



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΦΥΣΙΚΗΣ
ΤΟΜΕΑΣ ΦΥΣΙΚΗΣ ΠΕΡΙΒΑΛΛΟΝΤΟΣ-ΜΕΤΕΩΡΟΛΟΓΙΑΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:

ΒΕΛΤΙΩΣΗ ΤΗΣ ΠΡΟΓΝΩΣΗΣ ΤΟΥ ΑΝΕΜΟΥ ΜΕ ΤΗΝ ΕΦΑΡΜΟΓΗ ΒΑΥΕΣΙΑΝ ΚΑΛΜΑΝ ΦΙΛΤΡΩΝ

ΧΡΗΣΤΟΣ ΓΚΑΖΙΚΑΣ

ΑΡΙΘΜΟΣ ΜΗΤΡΩΟΥ: 201643

ΓΕΩΡΓΙΟΣ ΚΑΛΛΟΣ (ΕΠΙΒΛΕΠΩΝ), ΚΑΘΗΓΗΤΗΣ Ε.Κ.Π.Α
ΜΑΡΙΑ ΤΟΜΠΡΟΥ, ΚΑΘΗΓΗΤΡΙΑ Ε.Κ.Π.Α
ΕΛΕΝΑ ΦΛΟΚΑ, ΚΑΘΗΓΗΤΡΙΑ Ε.Κ.Π.Α

ΦΕΒΡΟΥΑΡΙΟΣ 2019

ΑΘΗΝΑ

ΠΡΟΛΟΓΟΣ

Η εργασία αυτή εκπονήθηκε στο πλαίσιο της διπλωματικής εργασίας του μεταπτυχιακού προγράμματος, «Φυσικής Περιβάλλοντος και Μετεωρολογίας», του Τμήματος Φυσικής του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών στον Τομέα της Φυσικής Περιβάλλοντος – Μετεωρολογίας.

Χρήστος Γκάζικας,

Φεβρουάριος 2019,

Αθήνα

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον Καθηγητή του Τομέα «Περιβάλλοντος-Μετεωρολογίας» κ. Γ. Κάλλο, τον Καθηγητή της Σχολής Ναυτικών Δοκίμων κ. Γ. Γαλάνη και την Ομάδα Ατμοσφαιρικών Μοντέλων και Πρόγνωσης Καιρού και ιδιαίτερα την Δρ. Χ. Καλογέρη και τον Χ. Σταθόπουλο για την βοήθειά τους στην χρήση των μοντέλων. Επίσης, θα ήθελα να ευχαριστήσω την Καθηγήτρια του Τομέα «Περιβάλλοντος-Μετεωρολογίας» κ. Ε. Φλόκα για τις χρήσιμες συστάσεις και τις παρατηρήσεις της.

Χρήστος Γκάζικας

Φεβρουάριος, 2019,

Αθήνα

Περίληψη

Τα αιολικά πάρκα για τη λειτουργία τους απαιτούν ακριβέστερες προγνώσεις του ανέμου για την εκτίμηση και πρόγνωση της παραγόμενης ενέργειας. Τα πρωτογενή αποτελέσματα των ατμοσφαιρικών μοντέλων (NWP) εμπεριέχουν μια σειρά σφαλμάτων από παράγοντες όπως οι αρχικές και οριακές συνθήκες, η δομή και η παραμετροποίηση του μοντέλου, η οριζόντια ανάλυση, η εξομάλυνση της ορογραφίας και των άλλων φυσιογραφικών χαρακτηριστικών.

Στην παρούσα εργασία χρησιμοποιούνται τα συνδυαστικά Bayesian Kalman φίλτρα προκειμένου να βελτιωθεί η πρόγνωση του ανέμου κοντά στην επιφάνεια του εδάφους. Τα συγκεκριμένα φίλτρα χρησιμοποιούν μη γραμμικές εξισώσεις σε συνδυασμό με τη συμπερασματολογία του Bayes.

Η εφαρμογή των φίλτρων έγινε για δύο διαφορετικές περιοχές αιολικών πάρκων της Κρήτης. Οι διορθώσεις έγιναν πάνω σε δεδομένα πρόγνωσης από το ατμοσφαιρικό μοντέλο υψηλής ανάλυσης, RAMS.

Μελετήθηκαν οι παράμετροι των φίλτρων και οι συνθήκες κάτω από τις οποίες αυτά εμφανίζουν τη μέγιστη αποδοτικότητα στη βελτίωση της πρόγνωσης του ανέμου. Δόθηκε έμφαση στην ανάδειξη της βελτίωσης αυτής, σε σχέση με τη χαμηλή υπολογιστική ισχύ η οποία απαιτούνταν. Ιδιαίτερη προσοχή δόθηκε στους χρονικούς ορίζοντες αξιοποίησης του φίλτρου, από όπου προέκυψε ότι στους πολύ βραχυπρόθεσμους τα αποτελέσματα ήταν πολύ καλά. Αποδείχθηκε ότι η διαθεσιμότητα μετρήσεων και αποτελεσμάτων του μετεωρολογικού μοντέλου με μικρότερα χρονικά βήματα βοηθάει σημαντικά στην επίτευξη καλύτερων ακόμη αποτελεσμάτων. Η δεκάλεπτη δειγματοληψία αποδείχθηκε η βέλτιστη.

Kalman filters, Bayesian inference, Post-processing, Numerical atmospheric modeling

Περιεχόμενα

ΠΕΡΙΛΗΨΗ.....	4
ΠΕΡΙΕΧΟΜΕΝΑ.....	5
ΕΙΣΑΓΩΓΗ	7
1.1. Στατιστική μετεπεξεργασία (statistical postprocessing)	7
1.2. Model Output Statistics (MOS)	8
1.3. Τεχνικές Machine Learning - Νευρωνικά δίκτυα	9
1.4. Στατιστική μετεπεξεργασία σε πιθανοτικά μοντέλα πρόγνωσης καιρού (Post-Processing of Ensemble Forecasts)	10
1.5. Analog Ensemble (AnEn)	11
1.6. Bayesian Model Averaging.....	11
1.7. Φίλτρα Kalman	12
1.7.1 Διόρθωση της παραμέτρου της ηλιακής ακτινοβολίας και εφαρμογές σε φωτοβολταϊκά	12
1.7.2 Βελτίωση των αποτελεσμάτων πρόγνωσης ταχύτητας του ανέμου από NWP μοντέλα	13
1.7.3 Εφαρμογή φίλτρων Kalman σε αποτελέσματα κυματικών αριθμητικών μοντέλων.....	14
1.7.4 Εφαρμογή φίλτρων Kalman στην πρόγνωση της παραμέτρου της θερμοκρασίας κοντά στην επιφάνεια του εδάφους.....	15
1.8. Στοιχεία θεωρίας Bayes- Στατιστική συμπερασματολογία.....	16
1.9. Συμπερασματολογία κατά Bayes	16
1.10. Η a-priori κατανομή (Prior Distribution).....	17
1.11. Βασικά χαρακτηριστικά της προσέγγισης κατά Bayes.....	17
1.12 Επάρκεια	18
1.13. Η Αρχή της Πιθανοφάνειας.....	18
1.14. Το Θεώρημα του Bayes (διακριτή περίπτωση).....	19
1.15. Το Θεώρημα του Bayes (περίπτωση κατανομών).....	19
1.16. Τα βασικά στάδια της θεωρίας του Bayes	20
1.17. Στατιστική του ανέμου-Γενικά	20
1.18. Κατανομή Weibull.....	20
1.19. Λογαριθμοκανονική κατανομή (Lognormal distribution).....	22
1.20. Γάμμα κατανομή	23
1.21. Στατιστική των ακραίων τιμών του ανέμου – Κατανομή Gumbel	23
1.21.1 Γενικευμένη κατανομή ακραίων τιμών	24
ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ	26
ΜΕΘΟΔΟΛΟΓΙΑ- ΛΕΛΟΜΕΝΑ	28
3.1. Γενικά	28
3.2. Περιγραφή του ατμοσφαιρικού μοντέλου πρόγνωσης καιρού RAMS ...	29
3.3. Το φίλτρο Kalman	31

3.3.1 Γενικά για το Kalman	31
3.3.2 Το Bayesian Kalman φίλτρο	33
3.3.3 Το μοντέλο της κατανομής Weibull	35
3.4. Μεγέθη ελέγχου απόδοσης	36
3.5. Περιοχές μελέτης	38
3.6. Δεδομένα Πρόγνωσης.....	38
ΕΦΑΡΜΟΓΗ ΤΩΝ ΦΙΛΤΡΩΝ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	40
4.1. Ανάλυση δοκιμών	40
4.2. Εξέταση βέλτιστης τάξης πολωνύμου	41
4.3. Εξέταση των βημάτων πρόγνωσης και της περιόδου «εκπαίδευσης»	45
4.3.1 Πάρκο «Αλκυόνη», Κίσαμος Χανίων	45
4.3.2 Πάρκο «Enercon», Κρυά Σητείας	57
4.3.3 Βέλτιστες τιμές	63
4.4. Συγκεντρωτικά αποτελέσματα για τα ωριαία δεδομένα	64
4.5. Εξέταση της απόδοσης του φίλτρου με πυκνότερο σετ δεδομένων	66
ΣΥΜΠΕΡΑΣΜΑΤΑ	73
ΒΙΒΛΙΟΓΡΑΦΙΑ	75

ΚΕΦΑΛΑΙΟ 1ο

ΕΙΣΑΓΩΓΗ

1.1. Στατιστική μετεπεξεργασία (statistical postprocessing)

Στην προσπάθεια να μοντελοποιηθεί η ατμόσφαιρα και να γίνει ακριβής πρόγνωση του καιρού και των μετεωρολογικών παραμέτρων, το κύριο εργαλείο που έχει αναπτυχθεί, αποτελούν τα αριθμητικά μοντέλα πρόγνωσης καιρού (NWP). Τα αριθμητικά μοντέλα από την απλή μορφή που είχαν αρχικά, εξελίχθηκαν σημαντικά, και έφτασαν έως την αλληλεπίδραση μεταξύ διαφορετικών τύπων μοντέλων, τα συζευγμένα μοντέλα (fully coupled models), τα οποία έχουν τη δυνατότητα να προσομοιώνουν καταστάσεις τόσο στην ατμόσφαιρα όσο και στον ωκεανό, συνδυαζόμενα μεταξύ τους και για διάφορες κλίμακες φαινομένων. Ωστόσο, παρά τη σημαντική εξέλιξη στην ανάπτυξη των μοντέλων αυτών, παραμένουν σημαντικά προβλήματα που αφορούν την ακρίβεια της πρόγνωσης.

Τα πρωτογενή αποτελέσματα των αριθμητικών μοντέλων εμπεριέχουν σφάλματα, τα οποία προέρχονται από διάφορες πηγές, όπως είναι τα δεδομένα που εισάγονται ως αρχικές συνθήκες, οι οριακές συνθήκες, η δομή και η παραμετροποίηση του μοντέλου. Για παράδειγμα, η οριζόντια ανάλυση ενός προγνωστικού μοντέλου σε συνδυασμό με την εξομάλυνση την ορογραφίας και των χαρακτηριστικών της επιφάνειας μπορούν να οδηγήσουν σε ελλιπή απεικόνιση των τοπικών φαινομένων της ροής. Κρίνεται απαραίτητο επομένως, να εφαρμοστούν κάποιες μέθοδοι προκειμένου να ποσοτικοποιηθούν και να μειωθούν τα σφάλματα αυτά.

Η στατιστική μετεπεξεργασία (statistical postprocessing) είναι μία μέθοδος βελτίωσης των αποτελεσμάτων που δίνουν τα αριθμητικά μοντέλα πρόγνωσης καιρού. Χρησιμοποιείται προκειμένου να διορθώσει τα συστηματικά σφάλματα ενός μοντέλου, όπως το bias και το RMSE, και μπορεί να λάβει υπόψιν του τοπικές επιδράσεις που δε φαίνονται πλήρως στις λύσεις του gridbox. Εκτός αυτού, μπορεί να χρησιμοποιηθεί και για την πρόγνωση παραμέτρων οι οποίες δεν υπολογίζονται γενικώς από ένα NWP μοντέλο.

Οι statistical postprocessing τεχνικές στη γενικότητα χρησιμοποιούν ένα αντιπροσωπευτικό στατιστικό σετ δεδομένων, τόσο από αποτελέσματα μοντέλου όσο και από παρατηρήσεις.

Υπάρχουν πολλές στατιστικές τεχνικές μετεπεξεργασίας. Διαφορετικές μέθοδοι μετεπεξεργασίας έχουν αναπτυχθεί για τις μετεωρολογικές προγνώσεις διαφορετικών παραμέτρων (π.χ. άνεμος), καθώς η κάθε παράμετρος έχει και διαφορετικές στατιστικές ιδιότητες. Κάποιες από τις πιο συνηθισμένες είναι η «κλασσική» κατηγορία των παλινδρομήσεων, όπως η πολλαπλή γραμμική παλινδρόμηση (multiple linear regression), το logistic regression, τα οποία προσπαθούν να βρουν το καλύτερο «ταίριασμα» από ένα σετ δεδομένων που χρησιμοποιείται προκειμένου να «εκπαιδεύσει» τον αλγόριθμο.

1.2. Model Output Statistics (MOS)

Μία ευρέως διαδεδομένη μέθοδος στατιστικής μετεπεξεργασίας είναι η προσέγγιση των MOS. Ουσιαστικά πρόκειται για μία ολόκληρη κατηγορία τεχνικών, οι οποίες χρησιμοποιούνται προκειμένου να βελτιώσουν την ικανότητα πρόγνωσης των αριθμητικών μοντέλων καιρού. Οι μέθοδοι των MOS περιλαμβάνουν ένα ευρύ φάσμα τεχνικών, από απλές διορθώσεις του bias έως σύνθετες παραμετρικές και μη παραμετρικές στατιστικές μεθόδους (Gangopadhyay et al. 2004). Κάποιες από αυτές περιλαμβάνουν, μεταξύ άλλων, εφαρμογές με πολλαπλή γραμμική παλινδρόμηση (Clark and Hay 2004), λογιστική παλινδρόμηση (logistic regression) (Hamill et al. 2004, Wilks 2009) και παλινδρόμηση τεταρτημορίων (quantile regression) (Bremnes 2004, Friederichs and Hense 2007).

Ορισμένα από τα πλεονεκτήματα που παρέχουν τα MOS είναι ότι ερμηνεύουν τα αριθμητικά μοντέλα πρόγνωσης βασιζόμενα σε ιστορικά δείγματα και καταφέρνουν κατά κάποιο τρόπο να μιμούνται τη «λογική» των προγνωστικών μοντέλων. Επίσης, μπορούν να προβλέψουν γεγονότα τα οποία ωθούνται από βαρομετρικά συστήματα συνοπτικής κλίμακας. Άλλο ένα πλεονέκτημα των MOS είναι η ικανότητα τους να ποσοτικοποιούν την αβεβαιότητα των αριθμητικών μοντέλων πρόγνωσης. Επιπλέον, λαμβάνουν υπόψη τους την «αλλοίωση» της ικανότητας ενός NWP μοντέλου όταν αυτό αυξάνει την προγνωστική του ανάλυση, όπως και τις τοπικές επιδράσεις, σε κάποιο βαθμό. Τέλος, έχουν τη δυνατότητα να ενσωματώνουν κλιματικά στοιχεία.

Από την άλλη, ορισμένα από τα βασικά τους μειονεκτήματα είναι ότι δεν μπορούν να προγνώσουν γεγονότα τα οποία ωθούνται από φαινόμενα μέσης κλίμακας, καθώς και να λάβουν υπόψιν τους κάποιον παράγοντα υπερεκτίμησης. Επίσης, δεν έχουν την δυνατότητα να διορθώνουν τα συστηματικά σφάλματα που σχετίζονται με το ανάγλυφο ή την συνοπτική κατάσταση. Επιπλέον, αδυνατούν να λαμβάνουν υπόψιν τους τις επιδράσεις τοπικής κλίμακας, όπως και τις μεταβολές στην παραμετροποίηση των αριθμητικών μοντέλων. Τέλος, σε περιπτώσεις όπου υπάρχουν ασυνήθιστα κλιματικά φαινόμενα, οι τεχνικές MOS δεν είναι σε θέση να τα συνεκτιμήσουν.

Συνοψίζοντας, οι τεχνικές MOS, ως στατιστικά μοντέλα, παρουσιάζουν μια σειρά από περιορισμούς, όπως το ότι απαιτούν ιστορικά δεδομένα παρατηρήσεων από τα σημεία του grid όπου κάνει τους υπολογισμούς το αριθμητικό μοντέλο. Ενώ ιδιαίτερα σημαντικό είναι το γεγονός ότι οι τεχνικές αυτές δεν ενδείκνυνται για βραχυπρόθεσμες περιόδους πρόγνωσης. Από την άλλη, αποτελούνται από μαθηματικά απλές εξισώσεις, η μη-γραμμικότητα μπορεί να μοντελοποιηθεί έμμεσα, μέσω των μεταβλητών και των μετασχηματισμών των ίδιων των αριθμητικών μοντέλων και καταφέρνουν να μειώσουν ορισμένα συστηματικά σφάλματα.

1.3. Τεχνικές Machine Learning - Νευρωνικά δίκτυα

Ορισμένοι θεωρούν τις Machine Learning τεχνικές ως έναν αποτελεσματικό τρόπο να βελτιώσουν την απόδοση ενός αριθμητικού μοντέλου, ιδιαίτερα για τις περιπτώσεις μεταβλητών, όπου η μη-γραμμικότητα υπερτερεί, όπως για παράδειγμα ο υετός (Oliveira, 2015). Η βασική ιδέα στην οποία στηρίζονται οι νευρωνικοί αλγόριθμοι είναι η μίμηση του τρόπου λειτουργίας των νευρώνων του ανθρωπίνου εγκεφάλου. Αποτελούνται από ένα σύστημα αλληλοσυνδεδεμένων κόμβων, οι οποίοι δέχονται κάποια δεδομένα εισόδου, εκτελούν τις κατά περίπτωση στατιστικές διεργασίες και στη συνέχεια περνούν την πληροφορία σε άλλους «νευρώνες» (κόμβους). Στο τέλος, κάθε κόμβος εκτιμά ένα στατιστικό βάρος για το «σήμα» εισόδου, και με μία διαδικασία «εκμάθησης» εκτελείται με διαφορετικούς επαναληπτικούς κύκλους κάθε φορά, «μαθαίνοντας» με αυτόν τον τρόπο ποιά εκτίμηση ταιριάζει καλύτερα με τα δεδομένα εισόδου.

Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί σε μία πληθώρα εφαρμογών, όπως είναι η ανίχνευση ακραίων καιρικών φαινομένων σε κλιματικά σετ δεδομένων (Liu et al.,

2016), η παραμετροποίηση sub-grid διαδικασιών σε μοντέλα παγκόσμιας κλίμακας (Gentine et al., 2018) αλλά και στην πρόγνωση της ηλιακής ακτινοβολίας (Wang et al., 2012) και ακραίων τιμών ανέμου (Lagerquist et al., 2017).

Ο συνδυασμός των αριθμητικών μοντέλων πρόγνωσης και των Machine Learning τεχνικών στοχεύει στην εκμετάλλευση των θετικών στοιχείων των δύο μεθόδων, προκειμένου να εξαχθούν βελτιωμένα αποτελέσματα πρόγνωσης. Σε αυτές τις περιπτώσεις, εκμεταλλεύονται τον μεγάλο όγκο των δεδομένων που εξάγονται από τα αριθμητικά μοντέλα, καθώς και τη μη-γραμμικότητα που εμπεριέχουν τα αποτελέσματα αυτά, προκειμένου να φτιάξουν «μονοπάτια» (patterns), τα οποία χρησιμοποιούνται για την «εκπαίδευση» των νευρωνικών δικτύων. Δεδομένα παρατηρήσεων επίσης μπορούν να χρησιμοποιηθούν για τον ίδιο λόγο (Oliveira, 2015). Στα θετικά στοιχεία των νευρωνικών δικτύων, αναφέρονται επίσης η ευκολία στη χρήση τους και η ευελιξία των αλγορίθμων τους. Ωστόσο, υπάρχουν πολλά ερωτηματικά σε σχέση με το κατά πόσον αλγόριθμοι τέτοιου τύπου μπορούν να δώσουν αποτελέσματα που ανταποκρίνονται στην πραγματικότητα, για φυσικές παραμέτρους, των οποίων τα αίτια της μεταβολής τους οφείλονται σε δυναμικούς παράγοντες.

1.4. Στατιστική μετεπεξεργασία σε πιθανοτικά μοντέλα πρόγνωσης καιρού (Post-Processing of Ensemble Forecasts)

Τα ensemble μοντέλα έχουν ως στόχο να παρέχουν αποτελέσματα τα οποία εξαρτώνται από την πυκνότητα πιθανότητας πιθανών μελλοντικών καταστάσεων της ατμόσφαιρας (Wilks and Hamill, 2007). Ιδανικά, η πιθανότητα κάθε φαινομένου θα έπρεπε να εκτιμάται απευθείας από την σχετική συχνότητα του φαινομένου, μέσα στο σύνολο. Ωστόσο, κάτι τέτοιο δε συμβαίνει στην πράξη, αφού οι προγνώσεις «μολύνονται» από τα συστηματικά σφάλματα και η διασπορά του συνόλου είναι μικρή (Hamill and Colucci 1997, 1998; Buizza et al. 2005). Στην προσπάθεια να περιοριστούν τα συστηματικά σφάλματα αυτών των μοντέλων αναπτύχθηκαν οι αντίστοιχες post-process τεχνικές. Ορισμένες από τις πιο σημαντικές μεθόδους που έχουν αναπτυχθεί για τα μοντέλα αυτού του τύπου είναι τα Ensemble model output statistics (EMOS), η nonhomogeneous regression (NR), η Nonhomogeneous Gaussian regression (NGR), το Gaussian ensemble dressing (GED), ως υποκατηγορίες (Wilks and Hamill, 2007).

1.5. Analog Ensemble (AnEn)

Το Analog Ensemble (AnEn) αποτελεί μία σχετικά νέα μέθοδο στατιστικής μετεπεξεργασίας, η οποία αρχικά προτάθηκε με σκοπό την πρόγνωση του ανέμου για εφαρμογές στην αιολική ενέργεια (Vannynne et al. 2013, Zhang et al. 2015). Η βασική ιδέα είναι να δημιουργηθεί μία πιθανοτική πρόγνωση βασισμένη σε αμιγώς ντετερμινιστικά δεδομένα (Delle Monache et al. 2013). Με τη μέθοδο αυτή χτίζεται ένα σύνολο προγνώσεων που βασίζονται σε προηγούμενες παρατηρήσεις (analog) για διάφορα φυσικά μεγέθη. Οι παρατηρήσεις αυτές έχουν επιλεγεί κατάλληλα από ένα σετ ιστορικών δεδομένων έτσι ώστε να αντιστοιχούν σε κάποιες κλασσικές (ντετερμινιστικές) προγνώσεις κάποιου αριθμητικού μοντέλου. Το AnEn παράγει μία συνάρτηση πυκνότητας πιθανότητας, η οποία έχει προέλθει από την παραδοχή ότι εάν έχουν βρεθεί τα κατάλληλα analogs, τότε τα σφάλματά τους θα είναι παρόμοια με το σφάλμα της πρόγνωσης που επιχειρείται. Ωστόσο, μη κατάλληλη επιλογή των σωστών analogs μπορεί να οδηγήσει σε πολύ μεγάλη διάδοση σφαλμάτων, κι επομένως σε λάθος προγνωστικά αποτελέσματα.

Οι τεχνικές αυτές έχουν εφαρμοστεί με επιτυχία σε βραχυπρόθεσμες προγνώσεις καιρού (Delle Monache et al. 2013).

1.6. Bayesian Model Averaging

Μία άλλη στατιστική τεχνική είναι το Bayesian Model Averaging (BMA), το οποίο συνδυάζει προγνώσεις από διαφορετικά σετ δεδομένων, προκειμένου να φτιάξει την κατάλληλη κατανομή πιθανοτήτων, εκμεταλλευόμενη την προσέγγιση του Bayes. Το BMA (Leamer 1978, Kass and Raftery 1993, Hoeting et al. 1999) αναπτύχθηκε ως μια μέθοδος όπου συνδυάζει συμπεράσματα και προγνωστικά αποτελέσματα από πολλαπλά στατιστικά μοντέλα και εφαρμόζεται σε μοντέλα τύπου γραμμικής πολλαπλής παλινδρόμησης. Στη συνέχεια επεκτάθηκε η εφαρμογή του και σε ensemble μοντέλα καιρού.

1.7. Φίλτρα Kalman

Μία από τις πιο επιτυχημένες μεθόδους που εφαρμόζονται προκειμένου να ξεπεραστούν οι περιορισμοί που επιβάλλουν τα αριθμητικά μοντέλα σε φαινόμενα μικρής κλίμακας, αποτελούν τα φίλτρα Kalman. Είναι ικανά να περιορίσουν τα συστηματικά σφάλματα των αριθμητικών μοντέλων σε μεγάλο βαθμό. Κάθε μετεωρολογική παράμετρος ωστόσο είναι διαφορετική και χρήζει ξεχωριστής μελέτης κάθε φορά, τόσο για την κλίμακα που εξετάζεται όσο και για το είδος των εξισώσεων που θα χρησιμοποιηθούν, λαμβάνοντας υπόψιν τις διαφορετικές στατιστικές ιδιότητες της κάθε παραμέτρου.

Πρόκειται για άλλη μία κατηγορία στατιστικών μοντέλων, τα οποία αποτελούνται από ένα σύνολο εξισώσεων, οι οποίες καταλήγουν σε μία μορφή επίλυσης τύπου ελαχίστων τετραγώνων. Στους αλγόριθμους αυτούς οι παρατηρήσεις συνδυάζονται περιοδικά με πρόσφατες σταθμισμένες προγνώσεις, ώστε να ελαχιστοποιηθούν τα συστηματικά σφάλματα.

Το κυριότερο πλεονέκτημα αυτής της μεθοδολογίας είναι η εύκολη προσαρμογή της στις όποιες μεταβολές των παρατηρήσεων, καθώς και το γεγονός ότι δεν είναι απαραίτητες μεγάλες χρονοσειρές παρελθοντικής πληροφορίας. Επίσης, ιδιαίτερα σημαντικό είναι το γεγονός ότι οι αλγόριθμοι αυτού του τύπου έχουν χαμηλό υπολογιστικό κόστος. Τα φίλτρα Kalman έχουν εφαρμοστεί με επιτυχία για διάφορες μετεωρολογικές παραμέτρους, αλλά και σε διάφορες εφαρμογές.

1.7.1 Διόρθωση της παραμέτρου της ηλιακής ακτινοβολίας και εφαρμογές σε φωτοβολταϊκά

Μία από τις εφαρμογές όπου τα φίλτρα Kalman μπορούν να φανούν ιδιαίτερος αποτελεσματικά είναι στην πρόγνωση μετεωρολογικών παραμέτρων που σχετίζονται με την ενέργεια. Οι περισσότερες τεχνικές μετεπεξεργασίας δεν τα καταφέρνουν καλά στην πρόγνωση των μικρών χρονικών οριζόντων, εν αντιθέσει με τα φίλτρα Kalman.

Στην περίπτωση της πρόγνωσης ενέργειας για φωτοβολταϊκά πάρκα, έχει τεράστια σημασία η όσο το δυνατόν ακριβέστερη πρόγνωση της ηλιακής ακτινοβολίας σε ενδοημερήσια βάση. Κάτι τέτοιο καθιστά ιδιαίτερα αποτελεσματικά τα φίλτρα Kalman, τα οποία με μικρές χρονοσειρές δεδομένων καταφέρνουν να μειώσουν συστηματικά σφάλματα. Υπάρχουν μελέτες όπου η μείωση του σφάλματος RMSE της

πρόγνωσης της ηλιακής ακτινοβολίας έφτασε σε κάποιες περιπτώσεις έως και 43% (Pelland et al., 2011), επηρεάζοντας ταυτόχρονα και την πρόγνωση της ενέργειας των φωτοβολταϊκών. Στην συγκεκριμένη περίπτωση, χρησιμοποιήθηκε μία προσέγγιση φίλτρων Kalman, όπου το bias είχε γραμμική εξάρτηση με την ηλιακή ακτινοβολία, η οποία προέκυπτε ύστερα από πρόγνωση.

1.7.2 Βελτίωση των αποτελεσμάτων πρόγνωσης ταχύτητας του ανέμου από

NWP μοντέλα

Η μοντελοποίηση της ταχύτητας του ανέμου παρουσιάζει αρκετές ιδιαιτερότητες, που έχουν σχέση με τις κλίμακες, την κατακόρυφη κατανομή του ανέμου κ.α. Στις προγνώσεις που γίνονται για εφαρμογές στην αιολική ενέργεια, τα μοντέλα που χρησιμοποιούνται συνήθως είναι περιορισμένης περιοχής (Limited Area Models). Όσον αφορά την παράμετρο του ανέμου, τα αριθμητικά μοντέλα παρουσιάζουν σημαντικά σφάλματα ιδιαίτερα κοντά στην επιφάνεια του εδάφους. Αυτό συμβαίνει τόσο λόγω της ίδιας της παραμετροποίησης του μοντέλου όσο και λόγω της αδυναμίας των μοντέλων να «πιάσουν» φαινόμενα πολύ μικρής κλίμακας (sub grid) (Louka et al., 2008).

Οι εξομαλύνσεις που γίνονται στην οριζόντια ανάλυση των μοντέλων έχουν ως αποτέλεσμα πολλές φορές, τα μοντέλα να μην μπορούν να λάβουν υπόψιν τους ατμοσφαιρικά φαινόμενα τοπικής κλίμακας, τα οποία είναι ιδιαίτερα σημαντικά για εφαρμογές όπως η πρόγνωση της αιολικής ενέργειας. Μία μορφή αντιμετώπισης αυτού του προβλήματος είναι η αύξηση της αναλυτικής ικανότητας του μοντέλου, με την συνεπακόλουθη αύξηση της υπολογιστικής ισχύος που αυτή απαιτεί. Ωστόσο, δεν είναι βέβαιο ότι ένα τέτοιο κόστος θα αξίζει την ποσοτική μεταβολή στην ακρίβεια της πρόγνωσης που αυτή θα επιφέρει. Η εφαρμογή των φίλτρων Kalman αποτελούν μία καλή λύση σε τέτοιες περιπτώσεις.

Τα φίλτρα Kalman έχουν εφαρμοστεί με επιτυχία σε μια σειρά από τέτοιες εφαρμογές που αφορούν τον άνεμο. Η εφαρμογή τους έδειξε ότι μπορεί να έχουν υψηλή αποδοτικότητα στην βελτίωση της πρόγνωσης, οδηγώντας έως και στην εκμηδένιση των συστηματικών σφαλμάτων (Louka et al., 2008, Stathopoulos et al. 2012). Ωστόσο, δεν έχει παρατηρηθεί σημαντική βελτίωση στη μεταβλητότητα του σφάλματος. Η κατανομή της ταχύτητας του ανέμου στα προγνωστικά αποτελέσματα, μετά την εφαρμογή των φίλτρων Kalman, είναι πολύ πιο κοντά στις παρατηρήσεις (Stathopoulos

et al. 2012). Πέρα από την «κλασσική» μετεωρολογική χρήση, αυτή η τεχνική μετεπεξεργασίας μπορεί να φανεί ιδιαίτερα αποτελεσματική στον τομέα της ενέργειας, αξιοποιώντας της διορθωμένες προγνώσεις του ανέμου σε μοντέλα πρόγνωσης ενέργειας, για μεγάλα χρονικά διαστήματα (Louka et al., 2008, Stathopoulos et al. 2012). Σημαντικό είναι επίσης το γεγονός, ότι η εφαρμογή των φίλτρων Kalman, ως μέθοδος μετεπεξεργασίας πάνω σε αποτελέσματα αριθμητικού μοντέλου πιο χαμηλής ανάλυσης, μπορεί να φέρει πιο ακριβή αποτελέσματα πρόγνωσης συγκριτικά με την επιλογή της αύξησης της χωρικής ανάλυσης του αριθμητικού μοντέλου (P. Louka et al., 2008).

Μία άλλη εφαρμογή στην οποία έχει εφαρμοστεί η μεθοδολογία των φίλτρων Kalman, είναι στην πρόγνωση των ριπών του ανέμου. Σε προηγούμενες μελέτες, τα προγνωστικά αποτελέσματα βελτιώθηκαν σημαντικά, μέσω της μείωσης των συστηματικών σφαλμάτων (Patlakas et al., 2017). Η διόρθωση μέσω αυτής της μεθοδολογίας μπορεί χαρακτηριστεί πολλά υποσχόμενη στην πρόγνωση των ριπών του ανέμου (Patlakas et al., 2017), ενώ αποδεικνύεται ότι τα φίλτρα Kalman αποτελούν ένα πολύ χρήσιμο εργαλείο για προγνωστικές δραστηριότητες επιχειρησιακού τύπου, όπως είναι τα αιολικά πάρκα κ.α.

1.7.3 Εφαρμογή φίλτρων Kalman σε αποτελέσματα κυματικών αριθμητικών μοντέλων

Η παράμετρος του κυματισμού και η ακριβής πρόγνωση του έχει παρουσιάσει έντονο ενδιαφέρον τα τελευταία χρόνια, λόγω των ποικίλων δραστηριοτήτων στη θάλασσα και των εμπορικών δυνατοτήτων που τις έχουν ακολουθήσει. Τα φίλτρα Kalman έχουν δοκιμαστεί και στην προσπάθεια βελτίωσης της πρόγνωσης του κυματισμού. Όπως και στα ατμοσφαιρικά μοντέλα, τα κυματικά μοντέλα αδυνατούν να λάβουν υπόψιν τους στους υπολογισμούς φαινόμενα τοπικής κυκλοφορίας. Όπως και στις άλλες παραμέτρους, η μέθοδος των φίλτρων Kalman εφαρμόζεται και σε αυτήν την περίπτωση για να μειωθεί η διακύμανση στην πρόγνωση. Το επιδιωκόμενο ποσοστό της μείωσης της διακύμανσης μπορεί να καθοριστεί από την κατάλληλη επιλογή στις παραμέτρους του φίλτρου (Galanis, 2011).

Σε προηγούμενες μελέτες έχει επιτευχθεί σημαντική μείωση στα συστηματικά σφάλματα των αριθμητικών μοντέλων, ανεξαρτήτως του αν τα μοντέλα υπερεκτιμούσαν ή υποεκτιμούσαν την παράμετρο πρόγνωσης. Με την κατάλληλη

εφαρμογή των φίλτρων αυτών, το bias είναι δυνατόν να απαλειφθεί πρακτικά, ενώ και άλλοι δείκτες μεταβλητότητας, όπως το RMSE, το MAPE να μειωθούν σημαντικά (Galanis, 2011, Galanis et al., 2014).

1.7.4 Εφαρμογή φίλτρων Kalman στην πρόγνωση της παραμέτρου της θερμοκρασίας κοντά στην επιφάνεια του εδάφους

Όπως αναφέρθηκε και πιο πάνω για τον άνεμο, τα αριθμητικά μοντέλα παρουσιάζουν ιδιαίτερα υψηλά σφάλματα κοντά στην επιφάνεια του εδάφους. Το ίδιο συμβαίνει και για τη μετεωρολογική παράμετρο της θερμοκρασίας. Η ακριβής πρόγνωση της θερμοκρασίας κοντά στην επιφάνεια παρουσιάζει ενδιαφέρον τόσο από μετεωρολογικής άποψης όσο κι από άποψης εφαρμογών όπως είναι η γεωργία και η παραγωγή ενέργειας από φωτοβολταϊκά.

Η εφαρμογή των φίλτρων Kalman για τη διόρθωση των προγνώσεων και αυτής της παραμέτρου έχει δοκιμαστεί σε μία σειρά από προηγούμενες μελέτες. Η μεθοδολογία του φίλτρου αυτού έχει καταφέρει να μειώσει τη μέση απόκλιση σε μεγάλο βαθμό, χωρίς ωστόσο να συμβεί το ίδιο και με άλλου είδους σφάλματα, όπως είναι το μέσο απόλυτο σφάλμα (Galanis & Anadranistakis, 2002). Επίσης, το φίλτρο Kalman έχει καταστεί αποτελεσματικό και στις προγνώσεις που αφορούν ακραίες τιμές θερμοκρασίας (Emmanouil et al., 2006). Ωστόσο, δεν μπορούμε να πούμε ότι αυτό συμβαίνει σε όλες τις περιπτώσεις καθώς υπεισέρχονται πολλοί παράγοντες, όπως για παράδειγμα το ότι το μέγεθος των συστηματικών σφαλμάτων στην περίπτωση της θερμοκρασίας μεταξύ άλλων εξαρτάται και από την εποχικότητα. Παρότι υπάρχουν αρκετά περιθώρια βελτίωσης, η μέθοδος των φίλτρων αυτών κρίνεται σε μεγάλο βαθμό ως αποτελεσματική και για αυτήν την παράμετρο.

1.8. Στοιχεία θεωρίας Bayes- Στατιστική συμπερασματολογία

Είναι ο κλάδος της Στατιστικής που έχει ως αντικείμενο την ανάπτυξη μεθόδων για την ανάλυση τιμών της τυχαίας μεταβλητής που μελετάται και την εξαγωγή συμπερασμάτων για αυτήν.

Η στατιστική συμπερασματολογία οδηγεί σε συμπεράσματα για την παράμετρο θ ενός στατιστικού πληθυσμού μέσω της παρατήρησης της μεταβλητής X , και τα βασικά συμπεράσματα βασίζονται στο ότι οι τιμές του θ που δίνουν μεγάλη πιθανότητα στην τιμή του x που παρατηρήθηκε, είναι πιο πιθανές απ'ότι εκείνες που δίνουν στο x μικρή πιθανότητα (αρχή της μέγιστης πιθανοφάνειας).

1.9. Συμπερασματολογία κατά Bayes

Το πλαίσιο στο οποίο κινείται η συμπερασματολογία κατά Bayes είναι παρόμοιο με αυτό της κλασικής στατιστικής (Δελλαπόρτας-Τσιαμυρτζής, 2004): υπάρχει η παράμετρος θ του πληθυσμού η οποία πρέπει να εκτιμηθεί, καθώς και η πιθανότητα $f(x|\theta)$ η οποία καθορίζει την πιθανότητα παρατήρησης διαφορετικών x , κάτω από διαφορετικές τιμές της παραμέτρου θ . Όμως η θεμελιώδης διαφορά είναι ότι το θ χρησιμοποιείται σαν τυχαία ποσότητα. Αν και η διαφορά αυτή μπορεί να φανεί όχι και τόσο ουσιαστική, οδηγεί σε μία τελείως διαφορετική προσέγγιση, ως προς την ερμηνεία, από αυτήν την κλασικής στατιστικής.

Τα κύρια χαρακτηριστικά της προσέγγισης κατά Bayes στην στατιστική συμπερασματολογία είναι:

1. Όλες οι άγνωστες ποσότητες αντιμετωπίζονται ως τυχαίες μεταβλητές, ενώ χρησιμοποιούνται κατανομές πιθανότητας για να περιγράψουν την κατάσταση της γνώσης μας (ή την γνώση μας) για τις άγνωστες αυτές ποσότητες.
2. Η συμπερασματολογία για τις άγνωστες ποσότητες γίνεται με βάση τον κανόνα του Bayes, που επιτρέπει την χρήση πιθανοτήτων δεσμευμένων επί των τιμών που παρατηρήθηκαν.

Ποιοτικά, η προσέγγιση κατά Bayes ξεκινά με μια κατανομή πιθανότητας η οποία περιγράφει το επίπεδο της γνώσης μας (την κατάσταση της γνώσης μας) αναφορικά με τις άγνωστες ποσότητες (συνήθως παραμέτρους) πριν συλλεγούν δεδομένα και στην

συνέχεια χρησιμοποιεί τα παρατηρηθέντα δεδομένα για να επανακαθορίσει (επανεξετάσει) την κατανομή αυτή.

Στην ουσία, η συμπερασματολογία μας θα βασιστεί στην $f(\theta|x)$ και όχι στην $f(x|\theta)$, δηλαδή στην πιθανότητα της κατανομής της παραμέτρου δεδομένης της x (δεδομένα) και όχι της x δεδομένης της παραμέτρου. Σε πολλές περιπτώσεις αυτό οδηγεί σε περισσότερο φυσικά συμπεράσματα σε σχέση με την κλασική στατιστική, για να μπορέσει όμως να επιτευχθεί αυτό θα πρέπει να καθοριστεί η *a-priori* κατανομή $f(\theta)$ (prior probability distribution), η οποία αντιπροσωπεύει «τις πεποιθήσεις» για την κατανομή του θ προτού αποκτήθει οποιαδήποτε πληροφορία για τα δεδομένα.

1.10. Η *a-priori* κατανομή (Prior Distribution)

Σχεδόν σε όλες τις περιπτώσεις, όταν προσπαθεί κανείς να εκτιμήσει την παράμετρο θ θα πρέπει να έχει κάποια γνώση ή κάποια πεποίθηση σχετικά με την τιμή της θ , προτού λάβει υπόψιν του τα δεδομένα. Η κατανομή αυτή περιγράφει την κατάσταση της γνώσης που έχει κάποιος για το θ πριν παρατηρήσει οποιαδήποτε δεδομένα. Δηλαδή, έχει κάποια γνώση για την διαδικασία την οποία μελετά προτού συλλέξει τα δεδομένα. Συνεπώς, σύμφωνα με τη συμπερασματολογία κατά Bayes, τα συμπεράσματα θα πρέπει να βασίζονται στον συνδυασμό της *a-priori* γνώσης με τα δεδομένα. Η συμπερασματολογία κατά Bayes είναι ο μηχανισμός ο οποίος εξάγει συμπεράσματα από αυτόν τον συνδυασμό.

1.11. Βασικά χαρακτηριστικά της προσέγγισης κατά Bayes

Μπορεί κανείς να εντοπίσει τέσσερα βασικά σημεία τα οποία χαρακτηρίζουν την θεωρία του Bayes σε σχέση με την κλασική στατιστική:

- **A-priori Πληροφορία (Prior Information):** Κάθε πρόβλημα είναι μοναδικό και έχει το δικό του περιεχόμενο. Από αυτό ακριβώς το περιεχόμενο εξάγονται *a-priori* πληροφορίες και είναι η διατύπωση και η εκμετάλλευση της προηγούμενης γνώσης που διαχωρίζουν την Μπευζιανή θεωρία από αυτήν της κλασικής στατιστικής.
- **Υποκειμενική Πιθανότητα (Subjective Probability):** Η κλασική στατιστική εξαρτάται από μία μακροχρόνια συχνότητα καθορισμού των πιθανοτήτων. Αντίθετα, η στατιστική κατά Bayes θέτει με σαφήνεια την ιδέα ότι όλες οι πιθανότητες είναι υποκειμενικές και εξαρτώνται από τις πεποιθήσεις του κάθε ατόμου και τις γνώσεις που μπορεί να έχει ο καθένας για μια δεδομένη «κατάσταση». Η συμπερασματολογία

της βασίζεται στην a-posteriori κατανομή (posterior distribution) $f(\theta|x)$, η μορφή της οποίας εξαρτάται (μέσω του θεωρήματος του Bayes) από τον τρόπο καθορισμού της a-priori κατανομής $f(\theta)$.

- **Συνέπεια (Self-Consistency):** Χρησιμοποιώντας την παράμετρο θ σαν τυχαία, όλη η ανάπτυξη της συμπερασματολογίας κατά Bayes πηγάζει και εξαρτάται μόνο από την θεωρία πιθανοτήτων. Αυτό έχει πολλά πλεονεκτήματα και σημαίνει πως όλα τα συμπεράσματα μπορούν να παρουσιαστούν με την μορφή πιθανοτήτων για την παράμετρο θ , πράγματι προκύπτουν άμεσα από την a-posteriori κατανομή.

- **Μη προσκόλληση σε «συνταγές»:** Επειδή η κλασική στατιστική δεν είναι σε θέση να χρησιμοποιήσει όρους πιθανοτήτων για την παράμετρο θ , έχουν αναπτυχθεί αρκετά κριτήρια με σκοπό να καθορίσουν πότε ένας συγκεκριμένος εκτιμητής θα μπορούσε να χαρακτηριστεί ως «καλός».

1.12 Επάρκεια

Στην ανάλυση κατά Bayes, η a-posteriori κατανομή εξαρτάται από τα δεδομένα μόνο μέσω μίας επαρκούς στατιστικής παραμέτρου. Ορίζεται στην στατιστική κατά Bayes η έννοια της επάρκειας. Μία στατιστική συνάρτηση δεδομένων $t(x)$ ονομάζεται επαρκής στατιστική συνάρτηση για την παράμετρο θ εάν:

$$f(\theta | t(x)) = f(\theta|x) \quad (1.1)$$

για κάθε a-priori κατανομή $f(\theta)$.

1.13. Η Αρχή της Πιθανοφάνειας

Η αρχή της πιθανοφάνειας αναφέρει ότι αν 2 πειράματα «μοιράζονται» την ίδια πιθανοφάνεια (αναλογικά), τότε η συμπερασματολογία μας σχετικά με την παράμετρο θ , θα πρέπει να είναι η ίδια σε κάθε περίπτωση. Με άλλα λόγια, όλα τα στοιχεία της συμπερασματολογίας θα πρέπει να είναι βασισμένα μονάχα στην συνάρτηση πιθανοφάνειας. Ένα σημαντικό πλεονέκτημα της Μπεϋζιανής θεωρίας είναι ότι οι τεχνικές οι οποίες χρησιμοποιούνται είναι άρρηκτα συνδεδεμένες με την αρχή της πιθανοφάνειας.

1.14. Το Θεώρημα του Bayes (διακριτή περίπτωση)

Στην βασική του μορφή το θεώρημα του Bayes είναι απλό και αφορά υπό συνθήκη πιθανότητες. Αν A και B είναι δύο ενδεχόμενα με $P(A) > 0$, τότε:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \quad (1.2)$$

Η χρησιμότητα του θεωρήματος του Bayes σε εφαρμογές πιθανοτήτων είναι ότι παρέχει την δυνατότητα αντιστροφής της «θέσης» των ενδεχομένων. Έτσι, γίνεται εμφανές πώς η πιθανότητα του $B|A$ σχετίζεται με την πιθανότητα του $A|B$. Μια μικρή προέκταση του θεωρήματος του Bayes μπορεί να γίνει, αν θεωρήσουμε τα ενδεχόμενα C_1, \dots, C_k , τα οποία διαμερίζουν ένα δειγματικό χώρο Ω , έτσι ώστε τα $C_i \cap C_j = \emptyset$ για κάθε $i \neq j$ και $C_1 \cup \dots \cup C_k = \Omega$. Σε αυτήν την περίπτωση θα έχουμε:

$$P(C_i|A) = \frac{P(A|C_i)P(C_i)}{\sum_{j=1}^k P(A|C_j)P(C_j)} \quad i=1, \dots, k \quad (1.3)$$

1.15. Το Θεώρημα του Bayes (περίπτωση κατανομών)

Το θεώρημα του Bayes σε όρους τυχαίων μεταβλητών με πυκνότητες που συμβολίζονται γενικά με f , παίρνει την εξής μορφή:

$$f(\theta|x) = \frac{f(\theta)f(\theta|x)}{\int f(\theta)f(x|\theta) d\theta} \quad (1.4)$$

Θα πρέπει να προσέξουμε ιδιαίτερα το γεγονός ότι από την στιγμή που ολοκληρώνουμε ως προς θ , ο παρανομαστής στο θεώρημα του Bayes είναι συνάρτηση μόνο ως προς x . Συνεπώς, για δεδομένες παρατηρήσεις x , ο παρανομαστής είναι σταθερά και ονομάζεται σταθερά κανονικοποίησης. Με βάση αυτά ένας εναλλακτικός τρόπος παρουσίασης του θεωρήματος του Bayes είναι ο εξής:

$$f(\theta|x) \propto f(\theta)f(x|\theta) \quad (1.5)$$

ή αλλιώς θα λέγαμε ότι η a-posteriori κατανομή (posterior distribution) είναι ανάλογη της a-priori κατανομής (prior distribution) πολλαπλασιαζόμενης με την συνάρτηση πιθανοφάνειας (likelihood function).

Η a-priori και η a-posteriori κατανομές είναι έννοιες σχετικές, δηλαδή η a-posteriori μπορεί να γίνει η a-priori σε επόμενο βήμα. Αυτό επιτυγχάνεται μόνο μέσω του θεωρήματος Bayes, εφόσον ορισθεί αρχικά η a-priori κατανομή. Κάτι τέτοιο συνιστά

την ευελιξία της μεθόδου, καθώς μπορούν στην πορεία να αλλάξουν οι υποθέσεις χωρίς να επηρεαστούν τα αποτελέσματα.

1.16. Τα βασικά στάδια της θεωρίας του Bayes

Τα βασικά στάδια της θεωρίας του Bayes συνοψίζονται στα εξής:

1. Καθορισμός του μοντέλου πιθανοφάνειας ($f(x|\theta)$).
2. Καθορισμός της εκ των υστέρων κατανομής $f(\theta)$.
3. Υπολογισμός της εκ των υστέρων κατανομής $f(\theta|x)$, από το θεώρημα του Bayes.
4. Εξαγωγή συμπερασμάτων από την εκ των υστέρων πληροφορία.

1.17. Στατιστική του ανέμου-Γενικά

Η ταχύτητα του ανέμου είναι μια από τις πλέον μεταβλητές και ασταθείς μετεωρολογικές παραμέτρους και η χρήση στατιστικών εργαλείων που περιγράφουν την εξέλιξη του είναι ιδιαίτερα σημαντική. Προκειμένου να μοντελοποιηθούν οι διακυμάνσεις της ταχύτητας του ανέμου έχουν χρησιμοποιηθεί διάφορες στατιστικές κατανομές, οι σημαντικότερες εκ των οποίων παρουσιάζονται παρακάτω.

1.18. Κατανομή Weibull

Η στατιστική κατανομή που θεωρείται πως προσεγγίζει καλύτερα την κατανομή των συχνοτήτων των ταχυτήτων του ανέμου, είναι η κατανομή Weibull. Η κατανομή αυτή είναι ασύμμετρη και διπαραμετρική.

Η συχνότητα εμφάνισης συγκεκριμένων τιμών της ταχύτητας του ανέμου μπορεί να περιγραφεί από την συνάρτηση πυκνότητας πιθανότητας (*probability density function*), $f(U)$, της ταχύτητας του ανέμου. Η πιθανότητα της ταχύτητας του ανέμου να βρεθεί μεταξύ δύο τιμών U_a και U_b :

$$p(U_a < U < U_b) = \int_{U_a}^{U_b} f(U) dU \quad (1.6)$$

Μπορεί επίσης να χρησιμοποιηθεί η συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς, τυχαίας μεταβλητής U είναι μια συνάρτηση $f(U)$, η οποία ορίζεται σ' ένα διάστημα πραγματικών αριθμών (a,b) και έχει τις εξής ιδιότητες:

$$f(U) \geq 0, \forall U \quad (1.7)$$

$$\int_a^b f(U) dU = 1 \quad (1.8)$$

Στην περίπτωση της κατανομής Weibull (Weibull distribution) η συνάρτηση πυκνότητας πιθανότητας μπορεί να προσεγγιστεί από:

$$P(u) = \left(\frac{k}{c}\right) \left(\frac{u}{c}\right)^{k-1} \exp\left[-\left(\frac{u}{c}\right)^k\right] \quad (1.9)$$

$$u \geq 0, c > 0, k > 0$$

όπου u η ταχύτητα του ανέμου, c η παράμετρος κλίμακας (scale parameter) ή χαρακτηριστική ταχύτητα του ανέμου (σε μονάδες ταχύτητας) και k η παράμετρος μορφής (shape parameter) που είναι αδιάστατο μέγεθος. Η ολοκλήρωση της (2.4) από 0 μέχρι κάποια τιμή του ανέμου $x > 0$ δίνει την αθροιστική συνάρτηση κατανομής (cumulative distribution function):

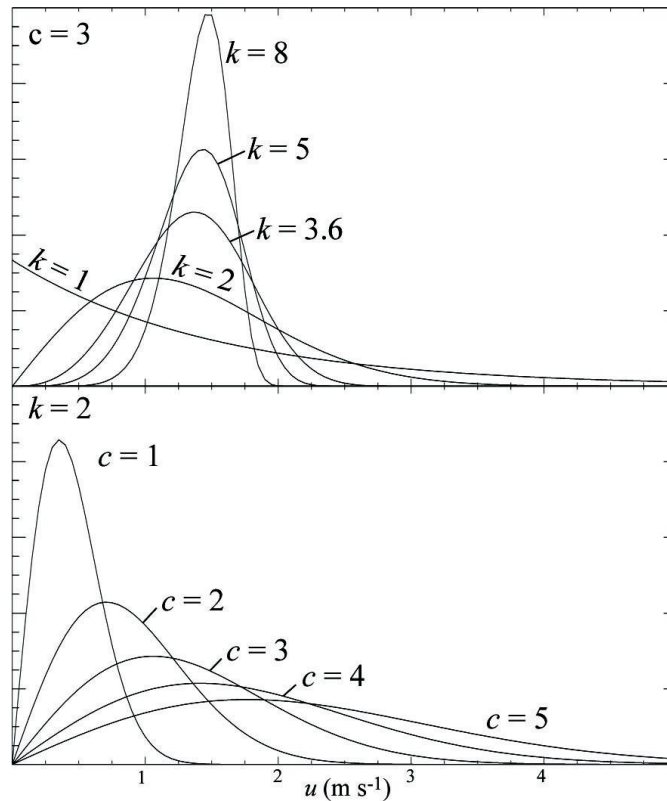
$$F_0(x) = 1 - \exp\left[-\left(\frac{x}{c}\right)^k\right], \quad x \geq 0 \quad (1.10)$$

Οι δύο παράμετροι της κατανομής c, k , η μέση ταχύτητα του ανέμου και η αντίστοιχη διασπορά συνδέονται μέσω της σχέσης:

$$\bar{u} = c \cdot \Gamma\left(1 + \frac{1}{k}\right) \quad (1.11)$$

όπου \bar{u} είναι η μέση ταχύτητα του ανέμου και $\Gamma(\cdot)$ η συνάρτηση γάμα.

Η παράμετρος κλίμακας c , καθορίζει το βαθμό της κατανομής και σχετίζεται με την μέση τιμή του ανέμου. Η παράμετρος μορφής k , είναι καθαρός αριθμός, καθορίζει το σχήμα της κατανομής (Σχήμα 2.1) και αποτελεί ένα μέτρο του πόσο «αιχμηρή» είναι η κατανομή (μεγάλες τιμές της k ισοδυναμούν με πιο «αιχμηρή» κατανομή, μεγαλύτερη συχνότητα εμφάνισης ταχυτήτων ανέμου γύρω από κάποια συγκεκριμένη τιμή)



Σχήμα 1.1: Η κατανομή Weibull για διάφορες τιμές της παραμέτρου μορφής και κλίμακας αντίστοιχα.

Όπως αποτυπώνεται και από το σχήμα, η συνάρτηση πυκνότητας πιθανότητας Weibull παρουσιάζει δεξιά λόξωση. Αυτό δείχνει ότι οι ισχυροί άνεμοι είναι πιο σπάνιοι από τους μέτριους και ασθενείς ανέμους. Με τη γνώση των δύο παραμέτρων της κατανομής Weibull σε συνάρτηση με την διεύθυνση του ανέμου παρέχεται όλη η πληροφορία για τη συμπεριφορά του ανέμου. Η κατανομή της διεύθυνσης του ανέμου έχει συνήθως δύο κέντρα και σχετίζεται αποκλειστικά με την μορφολογία του εδάφους της περιοχής. Αυτό σημαίνει ότι συνήθως ο άνεμος έρχεται από δύο καθορισμένες κατευθύνσεις.

Για $\kappa=2$, έχουμε την περίπτωση της κατανομής Rayleigh.

1.19. Λογαριθμοκανονική κατανομή (Lognormal distribution)

Γνωστή και ως κατανομή Galton, θεωρείται κατά περιπτώσεις κατάλληλη για την περιγραφή της ταχύτητας του ανέμου (Allouhi et al. 2017). Έχει την παρακάτω μορφή:

$$f(v) = \frac{1}{c v \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln(v)-\kappa}{c}\right)^2\right) \quad (1.12)$$

Όπου k η παράμετρος μορφής (shape parameter) και c η παράμετρος κλίμακας (scale parameter).

1.20. Γάμμα κατανομή

Η Γάμμα κατανομή είναι άλλη μια κατανομή που χρησιμοποιείται εκτεταμένα για την μοντελοποίηση του ανέμου. Το ενδιαφέρον της έγκειται στη σχέση της με την κανονική και την εκθετική κατανομή (Allouhi et al. 2017). Εκφράζεται ως:

$$f(v) = \frac{v^{k-1}}{\Gamma(k) c^k} \exp\left(-\frac{v}{c}\right) \quad (1.13)$$

Όπου Γ , η συνάρτηση Γάμμα, k η παράμετρος μορφής (shape parameter) και c η παράμετρος κλίμακας (scale parameter).

1.21. Στατιστική των ακραίων τιμών του ανέμου – Κατανομή Gumbel

Σημαντικό στοιχείο στη μελέτη του ανέμου, αποτελούν οι ακραίες τιμές της ταχύτητας του (ριπές ανέμου). Μία από τις κατάλληλες κατανομές για την περιγραφή των ακραίων τιμών αποτελεί η Gumbel.

Η συνάρτηση της πυκνότητας πιθανότητας της κατανομής Gumbel είναι η

$$P(u_e) = \frac{1}{\beta} \exp\left(-\frac{(u_e + \mu)}{\beta}\right) \exp\left(-\exp\left(-\frac{(u_e + \mu)}{\beta}\right)\right)$$

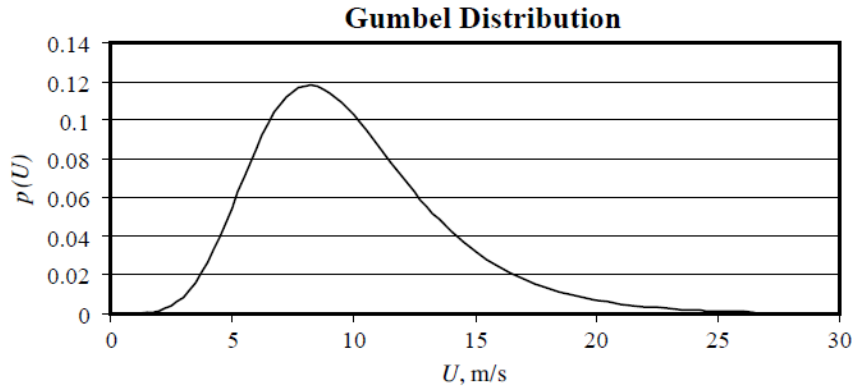
(1.14)

όπου u_e η extreme τιμή της ταχύτητας του ανέμου σε μια χρονική περίοδο, $\beta = (\sigma_e \sqrt{6})/\pi$, $\mu = \bar{u}_e - 0.577\beta$, \bar{u}_e η μέση τιμή των extreme τιμών, σ_e η τυπική απόκλιση των τιμών αυτών.

Η αθροιστική συνάρτηση κατανομής είναι η

$$F(u_e) = \exp\left(-\exp\left(-\frac{(u_e + \mu)}{\beta}\right)\right) \quad (1.14)$$

Στο Σχήμα 2.2 φαίνεται η συνάρτηση της πυκνότητας πιθανότητας Gumbel για μια μέση τιμή extreme ανέμου 10m/s και τυπική απόκλιση 4m/s.



Σχήμα 1.2: Η κατανομή Gumbel (Manwell J.F. et al., 2009).

1.21.1 Γενικευμένη κατανομή ακραίων τιμών

Μια τυχαία μεταβλητή Z λέμε ότι έχει γενικευμένη κατανομή ακραίων τιμών (Generalized Extreme Value distribution) με παράμετρο θέσης μ , παράμετρο κλίμακας $\sigma > 0$ και παράμετρο σχήματος ξ αν η αθροιστική συνάρτηση κατανομής της δίνεται από τον παρακάτω τύπο:

$$G_{\xi}(x) = \begin{cases} e^{-[1+\xi(x-\mu)/\sigma]^{-\frac{1}{\xi}}}, & \text{αν } \xi \neq 0 \\ e^{-e^{-(x-\mu)/\sigma}}, & \text{αν } \xi = 0. \end{cases} \quad (1.15)$$

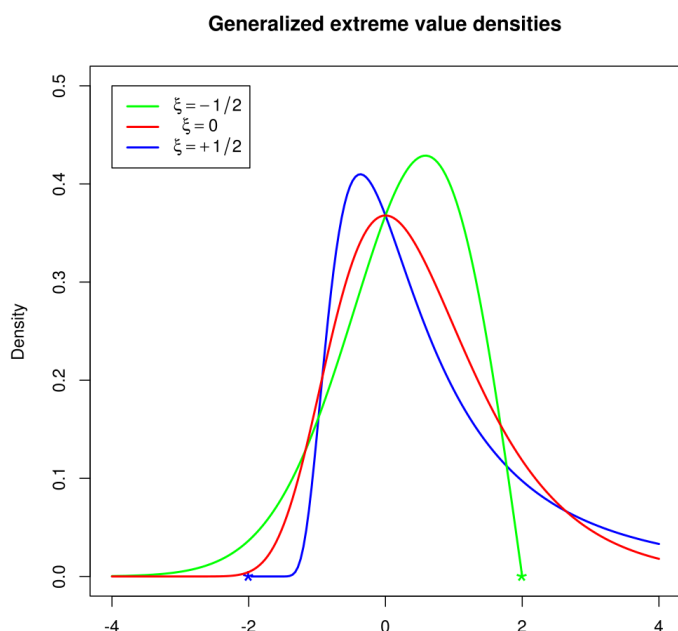
Η κατανομή αυτή είναι γνωστή ως γενικευμένη κατανομή ακραίων τιμών και συμβολίζεται $GEV(x; \mu, \sigma, \xi)$, $G_{\xi}(x)$ ή απλά $G(x)$. Η παράμετρος ξ είναι αυτή που καθορίζει τον τύπο των οριακών κατανομών. Για παράδειγμα η $GEV(1, 1, 1/\alpha)$ για $\alpha > 0$ θα είναι ίδιου τύπου με την $\Phi_{\alpha}(x)$ ενώ η $GEV(-1, -1, -1/\alpha)$ θα είναι ίδιου τύπου με την $\Psi_{\alpha}(x)$, αφού σε αυτές τις περιπτώσεις προκύπτουν οι τύποι που αντιστοιχούν στην Frechet και τη Weibull.

Παρατηρούμε ότι λαμβάνοντας το όριο στο μηδέν του πρώτου κλάδου της σχέσης 1.15, δηλαδή,

$$\lim_{\xi \rightarrow 0} e^{-[1+\xi(x-\mu)/\sigma]^{-\frac{1}{\xi}}}$$

καταλήγουμε στον δεύτερο κλάδο της σχέσης 1.15.

Η παράμετρος σχήματος ξ αναφέρεται συνήθως και ως δείκτης ουράς (tail index) της κατανομής. Αυτό διότι κατέχει σημαντικό ρόλο στον προσδιορισμό της συμπεριφοράς της ουράς της κατανομής. Παρακάτω φαίνεται ένα διάγραμμα της γενικευμένης κατανομής ακραίων τιμών για τιμές του ξ που αντιστοιχούν στις τρεις οριακές κατανομές από όπου φαίνεται ότι για μεγάλες τιμές του ξ έχουμε βαριά ουρά. Γενικά ισχύει ότι για τιμές του $\xi < 0$, $\xi = 0$ η G_ξ είναι ίδιου τύπου με την Weibull και την Gumbel αντίστοιχα.



Σχήμα 1.3: Γενικευμένη κατανομή ακραίων τιμών για διάφορες τιμές του ξ

Η συνάρτηση πυκνότητας πιθανότητας της γενικευμένης κατανομής ακραίων τιμών θα βρίσκεται παραγωγίζοντας τη σχέση 1.15.

$$\frac{1}{\sigma} \left[1 + \left(\frac{\xi(x-\mu)}{\sigma} \right) \right]^{-\frac{1}{\xi}-1} e^{-[1+\xi(x-\mu)/\sigma]^{\frac{1}{\xi}}}, \text{ αν } \xi \neq 0$$

$$g(x) = \left\{ \begin{array}{l} \frac{1}{\sigma} \left[1 + \left(\frac{\xi(x-\mu)}{\sigma} \right) \right]^{-\frac{1}{\xi}-1} e^{-[1+\xi(x-\mu)/\sigma]^{\frac{1}{\xi}}}, \text{ αν } \xi \neq 0 \\ \frac{1}{\sigma} e^{[-(x-\mu)/\sigma]} e^{-e^{-(x-\mu)/\sigma}}, \text{ αν } \xi = 0. \end{array} \right. \quad (1.16)$$

ΚΕΦΑΛΑΙΟ 2ο

ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Τα τελευταία χρόνια, τα αριθμητικά μοντέλα πρόγνωσης καιρού (NWP) έχουν αυξήσει κατά πολύ την ακρίβεια των προγνώσεων με τις υψηλής ανάλυσης προσομοιώσεις. Ωστόσο, οι προγνώσεις των μετεωρολογικών παραμέτρων σε τοπική κλίμακα αναδεικνύουν τους περιορισμούς των αριθμητικών μοντέλων, με συστηματικά ή μη σφάλματα τα οποία είναι ιδιαίτερα έντονα για τις περιοχές κοντά στην επιφάνεια του εδάφους. Πολλές παράμετροι υπεισέρχονται σε αυτά τα προβλήματα και οι οποίες σχετίζονται με την ίδια τη φυσική της ατμόσφαιρας και την παραμετροποίηση των μοντέλων, τα συστήματα μοντελοποίησης του κυματισμού, την ομαλοποίηση των χαρακτηριστικών της τοπογραφίας, τεχνικές δυσλειτουργίες στην επιτυχή προσομοίωση διαδικασιών μικρής κλίμακας, καθώς και δυναμικά προβλήματα στις αρχικές και οριακές συνθήκες. Αυτά τα προβλήματα κάνουν τη χρήση διαδικασιών μετεπεξεργασίας σημαντικό και αναπόσπαστο κομμάτι των ολοκληρωμένων συστημάτων πρόγνωσης.

Τα φίλτρα Kalman αποτελούν ένα πολύ σημαντικό και αξιόπιστο εργαλείο στην διόρθωση αποτελεσμάτων πρόγνωσης. Ωστόσο, δεν έχουν την ίδια αποτελεσματικότητα σε όλα τα είδη σφαλμάτων. Ενώ στη γενικότητα καταφέρνουν αρκετά μεγάλη μείωση στη μέση απόκλιση (bias), υπάρχουν αρκετά ερωτήματα σε σχέση με το εάν και κατά πόσον καταφέρνουν να μειώσουν τη μεταβλητότητα του σφάλματος (error variability). Επιπλέον, δεν υπάρχει σαφής εικόνα σε σχέση με τους χρονικούς ορίζοντες πρόγνωσης που αυτά μπορούν να έχουν τη μέγιστη απόδοση, όπως επίσης και με το κατά πόσον είναι ικανά να βελτιώσουν βραχυπρόθεσμες προγνώσεις.

Στην παρούσα εργασία θα εξετασθούν οι συνθήκες κάτω από ποιες μπορεί να προκύψει η βέλτιστη αξιοποίηση των συνδυασμένων Bayesian Kalman φίλτρων για την μετεωρολογική παράμετρο του ανέμου, σε περιοχές κοντά στην επιφάνεια του εδάφους. Συγκεκριμένα, σκοπός της εργασίας είναι :

α) Η εξέταση των παραμέτρων των συνδυασμένων αυτών φίλτρων και των συνθηκών κάτω από τις οποίες αυτά εμφανίζουν τη μέγιστη αποδοτικότητα, στην βελτίωση της πρόγνωσης του ανέμου.

β) Η διερεύνηση του κατά πόσον η διαθεσιμότητα μετρήσεων και αποτελεσμάτων του μετεωρολογικού μοντέλου με μικρότερα χρονικά βήματα, μπορεί να βοηθήσει στην επίτευξη καλύτερων αποτελεσμάτων.

γ) Η μελέτη των χρονικών οριζόντων στους οποίους τα αποτελέσματα μπορούν να είναι τα πλέον αξιόπιστα, καθώς επίσης και η εστίαση στις βραχυπρόθεσμες προγνώσεις, όπου μέχρι στιγμής οι προηγούμενες μελέτες δεν έχουν ιδιαίτερες αναφορές.

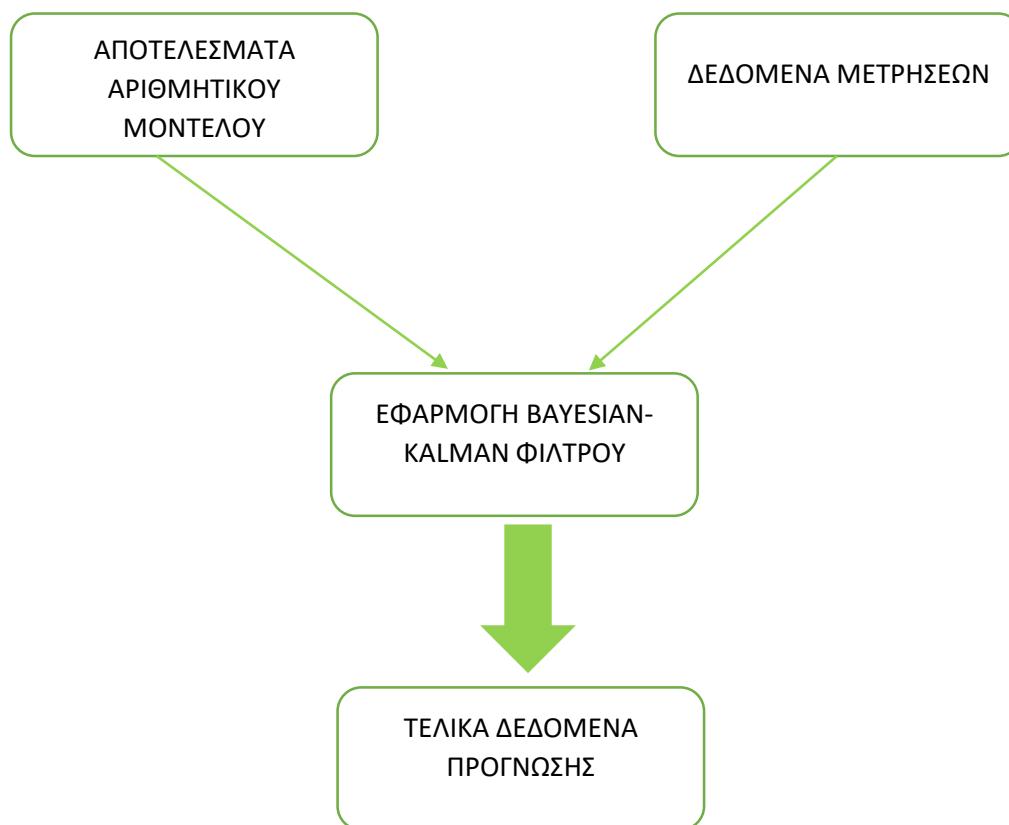
ΚΕΦΑΛΑΙΟ 3ο

ΜΕΘΟΔΟΛΟΓΙΑ- ΔΕΔΟΜΕΝΑ

3.1. Γενικά

Στο κεφάλαιο αυτό αναλύεται η μέθοδος που αναπτύχθηκε καθώς και η αναλυτική περιγραφή του φίλτρου Kalman που εξετάζει η παρούσα εργασία.

Τα δεδομένα που χρησιμοποιήθηκαν είναι δύο τύπων, τα προγνωστικά και οι παρατηρήσεις. Τα προγνωστικά προέρχονται από το ατμοσφαιρικό μοντέλο RAMS (Pielke et al., 1992), με χρήση του καταλληλότερου σημείου του grid στην περιοχή ενδιαφέροντος και για το ύψος των 45 μέτρων. Τα δεδομένα παρατηρήσεων προήλθαν από σύστημα SCADA αιολικών πάρκων που βρίσκονται στις εν λόγω περιοχές. Στη συνέχεια, τα δεδομένα αυτά εισάγονταν στον αλγόριθμο του φίλτρου Kalman προκειμένου να προκύψει η διόρθωση και να εξετάσουμε υπό ποιες συνθήκες έχουμε τις βέλτιστες διορθώσεις στα συστηματικά σφάλματα των αποτελεσμάτων του αριθμητικού μοντέλου.



Σχήμα 3.1 Σχηματική αναπαράσταση διαδικασίας

3.2. Περιγραφή του ατμοσφαιρικού μοντέλου πρόγνωσης καιρού RAMS

Το ατμοσφαιρικό μοντέλο RAMS (Regional Atmospheric Modeling System RAMS) είναι ένα ευρέως χρησιμοποιούμενο αριθμητικό μοντέλο, το οποίο αρχικά δημιουργήθηκε στο Colorado State University and Aster Division of Mission Research Inc. (<http://www.atmet.com>) για ερευνητικούς σκοπούς (Pielke et al., 1992). Η συνηθέστερη χρήση του αφορά στην προσομοίωση μιας πληθώρας ατμοσφαιρικών φαινομένων μέσης κλίμακας και χρησιμοποιείται από την Ομάδα Ατμοσφαιρικών Μοντέλων και Πρόγνωσης Καιρού (Atmospheric Modeling And Weather Forecasting Group) στις εγκαταστάσεις της στο Πανεπιστήμιο Αθηνών, παρέχοντας πρόγνωση καιρού μέχρι και σαράντα οχτώ ωρών στον Ελλαδικό χώρο (<http://forecast.uoa.gr>).

Οι Cotton et al. (2003) παρείχαν μια περιγραφή του συστήματος βασιζόμενοι στη νέα ανάπτυξη στο κομμάτι της φυσικής και των υπολογιστικών αλγορίθμων από το 1992. Το φυσικό τμήμα του μοντέλου, περιγράφει τα βασικά ατμοσφαιρικά φαινόμενα τα οποία περιλαμβάνουν μικροφυσικές διεργασίες στα σύννεφα. Το μοντέλο περιλαμβάνει ειδικότερα: επάλληλα πλέγματα αμφίδρομης διάταξης (two-way interactive nested grid structure), συντεταγμένες που ακολουθούν την τοπογραφία (terrain following height coordinates), διεργασίες ατμοσφαιρικής τυρβώδους διάχυσης, παραμετροποίηση της μικροφυσικής των νεφών (Walko et al., 1995; Harrington et al., Meyers et al. 1997), παραμετροποίηση της κατακόρυφης μεταφοράς (convective parameterization), (Tremback, 1990; Kain and Fritsch, 1993), σχήματα ακτινοβολίας νεφών (Chen and Cotton, 1983; Harrington et al. 1999), διαδικασίες διάχυσης και μεταφοράς υγρασίας, σχήμα παραμετροποίησης της επιφάνειας του εδάφους με τη δυνατότητα να εφαρμόζει πληροφορία της χρήσης γης και της υψής του εδάφους σε μεγάλη ανάλυση κλίμακα, μοντέλο θερμοκρασίας εδάφους και υγρασίας, υδρολογικό μοντέλο (Walko et al, 2000), μετατροπή των υδρατμών σε αρκετούς τύπους υδρομετεώρων και διάφορες επιλογές για τη χρήση οριακών συνθηκών. Οι κύριες συνιστώσες του είναι:

- (1) Ένα ατμοσφαιρικό μοντέλο, το οποίο εκτελεί τις προσομοιώσεις
- (2) Ένα σύστημα ανάλυσης δεδομένων, το οποίο προετοιμάζει τα αρχικά δεδομένα από τις παρατηρήσεις για να χρησιμοποιηθούν ως αρχικές και οριακές συνθήκες από το ατμοσφαιρικό μοντέλο.
- (3) Ένα σύστημα μετεπεξεργασίας, το οποίο οπτικοποιεί και αναλύει τα αποτελέσματα του ατμοσφαιρικού μοντέλου με μια ποικιλία λογισμικού οπτικοποίησης και άλλα προγράμματα υποστήριξης.

Το RAMS εφαρμόζεται συνήθως ως μοντέλο περιορισμένης περιοχής (Σχήμα 3.2). Δεν υπάρχει κατώτατο όριο για το μέγεθος της περιοχής ούτε για το πλέγμα πεπερασμένων διαφορών του μοντέλου. Με τον ίδιο κώδικα έχουν προσομοιωθεί φαινόμενα μικρής κλίμακας, όπως ανεμοστρόβιλοι και στρόβιλοι στο οριακό στρώμα, καθώς επίσης και μικρότερης κλίμακας τυρβώδεις ροές πάνω από κτήρια και αεροσύραγγες. Η εφαρμογή επάλληλων πλεγμάτων, που αλληλεπιδρούν αμφίδρομα, επιτρέπει την χρησιμοποίηση πλεγμάτων με υψηλή ανάλυση με σκοπό την επίλυση μικρής κλίμακας ατμοσφαιρικών συστημάτων, όπως τυφώνες, ενώ ταυτόχρονα επιλύει τα μεγάλης κλίμακας φαινόμενα στο πλέγμα χαμηλής ανάλυσης.



Σχήμα 3.2 Χάρτης απεικόνισης του μετεωρολογικού μοντέλου RAMS στην περιοχή της Μεσογείου.

3.3. Το φίλτρο Kalman

Το φίλτρο Kalman είναι ένας αναδρομικός αλγόριθμος επεξεργασίας δεδομένων. Θεωρείται ο καταλληλότερος γραμμικός εκτιμητής για δυναμικά συστήματα στα οποία ο θόρυβος και η διαταραχή είναι δυο αναπόσπαστοι παράγοντες.

3.3.1 Γενικά για το Kalman

Η προσέγγιση των φίλτρων Kalman αποτελεί μία από τις πιο επιτυχημένες μεθόδους που χρησιμοποιούνται για τη διόρθωση των αποτελεσμάτων πρόγνωσης των Αριθμητικών Μοντέλων Πρόγνωσης (NWPM), (Kalman 1960; Kalman and Bucy 1961; Bossanyi 1985; Persson, 1990; Dragulanesu, 1993; Kalnay, 2002; Galanis and Anadranistakis, 2002; Crochet, 2004; Giebel, 2000). Είναι στατιστικά μοντέλα τα οποία αποτελούνται από ένα σύνολο μαθηματικών εξισώσεων, που οδηγούν στην επίλυση της μεθόδου των ελαχίστων τετραγώνων. Τα φίλτρα Kalman είναι κατάλληλα για την εκτίμηση δυναμικών συστημάτων. Σε αυτά τα μοντέλα, οι παρατηρήσεις συνδυάζονται αναδρομικά με ένα πλήθος προσφάτων προγνωστικών αποτελεσμάτων, στα οποία εφαρμόζονται στατιστικά βάρη, ώστε να ελαχιστοποιηθούν οι αποκλίσεις μεταξύ των αντιστοίχων δεδομένων παρατήρησης-πρόγνωσης.

Ο όρος αναδρομικός αλγόριθμος ή φίλτρο αναφέρεται στην ιδιότητα του να χρησιμοποιεί μέρος προηγούμενης πληροφορίας χωρίς να χρειάζεται όλη η προσφερόμενη πληροφορία να αποθηκευτεί για να επεξεργαστεί ξανά, κάθε φορά που μια καινούργια μέτρηση παρουσιάζεται .

Η μεθοδολογία αυτή συγκεντρώνει μια σειρά από πλεονεκτήματα, τα οποία βοηθούν στην καλύτερη εκτίμηση της δυναμικής κατάστασης ενός συστήματος. Το βασικό πλεονέκτημα είναι η εύκολη προσαρμογή σε κάθε μεταβολή των δεδομένων παρατήρησης καθώς και το γεγονός ότι απαιτείται μικρό πλήθος προηγούμενης πληροφορίας. Ένα ακόμη σημαντικό πλεονέκτημα αυτών των στατιστικών μοντέλων είναι η μειωμένες απαιτήσεις σε υπολογιστική μνήμη (CPU memory). Σε γραμμικά περιβάλλοντα ο αλγόριθμος του Kalman αποτελεί τη βέλτιστη λύση. Η εφαρμογή του σε μετεωρολογικές εφαρμογές που παρουσιάζουν μη γραμμική ή ασυνεχή συμπεριφορά απαιτεί ιδιαίτερη προσαρμογή.

Η γενική μορφή ενός τέτοιου αλγορίθμου είναι ως εξής:

Ο βασικός στόχος αυτής της μεθόδου είναι η προσομοίωση μιας άγνωστης διαδικασίας \mathbf{x}_t , της οποίας η εξέλιξη στο χρόνο πρέπει να υπολογιστεί μέσω ενός πίνακα τιμών. Η χρονική μεταβολή της παραμέτρου αυτής περιγράφεται από την εξίσωση του συστήματος

$$\mathbf{x}_t = \mathbf{F}_t \cdot \mathbf{x}_{t-1} + \mathbf{w}_t \quad (3.1)$$

Επιπλέον, για κάθε χρονική στιγμή t έχουμε μία τιμή παρατήρησης \mathbf{y}_t . Η σχέση που συνδέει τις τιμές παρατήρησης με τις τιμές της άγνωστης διαδικασίας (εξίσωση συστήματος) λέγεται εξίσωση παρατηρήσεων και είναι:

$$\mathbf{y}_t = \mathbf{H}_t \cdot \mathbf{x}_t + \mathbf{v}_t \quad (3.2)$$

Όπου,

\mathbf{F}_t : ο πίνακας συστήματος

\mathbf{H}_t : ο πίνακας παρατηρήσεων

\mathbf{w}_t και \mathbf{v}_t : ανεξάρτητα τυχαία διανύσματα που ακολουθούν την Γκαουσιανή κατανομή, με μηδενική μέση τιμή (εξ ορισμού)

Τόσο οι πίνακες συστήματος (\mathbf{F}_t) και παρατηρήσεων (\mathbf{H}_t), όσο και οι πίνακες συνδιακύμανσης, \mathbf{W}_t και \mathbf{V}_t , των ανεξαρτήτων τυχαίων διανυσμάτων \mathbf{w}_t και \mathbf{v}_t , πρέπει να οριστούν πριν την εφαρμογή του φίλτρου.

Το φίλτρο Kalman παρέχει μια αναδρομική μέθοδο για τον υπολογισμό μιας άγνωστης παραμέτρου \mathbf{x}_t , τη χρονική στιγμή t , βασιζόμενη σε όλες τις προηγούμενες τιμές μέχρι τη χρονική στιγμή t . Μια πρώτη εκτίμηση του \mathbf{x}_t και του σφάλματός του, του πίνακα συνδιακύμανσης \mathbf{P}_t , βασισμένη σε προηγούμενες χρονικές τιμές, δίνεται από τις σχέσεις :

$$\mathbf{x}_{t/t-1} = \mathbf{F}_t \cdot \mathbf{x}_{t-1} \quad (3.3)$$

$$\mathbf{P}_{t/t-1} = \mathbf{F}_t \cdot \mathbf{P}_{t-1} \cdot \mathbf{F}_t^T + \mathbf{W}_t \quad (3.4)$$

Αυτή ακολουθείται από ένα αναβαθμισμένο βήμα, στο οποίο η παρατηρούμενη τιμή τη χρονική στιγμή t θα αναμειχθεί με την προηγούμενη πληροφορία :

$$\mathbf{x}_t = \mathbf{x}_{t/t-1} + \mathbf{K}_t \cdot (\mathbf{y}_t - \mathbf{H}_t \cdot \mathbf{x}_{t/t-1}) \quad (3.5)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \cdot \mathbf{H}_t) \cdot \mathbf{P}_{t/t-1} \quad (3.6)$$

Όπου,

I : ο μοναδιαίος πίνακας και

$$K_t = P_{t/t-1} \cdot H_t^T \cdot (H_t \cdot P_{t/t-1} \cdot H_t^T + V_t)^{-1} \quad (3.7)$$

είναι η προσθήκη του φίλτρου (*Kalman Gain*), μία πολύ σημαντική παράμετρος η οποία ρυθμίζει την ικανότητα του φίλτρου να προσαρμόζεται στις όποιες πιθανές μεταβολές των δεδομένων.

Οι εξισώσεις (3.5)-(3.7) ανανεώνουν τον αλγόριθμο Kalman μεταξύ δύο διαδοχικών χρονικών στιγμών $t-1$ και t .

Τα παραπάνω βήματα του αλγορίθμου Kalman αποτελούν τη γενική του μορφή και μπορούν να χρησιμοποιηθούν με ποικίλους τρόπους, ανάλογα με τις παραμέτρους που μελετώνται κάθε φορά.

3.3.2 Το Bayesian Kalman φίλτρο

Πρόκειται για ένα «υβριδικό» φίλτρο το οποίο χρησιμοποιεί μη γραμμικές συναρτήσεις, σε συνδυασμό με τη συμπερασματολογία του Bayes, προκειμένου να μειώσει τα πιθανά συστηματικά σφάλματα, τα οποία προκύπτουν από τα Αριθμητικά Μοντέλα Πρόγνωσης (NWPM). Αποτελεί μία επέκταση του κλασσικού φίλτρου Kalman.

Το στατιστικό αυτό μοντέλο προκύπτει από το συνδυασμό του μη γραμμικού φίλτρου Kalman και ενός Bayesian μοντέλου. Τα δύο συστατικά στοιχεία του υβριδικού αυτού μοντέλου, έχουν δύο διακριτούς ρόλους στο φιλτράρισμα των δεδομένων. Αρχικά, το κομμάτι του φίλτρου Kalman είναι αυτό που θα προσπαθήσει να εξαλείψει τα συστηματικά σφάλματα, τα οποία προέρχονται απευθείας από το Αριθμητικό Μοντέλο Πρόγνωσης (NWPM). Στη συνέχεια, το κομμάτι του Bayesian αλγορίθμου θα πάει να επιδράσει πάνω στα φιλτραρισμένα δεδομένα (που προέκυψαν από τον αλγόριθμο Kalman), στοχεύοντας στη μείωση του εναπομείναντος μη συστηματικού λευκού θορύβου.

Πιο αναλυτικά, έστω m_i η μέτρηση που δίνει ένα ατμοσφαιρικό μοντέλο τη χρονική στιγμή t , για μία συγκεκριμένη παράμετρο, όπως η θερμοκρασία ή η ταχύτητα του ανέμου, και y_t το σφάλμα στην πρόγνωση. Θεωρούμε ότι το σφάλμα εκφράζεται ως συνάρτηση του m_i και είναι ένα πολυώνυμο της μορφής :

$$y_t = x_{0,t} + x_{1,t} \cdot m_t + x_{2,t} \cdot m_t^2 + \dots + x_{n,t} \cdot m_t^n + v_t \quad (3.8)$$

Οι συντελεστές $x_{0,i}$ είναι οι παράμετροι που πρέπει να υπολογιστούν από το φίλτρο και v_t το Γκαουσιανό, μη συστηματικό σφάλμα της προηγούμενης διαδικασίας.

$$\mathbf{x}_t = [x_{0,t} \ x_{1,t} \ x_{2,t} \ \dots \ x_{n,t}]^T \quad (3.9)$$

Ο πίνακας παρατηρήσεων παίρνει τη μορφή

$$H_t = [1 \ m_t \ m_t^2 \ \dots \ m_t^n] \quad (3.10)$$

Ενώ ως πίνακας συστήματος, χρησιμοποιείται ο ταυτοτικός. Αντίστοιχα, οι εξισώσεις συστήματος και παρατηρήσεων (3.1) και (3.2) θα γίνουν :

$$\mathbf{x}_t = F_t \cdot \mathbf{x}_{t-1} + \mathbf{w}_t \quad (3.11)$$

$$y_t = H_t \cdot \mathbf{x}_t + v_t \quad (3.12)$$

Αξίζει να αναφερθεί ότι η τάξη του πολυωνύμου στην εξίσωση (3.8) μπορεί να διαφοροποιείται κατά περίπτωση. Δηλαδή, η βέλτιστη επιλογή εξαρτάται κάθε φορά από την περίπτωση που μελετάται, από το είδος της μετεωρολογικής παραμέτρου και από την τοποθεσία (Stathopoulos et al. 2013). Ωστόσο, έχει αποδειχθεί σε προηγούμενες μελέτες ότι τα πολυώνυμα 3^{ης} τάξης, έχουν πολύ μεγάλη συνεισφορά στη μείωση των συστηματικών σφαλμάτων με πολλή λίγη χρήση υπολογιστικού χρόνου (Galanis et al. 2011).

Τα \mathbf{w}_t και \mathbf{v}_t , όπως αναφέρεται και πιο πάνω, δίνουν το μη συστηματικό Γκαουσιανό σφάλμα του μοντέλου, με πίνακες συνδιακύμανσης, \mathbf{W}_t και \mathbf{V}_t . Πιο συγκεκριμένα, στην παρούσα εργασία οι πίνακες συνδιακύμανσης υπολογίζονται χρησιμοποιώντας τις τελευταίες 7 τιμές από τα σετ δεδομένων παρατηρήσεων και μοντέλου. Ο αριθμός αυτός των τιμών έχει επιλεγεί προκειμένου να καθιστά το φίλτρο αξιόπιστο και δυναμικά ευέλικτο, κι έχει προκύψει από προηγούμενες δοκιμές (Galanis et al., 2006, 2009, 2011).

$$V_{t_i} \equiv \frac{1}{6} \cdot \sum_{t=0}^6 \left((y_{t_i} - H_{t_i} x_{t_i}) - \left(\frac{\sum_{i=0}^6 (y_{t_i} - H_{t_i} x_{t_i})}{7} \right) \right)^2 \quad (3.13)$$

$$W_{t_i} \equiv \frac{1}{6} \cdot \sum_{t=0}^6 \left((x_{t_i+1} - x_{t_i}) - \left(\frac{\sum_{i=0}^6 (x_{t_i+1} - x_{t_i})}{7} \right) \right)^2 \quad (3.14)$$

Η ουσιαστική διαφορά του συγκεκριμένου φίλτρου σε σχέση με τα κλασσικά, είναι ότι συνδυάζει τις κλασσικές προσεγγίσεις Kalman με τη συμπερασματολογία του Bayes. Συγκεκριμένα, χρησιμοποιείται το επόμενο γραμμικό μοντέλο για τη βελτίωση των φιλτραρισμένων, μέσω Kalman δεδομένων:

$$\mathbf{k}_t = \mathbf{o}_t + \mathbf{v}_t \quad (3.15)$$

Όπου το k_t είναι η εξαγόμενη από το αριθμητικό μοντέλο τιμή, τη χρονική στιγμή t , και η οποία έχει υποστεί «φιλτράρισμα» (μεσω Kalman), o_t είναι η αντίστοιχη τιμή παρατήρησης και v_t ο εναπομείνας Γκαουσιανός μη συστηματικός λευκός θόρυβος.

3.3.3 Το μοντέλο της κατανομής Weibull

Η κατανομή Weibull, όπως αναφέρθηκε και πιο πάνω, θεωρείται η καταλληλότερη για την περιγραφή δεδομένων ταχύτητας ανέμου (χωρίς ωστόσο αυτό να είναι απόλυτο), λόγω του ότι η συνάρτηση κατανομής πιθανότητας είναι θετικά ασύμμετρη. Κάνουμε τις ακόλουθες παραδοχές για τη συνάρτηση πυκνότητας πιθανότητας που πρόκειται να χρησιμοποιήσουμε:

$$P(o_t) \sim W(a, b) \quad (3.16)$$

$$P(v_t) \sim N(0, \sigma_v^2) \quad (3.17)$$

Δηλαδή ότι οι τιμές των παρατηρήσεων ακολουθούν την κατανομή Weibull, ενώ οι τιμές του μη συστηματικού λευκού θορύβου την Κανονική κατανομή.

Σε αυτήν την περίπτωση, η posterior κατανομή για το μοντέλο Weibull παίρνει τη μορφή:

$$P(o_t | \mathbf{k}_t) = \frac{P(o_t | \mathbf{k}_t) P(o_t)}{\int P(o_t | \mathbf{k}_t) P(o_t) do_t} \Rightarrow$$

$$P(o_t | \mathbf{k}_t) = \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k_t - o_t)^2}{2\sigma^2}} \left(\frac{a}{b}\right) \left(\frac{o_t}{b}\right)^{(a-1)} e^{-\left(\frac{o_t}{b}\right)^a}}{\int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k_t - o_t)^2}{2\sigma^2}} \left(\frac{a}{b}\right) \left(\frac{o_t}{b}\right)^{(a-1)} e^{-\left(\frac{o_t}{b}\right)^a} do_t} \Rightarrow$$

$$P(O_t | k_t) = \frac{e^{-\frac{(k_t - O_t)^2}{2\sigma^2}} O_t^{(\alpha-1)} e^{-\left(\frac{O_t}{b}\right)^\alpha}}{\int e^{-\frac{(k_t - o_t)^2}{2\sigma^2}} o_t^{(\alpha-1)} e^{-\left(\frac{o_t}{b}\right)^\alpha} do_t} \quad (3.18)$$

Στον παρονομαστή της παραπάνω εξίσωσης υπάρχει ένα γενικευμένο ολοκλήρωμα για $\alpha < 1$, το οποίο ορίζεται έτσι ώστε η εκ των υστέρων κατανομή να είναι καλώς ορισμένη.

Εφόσον η εκ των υστέρων (posterior) κατανομή εμφανίζει ασυμμετρία, η εκτίμηση της **μέσης τιμής του O_t** γίνεται μέσω της σχέσης:

$$E(O_t) = \frac{\int e^{-\frac{(k_t - o_t)^2}{2\sigma^2}} o_t^\alpha e^{-\left(\frac{o_t}{b}\right)^\alpha} do_t}{\int e^{-\frac{(k_t - o_t)^2}{2\sigma^2}} o_t^{(\alpha-1)} e^{-\left(\frac{o_t}{b}\right)^\alpha} do_t} \quad (3.19)$$

3.4. Μεγέθη ελέγχου απόδοσης

Η χρήση στατιστικών μεθόδων ανάλυσης έχει σκοπό την εκτίμηση των σφαλμάτων της πρόγνωσης και την αξιολόγηση της. Σε μεγέθη συνεχών μεταβλητών, όπως είναι ο άνεμος, ενδιαφέρον παρουσιάζουν τα αριθμητικά σφάλματα, που είναι αποτέλεσμα διαφορετικών τιμών μεταξύ παρατήρησης και πρόγνωσης, καθώς και τα συστηματικά σφάλματα. Στα πλαίσια της διαδικασίας αξιολόγησης των συνεχών μεταβλητών της ταχύτητας του ανέμου έχει καταχωρηθεί μεγάλο πλήθος στατιστικών μεγεθών. Στη συνέχεια αναφέρονται τα μεγέθη που περιλαμβάνονται στη στατιστική ανάλυση του συνόλου των δοκιμών που πραγματοποιήθηκαν.

- **Μέση απόκλιση (Bias)**

Η μέση απόκλιση εκτιμά την αντιστοιχία ανάμεσα στη μέση τιμή της πρόγνωσης και της παρατήρησης. Το συγκεκριμένο μέγεθος υπολογίζει το άθροισμα διαφορών σε σύνολο N τιμών:

$$bias = \frac{1}{N} \sum_{i=1}^N (F_i - O_i) = \bar{F} - \bar{O} \quad (3.20)$$

Σε περίπτωση όπου $bias < 0$, το μοντέλο υποεκτιμά το σύνολο τιμών της συγκεκριμένης μεταβλητής, ενώ όταν το $bias > 0$, η τιμή συνεπάγεται την υπερεκτίμηση της αντίστοιχης τιμής παρατήρησης από το μοντέλο.

- **Τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Square Error-RMSE)**

Είναι από τα διαδεδομένα μεγέθη για την εκτίμηση ακρίβειας της πρόγνωσης και είναι η ρίζα του μέσου τετραγωνικού σφάλματος:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2} \quad (3.21)$$

Η συγκεκριμένη σχέση λαμβάνει τιμές στο διάστημα $[0, +\infty)$ και ιδανικά προγνωστικά πεδία προσεγγίζουν την τιμή 0. Το μέγεθος RMSE δεν εμφανίζεται αδιάστατο, καθώς διατηρεί τις μονάδες των υπό σύγκριση μετεωρολογικών μεταβλητών.

- **Δείκτης Nash-Sutcliffe (NS score)**

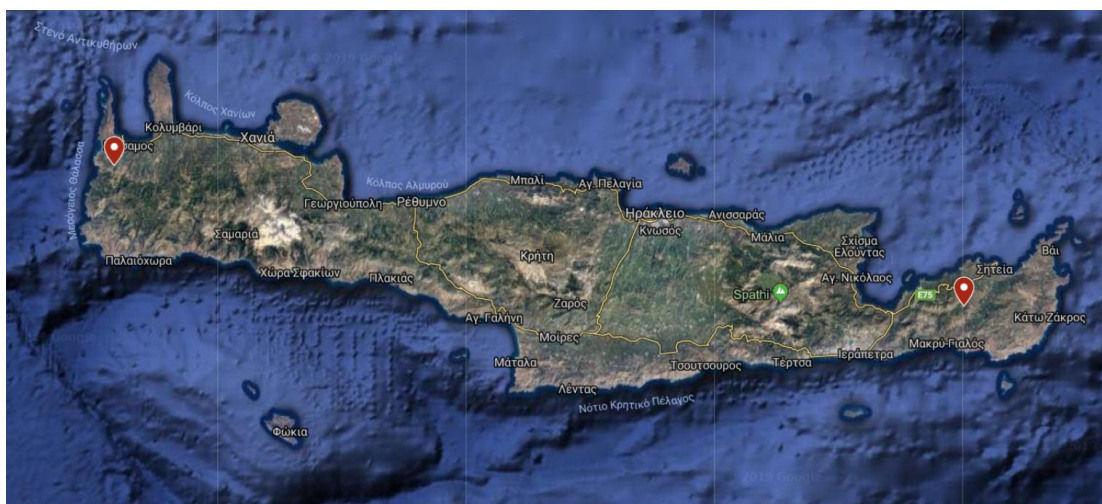
Ο δείκτης Nash-Sutcliffe χρησιμοποιείται προκειμένου να αξιολογηθεί η απόδοση του μοντέλου. Έχει τη μορφή:

$$NS = 1 - \frac{\sum_{i=1}^N (O_i - M_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (3.22)$$

Οι τιμές αυτής της παραμέτρου κυμαίνονται από $-\infty$ έως 1. Η τιμή 1 εκφράζει το ιδανικό, την ταύτιση μεταξύ των παρατηρήσεων και των προγνωστικών αποτελεσμάτων του μοντέλου. Η τιμή 0 σημαίνει ότι η ακρίβεια του μοντέλου προσεγγίζει τη μέση τιμή των παρατηρήσεων.

3.5. Περιοχές μελέτης

Η δοκιμή των φίλτρων Kalman έγινε από δεδομένα τα οποία προήλθαν από δύο αιολικά πάρκα στην Κρήτη. Το πρώτο αιολικό πάρκο, «Αλκυόνη», βρίσκεται στην δυτική πλευρά του νησιού στην περιοχή Κανιαβάς Κισσάμου, του νομού Χανίων (35°44'' Lat και 23°60'' Lon), εγκατεστημένο σε υψόμετρο 265m περίπου (Εικόνα 3.3). Ενώ το δεύτερο αιολικό πάρκο, «Enercon», είναι στο ανατολικό άκρο του νησιού, στην περιοχή Κρυά Σητείας, του νομού Λασιθίου (35°12'' Lat και 26°03'' Lon) σε υψόμετρο 450m (Εικόνα 3.3).



Εικόνα 3.3: Δορυφορική εικόνα του νησιού της Κρήτης, με επισήμανση των τοποθεσιών του αιολικού πάρκου «Αλκυόνη» (αριστερά) και του αιολικού πάρκου «Enercon» (δεξιά).

Τα δεδομένα παρατήρησης και για τα δύο αιολικά πάρκα καλύπτουν τις περιόδους από 03/09/2014 έως 24/02/2015. Οι παρατηρήσεις λαμβάνονταν ανά λεπτό σε τοπική ώρα (LTC) και περιελάμβαναν τιμές ταχύτητας και διεύθυνσης ανέμου σε m/s και μοίρες αντίστοιχα στο ύψος του πάρκου.

3.6. Δεδομένα Πρόγνωσης

Τα δεδομένα πρόγνωσης του ανέμου που χρησιμοποιήθηκαν προήλθαν από το μετεωρολογικό μοντέλο RAMS (παράγραφος 3.2) και αφορούν την περίοδο από 03/09/2014 έως 24/02/2015 (ωριαίες τιμές) και το διάστημα 30/10/2014 έως 4/11/2014 (τιμές ανά λεπτό). Η μετεωρολογική παράμετρος που εξετάστηκε ήταν η ταχύτητα του ανέμου. Ένα μετεωρολογικό μοντέλο, έχει τη δυνατότητα να υπολογίζει την ταχύτητα

του ανέμου σε διάφορα ύψη, στα διάφορα πλεγματικά σημεία. Η ανάλυση του μοντέλου που χρησιμοποιήθηκε ήταν 3km x 3km, ενώ οι τιμές του ανέμου που χρησιμοποιήθηκαν ήταν για το ύψος των 60 μέτρων. Στο επόμενο κεφάλαιο αναλύεται η εφαρμογή των Bayesian Kalman φίλτρων που χρησιμοποιήθηκαν προκειμένου να βελτιώσουν τα αποτελέσματα του μοντέλου πρόγνωσης, στις περιοχές ενδιαφέροντος.

ΚΕΦΑΛΑΙΟ 4ο

ΕΦΑΡΜΟΓΗ ΤΩΝ ΦΙΛΤΡΩΝ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1. Ανάλυση δοκιμών

Παρακάτω εξετάζονται τα αποτελέσματα όλων των παραμέτρων του φίλτρου που διερευνήθηκαν. Οι βασικές παράμετροι ήταν τα χρονικά βήματα υπολογισμού, τα βήματα «ιστορικότητας» (δηλαδή η περίοδος εκπαίδευσης του αλγορίθμου), ο βαθμός του πολυωνύμου και ο έλεγχος της επιλογής της κατάλληλης κατανομής.

Στους παρακάτω πίνακες ως *forecast* αναφέρονται τα χρονικά βήματα με τα οποία γινόταν η εκάστοτε διόρθωση, ως *history* τα βήματα «ιστορικότητας» (training period για τον αλγόριθμο) για μία ορισμένη χρονική περίοδο (forecasting period). Εντός των προγνωστικών περιόδων γίνεται κάθε φορά ένας αριθμός προγνώσεων (predictions) ανάλογα με το χρονικό βήμα «forecast», που έχει επιλεγεί κάθε φορά. Η εκτέλεση των διορθώσεων γίνεται διακριτά διαστήματα (αριθμός «forecast») στο διάστημα αυτό, κυλιόμενα προς τα μπρος χρονικά, και ο χρόνος «εκπαίδευσης» ανανεώνεται με τις πιο νέες τιμές. Μελετώνται ποσοτικά δύο τύποι σφαλμάτων, το bias (μέση απόκλιση) και το RMSE (τετραγωνική ρίζα μέσου τετραγωνικού σφάλματος), ενώ ο δείκτης Nash-Sutcliffe (NS), ο οποίος χρησιμοποιείται ως δείκτης αξιολόγησης της απόδοσης του μοντέλου, θα αξιοποιηθεί προκειμένου να δούμε πότε πετυχαίνουμε συνδυαστικά την καλύτερη μείωση σφαλμάτων. Ο έλεγχος της καταλληλότερης κατανομής έγινε με τη χρήση του κριτηρίου NlogL (negative of the likelihood).

Σε όλα τα παρακάτω γραφήματα παρουσιάζονται τρεις χρονοσειρές, με *μπλε* χρώμα είναι η χρονοσειρά των παρατηρήσεων, με *κόκκινο* η χρονοσειρά που προέκυψε ύστερα από πρόγνωση του μοντέλου και με *πράσινο* αυτή που προκύπτει ύστερα από τη διόρθωση του φίλτρου που εφαρμόστηκε.

4.2. Εξέταση βέλτιστης τάξης πολυωνύμου

Παρακάτω εξετάζεται ως παράμετρος του φίλτρου, η διάσταση του πολυωνύμου με την οποία το φίλτρο έχει τη βέλτιστη απόδοση. Εξετάστηκαν τρεις τάξεις του πολυωνύμου, 1^{ns} (dim=1), 2^{as} (dim=2) και 3^{ns} (dim=3).

Αναζητώντας την καταλληλότερη διάσταση, σύμφωνα με το σύνολο των περιπτώσεων οι οποίες εξετάστηκαν, μπορεί κανείς να καταλήξει στο συμπέρασμα ότι η 3^{ns} τάξη δεν έχει μεγάλη συνέπεια σε σχέση με τις άλλες δύο. Επίσης, όπως φαίνεται και στους Πίνακες 4.1 και 4.2 παρακάτω, σε ορισμένες περιπτώσεις η 3^{ns} τάξη μπορεί να μειώνει λίγο παραπάνω το bias, σε σχέση με την 2^{ns} τάξη, ωστόσο οι διαφορές είναι αμελητέες ενώ ταυτόχρονα δίνει λίγο μεγαλύτερες τιμές για το RMSE, σε σχέση με την 2^{ns} τάξη του πολυωνύμου. Η 1^{ns} τάξη δείχνει να δίνει καλύτερα αποτελέσματα σε σχέση με την 3^{ns} και μπορεί να θεωρηθεί αξιόπιστη. Επιπλέον, σε κάποιες περιπτώσεις το πολυώνυμο 1^{ns} τάξης μπορεί να μειώνει ελάχιστα παραπάνω το bias απ' ότι το πολυώνυμο 2^{as} τάξης, αλλά συνολικά υστερεί σε σχέση με την 2^{ns} . Οι διαφορές για το bias δεν είναι ουσιαστικές, καθώς το φίλτρο καταφέρνει σε όλες τις περιπτώσεις πολύ χαμηλές τιμές (σχεδόν εξαλείφεται) για αυτόν τον τύπο σφάλματος. Όσον αφορά το σφάλμα RMSE, οι διαφορές είναι λίγο περισσότερο έντονες και ιδιαίτερα για τα μεγαλύτερα προγνωστικά βήματα (f). Ο βαθμός του πολυωνύμου δείχνει να μην επηρεάζεται πολύ καθώς μειώνονται τα προγνωστικά βήματα (Πίνακες 4.1-4.2).

Στα σχήματα 4.1 έως 4.4 φαίνεται ποσοτικά πως από τα συνολικά αποτελέσματα, η 2^{ns} τάξη λειτούργησε αρκετά καλύτερα από την 3^{ns} , αλλά και καλύτερα από την 1^{ns} . Συγκεκριμένα, στο 68% των περιπτώσεων που αφορούν τα αποτελέσματα του bias, το 2^{ns} βαθμού πολυώνυμο έδωσε καλύτερα αποτελέσματα από το πολυώνυμο 3^{ns} βαθμού (Σχήμα 4.3), ενώ όσον αφορά το RMSE, στο 66% των περιπτώσεων έφερε καλύτερα αποτελέσματα σε σχέση με το πολυώνυμο 3^{ns} βαθμού (Σχήμα 4.4). Αντίστοιχα, στη σύγκριση μεταξύ 2^{ns} και 1^{ns} βαθμού, φαίνεται (Σχήμα 4.1) ότι στο 58% των συνολικών περιπτώσεων που αφορούν το σφάλμα bias, το 2^{ns} βαθμού πολυώνυμο έφερε καλύτερα αποτελέσματα. Τέλος, φαίνεται (Σχήμα 4.2) ότι και στη σύγκριση των συνολικών αποτελεσμάτων για το RMSE, στο 61% των περιπτώσεων το 2^{ns} βαθμού πολυώνυμο ήταν πιο αποτελεσματικό από αυτό του 1^{ns} . Καταλήγοντας μπορούμε να πούμε ότι παρότι σε πολλές περιπτώσεις οι διαφορές μπορεί να ήταν μικρές, η 2^{ns} τάξη του πολυωνύμου εμφανίζει ένα μικρό πλεονέκτημα σε σχέση με τις άλλες δύο. Επιπλέον,

όπως φάνηκε από τα αποτελέσματα, και στις τρεις περιπτώσεις είχαμε αρκετά καλές διορθώσεις τόσο στη μέση απόκλιση όσο και στο RMSE.

Πίνακας 4.1 Σύγκριση Αποτελεσμάτων φίλτρου για 3 τάξεις πολ/μου για το bias. Προγνωστική περίοδος: 48hr, Κατανομή: Weibull

BIAS/dim=2							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-4,88	-1,47	-0,07	0,48	0,51	0,38	0,39
h=12	-4,88	-1,34	0,03	0,20	0,26	0,16	0,19
h=6	-4,88	-0,43	-0,01	0,21	0,08	-0,02	-0,03
h=3	-4,88	-1,60	-0,56	-0,18	-0,04	-0,12	-0,07

BIAS/dim=1							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-4,88	-1,37	-0,16	0,28	0,33	0,28	0,30
h=12	-4,88	-1,34	-0,11	-0,01	0,14	0,09	0,12
h=6	-4,88	-0,50	-0,25	0,01	-0,03	-0,07	-0,07
h=3	-4,88	-1,68	-0,69	-0,29	-0,12	-0,12	-0,07

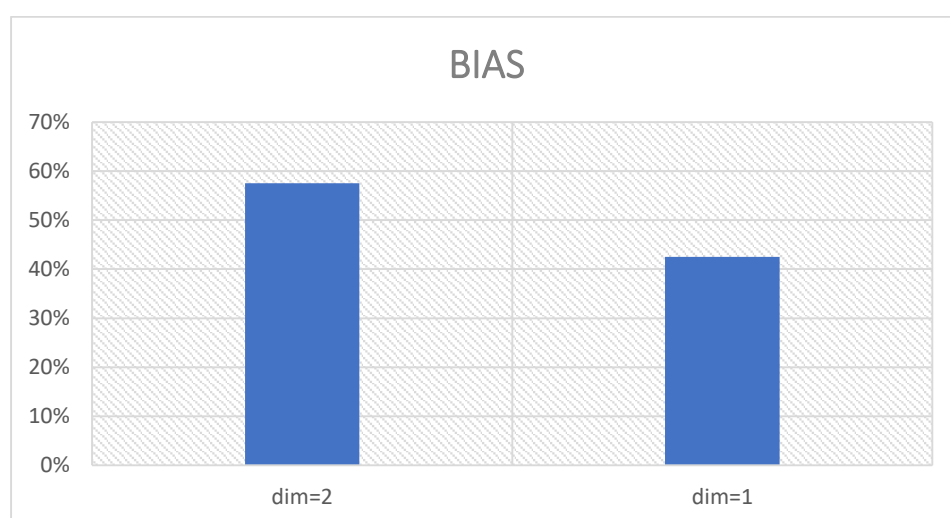
BIAS/dim=3							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-4,88	-2,99	-0,43	0,46	0,54	0,41	0,41
h=12	-4,88	-2,16	-0,24	0,25	0,28	0,19	0,20
h=6	-4,88	-1,08	-0,01	0,20	0,09	0,00	-0,01
h=3	-4,88	NaN	NaN	-0,20	-0,02	-0,12	-0,05

Πίνακας 4.2 Σύγκριση Αποτελεσμάτων φίλτρου για 3 τάξεις πολ/μου για το RMSE. Προγνωστική περίοδος: 48hr, Κατανομή: Weibull

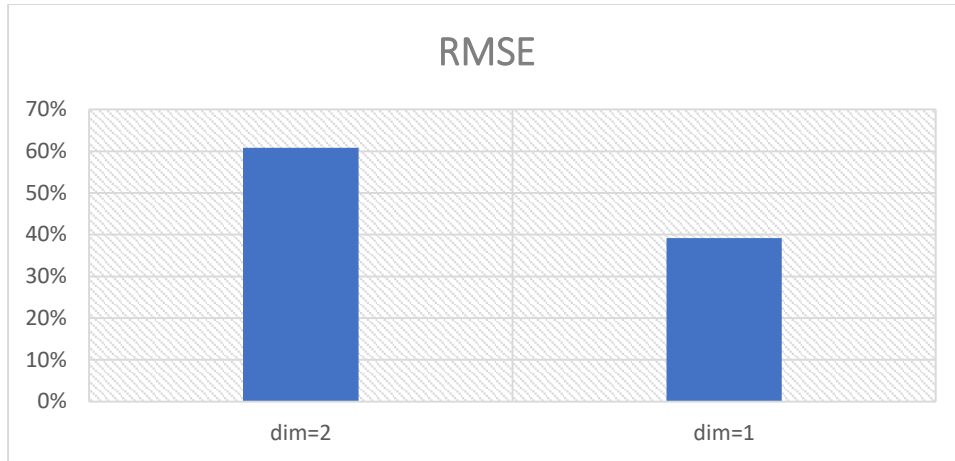
RMSE/dim=2							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	5,38	4,21	3,24	2,44	1,97	2,02	1,81
h=12	5,38	3,63	3,04	2,25	1,80	1,91	1,67
h=6	5,38	2,29	2,18	2,19	1,79	1,91	1,65
h=3	5,38	2,81	2,60	2,39	1,90	1,93	1,66

RMSE/dim=1							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	5,38	3,78	3,02	2,52	2,06	2,01	1,87
h=12	5,38	3,31	2,84	2,26	1,85	1,88	1,72
h=6	5,38	2,21	2,08	2,29	1,87	1,91	1,71
h=3	5,38	2,67	2,46	2,49	1,95	1,93	1,71

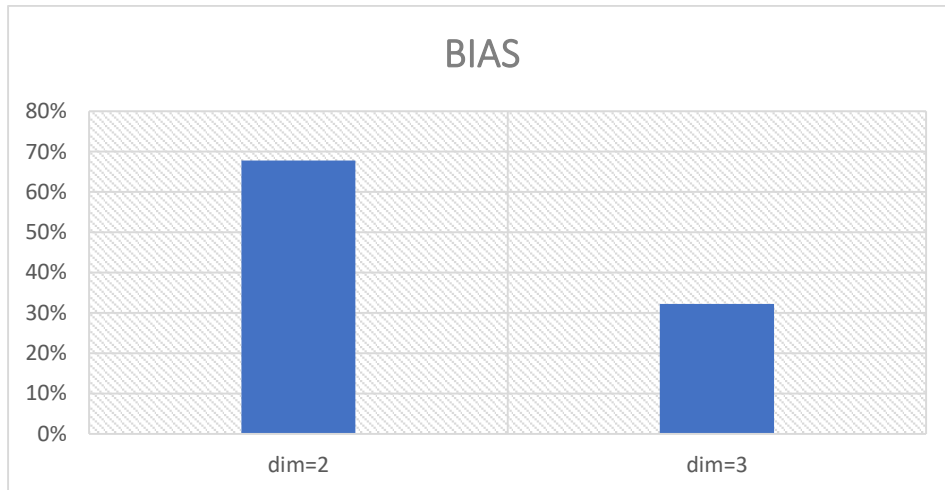
RMSE/dim=3							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	5,38	6,10	4,04	2,57	1,99	2,06	1,81
h=12	5,38	4,47	3,62	2,44	1,85	1,97	1,68
h=6	5,38	2,71	2,58	2,23	1,80	1,94	1,66
h=3	5,38	NaN	NaN	2,47	1,92	1,95	1,67



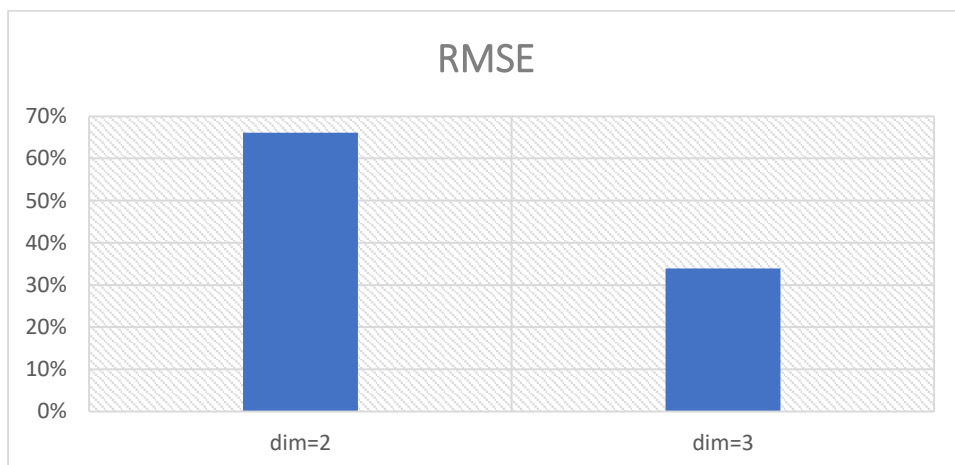
Σχήμα 4.1 Σύγκριση επιδόσεων μεταξύ της 2^{ης} με την 1^η τάξη πολυωνόμου για τα συνολικά αποτελέσματα του bias



Σχήμα 4.2 Σύγκριση επιδόσεων μεταξύ της 2^{ης} με την 1^η τάξη πολυωνύμου για τα συνολικά αποτελέσματα του *RMSE*



Σχήμα 4.3 Σύγκριση επιδόσεων μεταξύ της 2^{ης} με την 3^η τάξη πολυωνύμου για τα συνολικά αποτελέσματα του *bias*



Σχήμα 4.4 Σύγκριση επιδόσεων μεταξύ της 2^{ης} με την 3^η τάξη πολυωνύμου για τα συνολικά αποτελέσματα του *RMSE*

4.3. Εξέταση των βημάτων πρόγνωσης και της περιόδου «εκπαίδευσης»

Τα βήματα πρόγνωσης (*forecast*) μαζί με την περίοδο εκπαίδευσης (*history*) του αλγορίθμου αποτελούν τις πιο σημαντικές παραμέτρους του φίλτρου. Τα βήματα πρόγνωσης είναι ιδιαίτερα σημαντικά καθώς καθορίζουν τον χρονικό ορίζοντα αξιοπιστίας στον οποίο μπορεί κανείς να βασιστεί, αλλά και θέτουν τους περιορισμούς στα είδη των εφαρμογών στα οποία μπορεί να χρησιμοποιηθεί το φίλτρο. Τα βήματα «ιστορικότητας» ως παράμετρος, δηλαδή το χρονικό διάστημα το οποίο απαιτεί το φίλτρο προκειμένου να προσαρμοστεί στα δεδομένα που του δίνονται, είναι εξίσου σημαντικά, καθώς καθορίζουν την ευελιξία της μεθόδου και τις απαιτήσεις σε δεδομένα πριν χρησιμοποιηθεί το φίλτρο.

4.3.1 Πάρκο «Αλκυόνη», Κίσσαμος Χανίων

- *Προγνωστικές περίοδοι περίπου 2 ημερών*

Τα παρακάτω αποτελέσματα στην πρώτη περίπτωση (Σχήματα 4.7 και 4.8) προήλθαν από dataset της χρονικής περιόδου από 17/1/2015 έως 20/1/2015, η προγνωστική περίοδος (*forecasting period*) αφορά διάστημα 54 ωρών (στις 54 ώρες αλλάζει ο αριθμός των προγνώσεων ανάλογα με το χρονικό βήμα, f , που έχει επιλεγεί), χρησιμοποιήθηκε η κατανομή Weibull και το πολυώνυμο ήταν 2^ο βαθμού. Στη δεύτερη περίπτωση, η χρονική περίοδος αφορά το διάστημα από 24/11/2014 έως 26/1/2014, η προγνωστική περίοδος αφορά διάστημα 54 ωρών, χρησιμοποιήθηκε η κατανομή Weibull και το πολυώνυμο ήταν 2^ο βαθμού. Στην τρίτη περίπτωση, η χρονική περίοδος ήταν από 27/11/2014 έως 29/11/014, η κατανομή που ταίριαζε καλύτερα στα δεδομένα ήταν η lognormal και χρησιμοποιήθηκε και πάλι πολ/μο 2^ο βαθμού.

Πίνακας 4.3 Προγνωστική περίοδος: 54hr, Dataset :72hr, Κατανομή: Weibull, Βαθμός Πολ/μο: 2^{ου}

BIAS							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-6,77	0,49	0,42	-0,06	0,24	0,19	0,23
h=12	-6,77	1,11	0,61	0,07	0,21	0,16	0,16
h=6	-6,77	0,00	-0,19	0,21	0,31	0,16	0,14
h=3	-6,77	-0,18	0,31	0,63	0,29	0,16	0,10

RMSE							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	7,74	5,00	3,40	3,25	2,72	2,54	2,23
h=12	7,74	5,39	3,29	2,01	2,15	1,70	1,70
h=6	7,74	5,65	2,66	1,97	2,02	1,58	1,62
h=3	7,74	5,08	3,20	2,90	1,91	1,92	1,61

NS							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-3,74	-0,98	0,09	0,16	0,42	0,49	0,61
h=12	-3,74	-1,30	0,14	0,68	0,64	0,77	0,77
h=6	-3,74	-1,52	0,44	0,69	0,68	0,80	0,79
h=3	-3,74	-1,04	0,19	0,34	0,71	0,71	0,80

Πίνακας 4.4 Προγνωστική περίοδος: 52hr, Dataset :76hr, Κατανομή: , Βαθμός Πολ/μο: 2^{ου}

BIAS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	2,40	0,17	-0,34	-0,57	-0,34	-0,22	-0,18
h=12	2,40	-0,30	-0,56	-0,72	-0,50	-0,36	-0,31
h=6	2,40	-1,86	-0,74	-0,31	-0,28	-0,26	-0,20
h=3	2,40	-1,15	-0,28	-0,11	-0,29	-0,06	-0,10

RMSE							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	2,90	2,41	2,20	1,70	1,59	1,57	1,58
h=12	2,90	2,94	2,55	1,95	1,62	1,58	1,54
h=6	2,90	3,35	2,37	1,67	1,47	1,46	1,38
h=3	2,90	2,30	1,75	1,52	1,53	1,33	1,37

NS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-0,85	-0,27	-0,06	0,37	0,45	0,46	0,45
h=12	-0,85	-0,90	-0,43	0,16	0,42	0,45	0,48
h=6	-0,85	-1,46	-0,23	0,39	0,53	0,53	0,58
h=3	-0,85	-0,16	0,32	0,49	0,48	0,61	0,59

Πίνακας 4.5 Προγνωστική περίοδος: 48hr, Dataset :72hr, Κατανομή: Weibull, Βαθμός Πολ/μο: 2^{ου}

BIAS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-4,88	-1,47	-0,07	0,48	0,51	0,38	0,39
h=12	-4,88	-1,34	0,03	0,20	0,26	0,16	0,19
h=6	-4,88	-0,43	-0,01	0,21	0,08	-0,02	-0,03
h=3	-4,88	-1,60	-0,56	-0,18	-0,04	-0,12	-0,07

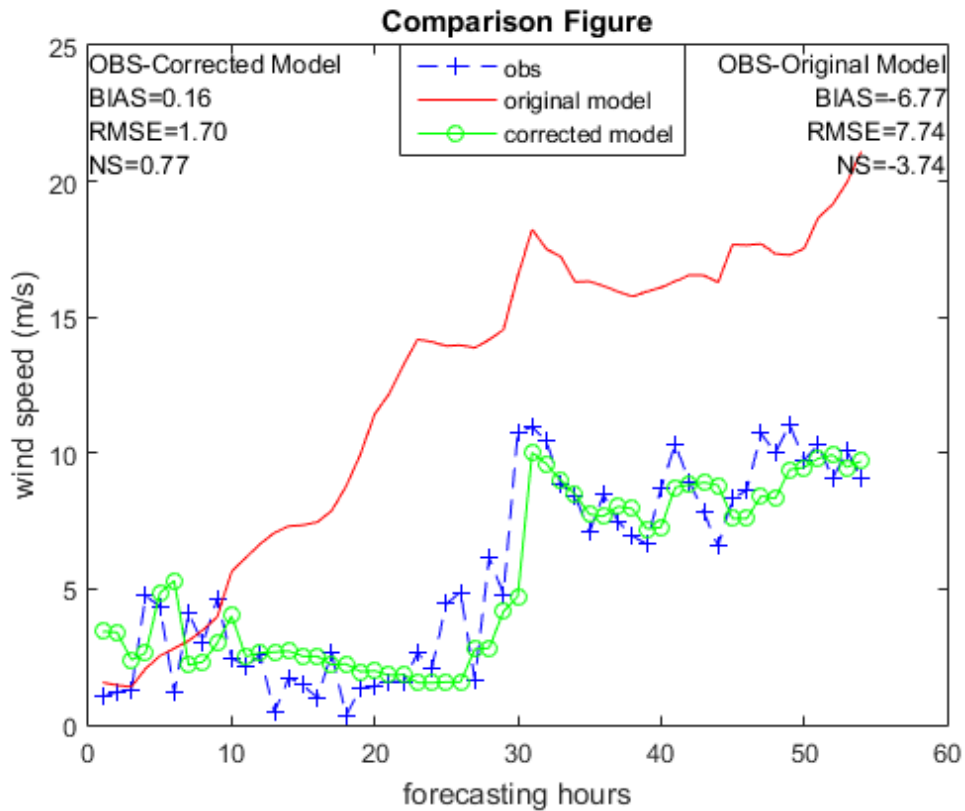
RMSE							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	5,38	4,21	3,24	2,44	1,97	2,02	1,81
h=12	5,38	3,63	3,04	2,25	1,80	1,91	1,67
h=6	5,38	2,29	2,18	2,19	1,79	1,91	1,65
h=3	5,38	2,81	2,60	2,39	1,90	1,93	1,66

NS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-4,97	-2,67	-1,17	-0,23	0,20	0,16	0,32
h=12	-4,97	-1,73	-0,91	-0,05	0,33	0,24	0,42
h=6	-4,97	-0,08	0,02	0,01	0,34	0,24	0,44
h=3	-4,97	-0,63	-0,39	-0,18	0,25	0,23	0,43

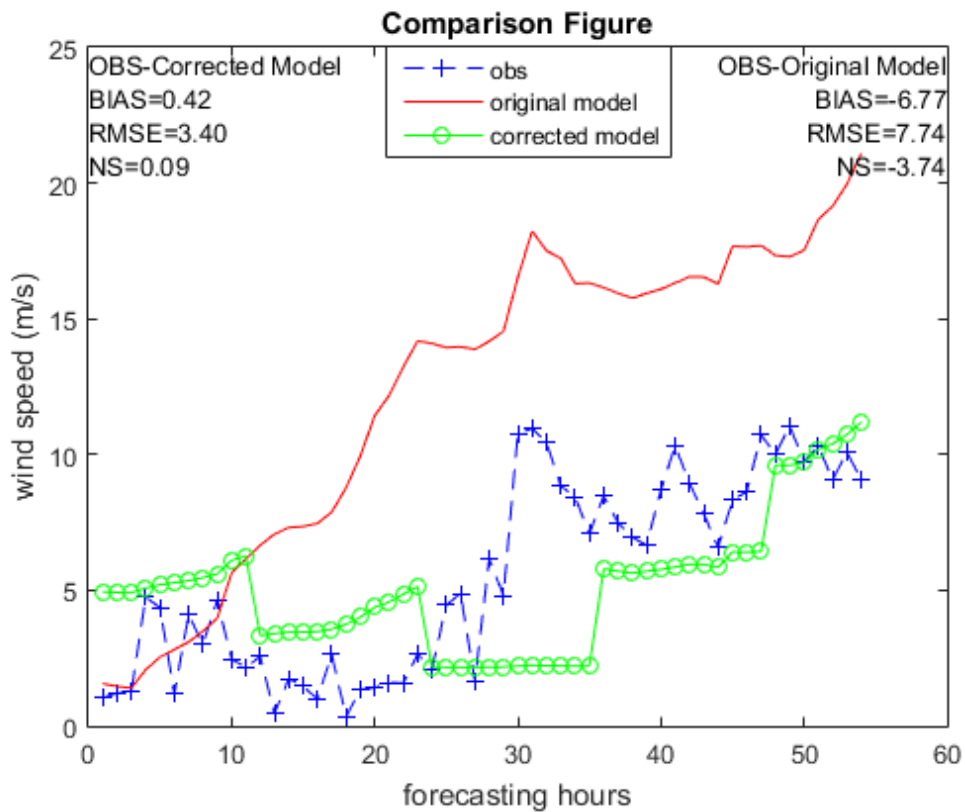
Όπως φαίνεται από τους Πίνακες 4.3 - 4.5 για τα προγνωστικά βήματα $f=6$, $f=3$, $f=2$ και $f=1$, τα αποτελέσματα είναι εξαιρετικά. Το bias έχει σχεδόν εξαλειφθεί και το RMSE έχει μειωθεί σε όλες τις περιπτώσεις πάνω από 45%, ενώ και ο δείκτης NS είναι σχεδόν σε όλες τις περιπτώσεις θετικός και σε ορισμένες περιπτώσεις πλησιάζει το 0.8. Στα μικρά αυτά προγνωστικά βήματα το φίλτρο καταφέρνει αρκετά καλά να «ακολουθήσει» τόσο τη φάση όσο και τις διακυμάνσεις, όπως φαίνεται κι από τις πράσινες καμπύλες στα Σχήματα 4.5 και 4.7. Σημαντικό είναι να αναφερθεί επίσης το γεγονός ότι στα μικρά προγνωστικά βήματα, όπως το $f=3$, η περίοδος εκπαίδευσης δεν πρέπει να είναι πολλή μεγάλη. Παρατηρείται καλύτερη συμπεριφορά εάν τα βήματα «ιστορικότητας» είναι έως $h=12$, σε σχέση με το $h=24$, κάτι που αποδεικνύεται και από τις επιδόσεις του δείκτη NS, όπως φαίνεται και στους Πίνακες 4.3-4.5. Η καμπύλη των διορθωμένων τιμών του ανέμου σχεδόν συμπίπτει με αυτή των παρατηρήσεων. Στις

τρεις αυτές περιπτώσεις (προγνωστικών περιόδων) το βήμα $f=12$ παρά τις μειώσεις στα σφάλματα που παρατηρούμε στους πίνακες, πρακτικά δεν έχει λειτουργήσει το φίλτρο. Στις περιπτώσεις αυτές παρατηρούνται κάποιες εξομοιωμένες, απότομες μεταβάσεις μεταξύ των χρονικών βημάτων, στην χρονοσειρά που έχει προκύψει ύστερα από την εφαρμογή του φίλτρου (Σχήματα 4.6). Αυτές οι απότομες μεταβάσεις είναι ένδειξη ότι το φίλτρο πρακτικά δεν έχει λειτουργήσει και προέρχονται από την Bayesian διόρθωση (σχέση 3.19), όπου στο προγνωστικό βήμα έχει υπολογιστεί ένας μέσος όρος ο οποίος έχει επικαλύψει όλες τις άλλες τιμές. Η αιτία ήταν ότι γινόταν κακή χρήση του συστήματος, ξεπερνώντας την περιοδικότητα του φίλτρου μέσω της κακής σχέσης μεταξύ προγνωστικών χρονικών βημάτων και χρονικών βημάτων «εκπαίδευσης» που εφαρμόζονταν. Όσον αφορά το προγνωστικό βήμα $f=24$, στα αποτελέσματα μπορεί να φαίνεται ότι υπάρχει ελάττωση τόσο του bias όσο και του RMSE (έως και 35%), ωστόσο τα αποτελέσματα αυτά δεν έχουν πραγματικό αντίκρισμα αφού όπως βλέπουμε στο σχήμα, το φίλτρο δεν έχει λειτουργήσει. Σε αυτήν την περίπτωση οι τιμές που δίνει το φίλτρο απλώς ακολουθούν τις τιμές του μοντέλου.

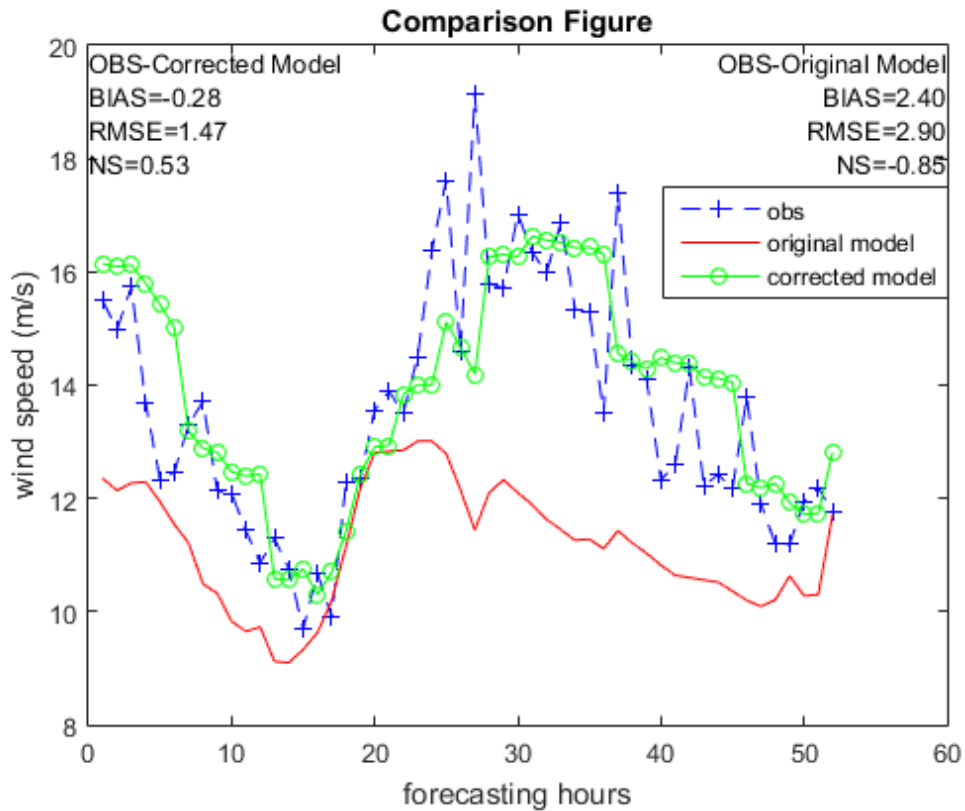
Εν κατακλείδι, μπορεί να εξαχθεί το συμπέρασμα ότι ένα σε δεδομένων αυτού του τύπου και αυτού του μεγέθους λειτουργεί άψογα για τα προγνωστικά βήματα από $f=6$ και κάτω. Τα βήματα $f=12$ και $f=24$ δεν λειτουργούν.



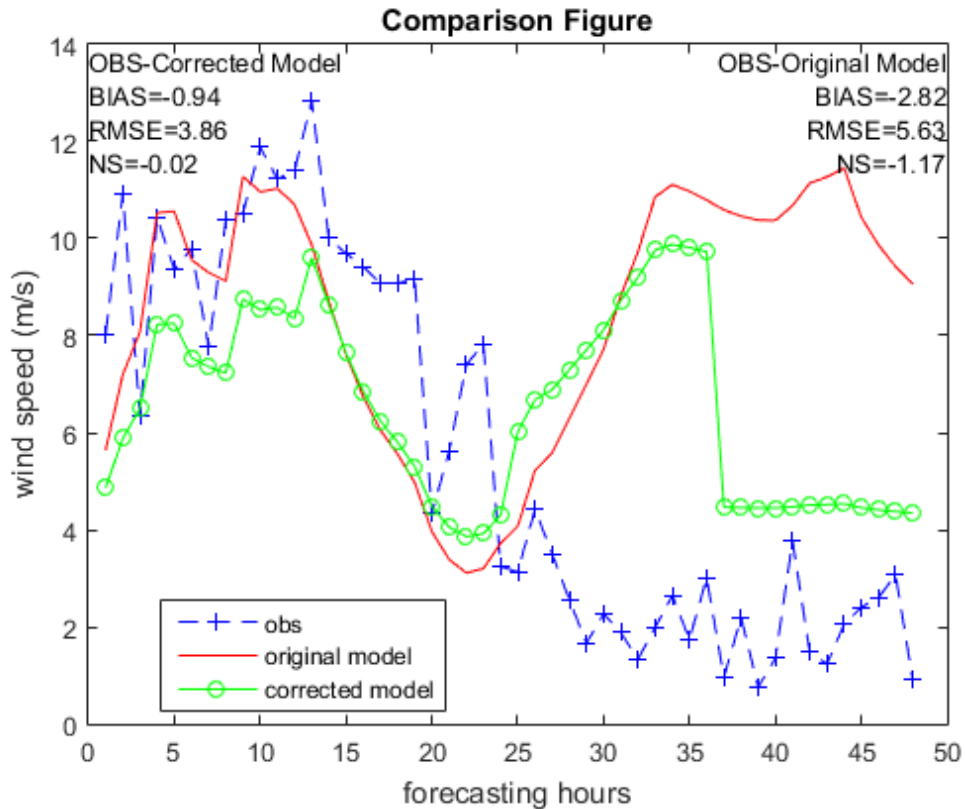
Σχήμα 4.5: Τα διορθωμένα αποτελέσματα για $f=2$ και $h=12$



Σχήμα 4.6: Τα διορθωμένα αποτελέσματα για $f=12$ και $h=24$



Σχήμα 4.7: Τα διορθωμένα αποτελέσματα για $f=3$ και $h=6$



Σχήμα 4.8: Τα διορθωμένα αποτελέσματα για $f=12$ και $h=24$

- **Προγνωστική περίοδος ενός μήνα περίπου**

Τα παρακάτω αποτελέσματα προήλθαν από σειτ δεδομένων 720 ωρών, κατά την περίοδο από 3/8/2014 έως την 29/8/2014. Η κατανομή η οποία ταίριαζε καλύτερα στα συγκεκριμένα δεδομένα ήταν η Weibull κατανομή, το πολυώνυμο ήταν 2^ο βαθμού, ενώ η προγνωστική περίοδος αφορά διάστημα 648 ωρών (27 ημερών). Διευκρινίζεται ότι η προγνωστική περίοδος αφορά συγκεκριμένο αριθμό προγνώσεων (predictions), ανάλογα με το χρονικό βήμα που έχει επιλεγεί, και όχι πρόγνωση 648 ωρών.

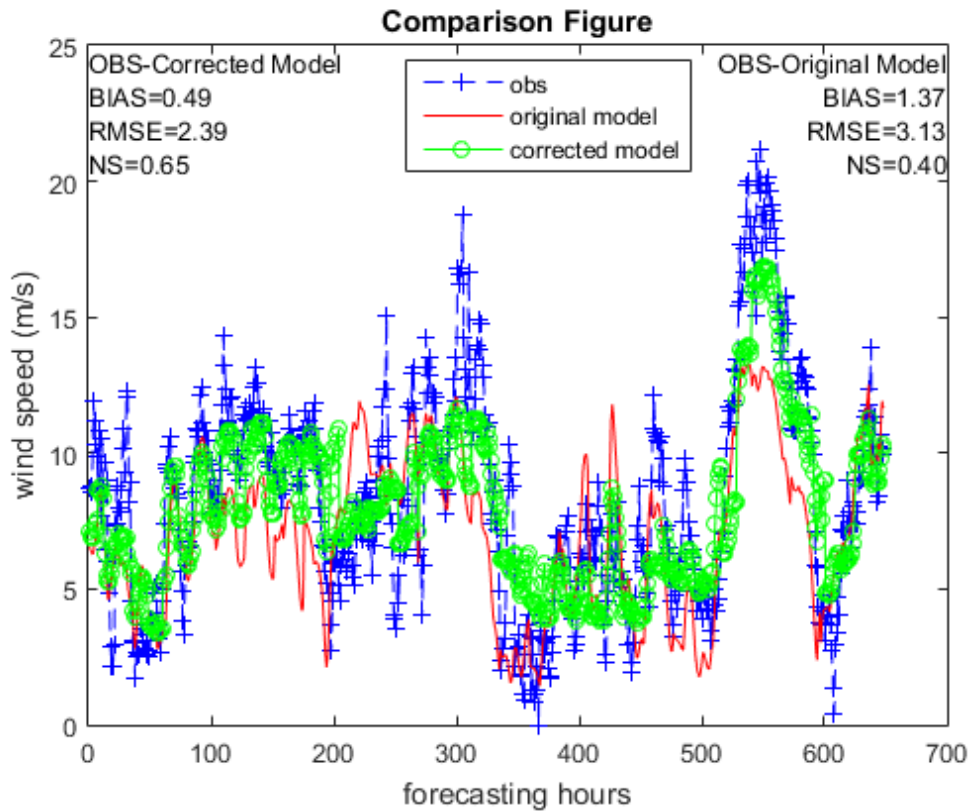
Πίνακας 4.6 Προγνωστική περίοδος: 648hr, Dataset :720hr, Κατανομή: Weibull, Βαθμός Πολ/μου: 2^{ος}

BIAS								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	1,37	-0,50	0,49	0,49	0,22	0,10	0,12	0,08
h=48	1,37	0,29	0,44	0,50	0,22	0,09	0,12	0,08
h=24	1,37	-0,52	0,41	0,52	0,20	0,09	0,11	0,07
h=12	1,37	0,61	0,53	0,41	0,21	0,09	0,11	0,07
h=6	1,37	-0,50	0,04	0,03	0,18	0,07	0,08	0,04
h=3	1,37	0,45	-0,03	-0,29	-0,10	0,03	-0,03	0,02

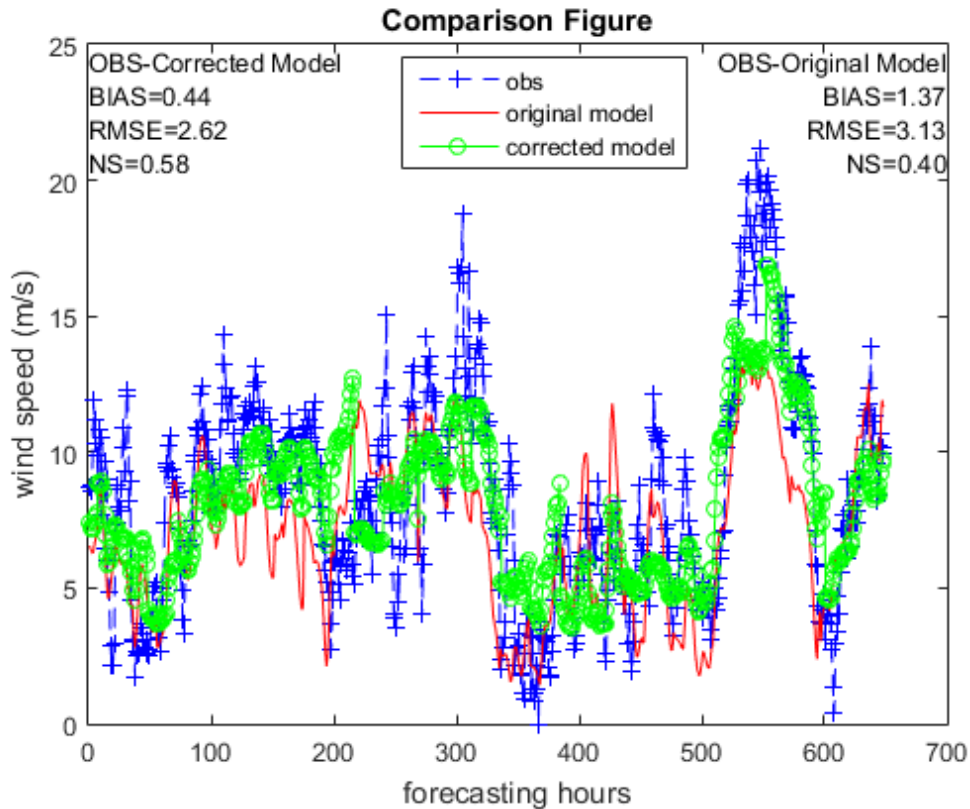
RMSE								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	3,13	3,96	2,67	2,39	2,24	2,08	1,94	1,76
h=48	3,13	4,58	2,62	2,38	2,23	2,06	1,90	1,72
h=24	3,13	3,75	2,56	2,47	2,20	1,97	1,80	1,62
h=12	3,13	3,57	3,18	2,52	2,19	1,93	1,76	1,57
h=6	3,13	3,96	3,11	2,68	2,17	1,91	1,72	1,51
h=3	3,13	3,34	3,11	2,74	2,54	1,91	1,68	1,40

NS								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	0,40	0,04	0,56	0,65	0,69	0,74	0,77	0,81
h=48	0,40	-0,28	0,58	0,65	0,70	0,74	0,78	0,82
h=24	0,40	0,14	0,60	0,63	0,70	0,76	0,80	0,84
h=12	0,40	0,22	0,38	0,61	0,71	0,77	0,81	0,85
h=6	0,40	0,04	0,41	0,56	0,71	0,78	0,82	0,86
h=3	0,40	0,32	0,41	0,54	0,61	0,78	0,83	0,88

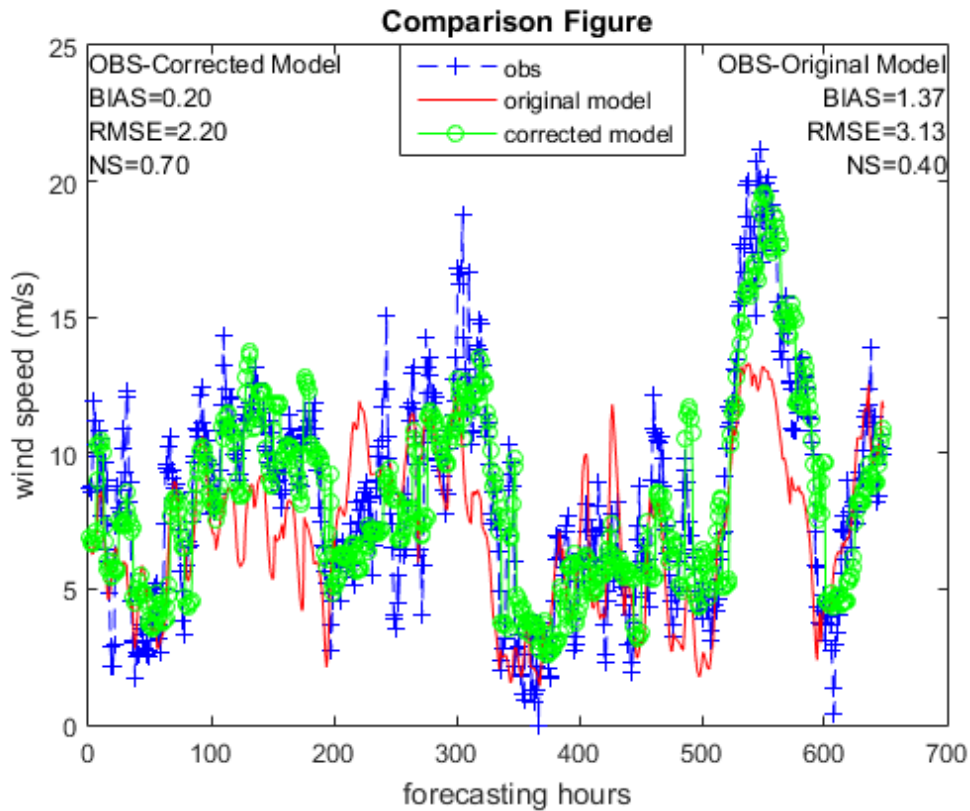
Το bias μειώθηκε σε όλες τις περιπτώσεις άνω του 55%, ενώ το RMSE σε όλες τις περιπτώσεις πλην του $f = 48$. Το προγνωστικό βήμα των 48 ωρών ($f=48$) μπορεί να μειώνει τη μέση απόκλιση, ωστόσο το RMSE αυξάνεται. Το bias σε όλες τις βέλτιστες περιπτώσεις έχει σχεδόν μηδενιστεί, ενώ το RMSE μειώνεται περισσότερο για $history > forecast$. Για $f=3,2,1$ τα βέλτιστα history είναι $h=12$, $h=6$ και $h=3$. Στην περίπτωση του $f=1$ το RMSE μειώνεται πάνω από 50%, ενώ το bias έχει σχεδόν εξαλειφθεί. Στις περιπτώσεις όπου το προγνωστικό βήμα είναι από 6 ώρες και κάτω, τα αποτελέσματα είναι εξαιρετικά για όλα τα βήματα ιστορικότητας, history, τόσο για το RMSE όσο και για το bias, το οποίο έχει εξαλειφθεί, κάτι το οποίο μαρτυρά και ο δείκτης NS, ο οποίος είναι κοντά στο 1. Τέλος, στις περιπτώσεις όπου $f=24$ μπορούμε να πούμε ότι τα αποτελέσματα είναι αξιόλογα για $h=48$, $h=72$ και $h=24$. Οι σχεδόν μηδενικές τιμές που φαίνονται για $f=24$ και $h=6$ ή $h=3$ είναι πλασματικές, αφού σε αυτές τις περιπτώσεις το φίλτρο δεν έχει λειτουργήσει κάτι το οποίο γίνεται αντιληπτό από το RMSE, το οποίο δεν έχει μειωθεί αλλά και από τον δείκτη NS ο οποίος αυξάνει αμυδρά. Αξίζει να σημειωθεί ότι σε ένα σετ δεδομένων μεγαλύτερου μεγέθους όπως αυτό φαίνεται ότι και το προγνωστικό βήμα των 24 ωρών να λειτουργεί, αλλά απαιτείται μία περίοδος «εκπαίδευσης» του αλγορίθμου τουλάχιστον διπλάσια (στην περίπτωση του Πίνακα 4.6 λειτούργησε για $h=48$ και $h=72$), για να γίνει μία διόρθωση της τάξης του 15% (Σχήμα 4.10). Σε όλες τις παραπάνω περιπτώσεις πρέπει να λάβουμε υπόψιν μας και το ιδιαίτερα χαμηλό σφάλμα που έδινε το μοντέλο. Όπως φαίνεται και στα παρακάτω γραφήματα, όσο πιο μικρό βήμα f έχουμε, το φίλτρο είναι ικανό να «πιάσει» και τις έντονες διακυμάνσεις. Στην πραγματικότητα είχαμε μία ήδη καλή πρόγνωση από το μοντέλο, την οποία το φίλτρο κατάφερε να φέρει ακόμα πιο κοντά στην πραγματικότητα.



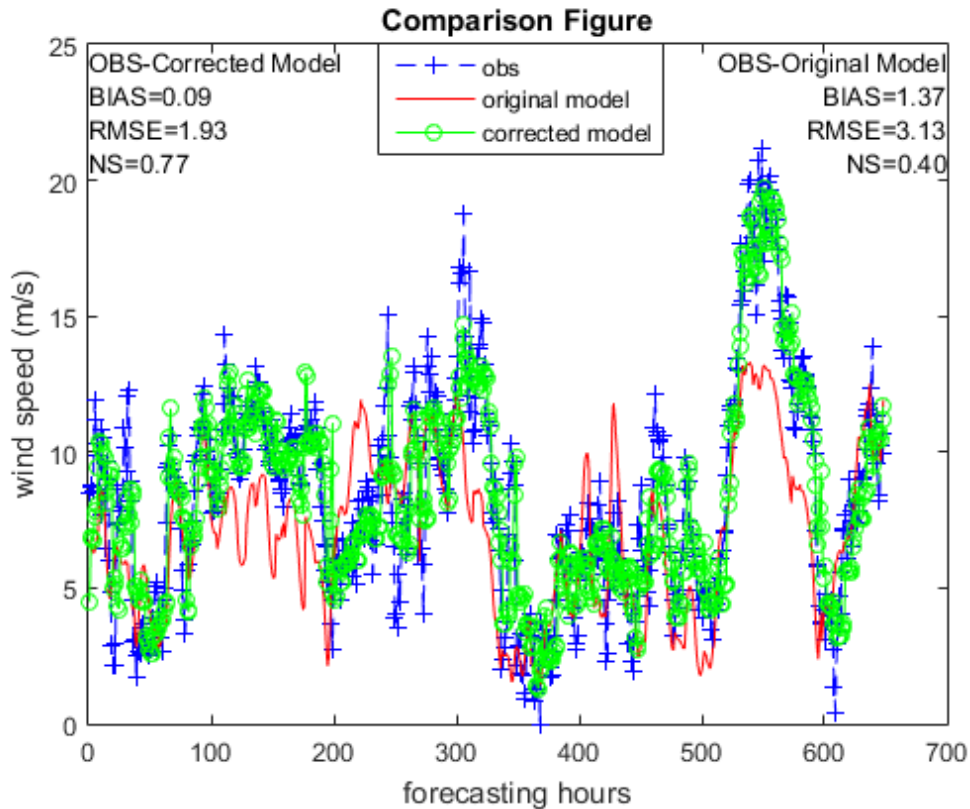
Γράφημα 4.9: Τα διορθωμένα αποτελέσματα για $f=12$ και $h=72$



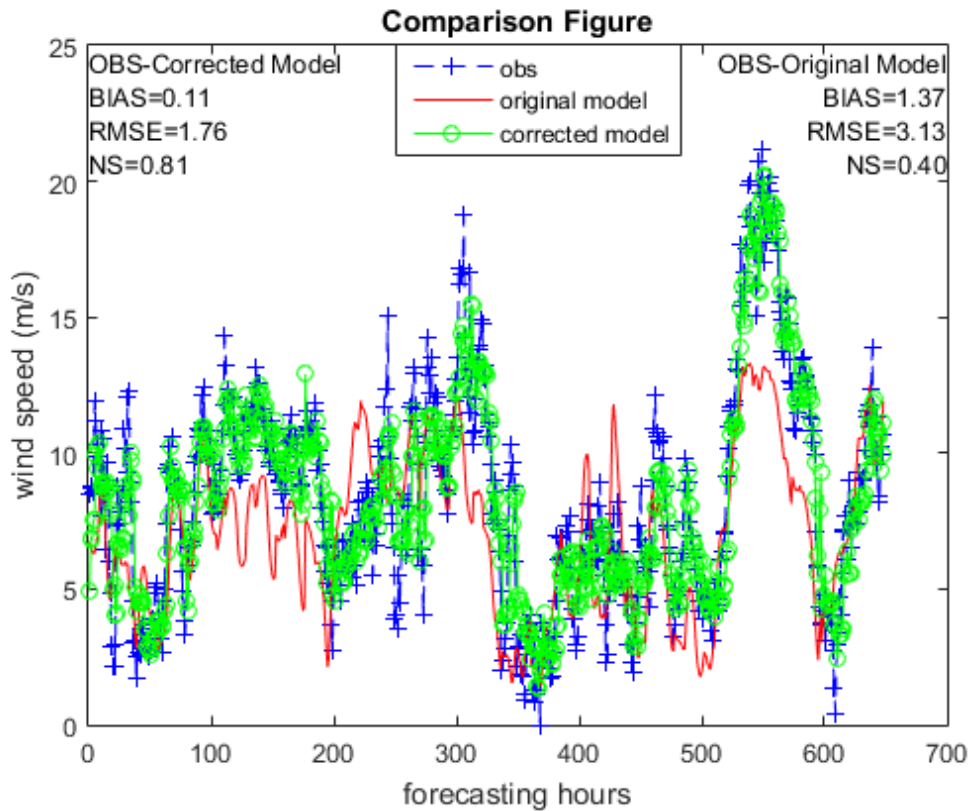
Σχήμα 4.10: Τα διορθωμένα αποτελέσματα για $f=24$ και $h=48$



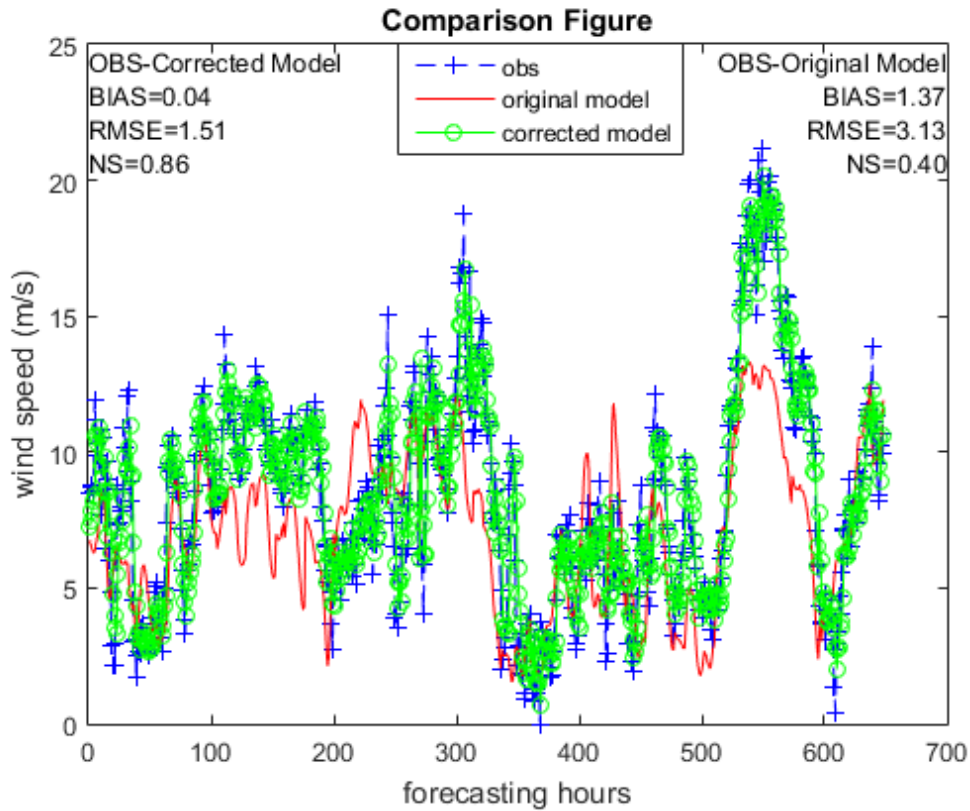
Σχήμα 4.11: Τα διορθωμένα αποτελέσματα για $f=6$ και $h=24$



Σχήμα 4.12: Τα διορθωμένα αποτελέσματα για $f=3$ και $h=12$



Σχήμα 4.13: Τα διορθωμένα αποτελέσματα για $f=2$ και $h=12$



Σχήμα 4.14: Τα διορθωμένα αποτελέσματα για $f=1$ και $h=6$

4.3.2 Πάρκο «Enercon», Κρύα Σητείας

- *Προγνωστικές περιόδους 3 ημερών περίπου*

Τα παρακάτω αποτελέσματα προήλθαν από σετ δεδομένων της περιόδου από 20/8/2014 έως 23/8/2014. Η κατανομή η οποία ταίριαζε καλύτερα στα συγκεκριμένα δεδομένα ήταν η Weibull (κριτήριο NlogL), το πολυώνυμο ήταν 2ου βαθμού, ενώ η προγνωστική περίοδος αφορά διάστημα 72 ωρών (3 ημερών). Ενώ για τη δεύτερη περίπτωση, το dataset αφορά την περίοδο 28/5/2015 έως 2/6/2015. Η κατανομή η οποία ταίριαζε καλύτερα στα δεδομένα ήταν η Weibull, το πολυώνυμο ήταν 2ου βαθμού, ενώ η προγνωστική περίοδος αφορά διάστημα 84 ωρών (3,5 ημερών).

Πίνακας 4.7 Προγνωστική περίοδος: 72hr, Dataset :96hr, Κατανομή: Weibull, Βαθμός Πολ/μου: 2^{ος}

BIAS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-3,04	-0,42	-0,1	-0,13	-0,04	-0,01	-0,07
h=12	-3,04	0,23	0,06	0,03	-0,07	0,04	-0,03
h=6	-3,04	0,75	0,12	0,11	-0,07	0,06	-0,01
h=3	-3,04	0,47	-0,23	0,08	-0,05	-0,13	-0,02

RMSE							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	3,81	2,26	2,78	2,42	2	1,69	1,62
h=12	3,81	2,05	2,24	1,84	2,13	1,8	1,71
h=6	3,81	2,2	1,9	1,93	1,9	1,86	1,74
h=3	3,81	2,49	2,16	2,04	1,85	2,05	1,72

NS							
history	forecast (hr)						
(hr)		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=24	-2,83	-0,34	-1,04	-0,55	-0,06	0,25	0,31
h=12	-2,83	-0,11	-0,32	0,11	-0,2	0,15	0,23
h=6	-2,83	-0,28	0,05	0,02	0,05	0,09	0,2
h=3	-2,83	-0,64	-0,24	-0,1	0,09	-0,11	0,22

Πίνακας 4.8 Προγνωστική περίοδος: 84hr, Dataset :132hr, Κατανομή: Weibull, Βαθμός Πολ/μου: 2^{ος}

BIAS							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=48	-2,44	-0,29	-0,15	0,03	0,02	0,22	0,22
h=24	-2,44	-0,06	-0,06	0,00	0,05	0,23	0,22
h=12	-2,44	-0,05	-0,16	-0,07	0,27	0,17	0,15
h=6	-2,44	0,49	0,32	-0,13	-0,09	0,07	0,06
h=3	-2,44	-0,59	0,33	0,07	-0,11	0,03	0,02

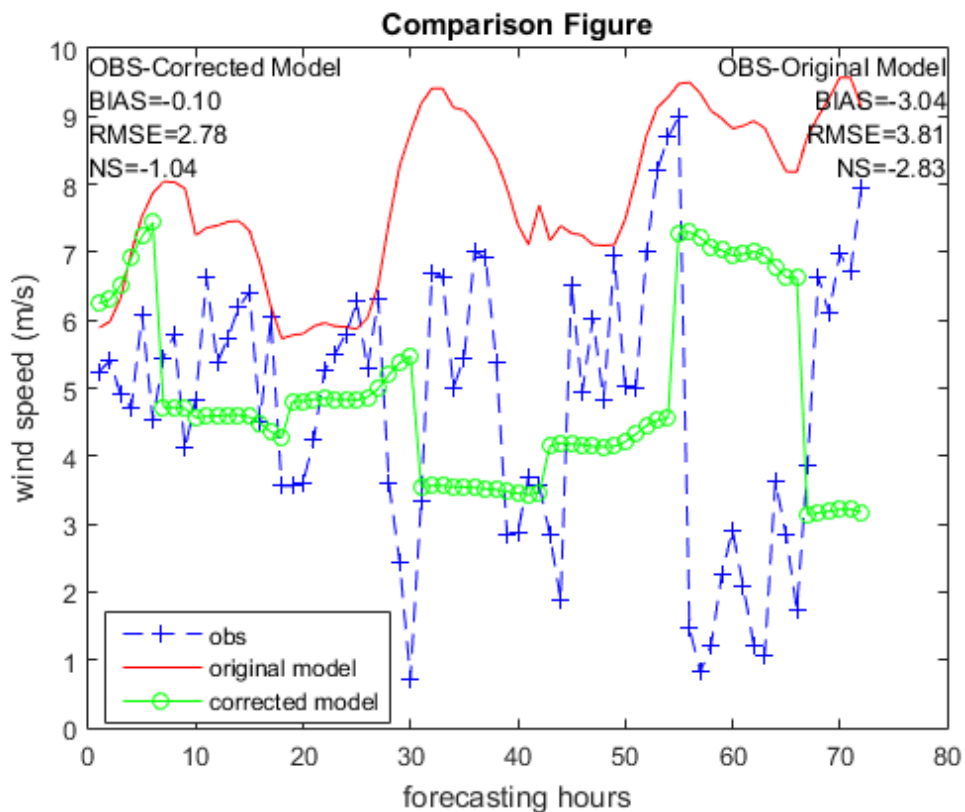
RMSE							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=48	3,10	2,29	1,93	1,69	1,72	1,73	1,69
h=24	3,10	2,15	1,82	1,60	1,65	1,63	1,57
h=12	3,10	1,98	1,70	1,55	1,57	1,59	1,53
h=6	3,10	1,70	1,50	1,58	1,67	1,61	1,54
h=3	3,10	1,86	2,34	1,93	1,76	1,66	1,57

NS							
<i>history</i>	<i>forecast (hr)</i>						
<i>(hr)</i>		f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected
h=48	-1,12	-0,16	0,18	0,37	0,35	0,34	0,37
h=24	-1,12	-0,02	0,27	0,43	0,40	0,42	0,46
h=12	-1,12	0,13	0,36	0,47	0,45	0,44	0,48
h=6	-1,12	0,36	0,50	0,45	0,38	0,45	0,48
h=3	-1,12	0,23	-0,21	0,17	0,32	0,39	0,45

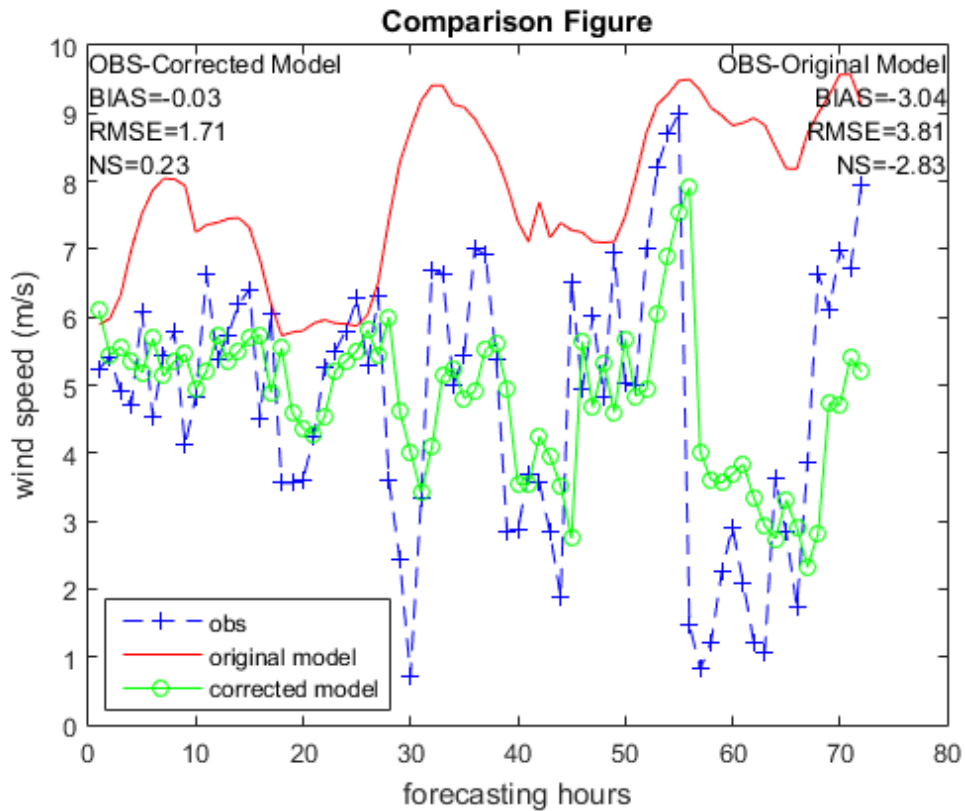
Στα παραπάνω αποτελέσματα (Πίνακες 4.7 - 4.8) φαίνεται ότι το bias έχει εξαλειφθεί σε όλες τις βέλτιστες περιπτώσεις, όπως και το RMSE σε όλες τις περιπτώσεις έχει μειωθεί πάνω από 45%. Πρόκειται για δύο περιπτώσεις όπου το μοντέλο (κόκκινη καμπύλη) έχει υπερεκτιμήσει τις τιμές του ανέμου. Όπως και στα αποτελέσματα της πρώτης περιοχής, των Χανίων, το φίλτρο έδωσε πολύ καλά αποτελέσματα για τα προγνωστικά βήματα των 6 ωρών και κάτω, κάτι το οποίο φαίνεται και από τον δείκτη NS, ο οποίος είναι αρκετά πάνω από το 0. Για το βήμα f=12, όπως και στα προηγούμενα αποτελέσματα, ενώ έχουμε ποσοτική μείωση στους δύο τύπους σφαλμάτων, αυτή είναι πλασματική. Το φίλτρο και πάλι δεν έχει λειτουργήσει και κάνουν την εμφάνιση τους

κάποια διακριτά, σταθερά διαστήματα (Σχήμα 4.15), τα οποία χαρακτηρίζονται από την ίδια τιμή. Η τιμή αυτή έχει προέλθει από τη διόρθωση του Bayesian τμήματος του φίλτρου, όπως αναφέρθηκε και προηγουμένως. Όσον αφορά το βήμα $f=24$, και πάλι το φίλτρο δεν έχει λειτουργήσει και οι διορθώσεις που φαίνονται στους Πίνακες 4.7 και 4.8 είναι φαινομενικές. Επιπλέον σε αυτήν την περίπτωση παρότι υπήρχε περίοδος εκπαίδευσης διπλάσια ($h=48$), το φίλτρο κατάφερε να δώσει μία εικόνα για τη μέση κατάσταση, χωρίς να πιάνει όμως καμία διακύμανση στις τιμές του ανέμου (Παράρτημα Β, Σχήμα).

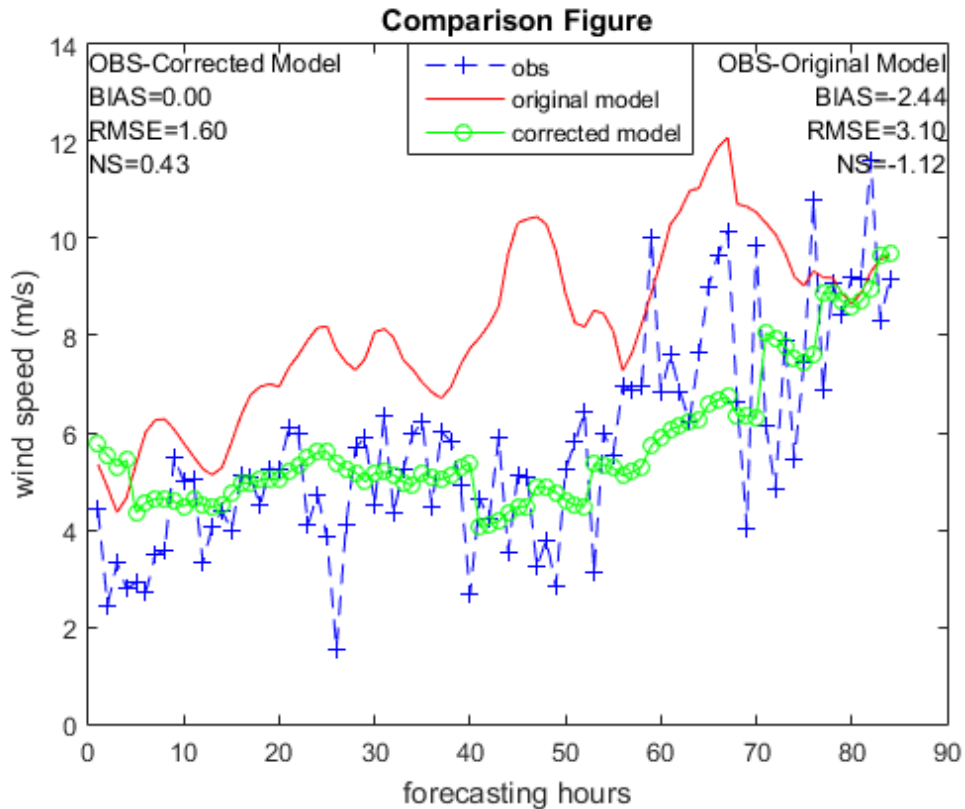
Τα αποτελέσματα που φαίνονται στους παραπάνω Πίνακες (4.7 – 4.8) δείχνουν ότι το φίλτρο λειτούργησε με τον ίδιο τρόπο και για την περιοχή της Σητείας, επομένως η αλλαγή της τοποθεσίας δεν επέφερε αλλαγές στη λειτουργία του φίλτρου.



Σχήμα 4.15: Τα διορθωμένα αποτελέσματα για $f=12$ και $h=24$



Σχήμα 4.16: Τα διορθωμένα αποτελέσματα για $f=1$ και $h=12$



Σχήμα 4.17: Τα διορθωμένα αποτελέσματα για $f=6$ και $h=24$

- *Προγνωστική περίοδος 1 εβδομάδας περίπου*

Τα παρακάτω αποτελέσματα προήλθαν από σει δεδομένων της περιόδου από 24/2/2015 έως 4/3/2015. Η κατανομή η οποία ταίριαζε καλύτερα στα συγκεκριμένα δεδομένα ήταν η Weibull, το πολώνυμο ήταν 2ου βαθμού, ενώ η προγνωστική περίοδος αφορά διάστημα 208 ωρών (8,5 ημερών).

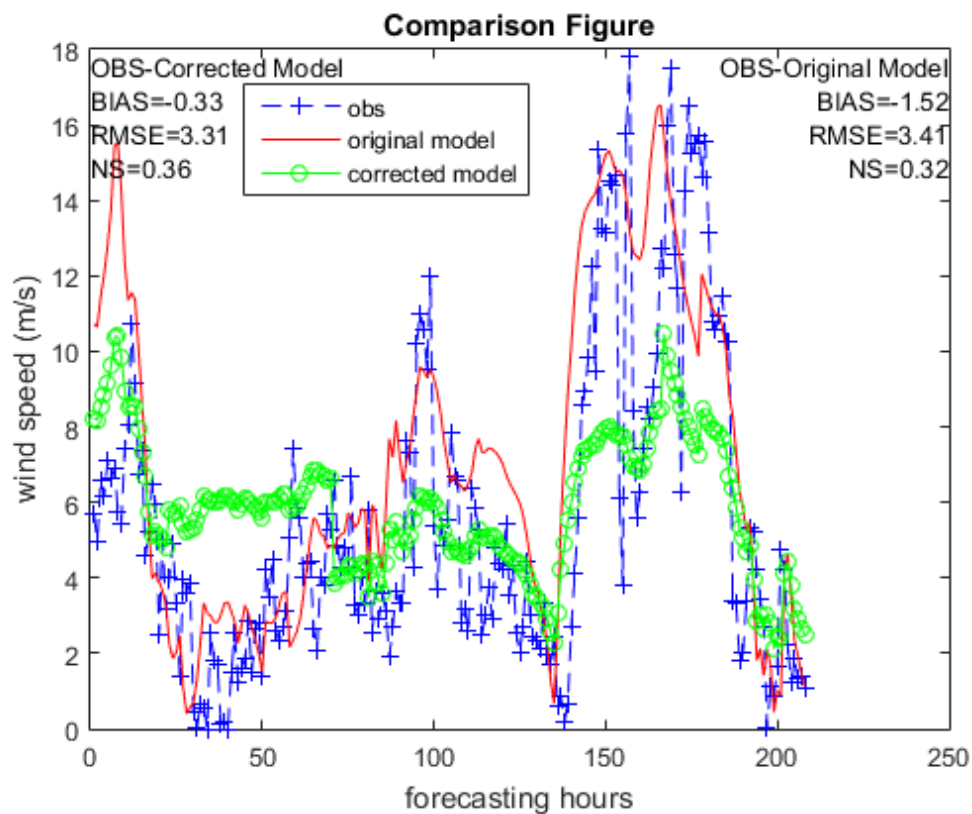
Πίνακας 4.9 Προγνωστική περίοδος: 208hr, Dataset :96hr, Κατανομή: Weibull, Βαθμός Πολ/μου: 2^{ος}

BIAS								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	-1,52	1,09	0,42	0,14	0,07	-0,14	-0,14	-0,24
h=48	-1,52	-0,33	0,18	-0,07	-0,09	-0,29	-0,26	-0,35
h=24	-1,52	1,33	0,41	0,20	0,07	-0,16	-0,13	-0,21
h=12	-1,52	0,19	-0,60	0,29	0,06	-0,10	-0,08	-0,14
h=6	-1,52	-0,81	-1,06	0,02	0,15	-0,02	-0,05	-0,09
h=3	-1,52	-1,57	-1,40	-1,19	-0,24	0,03	-0,16	-0,03

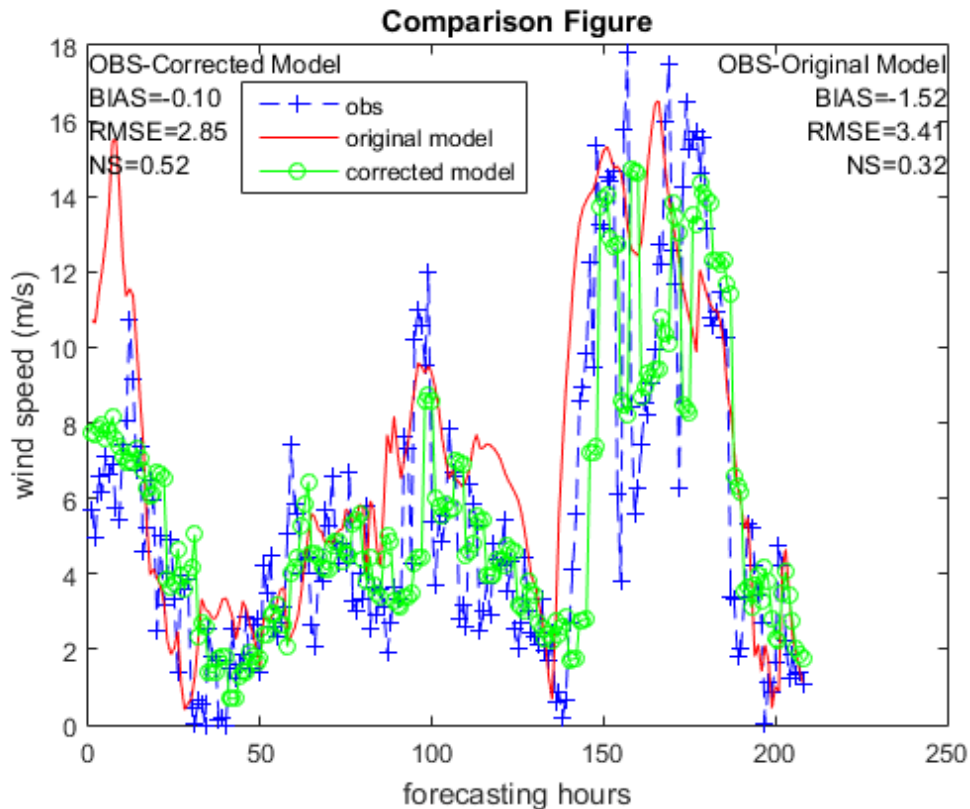
RMSE								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	3,41	3,98	3,65	3,53	3,43	3,13	2,71	2,65
h=48	3,41	3,31	3,66	3,58	3,44	3,18	2,77	2,71
h=24	3,41	4,77	3,84	3,56	3,24	3,01	2,60	2,53
h=12	3,41	4,31	3,87	3,67	3,24	2,85	2,48	2,33
h=6	3,41	4,04	3,31	3,52	2,94	2,70	2,36	2,21
h=3	3,41	4,14	3,33	3,14	3,16	2,79	2,72	2,24

NS								
history (hr)	forecast (hr)							
		f=48	f=24	f=12	f=6	f=3	f=2	f=1
	original	corrected	corrected	corrected	corrected	corrected	corrected	corrected
h=72	0,32	0,07	0,22	0,27	0,31	0,43	0,57	0,59
h=48	0,32	0,36	0,21	0,25	0,31	0,41	0,55	0,57
h=24	0,32	-0,33	0,13	0,26	0,38	0,47	0,60	0,62
h=12	0,32	-0,09	0,12	0,21	0,38	0,52	0,64	0,68
h=6	0,32	0,05	0,36	0,27	0,49	0,57	0,67	0,71
h=3	0,32	-0,01	0,35	0,42	0,41	0,54	0,57	0,71

Όπως και στις προηγούμενες περιπτώσεις που εξετάστηκαν, τα μικρά προγνωστικά βήματα είναι εξαιρετικά. Το bias έχει εξαλειφθεί σχεδόν (πάνω από 90% μείωση σε όλες τις βέλτιστες περιπτώσεις) και το RMSE μειώθηκε πάνω από 20% σε όλες τις βέλτιστες περιπτώσεις. Μία παρατήρηση άξια αναφοράς είναι ότι στα μικρά αυτά βήματα η υπερβολική αύξηση της περιόδου «εκπαίδευσης» δεν φέρνει καλύτερα αποτελέσματα, αλλά αντιθέτως στα $f=6$ τα κατάλληλα «history» είναι τα $h=6$, $h=12$ και $h=24$ (Πίνακας 4.9). Για το βήμα $f=48$, όπως φαίνεται στον Πίνακα 4.9, το RMSE μόνο σε μία περίπτωση εμφανίζει μικρότερη τιμή από το μοντέλο ($h=48$) και σε αυτήν την περίπτωση το φίλτρο ουσιαστικά δεν έχει λειτουργήσει (Σχήμα 4.18), δεν έχει ακολουθήσει καμία διακύμανση και απλώς έχει εκτιμήσει μία μέση κατάσταση για όλο το διάστημα. Ομοίως συνέβη και για τα βήματα $f=24$ και $f=12$.



Σχήμα 4.18: Τα διορθωμένα αποτελέσματα για $f=48$ και $h=48$



Σχήμα 4.19: Τα διορθωμένα αποτελέσματα για $f=3$ και $h=12$

4.3.3 Βέλτιστες τιμές

Ο πίνακας 4.10 δείχνει τα συγκεντρωτικά αποτελέσματα για τις βέλτιστες περιπτώσεις στο σύνολο των δοκιμών για την περιοχή των Χανίων. Όπως αναφέρθηκε και πιο πάνω, τα προγνωστικά βήματα άνω των 6 ωρών δεν είναι αξιόπιστα όσον αφορά τα αποτελέσματά τους, για τα μικρά σετ δεδομένων τα οποία χρησιμοποιήθηκαν (στις περιπτώσεις με μεγαλύτερα σετ λειτούργησαν και τα βήματα $f=24$ και $f=12$, σχήματα 4.9-4.10). Οι τιμές των βημάτων «ιστορικότητας», για τα βήματα $f=48$, $f=24$ και $f=12$ είναι πλασματικές, αφού το φίλτρο δεν λειτούργησε και επομένως δεν μπορούν να αξιοποιηθούν. Στα βήματα από τις 6 ώρες και κάτω, όπως φάνηκε και πιο πάνω, οι περίοδοι εκπαίδευσης των 12, των 6 και των 3 ωρών λειτουργούν εξίσου καλά. Αξίζει να αναφερθεί ότι αν αυξηθούν περαιτέρω οι περίοδοι «εκπαίδευσης» για αυτά τα προγνωστικά βήματα, αυτό θα λειτουργήσει ανασταλτικά για τα αποτελέσματα. Για $f=2$ και $f=1$ οι 6 ώρες «εκπαίδευσης» αρκούν.

Πίνακας 4.10 Συγκεντρωτικά αποτελέσματα για τις βέλτιστες τιμές των βημάτων πρόγνωσης και των περιόδων «εκπαίδευσης» του αλγορίθμου σε κάθε περίπτωση.

BIAS		RMSE		ΣΥΝΔΥΑΣΤΙΚΑ	
<i>forecast</i>	<i>history</i>	<i>forecast</i>	<i>history</i>	<i>forecast</i>	<i>history</i>
48	48	48	48	48	-
24	3	24	3	24	-
12	24,12	12	6,3	12	-
6	12	6	12,6,3	6	12,6
3	12,6,3	3	12,6,3	3	6,3
2	6,3	2	6,3	2	6,3
1	6,3	1	6,3	1	6,3

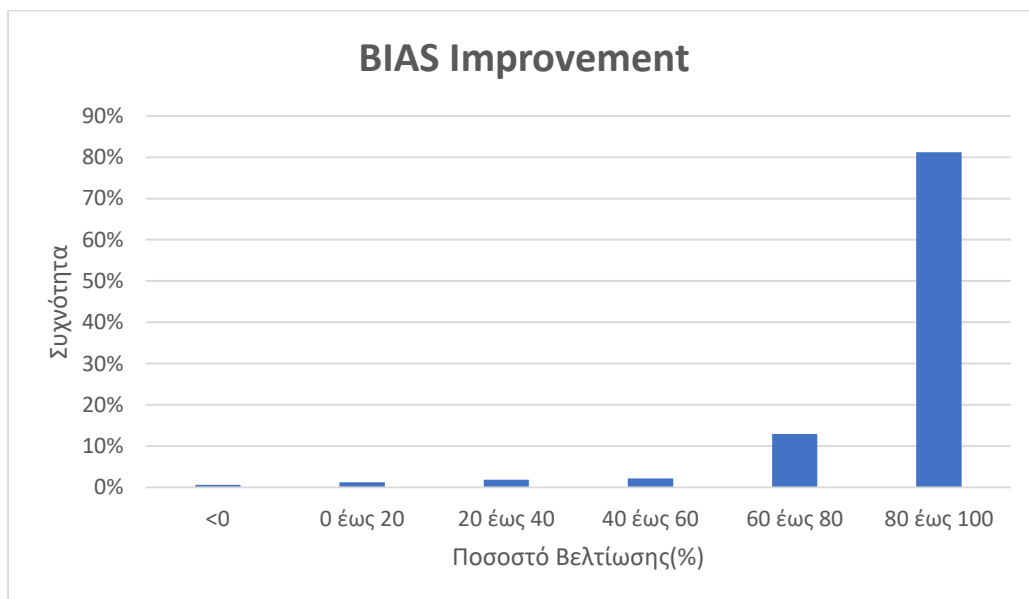
4.4. Συγκεντρωτικά αποτελέσματα για τα ωριαία δεδομένα

Τα σχήματα 4.20 και 4.21 απεικονίζουν τα ιστογράμματα των ποσοστών βελτίωσης για τα σφάλματα bias και RMSE, και αφορούν τα σύνολο των δοκιμών που εκτελέστηκαν για τα Χανιά. Όπως φαίνεται από το σχήμα 4.20, στο 81% των συνολικών δοκιμών το bias μειώθηκε κατά 80% έως 100% , δηλαδή σχεδόν εξαλείφθηκε. Στο 13% των περιπτώσεων το bias μειώθηκε από 60% έως 80%, και στο υπόλοιπο 7% ανήκουν όλα τα υπόλοιπα ποσοστά. Τα ποσοστά αυτά αποδεικνύουν ότι το η εφαρμογή του φίλτρου λειτουργεί με επιτυχία στην εξάλειψη της μέσης απόκλισης. Ωστόσο, επειδή η μέση απόκλιση εκφράζει τη διαφορά μεταξύ της μέσης τιμής της πρόγνωσης και της μέσης τιμής των παρατηρήσεων, δεν είναι απαραίτητο ότι η μείωση του συνεπάγεται και την καλή λειτουργία του φίλτρου. Για τον λόγο αυτό εξετάστηκε και ο τύπος του σφάλματος RMSE, το οποίο βοηθά στην εκτίμηση της ακρίβειας της πρόγνωσης. Στο σχήμα 4.21 φαίνεται ότι στο 45% των συνολικών δοκιμών που εκτελέστηκαν, η μείωση του σφάλματος RMSE κυμαίνεται μεταξύ 40% και 60%, και στο 19% των συνολικών περιπτώσεων από 60% έως 80%. Επίσης, το ποσοστό όπου το σφάλμα RMSE είτε αυξήθηκε είτε δεν κατάφερε να μειωθεί ήταν μόλις το 10% των συνολικών περιπτώσεων.

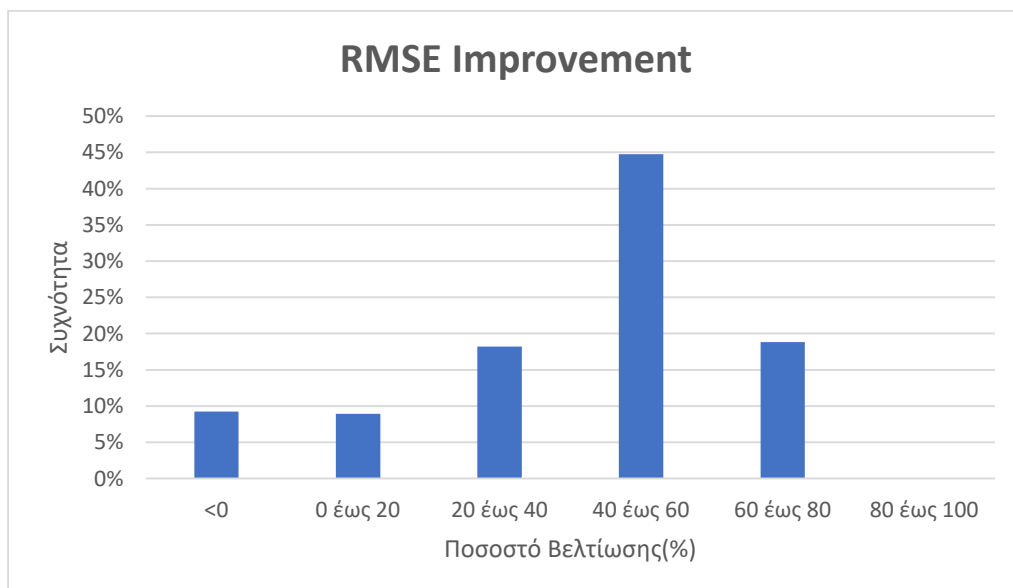
Τέλος, στο σχήμα 4.22 φαίνεται το ιστόγραμμα των ποσοστών βελτίωσης του δείκτη Nash-Sutcliffe, ο οποίος χρησιμοποιείται για την αξιολόγηση του μοντέλου. Όπως φαίνεται στο σχήμα 4.22, το μεγαλύτερο ποσοστό βελτίωσης σημειώθηκε για τα ποσοστά βελτίωσης 60% - 80% και αφορά το 36% των συνολικών αποτελεσμάτων. στο 31% των συνολικών περιπτώσεων επιτεύχθηκε βελτίωση του δείκτη NS από 80%

έως 100% (συγκεκριμένα το μεγαλύτερο ποσοστό ήταν 96% και αφορούσε τα προγνωστικά βήματα $f=1$ και $f=2$). Σε ποσοστό μόλις 10% ανήλθαν οι περιπτώσεις όπου το φίλτρο είχε αρνητικές επιδόσεις.

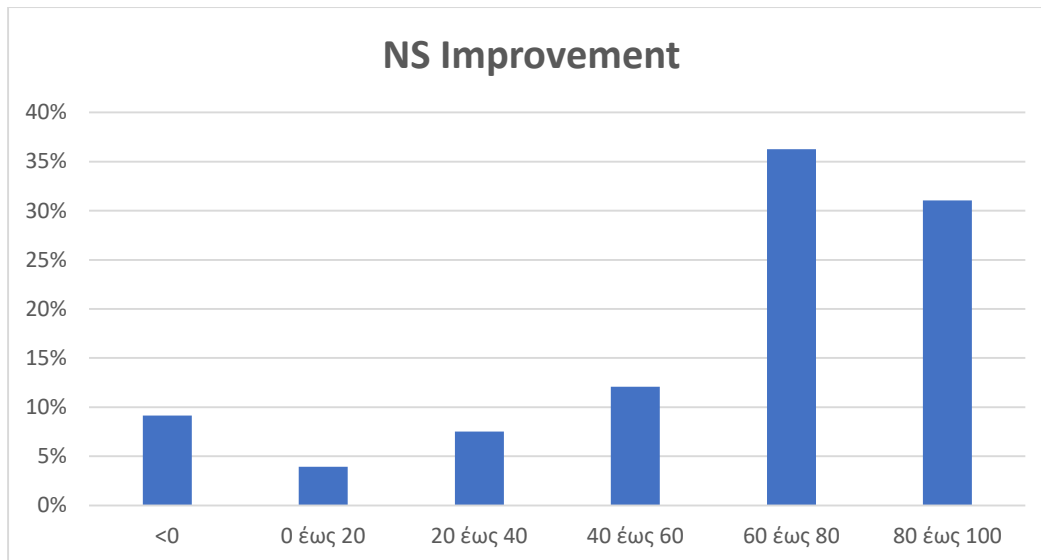
Συνολικά, σύμφωνα με το σύνολο των δοκιμών που εκτελέστηκαν, η εφαρμογή του φίλτρου φαίνεται ότι επιτυγχάνει σχεδόν να εξαλείψει το bias (σχεδόν 90% η μέση μείωση) και να μειώσει το RMSE κατά 40% και πλέον, κατά μέσο όρο. Στα αποτελέσματα αυτά θα πρέπει συνυπολογιστεί το γεγονός ότι επιτεύχθηκαν χρησιμοποιώντας πολύ μικρή υπολογιστική μνήμη.



Σχήμα 4.20: Ιστόγραμμα των ποσοστών βελτίωσης (%) για το bias, Χανιά



Σχήμα 4.21: Ιστόγραμμα των ποσοστών βελτίωσης (%) για το RMSE, Χανιά



Σχήμα 4.22: Ιστόγραμμα των ποσοστών βελτίωσης (%) για τον δείκτη NS, Χανιά

4.5. Εξέταση της απόδοσης του φίλτρου με πυκνότερο σετ δεδομένων

Προκειμένου να εξετασθεί η περαιτέρω βελτίωση στις διορθώσεις του φίλτρου και συνεπώς η αποδοτικότητά του, έγιναν δοκιμές που αφορούν τη χρονική περίοδο 27/11/2014 και 30/10/2014 έως 1/11/2014, αλλά αυτή τη φορά τα σετ δεδομένων ήταν χωρισμένα *ωριαία*, *ανά μισή ώρα*, *ανά 10 λεπτά*, *ανά 5 λεπτά*, *ανά 2 λεπτά* και *ανά 1 λεπτό*. Τα δεδομένα αυτά προήλθαν από το πάρκο «Αλκυόνη», της περιοχής του Κισιάμου Χανίων.

Όπως φαίνεται στους πίνακες 4.11 – 4.13 αλλά και στα σχήματα 4.24 – 4.26, η αύξηση της πυκνότητας των δεδομένων επιφέρει περαιτέρω μείωση στα σφάλματα. Αυτό συμβαίνει διότι με το πυκνότερο σετ δεδομένων, δηλαδή με μεγαλύτερο πλήθος δεδομένων για το ίδιο χρονικό διάστημα, το φίλτρο καταφέρνει να παίρνει με μεγαλύτερη ακρίβεια τους συντελεστές από την κατανομή που ταιριάζει καλύτερα στην κάθε περίπτωση, και τους οποίους στη συνέχεια χρησιμοποιεί στους υπολογισμούς για τη διόρθωση των τιμών του ανέμου.

Μεταξύ των διαφορετικών δειγμάτων δεδομένων παρατηρεί κανείς ότι μεταξύ τους οι διαφορές είναι μικρές. Παρόλα αυτά ξεχωρίζουν τα δείγματα των 10 λεπτών και των 5 λεπτών ως λίγο πιο αποδοτικά. Χαρακτηριστική είναι η περίπτωση που φαίνεται στον πίνακα 4.12, όπου το bias από την τιμή 3,95 μειώνεται σε -0,85, και το RMSE από την τιμή 4,52 μειώνεται σε 2,70, στο ωριαίο σετ δεδομένων και μετά την χρήση του

δεκαλέπτου σετ δεδομένων, το bias από 3,92 πέφτει στην τιμή -0,06, ενώ το RMSE από 4,50 σε 2,02. Επίσης, αξίζει να αναφερθεί ότι με τα πιο πυκνά σετ δεδομένων το φίλτρο καταφέρνει να αξιόπιστες διορθώσεις και στις τιμές των μεγαλύτερων βημάτων, όπως του $f=24$.

Σύμφωνα με τα σχήματα 4.23 και 4.24, όπου δείχνουν την ποσοστιαία βελτίωση (%) για το κάθε είδος δείγματος, για το σύνολο των δοκιμών που έγιναν, το πιο αποδοτικό δείγμα φαίνεται να είναι αυτό των 10 λεπτών. Επιφέρει μία επιπλέον μείωση κατά 13% κατά μέσο όρο, σε σχέση με τα ωριαία σετ δεδομένων, στο ήδη μειωμένο bias. Αντίστοιχα για τα αποτελέσματα του RMSE, για το δείγμα των 10 λεπτών, η επιπλέον βελτίωση είναι 16% κατά μέσο όρο. Αξίζει να αναφερθεί ότι το σετ δεδομένων των 30 λεπτών δεν επέφερε αλλαγές σε σχέση με το σετ των ωριαίων δεδομένων.

Γενικά, οι διαφορές μεταξύ των δειγμάτων είναι πολύ μικρές, όμως μπορεί να εξαχθεί το συμπέρασμα ότι το καλύτερο δείγμα είναι οριακά αυτό των 10 λεπτών. Τόσο λόγω της οριακής του υπεροχής στα αποτελέσματα έναντι των άλλων δειγμάτων, όσο και γιατί το δείγμα των 10 λεπτών δεν θα αυξήσει δραματικά το πλήθος των δεδομένων. Επίσης, όποιο δείγμα δεδομένων και να προτιμηθεί, το bias ελαττώνεται κατά περίπου 10% επιπλέον, ενώ το RMSE κατά 15% περίπου. Σε αυτά τα αποτελέσματα είναι καλό να αναλογιστεί κανείς ότι επρόκειτο για αποτελέσματα ενός μοντέλου υψηλής ανάλυσης (παράγραφος 3.2), καλώς ορισμένου και με μικρά σφάλματα.

Πίνακας 4.11 Συγκεντρωτικά αποτελέσματα για bias και RMSE, για διαφορετικά μεγέθη δειγματοληψίας. Στην τελευταία στήλη η βελτίωση σε σχέση με τα ωριαία δεδομένα. Με κίτρινο οι βέλτιστες περιπτώσεις. Η προγνωστική περίοδος ήταν 24 ώρες.

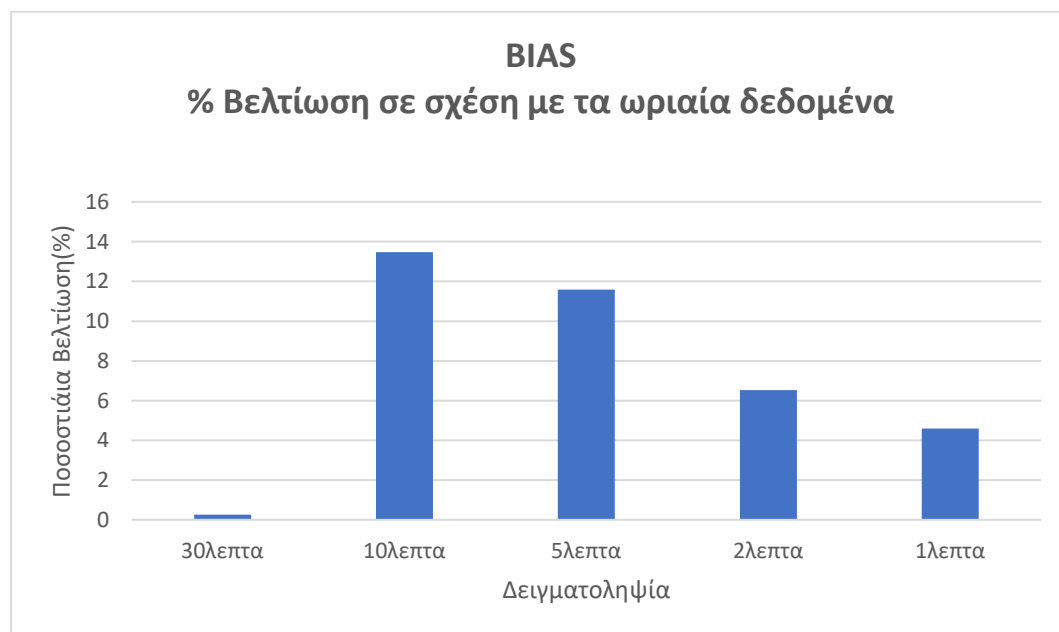
	bias			RMSE			βελτίωση σε σχέση με 1h%	
	original	corrected	% βελτίωση	original	corrected	% βελτίωση	BIAS	RMSE
ωριαία	2,04	0,64	68,63	2,95	2,75	6,78		
10λεπτα	2,26	-0,55	75,66	3,02	2,29	24,17	7,04	17,39
5λεπτα	2,26	-0,45	80,09	3,01	2,55	15,28	11,46	8,50
2λεπτα	2,26	-0,67	70,35	2,99	2,32	22,41	1,73	15,63
1λεπτα	2,26	-0,66	70,80	2,99	2,31	22,74	2,17	15,96

Πίνακας 4.12 Συγκεντρωτικά αποτελέσματα για bias και RMSE, για διαφορετικά μεγέθη δειγματοληψίας. Στην τελευταία στήλη η βελτίωση σε σχέση με τα ωριαία δεδομένα. Με κίτρινο οι βέλτιστες περιπτώσεις. Η προγνωστική περίοδος ήταν 72 ώρες.

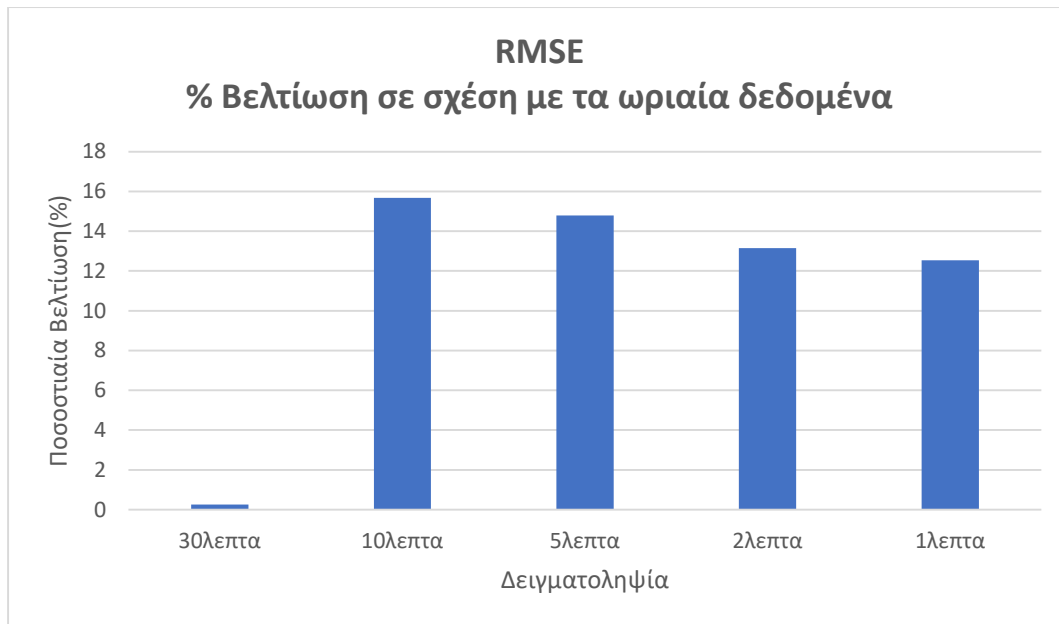
48h-24f	BIAS			RMSE			βελτιωση σε σχεση με 1h%	
	original	corrected	βελτιωση%	original	corrected	βελτιωση%	BIAS	RMSE
ωριαια	3,95	-0,58	85,32	4,52	2,70	40,27		
μισαωρα	3,88	-0,58	85,05	4,46	3,01	32,51	-0,26	-7,75
10λεπτα	3,92	-0,06	98,47	4,50	2,02	55,11	13,15	14,85
5λεπτα	3,92	0,50	87,24	4,50	1,93	57,11	1,93	16,85
2λεπτα	3,92	0,71	81,89	4,49	2,16	51,89	-3,43	11,63
1λεπτα	3,92	0,93	76,28	4,50	2,20	51,11	-9,04	10,85

Πίνακας 4.13 Συγκεντρωτικά αποτελέσματα για bias και RMSE, για διαφορετικά μεγέθη δειγματοληψίας. Στην τελευταία στήλη η βελτίωση σε σχέση με τα ωριαία δεδομένα. Με κίτρινο οι βέλτιστες περιπτώσεις. Η προγνωστική περίοδος ήταν 24 ώρες.

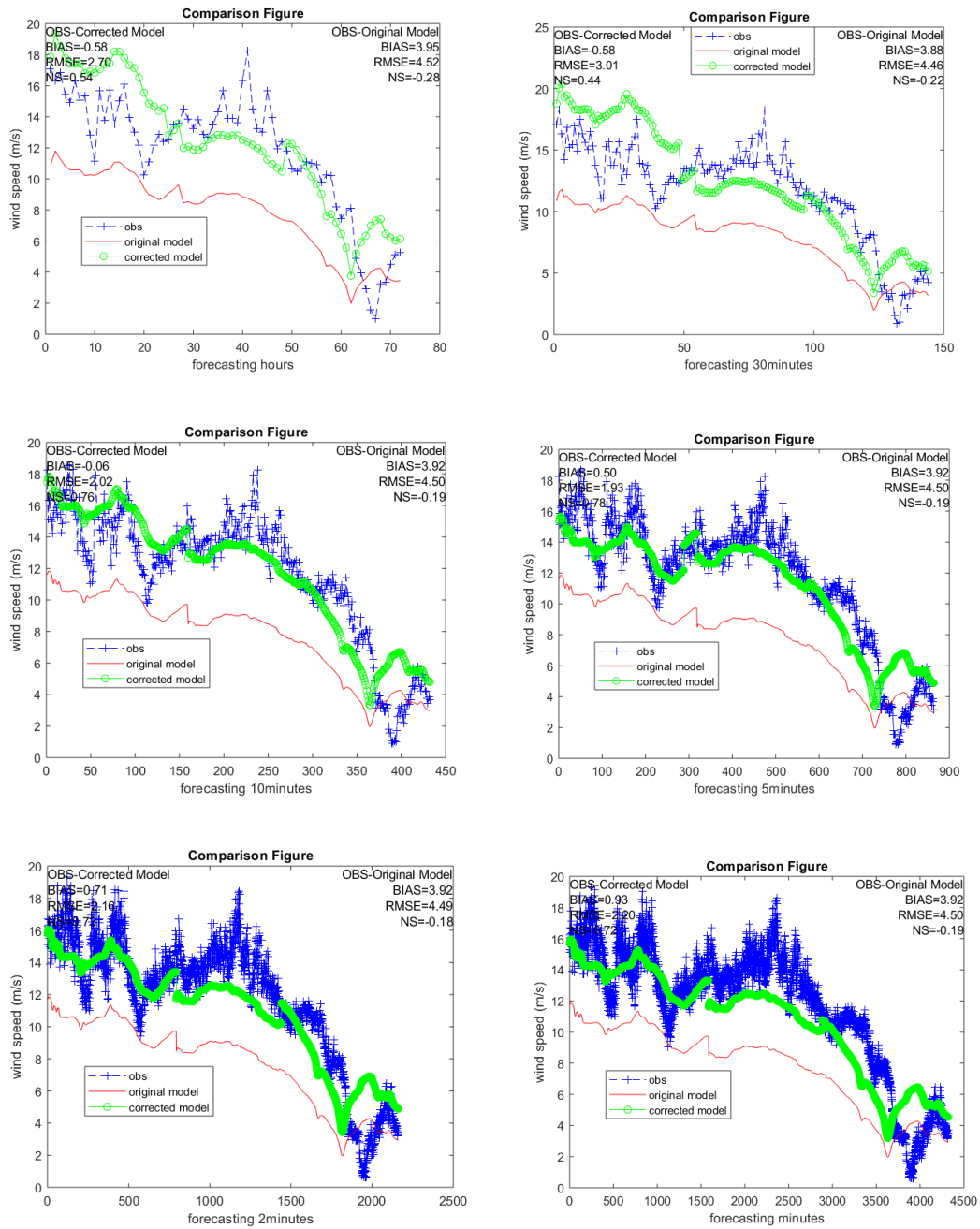
6h-3f	BIAS			RMSE			βελτιωση σε σχεση με 1h%	
	original	corrected	βελτιωση%	original	corrected	βελτιωση%	BIAS	RMSE
ωριαια	2,52	-0,20	92,06	3,06	2,03	33,66		
μισαωρα	2,59	-0,13	94,98	3,08	1,75	43,18	2,92	9,52
10λεπτα	2,65	-0,04	98,49	3,16	1,43	54,75	6,43	21,09
5λεπτα	2,68	-0,08	97,01	3,18	1,54	51,57	4,95	17,91
2λεπτα	2,80	0,04	98,57	3,33	1,82	45,35	6,51	11,69
1λεπτα	2,80	0,27	90,36	3,33	2,00	39,94	-1,71	6,28



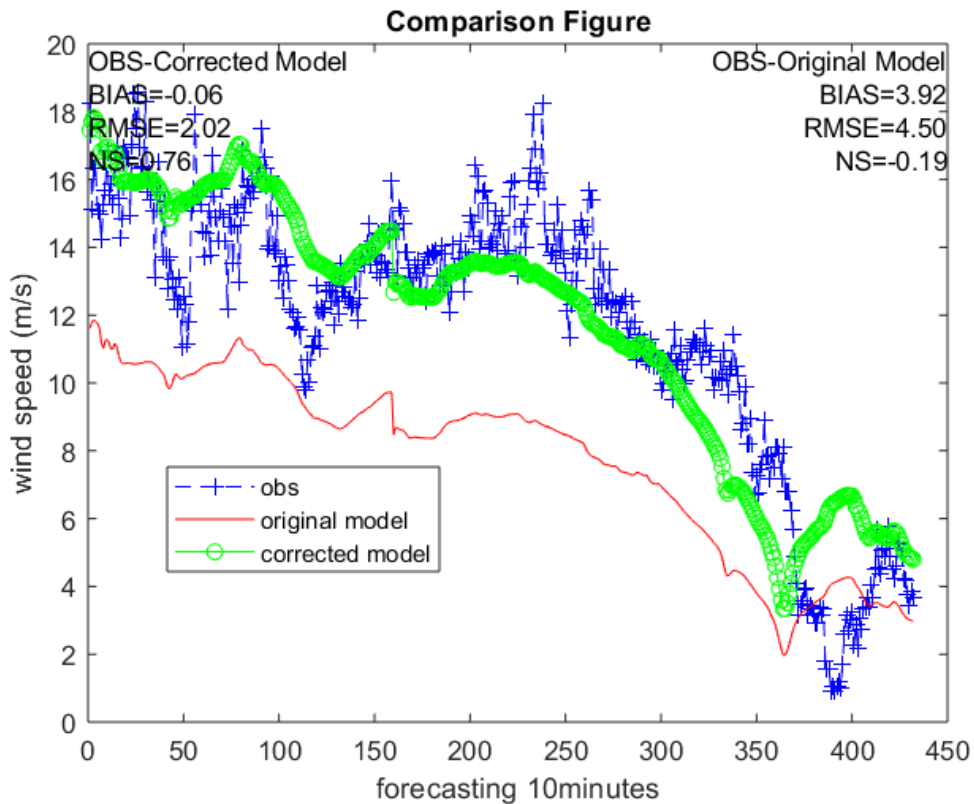
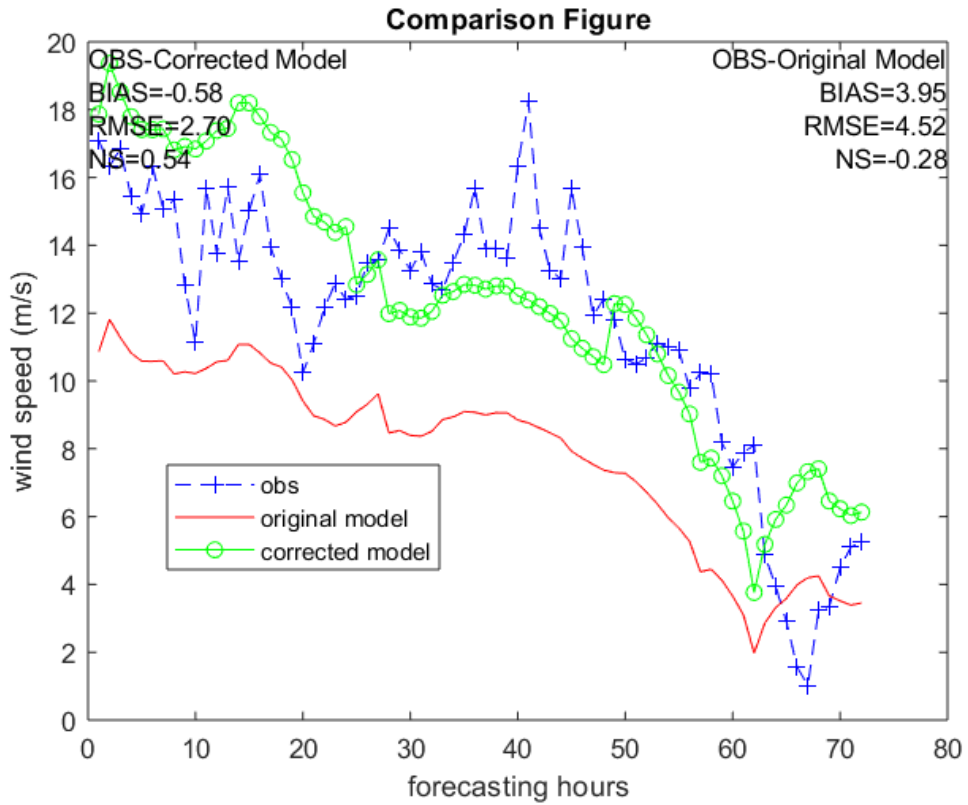
Σχήμα 4.23: Ποσοστιαία βελτίωση (%) του bias για τα διάφορα δείγματα δεδομένων σε σχέση με τα ωριαία



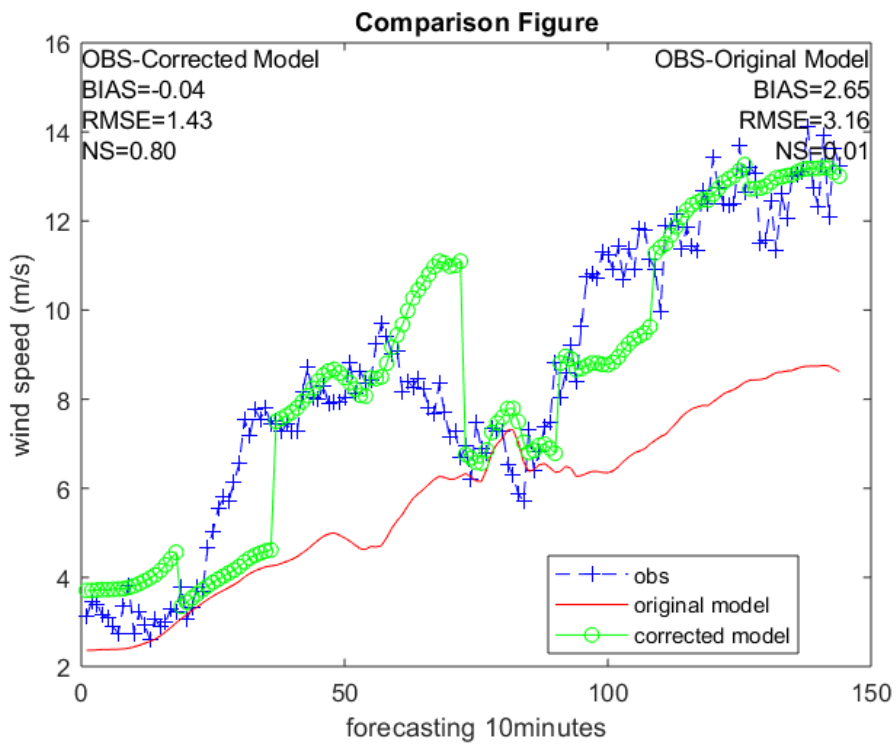
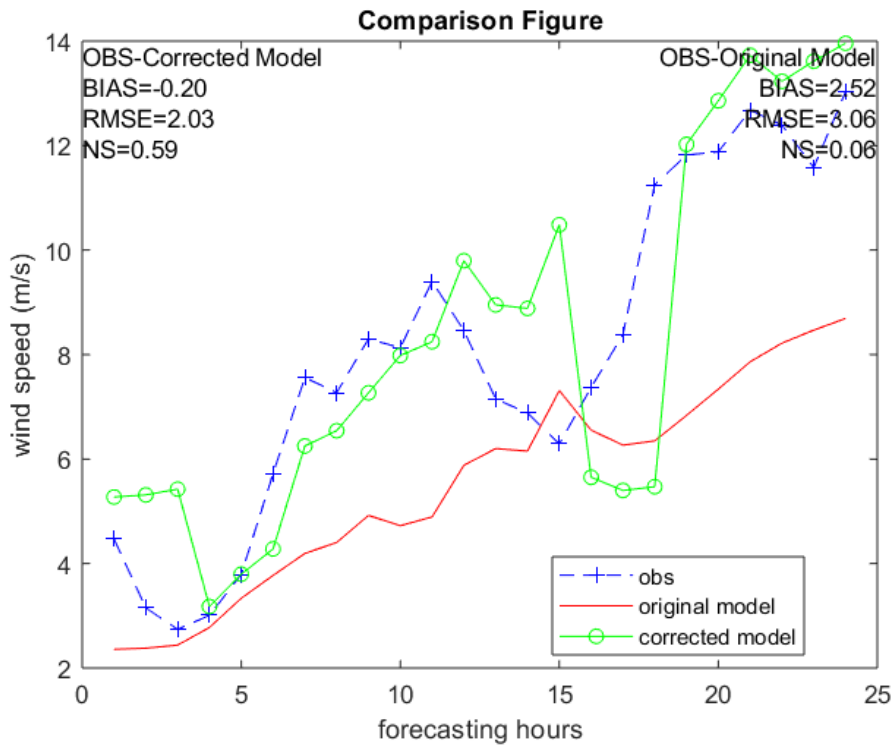
Σχήμα 4.23: Ποσοστιαία βελτίωση (%) του *RMSE* για τα διάφορα δείγματα δεδομένων σε σχέση με τα ωριαία



Σχήμα 4.24: Σύγκριση των αποτελεσμάτων για τα διαφορετικά σεί δεδομένων. Πάνω αριστερά τα ωριαία αποτελέσματα, δίπλα τα 30λεπτα, στη δεύτερη σειρά αριστερά τα 10λεπτα, δίπλα τα 5λεπτα και στην τελευταία σειρά από αριστερά τα 2λεπτα και δίπλα τα μονόλεπτα.



Σχήμα 4.25: Σύγκριση των αποτελεσμάτων για τα διαφορετικά σετ δεδομένων. Πάνω τα αποτελέσματα των ωριαίων δεδομένων και κάτω των δεκαλέπτων.



Σχήμα 4.26: Σύγκριση των αποτελεσμάτων για τα διαφορετικά σετ δεδομένων. Πάνω τα αποτελέσματα των ωριαίων δεδομένων και κάτω των δεκαλέπτων.

ΚΕΦΑΛΑΙΟ 5ο

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στο πλαίσιο της παρούσας εργασίας εξετάστηκαν τα συνδυασμένα Bayesian Kalman φίλτρα ως μέθοδος μετεπεξεργασίας των προγνωστικών αποτελεσμάτων μοντέλου υψηλής ανάλυσης, για την μετεωρολογική παράμετρο του ανέμου. Συγκεκριμένα, μελετήθηκαν οι παράμετροι των Bayesian Kalman φίλτρων, προκειμένου να βρεθούν οι συνθήκες κάτω από τις οποίες μπορούν να λειτουργήσουν με τη μέγιστη αποδοτικότητα.

Τα συμπεράσματα που προέκυψαν από τη συγκεκριμένη μελέτη συνοψίζονται στα εξής:

- Η κατανομή Weibull ταιριάζει καλύτερα στις περισσότερες περιπτώσεις των σετ δεδομένων που χρησιμοποιήθηκαν, όχι όμως σε όλες. Η extreme value κατανομή σε ορισμένες περιπτώσεις λειτουργεί καλύτερα από την Weibull. Συγκεκριμένα, στα μικρά σετ δεδομένων και στα μικρά προγνωστικά βήματα το φίλτρο μέσω της «extreme value» κατανομής καταφέρνει να περιγράψει τις διακυμάνσεις στην ταχύτητα του ανέμου αλλά και τη φάση.
- Το bias πρακτικά σχεδόν εξαλείφθηκε και το σφάλμα RMSE κατάφερε να μειωθεί κατά 40% - 60% κατά μέσο όρο στο σύνολο των δοκιμών που εκτελέστηκαν.
- Τα φίλτρα καταφέρνουν πολύ καλά αποτελέσματα στους πολύ βραχυπρόθεσμους χρονικούς ορίζοντες (very short-term forecasts), αξιοποιώντας την καλή προσαρμοστικότητα τους στα δεδομένα. Συγκεκριμένα στις προγνώσεις κάτω των 6 ωρών.
- Υπάρχουν όμως και περιπτώσεις όπου τα συνδυασμένα φίλτρα παρουσιάζουν προβλήματα: αυτά εστιάζονται στις περιπτώσεις όπου στη διορθωμένη χρονοσειρά εμφανίζονται κάποιες απότομες μεταβάσεις μεταξύ των προγνωστικών βημάτων. Αυτά προέρχονται από το κομμάτι της Bayesian διόρθωσης και αποτελούν μια εξομάλυνση των τιμών μέσω μίας μέσης τιμής.
- Η διαθεσιμότητα μετρήσεων και αποτελεσμάτων του μετεωρολογικού μοντέλου με μικρότερα χρονικά βήματα βοηθάει σημαντικά στην επίτευξη

καλύτερων ακόμη αποτελεσμάτων. Τα σετ δεδομένων των 10 λεπτών κρίνονται τα καταλληλότερα. Η δειγματοληψία των 10 λεπτών μπορεί να επιφέρει επιπλέον ελάττωση στο bias κατά 10% περίπου και στο RMSE κατά 15%, σε σχέση με τα ωριαία δεδομένα.

- Τα Bayesian Kalman φίλτρα επιφέρουν σημαντικές βελτιώσεις στα προγνωστικά δεδομένα, χρησιμοποιώντας πολύ χαμηλή υπολογιστική ισχύ.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Allouhi, O. Zamzoum, M.R. Islam, R. Saidur, T. Kousksou, A. Jamil, A. Derouich. 2017. Evaluation of wind energy potential in Morocco's coastal regions. *Renewable and Sustainable Energy Reviews* 72 (2017) 311–324.
- Papoulis and S. U. Pillai. *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill Higher Education, 2002.
- Adam S. Charles (2010), “Kalman Filtering: A Bayesian Approach”
- Bernardo, J., Smith, A., 2000. *Bayesian Theory*. John Wiley & Sons, LTD, New York, USA.
- Box, G., Tiao, G., 1992. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons, INC, New York, USA.
- Bremnes, J.B., 2004: Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Mon. Wea. Rev.*, 132, 338-347.
- Buizza, R., Houtekamer P.L., Toth Z., Pellerin G., Wei M., and Zhu Y., (2005). A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, Vol(13), 1076–1097.
- Chen, C. and W.R. Cotton, (1983), “A one-dimensional simulation of the stratocumulus-capped mixed layer”, *Boundary-layer Meteorol.*, 25, 289-321.
- Clark M.P. and L.E. Hay, 2004, Use of medium–range numerical weather prediction model output to produce forecasts of streamflow. *Journal of Hydrometeorology.*,15-32.
- Cotton, W.R., Pielke Sr., R.A., Walko, R.L., Liston, G.E., Tremback, C.J., Jiang, H., McAnelly, R.L., Harrington, J.Y., Nicholls, M.E., Carrio, G.G., McFadden, J.P., (2003), “RAMS 2001: current status and future directions”, *Meteor. Atmos. Phys.* 82, 5–29.
- Delle Monache, L., T. Eckel, D. Rife, and B. Nagarajan, 2013: Probabilistic weather prediction with an analog ensemble. *Mon. Wea. Rev.*, 141, 3498–3516.
- Delle Monache, L., T. Nipen, Y. Liu, G. Roux, and R. Stull, 2011: Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Wea. Rev.*, 139, 3554–3570.
- Dragulanescu, L., (1993), “Application des filtres Kalman pour ajuster les temperatures prognosees avec un modele numerique Meteor “, *Hydrol.* 23, 11–14.
- E. A. Bossanyi: Short-Term Wind Prediction Using Kalman Filters, *Wind Engineering* 9, 1–8, 1985.
- Emmanouil, G., Galanis, G., Kallos, G., 2006. Statistical methods for the prediction of night-time cooling and minimum temperature. *Meteorol. Appl.* 13 (2), 169–178.
- Friederichs, P., Mathes, A. and Hense, A., 2006: Statistical down-scaling of extreme precipitation events using censored quantile regression. *Mon. Wea. Rev.*

- Galanis G, Louka P, Katsafados P, Kallos G and Pytharoulis I. 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. *Annales Geophysicae*, 24, 2451–2460.
- Galanis, G. 2014. “Statistical post processes for the improvement of the results of numerical wave prediction models”. *Mathematics of Planet Earth Lecture Notes in Earth System Sciences 2014*, 561-564
- Galanis, G., Anadranistakis, M., 2002. A one dimensional Kalman filter for the correction of near surface temperature forecasts, *meteor. Appl* 9, 437–441.
- Gangopadhyay, S., M. Clark, K. Werner, D. Brandon, and B. Rajagopalan, 2004: Effects of spatial and temporal aggregation on the accuracy of statistically downscaled precipitation estimates in the upper Colorado River basin. *J. Hydrometeor.*, 5, 1192–1206.
- Gentine, P., Pritchard, M. S., Rasp, S., Reinaudi, G. and Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*.
- George Galanis, Evgenia Papageorgiou, Aristotelis Liakatas. 2017. A hybrid Bayesian Kalman filter and applications to numerical wind speed modeling. *Journal of Wind Engineering & Industrial Aerodynamics* 167 (2017) 1–22
- George Galanis, George Emmanouil, Christina Kalogeri and George Kallos. 2014. “Mathematical and physical models for the estimation of wind-wave power potential in the Eastern Mediterranean Sea”. *Mathematics of Planet Earth Lecture Notes in Earth System Sciences 2014*, 561-564
- Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta-RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, 125, 1312.
- Harrington JY, Reisin T, Cotton WR, Kreidenweis SM (1999), “Cloud resolving simulations of Arctic stratus. Part II: Transition-season clouds”, *Atmos Res* 55: 45–75.
- Harry R Glahn and Dale A. Lowry, ” The use of Model Output Statistics (MOS) in Objective Weather Forecasting”, 1972
- Hoeting, J.A., Madigan, D., Raftery, A.E. & Volinsky, C.T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, 14, 382- 417.
- Igor Oliveira and C. N. dos Santos, 2015. Post-processing of Model data with Machine-Learning Techniques for Operational Precipitation Forecast, Rio de Janeiro, RJ, Brazil 2015. IBM Research
- Jacks, E., J. B. Bower, V. J. Dagostaro, J. P. Dallavalle, M. C. Erickson, and J. C. Su, 1990: New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind. *Wea. Forecasting*, 5, 128–138
- Kain, J.S., and J.M. Fritsch, (1993), “Convective parameterization for mesoscale models: The Kain- Fritsch scheme. The representation of cumulus convection in numerical models”, *Meteor. Monogr.*, No. 24., Amer. Meteor. Soc., 165-170.
- Kalman RE. 1960. A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D*, 82, 35–45.
- Kalman, R.E., Bucy, R.S., (1961), “New results in linear filtering and prediction problems”, *Trans. ASME Ser. D* 83, 95–108.

- Kalnay, E., (2002), “Atmospheric Modeling, Data Assimilation and Predictability”, Cambridge University Press, Cambridge, 341pp.
- Kass, R.E. and Raftery A.E. (1993), “Bayes Factors and Model Uncertainty”, Technical Report 571, Carnegie Mellon University, Dept. of Statistics.
- Lagerquist, R., McGovern, A. and Smith, T. 2017. Machine learning for real-time prediction of damaging straight-line convective wind. *Weather and Forecasting*, 32, 2175–2193.
- Leamer, E.E. (1978), *Specification Searches: Ad Hoc Inference With Non-experimental Data*, New York: John Wiley.
- Liu, Y., Racah, E., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M. and Collins (2016). Application of deep convolutional neural networks for detecting extreme weather in climate datasets.
- Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., Kallos, G., 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J. Wind Eng. Ind. Aerodyn.* 96, 2348–2362.
- Mc Gowan, Rogers ,“Wind energy expailned, Manwell”.
- Meyers, M.P., R.L. Walko, J.Y Harrington, and W.R. Cotton, (1997), “New RAMS cloud microphysics parameterization. Part II: the two moment scheme”, *Atmos. Res.*, 45, 3-39.
- Pablo A. Mendoza, *Statistical Postprocessing of High-Resolution Regional Climate Model Output 2014*.
- Persson, A., (1990), “Kalman filtering a new approach to adaptive statistical interpretation of numerical meteorological forecasts”, *ECMWF Newsletter* (June).
- Pielke R. A. and Coauthors, (1992), “A comprehensive meteorological modeling system—RAMS”, *Meteor. Atmos. Phys.*, **49**, 69–91.
- Platon Patlakas, Eleni Drakaki, George Galanis, Christos Spyrou and George Kallos. 2017. Wind gust estimation by combining a numerical weather prediction model and statistical post-processing. *Energy Procedia* 125(2017) 190–198 .
- Robert E Kass, Adrian E Raftery, 1995. “Bayes factors”. *Journal of the American Statistical Association*, Vol.(90), No 430, 773-795.
- Sophie Pelland, George Galanis and George Kallos 2011. “Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model”. *Prog. Photovolt: Res. Appl.* (2011). Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/pip.1180
- Stathopoulos, C., Kaperoni, A., Galanis, G., Kallos, G., 2013. Wind power prediction based on numerical and statistical Models. *J. Wind Eng. Ind. Aerodyn.* 112, 25–38.
- Subhrendu Gangopadhyay, Martyn Clark, Lauren Hay, Balaji Rajagopalan, Robert Wilby , *The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields*, 2004.

- Tilmann Gneiting Heidelberg Institute for Theoretical Studies (HITS) Karlsruhe Institute of Technology (KIT) ECMWF, Statistical Post-Processing of Ensemble Forecasts: Current Developments and Future Directions, Reading, United Kingdom, February 11, 2015.
- Vanvyve, E., L. Delle Monache, A. J. Monaghan, and J. O. Pinto, 2013: Wind resource estimates with an analog ensemble approach. In preparation for Renewable Energy
- Walko, R.L., W.R Cotton, G. Feingold, and B. Stevens, (2000), “Efficient computation of vapor and heat diffusion between hydrometeors in a numerical model”, Atmos. Res., 53, 171-183.
- Walko, R.L., W.R. Cotton, M.P Meyers, and J.Y. Harrington (1995), “New RAMS cloud microphysics parameterization. Part I: The single-moment scheme”, Atmos. Res., 38, 29-62
- Wang, F., Mi, Z., Su, S. and Zhao, H. 2012. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies, 5, 1355–1370.
- Wilks Daniel S., Thomas M. Hamill , Comparison of Ensemble-MOS Methods Using GFS Reforecasts, 2007. American Meteorological Society. Vol.(135), 2379-2390
- Wilks, D., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press, Lake City.
- Wilks, D., 2009. Extending logistic regression to provide full-probability-distribution MOS forecasts. METEOROLOGICAL APPLICATIONS, Meteorol. Appl. 16: 361–368 (2009)
- Zhang Jie, Caroline Draxl, Thomas Hopson, Luca Delle Monache, EmilieVanvyve, Bri-Mathias Hodge. 2015. Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods. Applied Energy 156 (2015) 528–541.
- Πέτρος Δελλαπόρτας – Παναγιώτης Τσιαμυρτζής, Σημειώσεις μαθήματος “Στατιστική κατά Bayes”, Οικονομικό Πανεπιστήμιο Αθηνών, Τμήμα Στατιστικής (2004).