαν καλύτερα σφάλματα από την Naive. Η Holt παρουσιάζει σχεδόν σε όλες τις επαναλήψεις τα χειρότερα σφάλματα ανάμεσα στις πέντε αυτές μεθόδους, τόσο στο στάδιο με ημερήσια όσο και με την εβδομαδιαία εποχικότητα, βελτιώνοντας την επίδοσή της προς τις τελευταίες επαναλήψεις (έναντι μερικών και όχι όλων από τις άλλες μεθόδους). Ακόμη, όπως φαίνεται στην εικόνα 11 (συγκριτικά με την εικόνα 8) τα σφάλματα MdAPE έχουν πιο σταθερό επίπεδο τιμών στις αρχικές επαναλήψεις των μεθόδων, πριν αρχίσουν να αυξάνονται για $k=7$ και έπειτα.

Από τα διαγράμματα των εικόνων 9 και 12, με τα σφάλματα MAPE και MdAPE για το μοντέλο του Taylor, ή πρώτη πληροφορία που αποκομίζουμε είναι ότι η ακρίβεια της μεθόδου αυξάνεται και μειώνεται επαναλαμβανόμενα καθώς αυξάνει η τιμή του k . Παρουσιάζονται διακυμάνσεις του σφάλματος σε σχετικά σταθερά χρονικά διαστήματα, με εμφάνιση κορυφής ανά τρεις περίπου επαναλήψεις. Όσο αυξάνει όμως η τιμή του k , αυξάνεται και το επίπεδο των σφαλμάτων, υπάρχει δηλαδή μία θετική τάση, όσο μικρή και αν είναι. Η τάση αυτή φαίνεται καλύτερα από την παρατήρηση των σφαλμάτων σε ολόκληρο τον πίνακα (Appendix, πίνακας 1), όπου μπορεί να δει κανείς ότι η ελάχιστη τιμή κάθε φορά είναι μεγαλύτερη από ότι στον προηγούμενο κύκλο. Στο διάγραμμα της εικόνας 9 φαίνεται ότι το μεγαλύτερο από τα μέγιστα σφάλματα MAPE επιτυγχάνεται για $k=10$, ενώ στο διάγραμμα της εικόνας 12 το μέγιστο σφάλμα MdAPE παρουσιάζεται για $k=35$. Οι ελάχιστες τιμές τους από την άλλη πλευρά, εμφανίζονται για $k=8$ και για $k=4$ αντίστοιχα.

Υπάρχει λοιπόν, μια διαφορά ανάμεσα στα δύο σφάλματα, όσον αφορά την επανάληψη στην οποία αξιολογούν ως την περισσότερο και λιγότερο ακριβή. Στην περίπτωση του MAPE, τόσο η βέλτιστη όσο και η χειρότερη των αποδόσεων της μεθόδου βρίσκονται μόλις με μία επανάληψη διαφορά (για $k=9$) και για αρκετά μικρή τιμή του k . Στην περίπτωση του MdAPE όμως, η βέλτιστη απόδοση της μεθόδου κρίνεται ότι προκύπτει με τη χρήση μόλις τεσσάρων εβδομάδων παρατηρήσεων ($k=4$) ως in-sample και η χειρότερη απόδοση παρατηρείται για πολύ μεγαλύτερη χρονική περίοδο κύλισης ($k=35$).

Το πιο έντονο χαρακτηριστικό των παραπάνω πινάκων και διαγραμμάτων παρουσιάζεται στην απλοϊκή μέθοδο, καθώς για την εκτέλεση των προβλέψεων με ημερήσια εποχικότητα παρουσιάζει το ίδιο σφάλμα για όλα τα διαφορετικά διαστήματα στα οποία εφαρμόστηκε rolling forecasting. Το παράδοξο σε αυτή την περίπτωση είναι επειδή κάθε φορά κυλιούνται τα δεδομένα και υπολογίζονται νέοι -και διαφορετικοί σε τιμές- δείκτες εποχικότητας, όπως και διαφορετικές ημερήσιες προβλέψεις. Ωστόσο, το τελικό αποτέλεσμα μετά την επαναφορά της εποχικότητας στα αποτελέσματα της πρόβλεψης οδηγεί σε ίδιες τιμές τελικής πρόβλεψης εικοσιτετράωρου για όλες τις εφαρμογές του rolling.

Ακόμη, είναι εμφανής η υπεροχή της απλούστερης μεθόδου ακόμη και απέναντι στη Theta και την προσαρμογή του Taylor. Συνοπτικά παρουσιάζονται τα καλύτερα σφάλματα των μεθόδων στον πίνακα 6, προφανώς για συμπερίληψη μόνο της **εβδομαδιαίας** εποχικότητας στις απλές μεθόδους, καθώς όλων των μεθόδων τα βέλτιστα σφάλματα επιτυγχάνονται σε αυτό το στάδιο. Ο δείκτης k που αναφέρεται, είναι αυτός στον οποίο επιτυγχάνεται το ελάχιστο MAPE και για τη συγκεκριμένη επανάληψη παρουσιάζονται και τα υπόλοιπα σφάλματα.

Τα βέλτιστα σφάλματα (με rolling)					
Μέθοδοι	MAPE (%)	sMAPE (%)	MdAPE (%)	sMdAPE (%)	Rolling period (weeks)
Naive	3,461%	3,437%	2,103%	2,107%	$k=9$
	3,664%	3,642%	2,315%	2,313%	$k=5$
SES	3,696%	3,604%	2,317%	2,320%	$k=5$
Theta	3,712%	3,652%	2,377%	2,393%	$k=5$
Damped	3,763%	3,664%	2,321%	2,322%	$k=5$
Holt	3,926%	3,819%	2,542%	2,546%	$k=7$
Taylor	4,603%	4,589%	2,759%	2,759%	$k=8$

Πίνακας 6. Τα μικρότερα σφάλματα όλων των μεθόδων

Τα σφάλματα στον πίνακα έχουν καταχωρηθεί ιεραρχικά από το μικρότερο έως το μεγαλύτερο με βάση το σφάλμα MAPE. Η απλοϊκή μέθοδος, όπως φαίνεται, έχει το ελάχιστο σφάλμα, με το MAPE της να είναι μικρότερο κατά 6,36375% από το αμέσως μικρότερο MAPE (της μεθόδου SES) και αντίστοιχα για το sMAPE η ποσοστιαία μείωση ισούται με -4,64862%. Παρατηρείται βέβαια ότι η επίτευξη του βέλτιστου αποτελέσματος συμβαίνει στη μέθοδο Naive όταν αξιοποιούνται δεδομένα εννέα (9) εβδομάδων με τη μέθοδο κύλισης, ενώ για τις τρεις αμέσως επόμενες καλύτερες στατιστικές μεθόδους η καλύτερη απόδοσή τους παρουσιάζεται για χρήση πέντε (5) μόλις εβδομάδων από τα δεδομένα της αρχικής χρονοσειράς ως in-sample. Η μέθοδος Holt εμφανίζει σύμφωνα με το MAPE μεγαλύτερη

ακρίβεια για $k=7$ και για τη μέθοδο του Taylor βέλτιστο αποτέλεσμα επιτυγχάνεται με χρήση οκτώ (8) εβδομάδων αντίστοιχα. Αυτά πάντα για την περίπτωση εβδομαδιαίας εποχικότητας (δηλαδή με ΚΚΜΟ(168)). Ακόμη όμως και στην περίπτωση που λάβουμε τις τιμές σφαλμάτων και της απλοϊκής μεθόδου για πέντε (5) εβδομάδων rolling forecast, το MAPE είναι ακόμη το μικρότερο όλων των μεθόδων. Όχι όμως με τόση διαφορά όπως πριν, καθώς είναι μικρότερο μόλις κατά 0,86843%. Το sMAPE είναι πλέον λίγο μεγαλύτερο από το sMAPE της πρόβλεψης με SES (περίπου κατά 1%). Αν τώρα δει κανείς τα αποτελέσματα των υπόλοιπων μεθόδων για την υλοποίηση με $k=9$, θα παρατηρήσει ότι τα σφάλματά τους αυξάνονται πολύ περισσότερο από την αύξηση των σφαλμάτων της Naive για την υλοποίηση με $k=5$ (Πίνακες 2, 3, 4 και 5).

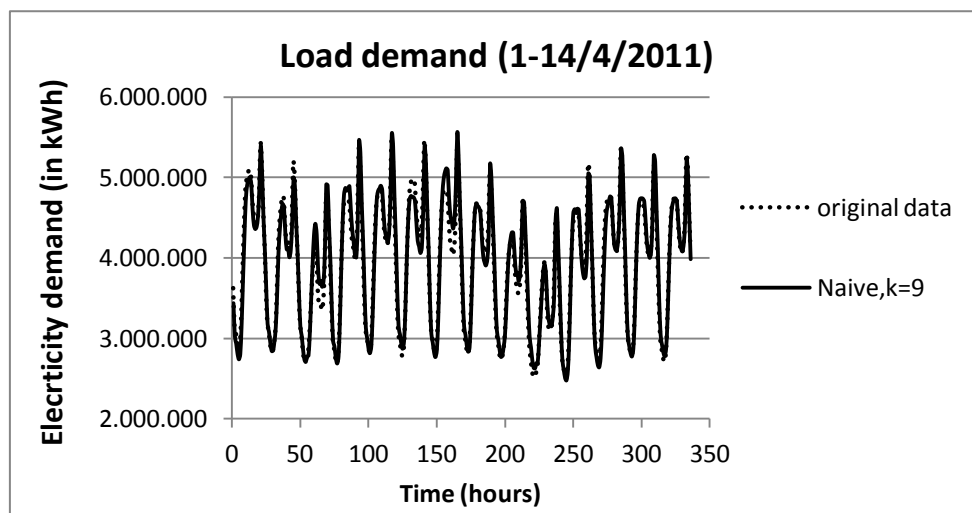
Είναι εντυπωσιακό το γεγονός ότι τα καλύτερα αποτελέσματα στην πρόβλεψη ζήτησης φορτίου προκύπτουν από την πιο απλή μέθοδο, με χρήση μόνο του κεντρικού κινητού μέσου όρου μήκους 168 (ωριαίων δεδομένων) για ενσωμάτωση της εβδομαδιαίας εποχικότητας της χρονοσειράς. Αναμέναμε ότι θα προκύψουν μέτρια σφάλματα από τη Naive, όχι όμως και τα καλύτερα σε σχέση με τις υπόλοιπες πολύ πιο αξιόλογες μεθόδους. Παρατηρώντας τις βέλτιστες τιμές των MdAPE και sMdAPE στον ίδιο πίνακα, σημειώνουμε κάποιες διαφορές. Η Naive, η SES και η Theta διατηρούν τις βέλτιστες τιμές των τεσσάρων σφαλμάτων για την ίδια χρονική περίοδο εφαρμογής της κυλιόμενης πρόβλεψης. Ενώ η Damped και η Holt πετυχαίνουν τα βέλτιστα MdAPE και sMdAPE για μόλις δύο εβδομάδων in-sample δεδομένα και η προσαρμογή του Taylor για τεσσάρων εβδομάδων δεδομένα, όπως αναφέρθηκε και παραπάνω.

Σε αντίθεση με τη μέθοδο Naive, η προσαρμογή του μοντέλου εκθετικής εξομάλυνσης Holt-Winters από τον Taylor, παρουσιάζει τα λιγότερο ικανοποιητικά αποτελέσματα, με το μικρότερο σφάλμα της (MAPE και sMAPE) να είναι κατά 33% περίπου μεγαλύτερο από το αντίστοιχο της Naive, τόσο για $k=8$ και για τις δύο μεθόδους, όσο και για $k=8$ για την μέθοδο του Taylor για $k=9$ για τη Naive. Η αδυναμία της μεθόδου αυτής να αποδώσει καλύτερα από τις υπόλοιπες σε συνδυασμό με τα μεγάλα σφάλματα των λοιπών στατιστικών μεθόδων στην περίπτωση του $e=24$ (ημερήσια εποχικότητα) οδηγεί στο συμπέρασμα ότι ενδεχομένως η συμπερίληψη στα μοντέλα της εποχικότητας σε ημερήσια βάση να μην είναι αποτελεσματική. Για μια καλύτερη εικόνα της επίδρασης των δύο εποχικοτήτων και στις υπόλοιπες μεθόδους ελέγξαμε σε κάθε μία (Naive, SES, κλπ) το συνδυασμό εβδομαδιαίας και ημερήσιας εποχικότητας (ξεχωριστή πρόβλεψη και συνδυασμός τελικής πρόβλεψης με ίσα βάρη), ο οποίος αποδείχθηκε ότι παράγει χειρότερα αποτελέσματα για τις στατιστικές μεθόδους. Χαρακτηριστικά για τη μέθοδο Naive, που πέτυχε το μικρότερο σφάλμα μόνο για $e=168$, με το συνδυασμό που αναφέραμε προέκυψαν τα εξής σφάλματα: MAPE=4,53932% και sMAPE=4,52492%. Ουσιαστικά, το αποτέλεσμα της πρόβλεψης και με τις δύο εποχικότητες οδηγεί σε σφάλμα περίπου ίσο με το μέσο όρο των δύο ξεχωριστών προβλέψεων (στο πρώτο στάδιο με ημερήσια εποχικότητα και στο δεύτερο στάδιο με εβδομαδιαία εποχικότητα). Εφόσον από τους παραπάνω πίνακες αποδείχθηκε ότι η ημερήσια εποχικότητα δεν συμβάλλει στην παραγωγή τόσο ικανοποιητικών προβλέψεων, είναι επόμενη η αρνητική επίδρασή της και στο μοντέλο του Taylor, με δύο εποχικότητες. Έτσι, είναι αρκετά λογική η διαφορά που παρουσιάζουν τα σφάλματα της μεθόδου του Taylor βάσει αυτής της παρατήρησης.

Ως δεύτερη καλύτερη μέθοδος πρόβλεψης παρουσιάζεται σύμφωνα με τα σφάλματα MAPE και sMAPE (στους πίνακες 2,3 και 6) η απλή εκθετική εξομάλυνση (SES) και επόμενες είναι η κλασική Theta, η Holt και η Damped. Οι απλούστερες μέθοδοι υπερισχύουν εν τέλει σε όλες τις προβλέψεις. Η τρίτη θέση της Theta με ιδιαίτερα ικανοποιητικά αποτελέσματα είναι κάτι που αναμέναμε εφόσον η μέθοδος αυτή χρησιμοποιεί διαφορετικά βάρη για την

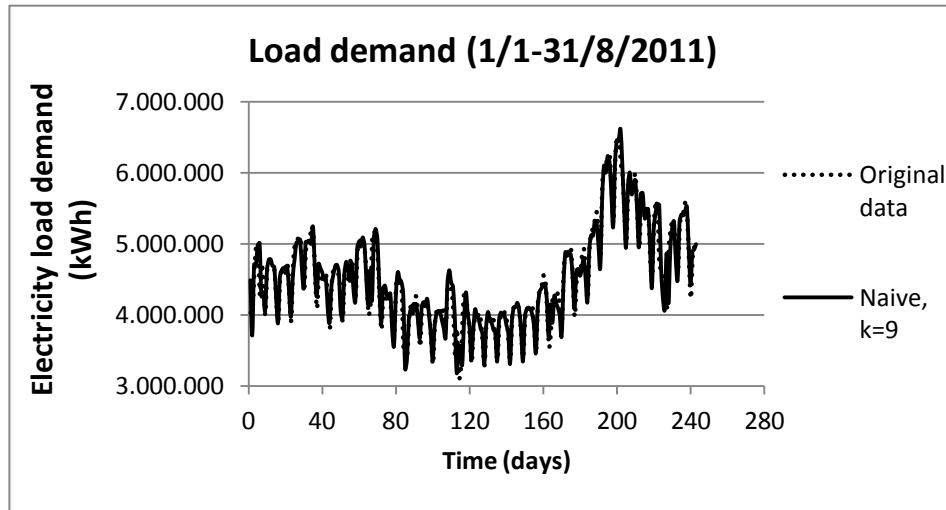
ενίσχυση της μακροπρόθεσμης ή της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς και εδώ με την εφαρμογή του κλασσικού μοντέλου διευκολύνεται η βραχυπρόθεσμη πρόβλεψη ζήτησης του φορτίου ηλεκτρισμού. Με πολύ μικρή διαφορά ακολουθεί η εκθετική εξομάλυνση φθίνουσας τάσης (Damped method), ενώ οι προβλέψεις του μοντέλου γραμμικής τάσης (Holt method) έχουν αποτελέσματα ανάμεσα σε αυτά των πρώτων και της προσαρμογής του Taylor, αρκετά χειρότερα από τα μεν και αρκετά καλύτερα από τα δε. Σύμφωνα με τα MdAPE και sMdAPE όμως (Πίνακες 4, 5), δεύτερη καλύτερη μέθοδος παρουσιάζεται η SES, έπειτα η Damped, η Holt και μετά η Theta, με την Taylor να καταλαμβάνει και πάλι την τελευταία θέση.

Στην εικόνα 13 παρουσιάζεται η πραγματική χρονοσειρά ζήτησης για τις δύο πρώτες εβδομάδες του Απρίλη του 2011, 1/4/2011(Παρασκευή) έως και 14/4/2011(Πέμπτη), καθώς και η προβλεπόμενη ζήτηση για τις ίδιες ημερομηνίες από την μέθοδο Naive, η οποία παράγει τα βέλτιστα αποτελέσματα όπως αναλύθηκε νωρίτερα (κύλιση 9 εβδομάδων και ενσωμάτωση εβδομαδιαίας εποχικότητας).



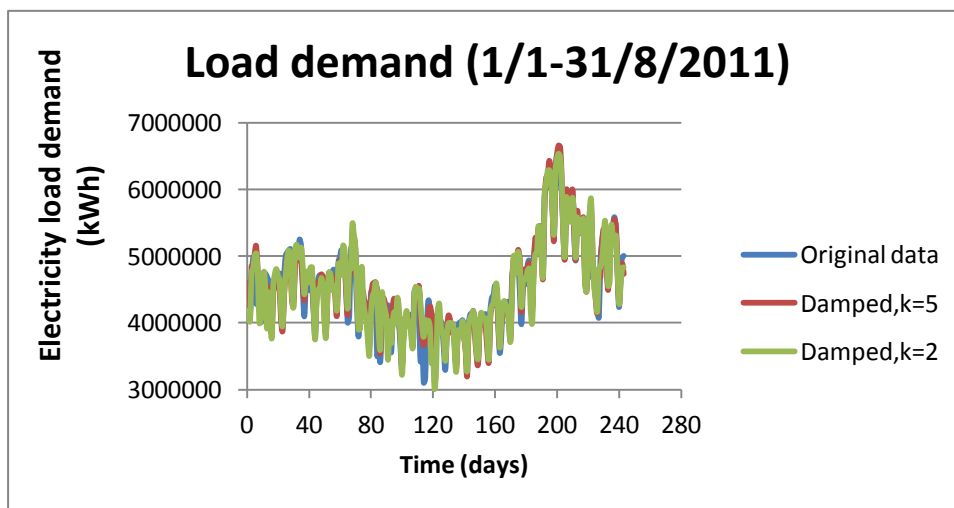
Εικόνα 13. Ζήτηση φορτίου ηλεκτρισμού για το διάστημα 1-14/4/2011 (πραγματική και βέλτιστη πρόβλεψη, Naive)

Στην εικόνα 14 παρουσιάζεται ολόκληρη η καμπύλη που παράγεται από τη μέθοδο Naive, με εβδομαδιαία εποχικότητα και για εφαρμογή της μεθόδου κύλισης για χρονική περίοδο ίση με εννέα εβδομάδες. Εύκολα καταλήγει κανείς στο ίδιο αποτέλεσμα με αυτό που έδειξαν και τα σφάλματα, ότι δηλαδή η απλοϊκή μέθοδος (και για τη συγκεκριμένη εκτέλεση) παράγει ιδιαίτερα ικανοποιητικές προβλέψεις.



Εικόνα 14. Ζήτηση φορτίου ηλεκτρισμού και για τους 8 μήνες του 2011 (πραγματική χρονοσειράς και βέλτιστη πρόβλεψη με Naive)

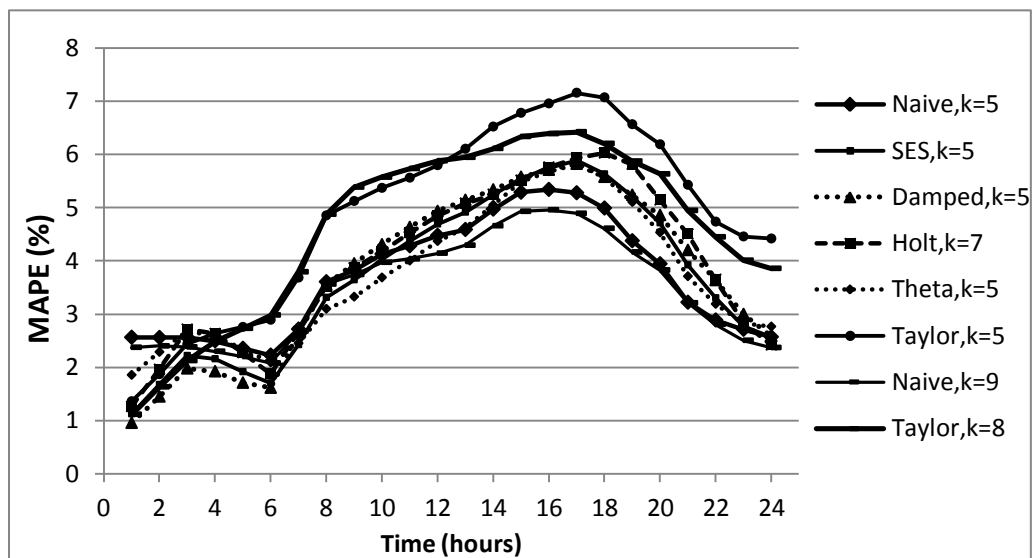
Στην επόμενη εικόνα παρατίθενται σε αντιστοιχία με την εικόνα 11, οι πραγματικές τιμές της χρονοσειράς των δεδομένων ζήτησης και οι προβλεπόμενες τιμές μέσω της μεθόδου Damped για τις επαναλήψεις με $k=5$ και $k=2$, στις οποίες παρατηρήθηκε διαφορετική αξιολόγηση από τα σφάλματα MAPE και MdAPE σύμφωνα με τον Πίνακα 6. Παρατηρώντας τα δεδομένα και των δύο επαναλήψεων της Damped σε σχέση με της πραγματικής χρονοσειράς, φαίνεται ότι η αξιολόγηση που λάβαμε από το δείκτη MAPE (και sMAPE μαζί) ανταποκρίνεται περισσότερο στην πραγματικότητα.



Εικόνα 15. Ζήτηση φορτίου ηλεκτρισμού και για τους 8 μήνες του 2011 (πραγματική και βέλτιστη προβλεπόμενη -μέσω Damped- χρονοσειρά)

Παρουσίαση σφαλμάτων σε 24ωρη βάση

Στη συνέχεια παρουσιάζονται για την περίοδο 1/1/2011 ως και 31/8/2011, τα out-of-sample σφάλματα MAPE σε εικοσιτετράωρη βάση και για όλες τις μεθόδους. Εκτός από τα σφάλματα για την καλύτερη υλοποίηση κάθε μοντέλου, πάντα σύμφωνα με το MAPE, (η Naive για $k=9$, οι SES, Damped, Holt, Theta για $k=5$ και η Taylor's adaption of Holt-Winters for double seasonality για $k=8$), παραθέτουμε ακόμη και για τις δύο μεθόδους με διαφορετικό k (Naive και προσαρμογή του Taylor) τα σφάλματά τους για $k=5$, ώστε να υπάρχει σύγκριση όλων των μεθόδων σε κοινή βάση.



Εικόνα 16. Μέσο ωριαίο σφάλμα MAPE για τις βέλτιστες προβλεπόμενες χρονοσειρές όλων των μεθόδων

Παρατηρούμε ότι κατά τις πρώτες ώρες της ημέρας (01:00 το βράδυ έως και 07:00 το πρωί) η απόκλιση των τιμών είναι της τάξης των δύο ποσοστιαίων μονάδων, με τα σφάλματα να λαμβάνουν τιμές στο διάστημα [1% , 3.5%]. Στη συνέχεια, καθώς αρχίζει να αυξάνεται η κατανάλωση ηλεκτρισμού, παρατηρείται και αύξηση του σφάλματος πρόβλεψης, με τις μέγιστες τιμές του να σημειώνονται στο χρονικό διάστημα 15:00 με 17:00 το απόγευμα. Οι τιμές αυτές κυμαίνονται από λίγο πάνω από 5% ως και λίγο πιο κάτω από 8%, πρόκειται επομένως για ιδιαίτερα μεγάλα σφάλματα. Έπειτα το σφάλμα πρόβλεψης αρχίζει να μειώνεται και πάλι για όλες τις μεθόδους, φτάνοντας κοντά στο 3% κατά την τελευταία ώρα της ημέρας. Συνεπώς υπάρχει αδυναμία όλων των μοντέλων να παράγουν καλές προβλέψεις για τις ώρες αιχμής της χρήσης ηλεκτρικής ενέργειας. Ένα ακόμη χαρακτηριστικό των αποτελεσμάτων είναι ότι το μοντέλο της Naive ($k=9$), πετυχαίνει καλύτερες προβλέψεις από τη μέθοδο Theta classic ($k=5$) εκτός από το διάστημα 7:00 έως και 11:00 το πρωί. Τα σφάλματα της μεθόδου του Taylor με διπλή εποχικότητα είναι μεγαλύτερα από των υπόλοιπων μεθόδων και για τις δύο περιπτώσεις, τόσο για $k=5$ όσο και για $k=8$. Τέλος, η διακύμανση των σφαλμάτων κάθε μεθόδου είναι διαφορετική, όπως παρουσιάζεται και στην επόμενη εικόνα (εικόνα 7). Η μεγαλύτερη διακύμανση παρατηρείται για τη μέθοδο Taylor (και για τις δύο εκτελέσεις της) και έπειτα για τη Damped. Και η μικρότερη διακύμανση παρουσιάζεται στην απλοϊκή μέθοδο (και στις δύο περιπτώσεις εκτέλεσής της) και στη συνέχεια στη μέθοδο Theta.

Να σημειωθεί ότι τα σφάλματα του πίνακα 7 δεν προκύπτουν ως ο καθολικός μέσος όρος των σφαλμάτων των ωριαίων προβλέψεων της χρονοσειράς για το 2011, αλλά όπως ειπώθηκε και παραπάνω πρόκειται για τα out-of-sample σφάλματα MAPE σε εικοσιτετράωρη βάση και για όλες τις μεθόδους. Συγκεκριμένα, παρουσιάζεται το ελάχιστο και το μέγιστο σφάλμα MAPE εκ των εικοσιτεσσάρων που υπολογίζονται, για την καλύτερη επανάληψη κάθε μεθόδου (και επιπλέον από μία επανάληψη για τα μοντέλα Naive και Taylor).

Methods	Naive (k=9)	SES (k=5)	Damped (k=5)	Holt (k=7)	Theta (k=5)	Naive (k=5)	Taylor (k=8)	Taylor (k=5)
Minimum MAPE	2,075%	1,121%	0,952%	1,257%	1,853%	2,225%	1,113%	1,358%
Maximum MAPE	4,958%	5,886%	5,824%	6,025%	5,810%	5,341%	6,425%	7,159%
Difference	2,883%	4,764%	4,872%	4,769%	3,958%	3,116%	5,312%	5,802%

Πίνακας 7. Ελάχιστο και μέγιστο σφάλμα MAPE των μεθόδων και η διακύμανσή του (ΜΟ των ωριαίων σφαλμάτων για όλο το 2011)

Μέχρι στιγμής παρουσιάζονται τα σφάλματα που προκύπτουν από τις εφαρμοζόμενες μεθόδους πρόβλεψης επί ολόκληρης της προβλεπόμενης χρονοσειράς, είτε βάση της χρονικής περιόδου εφαρμογής της κυλιόμενης πρόβλεψης σε εβδομάδες, είτε σε ημερήσια βάση για τους οκτώ μήνες του 2011, στους οποίους εξετάζεται η απόδοση των μοντέλων. Ακόμη, παρουσιάστηκαν τα σφάλματα για όλη τη διαθέσιμη χρονοσειρά του 2011 για κάθε ώρα της ημέρας (εικόνα 16). Τα σφάλματα για τόσο μεγάλο πλήθος δεδομένων –αυθυμνηθούμε ότι πρόκειται για 5832 παρατηρήσεις- είναι λογικό να μη λαμβάνουν ιδιαίτερα μικρές τιμές. Για να αντιληφθεί κανείς ωστόσο την ακρίβεια των μοντέλων πρόβλεψης σε ημερήσιο επίπεδο, υπολογίζεται για κάθε μέθοδο πρόβλεψης και για κάθε επανάληψή της, το ποσοστό των σφαλμάτων των ωριαίων προβλέψεων (MAPE και sMAPE) με τιμή μικρότερη του 3%. Τα αποτελέσματα της διαδικασίας αυτής βρίσκονται στον πίνακα 8.

Percentage of hourly forecasts with error < 3%						
Methods	Naive		SES		Holt	
Type of Errors	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
Worst case	42,970%	43,021%	33,573%	33,471%	30,247%	30,264%
Best case	63,872%	63,786%	59,294%	59,191%	58,539%	58,471%
Average percentage	50,733%	50,736%	47,041%	46,934%	43,839%	43,736%
Methods	Damped		Theta		Taylor	
Type of Errors	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
Worst case	35,168%	35,168%	30,881%	30,830%	41,615%	41,598%
Best case	59,002%	59,002%	58,693%	58,453%	54,252%	54,338%
Average percentage	47,416%	47,416%	44,940%	44,879%	47,205%	47,206%

Πίνακας 8. Ποσοστό πλήθους σφαλμάτων (MAPE, sMAPE) μικρότερο του 3%.

Το πρώτο από τα τρία ποσοστά είναι το ελάχιστο που παρατηρείται για κάθε μέθοδο, το δεύτερο είναι το μέγιστο και το τρίτο σφάλμα εκφράζει το μέσο όρο των ποσοστών που

μετρήθηκαν σε κάθε μέθοδο. Τα σφάλματα αυτά για τις μεθόδους Naive, SES, Damped, Holt και Theta υπολογίζονται για όλες τις επαναλήψεις που έχουν εκτελεσθεί, τόσο κατά το στάδιο της ενσωμάτωσης ημερήσιας εποχικότητας, όσο και στο στάδιο ενσωμάτωσης της εβδομαδιαίας εποχικότητας. Όπως έχει ήδη σχολιαστεί, οι προβλέψεις κατά το δεύτερο στάδιο (με εβδομαδιαία εποχικότητα) είναι πολύ καλύτερες από εκείνες του πρώτου σταδίου. Για αυτό το λόγο υπάρχει μεγάλη διαφορά ανάμεσα στις τιμές ελάχιστου και μέγιστου ποσοστού, αλλά και αρκετά μικρός μέσος όρος τους στην τρίτη γραμμή (συγκριτικά με το μέγιστο ποσοστό).

Ο πίνακας αυτός οδηγεί στα παρακάτω συμπεράσματα. Τόσο στην περίπτωση του ελάχιστου και του μέγιστου ποσοστού, όσο και στην περίπτωση του μέσου όρου του ποσοστού σφαλμάτων μικρότερων του 3% η απλοϊκή μέθοδος έχει το μεγαλύτερο ποσοστό ημερήσιων σφαλμάτων MAPE και sMAPE κάτω του 3%. Όμοια, η μέθοδοι Theta classic και Damped κατατάσσονται Τρίτη και τέταρτη αντίστοιχα σε όλες τις περιπτώσεις. Οι υπόλοιπες τρεις μέθοδοι εναλλάσσουν θέσεις για διαφορετικό υπολογισμό του ποσοστού. Αυτό που μας προκαλεί εντύπωση είναι ότι το προσαρμοσμένο μοντέλο του Taylor έχει ιδιαίτερα μεγάλο minimum ποσοστό ημερησίων σφαλμάτων μικρότερων του 3% συγκριτικά με τις υπόλοιπες μεθόδους (εκτός βέβαια από τη Naive). Αιτία για αυτή την παρατήρηση είναι ότι οι πιο απλές μέθοδοι έχουν δύο κατηγορίες σφαλμάτων, για ημερήσια και για εβδομαδιαία εποχικότητα ξεχωριστά, όπως σημειώθηκε νωρίτερα. Έτσι, το ποσοστό min percentage (%) των σφαλμάτων σε ημερήσια βάση με τιμή κάτω του 3% είναι σαφώς μικρότερο από τη μέθοδο Holt-Winters διπλής εποχικότητας, καθώς τα αντίστοιχα σφάλματά της «βελτιώνονται» από την επίδραση της εβδομαδιαίας εποχικότητας.

Με βάση όσα παρουσιάζονται στο κεφάλαιο αυτό, συμπεραίνουμε την ανάγκη για καλύτερη προσαρμογή των μοντέλων πρόβλεψης που ήδη υπάρχουν. Η ενσωμάτωση της ημερήσιας εποχικότητας στα μοντέλα απαιτεί περισσότερη προσοχή και μελέτη, καθώς τα αποτελέσματα που παράγονται από αυτήν ξεχωριστά αλλά και από συνδυασμό της με την εβδομαδιαία εποχικότητα οδηγούν σε μεγαλύτερα σφάλματα. Για τη μέθοδο του rolling forecast δεν απαιτείται ιδιαίτερα μεγάλο χρονικό διάστημα δεδομένων ώστε να ελαχιστοποιηθεί το σφάλμα κάθε μεθόδου. Το μικρότερο διάστημα που χρειάστηκε στην ως τώρα μελέτη είναι πέντε (5) εβδομάδες και το μεγαλύτερο εννέα (9) εβδομάδες (σύμφωνα με την αξιολόγηση του MAPE), μικρά διαστήματα και τα δύο, αν αναλογιστεί κανείς ότι τα διαθέσιμα φανερά στοιχεία αντιστοιχούν σε λίγο περισσότερες από 104 εβδομάδες στην αρχή του 2011. Ο λόγος για το γεγονός αυτό θεωρούμε ότι είναι η μεγάλη διαφοροποίηση των δεδομένων με το πέρασμα των μηνών και κατ' επέκταση των εποχών και ετών, οπότε η χρήση πρόσφατων σχετικά δεδομένων οδηγεί σε καλύτερες προβλέψεις, αφού περιγράφουν ακριβέστερα τη συμπεριφορά της χρονοσειράς για την τρέχουσα περίοδο που μελετάται.

Ως τώρα εξετάζεται η απόδοση των μοντέλων πρόβλεψης ζήτησης φορτίου μόνο με ενσωμάτωση της εποχικής συμπεριφοράς της χρονοσειράς που εξετάζεται. Βέβαια, εφαρμόζεται σε κάθε βήμα της διαδικασίας πρόβλεψης και για κάθε νέα εικοσιτετράωρη πρόβλεψη η μέθοδος της κύλισης των δεδομένων, όπως σε όλη την εργασία άλλωστε (σε όλα τα κεφάλαια 5, 6 και 7). Ωστόσο, η μελέτη της παρούσας εργασίας δε σταματάει εδώ, αλλά ακολουθεί το δεύτερο μέρος της, όπου επιχειρείται η ενσωμάτωση της επίδρασης των ειδικών ημερών στα μοντέλα που έχουν εξεταστεί ως τώρα.

Σημείωση: Στην αρχή του κεφαλαίου σημειώνεται στους πίνακες των σφαλμάτων ότι οι μέθοδοι Naive, SES, Damped, Holt και Theta classic δέχονται αποεποχικοποίηση των δεδομένων τους πριν την παραγωγή κάθε πρόβλεψης και ακόμη, εφαρμόζεται καθ' όλη τη διαδικασία η μέθοδος της κύλισης των δεδομένων. Και στην προσαρμογή του Taylor επί της μεθόδου Holt-Winters για συμπερίληψη και των δύο εποχικότητων εφαρμόζεται η μέθοδος της κύλισης των δεδομένων (rolling forecasting). Επομένως, παρότι στη συνέχεια του κεφαλαίου 6 δεν αναγράφεται κάθε φορά δίπλα στις μεθόδους το “deseason.&rolling” ή σκέτο το “rolling”, αυτό δε σημαίνει ότι πρόκειται για την απλή εφαρμογή των μεθόδων, αλλά ακολουθείται η διαδικασία που έχει αναλυθεί στην παράγραφο 5.2.

ΚΕΦΑΛΑΙΟ 7: ΕΠΙΔΡΑΣΗ ΕΙΔΙΚΩΝ ΗΜΕΡΩΝ

7.1 Επιλογή ειδικών ημερών

Παρατηρήσαμε στη μελέτη μας ότι τα σφάλματα σε μεμονωμένες ημέρες είναι ιδιαίτερα υψηλά. Με πιο προσεκτική ματιά βλέπουμε ότι αυτό ισχύει κυρίως σε ειδικές ημέρες της χρονοσειράς. Χαρακτηριστικά, οι ημέρες ειδικών αργιών και απεργιών της χώρας επηρεάζουν έντονα τη μεταβολή του φορτίου ζήτησης για τις ημέρες αυτές αλλά και για τις υπόλοιπες γύρω τους συγκριτικά με την αναμενόμενη τιμή που προκύπτει από απλή παρατήρηση των αντίστοιχων ωρών ίδιων ημερών γύρω από τις εξεταζόμενες ημέρες. Ανάλογα την ημέρα της εβδομάδας μάλιστα (καθημερινή, Σάββατο, Κυριακή), η επιρροή επί των ημερών που την περιβάλλουν κυμαίνεται από λίγες ώρες ως και αρκετές ημέρες πριν και μετά την ειδική ημέρα. Εστιάζουμε πλέον το ενδιαφέρον μας στην καταγραφή των ειδικών ημερών για την χρονική περίοδο που διαθέτουμε δεδομένα της χρονοσειράς και στη συνέχεια στην ενσωμάτωση της επίδρασής τους στις προβλέψεις που παράγουμε με τα υπάρχοντα μοντέλα.

Στις περισσότερες μελέτες βραχυπρόθεσμης πρόβλεψης ζήτησης φορτίου ηλεκτρικής ενέργειας, ενώ αναγνωρίζεται η επιρροή των ειδικών ημερών (αργίες, απεργίες, κλπ), ωστόσο δε λαμβάνονται υπόψη στη μοντελοποίηση των μεθόδων που ερευνούν. Είθισται να θεωρούνται κενές τιμές σε περιπτώσεις που αυτή η αντιμετώπιση δεν επηρεάζει αρνητικά το σύστημα, είτε εξομαλύνονται τα στοιχεία τους στην αρχική χρονοσειρά, χωρίς να πειραχτούν οι περιοδικότητες των δεδομένων, ώστε να μην επηρεάζεται η μεθοδολογία από τις διακυμάνσεις που προκαλούν. Θεωρούμε ότι αυτή η αντιμετώπιση δεν είναι κατάλληλη, εφόσον, όπως θα αποδειχτεί και στη συνέχεια, η επίδραση των ειδικών ημερών επί της πρόβλεψης ζήτησης φορτίου είναι ιδιαίτερα σημαντική και η αξιοποίησή της στα μοντέλα πρόβλεψης πολύ χρήσιμη. Έτσι, έπειτα από την απλή μελέτη των μοντέλων που υλοποιήσαμε σα να μην υπάρχουν ειδικές ημέρες (special days), θα προχωρήσουμε σε όσο το δυνατό πιο παραγωγική ένταξή τους στις προβλέψεις.

Το πρώτο βήμα μας είναι η καταγραφή των ειδικών ημερών, όπως εθνικές αργίες, θρησκευτικές εορτές, ειδικά εορταστικά προγράμματα στα καταστήματα και απεργίες για τη χρονική περίοδο στην οποία διαθέτουμε δεδομένα ζήτησης (και για τις 973 ημέρες, από 1/1/2009 ως και 31/8/2011). Με βάση την αναζήτηση που κάναμε, προέκυψαν οι παρακάτω ειδικές ημέρες για τη συγκεκριμένη χρονική περίοδο:

-Εθνικές αργίες (bank holidays), παρουσιάζονται στους πίνακες 9 και 10, καθώς στον πρώτο πίνακα είναι οι επίσημες αργίες και στον δεύτερο πίνακα έχουμε προσθέσει εμείς αργίες τις οποίες θεωρούμε σχεδόν εξίσου σημαντικές, καθώς έχουμε παρατηρήσει ότι επηρεάζουν τις τιμές ζήτησης των ημερών γύρω τους.

Αργίες	2009	2010	2011
<i>Πρωτοχρονιά (σταθερή)</i>	Πέμπτη, 1/1	Παρασκευή, 1/1	Σάββατο, 1/1
<i>Θεοφάνεια (σταθερή)</i>	Τρίτη, 6/1	Τετάρτη, 6/1	Πέμπτη, 6/1
<i>Κ. Δευτέρα (κινητή)</i>	Δευτέρα, 2/3	Δευτέρα, 15/2	Δευτέρα, 7/2
<i>25 Μάρτη (σταθερή)</i>	Τετάρτη, 25/3	Πέμπτη, 25/3	Παρασκευή, 25/3
<i>Μ. Παρασκευή (κινητή)</i>	Παρασκευή, 17/4	Παρασκευή, 2/4	Παρασκευή, 22/4
<i>Κ. Πάσχα (κινητή)</i>	Κυριακή, 19/4	Κυριακή, 4/4	Κυριακή, 24/4
<i>Δ. Πάσχα (κινητή)</i>	Δευτέρα, 20/4	Δευτέρα, 5/4	Δευτέρα, 25/4
<i>Πρωτομαγιά (σταθερή)</i>	Παρασκευή, 1/5	Σάββατο, 1/5	Κυριακή, 1/5
<i>Πεντηκοστή (κινητή)</i>	Κυριακή, 7/6	Κυριακή, 23/5	Κυριακή, 12/6
<i>Αγ. Πνεύματος (κινητή)</i>	Δευτέρα, 8/6	Δευτέρα, 24/5	Δευτέρα, 13/6
<i>15 Αυγούστου (σταθερή)</i>	Σάββατο, 15/8	Κυριακή, 15/8	Δευτέρα, 15/8
<i>28 Οκτώβρη (σταθερή)</i>	Τετάρτη, 28/10	Πέμπτη, 28/10	----
<i>Χριστούγεννα (σταθερή)</i>	Παρασκευή, 25/12	Σάββατο, 25/12	----
<i>26 Δεκέμβρη (σταθερή)</i>	Σάββατο, 26/12	Κυριακή, 26/12	----

Πίνακας 9. Επίσημες αργίες για τα έτη 2009, 2010, 2011

Αργίες	2009	2010	2011
<i>Κυριακή Βαΐων (κινητή)</i>	Κυριακή, 12/4	Κυριακή, 29/3	Κυριακή, 17/4
<i>Μ. Πέμπτη (κινητή)</i>	Πέμπτη, 16/4	Πέμπτη, 1/4	Πέμπτη, 21/4
<i>Μ. Σάββατο (κινητή)</i>	Σάββατο, 18/4	Σάββατο, 3/4	Σάββατο, 23/4
<i>Τ. Πάσχα (κινητή)</i>	Τρίτη, 21/4	Τρίτη, 5/4	Τρίτη, 26/4
<i>17 Νοέμβρη (σταθερή)</i>	Τρίτη, 17/11	Τετάρτη, 17/11	----

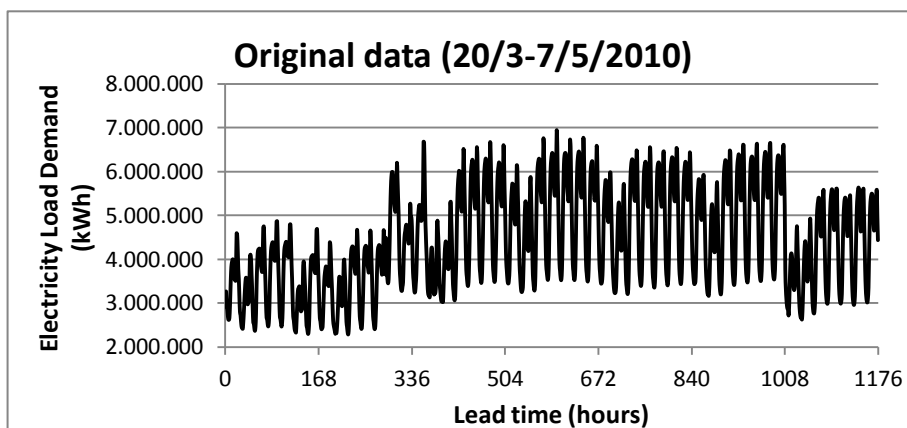
Πίνακας 10. Μη επίσημες αργίες για τα έτη 2009, 2010, 2011

-Απεργίες (strikes), παρουσιάζονται στον πίνακα 11. Τα στοιχεία για τις απεργίες στα τρία έτη που μας ενδιαφέρουν τα αντλήσαμε από τα αρχεία της ΓΣΕΕ (Γενική Συνομοσπονδία Εργατών Ελλάδας). Στην εικόνα 11 δεν έχουμε σημειώσει την ημέρα της Πρωτομαγιάς, εφόσον ήδη συμπεριλαμβάνεται στις επίσημες αργίες. Ακόμη, οι απεργίες που έχουμε συμπεριλάβει είναι μόνο οι εικοσιτετράωρες πανελλαδικές πανεργατικές απεργίες και όχι οι μικρότερης διάρκειας στάσεις εργασίας, καθώς οι δεύτερες δεν επηρεάζουν ιδιαίτερα τη ζήτηση φορτίου ηλεκτρισμού.

Απεργίες	2009	2010	2011
24ωρες πανελλαδικές πανεργατικές	Πέμπτη, 2/4	Τετάρτη, 24//2	Τετάρτη, 23/2
	----	Πέμπτη, 11/3	Τετάρτη, 11/5
	----	Τετάρτη, 5/5	Κυριακή, 15/5
	----	Πέμπτη, 20/5	Τρίτη, 28/6
	----	Τρίτη, 29/6	Τετάρτη, 29/6
	----	Πέμπτη, 8/7	----
	----	Τετάρτη, 15/12	----

Πίνακας 11. Απεργίες για τα έτη 2009, 2010, 2011

Πλέον, έχοντας στη διάθεσή μας τις ημερομηνίες των αργιών και απεργιών για τα έτη που μελετάμε, μπορούμε να παρατηρήσουμε καλύτερα την επίδρασή τους στη ζήτηση φορτίου και τις αποκλίσεις των προβλέψεων για τις αντίστοιχες ημερομηνίες. Ενδεικτικά παραθέτουμε (εικόνα 17) τη ζήτηση ηλεκτρικής ενέργειας για το χρονικό διάστημα από τις 20 Μαρτίου 2010 έως και τις 7 Μαΐου 2010. Επιλέχθηκε το συγκεκριμένο διάστημα, επειδή περιλαμβάνει πέντε από τις επίσημες αργίες, τέσσερις μη επίσημες αργίες και μία απεργία από αυτές που σημειώθηκαν στους πίνακες των προηγούμενων τριών εικόνων.



Εικόνα 17. Πραγματική ζήτηση φορτίου για τη χρονική περίοδο

Εύκολα διακρίνει κανείς τη μεταβολή του επιπέδου ζήτησης φορτίου για τις ειδικές ημέρες, καθώς είναι πολύ έντονη. Ειδικά στο αρχικό τμήμα της καμπύλης, οπότε και είναι όλες οι ημέρες των αργιών (επίσημων και μη) εκτός από την Πρωτομαγιά, η πτώση του επιπέδου ζήτησης είναι πολύ μεγάλη, φτάνοντας και τις $2 \cdot 10^6$ (kWh). Επιβεβαιώνεται λοιπόν η σημασία της ενσωμάτωσης των ειδικών ημερών κατά την πρόβλεψη ζήτησης ηλεκτρικού φορτίου, με την οποία θα ασχοληθούμε αμέσως στην παράγραφο 7.2.

7.2 Εισαγωγή ειδικών ημερών στα υπάρχοντα μοντέλα πρόβλεψης

Αποφασίζουμε να ελέγξουμε την επίδραση των ειδικών ημερών με περισσότερους από έναν τρόπους. Για κάθε ώρα της εκάστοτε ειδικής ημέρας ξεκινάμε υπολογίζοντας (βάσει των δεδομένων της πραγματικής χρονοσειράς) την ποσοστιαία μεταβολή επί του φορτίου ζήτησης της προηγούμενης ημέρας και της προηγούμενης εβδομάδας, που προκαλείται από

τη συγκεκριμένη αργία ή απεργία στη. Έχοντας για όλες τις ειδικές ημέρες αυτές τις δύο διαφορετικές κατηγορίες ποσοστών μεταβολής της ζήτησης φορτίου επί των δεδομένων της αρχικής χρονοσειράς, συνεχίζουμε με την ενσωμάτωσή τους στις τελικές προβλέψεις των μοντέλων που έχουμε ήδη παράγει. Δεν επαναλαμβάνουμε την παραγωγή προβλέψεων με «τρέξιμο» των μοντέλων μας, αλλά εφαρμόζουμε τα ανωτέρω ποσοστά επί των χρονοσειρών που έχουμε από το πρώτο στάδιο προβλέψεων. Οι συνδυασμοί των ποσοστιαίων μεταβολών που μελετάμε είναι οι εξής: προηγούμενη ημέρα, προηγούμενη εβδομάδα και προηγούμενη ημέρα και εβδομάδα.

Παραθέτουμε τους τύπους που χρησιμοποιούμε για τον υπολογισμό της τελικής πρόβλεψης στις ειδικές ημέρες, για καλύτερη κατανόηση της μεθοδολογίας αυτής, με τις προβλέψεις να αναφέρονται σε ωριαία δεδομένα για την εκάστοτε ειδική ημέρα (αργία ή απεργία). Η επεξήγηση των συμβολισμών που χρησιμοποιούνται στους ακόλουθους τύπους γίνεται ακριβώς μετά την παράθεση των τύπων.

$$F_{t,previous\ day} = (1 + P_{t,previous\ day}) \cdot D_{t,previous\ day}$$

$$F_{t,previous\ week} = (1 + P_{t,previous\ week}) \cdot D_{t,previous\ week}$$

$$F_{t,previous\ day+week} = Average(F_{t,previous\ day}, F_{t,previous\ week})$$

Οι συμβολισμοί των ανωτέρω σχέσεων παραγωγής τελικής πρόβλεψης για τις επιλεγμένες ειδικές ημέρες:

- ✓ $F_{t,previous\ day}$ = τελική ωριαία πρόβλεψη με βάση την αντίστοιχη ώρα της προηγούμενης ημέρας
- ✓ $F_{t,previous\ week}$ = τελική πρόβλεψη με βάση την αντίστοιχη ώρα της επόμενης εβδομάδας
- ✓ $F_{t,previous\ day+week}$ = τελική πρόβλεψη με βάση την αντίστοιχη ώρα της προηγούμενης ημέρας και εβδομάδας
- ✓ $P_{t,previous\ day}$ = ποσοστιαία μεταβολή της αντίστοιχης ώρας της προηγούμενης ημέρας
- ✓ $P_{t,previous\ week}$ = ποσοστιαία μεταβολή της αντίστοιχης ώρας της προηγούμενης εβδομάδας
- ✓ $D_{t,previous\ day}$ = αρχικά δεδομένα της αντίστοιχης ώρας της προηγούμενης ημέρας
- ✓ $D_{t,previous\ week}$ = αρχικά δεδομένα της αντίστοιχης ώρας της προηγούμενης εβδομάδας

Οι τελικές προβλέψεις για τις απλές ημέρες, δηλαδή για ημέρες που δεν ανήκουν στην κατηγορία των special days, παραμένουν ως έχουν από το πρώτο στάδιο πρόβλεψης. Οι ειδικές ημέρες αντιμετωπίζονται όπως μόλις αναφέρθηκε. Και το τελικό αποτέλεσμα προκύπτει ως ο συνδυασμός τους.

Στη συνέχεια θα μελετήσουμε διαφορετικούς συνδυασμούς των ειδικών ημερών, τους οποίους θα εντάξουμε στα μοντέλα πρόβλεψης που έχουμε ήδη αναλύσει, ώστε να αποφανθούμε το βέλτιστο συνδυασμό. Η μέθοδος όμως με την οποία συμπεριλαμβάνονται οι συνδυασμοί αυτοί παραμένει η ίδια, όπως περιγράφεται στις εξισώσεις παραπάνω. Για την ακρίβεια, οι συνδυασμοί που μελετάμε είναι οι εξής:

7.2.1 Ενσωμάτωση μόνο των αργιών

❖ Στην πρώτη περίπτωση, θεωρούμε μόνο τις επίσημες αργίες που παρουσιάζονται στον πίνακα της εικόνας 17, δηλαδή τις εξής: Πρωτοχρονιά, Θεοφάνεια, Καθαρή Δευτέρα, 25^η Μαρτίου, Μεγάλη Παρασκευή, Κυριακή του Πάσχα, Δευτέρα του Πάσχα, Πρωτομαγιά, Πεντηκοστή, Αγίου Πνεύματος και 15^η Αυγούστου (η κοίμηση της Θεοτόκου). Είναι εμφανές ότι οι ειδικές ημέρες μετά την Κοίμηση της Θεοτόκου δε χρησιμοποιούνται, επειδή για το έτος 2011 δεν υπάρχουν δεδομένα μετά τον Αύγουστο. Καταλήγουμε επομένως στις επίσημες αργίες που αναφέρθηκαν στην αρχή της παραγράφου και ο υπολογισμός των ζητούμενων ποσοστών (προηγούμενης-επόμενης ημέρας και εβδομάδας) γίνεται με το μέσο όρο των ποσοσטיών μεταβολών των εκάστοτε αργιών για τα έτη 2009 και 2010. Υπολογίζεται για παράδειγμα η ποσοστιαία μεταβολή από την προηγούμενη ημέρα για την ίδια αργία και για τα δύο έτη (2009, 2010), για κάθε ώρα της ημέρας και στο τέλος ο μέσος όρος τους. Το ίδιο γίνεται και για τις ποσοστιαίες μεταβολές από προηγούμενη εβδομάδα, κλπ. Στη συνέχεια, τα ποσοστά που έχουν υπολογιστεί ως μέσοι όροι εφαρμόζονται σε κάθε ειδική ημέρα του έτους 2011 (διαφορετικά ποσοστά για τις διαφορετικές αργίες) και με τη διαδικασία που προαναφέρθηκε παράγονται οι τελικές προβλέψεις ζήτησης φορτίου για τις επίσημες αργίες του 2011.

❖ Σε αυτή την περίπτωση χρησιμοποιούνται -εκτός από τις δέκα επίσημες αργίες της προηγούμενης περίπτωσης- και τέσσερις ακόμη από τις μη επίσημες αργίες που παραθέσαμε στον πίνακα της εικόνας 24, η Κυριακή των Βαΐων, η Μεγάλη Πέμπτη, το Μεγάλο Σάββατο και η Τρίτη του Πάσχα. Η 17^η Νοέμβρη μένει εκτός λίστας, για το λόγο ότι οι προβλέψεις του έτους 2011 φτάνουν μόλις μέχρι και την 31^η Αυγούστου. Και εδώ ο υπολογισμός των τεσσάρων κατηγοριών ποσοστών γίνεται με τη μέθοδο της προηγούμενης περίπτωσης, με το μέσο όρο δηλαδή των ποσοστών για την ίδια αργία, για τα δύο έτη (2009 και 2010).

7.2.2 Ενσωμάτωση μόνο των απεργιών

❖ Στην πρώτη περίπτωση που μελετάμε στις απεργίες, θεωρούμε όλες τις απεργίες των ετών 2009 και 2010. Για όλες τις ημέρες με απεργία (οι οκτώ ημέρες που φαίνονται στον πίνακα της εικόνας 25) υπολογίζονται τα ζητούμενα ποσοστά (προηγούμενης ημέρας ή εβδομάδας και επόμενης ημέρας ή εβδομάδας) και στη συνέχεια καταλήγουμε στα τελικά ποσοστά που χρησιμοποιούνται και προκύπτουν ως ο μέσος όρος των ποσοστών όλων των απεργιών. Ο μέσος όρος αυτός εφαρμόζεται σε όλες τις ημέρες με απεργίες του έτους 2011 (πέντε ημέρες με βάση την εικόνα 25) και έτσι λαμβάνονται οι τελικές προβλέψεις ζήτησης φορτίου για αυτές τις ημέρες. Οι υπόλοιπες ημέρες, στις οποίες δεν υπάρχουν απεργίες, διατηρούν ως τελικές προβλέψεις τις προβλέψεις του πρώτου σταδίου.

❖ Επόμενη περίπτωση είναι ο υπολογισμός του μέσου όρου των ποσοστών για ίδιες ημέρες της εβδομάδας στις οποίες έχουμε εμφάνιση απεργιών. Συγκεκριμένα, υπολογίζουμε το μέσο όρο των ποσοστών για Τρίτη, Τετάρτη και Πέμπτη, καθώς σε αυτές τις ημέρες παρατηρήθηκαν απεργίες για τα έτη 2009 και 2010. Για παράδειγμα, για τον υπολογισμό του ποσοστού για την Τετάρτη βρίσκουμε το Μέσο Όρο των ποσοστών των ημερών 24/2, 5/5 και 15/12 του 2010. Να τονίσουμε στο σημείο αυτό, ότι στην περίπτωση των απεργιών συνυπολογίζουμε όλες τις απεργίες που διαθέτουμε, ακόμη και αν αυτές συμβαίνουν έπειτα από την 31 Αυγούστου. Ο λόγος για τον συνυπολογισμό τους είναι η πολύ μικρή διάθεση δεδομένων για απεργίες, οπότε και παρατηρήσαμε ότι δεν επηρεάζει αρνητικά η συμπερίληψη όλων των απεργιών. Αντίθετα οδηγεί σε μικρότερα σφάλματα η συμπερίληψη της απεργίας της 15 Δεκεμβρίου του 2010. Παρατηρούμε ότι το έτος 2011 γίνεται απεργία

στις 15 Μαΐου και ημέρα Κυριακή. Εφόσον για την ημέρα αυτή της εβδομάδας δε διαθέτουμε στοιχεία από τα προηγούμενα έτη, δεν εφαρμόζουμε κάποιο ποσοστό επί των δεδομένων της και θεωρούμε ως τελική πρόβλεψη την ίδια που είχαμε υπολογίσει στο πρώτο στάδιο (σα να μην είναι ειδική ημέρα).

❖ Στην τελευταία περίπτωση στην αντιμετώπιση των απεργιών δεν υπολογίζουμε μέσο όρο. Έχοντας τις ποσοστιαίες μεταβολές για τις ημέρες απεργιών των ετών 2009 και 2010, συγκρίνουμε τις αντίστοιχες ημέρες και μάλιστα αντίστοιχων μηνών, με εκείνες του 2011. Παρατηρούμε ότι είναι παραπλήσιες, για παράδειγμα η απεργία την Τρίτη στις 28 Ιουνίου του 2011 μπορεί να αντιστοιχιστεί με την απεργία την Τρίτη στις 29 Ιουνίου του 2010. Με τον τρόπο αυτό, τοποθετούμε τα ποσοστά από τις ημέρες που ήδη έχουμε από τα προηγούμενα έτη απευθείας στις απεργίες του 2011, χωρίς να προβούμε σε περαιτέρω πράξεις. Και πάλι, η Κυριακή 15 Μαΐου θεωρείται ως απλή ημέρα χωρίς απεργία.

7.2.3 Ενσωμάτωση συνδυασμού αργιών και απεργιών

❖ Στο συνδυασμό αυτό υπάρχουν αρκετές περιπτώσεις προς μελέτη όπως είναι φανερό. Ασχολούμαστε ωστόσο κυρίως με δύο, τη συμπερίληψη των επίσημων αργιών και του μέσου όρου όλων των απεργιών (οι πρώτες περιπτώσεις από τα Α) και Β)) και τη συμπερίληψη όλων των αργιών (επίσημων και μη) και της δεύτερης περίπτωσης από το Β. Μελέτη έγινε και για τους υπόλοιπους συνδυασμούς, όπως επίσημες αργίες με τη δεύτερη περίπτωση των απεργιών και όλες οι αργίες με την πρώτη περίπτωση των απεργιών (ο μέσος όρος όλων των απεργιών για τα έτη 2009 και 2010). Δεν έχουμε περίπτωση που να συμπίπτουν απεργίες με αργίες, επομένως εφαρμόζονται οι εκάστοτε ποσοστιαίες μεταβολές στις αντίστοιχες ειδικές ημέρες (αργίες και απεργίες) και οι υπόλοιπες ημέρες διατηρούν τις προβλέψεις που υπήρχαν από το πρώτο μέρος της πρόβλεψης.

7.3 Αποτελέσματα νέων προβλέψεων

Στη συνέχεια παρουσιάζονται τα σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων που μελετήθηκαν προηγουμένως, έχοντας πλέον ενσωματωμένη την επίδραση των ειδικών ημερών, είτε χωριστά των αργιών (επίσημων και μη) ή χωριστά των απεργιών, είτε και με συνδυασμό των δύο ειδών ειδικών ημερών. Ο υπολογισμός των σφαλμάτων έγινε επί όλων των προβλεπόμενων τιμών ζήτησης για το 2011, για τις δύο εποχικότητες και για όλες τις εβδομάδες εφαρμογής rolling forecast.

7.3.1 Αποτελέσματα με επίδραση επίσημων αργιών

Ακολουθούν τα αποτελέσματα των σφαλμάτων που υπολογίζονται έπειτα από την ενσωμάτωση της επίδρασης των αργιών στις προβλέψεις που έχουν παραχθεί από τα κεφάλαια 5 και 6.

7.3.1.1 Επίσημες αργίες

Στους πίνακες των εικόνων 12 και 13 παρουσιάζεται τμήμα των σφαλμάτων για την τελική πρόβλεψη, στην περίπτωση ενσωμάτωσης μόνο των επίσημων αργιών (πρώτη περίπτωση της υποπαραγράφου 7.2.1). Τα αποτελέσματα για όλες τις εβδομάδες κυλιόμενης πρόβλεψης βρίσκονται στο Appendix (εικόνες A5 έως και A8). Στην περίπτωση αυτή, στους πίνακες 12 και 13 παρουσιάζονται τα σφάλματα μόνο για την περίπτωση εισαγωγής της επίδρασης της ποσοστιαίας μεταβολής με βάση την προηγούμενη ημέρα. Αυτό συμβαίνει

επειδή από όλους τους συνδυασμούς που μελετήθηκαν, αυτός αποδείχτηκε ο καλύτερος, έχοντας τα μικρότερα σφάλματα. Και στους επόμενους πίνακες, από πίνακα 14 έως και πίνακα 27, όπου θα παρουσιάζονται τα σφάλματα για τις υπόλοιπες κατηγορίες εισαγωγής των ειδικών ημερών (μόνο απεργίες, όλες οι αργίες, συνδυασμός των), θα παρατίθενται τα σφάλματα του καλύτερου συνδυασμού ενσωμάτωσης της ποσοστιαίας μεταβολής (βάσει προηγούμενης ημέρας ή εβδομάδας και το συνδυασμό τους). Η περίπτωση με τα βέλτιστα αποτελέσματα θα σημειώνεται στους πίνακες, δίπλα από το όνομα της κάθε μεθόδου (Naive, SES, κλπ).

		Naive previous day (official BH)		SES previous day (official BH)		Holt previous day (official BH)	
		MAPE(%)	SMAPE(%)	MAPE(%)	SMAPE(%)	MAPE(%)	SMAPE(%)
e=24	k=1	6,197	6,299	5,568	5,468	9,773	9,591
	k=2	6,197	6,299	5,522	5,451	8,559	8,465
	k=3	6,197	6,299	5,545	5,480	7,977	7,843
	k=4	6,197	6,299	5,567	5,501	7,699	7,649
	k=5	6,197	6,299	5,608	5,539	5,607	5,539
	k=6	6,197	6,299	5,672	5,598	5,667	5,593
e=168	k=4	3,493	3,543	3,363	3,365	3,384	3,387
	k=5	3,271	3,322	3,234	3,232	3,251	3,248
	k=6	3,144	3,194	3,286	3,282	3,299	3,293
	k=7	3,121	3,170	3,418	3,408	3,431	3,421
	k=8	3,131	3,175	3,523	3,509	3,534	3,518
	k=9	3,074	3,123	3,627	3,610	3,636	3,618
	k=10	3,087	3,137	3,748	3,727	3,751	3,728

Πίνακας 12. Σφάλματα MAPE και sMAPE με επίδραση των επίσημων αργιών (α)

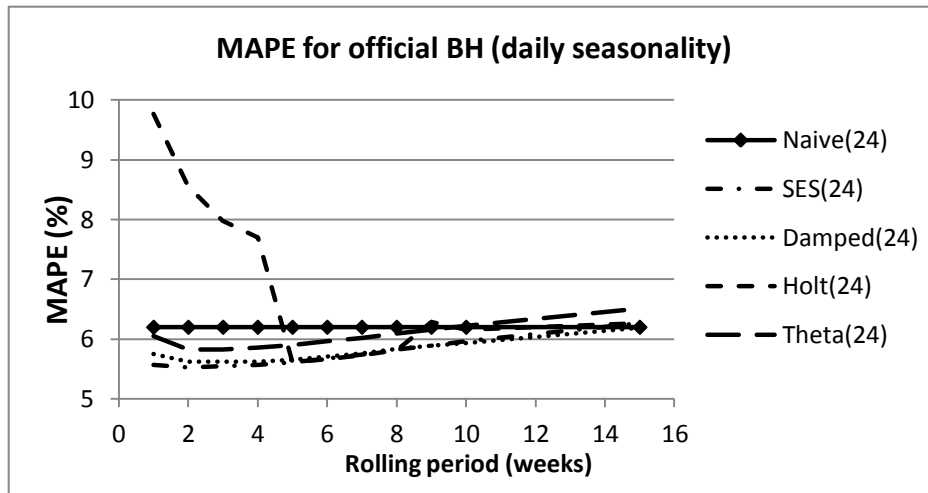
		Damped previous day (official BH)		Theta previous day (official BH)		Taylor previous day (official BH)	
		MAPE(%)	SMAPE(%)	MAPE(%)	SMAPE(%)	MAPE(%)	SMAPE(%)
e=24	k=1	5,751	5,631	6,047	5,996		
	k=2	5,623	5,534	5,823	5,796		
	k=3	5,619	5,535	5,829	5,809		
	k=4	5,622	5,537	5,856	5,839		
	k=5	5,653	5,564	5,899	5,883		
	k=6	5,711	5,614	5,959	5,944		
e=168	k=4	3,429	3,426	3,407	3,429	4,173	4,217
	k=5	3,296	3,288	3,287	3,307	4,422	4,537
	k=6	3,337	3,326	3,326	3,347	4,779	4,948
	k=7	3,453	3,436	3,452	3,472	4,431	4,438
	k=8	3,549	3,526	3,548	3,567	4,133	4,180
	k=9	3,653	3,625	3,644	3,664	4,739	4,774
	k=10	3,753	3,721	3,760	3,781	5,243	5,540

Πίνακας 13. Σφάλματα MAPE και sMAPE με επίδραση των επίσημων αργιών (β)

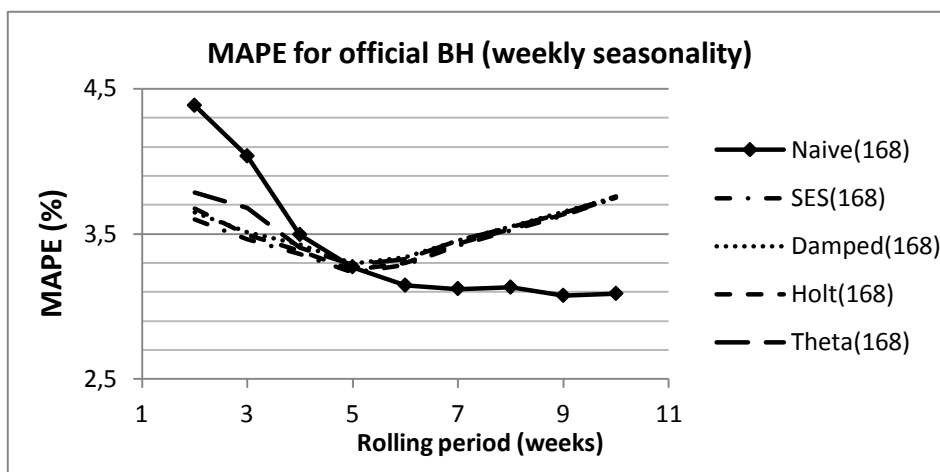
Να υπενθυμίσουμε ότι συμβολίζουμε με e το μήκος του κεντρικού κινητού μέσου όρου (ΚΚΜΟ) που υπολογίζεται στις στατιστικές μεθόδους για την ενσωμάτωση των εποχικότητων

χωριστά. Έτσι είναι $e=14$ (ημερήσια) ή $e=168$ (εβδομαδιαία). Ακόμη, με k συμβολίζεται το χρονικό διάστημα εφαρμογής της κύλισης στα δεδομένα, σε εβδομάδες. Εδώ δεν έχει γίνει νέα παραγωγή προβλέψεων όπως είπαμε και προηγουμένως, αλλά επέμβαση στις ήδη υπάρχουσες προβλέψεις από όλες τις στατιστικές μεθόδους, με βάση τις διαφορετικές ειδικές ημέρες.

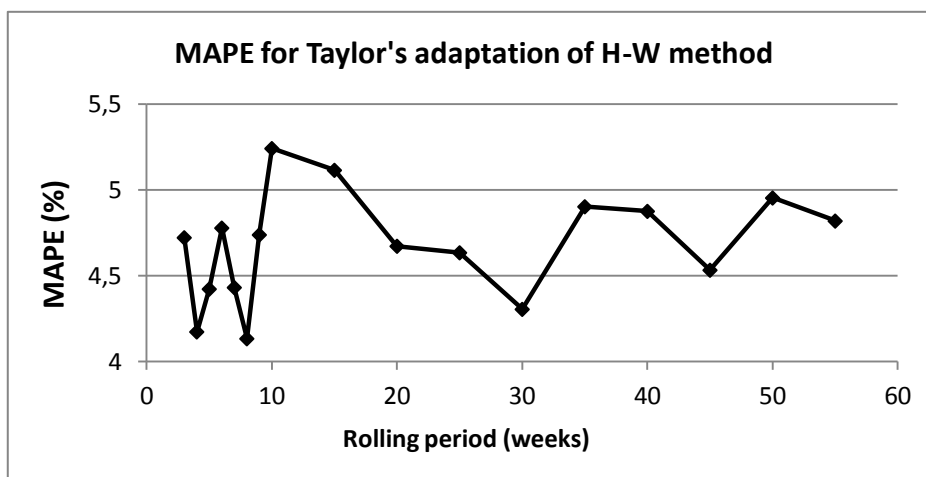
Στις εικόνες 17 ως και 19 παρατίθενται σε διαγράμματα για τα σφάλματα MAPE της μεθόδου που μελετάμε, δηλαδή με την επίδραση των επίσημων αργιών. Έτσι αποκτάμε μια καλύτερη αντίληψη των μεγεθών των σφαλμάτων.



Εικόνα 18. Σφάλμα MAPE με επίδραση των επίσημων αργιών



Εικόνα 19. Σφάλμα MAPE με επίδραση των επίσημων αργιών και εβδομαδιαία εποχικότητα



Εικόνα 20. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών

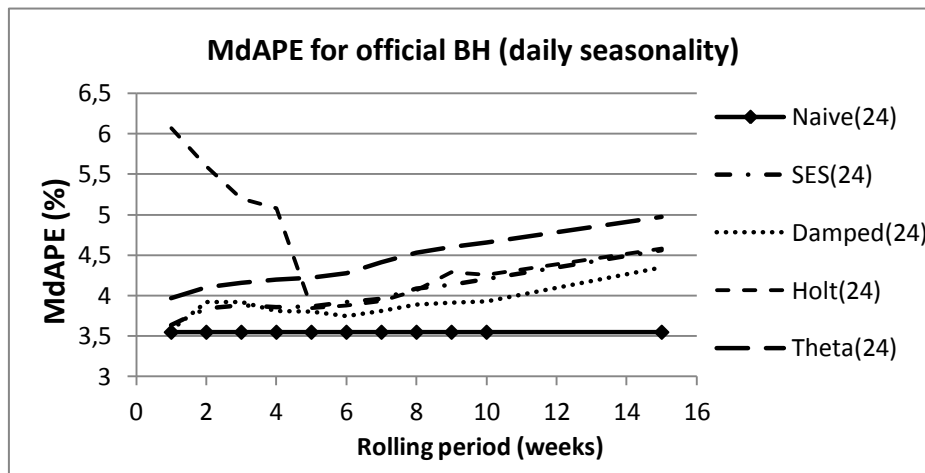
Στη συνέχεια (πίνακες 14 και 15) παρουσιάζονται τα σφάλματα MdAPE και sMdAPE αντίστοιχα με τα MAPE και sMAPE που είδαμε παραπάνω. Και έπειτα, στις εικόνες 20 έως και 22, τα διαγράμματά MdAPE με τα σφάλματα όλων των περιπτώσεων rolling forecasting για την παραγωγή πρόβλεψης.

		Naive previous day (official BH)		SES previous day (official BH)		Holt previous day (official BH)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,548	3,532	3,633	3,652	6,064	6,151
	k=2	3,548	3,532	3,842	3,878	5,603	5,675
	k=3	3,548	3,532	3,885	3,912	5,197	5,265
	k=4	3,548	3,532	3,856	3,894	5,083	5,121
	k=5	3,548	3,532	3,865	3,902	3,843	3,859
	k=6	3,548	3,532	3,919	3,937	3,881	3,890
e=168	k=4	2,352	2,347	2,259	2,261	2,253	2,259
	k=5	2,246	2,243	2,221	2,223	2,236	2,232
	k=6	2,104	2,112	2,272	2,274	2,295	2,301
	k=7	2,068	2,064	2,428	2,423	2,442	2,429
	k=8	2,070	2,071	2,523	2,522	2,525	2,529
	k=9	2,027	2,039	2,562	2,566	2,600	2,593
	k=10	2,040	2,039	2,690	2,695	2,689	2,692

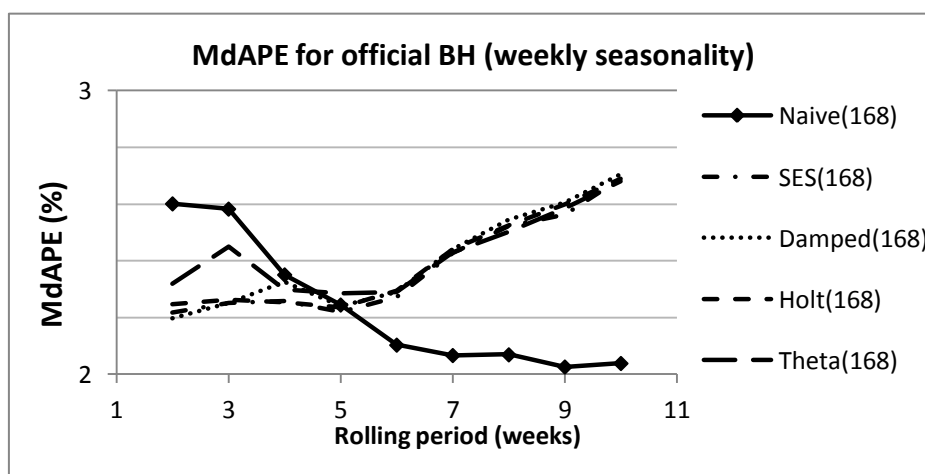
Πίνακας 14. Σφάλματα MdAPE και sMdAPE με επίδραση των επίσημων αργιών (α)

		Damped previous day (official BH)		Theta previous day (official BH)		Taylor previous day (official BH)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,574	3,591	3,966	3,994		
	k=2	3,918	3,949	4,103	4,126		
	k=3	3,925	3,939	4,160	4,191		
	k=4	3,813	3,830	4,196	4,218		
	k=5	3,801	3,831	4,225	4,237		
	k=6	3,744	3,771	4,276	4,298		
e=168	k=4	2,325	2,331	2,300	2,296	2,506	2,501
	k=5	2,241	2,239	2,287	2,282	2,663	2,660
	k=6	2,290	2,298	2,290	2,287	2,893	2,884
	k=7	2,440	2,450	2,429	2,426	2,785	2,786
	k=8	2,544	2,553	2,503	2,505	2,617	2,608
	k=9	2,605	2,601	2,585	2,598	2,913	2,911
	k=10	2,705	2,704	2,681	2,691	3,242	3,237

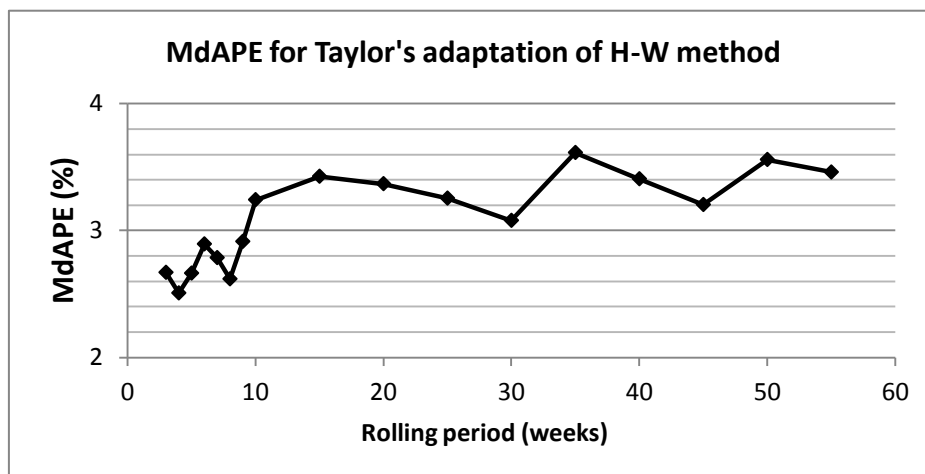
Πίνακας 15. Σφάλματα MdAPE και sMdAPE με επίδραση των επίσημων αργιών (β)



Εικόνα 21. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και ημερήσια εποχικότητα



Εικόνα 22. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και εβδομαδιαία εποχικότητα



Εικόνα 23. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών

7.3.1.2 Επίσημες και μη αργίες

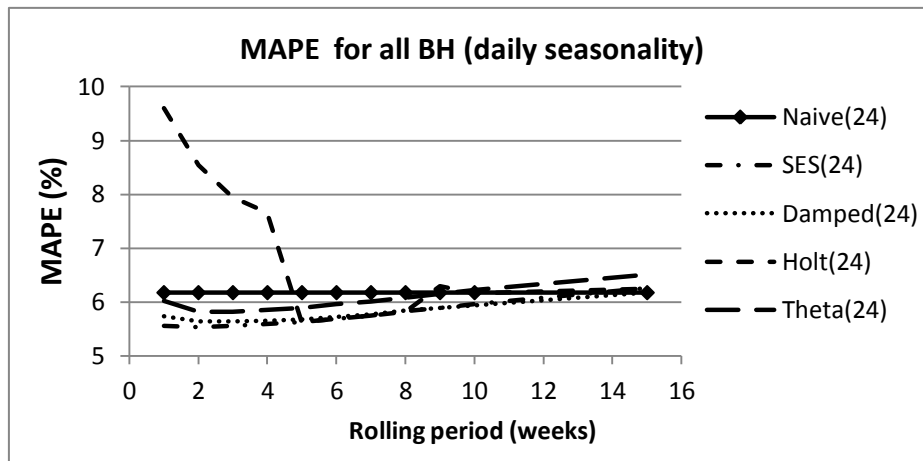
Στους πίνακες 16 και 17 παρουσιάζεται τμήμα των σφαλμάτων για την τελική πρόβλεψη, στην περίπτωση ενσωμάτωσης των επίσημων και των ανεπίσημων αργιών μαζί (δεύτερη περίπτωση της υποπαραγράφου 7.2.1). Και στα διαγράμματα των εικόνων 24, 25 και 26 εμφανίζονται τα σφάλματα MdAPE για όλα τα χρονικά διαστήματα υλοποίησης της κύλισης των δεδομένων. Τα αποτελέσματα για όλες τις εβδομάδες κυλιόμενης πρόβλεψης βρίσκονται στο Appendix (εικόνες A9 έως και A12).

		Naive previous day (all BH)		SES previous day (all BH)		Holt previous day (all BH)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	6,171	6,251	5,557	5,447	9,606	9,467
	k=2	6,171	6,251	5,534	5,455	8,540	8,448
	k=3	6,171	6,251	5,562	5,489	7,947	7,817
	k=4	6,171	6,251	5,588	5,514	7,664	7,619
	k=5	6,171	6,251	5,625	5,548	5,622	5,546
	k=6	6,171	6,251	5,689	5,609	5,684	5,603
e=168	k=4	3,452	3,478	3,387	3,381	3,407	3,401
	k=5	3,231	3,258	3,264	3,253	3,279	3,268
	k=6	3,100	3,127	3,317	3,304	3,328	3,315
	k=7	3,075	3,099	3,450	3,433	3,462	3,444
	k=8	3,048	3,073	3,556	3,535	3,566	3,544
	k=9	3,030	3,056	3,658	3,635	3,667	3,643
	k=10	3,044	3,069	3,771	3,744	3,778	3,749

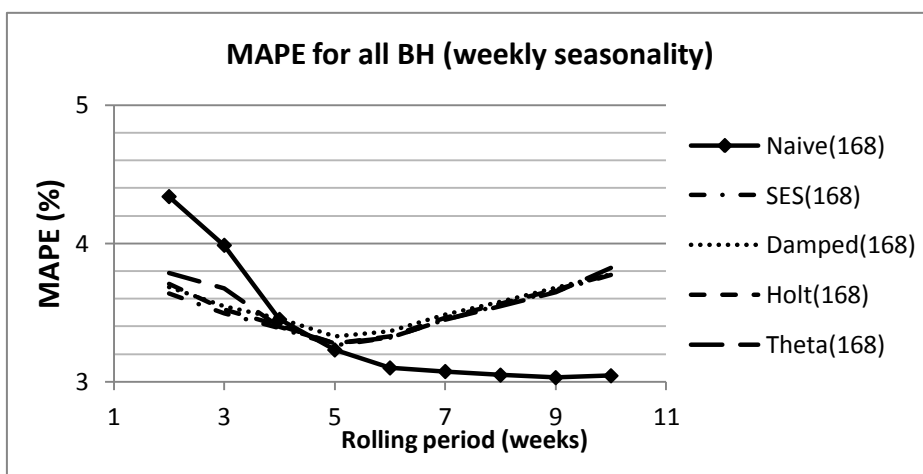
Πίνακας 16. Σφάλματα MAPE και sMAPE με επίδραση όλων των αργιών, επίσημων και μη (α)

		Damped previous day (all BH)		Theta previous day (all BH)		Taylor previous day (all BH)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	5,732	5,603	6,021	5,954		
	k=2	5,635	5,540	5,814	5,775		
	k=3	5,636	5,548	5,824	5,792		
	k=4	5,647	5,556	5,854	5,825		
	k=5	5,671	5,576	5,895	5,867		
	k=6	5,728	5,626	5,955	5,928		
e=168	k=4	3,455	3,445	3,397	3,405	4,185	4,227
	k=5	3,329	3,312	3,279	3,286	4,394	4,492
	k=6	3,366	3,348	3,321	3,329	4,776	4,937
	k=7	3,483	3,460	3,446	3,453	4,428	4,425
	k=8	3,579	3,550	3,548	3,555	4,143	4,181
	k=9	3,680	3,647	3,645	3,655	4,698	4,721
	k=10	3,777	3,739	3,821	3,824	5,142	5,411

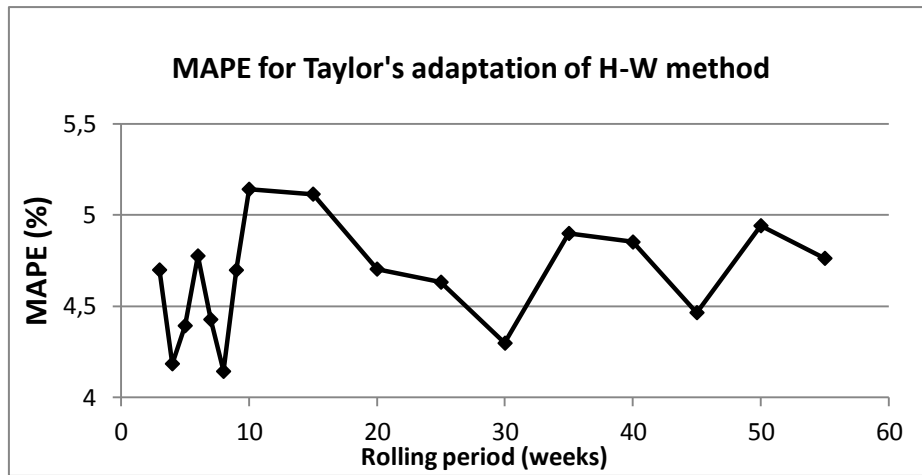
Πίνακας 17. Σφάλματα MAPE και sMAPE με επίδραση όλων των αργιών, επίσημων και μη (β)



Εικόνα 24. Σφάλμα MAPE με επίδραση επίσημων και μη αργιών και ημερήσια εποχικότητα



Εικόνα 25. Σφάλμα MAPE με επίδραση επίσημων και μη αργιών και εβδομαδιαία εποχικότητα



Εικόνα 26. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση επίσημων και μη αργιών

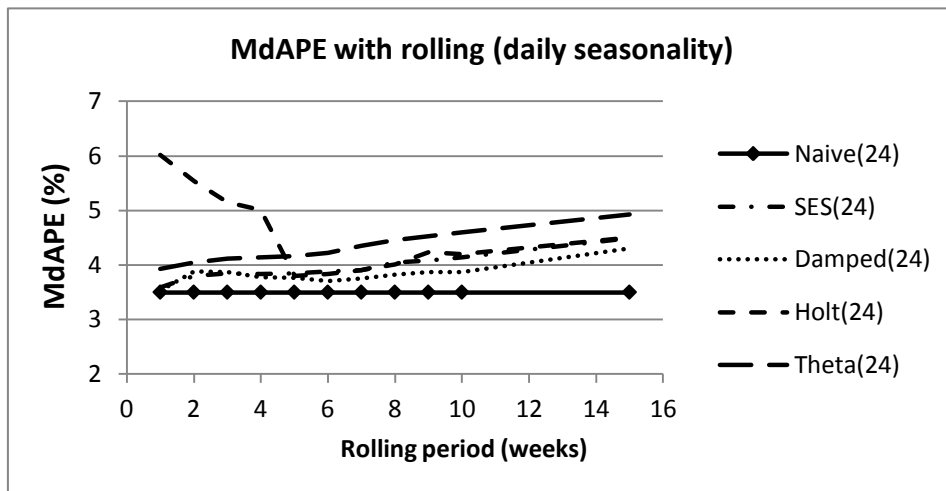
Ακολουθούν, στους πίνακες 18 και 19, ενδεικτικά σφάλματα MdAPE και sMdAPE για τις αντίστοιχες μεθόδους και περιόδους. Και έπειτα τα διαγράμματα των σφαλμάτων MdAPE (εικόνες 27, 28 και 29).

		Naive previous day (all BH)		SES previous day (all BH)		Holt previous day (all BH)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,499	3,497	3,592	3,609	6,013	6,075
	k=2	3,499	3,497	3,801	3,837	5,536	5,627
	k=3	3,499	3,497	3,843	3,888	5,152	5,211
	k=4	3,499	3,497	3,833	3,846	5,011	5,066
	k=5	3,499	3,497	3,842	3,847	3,801	3,808
	k=6	3,499	3,497	3,879	3,884	3,833	3,848
e=168	k=4	2,312	2,314	2,241	2,238	2,241	2,238
	k=5	2,228	2,227	2,189	2,197	2,217	2,212
	k=6	2,094	2,096	2,251	2,249	2,264	2,266
	k=7	2,042	2,053	2,411	2,400	2,413	2,414
	k=8	2,034	2,042	2,498	2,504	2,501	2,511
	k=9	2,009	2,008	2,540	2,545	2,575	2,569
	k=10	2,019	2,024	2,653	2,657	2,665	2,664

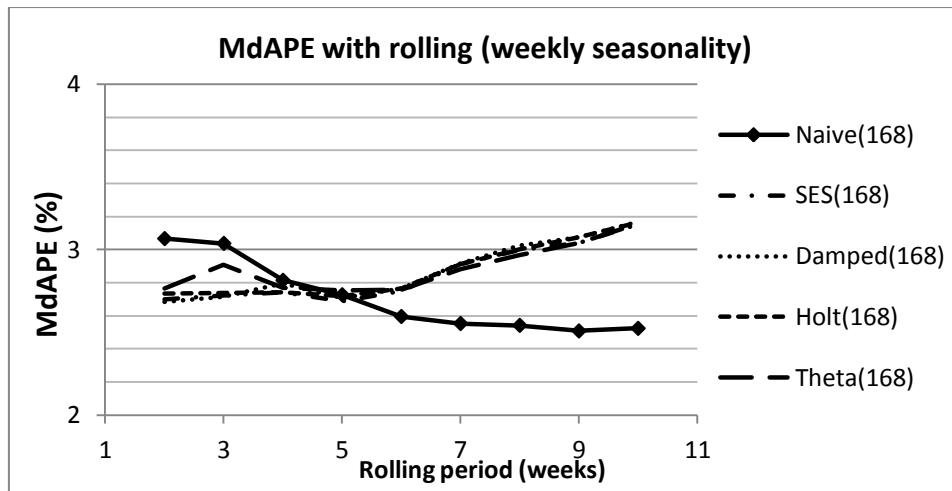
Πίνακας 18. Σφάλματα MdAPE και sMdAPE με επίδραση όλων των αργιών, επίσημων και μη (α)

		Damped previous day (all BH)		Theta previous day (all BH)		Taylor previous day (all BH)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,528	3,554	3,931	3,949		
	k=2	3,886	3,898	4,040	4,059		
	k=3	3,869	3,898	4,115	4,150		
	k=4	3,777	3,808	4,143	4,156		
	k=5	3,764	3,792	4,159	4,178		
	k=6	3,703	3,729	4,226	4,246		
e=168	k=4	2,297	2,311	2,270	2,270	2,481	2,469
	k=5	2,208	2,208	2,254	2,248	2,631	2,628
	k=6	2,264	2,263	2,258	2,261	2,838	2,848
	k=7	2,416	2,426	2,379	2,381	2,738	2,739
	k=8	2,523	2,526	2,466	2,470	2,577	2,569
	k=9	2,575	2,575	2,540	2,545	2,875	2,856
	k=10	2,669	2,683	2,655	2,661	3,194	3,197

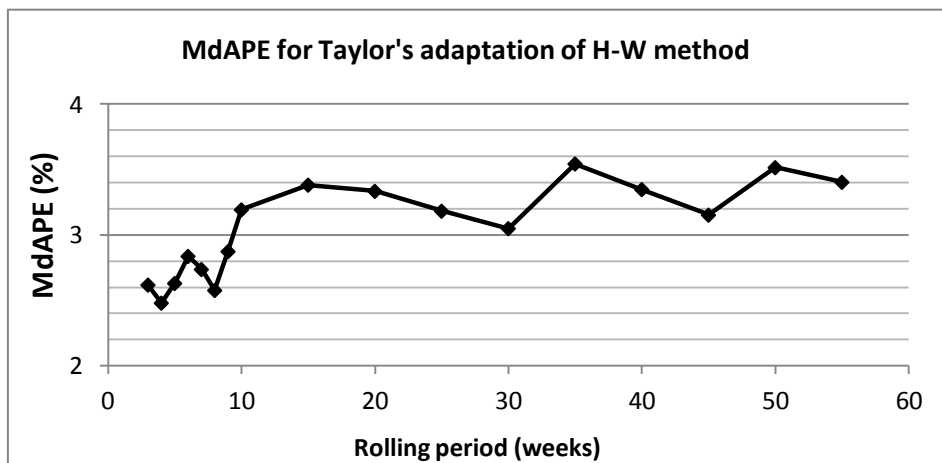
Πίνακας 19. Σφάλματα MdAPE και sMdAPE με επίδραση όλων των αργιών, επίσημων και μη (β)



Εικόνα 27. Σφάλμα MdAPE με επίδραση επίσημων και μη αργιών και ημερήσια εποχικότητα



Εικόνα 28. Σφάλμα MdAPE με επίδραση επίσημων και μη αργιών και εβδομαδιαία εποχικότητα



Εικόνα 29. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση επίσημων και μη αργιών

7.3.2 Αποτελέσματα με επίδραση των απεργιών (Μέσος όρος όλων των απεργιών των ετών 2009 και 2010)

Στη συνέχεια, παραθέτουμε τμήμα των σφαλμάτων που προέκυψαν από την ενσωμάτωση των ποσοστιαίων μεταβολών των απεργιών στις προβλέψεις. Συγκεκριμένα, χρησιμοποιείται η πρώτη περίπτωση της υποπαραγράφου 7.2.2. Η βέλτιστη εφαρμογή αυτού του συνδυασμού είναι για την περίπτωση εφαρμόζεται η ποσοστιαία μεταβολή (στην αρχική χρονοσειρά), για την προηγούμενη ημέρα και εβδομάδα, δηλαδή για την ποσοστιαία μεταβολή 24 (ημερήσια) και 168 (εβδομαδιαία) ωρών πριν αντίστοιχα. Τα αποτελέσματα για όλες τις εβδομάδες κυλιόμενης πρόβλεψης βρίσκονται στο Appendix (εικόνες A13 έως και A16).

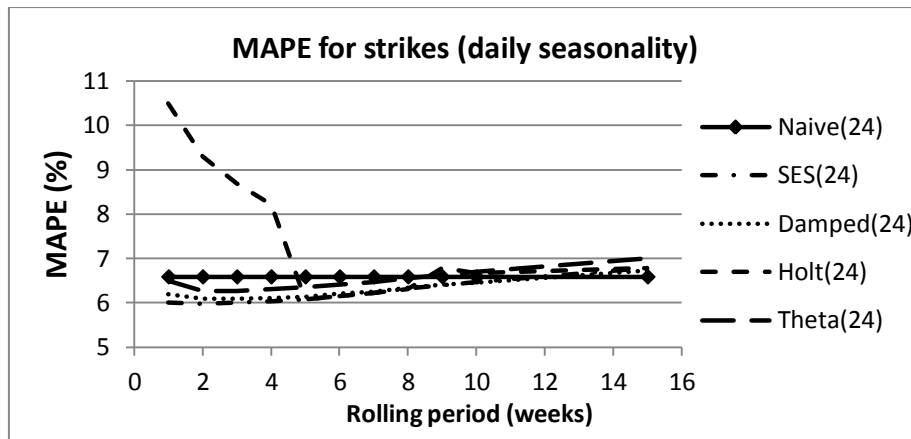
		Naive previous day+ week (strikes)		SES previous day+ week (strikes)		Holt previous day+ week (strikes)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	6,586	6,616	6,012	5,823	10,493	10,136
	k=2	6,586	6,616	5,973	5,811	9,293	9,020
	k=3	6,586	6,616	6,003	5,844	8,687	8,370
	k=4	6,586	6,616	6,035	5,873	8,216	8,060
	k=5	6,586	6,616	6,082	5,915	6,082	5,916
	k=6	6,586	6,616	6,156	5,983	6,153	5,979
e=168	k=4	3,870	3,848	3,778	3,693	3,796	3,712
	k=5	3,646	3,624	3,663	3,572	3,677	3,586
	k=6	3,520	3,497	3,741	3,642	3,752	3,653
	k=7	3,502	3,477	3,880	3,775	3,891	3,786
	k=8	3,466	3,442	3,997	3,884	4,007	3,893
	k=9	3,456	3,432	4,123	4,002	4,132	4,010
	k=10	3,463	3,439	4,240	4,113	4,247	4,119

Πίνακας 20. Σφάλματα MAPE και sMAPE με επίδραση του MO

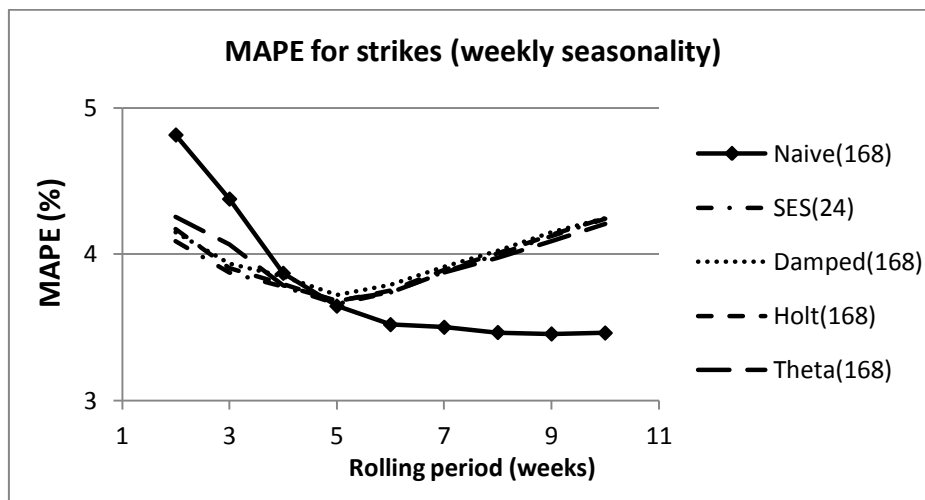
		Damped previous day+ week (strikes)		Theta previous day+ week (strikes)		Taylor previous day+ week (strikes)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	6,199	5,989	6,493	6,354		
	k=2	6,096	5,914	6,265	6,152		
	k=3	6,098	5,919	6,271	6,164		
	k=4	6,112	5,929	6,303	6,197		
	k=5	6,137	5,950	6,349	6,243		
	k=6	6,203	6,006	6,415	6,308		
e=168	k=4	3,847	3,756	3,788	3,732	4,636	4,569
	k=5	3,724	3,627	3,681	3,621	4,815	4,842
	k=6	3,792	3,685	3,739	3,676	5,225	5,304
	k=7	3,913	3,801	3,873	3,808	4,831	4,752
	k=8	4,024	3,902	3,975	3,907	4,577	4,563
	k=9	4,151	4,019	4,087	4,016	5,220	5,153
	k=10	4,248	4,111	4,208	4,135	5,290	5,323

Πίνακας 21. Σφάλματα MAPE και sMAPE με επίδραση του MO

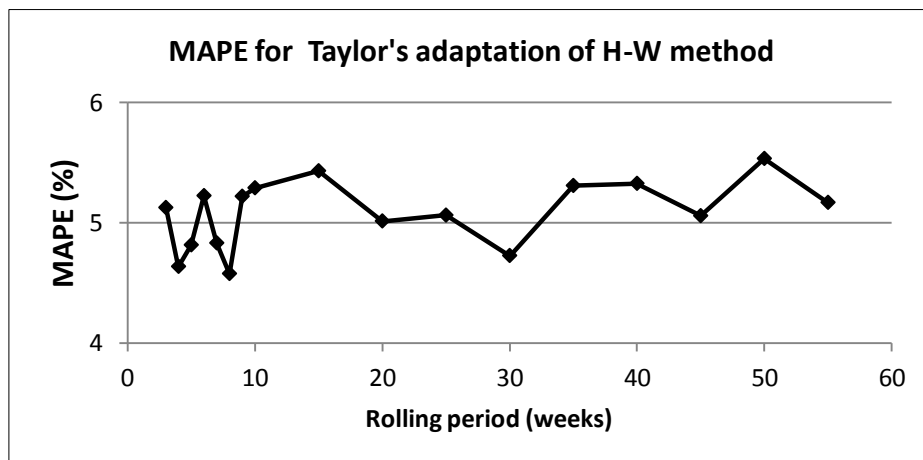
Ακολουθούν τα διαγράμματα με τα σφάλματα MAPE που σημειώθηκαν για τις μεθόδους υπό μελέτη, με την επίδραση της ποσοστιαίας μεταβολής όλων των απεργιών των ετών 2009 και 2010.



Εικόνα 30. Σφάλμα MAPE επίδραση του MO όλων των απεργιών και ημερήσια εποχικότητα



Εικόνα 31. Σφάλμα MAPE με επίδραση του MO όλων των απεργιών και εβδομαδιαία εποχικότητα



Εικόνα 32. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση του MO όλων των απεργιών

Στους πίνακες 22 και 23 παρουσιάζονται, όπως και στις προηγούμενες περιπτώσεις, τα αντίστοιχα σφάλματα MdAPE και sMdAPE και στις εικόνες 33, 34 και 35, τα διαγράμματα με τα αυτά τα σφάλματα για όλα τα χρονικά διαστήματα υλοποίησης της μεθόδους κύλισης των δεδομένων.

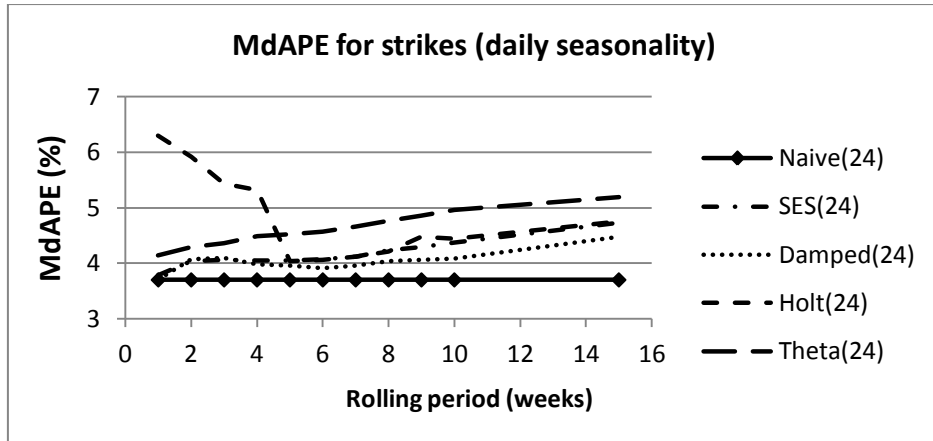
		Naive previous day+ week (strikes)		SES previous day+ week (strikes)		Holt previous day+ week (strikes)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,701	3,710	3,788	3,818	6,301	6,385
	k=2	3,701	3,710	4,042	4,077	5,917	5,980
	k=3	3,701	3,710	4,058	4,091	5,432	5,500
	k=4	3,701	3,710	4,046	4,078	5,319	5,348
	k=5	3,701	3,710	4,050	4,088	4,037	4,075
	k=6	3,701	3,710	4,072	4,099	4,066	4,102
e=168	k=4	2,426	2,424	2,289	2,295	2,287	2,295
	k=5	2,287	2,292	2,272	2,274	2,287	2,283
	k=6	2,179	2,179	2,350	2,352	2,350	2,356
	k=7	2,130	2,131	2,482	2,476	2,491	2,495
	k=8	2,123	2,131	2,598	2,601	2,602	2,617
	k=9	2,097	2,103	2,656	2,667	2,673	2,677
	k=10	2,106	2,106	2,769	2,781	2,777	2,779

Πίνακας 22. Σφάλματα MdAPE και sMdAPE επίδραση του MO όλων των απεργιών (α)

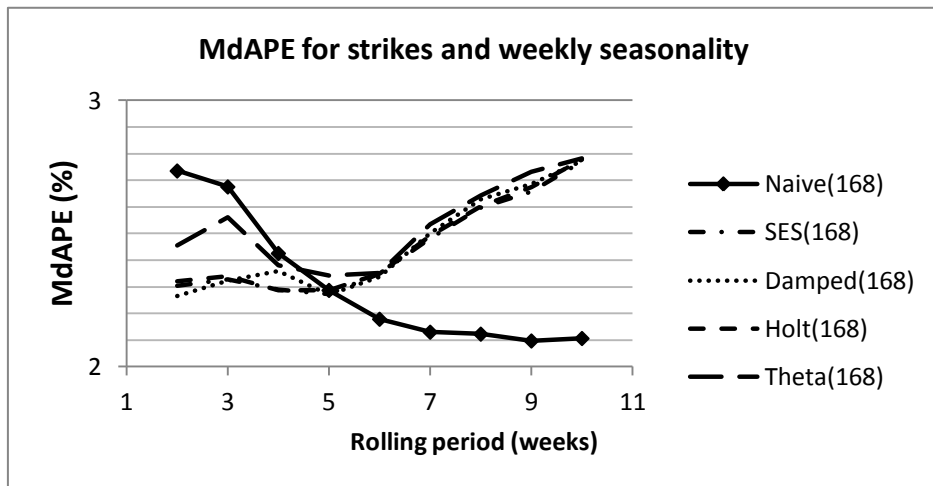
		Damped previous day+ week (strikes)		Theta previous day+ week (strikes)		Taylor previous day+ week (strikes)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,700	3,729	4,148	4,206		
	k=2	4,074	4,116	4,296	4,332		
	k=3	4,098	4,107	4,358	4,369		
	k=4	3,979	4,020	4,493	4,494		
	k=5	3,954	3,978	4,523	4,571		
	k=6	3,914	3,946	4,576	4,613		
e=168	k=4	2,358	2,359	2,380	2,394	2,559	2,561
	k=5	2,271	2,274	2,343	2,343	2,713	2,706
	k=6	2,337	2,342	2,353	2,352	2,995	2,994
	k=7	2,500	2,511	2,534	2,544	2,865	2,872
	k=8	2,627	2,630	2,642	2,649	2,708	2,703
	k=9	2,685	2,697	2,732	2,728	3,027	3,022
	k=10	2,773	2,773	2,781	2,797	3,288	3,287

Πίνακας 23. Σφάλματα MdAPE και sMdAPE επίδραση του MO όλων των απεργιών (β)

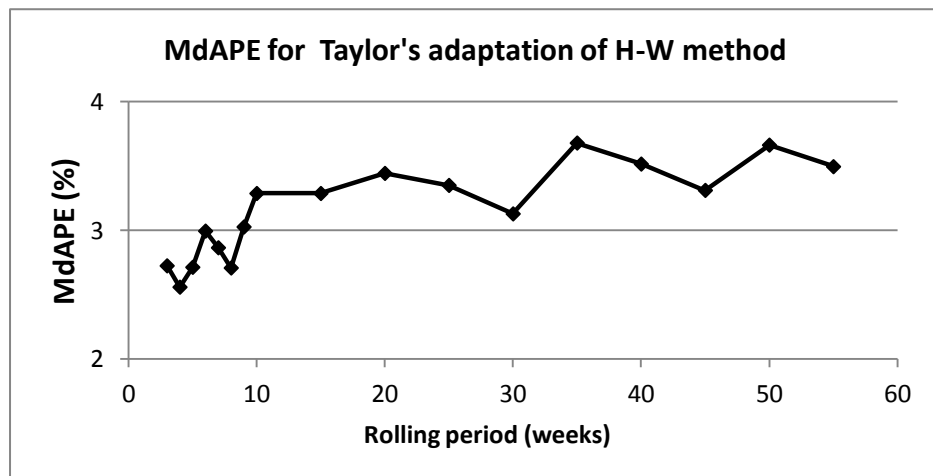
Και ακολουθούν οι εικόνες με τα αντίστοιχα MAPE.



Εικόνα 33. Σφάλμα MdAPE με επίδραση του MO όλων των απεργιών και ημερήσια εποχικότητα



Εικόνα 34. Σφάλμα MdAPE με επίδραση του MO όλων των απεργιών και εβδομαδιαία εποχικότητα



Εικόνα 35. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση του MO όλων των απεργιών

7.3.3 Αποτελέσματα με επίδραση αργιών και απεργιών

Στη συνέχεια παρουσιάζεται τμήμα των σφαλμάτων που προέκυψαν από την ενσωμάτωση της επίδρασης των ποσοστιαίων μεταβολών από το συνδυασμό των επίσημων αργιών και όλων των απεργιών των ετών 2009 και 2010. Στους πρώτους δύο πίνακες, 24 και 25 παρουσιάζεται με τμήμα των αποτελεσμάτων, ενώ στις εικόνες 36 έως και 38, παραθέτουμε διαγράμματα με όλα τα σφάλματα για καλύτερη εικόνα, καθώς γίνεται άμεση σύγκριση των μεθόδων. Τα αποτελέσματα για όλες τις εβδομάδες κυλιόμενης πρόβλεψης βρίσκονται στο Appendix (εικόνες A17 έως και A20).

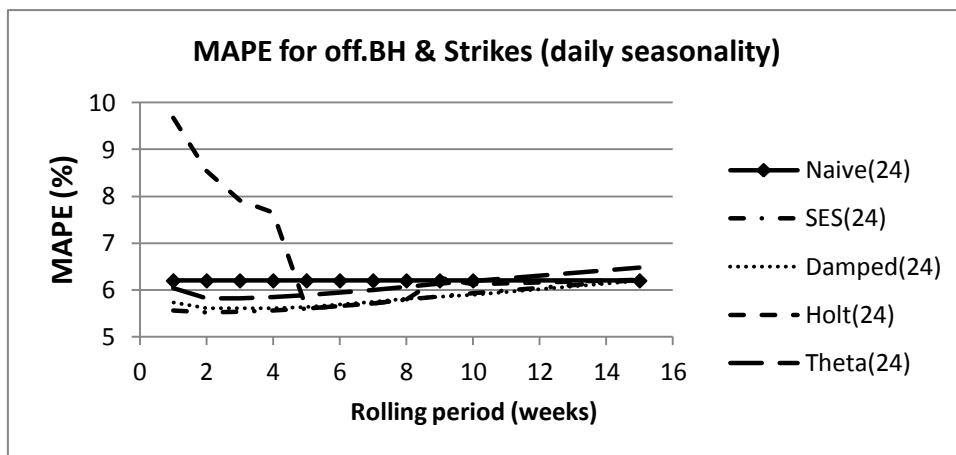
		Naive previous day (off.BH+ strikes)		SES previous day (off.BH+ strikes)		Holt previous day (off.BH+ strikes)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	6,195	6,299	5,557	5,461	9,670	9,485
	k=2	6,195	6,299	5,514	5,446	8,541	8,451
	k=3	6,195	6,299	5,533	5,470	7,907	7,779
	k=4	6,195	6,299	5,557	5,494	7,655	7,612
	k=5	6,195	6,299	5,594	5,528	5,594	5,529
	k=6	6,195	6,299	5,654	5,583	5,650	5,580
e=168	k=4	3,507	3,556	3,355	3,359	3,376	3,380
	k=5	3,289	3,341	3,237	3,236	3,253	3,252
	k=6	3,167	3,217	3,290	3,286	3,302	3,298
	k=7	3,147	3,196	3,420	3,412	3,433	3,424
	k=8	3,118	3,167	3,523	3,510	3,533	3,519
	k=9	3,105	3,155	3,629	3,613	3,638	3,621
	k=10	3,113	3,163	3,739	3,719	3,746	3,725

Πίνακας 24. Σφάλματα MAPE και sMAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (α)

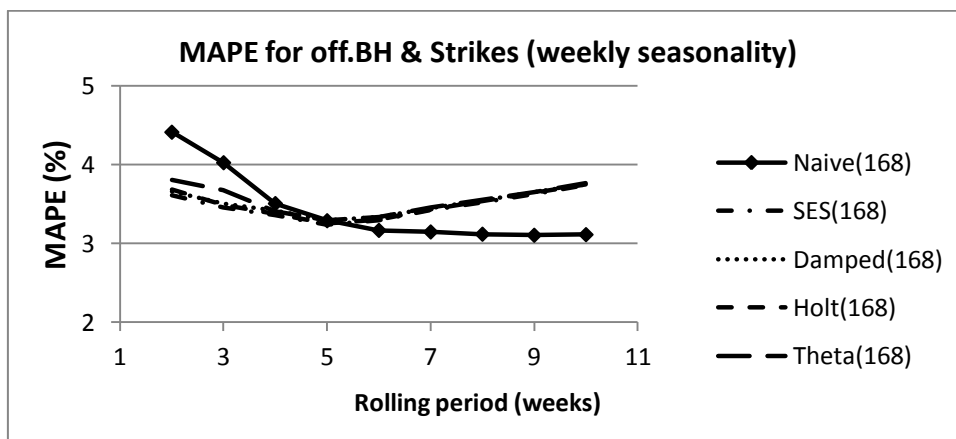
		Damped previous day (off.BH+ strikes)		Theta previous day (off.BH+ strikes)		Taylor previous day (off.BH+ strikes)	
		MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)	MAPE(%)	sMAPE(%)
e=24	k=1	5,730	5,613	6,037	5,989		
	k=2	5,613	5,527	5,817	5,793		
	k=3	5,604	5,524	5,817	5,800		
	k=4	5,613	5,531	5,844	5,829		
	k=5	5,639	5,554	5,883	5,870		
	k=6	5,693	5,600	5,940	5,927		
e=168	k=4	3,415	3,413	3,407	3,429	4,129	4,168
	k=5	3,293	3,287	3,292	3,312	4,411	4,527
	k=6	3,334	3,325	3,334	3,355	4,786	4,956
	k=7	3,450	3,435	3,459	3,479	4,422	4,432
	k=8	3,546	3,525	3,552	3,572	4,143	4,190
	k=9	3,653	3,627	3,648	3,669	4,752	4,787
	k=10	3,747	3,716	3,761	3,782	4,858	4,978

Πίνακας 25. Σφάλματα MAPE και sMAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (β)

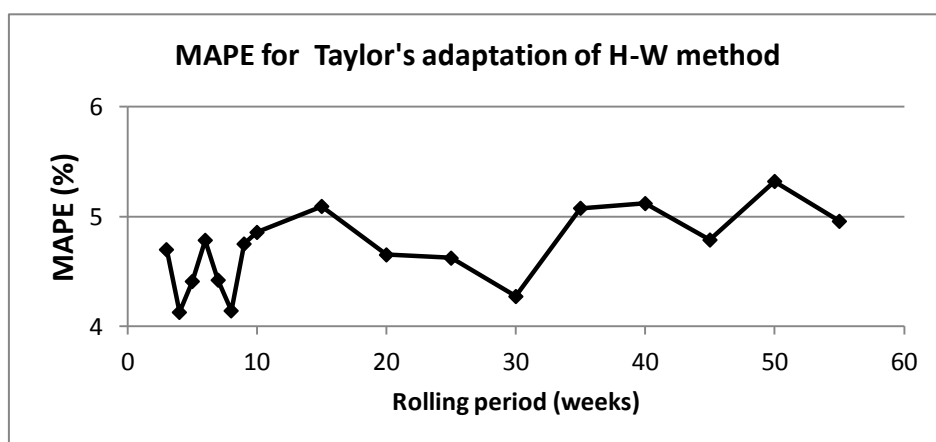
Και τα διαγράμματα που αναφέρθηκαν προηγουμένως.



Εικόνα 36. Σφάλμα MAPE με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (και ημερήσια εποχικότητα)



Εικόνα 37. Σφάλμα MAPE με επίδραση των επίσημων αργιών & όλων των απεργιών (και εβδομαδιαία εποχικότητα)



Εικόνα 38. Σφάλμα MAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών

Όπως και στους προηγούμενους συνδυασμούς, έτσι και εδώ παρουσιάζονται στους πίνακες 26 και 27 ορισμένα από τα σφάλματα MdAPE και sMdAPE και στη συνέχεια, έχουμε και τα διαγράμματα των σφαλμάτων MdAPE στις εικόνες 39, 40 και 41.

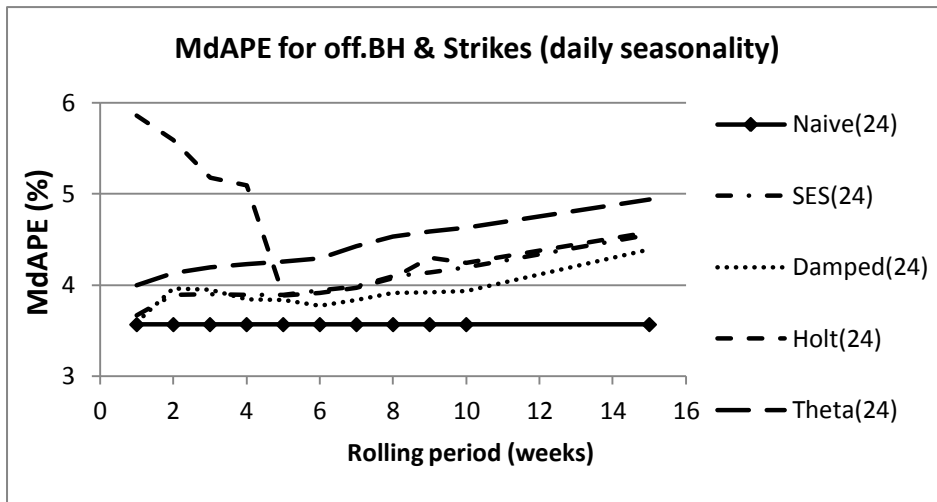
		NAIVE previous day (off.BH+ strikes)		SES previous day (off.BH+ strikes)		Holt previous day (off.BH+ strikes)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,565	3,545	3,665	3,686	5,861	5,927
	k=2	3,565	3,545	3,893	3,918	5,596	5,661
	k=3	3,565	3,545	3,901	3,948	5,178	5,246
	k=4	3,565	3,545	3,893	3,937	5,093	5,129
	k=5	3,565	3,545	3,892	3,924	3,885	3,906
	k=6	3,565	3,545	3,939	3,971	3,914	3,931
e=168	k=4	2,358	2,353	2,244	2,245	2,245	2,245
	k=5	2,252	2,252	2,219	2,219	2,234	2,230
	k=6	2,117	2,121	2,269	2,272	2,294	2,300
	k=7	2,082	2,084	2,429	2,423	2,437	2,428
	k=8	2,085	2,087	2,526	2,526	2,529	2,536
	k=9	2,063	2,062	2,578	2,577	2,603	2,606
	k=10	2,060	2,054	2,681	2,692	2,691	2,701

Πίνακας 26. Σφάλματα MdAPE και sMdAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (α)

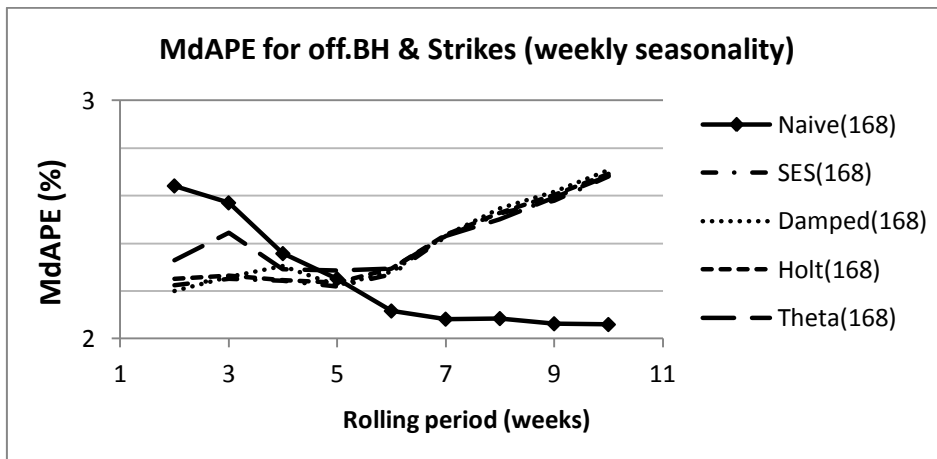
		Damped previous day (off.BH+ strikes)		Theta previous day (off.BH+ strikes)		Taylor previous day (off.BH+ strikes)	
		MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	3,589	3,614	3,997	4,030		
	k=2	3,962	4,001	4,127	4,176		
	k=3	3,949	3,975	4,195	4,216		
	k=4	3,841	3,880	4,232	4,255		
	k=5	3,833	3,859	4,261	4,273		
	k=6	3,772	3,804	4,290	4,326		
e=168	k=4	2,306	2,318	2,292	2,289	2,487	2,475
	k=5	2,224	2,228	2,286	2,280	2,649	2,649
	k=6	2,282	2,289	2,294	2,294	2,890	2,884
	k=7	2,433	2,444	2,431	2,427	2,788	2,794
	k=8	2,547	2,554	2,501	2,505	2,630	2,621
	k=9	2,616	2,612	2,593	2,604	2,921	2,929
	k=10	2,709	2,706	2,683	2,695	3,199	3,206

Πίνακας 27. Σφάλματα MdAPE και sMdAPE με την επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (α)

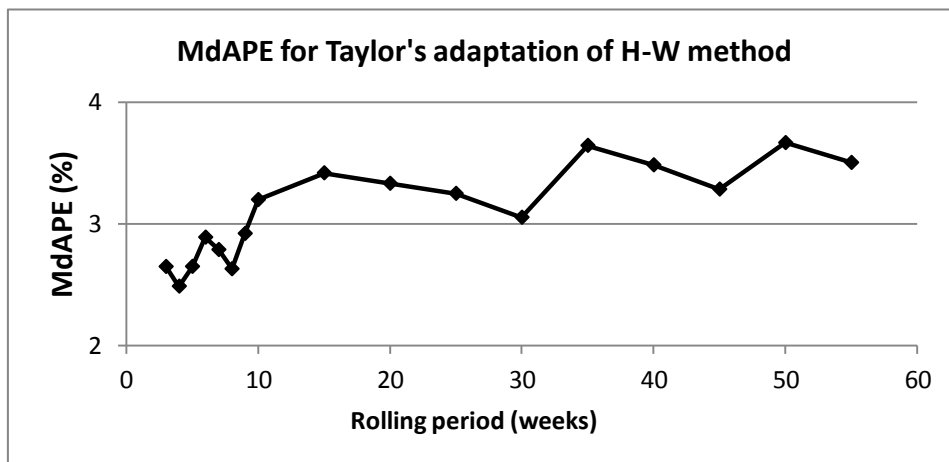
Και τα διαγράμματα των σφαλμάτων MdAPE, για τις όλες τις μεθόδους πρόβλεψης που μελετάμε και για όλες τις εβδομάδες που χρησιμοποιήθηκαν ως χρονική περίοδος εφαρμογής της μεθόδου κύλισης των δεδομένων:



Εικόνα 39. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (και ημερήσια εποχικότητα)



Εικόνα 40. Σφάλμα MdAPE με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών (και εβδομαδιαία εποχικότητα)



Εικόνα 41. Σφάλμα MdAPE για τη μέθοδο Holt-Winters exponential smoothing for double seasonality με επίδραση των επίσημων αργιών και του ΜΟ όλων των απεργιών

Η περίπτωση της συμπερίληψης όλων των αργιών (επίσημων και μη) και της δεύτερης περίπτωσης από το Β για τις αργίες (με υπολογισμό του μέσου όρου των ποσοστών για ίδιες ημέρες της εβδομάδας στις οποίες έχουμε εμφάνιση απεργιών) παρουσιάζεται στους πίνακες Α21 έως και Α24 στο Appendix, καθώς δεν παράγουν καλύτερο αποτέλεσμα συγκριτικά με τον πρώτο συνδυασμό αργιών και απεργιών που μόλις παρουσιάσαμε.

7.4 Ανάλυση αποτελεσμάτων

Μελετώντας τους πίνακες και τα διαγράμματα της παραγράφου 7.3 συμπεραίνει κανείς την εμφανή μείωση των σφαλμάτων επί όλων των δεδομένων της χρονοσειράς ζήτησης φορτίου για το χρονικό διάστημα 1/1/2011 έως και 31/8/2011 (243 ημέρες). Η μείωση στα σφάλματα ενδεχομένως να φαίνεται μικρή, όμως θα πρέπει κανείς να αναλογιστεί το μικρό πλήθος ειδικών ημερών που έχουν ενσωματωθεί συγκριτικά με το σύνολο των ημερών στις οποίες υπολογίζεται το τελικό σφάλμα. Συγκεκριμένα, στην περίπτωση επίσημων αργιών, έχει γίνει αλλαγή σε μόλις δέκα (10) ημέρες όπως έχουν αναφερθεί στην παράγραφο 7.2. Στην περίπτωση όλων των αργιών, ενσωματώθηκε η ποσοστιαία μεταβολή δεκατέσσερις (14) ημέρες. Στην ενσωμάτωση μόνο των απεργιών χρησιμοποιήθηκαν μόλις πέντε (5) ημέρες και τέλος, στο συνδυασμό των αργιών με τις απεργίες, οι ημέρες που συνυπολογίστηκαν είναι στην πρώτη μέθοδο και στη δεύτερη. Στον επόμενο πίνακα (πίνακας 28) παρουσιάζονται ενδεικτικά τα σφάλματα MAPE της μεθόδου Naive, υπολογισμένα μόνο για τις ειδικές ημέρες κάθε περίπτωσης που αναφέρθηκε. Σημειώνονται τα σφάλματα των ημερών αυτών έπειτα από την αρχική πρόβλεψη και μετά από την ενσωμάτωση των ποσοστιαίων μεταβολών των ειδικών ημερών. Και στη συνέχεια καταγράφεται η ποσοστιαία μεταβολή των σφαλμάτων επί των ειδικών ημερών και μόνο.

Error reduction (of MAPE for Naive method)										
Seasonality:	e=24		e=168							
Rolling period:	k=1,..15	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Initial MAPE for official BH:	14,303	14,578	13,457	13,705	13,553	13,470	13,512	13,354	13,430	13,399
MAPE with impact of official BH:	4,870	4,870	4,870	4,870	4,870	4,870	4,870	4,870	4,870	4,870
Percentage reduction of MAPE:	-65,950	-66,593	-63,809	-64,466	-64,066	-63,844	-63,957	-63,531	-63,736	-63,652
Actual reduction of MAPE:	-9,433	-9,708	-8,587	-8,835	-8,683	-8,600	-8,642	-8,484	-8,559	-8,528
Maximum % reduction:	-66,593									
Maximum actual reduction:	-9,708									
Initial MAPE for all BH:	13,137	13,623	12,859	12,949	12,820	12,809	12,877	12,686	12,777	12,753
MAPE with impact of all BH:	5,768	5,768	5,768	5,768	5,768	5,768	5,768	5,768	5,768	5,768
Percentage reduction of MAPE:	-56,093	-57,659	-55,145	-55,455	-55,008	-54,970	-55,207	-54,532	-54,856	-54,773
Actual reduction of MAPE:	-7,369	-7,855	-7,091	-7,181	-7,052	-7,041	-7,109	-6,918	-7,009	-6,985
Maximum % reduction:	-57,659									
Maximum actual reduction:	-7,855									

Πίνακας 28α. Σφάλματα MAPE για τη μέθοδο Naive πριν και μετά την επίδραση των ειδικών ημερών

Error reduction (of MAPE for Naive method)										
Seasonality:	e=24	e=168								
Rolling period:	k=1,..15	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Initial MAPE for Strikes:	3,413	1,993	3,842	2,699	2,458	2,235	2,066	1,946	1,813	2,083
MAPE with impact of Strikes:	1,571	1,571	1,571	1,571	1,571	1,571	1,571	1,571	1,571	1,571
Percentage reduction of MAPE:	-53,961	-21,147	-59,097	-41,778	-36,075	-29,687	-23,946	-19,249	-13,338	-24,574
Actual reduction of MAPE:	-1,842	-0,421	-2,270	-1,128	-0,887	-0,663	-0,495	-0,375	-0,242	-0,512
Maximum % reduction:	-59,097									
Maximum actual reduction:	-2,270									
Initial MAPE for off.BH+ Strikes:	10,900	10,645	10,452	10,266	10,086	9,959	9,935	9,789	9,799	9,863
MAPE with impact of off.BH+ Strikes:	4,851	4,851	4,851	4,851	4,851	4,851	4,851	4,851	4,851	4,851
Percentage reduction of MAPE:	-55,491	-54,427	-53,584	-52,742	-51,899	-51,285	-51,169	-50,441	-50,493	-50,810
Actual reduction of MAPE:	-6,048	-5,794	-5,601	-5,414	-5,234	-5,107	-5,084	-4,938	-4,948	-5,011
Maximum % reduction:	-55,491									
Maximum actual reduction:	-6,048									

Πίνακας 298β. Σφάλματα MAPE για τη μέθοδο Naive πριν και μετά την επίδραση των ειδικών ημερών

Είναι φανερό ότι η μεγαλύτερη μείωση σφάλματος επιτυγχάνεται για την περίπτωση εφαρμογής των επίσημων αργιών. Και πιο συγκεκριμένα, στην εφαρμογή Κεντρικού Κινητού Μέσου Όρου μήκους 168 ωρών (μία εβδομάδα) και κύλιση δεδομένων χρονικού διαστήματος δύο εβδομάδων. Ενώ τη μικρότερη μείωση σφάλματος παρουσιάζει η μέθοδος ενσωμάτωσης των επίσημων αργιών σε συνδυασμό με τις απεργίες, ίση με -55,49%. Το αρχικό επίπεδο σφάλματος στην περίπτωση αυτή είναι μικρότερο από εκείνο του σφάλματος για τις επίσημες αργίες, 10,9% έναντι 14,3% (για τις επίσημες αργίες), ωστόσο το τελικό σφάλμα είναι ελάχιστα μικρότερο, 4,85% έναντι 4,87% (για τις επίσημες αργίες). Γενικό συμπέρασμα είναι πως παρότι χρησιμοποιούνται περισσότερες ημέρες στην ενσωμάτωση όλων των αργιών, βλέπουμε ότι προκύπτουν γενικά χειρότερα σφάλματα επί του συνόλου των εκάστοτε ειδικών ημερών. Ακόμη, το μικρότερο σφάλμα εμφανίζεται για τις ημέρες των απεργιών, που θεωρείται και επόμενο, καθώς είναι ο μικρότερος αριθμός δεδομένων που εξετάζεται (ενώ για περισσότερες παρατηρήσεις σε διάσπαρτες ημέρες, αναμένεται μεγαλύτερο σφάλμα). Παρουσιάζει δε το δεύτερο μεγαλύτερο ποσοστό μείωσης σφάλματος MAPE έπειτα από την απλοϊκή μέθοδο. Όταν όμως εξετάζεται το σφάλμα σε όλη τη χρονοσειρά για την ίδια περίπτωση, τότε τα αποτελέσματα για ενσωμάτωση μόνο της επίδρασης των απεργιών, είναι λιγότερο αποδοτική η ενσωμάτωση της κατηγορίας αυτής των ειδικών ημερών, επειδή πρόκειται για πολύ μικρότερο αριθμό ημερών αναλογικά με το πλήθος που εξετάζεται, αλλά και συγκριτικά με το πλήθος των αργιών, η ενσωμάτωση των οποίων παράγει καλύτερες προβλέψεις.

Στον επόμενο πίνακα (πίνακας 29) σημειώνεται για κάθε συνδυασμό ειδικών ημερών που εξετάστηκε και για κάθε μοντέλο που υλοποιήθηκε, το ποσοστό των ωριαίων προβλέψεων (από τις 5832 παρατηρήσεις του 2011 που μελετάμε) στις οποίες τα σφάλματα MAPE και sMAPE έχουν τιμή μικρότερη του 3%.

Percentage of hourly forecasts with error < 3%							
Methods	Applied Combinations	Official Bank Holidays		All Bank Holidays		Strikes	
	Type of errors	MAPE	sMAPE	MAPE	sMAPE	MAPE	sMAPE
NAIVE	Worst case	44,959%	44,993%	45,422%	45,456%	43,587%	43,656%
	Best case	65,398%	65,432%	65,964%	65,878%	64,095%	64,009%
	Average percentage	52,585%	52,574%	53,050%	53,038%	51,233%	51,250%
SES	Worst case	35,511%	35,425%	35,940%	35,854%	34,482%	34,431%
	Best case	60,991%	60,888%	61,197%	61,111%	60,031%	59,962%
	Average percentage	48,810%	48,707%	49,162%	49,059%	47,719%	47,634%
DAMPED	Worst case	36,951%	36,797%	37,500%	37,346%	36,025%	35,905%
	Best case	60,648%	60,511%	60,974%	60,802%	59,808%	59,688%
	Average percentage	49,068%	48,928%	49,446%	49,300%	48,081%	47,957%
HOLT	Worst case	32,270%	32,287%	32,750%	32,767%	31,396%	31,430%
	Best case	60,974%	60,820%	61,300%	61,145%	60,117%	59,979%
	Average percentage	46,986%	46,880%	47,348%	47,245%	45,905%	45,806%
THETA	Worst case	32,956%	32,888%	33,556%	33,488%	31,842%	31,807%
	Best case	60,340%	60,082%	60,854%	60,597%	59,414%	59,225%
	Average percentage	46,800%	46,741%	47,326%	47,264%	45,652%	45,604%
TAYLOR's	Worst case	43,278%	43,227%	43,861%	43,810%	42,438%	42,438%
	Best case	56,070%	56,190%	56,361%	56,447%	55,161%	55,247%
	Average percentage	48,931%	48,925%	49,467%	49,465%	48,071%	48,088%

Πίνακας 29α. Ποσοστά ωριαίων προβλέψεων με σφάλμα (MAPE και sMAPE) μικρότερο του 3%, πριν και μετά από την ενσωμάτωση της επίδρασης των ειδικών ημερών

Percentage of hourly forecasts with error < 3%					
Methods	Applied Combinations	Off. BH and Strikes		Original Forecasts	
	Type of errors	MAPE	sMAPE	MAPE	sMAPE
NAIVE	Worst case	44,890%	44,890%	42,970%	43,021%
	Best case	65,089%	64,952%	63,872%	63,786%
	Average percentage	52,409%	52,361%	50,733%	50,736%
SES	Worst case	35,734%	35,648%	33,573%	33,471%
	Best case	60,940%	60,820%	59,294%	59,191%
	Average percentage	48,812%	48,680%	47,041%	46,934%
DAMPED	Worst case	36,660%	36,488%	35,168%	35,014%
	Best case	60,768%	60,597%	59,002%	58,865%
	Average percentage	49,024%	48,853%	47,416%	47,272%
HOLT	Worst case	32,733%	32,716%	30,247%	30,264%
	Best case	61,025%	60,837%	58,539%	58,471%
	Average percentage	47,035%	46,892%	43,839%	43,736%
THETA	Worst case	33,230%	33,128%	30,881%	30,830%
	Best case	60,374%	60,065%	58,693%	58,453%
	Average percentage	46,826%	46,728%	44,940%	44,879%
TAYLOR's	Worst case	42,644%	42,627%	41,615%	41,598%
	Best case	56,293%	56,361%	54,252%	54,338%
	Average percentage	48,810%	48,781%	47,205%	47,206%

Πίνακας 29β. Ποσοστά ωριαίων προβλέψεων με σφάλμα (MAPE και sMAPE) μικρότερο του 3%, πριν και μετά από την ενσωμάτωση της επίδρασης των ειδικών ημερών

Με βάση τα στοιχεία των προηγούμενων εικόνων παρατηρούμε ότι όλες οι περιπτώσεις έχουν ποσοστά μεγαλύτερα από τα αντίστοιχα της εικόνας 22 για την αρχική παραγωγή των προβλέψεων (χωρίς ενσωμάτωση των ειδικών ημερών). Άρα, η ενσωμάτωση της επίδρασης των ειδικών ημερών προκαλεί ελάττωση του επιπέδου του σφάλματος σε όλες τις εφαρμογές των μοντέλων πρόβλεψης. Ακόμη, η μεγαλύτερη τιμή ποσοστών των ημερών του 2011 με σφάλμα MAPE και sMAPE μικρότερο από 3%, εμφανίζεται για την περίπτωση των επίσημων και μη αργιών (δεύτερη στήλη σε κάθε περίπτωση). Γεγονός που έρχεται σε αντίθεση με τα αποτελέσματα των σφαλμάτων που παρουσιάστηκαν στην παράγραφο 7.1 και δείχνουν ότι η μικρότερες τιμές των σφαλμάτων επιτυγχάνονται για την περίπτωση της ενσωμάτωσης των επίσημων μόνο αργιών. Η διαφορά των ποσοστών είναι πολύ μικρή ανάμεσα στις δύο μεθόδους και θεωρούμε ότι οφείλεται στο ότι τα σφάλματα για την ενσωμάτωση των επίσημων αργιών είναι μικρότερα τα σφάλματα για την ενσωμάτωση όλων των αργιών (επίσημων και μη) σχεδόν σε κάθε περίπτωση. Έτσι, ακόμη και αν το πλήθος των ημερών με σφάλμα μικρότερο του 3% είναι για λίγο μεγαλύτερο για τη δεύτερη περίπτωση από το αντίστοιχο σφάλμα για την πρώτη, υπερτερεί η γενικότερα μικρή τιμή της πρώτης περίπτωσης. Ο συνδυασμός των επίσημων αργιών με τις απεργίες είναι τρίτος σε κατάταξη, ενώ τελευταία βρίσκεται η περίπτωση ενσωμάτωσης μόνο των απεργιών, όπως άλλωστε αναμέναμε.

Έχοντας υπόψη τις μειώσεις των σφαλμάτων, παραθέτουμε έπειτα συγκεντρωμένα τα καλύτερα σφάλματα MAPE, MdAPE, sMAPE και sMdAPE όλων των μεθόδων και για τους βασικούς συνδυασμούς που υλοποιήσαμε, καθώς και για την αρχική πρόβλεψη, από το κεφάλαιο 6.

Η παράθεση των σφαλμάτων γίνεται από την καλύτερη ως τη χειρότερη πρόβλεψη, σύμφωνα με το σφάλμα MAPE, για κάθε περίπτωση (1, 2, 3, 4 και 5). Παρατίθενται ξανά τα αποτελέσματα της αρχικής πρόβλεψης από το κεφάλαιο 6, ώστε να είναι ευκολότερη η σύγκριση με τις υπόλοιπες εκτελέσεις των προβλέψεων. Σε κάθε περίπτωση είναι τονισμένες οι τρεις μέθοδοι με τα πρώτα ελάχιστα σφάλματα MAPE. Για παράδειγμα, στην πρώτη περίπτωση (1) τα τρία πρώτα καλύτερα αποτελέσματα MAPE προκύπτουν από τις μεθόδους Naive, SES και Holt με τη σειρά, ενώ ακολουθούν οι επόμενες τρεις μέθοδοι (Theta, Damped και τελευταία η προσαρμογή του Taylor).

1)	Best methods with % impact of official BH	MAPE(%)	MdAPE(%)	sMAPE(%)	sMdAPE(%)
	Naive	3,074	2,027	3,123	2,039
	SES	3,234	2,219	3,232	2,219
	Holt	3,251	2,236	3,248	2,232
	Theta	3,287	2,287	3,307	2,282
	Damped	3,296	2,196	3,288	2,199
	Taylor	4,133	2,506	4,180	2,501
2)	Best methods with % impact of all BH	MAPE(%)	MdAPE(%)	sMAPE(%)	sMdAPE(%)
	Naive	3,030	2,009	3,056	2,008
	SES	3,264	2,189	3,253	2,197
	Holt	3,279	2,217	3,268	2,212
	Theta	3,279	2,254	3,286	2,248
	Damped	3,329	2,182	3,312	2,184
	Taylor	4,143	2,481	4,181	2,469
3)	Best methods official BH+ strikes	MAPE(%)	MdAPE(%)	sMAPE(%)	sMdAPE(%)
	Naive	3,105	2,060	3,155	2,054
	SES	3,237	2,219	3,236	2,219
	Holt	3,253	2,234	3,252	2,230
	Damped	3,293	2,200	3,287	2,200
	Theta	3,292	2,286	3,312	2,280
	Taylor	4,129	2,487	4,168	2,475
4)	Best methods with % impact of strikes (as total average)	MAPE(%)	MdAPE(%)	sMAPE(%)	sMdAPE(%)
	Naive	3,456	2,097	3,432	2,103
	SES	3,663	2,272	3,572	2,274
	Holt	3,677	2,287	3,586	2,283
	Theta	3,681	2,343	3,621	2,343
	Damped	3,724	2,265	3,627	2,263
	Taylor	4,577	2,559	4,563	2,561

Πίνακας 30α. Βέλτιστα σφάλματα μεθόδων για όλες τις προβλέψεις

5)	Best methods original (with rolling)	MAPE(%)	MdAPE(%)	sMAPE(%)	sMdAPE(%)
	Naive	3,461	2,103	3,437	2,107
	SES	3,696	2,317	3,604	2,320
	Theta	3,712	2,377	3,652	2,393
	Damped	3,763	2,303	3,664	2,299
	Holt	3,926	2,349	3,819	2,352
	Taylor	4,603	2,620	4,589	2,624

Πίνακας 30β. Βέλτιστα σφάλματα μεθόδων για όλες τις προβλέψεις

Συγκεκριμένα, η περίπτωση που παράγει το καλύτερο επίπεδο σφαλμάτων είναι η ενσωμάτωση της ποσοστιαίας μεταβολής των επίσημων αργιών (βάσει προηγούμενης ημέρας). Δεύτερη καλύτερη περίπτωση (με βάση το MAPE) και μηδαμινή διαφορά από την πρώτη, είναι για ενσωμάτωση όλων των αργιών (επίσημων και μη) και ακολουθούν ο συνδυασμός επίσημων αργιών και μέσου όρου όλων των απεργιών των ετών 2009 και 2010, όλες οι απεργίες, Τελευταία περίπτωση είναι ,όπως αναμέναμε, η αρχική πρόβλεψη χωρίς την συμπερίληψη της επίδρασης των ειδικών ημερών. Τα αποτελέσματα του συνδυασμού όλων των αργιών και της δεύτερης περίπτωσης απεργιών (βασισμένες στην ημέρα της απεργίας) είναι τα λιγότερο ικανοποιητικά από τις περιπτώσεις των ειδικών ημερών και δεν τα έχουμε συμπεριλάβει στον πίνακα 30.

Με βάση όλα τα παραπάνω, συμπεραίνουμε ότι η επίδραση των ειδικών ημερών στην πρόβλεψη ζήτησης ηλεκτρικού φορτίου είναι ιδιαίτερα σημαντική και ότι η αντιμετώπιση των δεδομένων τους ως κενές τιμές δεν είναι αντιπροσωπευτική. Αυτό φαίνεται κυρίως στον πίνακα 28, όπου παρουσιάζεται η μεταβολή των σφαλμάτων του απλοϊκού μοντέλου πρόβλεψης, ανάλογα με τη μέθοδο συμπερίληψης των ειδικών ημερών, αλλά και συγκρίνοντας τα ποσοστά από την αρχική πρόβλεψη, με εκείνα από την πρόβλεψη με τις ειδικές ημέρες, στον πίνακα 29.

Μία σημαντική παρατήρηση είναι η ελάχιστη τιμή σφάλματος που επιτύχαμε. Είναι για τη μέθοδο Naive (την απλοϊκή μέθοδο) και για εφαρμογή κύλισης δεδομένων με χρονικό διάστημα εννέα (9) εβδομάδων, όπως και στην αρχική πρόβλεψη. Ωστόσο, όμοια με τη μέθοδο Theta, οι βέλτιστες προβλέψεις δεν επιτυγχάνονται με την ενσωμάτωση των επίσημων αργιών, όπως συμβαίνει στις υπόλοιπες μεθόδους, αλλά για την περίπτωση ενσωμάτωσης όλων των αργιών. Η μείωση του σφάλματος ανάμεσα στις δύο αυτές περιπτώσεις είναι περίπου ίση με 1,4% για το MAPE και 2% για το MdAPE. Στη συνέχεια όμως που δοκιμάστηκε ο συνδυασμός όλων των αργιών με τις απεργίες το σφάλμα τους προέκυψε χειρότερο από εκείνο του συνδυασμού των επίσημων αργιών με τις απεργίες. Δηλαδή η μέθοδος με όλες τις αργίες έρχεται εν γένει δεύτερη σε σύγκριση με τη μέθοδο μόνο με τις επίσημες αργίες. Να σημειωθεί ότι έχει γίνει μελέτη όλων των συνδυασμών αργιών και απεργιών παρότι εδώ παρουσιάζεται μόλις ο ένας από αυτούς, των επίσημων αργιών με το μέσο όρο όλων των απεργιών (για τα έτη 2009 και 2010), ο οποίος παράγει τα καλύτερα αποτελέσματα από όλους τους συνδυασμούς αργιών με απεργίες.

Παρατηρώντας τον πίνακα 30, γίνεται αισθητή και η κατάταξη των τριών καλύτερων μεθόδων πρόβλεψης βραχυπρόθεσμης ζήτησης φορτίου ηλεκτρικής ενέργειας, καθώς έχουν σημειωθεί με πιο έντονα γράμματα σε κάθε περίπτωση. Ακολουθείται, λοιπόν, συγκεκριμένο μοτίβο κατά την παραγωγή προβλέψεων όταν ενσωματώνονται οι ειδικές

ημέρες, με πρώτη καλύτερη μέθοδο τη Naive, έπειτα τη SES και Τρίτη καλύτερη μέθοδο τη Holt. Η μόνη διαφορά σημειώνεται κατά την αρχική πρόβλεψη της ζήτησης του φορτίου, όπου οι πρώτες δύο καλύτερες μέθοδοι παραμένουν οι Naive και SES, αλλάζει ωστόσο η Τρίτη, όπου πλέον είναι το κλασικό μοντέλο Theta με ίσα βάρη.

Συνοψίζοντας, καταλήγουμε στο ότι το καλύτερο επίπεδο σφαλμάτων επιτυγχάνεται για όλες τις μεθόδους με την ενσωμάτωση της επίδρασης των επίσημων αργιών και το καλύτερο σφάλμα όλων των περιπτώσεων που μελετήθηκαν επιτυγχάνεται για την απλοϊκή μέθοδο (Naive), με την επίδραση όλων των αργιών που αναφέρθηκαν προηγουμένως.

ΚΕΦΑΛΑΙΟ 8: ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΡΟΕΚΤΑΣΕΙΣ

8.1 Συνολικά αποτελέσματα εργασίας

Η παρούσα μελέτη αποτελείται από δύο μέρη. Στο πρώτο μέρος, μέλημά μας είναι η σύγκριση απλών μεθόδων πρόβλεψης, με συμπερίληψη τουλάχιστον της μίας εκ των δύο βασικών εποχικότητων που παρουσιάζουν οι χρονοσειρές ζήτησης φορτίου ηλεκτρικής ενέργειας (ημερήσια και εβδομαδιαία εποχικότητα). Ενώ στο δεύτερο μέρος της μελέτης, στόχος μας είναι η ενσωμάτωση, στις προβλέψεις, των ειδικών ημερών που παρατηρείται ότι επηρεάζουν τη χρονοσειρά ζήτησης φορτίου και ο εντοπισμός της καλύτερης εξ' αυτών.

Εκτός από την επιλογή των βασικών μεθόδων που εξετάζονται και των ειδικών ημερών που ενσωματώνονται στις προβλέψεις (στο δεύτερο μέρος της εργασίας), απαραίτητη είναι η εύρεση των παραγόντων που πρέπει να ληφθούν υπόψη κατά την τροποποίηση των μοντέλων πρόβλεψης και η καλύτερη ένταξή τους στα μοντέλα. Στην περίπτωση μας, μεγάλη σημασία έχουν οι εποχικότητες της χρονοσειράς, καθώς δεν πρόκειται για μία μόνο, αλλά εμφανίζεται έντονη διπλή εποχικότητα. Επίσης, έχοντας διαθέσιμο μεγάλο πλήθος ωριαίων δεδομένων, τίθεται το ερώτημα αν είναι όλα αυτά τα δεδομένα σημαντικά για την παραγωγή όλων των προβλέψεων και η απάντηση στο ερώτημα δίνεται με εφαρμογή της μεθόδου κυλιόμενης πρόβλεψης, όπως εξηγείται και στη συνέχεια. Η γενική διαδικασία που ακολουθείται για τους παράγοντες που μόλις σημειώθηκαν αναφέρεται αμέσως μετά.

➤ Ένα γενικό συμπέρασμα της εργασίας, είναι η σημασία της συμπερίληψης της εποχικότητας της χρονοσειράς κατά την παραγωγή προβλέψεων σε όλα τα στάδια της εργασίας. Έπειτα από εξέταση τόσο της επίδρασης της ημερήσιας όσο και της εβδομαδιαίας εποχικότητας (με υπολογισμό του Κεντρικού Κινητού Μέσου Όρου, ΚΚΜΟ(24) και ΚΚΜΟ(168)), καταλήγουμε στο συμπέρασμα ότι η δεύτερη περίπτωση οδηγεί σε πολύ πιο αντιπροσωπευτικά αποτελέσματα και εμφανώς μειωμένα σφάλματα (σε όλους τους πίνακες σφαλμάτων μπορεί κανείς να το δει). Ο συνδυασμός των δύο εποχικότητων σε ένα μοντέλο παράγει λιγότερο ικανοποιητικές προβλέψεις. Αυτό προκύπτει από τις βασικές μεθόδους (Naive, SES, Holt, Damped, Theta) με την τροποποίηση που έχουμε εφαρμόσει πάνω τους, αλλά και από την μέθοδο του Taylor για διπλή εποχικότητα, η οποία παράγει τα χειρότερα αποτελέσματα στο σύνολο των περιπτώσεων.

➤ Ακόμη, πολύ σημαντική είναι και η εφαρμογή της μεθόδου κυλιόμενης πρόβλεψης σε κάθε βήμα της μελέτης. Επιλέγεται ένα σταθερό χρονικό διάστημα δεδομένων, τα οποία χρησιμοποιούνται για τον καθορισμό των συντελεστών των μοντέλων σε κάθε βήμα πρόβλεψης (για κάθε νέα εικοσιτετράωρη πρόβλεψη δηλαδή). Το μήκος του χρονικού διαστήματος κύλισης των δεδομένων καθορίζεται για κάθε μέθοδο ύστερα από έλεγχο αρκετών τιμών, ξεκινώντας από τον ελάχιστο δυνατό αριθμό εβδομάδων και αυξάνοντας την τιμή του μήκους του μέχρι την εύρεση του ελάχιστου σφάλματος. Έτσι, με τη χρήση της μεθόδου κυλιόμενης πρόβλεψης (rolling forecasting) μπορεί η εταιρεία παραγωγής ηλεκτρικής ενέργειας να ανταποκρίνεται άμεσα και αποτελεσματικά στις μεταβολές του περιβάλλοντος.

➤ Βάσει των στόχων της εργασίας, παρουσιάζονται στη συνέχεια (πίνακες 31 έως και 36) ταξινομημένες οι καλύτερες επαναλήψεις για όλες τις περιπτώσεις προβλέψεων για κάθε μέθοδο, σύμφωνα με τα σφάλματα MAPE και sMAPE.

Naive	MAPE(%)	sMAPE(%)
Impact of all BH	3,030%	3,056%
Impact of official BH	3,074%	3,123%
Impact of official BH + strikes	3,105%	3,155%
Impact of strikes	3,456%	3,432%
Original forecast	3,461%	3,437%

Πίνακας 31. Βέλτιστες εκτελέσεις της μεθόδου Naive

SES	MAPE(%)	sMAPE(%)
Impact of official BH	3,234%	3,232%
Impact of official BH + strikes	3,237%	3,236%
Impact of all BH	3,264%	3,253%
Impact of strikes	3,663%	3,572%
Original forecast	3,696%	3,604%

Πίνακας 32. Βέλτιστες εκτελέσεις της μεθόδου SES

Holt	MAPE(%)	sMAPE(%)
Impact of official BH	3,251%	3,248%
Impact of official BH + strikes	3,253%	3,252%
Impact of all BH	3,279%	3,268%
Impact of strikes	3,677%	3,586%
Original forecast	3,926%	3,819%

Πίνακας 33. Βέλτιστες εκτελέσεις της μεθόδου Holt

Theta	MAPE(%)	sMAPE(%)
Impact of all BH	3,279%	3,286%
Impact of official BH	3,287%	3,307%
Impact of official BH + strikes	3,292%	3,312%
Impact of strikes	3,681%	3,621%
Original forecast	3,712%	3,652%

Πίνακας 34. Βέλτιστες εκτελέσεις της μεθόδου Theta

Damped	MAPE(%)	sMAPE(%)
Impact of official BH + strikes	3,293%	3,287%
Impact of official BH	3,296%	3,288%
Impact of all BH	3,329%	3,312%
Impact of strikes	3,724%	3,627%
Original forecast	3,763%	3,664%

Πίνακας 35. Βέλτιστες εκτελέσεις της μεθόδου Damped

Taylor	MAPE(%)	sMAPE(%)
Impact of official BH + strikes	4,129%	4,168%
Impact of official BH	4,133%	4,180%
Impact of all BH	4,143%	4,181%
Impact of strikes	4,577%	4,563%
Original forecast	4,603%	4,589%

Πίνακας 36. Βέλτιστες εκτελέσεις της μεθόδου του Taylor

Αυτό που εύκολα αντιλαμβάνεται κανείς είναι ότι οι μέθοδοι SES και Holt έχουν την ίδια ταξινόμηση στις περιπτώσεις βέλτιστων σφαλμάτων τους. Συγκεκριμένα, το βέλτιστο σφάλμα επιτυγχάνεται με την ενσωμάτωση των επίσημων αργιών και ακολουθούν οι περιπτώσεις ενσωμάτωσης όλων των αργιών (επίσημων και μη), του συνδυασμού αργιών και απεργιών, μόνο των απεργιών και τέλος η αρχική πρόβλεψη χωρίς την επίδραση ειδικών ημερών. Όμοια, οι μέθοδοι Naive και Theta παρουσιάζουν την ίδια πορεία στην επίτευξη του βέλτιστου σφάλματος, σημειώνοντας το μεγαλύτερο σφάλμα κατά την αρχική πρόβλεψη (χωρίς ενσωμάτωση ειδικών ημερών) και καταλήγοντας στο μικρότερο σφάλμα για την περίπτωση ενσωμάτωσης όλων των απεργιών, επίσημων και μη. Τέλος, η μέθοδος Damped και η προσαρμογή του Taylor στη μέθοδο Holt-Winters για διπλή εποχικότητα έχουν διαφορετικά ταξινομημένες τις περιπτώσεις προβλέψεων σε σχέση με τις υπόλοιπες μεθόδους, πετυχαίνοντας βέλτιστη απόδοση για ενσωμάτωση του συνδυασμού επίσημων αργιών με απεργίες. Τα χειρότερα αποτελέσματα και πάλι σημειώνονται στην αρχική πρόβλεψη, όπου δεν εισάγονται οι ειδικές ημέρες. Ενώ για όλες τις μεθόδους, η δεύτερη επίδοση από το τέλος επιτυγχάνεται για την περίπτωση ενσωμάτωσης των απεργιών.

Σημαντικό συμπέρασμα από το πρώτο κιάλας μέρος της μελέτης είναι ότι οι απλούστερες μέθοδοι παράγουν τις καλύτερες προβλέψεις, σε αντίθεση με τις περισσότερο πολύπλοκες και απαιτητικές σε δεδομένα μεθόδους. Αυτή η θέση έχει υποστηριχθεί και παλαιότερα σε παρόμοιες μελέτες [Taylor J.W. 2003], παρόλα αυτά στην παρούσα εργασία αποδεικνύεται (πίνακας 6) ότι η πιο απλή από όλες τις μεθόδους, Naive, η οποία χρησιμοποιείται κατά κόρον ως μέθοδος benchmark, παράγει τις βέλτιστες προβλέψεις (στα πλαίσια πάντα των μεθόδων της εργασίας αυτής). Το συμπέρασμα αυτό επιβεβαιώνεται και στη συνέχεια, καθώς η Naive αποτελεί την πλέον αποδοτική μέθοδο σε όλες τις περιπτώσεις ενσωμάτωσης των ειδικών ημερών, όπως παρουσιάζεται συνοπτικά στον πίνακα 30. Ενδέχεται αρχικά η κατάταξη των μοντέλων πρόβλεψης να προκαλεί εντύπωση. Λαμβάνοντας όμως υπόψη τη μεγάλη διακύμανση των τιμών ζήτησης ηλεκτρικού φορτίου και το πολύ απλό μοντέλο της Naive, στο οποίο εφαρμόζεται η κυλιόμενη πρόβλεψη σε συνδυασμό με την επίδραση της εποχικότητας (και στο δεύτερο μέρος και η επίδραση των ειδικών ημερών), δεν είναι τόσο παράδοξα τα αποτελέσματα. Για μικρό μήκος κύλισης, δίνεται πολύ μεγάλη έμφαση στο

πρόσφατο παρελθόν, ενώ αγνοούνται παλαιότερες τιμές. Μάλιστα, η απλοϊκή μέθοδος δεν επεμβαίνει στην πορεία της χρονοσειράς όπως οι υπόλοιπες μέθοδοι. Έτσι, οδηγεί σε προβλέψεις βασισμένες στην τάση των τελευταίων εβδομάδων πριν από την ημέρα που επιθυμείται να προβλεφθεί και χωρίς μεγάλη αλλαγή του μοτίβου των διακυμάνσεων που παρουσιάζουν οι ως τώρα παρατηρήσεις. Άρα και σε καλύτερα αποτελέσματα.

Την απλοϊκή μέθοδο, Naive, ακολουθούν οι μέθοδοι SES και Holt, οι οποίες παράγουν τα δεύτερα και τρίτα καλύτερα αποτελέσματα σε κάθε περίπτωση πρόβλεψης για ενσωμάτωση των ειδικών ημερών. Ενώ κατά την αρχική παραγωγή πρόβλεψης, τρίτη κατατάσσεται η κλασική μέθοδος Theta, όπως αναφέρθηκε στο κεφάλαιο 7 (παράγραφος 7.4). Με βάση τα παραπάνω, συμπεραίνεται ότι οι μέθοδοι εκθετικής εξομάλυνσης είναι ιδιαίτερα αποδοτικές μέθοδοι, με ικανοποιητικές προβλέψεις και κρίνεται απόλυτα λογική η τάση των ερευνητών για τροποποίηση και αξιοποίησή τους στη βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου (παραδείγματα: [Gould et al., 2008], [Hyndman et al., 2008], [Taylor, 2003]).

Τέλος, η εργασία αυτή απέδειξε ότι οι ειδικές ημέρες επηρεάζουν τις προβλέψεις ζήτησης φορτίου ηλεκτρικής ενέργειας και μάλιστα, σε κάθε περίπτωση τρόπου ενσωμάτωσης των ειδικών ημερών που εξετάστηκε, τα αποτελέσματα ήταν πάντοτε πιο ακριβή από τις αντίστοιχες αρχικές προβλέψεις χωρίς την επίδραση αυτών. Ακόμη, όπως αναφέρθηκε προηγουμένως, δεν παράγουν όλες οι μέθοδοι τις βέλτιστες προβλέψεις τους για την ίδια περίπτωση ενσωμάτωσης ειδικών ημερών, αλλά γίνεται ταξινόμηση των περιπτώσεων ανά μέθοδο και των μεθόδων ανά περίπτωση ενσωμάτωσης ειδικών ημερών.

Επομένως επιτεύχθηκαν οι στόχοι της παρούσας μελέτης, εντοπίστηκε η καλύτερη μέθοδος ανάμεσα σε ένα σύνολο βασικών μεθόδων που ελέγχθηκαν, έχοντας πρώτα τροποποιηθεί κατάλληλα για να υποστηρίζουν το είδος της χρονοσειράς που εξετάζεται και η περίπτωση ενσωμάτωσης ειδικών ημερών με το καλύτερο επίπεδο σφαλμάτων σύμφωνα με το Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (MAPE).

8.2 Προεκτάσεις

Η βραχυπρόθεσμη πρόβλεψη ζήτησης φορτίου ηλεκτρισμού αποτελεί σημαντικό αντικείμενο του τομέα των προβλέψεων, καθώς είναι ιδιαίτερα κρίσιμη για τις εταιρείες παραγωγής ηλεκτρικής ενέργειας. Έτσι, η μείωση του σφάλματος πρόβλεψης στο ελάχιστο δυνατό είναι επιθυμητή ώστε τα ΣΗΕ να μη διατρέχουν κινδύνους, οι εταιρείες παραγωγής να μην έχουν μεγάλες απώλειες λόγω εσφαλμένων προβλέψεων (οικονομικής φύσης κυρίως) και για την πιο ακριβή πρόβλεψη της τιμής της ηλεκτρικής ενέργειας.

Όπως αναφέρθηκε στην αρχή της εργασίας, υπάρχουν περισσότεροι του ενός παράγοντες που επηρεάζουν τη ζήτηση φορτίου και είναι επιθυμητή η ενσωμάτωσή τους στα μοντέλα παραγωγής προβλέψεων. Μέχρι τώρα εξετάστηκε η συμπερίληψη των ειδικών ημερών στις προβλέψεις, με διαφορετικούς συνδυασμούς και για διάφορες ειδικές ημέρες. Κατά τη μελέτη της παραμέτρου αυτής, χρησιμοποιήθηκαν ακόμη μέθοδοι για τη συμπερίληψη της εποχικότητας και την κύλιση επί των διαθέσιμων δεδομένων που αξιοποιούνται. Θεωρούμε ότι έχουμε εφαρμόσει αποτελεσματικά τις μεθόδους που μόλις αναφέρθηκαν, παρόλα αυτά, υπάρχουν πάντοτε περιθώρια βελτίωσης.

Οι παράγοντες που επηρεάζουν τη ζήτηση ηλεκτρικής ενέργειας είναι εποχικοί, καιρικοί και οικονομικοί. Το μεγαλύτερο ρόλο στη διαμόρφωση της ζήτησης φορτίου ηλεκτρικής ενέργειας έχουν, έπειτα από τους εποχικούς παράγοντες, οι καιρικές συνθήκες της εκάστοτε

περιοχής που εξετάζεται. Αναφέραμε και στο κεφάλαιο 2 (παράγραφος 2.2) ότι η ενσωμάτωση των καιρικών παραγόντων είναι πολύ ενδιαφέρουσα στην πρόβλεψη ζήτησης φορτίου, έχει όμως ορισμένα έντονα μειονεκτήματα η μελέτη της. Οι καιρικές συνθήκες μιας περιοχής δεν είναι διαθέσιμες τη στιγμή που παράγονται οι προβλέψεις ζήτησης του ηλεκτρικού φορτίου και πρέπει να περιοριστεί κανείς στη χρήση προβλέψεων των καιρικών συνθηκών για το ίδιο χρονικό διάστημα που πρόκειται να γίνει η πρόβλεψη της ζήτησης φορτίου. Ακόμη, οι προβλεπόμενες τιμές των καιρικών συνθηκών δε δίνονται σε ωριαία βάση, αλλά συνηθίζεται να παράγονται προβλέψεις για τέσσερα τμήματα της ημέρας, όπως συμβαίνει στο www.meteo.gr. Ακόμη, η πρόβλεψη των καιρικών συνθηκών δε γίνεται καθολικά για μια ευρεία περιοχή, αλλά για τα επιμέρους τμήματά της. Χρειάζεται λοιπόν, να μελετηθεί προσεκτικά ο τρόπος συμπερίληψης των καιρικών συνθηκών στην παραγωγή προβλέψεων. Σημειώνουμε τέλος ότι βασικός και με διαφορά καιρικός παράγοντας όσον αφορά στην επιρροή της ζήτησης ηλεκτρικού φορτίου είναι η θερμοκρασία. Επομένως, στόχος είναι αρχικά η ενσωμάτωση αυτής στις προβλέψεις και μελλοντικά και των λοιπών καιρικών παραγόντων (υγρασία, άνεμοι, ηλιοφάνεια, κλπ).

Σχετικά με τη θερμοκρασία, γίνεται να χρησιμοποιηθούν ακέραιες όλες οι προβλεπόμενες τιμές της κατά την ενσωμάτωσή της στην πρόβλεψη ζητούμενου φορτίου ηλεκτρικής ενέργειας. Χωρίζονται επομένως οι εικοσιτέσσερις ωριαίες προβλέψεις ζήτησης φορτίου σε τέσσερις ομάδες, σύμφωνα με τα εξάωρα παραγόμενης πρόβλεψης της θερμοκρασίας (θεωρούμε ότι αξιοποιούνται οι προβλέψεις από το www.meteo.gr). Ένας ακόμη τρόπος, είναι να υπολογίζεται η μέση τιμή όλων των προβλέψεων της θερμοκρασίας και να ενσωματώνεται σε κάθε νέα ημερήσια πρόβλεψη (24 ωριαίων τιμών). Η αξιοποίηση του μέσου όρου της προβλεπόμενης θερμοκρασίας με μια πρώτη ματιά, θα οδηγήσει σε προβλέψεις με ικανοποιητικό ημερήσιο επίπεδο τιμών, χάνοντας όμως τις λεπτομέρειες των διακυμάνσεων ενδιάμεσα στην ημέρα. Τέλος, όπως στην παρούσα εργασία εξετάστηκε η ενσωμάτωση των ειδικών ημερών με βάση την ποσοστιαία μεταβολή τους από την προηγούμενη ημέρα ή εβδομάδα ή και συνδυασμό αυτών των δύο, έτσι και στην περίπτωση της ενσωμάτωσης της θερμοκρασίας μπορεί να εξεταστεί ο συνδυασμός της προβλεπόμενης θερμοκρασίας για την ημέρα που μας ενδιαφέρει, με την καταγεγραμμένη τιμή (ή τιμές) της θερμοκρασίας για την ίδια ημέρα της εβδομάδας, μία όμως εβδομάδα νωρίτερα, αλλά και μία ημέρα νωρίτερα. Μελέτες για την επίδραση των καιρικών συνθηκών στις προβλέψεις ζήτησης φορτίου ηλεκτρικής ενέργειας έχουν γίνει και συνεχίζουν να διενεργούνται (παραδείγματα: [Taylor and Buizza, 2003] και [Dordonnat et al., 2008]). Χρησιμοποιούνται βέβαια περισσότερο πολύπλοκα μοντέλα, καθώς από τη φύση τους έχουν δημιουργηθεί για τη συμπερίληψη πολλών ανεξάρτητων μεταβλητών για την παραγωγή της εξαρτημένης, εδώ ζήτησης φορτίου ηλεκτρισμού.

Η ενσωμάτωση ορισμένων από τους οικονομικούς παράγοντες αποτελεί ενδιαφέρον για προέκταση, καθώς το εξωτερικό περιβάλλον επηρεάζει σημαντικά τη συμπεριφορά του καταναλωτικού κοινού. Ένα πολύ καλό οικονομικό επίπεδο και περίοδοι εφορίας μιας περιοχής οδηγούν σε αυξημένη ζήτηση φορτίου και το αντίστροφο. Μπορεί να παρατηρηθεί ανάπτυξη στα τεχνολογικά επιτεύγματα όπως το ηλεκτρικό αυτοκίνητο, που συντελούν στην αύξηση της ζήτησης ηλεκτρικού φορτίου (αφού σε τέτοιες περιόδους δίνονται περισσότερες ευκαιρίες και χρηματοδοτήσεις για έρευνα). Ακόμη, οι οικονομικοί παράγοντες προκαλούν και δημογραφικές μεταβολές. Για παράδειγμα, σε περιόδους ανάπτυξης παρατηρείται αύξηση του αστικού πληθυσμού και αντίστοιχη μείωση του αγροτικού, γεγονός που επηρεάζει τη ζήτηση ηλεκτρικής ενέργειας κατά τόπους (και κατ' επέκταση στο σύνολο της περιοχής). Στην παρούσα εργασία δε μελετήθηκαν οι οικονομικοί παράγοντες, καθώς είναι ιδιαίτερα πολύπλοκη υπόθεση η ενσωμάτωσή τους στα μοντέλα πρόβλεψης, αποτελούν ωστόσο εν δυνάμει εξωτερικές μεταβλητές μοντέλων.

Το χαρακτηριστικό της ζήτησης φορτίου ηλεκτρικής ενέργειας, που αντιλαμβάνεται κανείς σε όλη την παρούσα μελέτη (αλλά και σε όλες τις μελέτες σχετικά με αυτό το θέμα), είναι ο πολύ σημαντικός ρόλος των εποχικότητων των χρονοσειρών ζήτησης. Η ύπαρξη περισσότερων από μιας εποχικότητας, οδηγεί στην έντονη αναζήτηση του βέλτιστου τρόπου ένταξης των πιο έντονων, έστω, από αυτές στα υπάρχοντα μοντέλα βραχυπρόθεσμης πρόβλεψης φορτίου. Ήδη έχουν αναπτυχθεί αξιολογικά μοντέλα πρόβλεψης με ενσωμάτωση δύο εποχικότητων, όπως η προσέγγιση του μοντέλου Holt-Winters από τον Taylor. Χρειάζονται όμως ακόμη περισσότερες βελτιώσεις των μεθοδολογιών αυτών, καθώς δεν καταφέρνουν πάντοτε να παράγουν ακριβείς προβλέψεις. Το μοτίβο που ακολουθήθηκε στη μελέτη μας για τη συμπερίληψη των εποχικότητων της χρονοσειράς ζήτησης φορτίου είναι ιδιαίτερα απλοϊκό, ενσωματώνοντας τις δύο εποχικότητες σε δύο ξεχωριστά στάδια, μόλις με χρήση του Κεντρικού Κινητού Μέσου Όρου μήκους $n=24$ και $n=168$ αντίστοιχα για ημερήσια και εβδομαδιαία εποχικότητα. Επομένως, μία ακόμη δυνατότητα για προέκταση της παρούσας μελέτης βρίσκεται στη διαχείριση των εποχικότητων της χρονοσειράς. Μπορεί να χρησιμοποιηθούν και επιπλέον δείκτες εξομάλυνσης ανάλογα με το είδος της εξεταζόμενης ημέρας μέσα στην εβδομάδα (για παράδειγμα: Δευτέρα, Τρίτη).

Τέλος, πολύ μεγάλη σημασία έχει η χρονοσειρά επί της οποίας εφαρμόζονται τα μελετώμενα μοντέλα. Θεωρούμε πολύ ενδιαφέρουσα την προοπτική εφαρμογής της μεθοδολογίας μας σε δεδομένα διαφορετικών εταιρειών, τόσο μεγαλύτερων όσο και μικρότερων από την εταιρεία που μελετήσαμε, λόγω της διαφορετικής εμβέλειάς τους. Ακόμη, σημασία έχει η μελέτη χρονοσειρών διαφορετικού μήκους. Στην περίπτωση μας έχουμε στοιχεία ωριαίας ζήτησης για διάστημα 973 ημερών, ενώ θα μπορούσαμε να έχουμε δεδομένα μεγαλύτερης χρονικής διάρκειας (για παράδειγμα πέντε έτη), δοσμένα ως τετράωρες παρατηρήσεις ή μικρότερης διάρκειας (για παράδειγμα δύο έτη) και σε βάση μισής ώρας. Επίσης, πολύ βασικό είναι τα δεδομένα ζήτησης να παρέχονται με βάση τη γεωγραφική κατανομή της ευρύτερης περιοχής που μελετάται. Αυτό θα συμβάλλει άμεσα στην πολύ καλή αξιοποίηση και των πληροφοριών της εκάστοτε υπο-περιοχής σχετικά με τις καιρικές συνθήκες που επικρατούν σε αυτή, διευκολύνοντας την ενσωμάτωση των καιρικών συνθηκών στη μεθοδολογία μας. Τέλος, κρίνεται σκόπιμη η εξέταση της ακρίβειας των μοντέλων που αναπτύχθηκαν στην παρούσα εργασία για ακόμη μεγαλύτερο εύρος ορίζοντα πρόβλεψης.

Βιβλιογραφία

- [1] Assimakopoulos V., Nikolopoulos N. (2000) 'The theta model: a decomposition approach to forecasting', *International Journal of Forecasting*, Vol. 16, No. 4, pp. 521-530
- [2] Box G.E.P., G.M. Jenkins (1970) 'Time series analysis: Forecasting and control', *San Francisco: Holden-Day*
- [3] Charytoniuk W., Chen M.S., Van Olinda P. (1998) 'Nonparametric Regression based short-term load forecasting', *IEEE Transactions on Power Systems*, Vol. 13, No.3, August 1998, pp. 725 – 730
- [4] Cho M.Y., Hwang J.C., Chen C.S. (1995) 'Customer Short-Term Load Forecasting by using ARIMA Transfer Function Model', *Proceedings of the International Conference on Energy Management and Power Delivery*, 1:317–322
- [5] Collopy F., Armstrong J. S. (1992) 'Rule-Based Forecasting: Development and Validation of an Expert Systems Approach to Combining Time Series Extrapolations', *Management Science*, 38 (10), 1394-1414
- [6] De Livera A.M., Hyndman R.J. (2009) 'Forecasting time series with complex seasonal patterns using exponential smoothing', *Working paper 15/09, Department of Econometrics and Business Statistics, Monash University*
- [7] Dehdashti A.S., Tudor J.R., Smith M.C. (1982) 'Forecasting of hourly load by pattern recognition – a deterministic approach', *IEEE Transactions on Power Apparatus and Systems*, Vol. 101, No.9, September, pp. 3290 – 3294
- [8] Dordonnat V., Koopman S.J., Ooms M., Dessertaine A., Collet J. (2008) 'An hourly periodic state space model for modelling French national electricity load', *International Journal of Forecasting*, 24, 566-587
- [9] Gould P.G., Koehler A.B., Ord J.K., Snyder R.D., Hyndman R.J., Vahid-Araghi F. (2008) 'Forecasting time series with multiple seasonal patterns', *European Journal of Operational Research*, 191, 207-222
- [10] Guo X.C., Chen Z.Y., Ge H.W., Liang Y.C. (2004) 'Short-term load forecasting using neural network with principal component analysis', *Proceedings of the International Conference on Machinery Learning and Cybernetics*, 6: 3365–3369
- [11] Haidda T. (1994) 'Regression based peak load using a transformation technique', *IEEE Transactions on Power Systems*, Vol. 9, No.4, November, pp.1788-1794
- [12] Hyndman R.J., Koehler A.B., (2006) 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, 21, 679-688
- [13] Hyndman R.J., Koehler A.B., Ord J.K. and Snyder R.D. (2008) 'Forecasting with exponential smoothing: the state space approach', *Springer-Verlag, Berlin*

- [14] Khotanzad A., Hwang R.C., Abaye A., Maratukulam D. (1995) 'An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities', *IEEE Transactions on Power Systems*, Vol. 10, No.3, August, pp. 1716 – 1722
- [15] Khotanzad A., Afkhami-Rohani R., Maratukulam D. (1998) 'ANNSTLF – Artificial neural network short-term load forecaster – Generation three', *IEEE Transactions on Power Systems*, Vol. 13, No.4, November, pp. 1413 – 1422
- [16] Koehler A.B. (2001) 'The asymmetry of the sAPE measure and other comments on the M3-competition', *International Journal of Forecasting*, 17, 537– 584
- [17] Makridakis S. (1993) 'Accuracy measures: Theoretical and practical concerns', *International Journal of Forecasting*, 9, 527– 529
- [18] Papalexopoulos A.D., Hesterberg T.C. (1990) 'A regression-based approach to short-term system load forecasting', *IEEE Transactions on Power Systems*, Vol. 5, No.4, November, pp. 1535– 1547
- [19] Pegels C.C. (1969) 'Exponential forecasting: some new variations', *Management Science*, 15, 311-315
- [20] Taylor J.W. (2003) 'Exponential smoothing with a damped multiplicative trend', *International Journal of Forecasting*, 19, 715-725
- [21] Taylor J.W. (2003) 'Short-term electricity demand forecasting using double seasonal exponential smoothing', *Journal of the Operational Research Society*, Vol. 54, pp. 779-805
- [22] Taylor J.W., Buizza R. (2003) 'Using weather ensemble predictions in electricity demand forecasting', *International Journal of Forecasting*, 19, 57-70
- [23] Taylor J.W., Menezes L.M., McSharry P.E. (2006) 'A comparison of univariate methods for forecasting electricity demand up to a day ahead', *International Journal of Forecasting*, 22, 1-16
- [24] Taylor J.W., (2008) 'An evaluation of methods for very short-term load forecasting using minute-by-minute British data', *International Journal of Forecasting*, 24, 645-658
- [25] Taylor J.W., McSharry (2008) 'Short-Term Load Forecasting Methods: An Evaluation Based on European Data', *IEEE Transactions on Power Systems*, 22, 2213-2219
- [26] Taylor J.W., Snyder R.D. (2012) 'Forecasting Intraday Time Series with Multiple Seasonal Cycles Using Parsimonious Seasonal Exponential Smoothing', *Omega*, Vol. 40, issue 6, 748-757
- [27] Taylor J.W. (2010) 'Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles', *International Journal of Forecasting*, 26, 627-646
- [28] Πετρόπουλος Φ., Ασημακόπουλος Β. (2011) 'Επιχειρησιακές Προβλέψεις', Αθήνα: Εκδόσεις Συμμετρία

Appendix

		Naive (deseasonal.&rolling)				SES (deseasonal.&rolling)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,624	6,651	3,780	3,772	6,060	5,866	3,842	3,883
	k=2	6,624	6,651	3,780	3,772	6,017	5,852	4,090	4,121
	k=3	6,624	6,651	3,780	3,772	6,052	5,890	4,123	4,167
	k=4	6,624	6,651	3,780	3,772	6,082	5,916	4,092	4,123
	k=5	6,624	6,651	3,780	3,772	6,133	5,963	4,089	4,134
	k=6	6,624	6,651	3,780	3,772	6,211	6,033	4,111	4,147
	k=7	6,624	6,651	3,780	3,772	6,284	6,100	4,171	4,206
	k=8	6,624	6,651	3,780	3,772	6,384	6,190	4,301	4,343
	k=9	6,624	6,651	3,780	3,772	6,468	6,265	4,397	4,406
	k=10	6,624	6,651	3,780	3,772	6,542	6,333	4,465	4,517
e=168	k=2	4,823	4,777	2,743	2,746	4,116	4,062	2,350	2,353
	k=3	4,423	4,396	2,727	2,735	3,915	3,850	2,361	2,371
	k=4	3,893	3,870	2,462	2,461	3,822	3,736	2,360	2,364
	k=5	3,664	3,642	2,315	2,313	3,696	3,604	2,317	2,320
	k=6	3,533	3,511	2,202	2,203	3,774	3,674	2,394	2,396
	k=7	3,512	3,487	2,137	2,140	3,914	3,808	2,527	2,531
	k=8	3,474	3,450	2,137	2,136	4,033	3,919	2,636	2,634
	k=9	3,461	3,437	2,103	2,107	4,157	4,035	2,708	2,714
	k=10	3,473	3,449	2,122	2,121	4,281	4,152	2,822	2,819

Πίνακας Α1. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με εφαρμογή rolling και ενσωμάτωση εποχικότητας (α)

		Holt (deseasonal.&rolling)				Damped (deseasonal.&rolling)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	10,633	10,278	6,501	6,624	6,257	6,043	3,765	3,817
	k=2	9,346	9,071	5,980	6,048	6,143	5,957	4,110	4,144
	k=3	8,794	8,471	5,537	5,622	6,150	5,967	4,156	4,170
	k=4	8,296	8,133	5,396	5,442	6,157	5,971	4,029	4,071
	k=5	7,667	7,615	5,149	5,196	6,188	5,997	3,998	4,033
	k=6	7,335	7,252	4,960	5,001	6,257	6,056	3,976	3,997
	k=7	7,335	7,252	4,753	4,783	6,317	6,108	4,031	4,043
	k=8	7,124	6,994	4,648	4,711	6,407	6,186	4,094	4,111
	k=9	6,863	6,725	4,546	4,581	6,475	6,242	4,119	4,144
	k=10	6,741	6,586	4,542	4,591	6,522	6,283	4,158	4,187
e=168	k=15	6,874	6,630	4,832	4,880	6,808	6,535	4,557	4,598
	k=2	4,198	4,145	2,349	2,352	4,181	4,122	2,303	2,299
	k=3	3,944	3,880	2,391	2,395	3,978	3,913	2,361	2,365
	k=4	4,753	4,583	2,624	2,632	3,898	3,805	2,404	2,405
	k=5	4,202	4,087	2,503	2,505	3,763	3,664	2,321	2,322
	k=6	4,125	3,980	2,468	2,474	3,831	3,723	2,391	2,383
	k=7	3,926	3,819	2,542	2,546	3,952	3,838	2,562	2,580
	k=8	4,044	3,928	2,660	2,658	4,063	3,939	2,667	2,679
k=9	4,167	4,043	2,709	2,720	4,187	4,053	2,726	2,734	
k=10	4,289	4,158	2,840	2,839	4,291	4,151	2,827	2,832	

Πίνακας Α2. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με εφαρμογή rolling και ενσωμάτωση εποχικότητας (β)

		Theta (deseasonal.&rolling)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,540	6,397	4,194	4,240
	k=2	6,308	6,192	4,356	4,388
	k=3	6,319	6,209	4,407	4,419
	k=4	6,352	6,243	4,464	4,462
	k=5	6,401	6,292	4,492	4,539
	k=6	6,470	6,360	4,532	4,583
	k=7	6,534	6,423	4,653	4,696
	k=8	6,619	6,506	4,772	4,823
	k=9	6,694	6,580	4,862	4,914
	k=10	6,767	6,652	4,982	5,002
e=168	k=15	7,078	6,958	5,238	5,309
	k=2	4,274	4,229	2,434	2,443
	k=3	4,107	4,060	2,566	2,569
	k=4	3,825	3,768	2,386	2,403
	k=5	3,712	3,652	2,377	2,393
	k=6	3,768	3,704	2,385	2,386
	k=7	3,903	3,837	2,533	2,543
	k=8	4,007	3,939	2,645	2,647
k=9	4,120	4,048	2,719	2,721	
k=10	4,243	4,170	2,839	2,845	

Πίνακας A3. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με εφαρμογή rolling και ενσωμάτωση εποχικότητας (γ)

	Taylor(deseasonal.&rolling)			
	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	5,184	5,209	2,798	2,805
k=4	4,716	4,654	2,620	2,624
k=5	4,863	4,888	2,791	2,786
k=6	5,254	5,332	3,063	3,061
k=7	4,877	4,795	2,923	2,917
k=8	4,603	4,589	2,759	2,759
k=9	5,243	5,175	3,070	3,082
k=10	5,711	5,922	3,402	3,406
k=15	5,489	5,442	3,634	3,613
k=20	5,070	5,048	3,516	3,531
k=25	5,110	5,071	3,423	3,417
k=30	4,794	4,701	3,237	3,240
k=35	5,322	5,275	3,743	3,761
k=40	5,331	5,280	3,588	3,585
k=45	5,049	5,009	3,359	3,370
k=50	5,557	5,491	3,739	3,749
k=55	5,212	5,228	3,599	3,602

Πίνακας A4. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με εφαρμογή rolling

		Naive previous day (official BH)				SES previous day (official BH)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,197	6,299	3,548	3,532	5,568	5,468	3,633	3,652
	k=2	6,197	6,299	3,548	3,532	5,522	5,451	3,842	3,878
	k=3	6,197	6,299	3,548	3,532	5,545	5,480	3,885	3,912
	k=4	6,197	6,299	3,548	3,532	5,567	5,501	3,856	3,894
	k=5	6,197	6,299	3,548	3,532	5,608	5,539	3,865	3,902
	k=6	6,197	6,299	3,548	3,532	5,672	5,598	3,919	3,937
	k=7	6,197	6,299	3,548	3,532	5,737	5,658	3,956	3,979
	k=8	6,197	6,299	3,548	3,532	5,820	5,736	4,088	4,123
	k=9	6,197	6,299	3,548	3,532	5,892	5,802	4,136	4,165
	k=10	6,197	6,299	3,548	3,532	5,961	5,867	4,206	4,238
	k=15	6,197	6,299	3,548	3,532	6,267	6,147	4,559	4,589
e=168	k=2	4,383	4,409	2,602	2,603	3,598	3,634	2,219	2,219
	k=3	4,035	4,078	2,584	2,584	3,461	3,483	2,252	2,257
	k=4	3,493	3,543	2,352	2,347	3,363	3,365	2,259	2,261
	k=5	3,271	3,322	2,246	2,243	3,234	3,232	2,221	2,223
	k=6	3,144	3,194	2,104	2,112	3,286	3,282	2,272	2,274
	k=7	3,121	3,170	2,068	2,064	3,418	3,408	2,428	2,423
	k=8	3,131	3,175	2,070	2,071	3,523	3,509	2,523	2,522
	k=9	3,074	3,123	2,027	2,039	3,627	3,610	2,562	2,566
	k=10	3,087	3,137	2,040	2,039	3,748	3,727	2,690	2,695

Πίνακας Α5. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών (α)

		Holt previous day (official BH)				Damped previous day (official BH)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	9,773	9,591	6,064	6,151	5,751	5,631	3,574	3,591
	k=2	8,559	8,465	5,603	5,675	5,623	5,534	3,918	3,949
	k=3	7,977	7,843	5,197	5,265	5,619	5,535	3,925	3,939
	k=4	7,699	7,649	5,083	5,121	5,622	5,537	3,813	3,830
	k=5	5,607	5,539	3,843	3,859	5,653	5,564	3,801	3,831
	k=6	5,667	5,593	3,881	3,890	5,711	5,614	3,744	3,771
	k=7	5,733	5,654	3,936	3,961	5,763	5,659	3,808	3,829
	k=8	5,816	5,730	4,065	4,089	5,839	5,727	3,887	3,902
	k=9	6,287	6,262	4,292	4,339	5,893	5,774	3,909	3,947
	k=10	6,161	6,120	4,260	4,298	5,937	5,812	3,927	3,957
	k=15	6,263	6,141	4,578	4,608	6,191	6,039	4,345	4,376
e=168	k=2	3,673	3,711	2,247	2,251	3,650	3,689	2,196	2,199
	k=3	3,493	3,516	2,263	2,267	3,512	3,538	2,251	2,261
	k=4	3,384	3,387	2,253	2,259	3,429	3,426	2,325	2,331
	k=5	3,251	3,248	2,236	2,232	3,296	3,288	2,241	2,239
	k=6	3,299	3,293	2,295	2,301	3,337	3,326	2,290	2,298
	k=7	3,431	3,421	2,442	2,429	3,453	3,436	2,440	2,450
	k=8	3,534	3,518	2,525	2,529	3,549	3,526	2,544	2,553
	k=9	3,636	3,618	2,600	2,593	3,653	3,625	2,605	2,601
	k=10	3,751	3,728	2,689	2,692	3,753	3,721	2,705	2,704

Πίνακας Α6. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών (β)

		THETA previous day (official BH)			
		MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,047	5,996	3,966	3,994
	k=2	5,823	5,796	4,103	4,126
	k=3	5,829	5,809	4,160	4,191
	k=4	5,856	5,839	4,196	4,218
	k=5	5,899	5,883	4,225	4,237
	k=6	5,959	5,944	4,276	4,298
	k=7	6,017	6,002	4,409	4,449
	k=8	6,090	6,076	4,527	4,568
	k=9	6,157	6,143	4,601	4,656
	k=10	6,226	6,212	4,653	4,686
	k=15	6,515	6,501	4,972	5,013
e=168	k=2	3,786	3,819	2,319	2,305
	k=3	3,681	3,711	2,450	2,454
	k=4	3,407	3,429	2,300	2,296
	k=5	3,287	3,307	2,287	2,282
	k=6	3,326	3,347	2,290	2,287
	k=7	3,452	3,472	2,429	2,426
	k=8	3,548	3,567	2,503	2,505
	k=9	3,644	3,664	2,585	2,598
	k=10	3,760	3,781	2,681	2,691

Πίνακας A7. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών (γ)

	TAYLOR previous day (official BH)			
	MAPE(%)	sMAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	4,722	4,833	2,669	2,663
k=4	4,173	4,217	2,506	2,501
k=5	4,422	4,537	2,663	2,660
k=6	4,779	4,948	2,893	2,884
k=7	4,431	4,438	2,785	2,786
k=8	4,133	4,180	2,617	2,608
k=9	4,739	4,774	2,913	2,911
k=10	5,243	5,540	3,242	3,237
k=15	5,116	5,136	3,426	3,437
k=20	4,673	4,727	3,368	3,376
k=25	4,634	4,685	3,254	3,256
k=30	4,304	4,302	3,078	3,092
k=35	4,903	4,935	3,613	3,593
k=40	4,876	4,909	3,406	3,421
k=45	4,533	4,590	3,203	3,202
k=50	4,954	4,986	3,558	3,547
k=55	4,819	4,910	3,460	3,462

Πίνακας A8. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με επίδραση των επίσημων αργιών

		Naive previous day (all BH)				SES previous day (all BH)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,171	6,251	3,499	3,497	5,557	5,447	3,592	3,609
	k=2	6,171	6,251	3,499	3,497	5,534	5,455	3,801	3,837
	k=3	6,171	6,251	3,499	3,497	5,562	5,489	3,843	3,888
	k=4	6,171	6,251	3,499	3,497	5,588	5,514	3,833	3,846
	k=5	6,171	6,251	3,499	3,497	5,625	5,548	3,842	3,847
	k=6	6,171	6,251	3,499	3,497	5,689	5,609	3,879	3,884
	k=7	6,171	6,251	3,499	3,497	5,750	5,665	3,908	3,936
	k=8	6,171	6,251	3,499	3,497	5,826	5,737	4,042	4,064
	k=9	6,171	6,251	3,499	3,497	5,894	5,800	4,083	4,114
	k=10	6,171	6,251	3,499	3,497	5,962	5,864	4,138	4,172
e=168	k=2	4,339	4,341	2,562	2,566	3,637	3,664	2,199	2,202
	k=3	3,987	4,005	2,541	2,537	3,493	3,504	2,224	2,221
	k=4	3,452	3,478	2,312	2,314	3,387	3,381	2,241	2,238
	k=5	3,231	3,258	2,228	2,227	3,264	3,253	2,189	2,197
	k=6	3,100	3,127	2,094	2,096	3,317	3,304	2,251	2,249
	k=7	3,075	3,099	2,042	2,053	3,450	3,433	2,411	2,400
	k=8	3,048	3,073	2,034	2,042	3,556	3,535	2,498	2,504
	k=9	3,030	3,056	2,009	2,008	3,658	3,635	2,540	2,545
k=10	3,044	3,069	2,019	2,024	3,771	3,744	2,653	2,657	

Πίνακας Α9. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών (α)

		Holt previous day (all BH)				Damped previous day (all BH)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	9,606	9,467	6,013	6,075	5,732	5,603	3,528	3,554
	k=2	8,540	8,448	5,536	5,627	5,635	5,540	3,886	3,898
	k=3	7,947	7,817	5,152	5,211	5,636	5,548	3,869	3,898
	k=4	7,664	7,619	5,011	5,066	5,647	5,556	3,777	3,808
	k=5	5,622	5,546	3,801	3,808	5,671	5,576	3,764	3,792
	k=6	5,684	5,603	3,833	3,848	5,728	5,626	3,703	3,729
	k=7	5,745	5,660	3,900	3,910	5,774	5,666	3,752	3,773
	k=8	5,820	5,730	4,016	4,048	5,842	5,726	3,825	3,845
	k=9	6,290	6,261	4,236	4,276	5,893	5,770	3,870	3,895
	k=10	6,160	6,115	4,203	4,222	5,935	5,807	3,872	3,895
e=168	k=2	6,250	6,127	4,502	4,549	6,177	6,024	4,300	4,317
	k=3	3,708	3,737	2,233	2,237	3,684	3,715	2,182	2,184
	k=4	3,523	3,536	2,239	2,237	3,545	3,563	2,216	2,217
	k=5	3,407	3,401	2,241	2,238	3,455	3,445	2,297	2,311
	k=6	3,279	3,268	2,217	2,212	3,329	3,312	2,208	2,208
	k=7	3,328	3,315	2,264	2,266	3,366	3,348	2,264	2,263
	k=8	3,462	3,444	2,413	2,414	3,483	3,460	2,416	2,426
	k=9	3,566	3,544	2,501	2,511	3,579	3,550	2,523	2,526
k=10	3,667	3,643	2,575	2,569	3,680	3,647	2,575	2,575	
k=10	3,778	3,749	2,665	2,664	3,777	3,739	2,669	2,683	

Πίνακας Α10. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών (β)

		Theta previous day (all BH)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,021	5,954	3,931	3,949
	k=2	5,814	5,775	4,040	4,059
	k=3	5,824	5,792	4,115	4,150
	k=4	5,854	5,825	4,143	4,156
	k=5	5,895	5,867	4,159	4,178
	k=6	5,955	5,928	4,226	4,246
	k=7	6,012	5,985	4,351	4,399
	k=8	6,082	6,057	4,460	4,493
	k=9	6,148	6,124	4,522	4,564
	k=10	6,218	6,195	4,596	4,659
	k=15	6,505	6,483	4,925	4,967
e=168	k=2	3,787	3,806	2,266	2,277
	k=3	3,675	3,690	2,409	2,416
	k=4	3,397	3,405	2,270	2,270
	k=5	3,279	3,286	2,254	2,248
	k=6	3,321	3,329	2,258	2,261
	k=7	3,446	3,453	2,379	2,381
	k=8	3,548	3,555	2,466	2,470
	k=9	3,645	3,655	2,540	2,545
	k=10	3,821	3,824	2,655	2,661

Πίνακας A11. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών (γ)

	Taylor previous day (all BH)			
	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	4,700	4,793	2,619	2,612
k=4	4,185	4,227	2,481	2,469
k=5	4,394	4,492	2,631	2,628
k=6	4,776	4,937	2,838	2,848
k=7	4,428	4,425	2,738	2,739
k=8	4,143	4,181	2,577	2,569
k=9	4,698	4,721	2,875	2,856
k=10	5,142	5,411	3,194	3,197
k=15	5,115	5,120	3,383	3,384
k=20	4,704	4,748	3,335	3,354
k=25	4,632	4,671	3,184	3,181
k=30	4,298	4,276	3,050	3,046
k=35	4,900	4,923	3,545	3,551
k=40	4,853	4,872	3,349	3,349
k=45	4,465	4,503	3,154	3,139
k=50	4,942	4,957	3,516	3,515
k=55	4,763	4,831	3,404	3,390

Πίνακας A12. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με επίδραση όλων των αργιών

		Naive previous day+ week (strikes)				SES previous day+ week (strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,586	6,616	3,701	3,710	6,012	5,823	3,788	3,818
	k=2	6,586	6,616	3,701	3,710	5,973	5,811	4,042	4,077
	k=3	6,586	6,616	3,701	3,710	6,003	5,844	4,058	4,091
	k=4	6,586	6,616	3,701	3,710	6,035	5,873	4,046	4,078
	k=5	6,586	6,616	3,701	3,710	6,082	5,915	4,050	4,088
	k=6	6,586	6,616	3,701	3,710	6,156	5,983	4,072	4,099
	k=7	6,586	6,616	3,701	3,710	6,225	6,045	4,123	4,154
	k=8	6,586	6,616	3,701	3,710	6,319	6,129	4,233	4,282
	k=9	6,586	6,616	3,701	3,710	6,397	6,200	4,288	4,317
	k=10	6,586	6,616	3,701	3,710	6,467	6,263	4,371	4,380
e=168	k=2	4,814	4,769	2,735	2,729	4,090	4,036	2,303	2,303
	k=3	4,377	4,351	2,675	2,667	3,876	3,812	2,327	2,338
	k=4	3,870	3,848	2,426	2,424	3,778	3,693	2,289	2,295
	k=5	3,646	3,624	2,287	2,292	3,663	3,572	2,272	2,274
	k=6	3,520	3,497	2,179	2,179	3,741	3,642	2,350	2,352
	k=7	3,502	3,477	2,130	2,131	3,880	3,775	2,482	2,476
	k=8	3,466	3,442	2,123	2,131	3,997	3,884	2,598	2,601
	k=9	3,456	3,432	2,097	2,103	4,123	4,002	2,656	2,667
	k=10	3,463	3,439	2,106	2,106	4,240	4,113	2,769	2,781

Πίνακας Α13. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση του MO όλων των απεργιών (α)

		Holt previous day+ week (strikes)				Damped previous day+ week (strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	10,493	10,136	6,301	6,385	6,199	5,989	3,700	3,729
	k=2	9,293	9,020	5,917	5,980	6,096	5,914	4,074	4,116
	k=3	8,687	8,370	5,432	5,500	6,098	5,919	4,098	4,107
	k=4	8,216	8,060	5,319	5,348	6,112	5,929	3,979	4,020
	k=5	6,082	5,916	4,037	4,075	6,137	5,950	3,954	3,978
	k=6	6,153	5,979	4,066	4,102	6,203	6,006	3,914	3,946
	k=7	6,222	6,041	4,117	4,148	6,257	6,052	3,960	4,010
	k=8	6,315	6,124	4,217	4,266	6,343	6,125	4,036	4,061
	k=9	6,793	6,660	4,478	4,498	6,404	6,177	4,061	4,080
	k=10	6,667	6,517	4,440	4,446	6,449	6,214	4,086	4,115
e=168	k=2	4,171	4,119	2,321	2,311	4,152	4,094	2,265	2,263
	k=3	3,905	3,843	2,339	2,350	3,936	3,873	2,322	2,323
	k=4	3,796	3,712	2,287	2,295	3,847	3,756	2,358	2,359
	k=5	3,677	3,586	2,287	2,283	3,724	3,627	2,271	2,274
	k=6	3,752	3,653	2,350	2,356	3,792	3,685	2,337	2,342
	k=7	3,891	3,786	2,491	2,495	3,913	3,801	2,500	2,511
	k=8	4,007	3,893	2,602	2,617	4,024	3,902	2,627	2,630
	k=9	4,132	4,010	2,673	2,677	4,151	4,019	2,685	2,697
	k=10	4,247	4,119	2,777	2,779	4,248	4,111	2,773	2,773

Πίνακας Α14. Σφάλματα MAPE, sMAPE, MdAPE & sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση του MO όλων των απεργιών (β)

		Theta previous day+ week (strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,493	6,354	4,148	4,206
	k=2	6,265	6,152	4,296	4,332
	k=3	6,271	6,164	4,358	4,369
	k=4	6,303	6,197	4,493	4,494
	k=5	6,349	6,243	4,523	4,571
	k=6	6,415	6,308	4,576	4,613
	k=7	6,474	6,367	4,662	4,707
	k=8	6,556	6,446	4,772	4,816
	k=9	6,628	6,517	4,854	4,904
	k=10	6,697	6,585	4,961	4,974
e=168	k=2	4,255	4,212	2,455	2,454
	k=3	4,067	4,022	2,561	2,568
	k=4	3,788	3,732	2,380	2,394
	k=5	3,681	3,621	2,343	2,343
	k=6	3,739	3,676	2,353	2,352
	k=7	3,873	3,808	2,534	2,544
	k=8	3,975	3,907	2,642	2,649
	k=9	4,087	4,016	2,732	2,728
	k=10	4,208	4,135	2,781	2,797

Πίνακας A15. Σφάλματα MAPE, sMAPE, MdAPE & sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση του MO όλων των απεργιών (γ)

	Taylor previous day+ week (strikes)			
	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	5,126	5,153	2,725	2,736
k=4	4,636	4,569	2,559	2,561
k=5	4,815	4,842	2,713	2,706
k=6	5,225	5,304	2,995	2,994
k=7	4,831	4,752	2,865	2,872
k=8	4,577	4,563	2,708	2,703
k=9	5,220	5,153	3,027	3,022
k=10	5,290	5,323	3,288	3,287
k=15	5,431	5,386	3,288	3,287
k=20	5,014	4,991	3,444	3,449
k=25	5,062	5,025	3,350	3,330
k=30	4,727	4,633	3,129	3,138
k=35	5,272	5,226	3,680	3,683
k=40	5,290	5,240	3,516	3,519
k=45	5,021	4,982	3,312	3,302
k=50	5,497	5,433	3,664	3,648
k=55	5,132	5,151	3,496	3,510

Πίνακας A16. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με επίδραση του MO όλων των απεργιών

		Naive previous day (official BH+ strikes)				SES previous day (official BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,195	6,299	3,565	3,545	5,557	5,461	3,665	3,686
	k=2	6,195	6,299	3,565	3,545	5,514	5,446	3,893	3,918
	k=3	6,195	6,299	3,565	3,545	5,533	5,470	3,901	3,948
	k=4	6,195	6,299	3,565	3,545	5,557	5,494	3,893	3,937
	k=5	6,195	6,299	3,565	3,545	5,594	5,528	3,892	3,924
	k=6	6,195	6,299	3,565	3,545	5,654	5,583	3,939	3,971
	k=7	6,195	6,299	3,565	3,545	5,714	5,639	3,982	3,999
	k=8	6,195	6,299	3,565	3,545	5,792	5,711	4,095	4,132
	k=9	6,195	6,299	3,565	3,545	5,858	5,773	4,138	4,163
	k=10	6,195	6,299	3,565	3,545	5,923	5,833	4,195	4,229
e=168	k=2	4,411	4,438	2,642	2,648	3,607	3,645	2,225	2,223
	k=3	4,024	4,070	2,571	2,574	3,458	3,482	2,250	2,255
	k=4	3,507	3,556	2,358	2,353	3,355	3,359	2,244	2,245
	k=5	3,289	3,341	2,252	2,252	3,237	3,236	2,219	2,219
	k=6	3,167	3,217	2,117	2,121	3,290	3,286	2,269	2,272
	k=7	3,147	3,196	2,082	2,084	3,420	3,412	2,429	2,423
	k=8	3,118	3,167	2,085	2,087	3,523	3,510	2,526	2,526
	k=9	3,105	3,155	2,063	2,062	3,629	3,613	2,578	2,577
k=10	3,113	3,163	2,060	2,054	3,739	3,719	2,681	2,692	

Πίνακας A17. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (α)

		Holt previous day (official BH+ strikes)				Damped previous day (official BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	9,670	9,485	5,861	5,927	5,730	5,613	3,589	3,614
	k=2	8,541	8,451	5,596	5,661	5,613	5,527	3,962	4,001
	k=3	7,907	7,779	5,178	5,246	5,604	5,524	3,949	3,975
	k=4	7,655	7,612	5,093	5,129	5,613	5,531	3,841	3,880
	k=5	5,594	5,529	3,885	3,906	5,639	5,554	3,833	3,859
	k=6	5,650	5,580	3,914	3,931	5,693	5,600	3,772	3,804
	k=7	5,711	5,635	3,972	3,990	5,739	5,640	3,835	3,861
	k=8	5,788	5,707	4,076	4,098	5,810	5,703	3,911	3,936
	k=9	6,254	6,234	4,301	4,343	5,859	5,745	3,922	3,960
	k=10	6,123	6,087	4,244	4,289	5,900	5,780	3,934	3,962
e=168	k=15	6,211	6,096	4,578	4,608	6,196	6,046	4,392	4,424
	k=2	3,682	3,721	2,251	2,256	3,658	3,698	2,200	2,200
	k=3	3,490	3,515	2,264	2,268	3,506	3,535	2,258	2,266
	k=4	3,376	3,380	2,245	2,245	3,415	3,413	2,306	2,318
	k=5	3,253	3,252	2,234	2,230	3,293	3,287	2,224	2,228
	k=6	3,302	3,298	2,294	2,300	3,334	3,325	2,282	2,289
	k=7	3,433	3,424	2,437	2,428	3,450	3,435	2,433	2,444
	k=8	3,533	3,519	2,529	2,536	3,546	3,525	2,547	2,554
k=9	3,638	3,621	2,603	2,606	3,653	3,627	2,616	2,612	
k=10	3,746	3,725	2,691	2,701	3,747	3,716	2,709	2,706	

Πίνακας A18. Σφάλματα MAPE, sMAPE, MdAPE & sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (β)

		Theta previous day (official BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,037	5,989	3,997	4,030
	k=2	5,817	5,793	4,127	4,176
	k=3	5,817	5,800	4,195	4,216
	k=4	5,844	5,829	4,232	4,255
	k=5	5,883	5,870	4,261	4,273
	k=6	5,940	5,927	4,290	4,326
	k=7	5,994	5,982	4,427	4,455
	k=8	6,064	6,053	4,531	4,555
	k=9	6,127	6,116	4,585	4,635
	k=10	6,193	6,182	4,631	4,675
	k=15	6,470	6,460	4,939	4,980
e=168	k=2	3,804	3,838	2,329	2,314
	k=3	3,678	3,710	2,444	2,447
	k=4	3,407	3,429	2,292	2,289
	k=5	3,292	3,312	2,286	2,280
	k=6	3,334	3,355	2,294	2,294
	k=7	3,459	3,479	2,431	2,427
	k=8	3,552	3,572	2,501	2,505
	k=9	3,648	3,669	2,593	2,604
	k=10	3,761	3,782	2,683	2,695

Πίνακας A19. Σφάλματα MAPE, sMAPE, MdAPE & sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση των επίσημων αργιών και του MO όλων των απεργιών (γ)

	Taylor previous day (official BH+ strikes)			
	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	4,700	4,813	2,648	2,656
k=4	4,129	4,168	2,487	2,475
k=5	4,411	4,527	2,649	2,649
k=6	4,786	4,956	2,890	2,884
k=7	4,422	4,432	2,788	2,794
k=8	4,143	4,190	2,630	2,621
k=9	4,752	4,787	2,921	2,929
k=10	4,858	4,978	3,199	3,206
k=15	5,094	5,117	3,419	3,430
k=20	4,653	4,706	3,333	3,347
k=25	4,623	4,676	3,250	3,253
k=30	4,274	4,271	3,053	3,052
k=35	5,077	5,076	3,645	3,651
k=40	5,120	5,108	3,485	3,476
k=45	4,788	4,799	3,285	3,284
k=50	5,322	5,311	3,668	3,663
k=55	4,958	5,012	3,505	3,511

Πίνακας A20. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με επίδραση των επίσημων αργιών και του MO όλων των απεργιών

		Naive previous day (all BH+ strikes)				SES previous day (all BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,416	6,479	3,668	3,665	5,853	5,694	3,750	3,781
	k=2	6,416	6,479	3,668	3,665	5,815	5,684	3,984	4,004
	k=3	6,416	6,479	3,668	3,665	5,841	5,715	4,019	4,044
	k=4	6,416	6,479	3,668	3,665	5,862	5,734	3,999	4,025
	k=5	6,416	6,479	3,668	3,665	5,903	5,772	3,991	4,012
	k=6	6,416	6,479	3,668	3,665	5,974	5,836	4,020	4,043
	k=7	6,416	6,479	3,668	3,665	6,040	5,896	4,069	4,096
	k=8	6,416	6,479	3,668	3,665	6,131	5,978	4,183	4,211
	k=9	6,416	6,479	3,668	3,665	6,214	6,052	4,240	4,280
	k=10	6,416	6,479	3,668	3,665	6,293	6,124	4,340	4,357
	k=15	6,416	6,479	3,668	3,665	6,624	6,422	4,685	4,731
e=168	k=2	4,676	4,657	2,684	2,690	3,934	3,913	2,302	2,303
	k=3	4,264	4,265	2,686	2,677	3,717	3,683	2,333	2,339
	k=4	3,700	3,712	2,425	2,423	3,587	3,542	2,308	2,310
	k=5	3,471	3,483	2,284	2,289	3,451	3,400	2,278	2,278
	k=6	3,356	3,367	2,176	2,178	3,526	3,469	2,360	2,361
	k=7	3,329	3,338	2,116	2,115	3,662	3,599	2,479	2,471
	k=8	3,291	3,301	2,109	2,111	3,774	3,705	2,584	2,588
	k=9	3,276	3,286	2,082	2,088	3,896	3,819	2,643	2,645
	k=10	3,288	3,298	2,090	2,093	4,025	3,941	2,763	2,774

Πίνακας Α21. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών και των απεργιών (α)

		Holt previous day (all BH+ strikes)				Damped previous day (all BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	10,409	10,092	6,337	6,437	6,051	5,871	3,669	3,712
	k=2	9,142	8,901	5,809	5,879	5,942	5,789	4,036	4,071
	k=3	8,585	8,297	5,400	5,459	5,942	5,794	4,059	4,077
	k=4	8,076	7,951	5,262	5,284	5,940	5,790	3,936	3,966
	k=5	5,901	5,770	3,979	4,020	5,958	5,805	3,909	3,932
	k=6	5,968	5,831	4,016	4,041	6,020	5,858	3,858	3,899
	k=7	6,034	5,890	4,049	4,089	6,071	5,901	3,912	3,945
	k=8	6,869	6,781	4,505	4,564	6,151	5,971	3,979	3,998
	k=9	6,607	6,511	4,427	4,445	6,218	6,026	4,006	4,017
	k=10	6,491	6,376	4,386	4,405	6,272	6,072	4,043	4,081
	k=15	6,619	6,415	4,704	4,722	6,555	6,321	4,447	4,472
e=168	k=2	4,015	3,995	2,324	2,313	3,983	3,960	2,267	2,272
	k=3	3,750	3,717	2,344	2,352	3,774	3,741	2,319	2,323
	k=4	3,607	3,563	2,317	2,311	3,655	3,604	2,370	2,376
	k=5	3,465	3,415	2,289	2,285	3,509	3,454	2,289	2,290
	k=6	3,537	3,480	2,363	2,359	3,574	3,512	2,351	2,357
	k=7	3,673	3,610	2,487	2,492	3,693	3,624	2,502	2,511
	k=8	3,784	3,713	2,593	2,595	3,797	3,719	2,613	2,619
	k=9	3,905	3,826	2,667	2,670	3,919	3,833	2,675	2,684
	k=10	4,081	3,991	2,800	2,804	4,030	3,936	2,773	2,774

Πίνακας Α22. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών και των απεργιών (β)

		Theta previous day (all BH+ strikes)			
		MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
e=24	k=1	6,310	6,205	4,095	4,137
	k=2	6,093	6,012	4,228	4,273
	k=3	6,098	6,025	4,296	4,313
	k=4	6,124	6,054	4,341	4,377
	k=5	6,166	6,097	4,363	4,402
	k=6	6,230	6,160	4,424	4,473
	k=7	6,289	6,219	4,537	4,594
	k=8	6,368	6,295	4,666	4,705
	k=9	6,442	6,369	4,761	4,803
	k=10	6,518	6,443	4,823	4,866
	k=15	6,823	6,743	5,094	5,151
e=168	k=2	4,111	4,096	2,400	2,397
	k=3	3,925	3,907	2,525	2,527
	k=4	3,608	3,587	2,346	2,349
	k=5	3,485	3,462	2,344	2,343
	k=6	3,540	3,516	2,348	2,348
	k=7	3,671	3,644	2,490	2,486
	k=8	3,771	3,743	2,569	2,576
	k=9	3,881	3,850	2,664	2,653
	k=10	4,007	3,974	2,770	2,771

Πίνακας A23. Σφάλματα MAPE, sMAPE, MdAPE & sMdAPE των μεθόδων Naive, SES, Damped, Holt και Theta με επίδραση όλων των αργιών και των απεργιών (γ)

	Taylor previous day (all BH+ strikes)			
	MAPE(%)	MAPE(%)	MdAPE(%)	sMdAPE(%)
k=3	4,953	5,011	2,731	2,740
k=4	4,433	4,429	2,576	2,568
k=5	4,620	4,690	2,724	2,719
k=6	5,047	5,166	3,004	2,996
k=7	4,655	4,619	2,845	2,847
k=8	4,351	4,384	2,687	2,694
k=9	4,963	4,942	2,977	2,975
k=10	5,411	5,668	3,296	3,297
k=15	5,267	5,257	3,498	3,496
k=20	4,870	4,876	3,440	3,444
k=25	4,900	4,900	3,339	3,327
k=30	4,540	4,475	3,140	3,147
k=35	5,077	5,076	3,645	3,651
k=40	5,120	5,108	3,485	3,476
k=45	4,788	4,799	3,285	3,284
k=50	5,322	5,311	3,668	3,663
k=55	4,958	5,012	3,505	3,511

Πίνακας A24. Σφάλματα MAPE, sMAPE, MdAPE και sMdAPE για το μοντέλο Holt-Winters για διπλή εποχικότητα, με επίδραση όλων των αργιών και των απεργιών