



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ
ΤΟΜΕΑΣ ΜΕΤΕΩΡΟΛΟΓΙΑΣ ΚΑΙ ΚΛΙΜΑΤΟΛΟΓΙΑΣ



ΓΕΩΡΓΙΟΣ Κ. ΚΑΡΟΥΤΣΟΣ
Πτυχιούχος Μαθηματικός

**ΕΦΑΡΜΟΓΗ ΦΙΛΤΡΟΥ KALMAN
ΣΕ ΠΡΟΓΝΩΣΕΙΣ ΘΕΡΜΟΚΡΑΣΙΑΣ
ΤΡΙΩΝ ΑΡΙΘΜΗΤΙΚΩΝ ΜΟΝΤΕΛΩΝ ΠΡΟΓΝΩΣΗΣ
ΚΑΙΡΟΥ**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ ΕΙΔΙΚΕΥΣΗΣ

ΘΕΣΣΑΛΟΝΙΚΗ
2013

ΓΕΩΡΓΙΟΣ Κ. ΚΑΡΟΥΤΣΟΣ
Πτυχιούχος Μαθηματικός

ΕΦΑΡΜΟΓΗ ΦΙΛΤΡΟΥ KALMAN ΣΕ ΠΡΟΓΝΩΣΕΙΣ ΘΕΡΜΟΚΡΑΣΙΑΣ
ΤΡΙΩΝ ΑΡΙΘΜΗΤΙΚΩΝ ΜΟΝΤΕΛΩΝ ΠΡΟΓΝΩΣΗΣ ΚΑΙΡΟΥ

Υποβλήθηκε στο Τμήμα Γεωλογίας στα πλαίσια του Μεταπτυχιακού Προγράμματος
Σπουδών 'Μετεωρολογία, Κλιματολογία και Ατμοσφαιρικό Περιβάλλον'
Τομέας Μετεωρολογίας και Κλιματολογίας
Ημερομηνία Προφορικής Εξέτασης:

Τριμελής Συμβουλευτική Επιτροπή

Θεόδωρος Καρακώστας, Καθηγητής, Επιβλέπων

Ιωάννης Πυθαρούλης, Επίκουρος Καθηγητής

Πρόδρομος Ζάνης, Επίκουρος Καθηγητής

Αριθμός Παραρτήματος Επιστημονικής επετηρίδας Τμήματος Γεωλογίας Ν°

© Γεώργιος Κ. Καρούτσος, 2013

Με επιφύλαξη παντός δικαιώματος. All right reserved.

ΕΦΑΡΜΟΓΗ ΦΙΛΤΡΟΥ KALMAN ΣΕ ΠΡΟΓΝΩΣΕΙΣ ΘΕΡΜΟΚΡΑΣΙΑΣ ΤΡΙΩΝ ΑΡΙΘΜΗΤΙΚΩΝ ΜΟΝΤΕΛΩΝ ΠΡΟΓΝΩΣΗΣ ΚΑΙΡΟΥ

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευτεί ότι εκφράζουν τις επίσημες θέσεις του Α.Π.Θ.

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε στον τομέα Μετεωρολογίας και Κλιματολογίας του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης. Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου Θεόδωρο Καρακώστα για την παρότρυνσή του να ασχοληθώ με ένα πολύ ενδιαφέρον θέμα που συνδυάζει την μετεωρολογία, με τα μαθηματικά και ειδικά την αριθμητική πρόγνωση του καιρού, καθώς επίσης και για την πολύτιμη βοήθεια και τις συμβουλές του σε όλες τις φάσεις της εκπόνησης της διπλωματικής εργασίας αλλά και γενικότερα κατά τη διάρκεια των σπουδών μου στον τομέα. Επίσης, θα ήθελα να ευχαριστήσω τον καθηγητή Ιωάννη Πυθαρούλη ο οποίος ήταν πάντα διαθέσιμος για διευκρινήσεις, απορίες αλλά και χρήσιμες πληροφορίες σχετικά με την εργασία. Έπειτα, θα ήθελα να ευχαριστώ τους καθηγητές του τομέα Μετεωρολογίας και Κλιματολογίας που με καθοδήγησαν τα τελευταία δύο χρόνια στο πολύ ενδιαφέρον και ευρύ αντικείμενο της μετεωρολογίας και της κλιματολογίας. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την στήριξη και την συμπαράστασή τους σε όλη τη διάρκεια των σπουδών μου.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ	3
ΠΕΡΙΕΧΟΜΕΝΑ	4
ΚΕΦΑΛΑΙΟ 1 - Εισαγωγή	6
1.1 Εισαγωγή	7
1.2 Ατμοσφαιρικά μοντέλα	7
1.3 Ιστορική επισκόπηση	9
1.4 Σφάλματα πρόγνωσης	13
1.5 Σκοπός της εργασίας	16
ΚΕΦΑΛΑΙΟ 2 - Το φίλτρο Kalman	20
2.1 Επεξεργασία σήματος	21
2.2 Κατηγορίες φίλτρων	22
2.2.1 Αναδρομικά φίλτρα	24
2.3 Το φίλτρο Wiener	24
2.4 Το φίλτρο Kalman	25
2.4.1 Ο Rudolf Emil Kalman	26
2.4.2 Το θεωρητικό μέρος	27
2.4.3 Ο Αλγόριθμος και οι μαθηματικές σχέσεις	29
ΚΕΦΑΛΑΙΟ 3 - Το φίλτρο Kalman στην Μετεωρολογία	35
3.1 Εισαγωγή	36
3.2 Οι εφαρμογές της μεθόδου, πλεονεκτήματα και μειονεκτήματα	36
ΚΕΦΑΛΑΙΟ 4 - Δεδομένα και Μεθοδολογία	42
4.1 Εισαγωγή	43
4.2 Δεδομένα	43

4.2.1 Το αριθμητικό μοντέλο Weather Research and Forecasting (WRF)	44
4.2.2 Το αριθμητικό μοντέλο πρόγνωσης καιρού Skiron/Eta	45
4.2.3 Το αριθμητικό μοντέλο πρόγνωσης καιρού Global Forecasting System	47
4.3 Μεθοδολογία	48
ΚΕΦΑΛΑΙΟ 5 - Αποτελέσματα.	55
5.1 Ανάλυση των σφαλμάτων της πρόγνωσης	56
5.2 Ανάλυση των σφαλμάτων μετά την εφαρμογή του φίλτρου	64
5.2.1 Επιλογή βέλτιστης μεθόδου	64
5.2.2 Ετήσιο και μηνιαίο σφάλμα της πρόγνωσης με χρήση του φίλτρου Kalman	69
5.3 Συμπεράσματα	81
ΠΕΡΙΛΗΨΗ	83
ABSTRACT	84
BIBΛΙΟΓΡΑΦΙΑ	85

Κεφάλαιο 1

Εισαγωγή

Κεφάλαιο 1

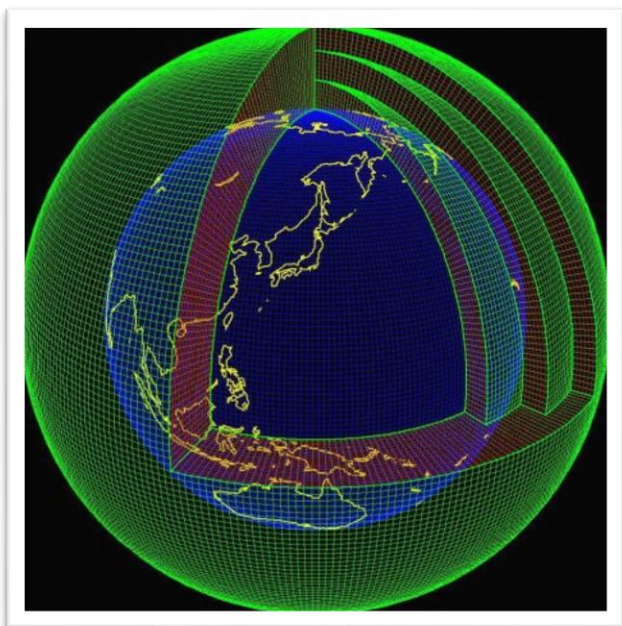
1.1 Εισαγωγή

Οι καιρικές συνθήκες ανέκαθεν επηρέαζαν τη ζωή και τις ασχολίες του ανθρώπου. Ακόμα και απλά πράγματα της καθημερινής ζωής του ανθρώπου, όπως η ένδυση, εξαρτώνται σε μεγάλο βαθμό από τις καιρικές συνθήκες, αλλά και όλες οι υπαίθριες δραστηριότητες επηρεάζονται άμεσα από τον καιρό. Επιπλέον, η επιτυχημένη καλλιέργεια της γης, η οποία είναι αναγκαία για την επιβίωσή του μπορεί να καταστραφεί σε μεγάλο βαθμό, ή και ολοκληρωτικά, από τη δύναμη της φύσης. Ο καιρός έχει καθορίσει ακόμα και την ίδια την ιστορία του ανθρώπινου είδους, είτε λόγω των καταστρεπτικών του συνεπειών εξαιτίας ακραίων φαινομένων που ο άνθρωπος δεν ήταν έτοιμος να αντιμετωπίσει όπως η ξηρασία στην Κίνα στα έτη 1876-1879 που είχε ως αποτέλεσμα περισσότερα από 10 εκατομμύρια νεκρούς (Zhi Xin Hao et al, 2010), είτε διότι έχει καθορίσει αρκετές φορές την έκβαση πολεμικών αναμετρήσεων με χαρακτηριστικά παραδείγματα την αποτυχία των Οθωμανών στην πολιορκία της Βιέννης το φθινόπωρο του 1529, την καταστροφή της Ισπανικής Αρμάδας το 1588, την ήττα του Ναπολέοντα στο Βατερλό το 1815 (<http://en.wikipedia.org>) και όχι μόνο. Έτσι έχει γίνει κατανοητή η αναγκαιότητα της σωστής πρόγνωσης των καιρικών φαινομένων χάρη στην οποία είναι δυνατός ο καλύτερος προγραμματισμός διάφορων δραστηριοτήτων.

1.2 Ατμοσφαιρικά μοντέλα

Σήμερα, η πρόγνωση του καιρού γίνεται με τη βοήθεια ατμοσφαιρικών μοντέλων. Ένα σύνολο διαφορικών εξισώσεων μετεωρολογικών μεταβλητών με μερικές παραγώγους ως προς το χρόνο λύνεται με αριθμητικές ή φασματικές μεθόδους, πάνω σε ένα τρισδιάστατο πλέγμα χωρικά διακριτών σημείων (Σχήμα 1), προσομοιώνοντας έτσι τις μεγάλης κλίμακας κινήσεις του αέρα στην ατμόσφαιρα, στη βάση των αρχών διατήρησης της ορμής, της μάζας, της ενέργειας και της

υγρασίας. Παράλληλα, ορισμένες φυσικές διαδικασίες, όπως η τυρβώδης ροή, οι ροές ενέργειας και υγρασίας, από και προς το έδαφος, η ροή των μικρού και μεγάλου μήκους κύματος ακτινοβολιών, οι μεγάλης κλίμακας διαδικασίες υγραποίησης και σχηματισμού νεφών, προσομοιώνονται με τη βοήθεια παραμετροποιήσεων, όταν είναι αδύνατον να αναλυθούν αυτές από το πλέγμα που χρησιμοποιείται. Για την αρχικοποίηση του μοντέλου χρησιμοποιούνται δεδομένα από παρατηρήσεις επιφάνειας και ανώτερης ατμόσφαιρας, αφού πρώτα υποστούν επεξεργασία από συστήματα αφομοίωσης δεδομένων, ενώ πλευρικές οριακές συνθήκες παρέχονται στην περιφέρειά του για όλες τις προγνωστικές μεταβλητές, για να είναι δυνατοί οι υπολογισμοί των προγνωστικών εξισώσεων. Ανάλογα με την έκταση της περιοχής ολοκλήρωσης, τα μοντέλα χωρίζονται σε δύο κατηγορίες: τα παγκόσμια, τα οποία περιλαμβάνουν ολόκληρη την επιφάνεια της γης και τα μοντέλα περιορισμένης περιοχής, τα οποία μπορούν να καλύπτουν πολύ μεγάλες εκτάσεις, όπως ολόκληρες ηπείρους, ή πολύ μικρές περιοχές, όπως ένα νομό. Η ανάγκη ύπαρξης των τελευταίων έγκειται στη δυσκολία πραγματοποίησης υψηλής ανάλυσης προσομοιώσεων για όλη τη γη λόγω του μεγάλου όγκου των υπολογισμών.



Σχήμα 1: Η περιοχή ολοκλήρωσης ενός παγκόσμιου αριθμητικού μοντέλου πρόγνωσης καιρού. Πηγή: <http://www.jma.go.jp>.

Χαρακτηριστικά παραδείγματα αριθμητικών μοντέλων πρόγνωσης καιρού είναι το μοντέλο του Ευρωπαϊκού Κέντρου Μεσοπρόθεσμων Προγνώσεων (ECMWF - Integrated Forecast System) (<http://www.ecmwf.int/>), το μοντέλο GFS της Εθνικής Υπηρεσίας Ωκεανών και Ατμόσφαιρας των Η.Π.Α (NOAA) (<http://www.emc.ncep.noaa.gov/index.php?branch=GFS>), το μοντέλο του Καναδικού Μετεωρολογικού Κέντρου (CMC - Global Environmental Multiscale Model) (http://collaboration.cmc.ec.gc.ca/science/rpn/gef_html_public). Στην Ελλάδα έχει αναπτυχθεί σε ερευνητικό επίπεδο το μοντέλο Σκίρων από το Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (<http://forecast.uoa.gr>). Εκτός από αυτά, υπάρχει μια πλειάδα άλλων που λειτουργούν σε επιχειρησιακό ή ερευνητικό επίπεδο.

1.3 Ιστορική Επισκόπηση

Η αξία της πρόγνωσης του καιρού έγινε αντιληπτή από τα αρχαία ακόμα χρόνια. Προσπάθειες για την επίτευξη του στόχου αυτού έγιναν από τους Βαβυλώνιους, τους Έλληνες, τους Κινέζους και τους Άραβες. Ωστόσο, όλες οι μέθοδοι που χρησιμοποιήθηκαν κατά καιρούς ήταν εμπειρικές, είχαν μικρή απόδοση και η αξιοπιστία τους ήταν περιορισμένη. Η εφεύρεση του θερμομέτρου και του βαρομέτρου, καθώς και η εφεύρεση του τηλέγραφου το 1835, έδωσαν τη δυνατότητα μέτρησης της θερμοκρασίας και της ατμοσφαιρικής πίεσης και τη σχεδόν στιγμιαία ανταλλαγή πληροφοριών σχετικές με τις καιρικές συνθήκες.

Το Νοεμβρίου 1854 κατά τη διάρκεια του Κριμαϊκού Πολέμου μια κακοκαιρία επηρέασε τη περιοχή του Εύξεινου Πόντου προκαλώντας μεγάλες ζημιές στον Αγγλικό και το Γαλλικό στόλο. Ο Γάλλος Υπουργός των Στρατιωτικών με αφορμή το γεγονός αυτό ζήτησε από τον διευθυντή του Αστεροσκοπείου των Παρισίων Urbain Le Verrier να διενεργήσει μια έρευνα ώστε να φανεί εάν θα μπορούσε να έχει γίνει πρόβλεψη της κακοκαιρίας. Ο Le Verrier συλλέγοντας παρατηρήσεις από διακόσιους και πλέον μετεωρολογικούς σταθμούς της Ευρώπης και άλλες αναφορές διαπίστωσε ότι η μεγάλη ατμοσφαιρική διαταραχή που επηρέασε τον Εύξεινο Πόντο είχε προηγουμένως διατρέξει από τα δυτικά προς τα ανατολικά όλη την Ευρώπη. Κατέληξε λοιπόν στο συμπέρασμα ότι η πρόγνωση του καιρού είναι εφικτή αν υπάρχει συχνή

καταγραφή μετεωρολογικών παρατηρήσεων από πολλές περιοχές και άμεση αποστολή των συλλεγμένων στοιχείων στα ειδικά μετεωρολογικά κέντρα (Καρακώστας, 2010).

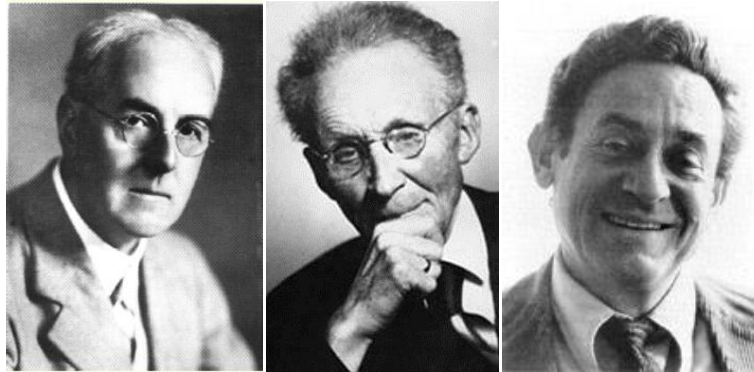
Στις αρχές του 20^{ου} αιώνα ο Νορβηγός Vilhelm Bjerknes (Σχήμα 3) ισχυρίστηκε ότι η επιστήμη της μετεωρολογίας είχε τέτοια εξέλιξη που πλέον επέτρεπε την πρόγνωση του καιρού με τη βοήθεια των μαθηματικών. Για αυτό το λόγο ανέπτυξε ένα σύνολο 7 διαφορεικών εξισώσεων, η λύση των οποίων θα προέβλεπε εξ αρχής τις μεγάλης κλίμακας κινήσεις της ατμόσφαιρας. Ο ίδιος πρότεινε μία λύση των εξισώσεων βασισμένη στην ανάλυση των χαρτών καιρού.

Το 1922 ο Lewis Fry Richardson (Σχήμα 3) δημοσίευσε ένα βιβλίο στο οποίο παρουσίασε την προσπάθειά του να κάνει μία πρόγνωση καιρού (Lynch, 1999). Υπέθεσε ότι η κατάσταση της ατμόσφαιρας μπορεί να περιγραφεί σε κάθε σημείο από 7 αριθμούς που θα περιγράφουν τις τιμές της ατμοσφαιρικής πίεσης, της θερμοκρασίας, της πυκνότητας, της υγρασίας και της ταχύτητας του ανέμου σε διεύθυνση βορρά-νότου, ανατολής-δύσης αλλά και την κτακόρυφη συνιστώσα της. Χρησιμοποίησε το μαθηματικό μοντέλο του V. Bjerknes το οποίο περιέγραφε τις μεταβολές στις τιμές αυτών των μεταβλητών με τη βοήθεια διαφορικών εξισώσεων. Για την επίλυση αυτών των εξισώσεων διαίρεσε την ατμόσφαιρα σε στήλες οι οποίες είχαν μήκος 3° και πλάτος 200Km. Έτσι σχηματίστηκαν $120 \times 100 = 12000$ στήλες, που κάλυπταν όλη την υδρόγειο. Κάθε στήλη χωρίστηκε σε πέντε κελιά, στο κέντρο των οποίων δίνονταν οι τιμές των μεταβλητών. Για αρχικές συνθήκες χρησιμοποίησε 12 ραδιοβολίσσεις και άλλες 18 αναφορές για ανέμους της ανώτερης ατμόσφαιρας για τις 20 Μαΐου 1910 στις 7 UTC. Τέλος οι διαφορικές εξισώσεις λύθηκαν προσεγγιστικά με τη μέθοδο των πεπερασμένων διαφορών. Όπως αναφέρει ο ίδιος, χρειάστηκε 6 εβδομάδες για να ολοκληρώσει με το χέρι τους υπολογισμούς για τις δύο στήλες στο κέντρο της Ευρώπης, ώστε να υπολογίσει τις μεταβολές στις τιμές της μάζας και του ανέμου 6 ώρες μετά τις αρχικές συνθήκες.



Σχήμα 2: Το εργαστήριο πρόγνωσης του καιρού όπως το φαντάστηκε ο L.F. Richardson. Πηγή: <http://celebrating200years.noaa.gov>.

Το αποτέλεσμα ήταν εκτός πραγματικότητας, καθώς προέβλεπε αύξηση της ατμοσφαιρικής πίεσης στην επιφάνεια κατά 145hPa, γεγονός όμως που όπως αργότερα αποδείχθηκε, οφειλόταν στις ανεπάρκειες και ανισορροπίες των αρχικών δεδομένων και όχι στη μέθοδο που χρησιμοποιήθηκε (Lynch, 1993). Αυτή ήταν η πρώτη απόπειρα αριθμητικής πρόγνωσης του καιρού.



Σχήμα 3: Από αριστερά προς τα δεξιά οι L.F. Richardson, V. Bjerknes και J. G. Charney. Πηγή: <http://www.nap.edu>.

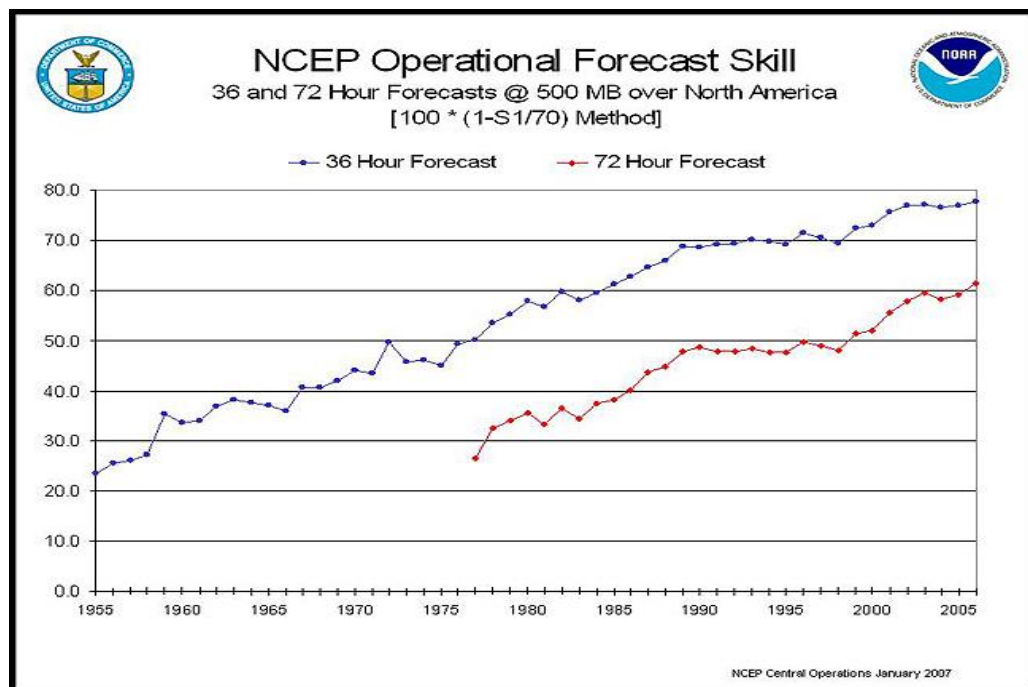
Για χρόνια η μέθοδος που προτάθηκε από τον Richardson δεν βρήκε ανταπόκριση, είτε λόγω του αποτυχημένου αποτελέσματος, είτε λόγω του μεγάλου όγκου υπολογισμών που χρειαζόταν για να πραγματοποιηθεί και οι οποίοι απαιτούσαν υπερβολικό χρόνο. Όταν εφευρέθηκε ο πρώτος υπολογιστής, ο ENIAC, τότε άρχισαν να γίνονται και πάλι προγνώσεις καιρού με τη βοήθεια αυτού. Οι J. Charney (Σχήμα 3), R. Fjortoft, και J. von Neumann (1950) πρότειναν μία μέθοδο με την οποία γινόταν αριθμητική επίλυση της εξίσωσης του στροβιλισμού πάνω από μία περιοχή της γης, θεωρώντας την ατμόσφαιρα βαροτροπική. Τα σημεία του πλέγματος που χρησιμοποιήθηκε είχαν απόσταση 736Km, ενώ συνολικά το πλέγμα είχε $15 \times 18 = 270$ σημεία και κάλυπτε ολόκληρη τη Βόρεια Αμερική και ένα μέρος του Ατλαντικού ωκεανού. Το μοναδικό κατακόρυφο επίπεδο που υπήρχε ήταν αυτό στα 500hPa. Τα λάθη που παρατηρήθηκαν αποδόθηκαν εν μέρει στη χαμηλή χωρική διακριτοποίηση, αλλά και στην υπόθεση ότι η ατμόσφαιρα είναι βαροτροπική. Ο χρόνος που απαιτούνταν για την 24ωρη πρόγνωση ήταν περίπου 24 ώρες πράγμα που σήμαινε ότι οι υπολογισμοί γίνονταν παράλληλα με την εξέλιξη των φαινομένων που προσομοίωναν. Η εργασία αυτή των J. Charney, R. Fjortoft, και J. von Neumann απέδειξε ωστόσο ότι η αριθμητική πρόγνωση καιρού ήταν εφικτή.

Στα επόμενα χρόνια άρχισε η αριθμητική πρόγνωση καιρού να αναπτύσσεται (Σχήμα 4). Με τη συνεχόμενη αναβάθμιση των υπολογιστών ήταν δυνατόν να γίνονται συνεχώς καλύτερες προγνώσεις. Τα σημεία πλέγματος των πρώτων αριθμητικών μοντέλων πρόγνωσης καιρού στις ΗΠΑ κάλυπταν μόνο το βόρειο ημισφαίριο και αυτό γιατί το υπολογιστικό κόστος ήταν ακόμα πολύ μεγάλο, αλλά και γιατί ήταν περιορισμένα τα δεδομένα που χρησιμοποιούνταν ως αρχικές συνθήκες. Στις δεκαετίες του '60 και του '70 αυξήθηκε η περιοχή ολοκλήρωσης, καθώς και η χωρική της διακριτοποίηση, ενώ έγιναν διαθέσιμες πλέον περισσότερες πηγές δεδομένων, αφού εκτός των άλλων προστέθηκαν και οι δορυφορικές παρατηρήσεις. Όμως παράλληλα με την εξέλιξη της τεχνολογίας, έγινε σημαντική πρόοδος και στην περιγραφή των φυσικών διεργασιών της ατμόσφαιρας. Στις ίδιες τις βασικές εξισώσεις γίνονται πλέον λιγότερες απλοποιήσεις και χρησιμοποιούνται πιο πλήρεις εξισώσεις για τη θερμοδυναμική της ατμόσφαιρας. Εκτός από τις βασικές εξισώσεις, στα νέα μοντέλα χρησιμοποιούνται και διάφορες παραμετροποιήσεις που περιγράφουν φυσικές διεργασίες, οι οποίες είναι αδύνατο να αναλυθούν από το πλέγμα του μοντέλου.

Το 1975 ιδρύθηκε το Ευρωπαϊκό Κέντρο Μεσοπρόθεσμων Προγνώσεων (ECMWF) αφού είχε γίνει κατανοητή η ανάγκη συσσώρευσης των επιστημονικών και τεχνικών πόρων των μετεωρολογικών υπηρεσιών και των θεσμικών οργάνων της Ευρώπης για την παραγωγή μεσοπρόθεσμων προγνώσεων καιρού, τα οικονομικά και κοινωνικά οφέλη των οποίων θα ήταν ανεκτίμητα.

Σήμερα, επιχειρησιακά κέντρα αριθμητικής πρόγνωσης καιρού, σε πολλά μέρη του κόσμου, χρησιμοποιούν καθημερινά ένα πλήθος μοντέλων. Τα μοντέλα αυτά παράγουν ένα μεγάλο εύρος προϊόντων και υπηρεσιών. Εκτός της κλασικής πρόγνωσης, τα μοντέλα υποστηρίζουν πολλές άλλες επιχειρησιακές και ερευνητικές δραστηριότητες όπως η προσομοίωση πλημμύρων, η παραγωγή δεικτών επικινδυνότητας για πυρκαγιές, η προσομοίωση μεταφοράς σκόνης από ερημικές περιοχές και άλλες. Εκατομμύρια παρατηρήσεις χρησιμοποιούνται, ώστε να αποτυπωθεί μια ολοκληρωμένη και όσο το δυνατόν με μεγαλύτερη λεπτομέρεια εικόνα της κατάστασης της ατμόσφαιρας. Η πλειονότητα αυτών των παρατηρήσεων προέρχεται από δορυφόρους και αφού υποβληθούν σε επεξεργασία αρχικοποιούν τα μοντέλα. Η

ταχύτητα των υπολογιστών επιτρέπει την πραγματοποίηση 14 τρισεκατομμυρίων πράξεων κάθε δευτερόλεπτο, επιτρέποντας έτσι την προσομοίωση αρκετών ημερών μέσα σε λίγες ώρες.



Σχήμα 4: Ικανότητα πρόγνωσης 36 και 72 ωρών της επιφάνειας των 500hPa . NCEP Central Operations January 2007. Πηγή: <http://celebrating200years.noaa.gov> .

1.4 Σφάλματα Πρόγνωσης

Η ανάπτυξη της πρόγνωσης των καιρικών συνθηκών με τη μέθοδο της αριθμητικής πρόγνωσης τα τελευταία 50 χρόνια είναι όπως φαίνεται ραγδαία. Παρά τη μεγάλη ανάπτυξη, αλλά και την αύξηση του ποσοστού επιτυχίας, τα διάφορα μοντέλα παρουσιάζουν σφάλματα κατά την προσομοίωση των διεργασιών που λαμβάνουν χώρα στην ατμόσφαιρα. Οι πηγές αυτών των σφαλμάτων μπορούν να συμπυκνωθούν στις παρακάτω κατηγορίες. Υπάρχουν τα σφάλματα που οφείλονται: (α) στις αρχικές συνθήκες που χρησιμοποιούνται και (β) τα λάθη που

προκαλούνται κατά το σχηματισμό και το τρέξιμο του μαθηματικού μοντέλου και οφείλονται στις μεθόδους που χρησιμοποιούνται.

Τα σφάλματα που οφείλονται στις αρχικές συνθήκες, δηλαδή τις τιμές ορισμένων μεταβλητών με τις οποίες περιγράφεται η εικόνα της κατάστασης της ατμόσφαιρας τη στιγμή από την οποία θα ξεκινήσει η προσομοίωση, προκαλούνται για τους εξής λόγους. Πρώτον, διότι το πλήθος των παρατηρήσεων που είναι διαθέσιμες δεν είναι, ούτε αρκετά πυκνό, ούτε ομοιόμορφα κατανεμημένο. Υπάρχουν περιοχές στις οποίες έχουν γίνει αρκετές παρατηρήσεις, αλλά και άλλες στις οποίες οι τιμές ακόμα και των πιο σημαντικών για το μοντέλο μεταβλητών, είναι άγνωστες. Αυτό συμβαίνει ιδιαίτερα πάνω από θαλάσσιες επιφάνειες και περιοχές αραιοκατοικημένες ή τελείως ακατοίκητες. Τα κενά που υπάρχουν καλύπτονται κατά ένα πολύ μεγάλο μέρος από τις δορυφορικές παρατηρήσεις. Πρόβλημα δημιουργείται και από τη μικρή στις περισσότερες περιπτώσεις συχνότητα ανανέωσης των δεδομένων. Ένας επιπλέον λόγος για τον οποίο οι αρχικές συνθήκες δημιουργούν σφάλματα στην πρόγνωση είναι το γεγονός ότι οι μετρήσεις περιέχουν τιμές που δεν ανταποκρίνονται πλήρως στην πραγματικότητα, αλλά έχουν ένα ποσοστό λάθους, είτε λόγω αδυναμίας του ίδιου του οργάνου, είτε λόγω κακής μετάδοσης των δεδομένων.

Για την άλλη πηγή σφαλμάτων μπορεί να είναι και το ίδιο το μαθηματικό μοντέλο που χρησιμοποιείται. Αυτό συμβαίνει για τους παρακάτω λόγους: πρώτον γιατί οι εξισώσεις που εφαρμόζονται είναι ελλειπείς, αφού έχουν υποστεί απλοποιήσεις, ώστε να μειωθεί το υπολογιστικό κόστος. Όροι στις εξισώσεις οι οποίοι συνεισφέρουν στο τελικό αποτέλεσμα σε πολύ μικρό ποσοστό θεωρούνται αμελητέοι και διαγράφονται αφήνοντας λιγότερους για υπολογισμό. Δεύτερον, γιατί οι λύσεις που προκύπτουν είναι προσεγγιστικές και όχι αναλυτικές, καθώς είναι αδύνατη η επίλυση των μη γραμμικών διαφορικών εξισώσεων με μερικές παραγώγους με τις γνωστές αναλυτικές μεθόδους. Επιπλέον, για να είναι εφικτοί οι υπολογισμοί, οι συνεχείς συναρτήσεις στις εξισώσεις αναπαριστώνται από ένα πεπερασμένο πλήθος σημειακών τιμών, γεγονός που οδηγεί σε σφάλματα αποκοπής. Προβλήματα στην πρόγνωση δημιουργεί και η όχι ακριβής απεικόνιση των υψομετρικών μεταβολών της επιφάνειας της γης, η οποία σε μεγάλο βαθμό εξαρτάται από τη χωρική διακριτοποίηση των σημείων του πλέγματος. Τέλος, αιτία σφαλμάτων αποτελούν και οι πλευρικές οριακές συνθήκες (lateral boundary

conditions), καθώς αυτές μπορούν να είναι διαθέσιμες, είτε από κάποιο σύστημα αφομοίωσης δεδομένων, είτε από κάποιο άλλο μοντέλο μεγαλύτερης περιοχής, όπως είναι ένα παγκόσμιο μοντέλο.

Οι παραπάνω αιτίες σφαλμάτων δημιουργούν λάθη κατά τη διάρκεια των υπολογισμών. Ακόμα και αν εξαλείφονταν όλες οι παραπάνω πηγές σφαλμάτων και πάλι θα υπήρχε ένα μέρος του σφάλματος στο τέλος. Αυτό συμβαίνει γιατί υπάρχουν ατέλειες στην παραμετροποίηση των φυσικών διεργασιών. Για παράδειγμα, δεν είναι δυνατή η πλήρης περιγραφή των μεγάλης και μέσης κλίμακας διαδικασιών σχηματισμού βροχόπτωσης, αλλά και οι διαδικασίες σχηματισμού υετού μέσα στα κατακόρυφης ανάπτυξης νέφη, με αποτέλεσμα να μην προσομοιώνονται με ακρίβεια οι νεφικοί σχηματισμοί. Σε παρόμοια λάθη στην πρόγνωση μπορεί να οδηγήσουν οι εξισώσεις που περιγράφουν τις ροές υγρασίας, ενέργειας και ορμής, τόσο μεταξύ του εδάφους και του ατμοσφαιρικού αέρα, όσο και μεταξύ του ατμοσφαιρικού οριακού στρώματος και της ελεύθερης ατμόσφαιρας, καθώς δεν περιγράφουν πλήρως την πραγματικότητα.

Όλες οι παραπάνω πηγές σφαλμάτων συνεισφέρουν, άλλες σε μικρότερο, και άλλες σε μεγαλύτερο βαθμό, στη διαμόρφωση του τελικού σφάλματος της πρόγνωσης. Κάθε μεταβλητή ωστόσο δεν παρουσιάζει το ίδιο σφάλμα με όλες τις άλλες, αλλά επηρεάζεται σε διαφορετικό βαθμό. Τα μεγαλύτερα σφάλματα παρατηρούνται συνήθως στις μεταβλητές της επιφάνειας της γης και αυτό γιατί εκείνες οι μεταβλητές βρίσκονται στα όρια του πλέγματος και οι τιμές τους εξαρτώνται σε μεγάλο βαθμό από τον τρόπο που περιγράφονται οι αλληλεπιδράσεις ανάμεσα στο έδαφος και τον ατμοσφαιρικό αέρα και τις πλευρικές οριακές συνθήκες. Η καταγραφή και μελέτη των σφαλμάτων της πρόγνωσης είναι πολύ σημαντική, ώστε να φανεί το μέγεθος και οι πιο σημαντικές αιτίες του, καθώς αν είναι γνωστά αυτά, μπορεί να γίνει μία βελτίωση που θα αυξήσει το ποσοστό επιτυχίας ενός μοντέλου.

Οι (Katsafados et al., 2011) παρουσίασαν μια μελέτη της απόδοσης των προγνώσεων του μοντέλου WRF (Weather Research and Forecasting Model) για την περιοχή της Μεσογείου στην οποία φαίνεται το μέγεθος των σφαλμάτων σε ορισμένες μεταβλητές. Για την εργασία αυτή χρησιμοποιήθηκαν δεδομένα επιφάνειας (ατμοσφαιρικής πίεσης στο ύψος της μέσης στάθμης της θάλασσας, θερμοκρασίας στα 2m και ταχύτητες ανέμου στα 10m) από 900 περίπου

μετεωρολογικούς σταθμούς, ώστε να ελεγχθούν και να συγκριθούν με τις προβλέψεις του μοντέλου για δύο συνεχή χρόνια (2009 και 2010). Από τα αποτελέσματα προέκυψε, ότι η ρίζα του μέσου τετραγωνικού σφάλματος της ατμοσφαιρικής πίεσης στο ύψος της μέσης στάθμης της θάλασσας ήταν μεταξύ 1 hPa και 2.5 hPa και αυξανόταν συνεχώς με την προγνωστική ώρα, ενώ παρατηρήθηκε υποτίμηση της πραγματικής τιμής σε ένα εύρος από 0.1 hPa ως 0.5hPa. Το αντίστοιχο σφάλμα της θερμοκρασίας ήταν 2.7°C για το σύνολο της περιόδου χωρίς να μεταβάλλεται με την προγνωστική ώρα ή την εποχή. Κατά τη διάρκεια της ημέρας παρατηρήθηκε υπερεκτίμηση της πραγματικής τιμής της θερμοκρασίας κατά 0.5 °C, αλλά κατά τη νύχτα υποεκτίμηση κατά 0.7 °C. Για την ταχύτητα του ανέμου σε ύψος 10m από το έδαφος η τιμή της ρίζας του μέσου τετραγωνικού σφάλματος ήταν 2.5m/s – 3m/s, με τις τιμές να είναι μεγαλύτερες το χειμώνα και το φθινόπωρο από ότι την άνοιξη και το καλοκαίρι.

1.5 Σκοπός της Εργασίας

Η ανάλυση και ελαχιστοποίηση των σφαλμάτων μιας πρόγνωσης είναι ένα αντικείμενο έρευνας και μελέτης ιδιαίτερα σύγχρονο. Υπάρχουν τρόποι με τους οποίους γίνεται προσπάθεια να αντιμετωπιστεί αυτό το πρόβλημα. Ένας από αυτούς είναι η επιλογή υψηλότερης ανάλυσης του μοντέλου, τόσο χωρικά με αύξηση των σημείων πλέγματος της περιοχής ολοκλήρωσης, όσο και χρονικά με τη μείωση του χρονικού βήματος ανάμεσα στους υπολογισμούς. Έτσι τα χαρακτηριστικά των μικρής κλίμακας ροών είναι δυνατόν να προσομοιωθούν με μεγαλύτερη λεπτομέρεια. Παρόλα αυτά όμως, η υψηλότερη ανάλυση που θα χρησιμοποιηθεί δεν εγγυάται ότι η βελτίωση θα είναι τόσο μεγάλη. Ακόμα και αν υπάρχει όμως σημαντική βελτίωση, αυτή αντισταθμίζεται από το μεγαλύτερο υπολογιστικό κόστος, που είναι απαραίτητο για την ολοκλήρωση των υπολογισμών.

Παράδειγμα όπου φαίνεται ότι η αύξηση της ανάλυσης ενός μοντέλου δεν είναι ικανή συνθήκη για τη μείωση του σφάλματος είναι η εργασία των Kotroni and Lagouvardos (2004). Σε αυτή έγινε αποτίμηση της επίδοσης του πέμπτης γενιάς μέσης κλίμακας μοντέλου (MM5) του Εθνικού Κέντρου Έρευνας της Ατμόσφαιρας των ΗΠΑ (NCAR) για την περιοχή της Αθήνας. Χρησιμοποιήθηκαν δύο διαφορετικές αναλύσεις, μία των 8Km και μία υψηλότερης ανάλυσης

των 2Km, ενώ η περίοδος μελέτης ήταν το έτος 2002. Οι μεταβλητές που εξετάστηκαν ήταν η θερμοκρασία του αέρα στα 2m και η ταχύτητα του ανέμου στα 10m, ενώ παρατηρήσεις των τιμών αυτών των μεταβλητών πάρθηκαν από 4 μετεωρολογικούς σταθμούς των Αθηνών και της ευρύτερης περιοχής. Τα αποτελέσματα έδειξαν ότι το μέσο απόλυτο σφάλμα των τιμών της θερμοκρασίας κοντά στην επιφάνεια ήταν μεταξύ 1.32°C και 3.38°C όταν χρησιμοποιήθηκε η ανάλυση των 8Km, ενώ για την υψηλότερη ανάλυση το μέσο απόλυτο σφάλμα κυμαινόταν μεταξύ 1.37°C και 3.34°C, δίνοντας μία πολύ μικρή βελτίωση. Οι τιμές της ταχύτητας του ανέμου παρουσίασαν μέσο απόλυτο σφάλμα σε ένα εύρος τιμών από 1.56 ως 2.43 m/s όταν χρησιμοποιήθηκε η ανάλυση των 8Km, ενώ το σφάλμα αυτό μειώθηκε μόλις κατά 0.16 m/s με την υψηλότερη ανάλυση.

Μία δεύτερη εργασία στη οποία έγινε υπολογισμός του σφάλματος ορισμένων μεταβλητών της επιφάνειας είναι αυτή των (Zielonka et al., 2008) από το πανεπιστήμιο της Πενσυλβάνιας των ΗΠΑ και στην οποία έκαναν μία σύγκριση των τιμών των μεταβλητών της επιφάνειας μεταξύ του MM5 και του WRF. Οι προσομοιώσεις έγιναν για το μήνα Μάιο του 2008. Τα αποτελέσματα έδειξαν ότι το μέσο απόλυτο σφάλμα για τη θερμοκρασία του αέρα ήταν μεγαλύτερο για το WRF, αφού κατά τη διάρκεια της 60-ωρης πρόγνωσης κυμαινόταν από 2°C ως 3°C, ενώ για το MM5 ήταν 1°C. Όταν χρησιμοποιήθηκε υψηλότερη ανάλυση δεν υπήρξε μείωση των σφαλμάτων αυτών. Το μέσο απόλυτο σφάλμα της ταχύτητας του ανέμου ήταν παραπλήσιο για τα δύο μοντέλα και κυμαινόταν από 2.6m/s ως 3.5m/s.

Ένας εναλλακτικός τρόπος μείωσης των σφαλμάτων και αύξησης της απόδοσης ενός αριθμητικού μοντέλου πρόγνωσης καιρού είναι η εφαρμογή στατιστικών μεθόδων στις τιμές των μεταβλητών που εξάγει το μοντέλο. Τέτοιες μέθοδοι είναι η τέλεια προγνωστική μέθοδος (perfect prognosis method) (Klein and Lewis, 1970) αλλά και η στατιστική ανάλυση της εξόδου του μοντέλου (model output statistics) (Glahn and Lowry, 1972) οι οποίες εφαρμόζονται κυρίως σε μεταβλητές επιφάνειας. Στόχος αυτών είναι η ανάπτυξη εξισώσεων μεταξύ των παρατηρούμενων τιμών μιας μεταβλητής και αντικειμενικών προγνωστικών βοηθημάτων. Δημιουργούνται δηλαδή κατάλληλες στατιστικές σχέσεις ανάμεσα στις προγνωστικές μεταβλητές και τις παρατηρούμενες. Οι σχέσεις μπορούν να προσδιορίζονται με γραμμική ή μη γραμμική παλινδρόμηση, πολλαπλή διακριτική ανάλυση, ή άλλες στατιστικές μεθόδους. Τόσο όμως η μία, όσο και η άλλη μέθοδος, απαιτούν μία μεγάλη συλλογή δεδομένων, ενώ η

εγκυρότητα των εξισώσεων χάνεται όταν χρειαστεί να γίνει μία αλλαγή στα χαρακτηριστικά του μοντέλου και γίνεται αναγκαία μία νέα βαθμονόμηση.

Μία άλλη στατιστική μέθοδος η οποία μπορεί να διορθώσει τις τιμές ενός μοντέλου, είναι η εφαρμογή του φίλτρου Kalman. Αποτελείται από ένα σύνολο μαθηματικών εξισώσεων, που παρέχουν μία αποτελεσματική υπολογιστική (αναδρομική) λύση, η οποία ελαχιστοποιεί τη μέση τιμή του τετραγωνικού σφάλματος, με πολύ μικρό υπολογιστικό κόστος και εύκολη προσαρμογή σε οποιαδήποτε μεταβολή των παρατηρήσεων (Welch and Bishop, 2004). Το φίλτρο Kalman προσδιορίστηκε και δημοσιεύτηκε το 1960 από τον Αμερικανό, Ουγγρικής καταγωγής, ηλεκτρολόγο μηχανικό Rudolf Emil Kalman, ως μία αναδρομική λύση στο πρόβλημα της απαλοιφής ανεπιθύμητου θορύβου από μετρήσεις μεταβλητών που περιγράφουν μια διαδικασία. Μία από τις πρώτες εφαρμογές του ήταν στο σύστημα πλοήγησης για το πρόγραμμα Apollo της NASA (www.ion.org/museum/item_view.cfm?cid=6&scid=5&iid=293). Στη Μετεωρολογία, από τις πρώτες προσπάθειες εφαρμογής του φίλτρου Kalman είναι αυτή του Pearsson και του Simonsen στις αρχές της δεκαετίας του 1990 (Homleid, 2004). Στόχος τους ήταν η βελτίωση της εκτίμησης της πραγματικής τιμής της θερμοκρασίας που γινόταν από ένα αριθμητικό μοντέλο πρόγνωσης καιρού.

Η ιδέα πίσω από το φίλτρο Kalman μπορεί να φανεί σε ένα απλό πρόβλημα (Mackenzie, 2003). Αν για παράδειγμα είναι διαθέσιμες δύο εκτιμήσεις, x_1 και x_2 , της πραγματικής τιμής μιας μεταβλητής x , τότε στην ερώτηση ποιος συνδυασμός των x_1 και x_2 προσεγγίζει καλύτερα την πραγματική τιμή της x , η απάντηση εξαρτάται από το πόσο μεγάλη αβεβαιότητα υπάρχει σε κάθε μία εκτίμηση. Καθώς η εκτίμηση της αβεβαιότητας της μέτρησης δίνεται από τη διασπορά, αυτή θα είναι σ_1 και σ_2 αντίστοιχα. Είναι εύκολο να βρεθεί ότι ο συνδυασμός των x_1 και x_2 με την μικρότερη δυνατή διασπορά είναι:

$$x = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} x_2$$

Αυτή η εξίσωση είναι απόλυτα συνεπής, καθώς αν μία από τις διασπορές (π.χ η σ_2) είναι άπειρη, τότε θα πρέπει να χρησιμοποιηθεί η άλλη μεταβλητή (δηλαδή η x_1), ως εκτίμηση της πραγματικής κατάστασης. Αλλά, αν και οι δύο διακυμάνσεις είναι πεπερασμένες, τότε ο τύπος δίνει τα σωστά βάρη για το καλύτερο αποτέλεσμα.

Σε αυτήν την εργασία γίνεται εφαρμογή του φίλτρου Kalman σε τιμές θερμοκρασίας του ατμοσφαιρικού αέρα. Χρησιμοποιούνται δεδομένα από τρία αριθμητικά μοντέλα πρόγνωσης καιρού (WRF, GFS και Skiron/Eta), ώστε να βρεθεί η μείωση του σφάλματος που μπορεί να παρατηρηθεί με αυτή τη μέθοδο κάθε φορά. Καταρχήν, γίνεται μία σύγκριση ανάμεσα σε παραλλαγές που έχουν κατά καιρούς προταθεί, ώστε να φανεί αν υπάρχει ουσιαστική διαφορά ανάμεσα σε αυτές. Στόχος πάντα είναι η αφαίρεση, όσο το δυνατόν μεγαλύτερου τμήματος του συστηματικού λάθους της πρόγνωσης, που θα οδηγεί σε μηδενισμό της μέσης τιμής του σφάλματος της πρόγνωσης. Με αυτό τον τρόπο, η τιμή της θερμοκρασίας που εξάγεται από το μοντέλο θα τροποποιείται, ώστε να πλησιάζει περισσότερο την πραγματική τιμή, αφού αυτή θα είναι απαλλαγμένη από ένα μέρος του λάθους που οφείλεται στις αρχικές συνθήκες που χρησιμοποιούνται και στον αλγόριθμο της πρόγνωσης.

Κεφάλαιο 2

Το φίλτρο Kalman

Κεφάλαιο 2

2.1 Επεξεργασία σήματος

Συνήθως, ο όρος φίλτρο υποδηλώνει μια διαδικασία απομάκρυνσης μη επιθυμητών στοιχείων. Από την άλλη, ο όρος "σήμα" χρησιμοποιείται κυρίως στις τηλεπικοινωνίες για την περιγραφή μιας "πληροφορίας", η οποία μεταδίδεται από ένα σημείο σε κάποιο άλλο (Τίγκελης & Φραντζεσκάκης, 2008). Γενικότερα, ως σήμα θεωρείται κάθε συνάρτηση του χρόνου ή άλλης μεταβλητής, που αντιπροσωπεύει κάποια φυσική ποσότητα, η οποία εμφανίζει ενδιαφέρον. Έτσι, σήμα είναι η ανθρώπινη ομιλία, μια εικόνα, κάποιο ηλεκτρικό μέγεθος (τάση, ένταση, κ.λ.π), η ηχώ του ραντάρ, η θερμοκρασία του αέρα, κ.λ.π. Υπάρχουν πολλοί τύποι σημάτων ανάλογα με το μέσο και τον τρόπο μετάδοσης.

Στην επεξεργασία σήματος, ένα φίλτρο είναι μια συσκευή ή διαδικασία που αφαιρεί από ένα σήμα ορισμένες ανεπιθύμητες συχνότητες, ενώ επιτρέπει τις υπόλοιπες να περάσουν (Μουστακίδης, 2003). Στόχος είναι να αποκοπούν σήματα που παρεμβάλλονται και προκαλούν θόρυβο, ώστε να παραμείνουν μόνο τα χρήσιμα κομμάτια ενός σήματος, όπως οι συνιστώσες που βρίσκονται σε μια συγκεκριμένη περιοχή συχνοτήτων. Δηλαδή οι μέθοδοι επεξεργασίας στην πλειοψηφία τους στηρίζονται στην υπόθεση ότι το σήμα πληροφορίας και ο θόρυβος δεν έχουν κοινές συχνότητες και επομένως διαφοροποιούνται μέσω των συχνοτήτων τους. Η υιοθέτηση αυτής της υπόθεσης επιτρέπει την πλήρη απομάκρυνση του θορύβου με απλή γνώση των συχνοτικών ζωνών του σήματος πληροφορίας και θορύβου, δίχως να είναι απαραίτητη η ακριβής γνώση του συχνοτικού περιεχομένου των σημάτων, με τη χρήση απλών γραμμικών, χρονικά σταθερών, συστημάτων. Τα συστήματα αυτά καλούνται φίλτρα και η αντίστοιχη επεξεργασία φιλτράρισμα. Εκτός από το πεδίο των συχνοτήτων η επεξεργασία σήματος μπορεί να γίνει και στο πεδίο του χρόνου, ή στο χώρο.

Η συμπεριφορά ενός φίλτρου στο πεδίο των συχνοτήτων περιγράφεται μαθηματικώς με τη συνάρτηση μεταφοράς. Αυτή ορίζεται να είναι ο μετασχηματισμός Laplace του λόγου της συνάρτησης εξόδου του σήματος προς την αντίστοιχη της εισόδου. Η συνάρτηση μεταφοράς

ορίζει την απόκριση του φίλτρου σε οποιοδήποτε σήμα εισόδου, αλλά μπορεί να χρησιμοποιηθεί για να καθορίσει και πολλά άλλα χαρακτηριστικά του φίλτρου.

Ωστόσο, τα φίλτρα δεν χρησιμοποιούνται μόνο στο πεδίο των συχνοτήτων, αλλά μπορούν να χρησιμοποιηθούν και στο πεδίο του χρόνου. Αυτό γίνεται με την υιοθέτηση του χώρου καταστάσεων. Μια αναπαράσταση του χώρου καταστάσεων είναι ένα μαθηματικό μοντέλο ενός φυσικού συστήματος που περιγράφεται από ένα σύνολο μεταβλητών εισόδου, εξόδου και κατάστασης, οι οποίες αλληλοσυνδέονται με πρώτης τάξης διαφορικές εξισώσεις και για συντομία χώρου εκφράζονται ως διανύσματα. Αν το δυναμικό σύστημα είναι γραμμικό και χρονικά αναλλοίωτο, οι διαφορικές και αλγεβρικές εξισώσεις μπορούν να γραφούν σε μορφή πινάκων.

Η αναπαράσταση του χώρου καταστάσεων (επίσης γνωστή ως η «προσέγγιση στο πεδίο του χρόνου») παρέχει ένα βολικό και συμπαγή τρόπο για την κωδικοποίηση και την ανάλυση συστημάτων με πολλαπλές εισόδους και εξόδους, καθώς για ένα σύστημα με p εισόδους και q εξόδους, θα χρειάζονταν να γραφούν $q \times p$ μετασχηματισμοί Laplace για να κωδικοποιηθεί ολόκληρη η πληροφορία. Σε αντίθεση με την προσέγγιση στο πεδίο των συχνοτήτων, η χρήση της αναπαράστασης του χώρου καταστάσεων δεν περιορίζεται σε συστήματα με γραμμικές συνιστώσες και μηδενικές αρχικές συνθήκες. Ο χώρος κατάστασης είναι εκείνος του οποίου οι άξονες είναι οι μεταβλητές κατάστασης. Η κατάσταση του συστήματος μπορεί να αναπαρασταθεί ως ένα διάνυσμα εντός του εν λόγω χώρου.

2.2 Κατηγορίες φίλτρων

Υπάρχουν δυο κύρια είδη φίλτρων, τα αναλογικά και τα ψηφιακά (Smith, 1999). Είναι εντελώς διαφορετικά στη φυσική τους δομή και στον τρόπο με τον οποίο λειτουργούν. Κατά την αναλογική επεξεργασία σήματος ένα αναλογικό φίλτρο χρησιμοποιεί ηλεκτρονικά κυκλώματα που αποτελούνται από συστατικά όπως αντιστάσεις, πυκνωτές και ενισχυτές για να παραχθεί η απαιτούμενη επίδραση φίλτρου. Τέτοια κυκλώματα φίλτρων χρησιμοποιούνται ευρέως σε εφαρμογές όπως η μετάδοση, ενίσχυση και λήψη τηλεοπτικού σήματος, η επεξεργασία ήχου και

σε άλλες περιπτώσεις. Υπάρχουν καλά καθορισμένες τεχνικές για σχεδίαση κυκλώματος αναλογικού φίλτρου, για δοσμένες απαιτήσεις. Η επεξεργασία του σήματος είναι άμεση, δίχως καμία καθυστέρηση.

Από την άλλη, με ένα ψηφιακό φίλτρο επιτυγχάνεται η ψηφιακή επεξεργασία σήματος. Ο σκοπός της ψηφιακής επεξεργασίας σήματος συνήθως είναι η μέτρηση, το φιλτράρισμα ή η συμπίεση συνεχόμενων αναλογικών φυσικών σημάτων. Το πρώτο βήμα συνήθως είναι η μετατροπή του σήματος από αναλογικό σε ψηφιακό, μέσω δειγματοληψίας από έναν μετατροπέα (analog-to-digital converter, ADC), ο οποίος μετατρέπει το αναλογικό σήμα σε μια ακολουθία από αριθμούς. Αν το ζητούμενο σήμα εξόδου είναι επίσης αναλογικό, τότε χρειάζεται ένας μετατροπέας που θα μετατρέψει στο τέλος της διαδικασίας το σήμα από ψηφιακό σε αναλογικό (digital-to-analog converter, DAC). Αν και αυτή η διαδικασία μπορεί να είναι πιο πολύπλοκη σε σχέση με την αναλογική επεξεργασία και έχει διακριτό πεδίο τιμών, η χρήση υπολογιστικής ισχύος στην ψηφιακή επεξεργασία σήματος έχει πολλά πλεονεκτήματα σε σχέση με την αναλογική επεξεργασία σήματος σε πολλές εφαρμογές, όπως ο εντοπισμός και η διόρθωση λαθών στις επικοινωνίες και η συμπίεση δεδομένων. Ένα ψηφιακό φίλτρο χρησιμοποιεί ένα ψηφιακό επεξεργαστή για να εκτελέσει αριθμητικούς υπολογισμούς στις τιμές του σήματος. Ο επεξεργαστής μπορεί να είναι ένας γενικού σκοπού υπολογιστής, όπως ένα PC. Ένα ψηφιακό φίλτρο είναι σε θέση να προγραμματισθεί. Η λειτουργία του καθορίζεται από ένα πρόγραμμα στη μνήμη του επεξεργαστή. Αυτό σημαίνει ότι το ψηφιακό φίλτρο μπορεί να αλλαχθεί εύκολα χωρίς να επηρεαστεί το κύκλωμα (hardware), ενώ αντίθετα ένα αναλογικό φίλτρο, μπορεί μόνο να αλλαχθεί ξανασχεδιάζοντας το κύκλωμα του φίλτρου. Κάποια πεδία όπου χρησιμοποιείται η ψηφιακή επεξεργασία σήματος είναι η επεξεργασία σημάτων ήχου, η επεξεργασία σημάτων φωνής, η επεξεργασία σημάτων από σόναρ, ραντάρ και συστοιχίες αισθητήρων, η εκτίμηση φάσματος (spectral estimation), η στατιστική επεξεργασία σήματος, η ψηφιακή επεξεργασία εικόνας, η επεξεργασία σήματος στις τηλεπικοινωνίες, ο έλεγχος συστημάτων, η πρόγνωση καιρού, οι οικονομικές προβλέψεις, η επεξεργασία δεδομένων από σεισμούς, η ανάλυση και ο έλεγχος βιομηχανικών διεργασιών, η ιατρική απεικόνιση (medical imaging) όπως οι αξονικές και οι μαγνητικές τομογραφίες, η συμπίεση MP3, τα γραφικά υπολογιστών και πολλά άλλα.

Μια άλλη κατηγοριοποίηση των φίλτρων γίνεται με βάση το πως αυτά επηρεάζουν τις διάφορες συχνότητες ενός σήματος (Lacanette, 1991). Έτσι ένα φίλτρο ονομάζεται χαμηλοπερατό ή βαθυπερατό (low-pass) αν αφήνει ανεμπόδιστες τις συχνότητες μέχρι μια ορισμένη τιμή ενώ αποκόπτει όλες τις συχνότητες που έχουν τιμή μεγαλύτερη από αυτή. Από την άλλη, ένα ανωπερατό φίλτρο (high-pass) επιτρέπει να περάσουν μόνο οι υψηλές συχνότητες ενώ αποκόπτει όσες έχουν τιμές μικρές. Ένα ζωνοπερατό φίλτρο (band pass) επιτρέπει να περάσουν συχνότητες με τιμές που βρίσκονται σε μία ζώνη του φάσματος, ενώ τα φίλτρα αποκοπής ζώνης αποκόπτουν τις συχνότητες που βρίσκονται σε μία συγκεκριμένη ζώνη. Τέλος, τα ολοπερατά φίλτρα δεν αποκόπτουν καμία συχνότητα, αλλά επιτρέπουν σε όλες να περάσουν αλλάζοντας μόνον τη φάση του σήματος.

2.2.1 Αναδρομικά φίλτρα

Τα ψηφιακά φίλτρα χωρίζονται σε δύο κατηγορίες, τα αναδρομικά και τα μη αναδρομικά (Smith, 1999). Ένα αναδρομικό φίλτρο είναι αυτό, το οποίο χρησιμοποιεί ως εισαγόμενες τιμές και προηγούμενες εξαγόμενες τιμές. Αυτές, όπως και οι προηγούμενες εισαγόμενες τιμές, αποθηκεύονται στη μνήμη του επεξεργαστή. Η λέξη αναδρομικός κυριολεκτικά σημαίνει "επιστρέφω πίσω" και αναφέρεται στο γεγονός ότι τιμές που έχουν υπολογιστεί στο παρελθόν επιστρέφουν πίσω ως είσοδοι για τον υπολογισμό των νέων τιμών. Από τον ορισμό αυτό, μπορεί να φαίνεται ότι τα αναδρομικά φίλτρα απαιτούν περισσότερους υπολογισμούς, από τη στιγμή που υπάρχουν προηγούμενοι εξαγόμενοι όροι στην έκφραση του φίλτρου όπως και εισαγόμενοι όροι, όμως ισχύει συνήθως το αντίθετο. Για να επιτευχθεί μια δοσμένη απόκριση συχνότητας, με ένα αναδρομικό φίλτρο, γενικά απαιτούνται λιγότεροι όροι που πρέπει να υπολογιστούν από τον επεξεργαστή, σε σύγκριση με ένα μη αναδρομικό φίλτρο.

2.3 Το φίλτρο Wiener

Το φίλτρο Wiener είναι ένα φίλτρο που προτάθηκε από τον Norbert Wiener κατά τη διάρκεια της δεκαετίας του 1940 και δημοσιεύτηκε το 1949 (Smith, 1999). Ο σκοπός του είναι να μειώσει το ποσό της παρουσίας θορύβου σε ένα σήμα σε σχέση με μια εκτίμηση του

επιθυμητού σήματος και βασίζεται σε μια στατιστική προσέγγιση. Το ισοδύναμο για διακριτού χρόνου καταστάσεις του έργου του Wiener προήλθε από τον Andrey Kolmogorov και δημοσιεύθηκε το 1941. Συνήθως τα φίλτρα σχεδιάζονται για να πετύχουν μια επιθυμητή απόκριση συχνότητας βασισμένα στη λογική, ότι οι συχνότητες που περιέχουν περισσότερο σήμα παρά θόρυβο θα πρέπει να διατηρηθούν, ενώ όσες περιέχουν περισσότερο θόρυβο να αφαιρεθούν. Ωστόσο, ο σχεδιασμός του φίλτρου Wiener στηρίζεται σε μια διαφορετική προσέγγιση. Από τη μία γίνεται η υπόθεση ότι είναι γνωστές οι φασματικές ιδιότητες του αρχικού σήματος και του θορύβου και από την άλλη αναζητείται το γραμμικό χρονικά αναλλοίωτο φίλτρο, του οποίου η έξοδος θα είναι όσο το δυνατόν πιο κοντά στο επιθυμητό σήμα. Για κάθε συχνότητα καθορίζεται το ποσοστό του περιεχόμενου σήματος και το αντίστοιχο του θορύβου και η επιθυμητή απόκριση συχνότητας $H(f)$ λαμβάνεται από τον τύπο:

$$H(f) = \frac{S(f)^2}{S(f)^2 + N(f)^2}$$

όπου $S(f)$ το επιθυμητό σήμα της συχνότητας και $N(f)$ ο θόρυβος που περιέχεται στη συχνότητα. Όπως φαίνεται, το κριτήριο για την επιθυμητή έξοδο είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος, ένα σφάλμα που ορίζεται ως η διαφορά του πραγματικού σήματος από αυτό που λαμβάνεται από την έξοδο του φίλτρου. Υπό αυτή την έννοια το φίλτρο Wiener θεωρείται βέλτιστο.

Αν τώρα αντί να χρησιμοποιηθεί το πεδίο των συχνοτήτων χρησιμοποιηθεί το πεδίο του χρόνου δηλαδή μια αναπαράσταση του χώρου καταστάσεων τότε προκύπτει το φίλτρο Kalman.

2.4 Το φίλτρο Kalman

Σύμφωνα με τους Welch and Bishop (2004) το φίλτρο Kalman είναι ένα σύνολο από μαθηματικές εξισώσεις που παρέχει μία αποτελεσματική υπολογιστική (αναδρομική) λύση για την εκτίμηση της κατάστασης μιας διαδικασίας, με τέτοιο τρόπο που να ελαχιστοποιείται η μέση τιμή του τετραγωνικού σφάλματος της εκτίμησης της κατάστασης. Το φίλτρο είναι πολύ ισχυρό

σε κάποιες περιπτώσεις, καθώς υποστηρίζει εκτιμήσεις της κατάστασης για χρονικές στιγμές του παρελθόντος, του παρόντος, ακόμη και του μέλλοντος, και αυτό μπορεί να το κάνει ακόμα και όταν η ακριβής φύση του συστήματος είναι άγνωστη.

Το φιλτράρισμα δεδομένων με τη μέθοδο του Kalman είναι μια τεχνική για την εκτίμηση της κατάστασης ενός δυναμικού συστήματος από μια ακολουθία μετρήσεων που περιέχει ελλείψεις, ή θόρυβο. Δεν χρειάζονται απευθείας μετρήσεις των μεταβλητών κατάστασης, αλλά αρκεί η μέτρησή τους μέσω μιας γραμμικής λειτουργικής σχέσης. Θεωρητικά, είναι ένας εκτιμητής του γραμμικού προβλήματος ελαχίστων τετραγώνων, όπου καλείται κανείς να υπολογίσει τη στιγμιαία “κατάσταση” (state) ενός γραμμικού δυναμικού συστήματος που διαταράσσεται από το “λευκό θόρυβο” (white noise), δηλαδή από μετρήσεις που σχετίζονται γραμμικά με τη δυναμική κατάσταση του συστήματος, αλλά είναι αλλοιωμένες από το “λευκό θόρυβο”. Ο εκτιμητής είναι στατιστικά βέλτιστος ως προς οποιαδήποτε συνάρτηση του τετραγώνου του σφάλματος εκτίμησης.

2.4.1 Ο Rudolf Emil Kalman

Η μέθοδος αυτή προτάθηκε από τον Rudolf Emil Kalman (Σχήμα 5) το 1960 όταν δημοσίευσε μία εργασία του με τίτλο “A New Approach to Linear Filtering and Prediction Problems” (Hartmann, 2013). Ο R. E. Kalman γεννήθηκε στη Βουδαπέστη το 1930, και μετανάστευσε με την οικογένειά του για τις ΗΠΑ το 1943 (<http://emuseum.kyotoprize.org>). Πήγε στο MIT και ολοκλήρωσε το διδακτορικό του στο Columbia το 1957. Μελέτησε τη δουλειά του Wiener στην επεξεργασία σήματος και προσπάθησε επιτυχημένα να την εφαρμόσει από το πεδίο των συχνοτήτων στο πεδίο του χρόνου (χώρος κατάστασης) και έτσι προέκυψε το φίλτρο Kalman. Το φίλτρο Wiener χρησιμοποιείται με αναλογικά ηλεκτρονικά, ενώ αντίθετα το φίλτρο Kalman είναι ιδανικό για την διαχείριση ψηφιακών δεδομένων. Το φίλτρο Kalman χρησιμοποιήθηκε ως μέρος του συστήματος πλοήγησης στο πρόγραμμα Apollo της Εθνική Υπηρεσία Αεροναυτικής και Διαστήματος των Η.Π.Α (NASA), που ήταν και μία από τις πρώτες του εφαρμογές.



Σχήμα 5: Rudolf Emil Kalman (1930-).

Πηγή: <http://emuseum.kyotoprize.org>

2.4.2 Το θεωρητικό μέρος

Για να γίνουν πιο κατανοητές οι αρχές πάνω στις οποίες στηρίζεται το σκεπτικό αυτής της τεχνικής ας θεωρηθεί το επόμενο απλό πρόβλημα συνδυασμού δύο ή περισσότερων εκτιμήσεων μιας ποσότητας των οποίων η διακύμανση του σφάλματος είναι γνωστή, ώστε να υπολογιστεί με μεγαλύτερη ακρίβεια η πραγματική τιμή της ποσότητας (Maybeck, 1979). Έστω λοιπόν ότι είναι διαθέσιμη μία εκτίμηση χ_1 της ποσότητας χ , η οποία περιέχει σφάλμα ε_1 . Δηλαδή:

$$\chi_1 - \chi = \varepsilon_1$$

Γίνεται η υπόθεση ότι για το σφάλμα ε_1 ισχύει $E\{\varepsilon_1\}=0$, όπου $E\{y\}$ είναι η αναμενόμενη τιμή του y και ότι η διασπορά του σφάλματος είναι $E\{\varepsilon_1^2\} = \sigma_1^2$. Έστω τώρα μία δεύτερη εκτίμηση χ_2 η οποία περιέχει ένα σφάλμα ε_2 , δηλαδή:

$$\chi_2 - \chi = \varepsilon_2$$

και ισχύει για αυτή την εκτίμηση ότι $E\{\varepsilon_2\}=0$ και η διασπορά του σφάλματος είναι $E\{\varepsilon_2^2\} = \sigma_2^2$. Σκοπός τώρα είναι να συνδυαστούν οι δύο αυτές εκτιμήσεις ώστε να γίνει μια όσο το δυνατόν καλύτερη εκτίμηση της πραγματικής τιμής. Η συνάρτηση

$$S(\chi) = \frac{(\chi_1 - \chi)^2}{\sigma_1^2} + \frac{(\chi_2 - \chi)^2}{\sigma_2^2}$$

περιγράφει το μέγεθος του τετραγωνικού σφάλματος που προκύπτει από μία εκτίμηση η οποία στηρίζεται στις δύο διαθέσιμες παρατηρήσεις. Είναι αυτή η συνάρτηση που πρέπει να ελαχιστοποιηθεί για να βρεθούν τα βάρη με τα οποία θα σταθμιστούν οι δύο εκτιμήσεις ώστε να προκύψει με βάση αυτές μία καλύτερη. Το ελάχιστο της συνάρτησης είναι εκείνο το $\hat{\chi}$ για το οποίο ισχύει:

$$\begin{aligned} \frac{\partial S(\chi)}{\partial \chi} &= 0 \Leftrightarrow \\ -\frac{2(\chi_1 - \chi)}{\sigma_1^2} - \frac{2(\chi_2 - \chi)}{\sigma_2^2} &= 0 \Leftrightarrow \\ \frac{(\chi_1 - \chi)}{\sigma_1^2} + \frac{(\chi_2 - \chi)}{\sigma_2^2} &= 0 \Leftrightarrow \end{aligned}$$

Επομένως, η τιμή που πλησιάζει περισσότερο την πραγματικότητα, χρησιμοποιώντας τα διαθέσιμα δεδομένα, είναι η

$$\begin{aligned} \hat{\chi} &= \frac{\frac{\chi_1}{\sigma_1^2} + \frac{\chi_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \Leftrightarrow \\ \hat{\chi} &= \widehat{\sigma^2} \left(\frac{\chi_1}{\sigma_1^2} + \frac{\chi_2}{\sigma_2^2} \right) \end{aligned}$$

$$\text{όπου } \widehat{\sigma^2} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1}$$

Από τη σχέση αυτή φαίνεται ότι, αν η αξιοπιστία κάποιας από τις δύο εκτιμήσεις είναι πολύ μικρή, δηλαδή έχει μεγάλη πιθανότητα να είναι μακριά από την πραγματικότητα, τότε αυτή λαμβάνεται λιγότερο υπόψη και η τελική εκτίμηση βασίζεται περισσότερο στην πρώτη. Αν η ακρίβεια και των δύο εκτιμήσεων είναι ίδια, τότε απλά λαμβάνεται μία μέση τιμή αυτών, η οποία λογίζεται ως η βέλτιστη εκτίμηση.

Αν τώρα γίνει διαθέσιμη και μία τρίτη εκτίμηση χ_3 , τότε είναι δυνατόν να ληφθεί και αυτή υπόψη και να βελτιωθεί και άλλο η εκτίμηση της ποσότητας χ , αρκεί να είναι γνωστή η ακρίβειά της μέσω της διασποράς της σ_3^2 . Τότε η νέα βέλτιστη εκτίμηση $\hat{\chi}_v$ της ποσότητας χ θα δίνεται από τον τύπο:

$$\hat{\chi}_v = \frac{\frac{\hat{\chi}}{\sigma^2} + \frac{\chi_3}{\sigma_3^2}}{\frac{1}{\sigma^2} + \frac{1}{\sigma_3^2}}$$

Αυτή η διαδικασία μπορεί να συνεχιστεί αν γίνονται διαθέσιμες νέες εκτιμήσεις της πραγματικής τιμής του χ , αρκεί αυτές να συνοδεύονται από τη διασπορά του σφάλματος που περιέχουν. Αυτή είναι η βάση της διαδοχικής εκτίμησης (sequential estimation), η οποία είναι η διαδικασία πάνω στην οποία αναπτύχθηκε το φίλτρο Kalman.

2.4.3 Ο Αλγόριθμος και οι μαθηματικές σχέσεις

Ο στόχος λοιπόν είναι να ελαχιστοποιηθεί το σφάλμα της εκτίμησης της πραγματικής κατάστασης X_k ενός συστήματος, αν υπάρχει μια ακολουθία παρατηρήσεων $y_0, y_1, \dots, y_{k-1}$ που όμως δεν είναι απόλυτα ακριβείς (Reid, 2001). Γίνεται η υπόθεση ότι το σύστημα μπορεί να περιγραφεί από την ακόλουθη γραμμική εξίσωση:

$$x_{k+1} = F_k x_k + G_k u_k + w_k$$

όπου το διάνυσμα $x_k \in \mathbb{R}^n$ περιγράφει το χώρο κατάστασης τη χρονική στιγμή k , το $u_k \in \mathbb{R}^m$ είναι ένα διάνυσμα ελέγχου, το w_k είναι το σφάλμα του συστήματος, G_k είναι ο πίνακας ελέγχου και F_k ο πίνακας μετάβασης.

Επίσης, γίνεται η υπόθεση ότι οι μεταβλητές του χώρου κατάστασης δεν είναι άμεσα παρατηρήσιμες, αλλά συνδέονται γραμμικά με ένα μετρήσιμο διάνυσμα $z_k \in \mathbb{R}^l$ σύμφωνα με την εξίσωση:

$$z_k = H_k x_k + v_k$$

όπου H_k ο πίνακας παρατήρησης ή μέτρησης και v_k είναι το σφάλμα της μέτρησης.

Θεωρείται ότι το σφάλμα της διαδικασίας και το σφάλμα της μέτρησης είναι ανεξάρτητα μεταξύ τους και έχουν μηδενική μέση τιμή ενώ είναι γνωστοί οι πίνακες συνδιασποράς τους. Δηλαδή $E[w_k w_\lambda^T] = E[v_k v_\lambda^T] = 0$, αν $k \neq \lambda$, $E[w_k w_\lambda^T] = Q_k$, $E[v_k v_\lambda^T] = R_k$, αν $k = \lambda$ ενώ $E[w_k v_\lambda^T] = 0$, για κάθε k και λ . Επίσης, θεωρείται ότι η αρχική κατάσταση του συστήματος x_0 είναι ένα τυχαίο διάνυσμα που είναι ανεξάρτητο από το σφάλμα της διαδικασίας ή της μέτρησης με γνωστή μέση τιμή $\hat{x}_{0|0} = E[x_0]$ και πίνακα συνδιασποράς:

$$P_{0|0} = E[(\hat{x}_{0|0} - x_0)(\hat{x}_{0|0} - x_0)^T].$$

Με δεδομένες τις παραπάνω υποθέσεις, ο σκοπός του αλγόριθμου είναι να καθοριστεί ο εκτιμητής του φίλτρου για τη χρονική στιγμή $k+1$, ο οποίος αποτελεί τη βέλτιστη προσέγγιση της πραγματικής τιμής της κατάστασης $\hat{x}_{k+1}$ υπό την έννοια ότι ελαχιστοποιεί τη συνάρτηση του αναμενόμενου τετραγωνικού σφάλματος

$$\Sigma(\hat{x}_{k+1}) = E[(x_{k+1} - \hat{x}_{k+1})(x_{k+1} - \hat{x}_{k+1})^T]$$

Η εκτίμηση της κατάστασης τη χρονική στιγμή $k+1$ η οποία στηρίζεται στις μετρήσεις μέχρι και τη χρονική στιγμή k , συμβολίζεται με $\hat{x}_{k+1|k}$ και ουσιαστικά είναι μία πρόβλεψη της κατάστασης. Για αυτήν ισχύει ότι:

$$\begin{aligned} \hat{x}_{k+1|k} &= E[\hat{x}_{k+1|k} | z_1, \dots, z_k] = E[x_{k+1} | Z^k] = E[F_k x_k + G_k u_k + w_k | Z^k] \\ &= F_k E[x_k | Z^k] + G_k u_k + E[w_k | Z^k] = F_k \hat{x}_{k|k} + G_k u_k \end{aligned}$$

αφού $E[w_k | Z^k] = 0$ ενώ η διασπορά της εκτίμησης $P_{k+1|k}$ είναι το μέσο τετραγωνικό σφάλμα του εκτιμητή $\hat{x}_{k+1|k}$ και άρα:

$$\begin{aligned}
P_{\kappa+1|\kappa} &= E \left[(x_{\kappa+1} - \hat{x}_{\kappa+1|\kappa})(x_{\kappa+1} - \hat{x}_{\kappa+1|\kappa})^T \middle| Z^\kappa \right] \\
&= F_\kappa E \left[(x_\kappa - x_{\kappa|\kappa})(x_\kappa - x_{\kappa|\kappa})^T \middle| Z^\kappa \right] F_\kappa^T + E[w_\kappa w_\kappa^T] = F_\kappa P_{\kappa|\kappa} F_\kappa^T + Q_\kappa
\end{aligned}$$

Αν τώρα μια νέα μέτρηση $Z_{\kappa+1}$ γίνει διαθέσιμη τότε αυτή η πληροφορία μπορεί να χρησιμοποιηθεί για τη διόρθωση της πρώτης πρόβλεψης και την εύρεση μιας νέας εκτίμησης $\hat{x}_{\kappa+1|\kappa+1}$ που βασίζεται σε μετρήσεις ως τη χρονική στιγμή $\kappa+1$. Αυτή η νέα εκτίμηση θα είναι το σταθμισμένο άθροισμα της πρώτης πρόβλεψης $\hat{x}_{\kappa+1|\kappa}$ και της νέας παρατήρησης $Z_{\kappa+1}$.

$$\hat{x}_{\kappa+1|\kappa+1} = K'_{\kappa+1} \hat{x}_{\kappa+1|\kappa} + K_{\kappa+1} Z_{\kappa+1}$$

όπου $K'_{\kappa+1}$ και $K_{\kappa+1}$ τα ζητούμενα βάρη που πρέπει να ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα της εκτίμησης, όπου το σφάλμα της εκτίμησης δίνεται από τον τύπο:

$$\tilde{x}_{\kappa+1|\kappa+1} = \hat{x}_{\kappa+1|\kappa} - x_{\kappa+1}$$

Προκειμένου όμως ο εκτιμητής να είναι αμερόληπτος, χρειάζεται να ισχύει η σχέση:

$$E[\hat{x}_{\kappa+1|\kappa+1}] = E[x_{\kappa+1}]$$

Άρα θα είναι:

$$\begin{aligned}
E[\hat{x}_{\kappa+1|\kappa+1}] &= E[K'_{\kappa+1} \hat{x}_{\kappa+1|\kappa} + K_{\kappa+1} Z_{\kappa+1}] \\
&= E[K'_{\kappa+1} \hat{x}_{\kappa+1|\kappa} + K_{\kappa+1} H_{\kappa+1} x_{\kappa+1} + K_{\kappa+1} v_{\kappa+1}] \\
&= K'_{\kappa+1} E[\hat{x}_{\kappa+1|\kappa}] + K_{\kappa+1} H_{\kappa+1} E[x_{\kappa+1}] + K_{\kappa+1} E[v_{\kappa+1}] \\
&= K'_{\kappa+1} E[F_\kappa \hat{x}_{\kappa|\kappa} + G_\kappa u_\kappa] + K_{\kappa+1} H_{\kappa+1} E[x_{\kappa+1}] \\
&= K'_{\kappa+1} \{F_\kappa E[\hat{x}_{\kappa|\kappa}] + G_\kappa u_\kappa\} + K_{\kappa+1} H_{\kappa+1} E[x_{\kappa+1}] \\
&= K'_{\kappa+1} E[x_{\kappa+1}] + K_{\kappa+1} H_{\kappa+1} E[x_{\kappa+1}] = \{K'_{\kappa+1} + K_{\kappa+1} H_{\kappa+1}\} E[x_{\kappa+1}]
\end{aligned}$$

Επομένως λόγω του ότι ο εκτιμητής είναι αμερόληπτος προκύπτει:

$$K'_{k+1} + K_{k+1}H_{k+1} = I \Leftrightarrow K'_{k+1} = I - K_{k+1}H_{k+1}$$

και άρα η σχέση που δίνει τη νέα εκτίμηση θα πάρει τη μορφή:

$$\begin{aligned}\hat{x}_{k+1|k+1} &= K'_{k+1}\hat{x}_{k+1|k} + K_{k+1}z_{k+1} = \{I - K_{k+1}H_{k+1}\}\hat{x}_{k+1|k} + K_{k+1}z_{k+1} \\ &= \hat{x}_{k+1|k} + K_{k+1}\{z_{k+1} - H_{k+1}\hat{x}_{k+1|k}\}\end{aligned}$$

Από τη σχέση αυτή προκύπτει ότι η νέα εκτίμηση είναι το άθροισμα της πρώτης πρόβλεψης και ενός κλάσματος της διαφοράς ανάμεσα στην πρόβλεψη και την πραγματική μέτρηση. Ο πίνακας K_{k+1} ονομάζεται κέρδος Kalman.

Επόμενο βήμα είναι ο προσδιορισμός της συνδιασποράς του σφάλματος της νέας εκτίμησης $P_{k+1|k+1}$.

$$\begin{aligned}P_{k+1|k+1} &= E[\tilde{x}_{k+1|k+1}\tilde{x}_{k+1|k+1}^T | Z^k] = E[(x_{k+1} - \hat{x}_{k+1|k+1})(x_{k+1} - \hat{x}_{k+1|k+1})^T] \\ &= (I - K_{k+1}H_{k+1})E[\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T](I - K_{k+1}H_{k+1})^T \\ &\quad + K_{k+1}E[v_{k+1}v_{k+1}^T]K_{k+1}^T + 2(I - K_{k+1}H_{k+1})E[\tilde{x}_{k+1|k}v_{k+1}^T]K_{k+1}^T\end{aligned}$$

Έχοντας υπόψη ότι ισχύουν οι παρακάτω σχέσεις

$$E[v_{k+1}v_{k+1}^T] = R_{k+1}$$

$$E[\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T] = E[(\hat{x}_{k+1|k} - x_{k+1})(\hat{x}_{k+1|k} - x_{k+1})^T] = P_{k+1|k}$$

$$E[\tilde{x}_{k+1|k}v_{k+1}^T] = E[(\hat{x}_{k+1|k} - x_{k+1})v_{k+1}^T] = 0$$

η σχέση που δίνει το ζητούμενο γίνεται:

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T$$

Τέλος έμεινε να υπολογιστεί το κέρδος Kalman. Ο προσδιορισμός και αυτού γίνεται στη βάση της ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος της νέας εκτίμησης. Αν οι τιμές του πίνακα K_{k+1} μπορούν να περιγραφούν από τη συνάρτηση:

$$\begin{aligned}
S(K_{\kappa+1}) &= E(\tilde{x}_{\kappa+1|\kappa+1}\tilde{x}_{\kappa+1|\kappa+1}^T | Z^{\kappa+1}) = \text{trace}(E[\tilde{x}_{\kappa+1|\kappa+1}\tilde{x}_{\kappa+1|\kappa+1}^T | Z^{\kappa+1}]) \\
&= \text{trace}(P_{\kappa+1|\kappa+1}) \\
&= \text{trace}\{(I - K_{\kappa+1}H_{\kappa+1})P_{\kappa+1|\kappa}(I - K_{\kappa+1}H_{\kappa+1})^T + K_{\kappa+1}R_{\kappa+1}K_{\kappa+1}^T\}
\end{aligned}$$

όπου $\text{trace}(A)$ το ίχνος ενός πίνακα A , τότε για να βρεθεί το ελάχιστο της συνάρτησης $S(K_{\kappa+1})$ αρκεί να λυθεί η εξίσωση:

$$\frac{\partial S(K_{\kappa+1})}{\partial K_{\kappa+1}} = 0$$

χρησιμοποιώντας την ιδιότητα ενός συμμετρικού πίνακα B :

$$\frac{\partial}{\partial A}(\text{trace}(ABA^T)) = 2AB$$

επομένως

$$\frac{\partial S(K_{\kappa+1})}{\partial K_{\kappa+1}} = -2(I - K_{\kappa+1}H_{\kappa+1})P_{\kappa+1|\kappa}H_{\kappa+1}^T + 2K_{\kappa+1}R_{\kappa+1} = 0 \Leftrightarrow$$

$$K_{\kappa+1} = P_{\kappa+1|\kappa}H_{\kappa+1}^T[H_{\kappa+1}P_{\kappa+1|\kappa}H_{\kappa+1}^T + R_{\kappa+1}]^{-1}.$$

Συνοψίζοντας τις εξισώσεις του αλγόριθμου προκύπτουν πέντε σχέσεις που χωρίζουν τον αλγόριθμο σε δύο φάσεις. Η πρώτη είναι η φάση της πρόβλεψης στην οποία γίνεται μία πρώτη εκτίμηση της κατάστασης, καθώς και της διασποράς του σφάλματος αυτής, τη χρονική στιγμή $\kappa+1$. Περιγράφεται από τις εξής δύο εξισώσεις:

$$\hat{x}_{\kappa+1|\kappa} = F_{\kappa}\hat{x}_{\kappa|\kappa} + G_{\kappa}u_{\kappa} \quad (2.1)$$

$$P_{\kappa+1|\kappa} = F_{\kappa}P_{\kappa|\kappa}F_{\kappa}^T + Q_{\kappa} \quad (2.2)$$

Η δεύτερη φάση είναι η αυτή της ενημέρωσης-διόρθωσης της αρχικής εκτίμησης, αφού πλέον γίνεται διαθέσιμη μία νέα μέτρηση της κατάστασης τη χρονική στιγμή $k+1$. Έτσι γίνεται μία νέα βελτιωμένη εκτίμηση και υπολογίζεται η διασπορά του σφάλματος αυτής. Η φάση αυτή περιγράφεται από τις παρακάτω τρεις εξισώσεις:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\{z_{k+1} - H_{k+1}\hat{x}_{k+1|k}\} \quad (2.3)$$

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T \quad (2.4)$$

$$K_{k+1} = P_{k+1|k}H_{k+1}^T[H_{k+1}P_{k+1|k}H_{k+1}^T + R_{k+1}]^{-1} \quad (2.5)$$

Οι εξισώσεις αυτές,, μαζί με τους πίνακες συνδιασποράς του σφάλματος της διαδικασίας και της μέτρησης και τις αρχικές συνθήκες, συνθέτουν τον αλγόριθμο του αναδρομικού φίλτρου Kalman για διακριτού χρόνου κατάσταση.

Το συμπέρασμα είναι ότι για τον υπολογισμό του εκτιμητή του φίλτρου Kalman λαμβάνονται υπόψη τα γνωστά δυναμικά χαρακτηριστικά του συστήματος, η ικανότητα του συστήματος να περιγράψει επ' ακριβώς τη διαδικασία, ότι είναι γνωστό γύρω από τα χαρακτηριστικά των σφαλμάτων των παρατηρήσεων και τις σχέσεις που συνδέουν τις μετρήσεις με τις μεταβλητές που περιγράφουν την κατάσταση του συστήματος. Το αποτέλεσμα των υπολογισμών είναι η εξαγωγή ενός βέλτιστου εκτιμητή της κατάστασης του συστήματος που συνοδεύεται από μία εκτίμηση του σφάλματος που περιέχει.

Κεφάλαιο 3

Το φίλτρο Kalman στην Μετεωρολογία

Κεφάλαιο 3

3.1 Εισαγωγή

Από το 1961 όταν ο R. E. Kalman δημοσίευσε την εργασία του με την οποία παρουσίασε τη λύση που πρότεινε ο ίδιος για την επίλυση του προβλήματος του Wiener στο πεδίο του χρόνου, το φίλτρο Kalman, όπως ονομάστηκε αυτή η λύση, χρησιμοποιήθηκε σε πολλές εφαρμογές. Στη Μετεωρολογία από τις πρώτες προσπάθειες εφαρμογής του φίλτρου Kalman είναι αυτή του Pearsson το 1991 και του Simonsen το ίδιο έτος (Homleid, 2004). Στόχος τους ήταν η βελτίωση της εκτίμησης της πραγματικής τιμής της θερμοκρασίας που γινόταν από ένα αριθμητικό μοντέλο πρόγνωσης καιρού.

Η εκτίμηση της πραγματικής τιμής της θερμοκρασίας του ατμοσφαιρικού αέρα δεν ήταν η μόνη εφαρμογή στην οποία χρησιμοποιήθηκε το φίλτρο Kalman στον τομέα της Μετεωρολογίας. Ένα άλλο πρόβλημα παρόμοιο με αυτό της θερμοκρασίας είναι η εκτίμηση της πραγματικής ταχύτητας του ανέμου (Louka, 2008), της σχετικής υγρασίας του αέρα (Anadranistakis, 2004), αλλά και της ολικής ηλιακής ακτινοβολίας (Jingying Fu, 2013). Έτσι, με παρόμοιο τρόπο εφαρμόστηκε αυτή η μέθοδος για τη βελτίωση των προγνώσεων αυτών των μεταβλητών. Το ίδιο έγινε και για την εκτίμηση της ποσότητας όζοντος στα κατώτερα στρώματα της ατμόσφαιρας (ατμοσφαιρικό οριακό στρώμα) (DelleMonache, 2006), (DelleMonache, 2006). Μια άλλη εφαρμογή του φίλτρου είναι στη διαδικασία αφομοίωσης δεδομένων για τη ομαλή λειτουργία ενός αριθμητικού μοντέλου πρόγνωσης καιρού (Houtekamer, 1998).

3.2 Οι εφαρμογές της μεθόδου, πλεονεκτήματα και μειονεκτήματα

Ο Pearsson (1991) επιχείρησε με επιτυχία να βελτιώσει τις προγνώσεις της θερμοκρασίας του αέρα κοντά στην επιφάνεια του εδάφους στην πόλη Lulea, που βρίσκεται στα βόρεια τμήματα της Σουηδίας, χρησιμοποιώντας παρατηρήσεις για μία περίοδο δύο μηνών (Νοέμβριος -

Δεκέμβριος 1988), κατά τη διάρκεια της οποίας παρατηρήθηκαν ασυνήθιστα υψηλές για την εποχή θερμοκρασίες. Η χρήση του φίλτρου μείωσε τα λάθη της πρόγνωσης κατά 0.5°C με 0.8°C και πέτυχε μία πιο καλή προσομοίωση της πραγματικότητας, δίχως την ανάγκη πολύ μεγάλης βάσης δεδομένων από το παρελθόν.

Σε αυτή τη λογική δηλαδή της χρήσης μικρότερης βάσης δεδομένων, η οποία δεν θα καταστρέφεται από αναγκαίες αλλαγές στη δομή του αριθμητικού μοντέλου πρόγνωσης καιρού, όπως παρατηρούνταν με τη χρήση της μεθόδου Model Output Statistics – MOS, πειραματίστηκε ο C. Simonsen (1991) του Μετεωρολογικού Ινστιτούτου της Δανίας χρησιμοποιώντας το φίλτρο Kalman. Χρησιμοποίησε δεδομένα θερμοκρασίας στα 2m πάνω από το έδαφος για την περίοδο από τον Δεκέμβριο 1990 ως τον Ιανουάριο 1991 από το Υψηλής Ανάλυσης Περιορισμένης Περιοχής Μοντέλο (Denmark – High Resolution Limited Area Model - DK-HIRLAM) και δεδομένα από παρατηρήσεις για την πόλη De Bilt της Ολλανδίας. Αποδείχθηκε ότι το μοντέλο το οποίο είχε αρχίσει να λειτουργεί από την 1/12/1990 υπερεκτιμούσε την πραγματική τιμή της θερμοκρασίας και η απόλυτη τιμή του σφάλματός του ήταν ίση με το 70% του σφάλματος μιας πρόγνωσης, η οποία ορίστηκε να είναι η μέση τιμή της πρόγνωσης του μοντέλου για την περίοδο μελέτης. Αντίθετα, μετά τη εφαρμογή του φίλτρου, το σφάλμα αυτό μειώθηκε στο 50% περίπου. Έτσι δόθηκε μια πολύ ικανοποιητική λύση στις περιπτώσεις εκείνες όπου δεν είναι διαθέσιμη μία μεγάλη σειρά δεδομένων για την εφαρμογή μεθόδων που απαιτούν κάτι τέτοιο όπως η MOS.

Μία άλλη χαρακτηριστική εργασία στην οποία φαίνεται η δυναμική της μεθόδου είναι αυτή των Kotroni and Lagouvardos (2004) από το Εθνικό Αστεροσκοπείο Αθηνών. Σε αυτή όπως και στις προηγούμενες γίνεται προσπάθεια βελτίωσης της εκτίμησης της πραγματικής τιμής της θερμοκρασίας του αέρα. Για το σκοπό αυτό χρησιμοποιήθηκε το Πέμπτης γενιάς Μέσης κλίμακας Μοντέλο (Mesoscale Model 5 - MM5) στην ευρύτερη περιοχή των Αθηνών και έγιναν διαθέσιμες παρατηρήσεις από τέσσερις μετεωρολογικούς σταθμούς που βρίσκονταν στο Λεκανοπέδιο της Αττικής. Η περίοδος μελέτης ήταν το έτος 2002. Χρησιμοποιήθηκαν δύο διαφορετικά πλέγματα της περιοχής ολοκλήρωσης, ένα με χαμηλή χωρική ανάλυση (8Km) και ένα άλλο με υψηλότερη χωρική ανάλυση (2Km). Ο κύριος στόχος του φίλτρου ήταν η ελαχιστοποίηση του σφάλματος του μοντέλου και από τη στατιστική ανάλυση φάνηκε ότι μετά την εφαρμογή του αυτό σχεδόν μηδενίστηκε σε όλες τις περιπτώσεις (0.02-0.13). Από την άλλη,

τα αποτελέσματα για την απόλυτη τιμή του σφάλματος ήταν εντυπωσιακά, καθώς η βελτίωση της τιμής αυτής με τη χρήση υψηλότερης χωρικής ανάλυσης ήταν μικρότερη από τη χρήση του φίλτρου στις τιμές θερμοκρασίας της περιοχής με τη χαμηλότερη χωρική ανάλυση. Για παράδειγμα, για το σταθμό που βρίσκεται στο κέντρο της Αθήνας (Θησείο), η απόλυτη τιμή του σφάλματος της πρόγνωσης της θερμοκρασίας για την 42η προγνωστική ώρα, χρησιμοποιώντας τη χαμηλή ανάλυση, ήταν 2.96°C , ενώ μετά τη χρήση του φίλτρου αυτή περιοριζόταν στους 1.19°C . Από την άλλη, με την υψηλότερη χωρική ανάλυση η απόλυτη τιμή του σφάλματος ήταν 2.08°C και μετά τη χρήση του φίλτρου περιοριζόταν στους 1.14°C . Όμοια ήταν τα αποτελέσματα για τους άλλους σταθμούς. Έτσι σε αυτόν του Ελληνικού η απόλυτη τιμή του σφάλματος της θερμοκρασίας για την 36η προγνωστική ώρα ήταν 2.7°C για την χαμηλή ανάλυση και 3.0°C για την υψηλή, ενώ μετά τη χρήση του φίλτρου αυτές οι τιμές γίνονταν 1.32°C και 1.31°C , αντίστοιχα. Φαίνεται λοιπόν ότι το όποιο συστηματικό λάθος παρουσίαζε η πρόγνωση και θα μπορούσε να οφείλεται σε λόγους όπως η κακή προσομοίωση της τοπογραφίας, αφαιρούνταν πιο εύκολα με το φίλτρο παρά με την αύξηση της χωρικής ανάλυσης.

Ο Crochet (2004) χρησιμοποίησε προγνώσεις από το μοντέλο του Ευρωπαϊκού Κέντρου Μεσοπρόθεσμων Προγνώσεων για το δίμηνο Ιανουαρίου-Φεβρουαρίου 2001 και παρατηρήσεις από 11 μετεωρολογικούς σταθμούς της Ισλανδίας. Εφάρμοσε το φίλτρο Kalman σε προγνώσεις θερμοκρασίας 2m και ταχύτητες ανέμου και κατέληξε στο συμπέρασμα ότι, ενώ η μέση τιμή του σφάλματος της πρόγνωσης μηδενίζεται, εντούτοις όλες οι νέες τιμές δεν είναι καλύτερες από αυτές του μοντέλου. Αντίθετα, υπάρχουν φορές που το φίλτρο βελτίωσε τις τιμές του μοντέλου, αλλά και άλλες στις οποίες αύξησε το σφάλμα. Γενικά, όπου το σφάλμα είχε ένα συστηματικό χαρακτήρα, παρατηρούνταν καλύτερα αποτελέσματα και σημαντική μείωση της ρίζας του μέσου τετραγωνικού σφάλματος. Σε άλλες περιπτώσεις, όπου το σφάλμα ήταν τυχαίο, η βελτίωση μέσω του φίλτρου ήταν περιορισμένη.

Οι Louka et al. (2008) χρησιμοποίησαν δύο περιορισμένης περιοχής μοντέλα πρόγνωσης καιρού το Skiron/Eta και το Περιχικό Ατμοσφαιρικό Σύστημα (Regional Atmospheric Modelling System - RAMs v4.3.0) και δεδομένα από ένα αιολικό πάρκο της Κρήτης για να αξιολογήσουν την απόδοση του φίλτρου Kalman στην πρόγνωση της ταχύτητας του ανέμου. Έδειξαν ότι δεν είναι αναγκαία η χρήση ακριβών υπολογιστικά συστημάτων τα οποία θα

παράγουν υψηλής χωρικής ανάλυσης προγνώσεις (<6Km) για εφαρμογές που σχετίζονται με την παραγωγή αιολικής ενέργειας. Αντίθετα, η πρόγνωση της ταχύτητας του ανέμου θα μπορούσε να πραγματοποιηθεί με μεγάλη ακρίβεια, χάρη στο συνδυασμό ενός μέτριας χωρικής ανάλυσης αριθμητικού μοντέλου και μιας στατιστικής μεθόδου, όπως το φίλτρο Kalman.

Μια εργασία στην οποία εκτός από τη βελτίωση της πρόγνωσης της θερμοκρασίας υπήρχε ανάγκη βελτίωσης και της πρόγνωσης της υγρασίας του αέρα για λόγους που σχετίζονται με την καλλιέργεια της γης ήταν αυτή των Anadranistakis et al. (2004). Σε αυτή χρησιμοποιήθηκε το Περιορισμένης Περιοχής Μοντέλο της Μπολόνια (Bologna Limited Area Model - BoLAM) με το οποίο πραγματοποιήθηκαν τριήμερες προγνώσεις για όλο το 2002. Πραγματικά δεδομένα χρησιμοποιήθηκαν από τον αυτόματο μετεωρολογικό σταθμό του Νευροκοπίου Δράμας, ο οποίος βρίσκεται σε μικρή απόσταση από το χωριό. Τα λάθη του μοντέλου αποδόθηκαν στην κακή αναπαράσταση του ανάγλυφου της περιοχής, η οποία είναι μια μικρή κοιλάδα περιτοιχισμένη από ψηλά βουνά. Η σημαντική υπερεκτίμηση της πραγματικής θερμοκρασίας που παρατηρήθηκε κατά τη διάρκεια του χειμώνα, σχετίστηκε με το γεγονός ότι δεν λαμβάνονταν υπόψη η χιονοκάλυψη της περιοχής από το μοντέλο. Η μείωση της μέσης τιμής της απόλυτης τιμής του σφάλματος της θερμοκρασίας κατά 2.0°C περίπου και της υγρασίας κατά 10% μετά την εφαρμογή του φίλτρου απέδειξε ότι τα σφάλματα του μοντέλου ήταν συστηματικά, καθώς η μείωση δεν παρατηρήθηκε μόνο κάποιες ώρες, αλλά σε ολόκληρο το 24ωρο και οδήγησε σε πιο επιτυχημένες προγνώσεις.

Το αδύνατο σημείο της μεθόδου φάνηκε με την εργασία του Homleid (1994). Οι προγνώσεις που πραγματοποιήθηκαν με το περιορισμένης περιοχής μοντέλο LAM50 στο Μετεωρολογικό Ινστιτούτο της Νορβηγίας για το έτος 1993 συνδυάστηκαν με παρατηρήσεις από ένα μεγάλο αριθμό μετεωρολογικών σταθμών της Νορβηγίας. Παρόλο που η μέση τιμή του σφάλματος της πρόγνωσης σχεδόν μηδενίστηκε, η τιμή της τυπικής απόκλισης παρέμεινε περίπου ίδια. Επίσης, σε κάποιες περιπτώσεις κατά τη διάρκεια του χειμώνα φάνηκε ότι δεν υπήρχε καμία βελτίωση της πρόγνωσης, αλλά αντίθετα παρατηρούνταν αύξηση του σφάλματος. Το γεγονός αυτό αποδόθηκε στις απότομες αλλαγές του καιρού την εποχή αυτή, που είχαν ως αποτέλεσμα τη διαφορετική συμπεριφορά του μοντέλου από μέρα σε μέρα. Έτσι υπήρχαν περιπτώσεις που η μέθοδος δημιουργούσε περισσότερα προβλήματα από αυτά που έλυσε.

Σε ότι αφορά την εκτίμηση της συγκέντρωσης όζοντος στα κατώτερα στρώματα της ατμόσφαιρας χαρακτηριστική είναι η εργασία των Delle Monache et al. (2006). Σε αυτή χρησιμοποιήθηκαν 8 διαφορετικά μοντέλα που παρήγαγαν εκτιμήσεις της συγκέντρωσης του όζοντος για μια περίοδο 50 ημερών κατά τη διάρκεια του καλοκαιριού του 2004. Η περιοχή μελέτης κάλυπτε το βορειοανατολικό τμήμα των Η.Π.Α και ένα μέρος της νότιας περιοχής του Καναδά. Μετρήσεις της πραγματικής τιμής αυτής της μεταβλητής προήλθαν από περίπου 360 σταθμούς. Αποδείχθηκε ότι οι προγνώσεις βελτιώνονταν σημαντικά μετά την εφαρμογή του φίλτρου. Βελτιώθηκε ακόμα η ικανότητα πρόβλεψης του ημερήσιου μέγιστου της συγκέντρωσης όζοντος, ενώ παράλληλα μειώθηκε η χρονική διαφορά ανάμεσα στο μέγιστο της πρόγνωσης και το πραγματικό.

Τέλος δύο εργασίες στις οποίες γίνεται εκτίμηση του σφάλματος της πρόγνωσης του μοντέλου σε τιμές θερμοκρασίας του αέρα και ταχύτητας του ανέμου είναι αυτές των Galanis & Anadranistakis (2002) και Galanis et al. (2006) των οποίων οι μέθοδοι που παρουσιάζονται εκεί υιοθετήθηκαν και εφαρμόζονται σε αυτή την εργασία.

Στην πρώτη, χρησιμοποιούνται ημερήσιες μέγιστες και ελάχιστες θερμοκρασίες από προγνώσεις του μοντέλου Skiron και παράλληλα πραγματικές μετρήσεις ημερήσιας μέγιστης και ελάχιστης θερμοκρασίας του αέρα από παρατηρήσεις τριών μετεωρολογικών σταθμών της Αθήνας, της Θεσσαλονίκης και του Ηρακλείου. Για τη μέγιστη ημερήσια θερμοκρασία χρησιμοποιήθηκε η τιμή του μοντέλου κατά την 36^η προγνωστική ώρα, καθώς αυτό αρχικοποιήθηκε στις 00 UTC και για την ελάχιστη θερμοκρασία η μικρότερη τιμή από τις τιμές του μοντέλου που αναφέρονται στην 27^η, 28^η και 29^η προγνωστική ώρα. Περίοδος μελέτης ήταν το εξάμηνο από τον Ιούλιο ως τον Δεκέμβριο του 2000. Παρατηρήθηκε μια συστηματική υπερεκτίμηση της μέγιστης θερμοκρασίας, η οποία αποδόθηκε στο γεγονός ότι σε κάποιους σταθμούς το πλησιέστερο σημείο βρισκόταν πάνω από τη θάλασσα και η μέση τιμή του σφάλματος ήταν από 2.2°C ως και 5.2 °C, ενώ από την άλλη η μέση τιμή του σφάλματος της ελάχιστης θερμοκρασίας ήταν αρκετά μικρότερη (από 0.7 °C ως 2.3 °C). Η εφαρμογή του φίλτρου οδήγησε στην ελάττωση του σφάλματος, η οποία ήταν μεγαλύτερη στις τιμές της μέγιστης θερμοκρασίας και μικρότερη σε αυτές της ελάχιστης. Η μορφή του φίλτρου που προτάθηκε ήταν αρκετά απλή και χωρίς να απαιτείται μεγάλο υπολογιστικό κόστος.

Στη δεύτερη εργασία χρησιμοποιείται ένα πολυώνυμο για να περιγράψει το σφάλμα του μοντέλου. Για αυτό το πολυώνυμο ελέγχεται ο βαθμός εκείνος που δίνει τα καλύτερα αποτελέσματα, καθώς ο βαθμός του πολυωνύμου είναι αυτός που ελέγχει τη διάσταση του φίλτρου που χρησιμοποιείται. Οι μεταβλητές που εξετάστηκαν ήταν η θερμοκρασία του αέρα σε απόσταση 2m από την επιφάνεια του εδάφους και η ταχύτητα του ανέμου σε ύψος 10m. Το αριθμητικό μοντέλο πρόγνωσης καιρού που χρησιμοποιήθηκε ήταν το Skiron/Eta, ενώ μετρήσεις των πραγματικών τιμών των μεταβλητών έγιναν διαθέσιμες από παρατηρήσεις, από δύο μετεωρολογικούς σταθμούς, ο ένας από τους οποίους βρίσκεται στην περιοχή Alaiz της Ισπανίας και ο άλλος στην περιοχή Ρόκας της Κρήτης. Αποδείχθηκε ότι την καλύτερη απόδοση, τόσο για τη θερμοκρασία του αέρα, όσο και για την ταχύτητα του ανέμου, παρουσίαζε το φίλτρο Kalman, όταν χρησιμοποιούνταν ένα πολυώνυμο με βαθμό 2 ή 3, καθώς τα αντίστοιχα πολυώνυμα που είχαν μεγαλύτερο βαθμό παρουσίαζαν αστάθειες, ενώ χρειάζονταν και πολύ περισσότερους υπολογισμούς. Για την εκτίμηση των πινάκων διασποράς των σφαλμάτων αποδείχθηκε ότι η βέλτιστη περίοδος παρατηρήσεων είναι αυτή των 7 ημερών.

Κεφάλαιο 4

Δεδομένα και Μεθοδολογία

Κεφάλαιο 4

4.1 Εισαγωγή

Στην εργασία αυτή εφαρμόζεται το φίλτρο Kalman σε προγνώσεις θερμοκρασίας τριών διαφορετικών αριθμητικών μοντέλων πρόγνωσης καιρού προκειμένου να εξεταστεί το μέγεθος της βελτίωσης που προσφέρει η εφαρμογή της μεθόδου αυτής. Για το σκοπό αυτό χρησιμοποιήθηκαν δύο διαφορετικές προσεγγίσεις. Στόχος κάθε μίας είναι η εξάλειψη του συστηματικού τμήματος του σφάλματος της πρόγνωσης και για το σκοπό αυτό χρησιμοποιούνται οι τιμές των σφαλμάτων της πρόγνωσης μια μικρής περιόδου ορισμένων ημερών. Η εφαρμογή του φίλτρου έγινε για την περίοδο, από τον Ιούνιο 2011 ως και τον Ιούλιο 2012. Για τη στατιστική μελέτη συμπεριλήφθηκαν μόνο οι 12 τελευταίοι μήνες (Αύγουστος 2011 – Ιούλιος 2012), ώστε να μην υπάρχει καμία εξάρτηση από τα αρχικά δεδομένα. Τα μοντέλα των οποίων οι προγνώσεις χρησιμοποιήθηκαν σε αυτή την εργασία είναι τα εξής: το Weather Research and Forecasting (WRF), το μοντέλο Skiron/Eta και το Global Forecasting System (GFS).

4.2 Δεδομένα

Για το διάστημα της περιόδου μελέτης έγιναν διαθέσιμα πραγματικά δεδομένα που αφορούσαν τιμές θερμοκρασίας σε απόσταση 2m από την επιφάνεια του εδάφους από το μετεωρολογικό σταθμό που βρίσκεται στο αεροδρόμιο της Θεσσαλονίκης. Ο σταθμός έχει συντεταγμένες 40.527° N, 22.971° E (Σχήμα 6) και βρίσκεται σε ύψος 4m από την μέση στάθμη θάλασσας. Για κάθε μία μέρα είναι διαθέσιμες 8 μετρήσεις θερμοκρασίας στις 00, 03, 06, 09, 12, 15, 18 και 21 UTC.

4.2.1 Το αριθμητικό μοντέλο Weather Research and Forecasting (WRF)

Το αριθμητικό μοντέλο πρόγνωσης καιρού Weather Research and Forecasting (WRF) αναπτύχθηκε κατόπιν συνεργασίας ανάμεσα σε περισσότερους από 150 οργανισμούς και ερευνητικά κέντρα των Η.Π.Α και όχι μόνο, ανάμεσα στα οποία πρωταρχικό ρόλο είχαν το Εθνικό Κέντρο Έρευνας της Ατμόσφαιρας των Η.Π.Α (National Center for Atmospheric Research - NCAR) και η Εθνική Υπηρεσία Ωκεανών και Ατμόσφαιρας των Η.Π.Α (National Oceanic and Atmospheric Administration - NOAA). Το μοντέλο, το οποίο είναι ανοιχτού κώδικα, λειτουργεί επιχειρησιακά σε καθημερινή βάση στον Τομέα Μετεωρολογίας και Κλιματολογίας του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης. Οι αρχικές και οριακές συνθήκες προέρχονται από τον προγνωστικό κύκλο 12UTC του Παγκόσμιου Προγνωστικού Συστήματος (Global Forecasting System - GFS) σε ανάλυση $0.5^{\circ} \times 0.5^{\circ}$ και χρονική διακριτοποίηση των αποτελεσμάτων 3 ώρες. Το πεδίο ολοκλήρωσης των προσομοιώσεων αποτελεί μία περιοχή με χαμηλή χωρική διακριτοποίηση 15×15 Km που περιλαμβάνει την Ευρώπη, ενώ υπάρχει και μια δεύτερη περιοχή που περιλαμβάνει την Ελλάδα με υψηλότερη χωρική διακριτοποίηση 5×5 Km. Τα κατακόρυφα επίπεδα είναι τριάντα εννιά (39), ο χρονικός ορίζοντας των προγνώσεων είναι εβδομήντα δύο (72) ώρες, ενώ ο δυναμικός πυρήνας που χρησιμοποιείται είναι ο Advanced Research WRF - ARW.

Τα σχήματα παραμετροποιήσεων με τα οποία προσομοιώνονται οι φυσικές διεργασίες είναι τα εξής. Για την μικροφυσική των νεφών χρησιμοποιείται το σχήμα Ferrier (new Eta) (Ferrier et al., 2002), ενώ για το ισοζύγιο των ακτινοβολιών χρησιμοποιείται το σχήμα ταχείας μεταφοράς ακτινοβολίας (Rapid Radiative Transfer Model - RRTMG) (Mlawer et al., 1997) και το χρονικό βήμα σε λεπτά στο οποίο υπολογίζονται οι νέες τιμές για τις μικρού και μεγάλου μήκους ακτινοβολίες είναι 15min. Η φυσική του επιφανειακού στρώματος παραμετροποιείται με το σχήμα Monin-Obukhov-Janjic (Monin and Obukhov, 1954), (Janjic, 1996, 2002), ενώ υπομοντέλο εδάφους είναι το ενοποιημένο NOAH (Chen and Dudhia, 2001). Για την τυρβώδη κινητική ενέργεια του πλανητικού οριακού στρώματος χρησιμοποιείται το σχήμα Mellor-Yamada-Janjic (Mellor and Yamada, 1982), (Janjic, 1990, 1996, 2002) και η τυρβώδης κινητική

ενέργεια υπολογίζεται σε κάθε χρονικό βήμα. Για την κατακόρυφη αστάθεια χρησιμοποιείται το σχήμα Betts-Miller-Janjic (Janjic, 1994, 2000), (Betts, 1986, Betts and Miller, 1986) και αυτές υπολογίζονται κάθε 5 λεπτά. Οι επιφανειακές ροές θερμότητας και υγρασίας, καθώς και η επίδραση των νεφών στο οπτικό βάθος για την ακτινοβολία συνυπολογίζονται, αν όμως υπάρχει στρώμα από χιόνι, οι επιδράσεις αυτού δεν συνυπολογίζονται.

Ο χρόνος αρχικοποίησης των προγνώσεων ήταν καθημερινά στις 12UTC και ο προγνωστικός ορίζοντας ήταν εβδομήντα δύο 72 ώρες. Τα δεδομένα αποθηκεύονται κάθε έξι ώρες (00, 06, 12 και 18 UTC) και αυτά που χρησιμοποιήθηκαν καλύπτουν την περίοδο από τον Ιούνιο του 2011 ως τον Ιούλιο του 2012 και περιγράφουν την θερμοκρασία σε απόσταση 2m από την επιφάνεια του εδάφους για το πλησιέστερο στο αεροδρόμιο της Θεσσαλονίκης σημείο του πλέγματος της περιοχής ολοκλήρωσης με την υψηλότερη χωρική διακριτοποίηση του οποίου οι συντεταγμένες είναι 40.49775314° N, 22.96032715° E (Σχήμα 6).

4.2.2 Το αριθμητικό μοντέλο πρόγνωσης καιρού Skiron/Eta

Το αριθμητικό μοντέλο πρόγνωσης καιρού Skiron/Eta αναπτύχθηκε στο Πανεπιστήμιο Αθηνών από την Ομάδα Ατμοσφαιρικών Μοντέλων και Πρόγνωσης Καιρού (Kallos et al., 1997) και είναι βασισμένο στο μοντέλο Eta/NCEP (Janjic, 1994). Είναι κατάλληλο για μέσης κλίμακας προσομοιώσεις σε περιοχές με ποικίλα φυσιογραφικά χαρακτηριστικά ενώ το γεγονός ότι το μοντέλο εκτός της υδροστατικής προσέγγισης περιλαμβάνει και την επιλογή να ολοκληρώνεται μη-υδροστατικά, εμφανίζεται να είναι πολύ αποτελεσματικό σε εφαρμογές αριθμητικής πρόγνωσης στις οποίες απαιτείται υψηλή χωρική ανάλυση των προσομοιώσεων, αφού τότε δίνει γενικά καλύτερες και ομαλότερες λύσεις από το υδροστατικό. Το πλέγμα που χρησιμοποιείται για την αναπαράσταση των μεταβλητών είναι το πλέγμα-E του Arakawa ενώ ως κατακόρυφη μεταβλητή χρησιμοποιείται η *eta*.

Κατά την καθημερινή επιχειρησιακή λειτουργία του μοντέλου στον Τομέα Μετεωρολογίας και Κλιματολογίας του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης παράγονται προγνώσεις για την περιοχή της Κεντρικής και Βόρειας Ελλάδας με οριζόντια

χωρική διακριτοποίηση $0.05^{\circ} \times 0.05^{\circ}$ και 38 κατακόρυφα επίπεδα που φτάνουν μέχρι την επιφάνεια των 25hPa, για μία περίοδο τριών ημερών (Pytharoulis, 2009). Οι αρχικές και οριακές συνθήκες προέρχονται και αυτές από τον προγνωστικό κύκλο 12UTC του Παγκόσμιου Προγνωστικού Συστήματος (Global Forecasting System - GFS) σε ανάλυση $0.5^{\circ} \times 0.5^{\circ}$ και χρονική διακριτοποίηση 3 ωρών.

Για τις φυσικές διεργασίες που δεν είναι δυνατόν να αναπαρασταθούν, παρά την υψηλή ανάλυση του μοντέλου χρησιμοποιούνται τα παρακάτω σχήματα παραμετροποίησης. Για την παραμετροποίηση του ατμοσφαιρικού οριακού στρώματος χρησιμοποιείται το σχήμα των Mellor and Yamada (1982), ενώ ένα δεύτερης τάξης σχήμα των Mellor και Yamada χρησιμοποιείται για την επιφάνεια του εδάφους. Ένα σχήμα που περιγράφει το ιξώδες υπόστρωμα χρησιμοποιείται πάνω από το έδαφος και τις υδάτινες επιφάνειες, προκειμένου να βελτιωθεί ο υπολογισμός των ροών της επιφάνειας, ενώ η θερμοκρασία και η υγρασία του εδάφους υπολογίζονται σε έξι στρώματα που εκτείνονται από την επιφάνεια μέχρι και 2.55m κάτω από αυτή. Η προσομοίωση των διεργασιών σχηματισμού νεφών και υετού βασίζεται στο σύστημα των Zhao and Carr (1997), με το οποίο γίνεται ο υπολογισμός δύο τύπων υετού, της βροχόπτωσης και του χιονιού, ενώ ο σχηματισμός νεφών και υετού λόγω ανωμεταφοράς γίνεται με το σχήμα Betts-Miller-Janjic. Τέλος, οι ροές ακτινοβολίας υπολογίζονται με το σχήμα ακτινοβολίας Geophysical Fluid Dynamics Laboratory – GFDL (Fels and Schwarzkopf, 1975, Schwarzkopf and Fels 1991).

Ο χρόνος αρχικοποίησης των προγνώσεων ήταν καθημερινά στις 12UTC και ο προγνωστικός ορίζοντας ήταν εβδομήντα δύο 72 ώρες. Τα δεδομένα αποθηκεύονται κάθε τρεις ώρες (00, 03, 06, 09, 12, 15, 18 και 21 UTC) και αυτά που χρησιμοποιήθηκαν καλύπτουν την περίοδο από τον Ιούνιο του 2011 ως τον Ιούλιο του 2012 και περιγράφουν τη θερμοκρασία σε απόσταση 2m από την επιφάνεια του εδάφους για το πλησιέστερο στο αεροδρόμιο της Θεσσαλονίκης σημείο του πλέγματος της περιοχής ολοκλήρωσης του οποίου οι συντεταγμένες είναι 40.54598° N, 23.01362° E (Σχήμα 6).

4.2.3 Το αριθμητικό μοντέλο πρόγνωσης καιρού Global Forecasting System

Το Παγκόσμιο Προγνωστικό Σύστημα (Global Forecasting System - GFS) είναι ένα παγκόσμιο σύστημα αριθμητικής πρόγνωσης καιρού το οποίο διευθύνεται από την Εθνική Υπηρεσία Ωκεανών και Ατμόσφαιρας των Η.Π.Α (National Oceanic and Atmospheric Administration - NOAA). Αυτό το μαθηματικό μοντέλο λειτουργεί τέσσερις φορές την ημέρα κατά τις συνοπτικές ώρες 00, 06, 12 και 18 UTC και παράγει προβλέψεις έως και 16 ημέρες, αλλά με μειωμένη χωρική και χρονική ανάλυση με την πάροδο του χρόνου. Είναι ευρέως αποδεκτό ότι πέραν των 7 ημερών η πρόβλεψη είναι πολύ γενική και δεν είναι πολύ ακριβής. Μαζί με το Προγνωστικό Σύστημα του Ευρωπαϊκού Κέντρου Μεσοπρόθεσμων Προγνώσεων (ECMWF), το μοντέλο της Καναδικής Μετεωρολογικής Υπηρεσίας Global Environmental Multiscale (GEM), των οποίων οι προγνώσεις είναι δεκαήμερες και το μοντέλο του Ναυτικού Εργαστηρίου Έρευνας των Η.Π.Α Navy Global Environmental Model (NAVGEN) (<http://www.nrlmry.navy.mil/metoc/nogaps/>), το οποίο δίνει πρόγνωση του καιρού για οκτώ ημέρες, είναι ένα από τα τέσσερα υψηλής ανάλυσης, παγκόσμια μοντέλα γενικής χρήσης.

Το μοντέλο βασίζεται στις συνήθεις εξισώσεις της διατήρησης της μάζας, ορμής, ενέργειας, υγρασίας (<http://www.emc.ncep.noaa.gov>). Η χωρική ανάλυση του μοντέλου είναι της τάξης των 27Km περίπου (T574) και διαθέτει 64 κατακόρυφα επίπεδα σε ένα υβριδικό σύστημα στο οποίο η κατακόρυφη συντεταγμένη χαμηλά στην ατμόσφαιρα είναι η σίγμα ενώ ψηλότερα χρησιμοποιείται η πίεση. Τα δεδομένα δίνονται με ανάλυση $0.5^{\circ} \times 0.5^{\circ}$ και χρονικό βήμα 3 ωρών. Από αυτά χρησιμοποιήθηκαν στην εργασία οι τιμές θερμοκρασίας του αέρα σε απόσταση 2m από την επιφάνεια του εδάφους για το πλησιέστερο στο αεροδρόμιο της Θεσσαλονίκης σημείο το οποίο έχει συντεταγμένες 40.5° N, 23.0° E (Σχήμα 6) για την περίοδο από τον Ιούνιο 2011 ως τον Ιούλιο 2012.



Σχήμα 6: Χάρτης με τα σημεία των τριών μοντέλων και του μετεωρολογικού σταθμού.

4.3 Μεθοδολογία

Αφού λοιπόν συγκεντρώθηκαν οι τιμές θερμοκρασίας στα 2m πάνω από την επιφάνεια του εδάφους, τόσο από το μετεωρολογικό σταθμό του αεροδρομίου της Θεσσαλονίκης, όσο και από το πλησιέστερο στο αεροδρόμιο σημείο του πλέγματος της περιοχής ολοκλήρωσης των τριών μοντέλων, υπολογίστηκαν αρχικά τα σφάλματα των μοντέλων. Η στατιστική μελέτη των σφαλμάτων έγινε με τη βοήθεια ορισμένων στατιστικών μεγεθών όπως: η μέση τιμή του σφάλματος και της απόλυτης τιμής του σφάλματος, καθώς και η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (Wilks, 1995). Αυτά τα μεγέθη υπολογίζονται με τη χρήση των παρακάτω τύπων

$$Bias = \frac{1}{N} \cdot \sum_{i=1}^N (M_i - O_i)$$

$$Absolute\ Bias = \frac{1}{N} \cdot \sum_{i=1}^N |M_i - O_i|$$

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (M_i - O_i)^2}$$

όπου M_i είναι η τιμή της θερμοκρασίας από το αριθμητικό μοντέλο πρόγνωσης και O_i η τιμή που παρατηρήθηκε κατά την ίδια ακριβώς ώρα στο μετεωρολογικό σταθμό. Τα μεγέθη αυτά υπολογίστηκαν αρχικά για κάθε προγνωστική ώρα ξεχωριστά, τόσο για την περίοδο των 12 μηνών (Αύγουστος 2011 - Ιούλιος 2012), όσο και για κάθε μήνα ξεχωριστά και έπειτα για κάθε 24ωρο πρόγνωσης, όπου και σε αυτή την περίπτωση έγιναν υπολογισμοί, τόσο για την περίοδο όλου του έτους, όσο και για κάθε μήνα.

Χρησιμοποιήθηκαν δύο παραπλήσιες μορφές του φίλτρου Kalman. Τόσο στη μία όσο και στην άλλη μορφή, το φίλτρο εφαρμόζεται σε κάθε προγνωστική ώρα ξεχωριστά. Η διαφορά που υπάρχει μεταξύ τους αφορά στο πως διαχειρίζονται το σφάλμα του μοντέλου. Στην πρώτη περίπτωση χρησιμοποιήθηκε μόνο μία παράμετρος πράγμα που οδήγησε σε μια αρκετά απλή μορφή του φίλτρου (Galanis & Anadranistakis, 2002). Το γεγονός αυτό προσφέρει μία σχετικά εύκολη και υπολογιστικά φθηνή εκτίμηση του επόμενου σφάλματος του μοντέλου. Η μονοδιάστατη αυτή μέθοδος θα αναφέρεται στο εξής ως μέθοδος 1. Στη μέθοδο 1 λοιπόν οι εξισώσεις συστήματος και μέτρησης έχουν την παρακάτω μορφή:

$$x_k = x_{k-1} + w_k$$

$$z_k = x_k + v_k = M_k + O_k$$

Σε αυτούς τους τύπους έχει οριστεί σαν Z_k η διαφορά ανάμεσα στην πραγματική τιμή της θερμοκρασίας του αέρα (O_k) και αυτή του μοντέλου (M_k) σε μια συγκεκριμένη χρονική στιγμή. Το συστηματικό μέρος αυτής της διαφοράς συμβολίζεται με x_k . Στον πίνακα μετάβασης έχει δοθεί η τιμή $F=1$, ενώ ίδια τιμή δόθηκε και στον πίνακα H , που είναι ο πίνακας παρατήρησης. Οι μεταβλητές w_k και v_k εκφράζουν την τιμή της μεταβολής του συστηματικού μέρους του

σφάλματος της πρόγνωσης του μοντέλου από τη μία χρονική στιγμή στην επόμενη και του τυχαίου μέρους του σφάλματος της πρόγνωσης του μοντέλου, αντίστοιχα.

Με βάση τα παραπάνω, οι εξισώσεις του φίλτρου Kalman για τη μέθοδο 1 διαμορφώνονται ως εξής:

Εξισώσεις πρόβλεψης:

$$x_{k+1|k} = x_{k|k} \quad (4.1)$$

$$P_{k+1|k} = P_{k|k} + W_{k+1} \quad (4.2)$$

Εξισώσεις διόρθωσης:

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1} \cdot (z_k - x_{k+1|k}) \quad (4.3)$$

$$P_{k+1|k+1} = (1 - K_{k+1}) \cdot P_{k+1|k} \quad (4.4)$$

όπου

$$K_{k+1} = \frac{P_{k+1|k}}{P_{k+1|k} + R_{k+1}} \quad (4.5)$$

Η σχέση (4.4) προκύπτει από την (2.4) λαμβάνοντας υπόψη ότι $H=1$ και την (4.5) άρα θα είναι:

$$\begin{aligned} P_{k+1|k+1} &= (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T \\ &= (1 - K_{k+1})^2 P_{k+1|k} + K_{k+1}^2 R_{k+1} \\ &= (1 - K_{k+1})^2 P_{k+1|k} + K_{k+1}^2 \left(\frac{P_{k+1|k}}{K_{k+1}} - P_{k+1|k} \right) \\ &= P_{k+1|k} - 2K_{k+1}P_{k+1|k} + K_{k+1}^2 P_{k+1|k} + K_{k+1}P_{k+1|k} - K_{k+1}^2 P_{k+1|k} \\ &= P_{k+1|k} - K_{k+1}P_{k+1|k} = (1 - K_{k+1}) \cdot P_{k+1|k} \end{aligned}$$

Στις παραπάνω εξισώσεις, τα W_{k+1} και V_{k+1} είναι οι πίνακες διασπορών των w_k και v_k , αντίστοιχα. Καθώς το φίλτρο έχει διάσταση ένα (1) οι πίνακες αυτοί είναι ουσιαστικά αριθμοί (1x1) και η τιμή τους δίνεται από τους τύπους:

$$W_{\kappa+1} = \frac{1}{N-1} \cdot \sum_{i=0}^{N-1} (w_{\kappa-i} - \frac{1}{N} \cdot \sum_{i=0}^{N-1} w_{\kappa-i})^2$$

$$R_{\kappa+1} = \frac{1}{N-1} \cdot \sum_{i=0}^{N-1} (v_{\kappa-i} - \frac{1}{N} \cdot \sum_{i=0}^{N-1} v_{\kappa-i})^2$$

όπου N είναι ο αριθμός των ημερών που λαμβάνονται υπόψη για τον υπολογισμό του συστηματικού μέρους του σφάλματος της πρόγνωσης τη συγκεκριμένη προγνωστική ώρα. Δοκιμάστηκαν πέντε (5) διαφορετικές τιμές για την μεταβλητή N . Αρχικά η εκτίμηση του σφάλματος της πρόγνωσης βασίστηκε στα σφάλματα του μοντέλου τις τέσσερις (4) προηγούμενες ημέρες, ενώ αμέσως μετά χρησιμοποιήθηκαν τα σφάλματα των επτά (7), δέκα (10), δεκαπέντε (15) και τριάντα (30) προηγούμενων ημερών.

Σαν αρχικές τιμές, στις μεταβλητές $x_{0|0}$ και $P_{0|0}$ δόθηκαν οι τιμές 0 και 4, αντίστοιχα, ωστόσο αυτό έχει μικρή σημασία καθώς από τη μία η τιμή της $P_{0|0}$ δείχνει ότι δεν είναι αρκετά ακριβής η πρώτη εκτίμηση του σφάλματος $x_{0|0}$, ενώ από την άλλη η στατιστική μελέτη της απόδοσης του φίλτρου ξεκινά δύο μήνες μετά την αρχικοποίηση. Έτσι επιτυγχάνεται μηδενική εξάρτηση των αποτελεσμάτων από τις αρχικές τιμές.

Μετά την ολοκλήρωση των υπολογισμών η τιμή $x_{\kappa+1|\kappa+1}$ που προκύπτει αποτελεί τη βέλτιστη εκτίμηση του συστηματικού μέρους του σφάλματος της επόμενης πρόγνωσης. Αυτή η εκτίμηση συνοδεύεται από ένα σφάλμα του οποίου η εκτίμηση της διασποράς περιγράφεται από την τιμή της μεταβλητής $P_{\kappa+1|\kappa+1}$. Η τιμή $x_{\kappa+1|\kappa+1}$ επομένως αφαιρείται από την επόμενη έξοδο του μοντέλου που αναφέρεται στη συγκεκριμένη προγνωστική ώρα, καθώς ισχύει

$$z_{\kappa+1} = M_{\kappa+1} - O_{\kappa+1} = x_{\kappa+1} + v_{\kappa+1}$$

και αφού η διαφορά ανάμεσα στην πρόγνωση και την πραγματικότητα προσεγγίζεται με την τιμή $x_{\kappa+1|\kappa+1}$ θα είναι

$$M_{\kappa+1} - O_{\kappa+1} \approx x_{\kappa+1|\kappa+1}$$

δηλαδή

$$O_{k+1} \approx M_{k+1} - x_{k+1|k+1}$$

που σημαίνει ότι η πραγματική τιμή προσεγγίζεται, αν από την τιμή του μοντέλου αφαιρεθεί η εκτίμηση του σφάλματος.

Στη δεύτερη περίπτωση χρησιμοποιήθηκε ένα λίγο διαφορετικό σκεπτικό. Όπως στην προηγούμενη περίπτωση, έτσι και σε αυτή, το φίλτρο εφαρμόζεται σε κάθε προγνωστική ώρα ξεχωριστά. Η υπόθεση που γίνεται αυτή τη φορά είναι ότι το σφάλμα της πρόγνωσης σε μία συγκεκριμένη ημέρα μπορεί να περιγραφεί ως ένα πολυώνυμο στο οποίο η μεταβλητή είναι η έξοδος του μοντέλου, ενώ οι συντελεστές αυτού είναι οι αριθμοί $a_{l,k}$ οι οποίοι μέσα από την εφαρμογή του φίλτρου πρέπει να υπολογιστούν. Θεωρείται ότι η παραπάνω διαδικασία δεν περιγράφει επ' ακριβώς το σφάλμα της πρόγνωσης, αλλά περιέχει ένα μη συστηματικό λάθος (Galanis et al., 2006). Δηλαδή:

$$z_k = a_{0,k} + a_{1,k} \cdot M_k + a_{2,k} \cdot M_k^2 + a_{3,k} \cdot M_k^3 + \dots + v_k$$

όπου M_k η τιμή της θερμοκρασίας του αέρα από το μοντέλο.

Αυτή η σχέση είναι η εξίσωση παρατήρησης του φίλτρου. Η εξίσωση συστήματος είναι:

$$x_{k+1} = x_k + w_k$$

Με βάση τα παραπάνω προκύπτει ότι ο πίνακας παρατήρησης θα είναι

$$H_k = [1 \ M_k \ M_k^2 \ \dots]$$

ενώ το διάνυσμα x_k θα είναι

$$x_k = [a_{0,k} \ a_{1,k} \ a_{2,k} \ \dots]^T$$

και ο πίνακας μετάβασης F θα είναι ο μοναδιαίος πίνακας.

Οι εξισώσεις πρόβλεψης γίνονται:

$$x_{k+1|k} = x_{k|k} \quad (4.6)$$

$$P_{k+1|k} = P_{k|k} + W_{k+1} \quad (4.7)$$

και οι εξισώσεις διόρθωσης:

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}\{z_k - H_{k+1}x_{k+1|k}\} \quad (4.8)$$

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T \quad (4.9)$$

$$K_{k+1} = P_{k+1|k}H_{k+1}^T [H_{k+1}P_{k+1|k}H_{k+1}^T + R_{k+1}]^{-1} \quad (4.10)$$

Στις παραπάνω εξισώσεις τα W_{k+1} και V_{k+1} είναι οι πίνακες διασπορών των w_k και v_k αντίστοιχα και οι τιμές τους δίνονται από τους τύπους:

$$W_{k+1} = \frac{1}{N-1} \cdot \sum_{i=0}^{N-1} (w_{k-i} - \frac{1}{N} \cdot \sum_{i=0}^{N-1} w_{k-i})^2$$

$$R_{k+1} = \frac{1}{N-1} \cdot \sum_{i=0}^{N-1} (v_{k-i} - \frac{1}{N} \cdot \sum_{i=0}^{N-1} v_{k-i})^2$$

όπου N είναι αριθμός των ημερών που λαμβάνονται υπόψη για τον υπολογισμό αυτών.

Σαν αρχικές τιμές στους πίνακες $x_{0|0}$ και $P_{0|0}$ δόθηκαν οι τιμές:

$$x_{0|0} = [0 \ 0 \ 0 \ \dots]^T$$

$$P_{0|0} = \begin{pmatrix} 4 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 4 \end{pmatrix}$$

οι οποίες υποδεικνύουν ότι δεν είναι αρκετά ακριβής η πρώτη εκτίμηση του σφάλματος $x_{0|0}$ και οδηγούν σε γρήγορη προσαρμογή στα πραγματικά δεδομένα. Παράλληλα στους πίνακες W_k και R_k δόθηκαν οι τιμές

$$W_{\kappa} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

$$R_{\kappa} = 6$$

Με αυτές τις τιμές έγινε η εκτίμηση του επόμενου σφάλματος του μοντέλου.

Στην εργασία αυτή χρησιμοποιήθηκαν ένα μηδενικού βαθμού, ένα πρωτοβάθμιο και ένα δευτεροβάθμιο πολυώνυμο. Στην πρώτη περίπτωση η μέθοδος ονομάστηκε μέθοδος 2α, στην δεύτερη μέθοδος 2β και στην τρίτη μέθοδος 2γ. Και για τις τρεις αυτές εφαρμογές για τον υπολογισμό των $W_{\kappa+1}$ και $R_{\kappa+1}$ δοκιμάστηκαν πέντε (5) διαφορετικές τιμές για τη μεταβλητή N. Αρχικά η εκτίμηση του σφάλματος της πρόγνωσης βασίστηκε στα σφάλματα του μοντέλου τις τέσσερις (4) προηγούμενες ημέρες, ενώ αμέσως μετά χρησιμοποιήθηκαν τα σφάλματα των επτά (7), δέκα (10), δεκαπέντε (15) και τριάντα (30) προηγούμενων ημερών.

Κεφάλαιο 5

Αποτελέσματα

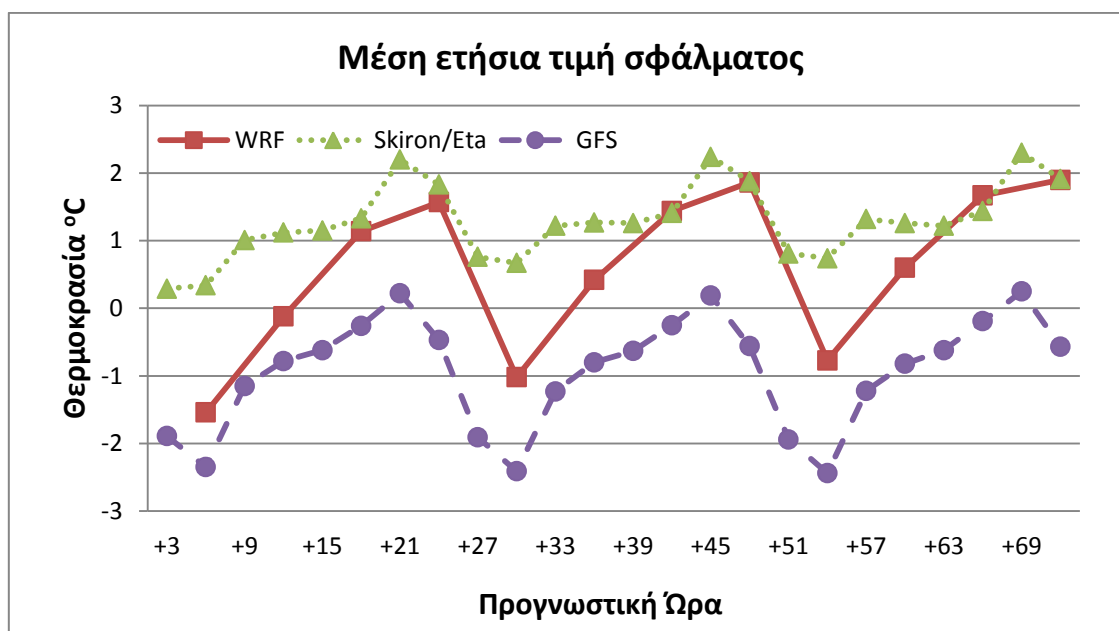
Κεφάλαιο 5

5.1 Ανάλυση των σφαλμάτων της πρόγνωσης

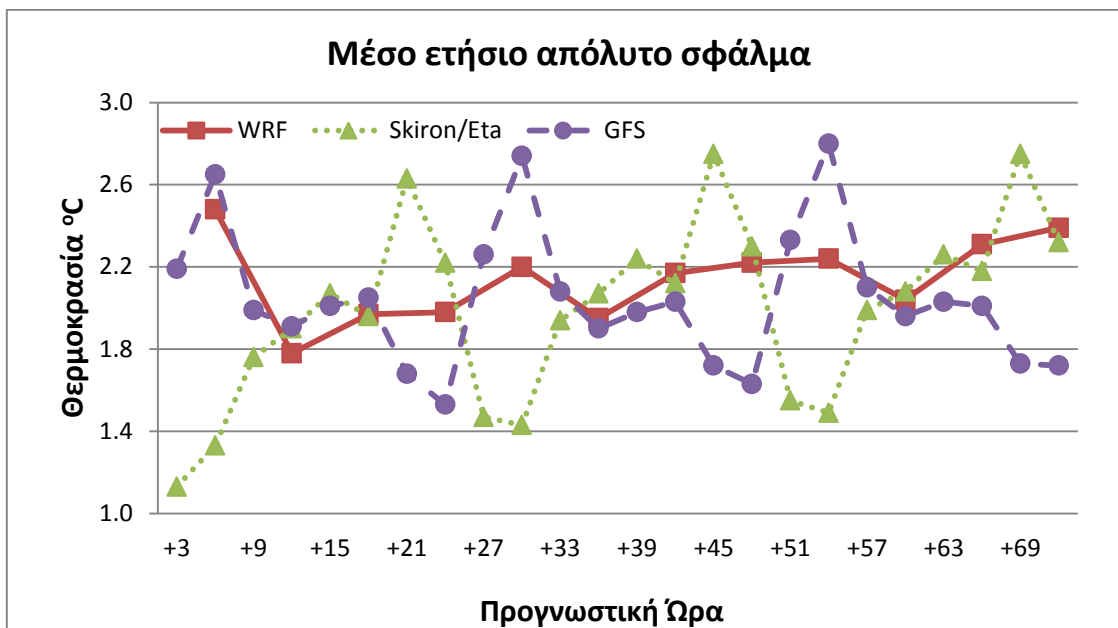
Όπως αναφέρθηκε και παραπάνω, ενώ το φίλτρο Kalman εφαρμόστηκε στις προγνώσεις της περιόδου Ιουνίου 2011 ως και τον Ιούλιο 2012, εντούτοις τα αποτελέσματα παρουσιάζονται μόνο για το τελευταίο δωδεκάμηνο (Αύγουστος 2011 – Ιούλιος 2012), έτσι ώστε να μην επηρεάζονται από τις αρχικές τιμές που δόθηκαν. Από τη στατιστική μελέτη των σφαλμάτων των τριών μοντέλων σε αυτό το δωδεκάμηνο πριν την εφαρμογή του φίλτρου Kalman παρατηρήθηκαν τα εξής. Η μέση τιμή του σφάλματος για το μοντέλο WRF ήταν 0.6°C , για το GFS -0.94°C και για το Skiron/Eta 1.29°C . Τη μικρότερη ωστόσο μέση τιμή του απόλυτου σφάλματος παρουσίασε το αριθμητικό μοντέλο Skiron/Eta με τιμή 2.0°C , η μέση τιμή του απόλυτου σφάλματος για το μοντέλο GFS ήταν 2.04°C , ενώ για το WRF ήταν 2.14°C . Από την άλλη, η ρίζα του μέσου τετραγωνικού σφάλματος είχε μεγαλύτερη τιμή για το WRF και ήταν ίση με 2.68°C , για το GFS ήταν 2.48°C και τέλος για το Skiron/Eta 2.46°C . Φαίνεται λοιπόν ότι κατά μέσο όρο και τα τρία μοντέλα έχουν περίπου ίδια τιμή απόλυτου σφάλματος, κοντά στους 2.0°C , αλλά η συμπεριφορά τους είναι διαφορετική, καθώς το GFS παρουσιάζει μια μικρή υποεκτίμηση της πραγματικής θερμοκρασίας του αέρα, το WRF μια μικρή υπερεκτίμηση, ενώ το Skiron/Eta μια λίγο μεγαλύτερη (κατά 0.7°C) υπερεκτίμηση από το WRF.

Στα Σχήματα 7, 8 και 9 φαίνονται οι μέσες ετήσιες τιμές του σφάλματος και του απόλυτου σφάλματος καθώς και της ρίζας του μέσου ετήσιου τετραγωνικού σφάλματος για κάθε προγνωστική ώρα ξεχωριστά και για τα τρία μοντέλα. Ξεκινώντας από το GFS, φαίνεται ότι αυτό παρουσιάζει υποεκτίμηση της πραγματικής θερμοκρασίας κατά από το μεσημέρι (+24, +48, +72), τις απογευματινές (+03, +06, +27, +30, +51, +54) όπως και τις βραδινές (+09, +12, +15, +33, +36, +39, +57, +60, +63) ώρες με τιμές μικρότερες από -1.0°C που φτάνουν ως και τους -2.44°C . Η υποεκτίμηση αυτή είναι μικρότερη τις υπόλοιπες ώρες της ημέρας και σε κάποιες προγνωστικές ώρες παρατηρείται μια μικρή υπερεκτίμηση. Η μέση ετήσια τιμή του απόλυτου σφάλματος είναι συνήθως πάνω από τους 2.2°C κατά τη διάρκεια των απογευματινών ωρών και

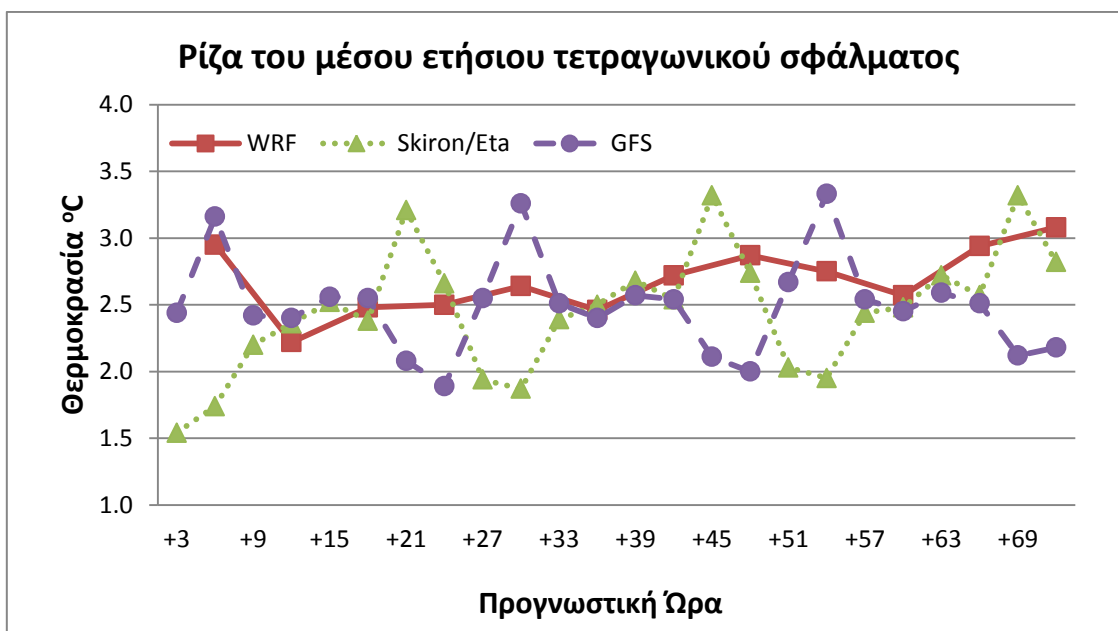
κάτω από τους 2.0 °C τις πρωινές ώρες (+18, +21, +42, +45, +66, +69) της ημέρας. Από την άλλη, το μοντέλο Skiron/Eta παρουσιάζει σε όλες τις προγνωστικές ώρες υπερεκτίμηση της πραγματικής θερμοκρασίας. Αυτή η υπερεκτίμηση είναι μεγαλύτερη από τις πρώτες πρωινές μέχρι τις μεσημεριανές ώρες και μικρότερη κατά τις απογευματινές και βραδινές ώρες. Η μέση τιμή του απόλυτου σφάλματος παρουσιάζει μικρότερες τιμές κατά τις απογευματινές ώρες και μέγιστες τιμές κατά τις πρώτες νυχτερινές ώρες. Τέλος το μοντέλο WRF παρουσιάζει υποεκτίμηση κατά τις απογευματινές και βραδινές ώρες και υπερεκτίμηση κατά τη διάρκεια της υπόλοιπης ημέρας. Αυτή η υπερεκτίμηση είναι μεγαλύτερη το μεσημέρι. Τις μεγαλύτερες τιμές απόλυτου σφάλματος παρουσιάζει το WRF κατά τις πρώτες πρωινές και μεσημεριανές ώρες φτάνοντας μέχρι και τους 2.4°C. Η κύμανση της ρίζας του μέσου ετήσιου τετραγωνικού σφάλματος κατά τη διάρκεια του δωδεκάμηνου είναι όμοια με αυτή του μέσου ετήσιου απόλυτου σφάλματος και στα τρία μοντέλα, με τιμές που κυμαίνονται από 1.5°C ως 3.4°C για το Skiron/Eta, από 1.9°C ως 3.4°C για το GFS και από 2.2°C ως 3.1°C για το WRF.



Σχήμα 7: Μέση ετήσια τιμή σφάλματος της θερμοκρασίας ανά μοντέλο και προγνωστική ώρα.



Σχήμα 8: Μέση ετήσια απόλυτη τιμή σφάλματος της θερμοκρασίας ανά μοντέλο και προγνωστική ώρα.

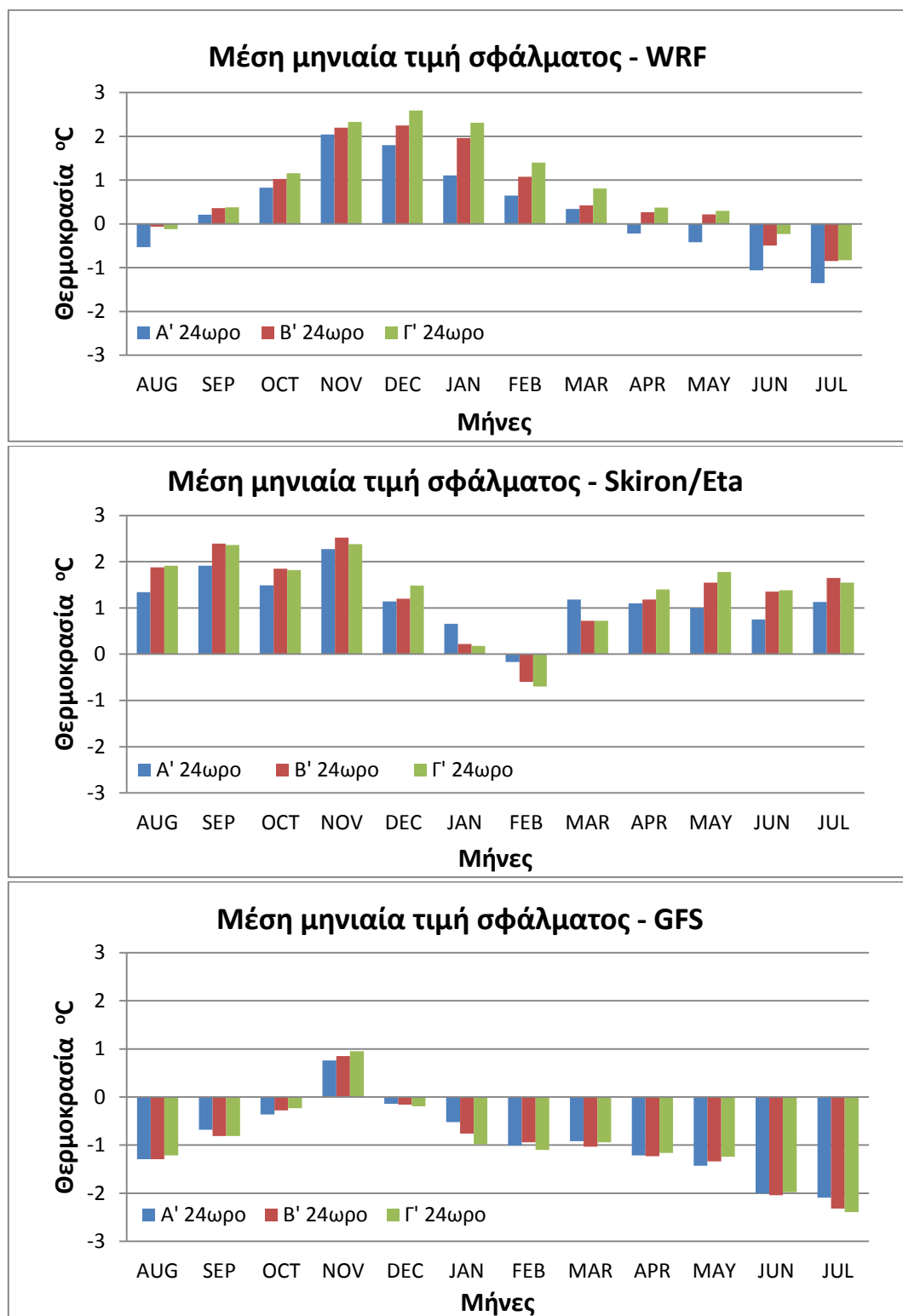


Σχήμα 9: Ρίζα του μέσου ετήσιου τετραγωνικού σφάλματος της θερμοκρασίας ανά μοντέλο και προγνωστική ώρα.

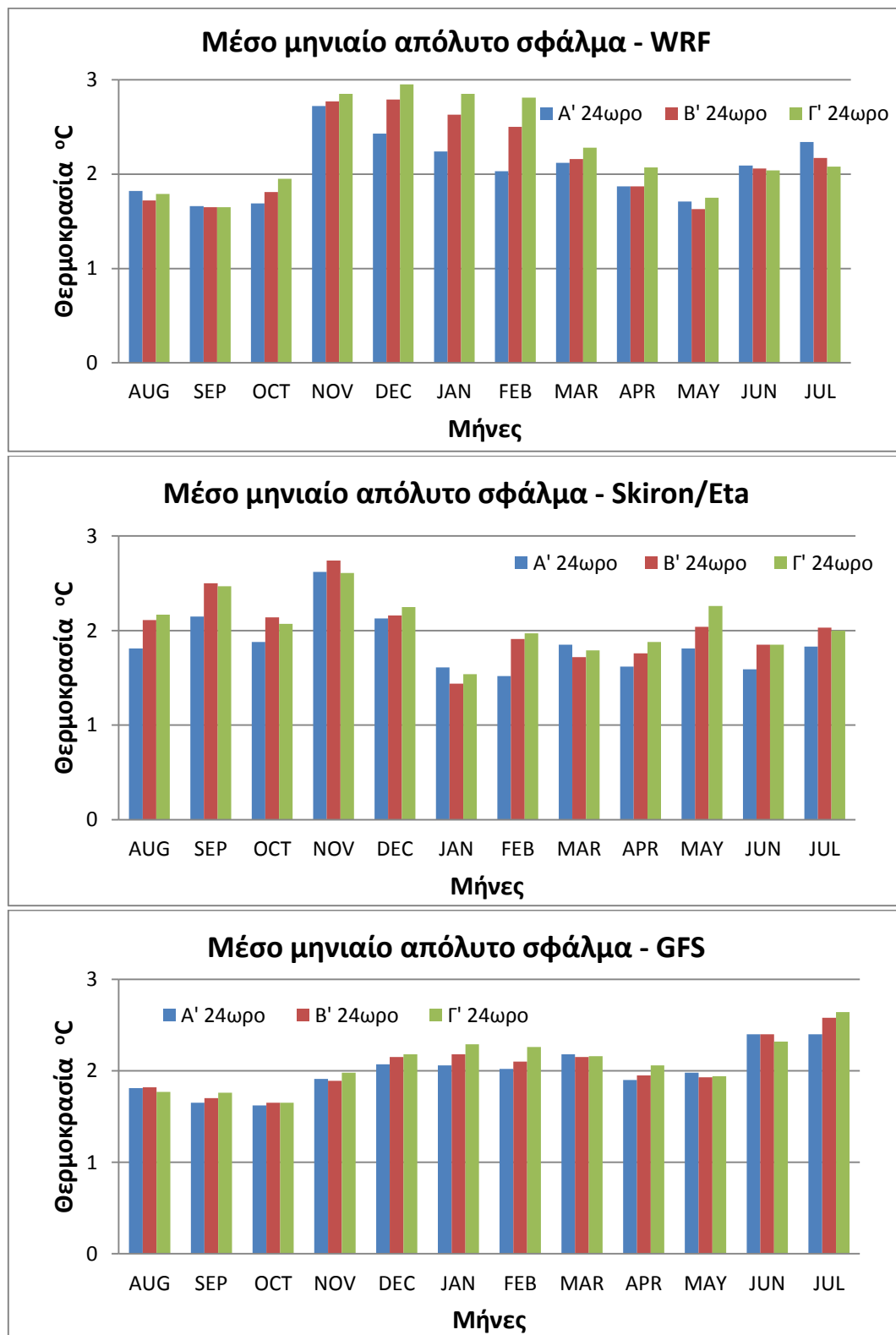
Εξετάζοντας τώρα κάθε μήνα ξεχωριστά σε αυτό το δωδεκάμηνο παρατηρούνται τα εξής (Σχήμα 10). Για το μοντέλο WRF η μέση τιμή σφάλματος είναι μεγαλύτερη τους μήνες Νοέμβριο και Δεκέμβριο (2.1°C) ενώ μικρότερη είναι κατά τους θερινούς μήνες με ελάχιστο τον Ιούλιο (-0.9°C). Εκτός από τους θερινούς μήνες και το Μάιο το WRF παρουσιάζει υπερεκτίμηση της θερμοκρασίας. Η τιμή του μέσου μηνιαίου απόλυτου σφάλματος είναι μεγαλύτερη από 2.0°C από το Νοέμβριο ως το Μάρτιο και τον Ιούλιο με μέγιστο το Νοέμβριο (2.7°C). Την ελάχιστη τιμή παρουσιάζει το Μάιο (1.6°C). Από την άλλη η ρίζα του μέσου μηνιαίου τετραγωνικού σφάλματος είναι ελάχιστη το Σεπτέμβριο (2.0°C) και μέγιστη το Δεκέμβριο (3.2°C), ενώ παρουσιάζει και ένα δευτερεύων μέγιστο τον Ιούλιο (2.6°C).

Για το μοντέλο GFS παρατηρείται μια συνεχής υποεκτίμηση της πραγματικής θερμοκρασίας σε όλη τη διάρκεια του έτους εκτός από το Νοέμβριο, όπου και η μέση τιμή σφάλματος είναι 0.8°C . Η μέγιστη υποεκτίμηση παρατηρείται τον Ιούλιο (-2.1°C), ενώ στους υπόλοιπους μήνες παρατηρείται μια απλή κύμανση της μέσης τιμής του σφάλματος από τη μία ακραία τιμή στην άλλη. Όπως φαίνεται στο Σχήμα 10, σε κάποιους μήνες, το τρίτο ή το δεύτερο 24ωρο της πρόγνωσης είναι κατά μέσο όρο πλησιέστερα στην πραγματική τιμή της θερμοκρασίας από ότι το πρώτο 24ωρο. Η μέση τιμή του απόλυτου σφάλματος είναι μεγαλύτερη από 2.0°C από το Δεκέμβριο ως το Μάρτιο και το δίμηνο Ιουνίου-Ιουλίου όπου και φτάνει τη μέγιστη τιμή (2.5°C). Αντίθετα, τους υπόλοιπους μήνες έχει τιμή κάτω από 2.0°C με ελάχιστο τον Οκτώβριο (1.7°C). Δύο μέγιστα παρουσιάζει και η τιμή της ρίζας του μέσου μηνιαίου τετραγωνικού σφάλματος. Ένα το Φεβρουάριο (2.7°C) και ένα το δίμηνο Ιουνίου-Ιουλίου (3.0°C). Σε όλους τους υπόλοιπους μήνες η τιμή της είναι πάνω από τους 2.0°C , με ελάχιστο τον Οκτώβριο.

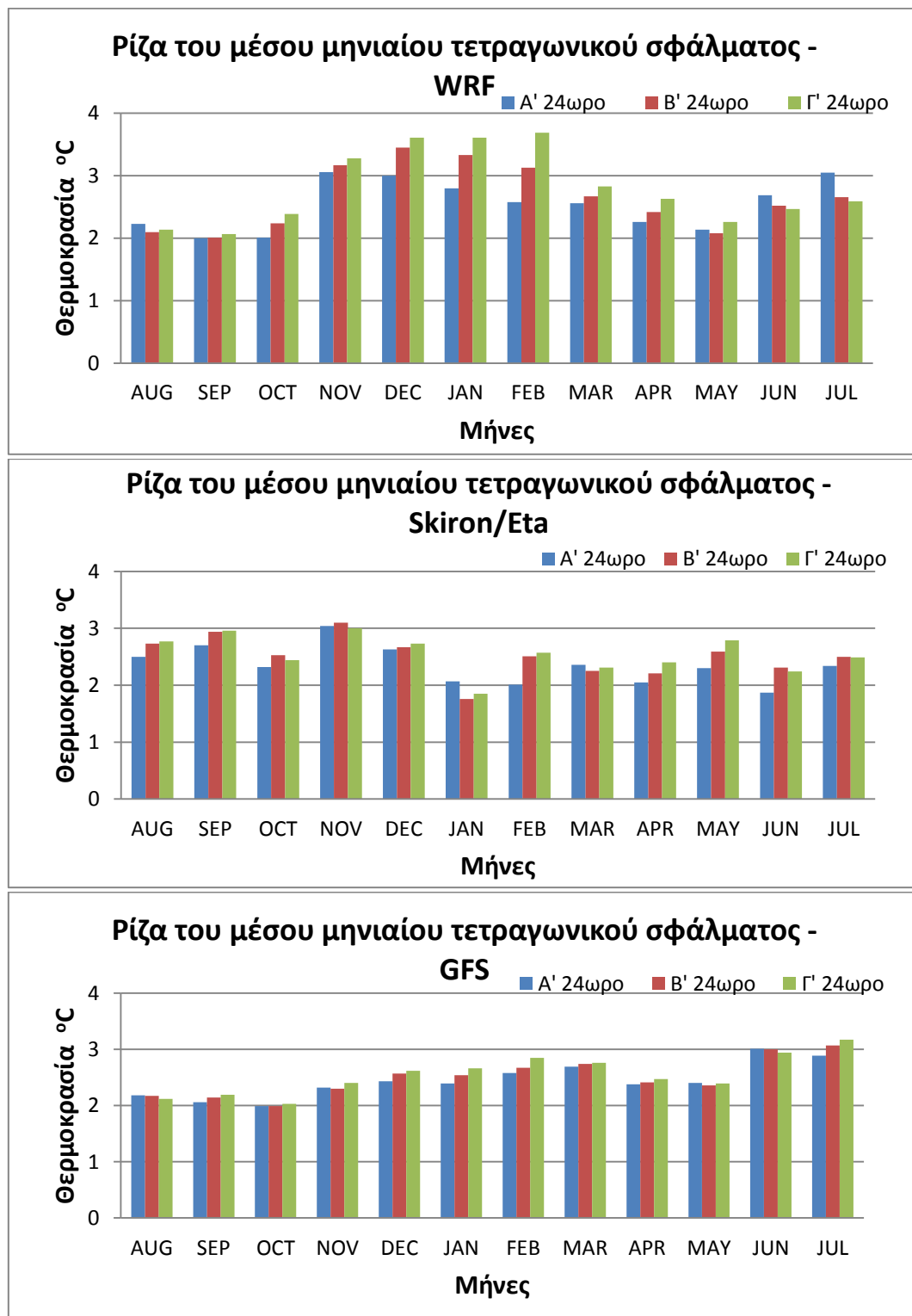
Για το μοντέλο Skiron/Eta η μέση μηνιαία τιμή σφάλματος είναι μεγαλύτερη κατά τους φθινοπωρινούς μήνες με τιμές μεγαλύτερες από 1.5°C και το Νοέμβριο παρουσιάζει μέγιστο με 2.2°C . Ο Φεβρουάριος είναι ο μοναδικός μήνας που παρουσιάζει υποεκτίμηση της θερμοκρασίας (-0.4°C), ενώ στους υπόλοιπους μήνες η τιμή αυτή είναι κοντά στον 1.0°C . Η ελάχιστη μέση μηνιαία τιμή απόλυτου σφάλματος παρατηρείται τον Ιανουάριο (1.5°C). Εκτός από τους φθινοπωρινούς μήνες και το Μάιο, η τιμή αυτή είναι μικρότερη από τους 2.0°C . Το μέγιστο παρατηρείται το Νοέμβριο (2.5°C).



Σχήμα 10: Μέση μηνιαία τιμή σφάλματος ανά 24ωρο πρόγνωσης για κάθε μοντέλο.



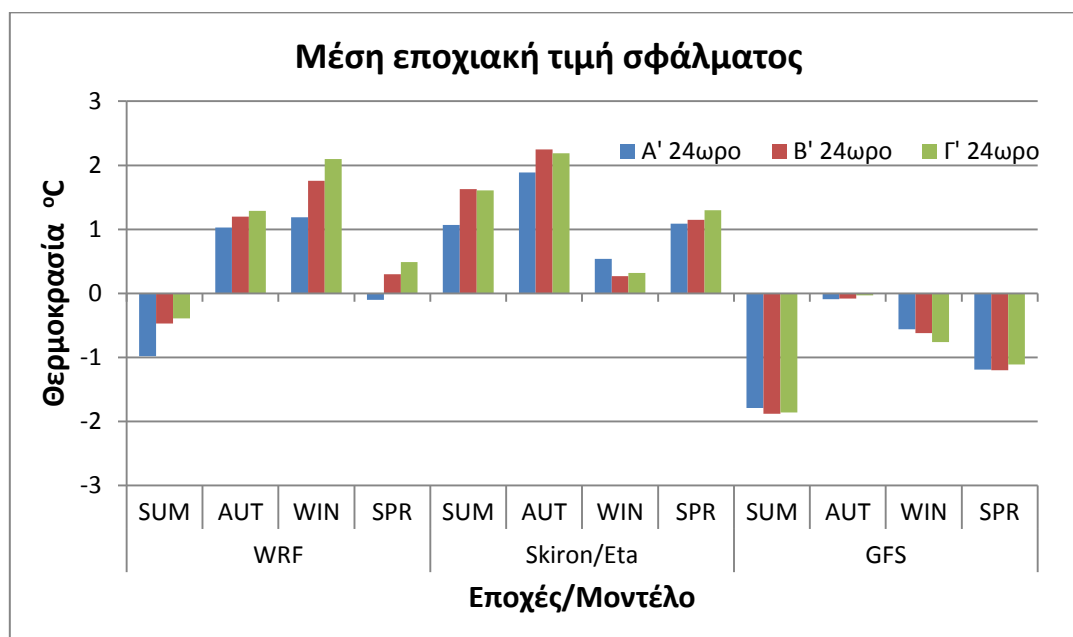
Σχήμα 11: Μέση μηνιαία τιμή απόλυτου σφάλματος ανά 24ωρο πρόγνωσης για κάθε μοντέλο.



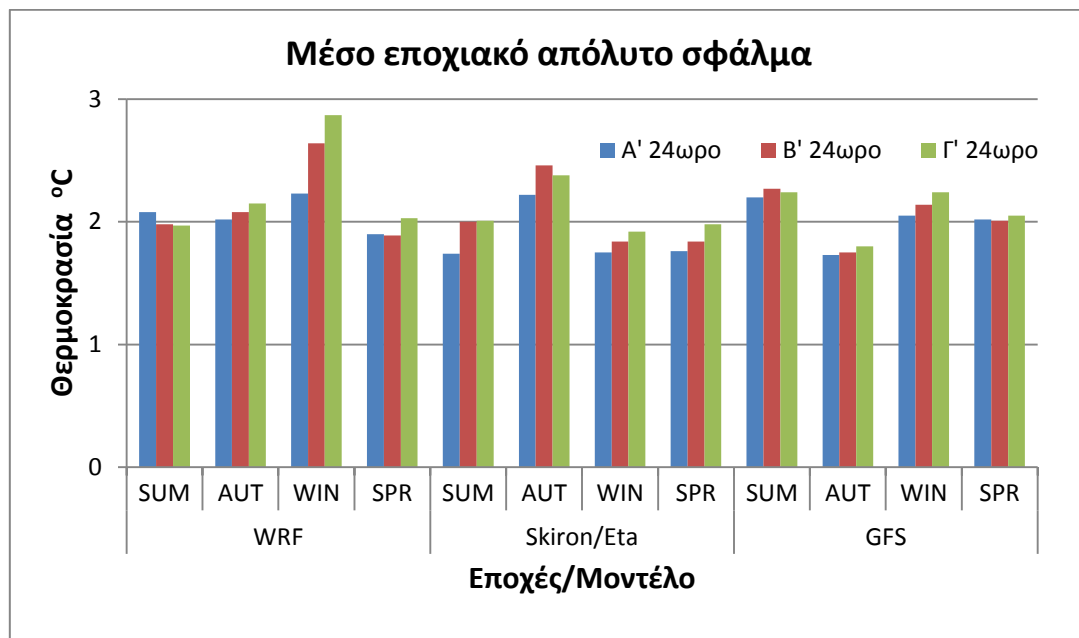
Σχήμα 12: Ρίζα του μέσου μηνιαίου τετραγωνικού σφάλματος ανά 24ωρο πρόγνωσης για κάθε μοντέλο.

Η τιμή της ρίζας του μέσου μηνιαίου τετραγωνικού σφάλματος είναι μεταξύ των 2.0 °C και 3.0 °C, με μέγιστο το Νοέμβριο. Εξαιρέση αποτελεί ο Ιανουάριος με τιμή 1.8 °C.

Στα Σχήματα 13 και 14 φαίνονται οι μέσες εποχιακές τιμές του σφάλματος και του απόλυτου σφάλματος για κάθε μοντέλο. Όπως φαίνεται, το GFS πραγματοποιεί καλύτερες προγνώσεις κατά τη διάρκεια του φθινοπώρου, ενώ το WRF και το Skiron/Eta κατά την άνοιξη και το καλοκαίρι. Αντίθετα μεγαλύτερα σφάλματα παρουσιάζουν το GFS κατά το καλοκαίρι, το WRF κατά το χειμώνα και το Skiron/Eta κατά το φθινόπωρο.



Σχήμα 13: Μέση εποχιακή τιμή σφάλματος ανά 24ωρο πρόγνωσης για κάθε μοντέλο.



Σχήμα 14: Μέση εποχιακή τιμή απόλυτου σφάλματος ανά 24ωρο πρόγνωσης για κάθε μοντέλο.

5.2 Ανάλυση των σφαλμάτων μετά την εφαρμογή του φίλτρου

5.2.1 Επιλογή βέλτιστης μεθόδου

Κάθε μία από τις τέσσερις διαφορετικές παραλλαγές της εφαρμογής του φίλτρου Kalman που χρησιμοποιήθηκαν (μέθοδος 1, μέθοδος 2α, μέθοδος 2β και μέθοδος 2γ), είχε στόχο την εκτίμηση του συστηματικού μέρους του σφάλματος της επόμενης πρόγνωσης της θερμοκρασίας. Η διαφορά ανάμεσά τους βρίσκεται στο γεγονός ότι στην πρώτη γίνεται απευθείας εκτίμηση του συστηματικού σφάλματος της πρόγνωσης, ενώ στις άλλες αυτό συμβαίνει μέσω ενός πολυωνύμου του οποίου η μεταβλητή είναι η έξοδος του μοντέλου (τιμή θερμοκρασίας) και οι συντελεστές αυτού υπολογίζονται από τον αλγόριθμο. Το πολυώνυμο αυτό είναι μηδενικού βαθμού στην μέθοδο 2α, πρωτοβάθμιο στην 2β και δευτεροβάθμιο στην 2γ. Ο υπολογισμός των εκτιμήσεων του σφάλματος έγινε κάθε φορά με τη βοήθεια των σφαλμάτων των προηγούμενων τεσσάρων (4), επτά (7), δέκα (10), δεκαπέντε (15) και τριάντα (30) ημερών. Έτσι συνολικά έγιναν είκοσι (20) εφαρμογές του φίλτρου.

Στα Σχήματα 15 και 16 απεικονίζονται η μέση ετήσια τιμή σφάλματος της πρόγνωσης και το μέσο ετήσιο απόλυτο σφάλμα, μετά την εφαρμογή του φίλτρου Kalman, για κάθε μοντέλο και εφαρμογή. Από αυτά φαίνεται ότι σε όλες σχεδόν τις περιπτώσεις η μέση τιμή του σφάλματος σχεδόν μηδενίζεται, ενώ το μεγαλύτερο σφάλμα (κατά απόλυτη τιμή) που παρατηρείται είναι 0.17°C . Ο στόχος της μεθόδου επιτυγχάνεται απόλυτα στις περισσότερες περιπτώσεις. Οι μοναδικές περιπτώσεις που η μέση τιμή του σφάλματος ξεπέρασε τους 0.1°C παρατηρήθηκαν στις μεθόδους 2α και 2β για το μοντέλο GFS και 2β για το WRF.

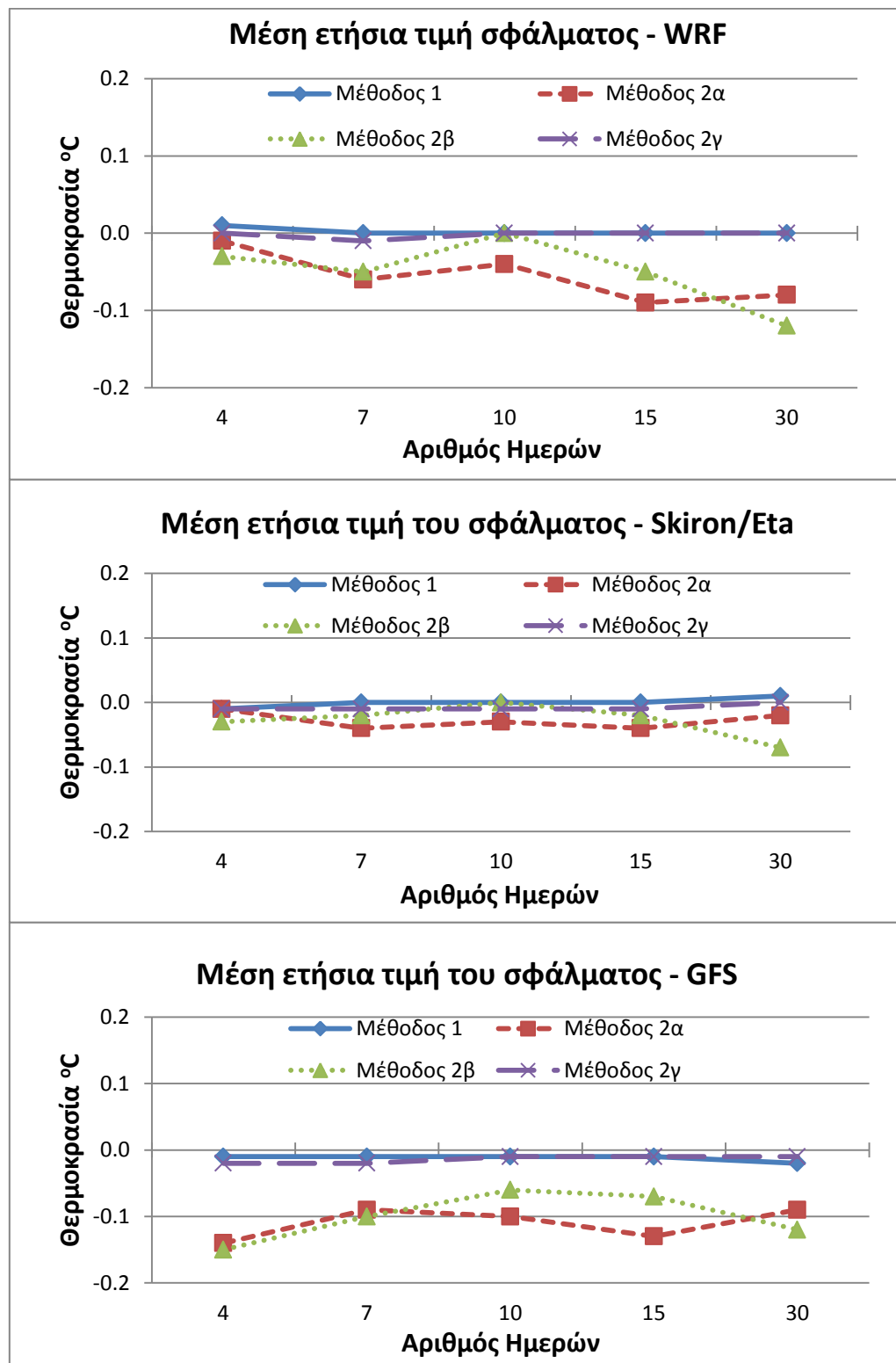
Η μέθοδος που οδήγησε σε καλύτερα αποτελέσματα προκύπτει από τη μελέτη του μέσου ετήσιου απόλυτου σφάλματος (Σχήμα 16). Παρατηρείται ότι το μεγαλύτερο μέσο ετήσιο απόλυτο σφάλμα παρουσιάζει η μέθοδος 2γ, ανεξάρτητα από τον αριθμό των ημερών που χρησιμοποιούνται και το μοντέλο. Το αμέσως μεγαλύτερο μέσο ετήσιο απόλυτο σφάλμα δίνει η μέθοδος 2α, ενώ η μέθοδος 2β έχει σε κάθε περίπτωση μικρότερο απόλυτο σφάλμα από τις άλλες δύο. Προκύπτει δηλαδή ότι το πολυώνυμο που σε αυτή την περίπτωση δίνει τα καλύτερα αποτελέσματα είναι το πρωτοβάθμιο. Οι διαφορές πάντως ανάμεσα στις διαφορετικές τιμές των τριών αυτών μεθόδων δεν είναι ποτέ μεγαλύτερες από 0.2°C . Η μέθοδος 1 από την άλλη, παρουσιάζει τη μικρότερη τιμή απόλυτου σφάλματος όταν λαμβάνονται υπόψη οι τιμές θερμοκρασίας των προηγούμενων δέκα (10) ημερών. Σε κάθε μοντέλο, αυτές οι τιμές είναι μικρότερες από οποιαδήποτε άλλη εφαρμογή. Για λιγότερες από δέκα (10), ή περισσότερες από δέκα (10) ημέρες, η μέση τιμή του απόλυτου σφάλματος είναι ίση ή μεγαλύτερη από αυτή της μεθόδου 2β. Φαίνεται λοιπόν ότι η μέθοδος που παρουσιάζει τη μικρότερη μέση τιμή απόλυτου σφάλματος είναι η μέθοδος 1 χρησιμοποιώντας δεδομένα από τις προηγούμενες δέκα (10) ημέρες.

Η ύπαρξη απόλυτου σφάλματος μετά τη χρήση του φίλτρου Kalman οφείλεται σε δύο λόγους. Πρώτον, υπάρχουν περιπτώσεις που το φίλτρο μείωσε το σφάλμα της πρόγνωσης αλλά δεν το εκμηδένισε και όπως θα φανεί πιο κάτω, αυτές είναι οι περισσότερες. Επίσης, υπάρχουν περιπτώσεις που η χρήση του φίλτρου αύξησε το σφάλμα. Στον Πίνακα 1 παρουσιάζονται για κάθε μέθοδο και μοντέλο τα ποσοστά των περιπτώσεων στις οποίες παρατηρήθηκε βελτίωση της πρόγνωσης της θερμοκρασίας έπειτα από τη χρήση του φίλτρου. Το ποσοστό αυτό κυμαίνεται από 58% ως 63% για το WRF, από 58% ως 65% για το Skiron/Eta και από 59% ως 62% για το

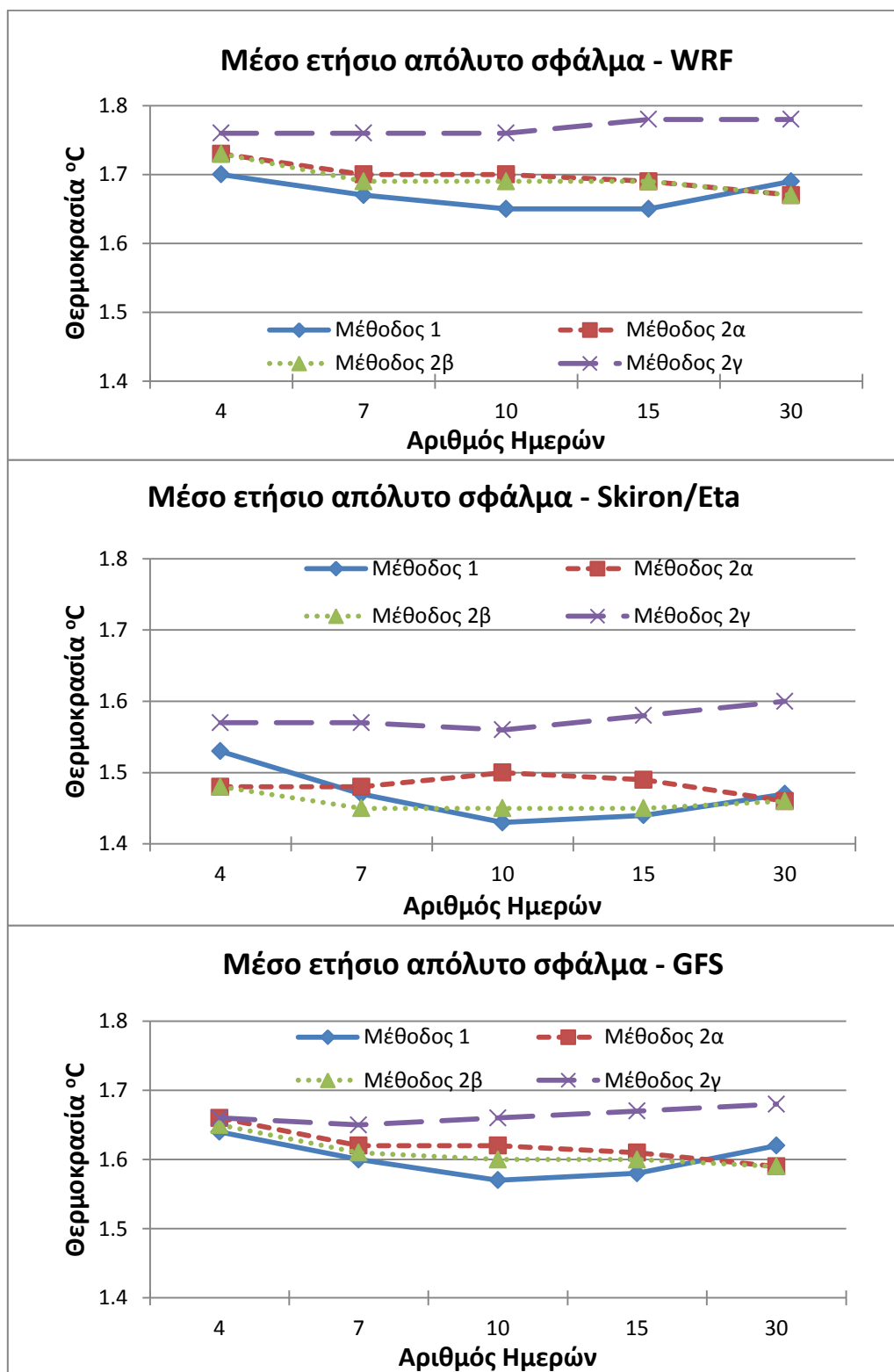
GFS, ανάλογα με τη μέθοδο. Το συμπέρασμα είναι ότι με τη χρήση του φίλτρου επιτυγχάνεται βελτίωση της πρόγνωσης της τάξεως του 60% των περιπτώσεων για κάθε μοντέλο. Στις υπόλοιπες περιπτώσεις η εκτίμηση του σφάλματος προκαλεί μεγαλύτερο λάθος από αυτό του μοντέλου. Τα μικρότερα ποσοστά παρουσιάζει η μέθοδος 2γ, η οποία σε καμία περίπτωση δεν ξεπερνά το 60%. Η μέθοδος 2α και η μέθοδος 2β έχουν μεγαλύτερα ποσοστά όταν χρησιμοποιούνται δεδομένα δεκαπέντε (15) ή τριάντα (30) ημερών, ενώ η μέθοδος 1 όταν χρησιμοποιούνται δεδομένα δέκα (10) ή δεκαπέντε (15) ημερών. Οι διαφορές ανάμεσα στα καλύτερα ποσοστά των μεθόδων 1, 2α και 2β είναι της τάξης του 0.1% με 0.6% και οι οποίες μπορούν να θεωρηθούν ασήμαντες, καθώς σε ένα σύνολο οχτώ χιλιάδων διακοσίων (8200) περίπου περιπτώσεων για το Skiron/Eta, οχτώ χιλιάδων επτακοσίων (8700) για το GFS και τεσσάρων χιλιάδων διακοσίων (4200) για το WRF το 0.5% αντιπροσωπεύει γύρω στις δέκα (10) με είκοσι (20) προγνώσεις ανάλογα με το μοντέλο.

Πίνακας 1: Ποσοστά επιτυχίας φίλτρου Kalman ανά μέθοδο και μοντέλο.

Μέθοδος	Αριθμός Ημερών	WRF	Skiron/Eta	GFS
1	4	59.4	60.5	60.4
1	7	61.8	62.8	61.2
1	10	62.8	64.2	61.8
1	15	62.6	64.7	61.2
1	30	61.7	63.4	60.8
2α	4	60.9	63.0	59.4
2α	7	62.0	63.5	59.8
2α	10	62.8	63.0	60.5
2α	15	62.7	63.2	61.1
2α	30	62.9	63.6	61.4
2β	4	60.7	62.7	59.5
2β	7	62.0	63.7	60.5
2β	10	62.4	64.2	60.5
2β	15	62.8	63.9	60.9
2β	30	62.7	63.5	62.0
2γ	4	58.1	59.1	59.2
2γ	7	58.3	58.8	59.9
2γ	10	58.0	58.8	59.2
2γ	15	58.1	58.3	59.2
2γ	30	58.0	58.1	58.8



Σχήμα 15: Μέση ετήσια τιμή σφάλματος της πρόγνωσης μετά την εφαρμογή του φίλτρου Kalman για κάθε μοντέλο και μέθοδο.



Σχήμα 16: Μέσο ετήσιο απόλυτο σφάλμα της πρόγνωσης μετά την εφαρμογή του φίλτρου Kalman για κάθε μοντέλο και μέθοδο.

Συνοψίζοντας τα παραπάνω φαίνεται ότι η μέθοδος 1, με δεδομένα σφαλμάτων δέκα (10) ημερών (μέθοδος 1-10), παρουσιάζει τα καλύτερα αποτελέσματα από κάθε άλλη μέθοδο. Αυτό συμβαίνει, καθώς έχει ένα από τα μεγαλύτερα ποσοστά επιτυχίας σε κάθε μοντέλο. Επιπλέον παρουσιάζει το μικρότερο ελάχιστο απόλυτο ετήσιο σφάλμα και στα τρία μοντέλα, ενώ παράλληλα η τιμή του μέσου σφάλματος είναι μηδέν. Έτσι δείχνει ότι είναι αυτή που λειτουργεί καλύτερα ανεξάρτητα από το μοντέλο. Από τις άλλες τρεις μεθόδους καλύτερη φαίνεται ότι είναι η 2β με χρήση δεδομένων τριάντα (30) ημερών (μέθοδος 2β-30). Παρόλα αυτά οι διαφορές ανάμεσά τους είναι πολύ μικρές.

Για να εξεταστούν οι περιπτώσεις στις οποίες επιτυγχάνεται βελτίωση της πρόγνωσης και το μέγεθος αυτής της βελτίωσης, αλλά και να φανεί πότε το φίλτρο αδυνατεί να εκτιμήσει σωστά το σφάλμα, παρουσιάζονται τα σχήματα με τα ετήσια και μηνιαία σφάλματα. Σε αυτά φαίνονται οι τιμές των μεθόδων 1-10 και 2β-10. Η πρώτη επιλέχθηκε καθώς είναι η καλύτερη και η δεύτερη για λόγους σύγκρισης.

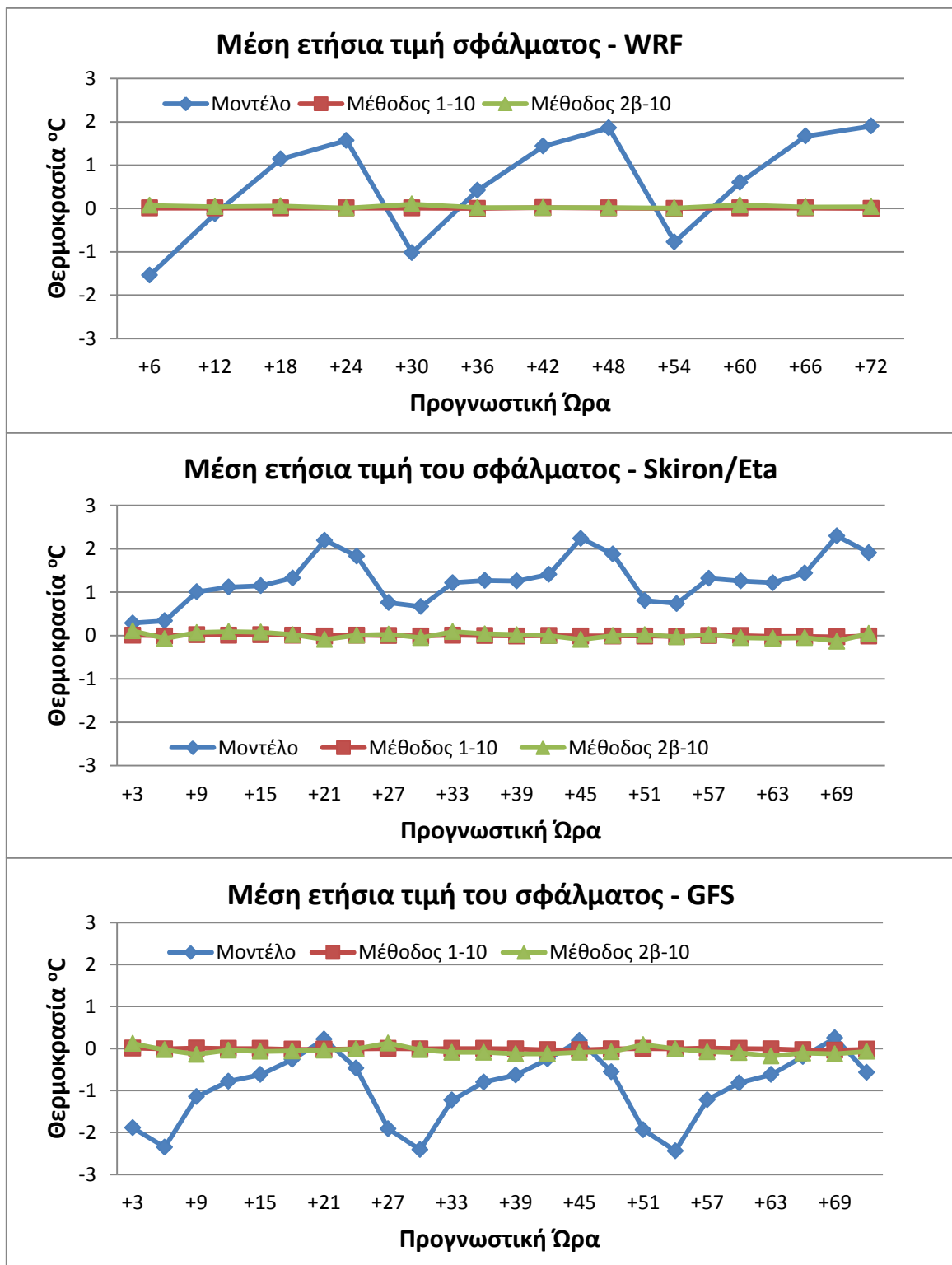
5.2.2 Ετήσιο και μηνιαίο σφάλμα της πρόγνωσης με χρήση του φίλτρου Kalman

Στο Σχήμα 17 φαίνεται η μέση ετήσια τιμή σφάλματος ανά προγνωστική ώρα για κάθε μοντέλο για τις μεθόδους 1-10 και 2β-10. Είναι ξεκάθαρο ότι ανεξάρτητα από την προγνωστική ώρα και την τιμή του μέσου σφάλματος σε αυτή για το δωδεκάμηνο, το φίλτρο μηδενίζει αυτή την τιμή. Διαφορά δεν υπάρχει ούτε ανάμεσα στις δύο μεθόδους αποδεικνύοντας ότι λειτουργούν το ίδιο καλά ως προς το στόχο που υπήρχε.

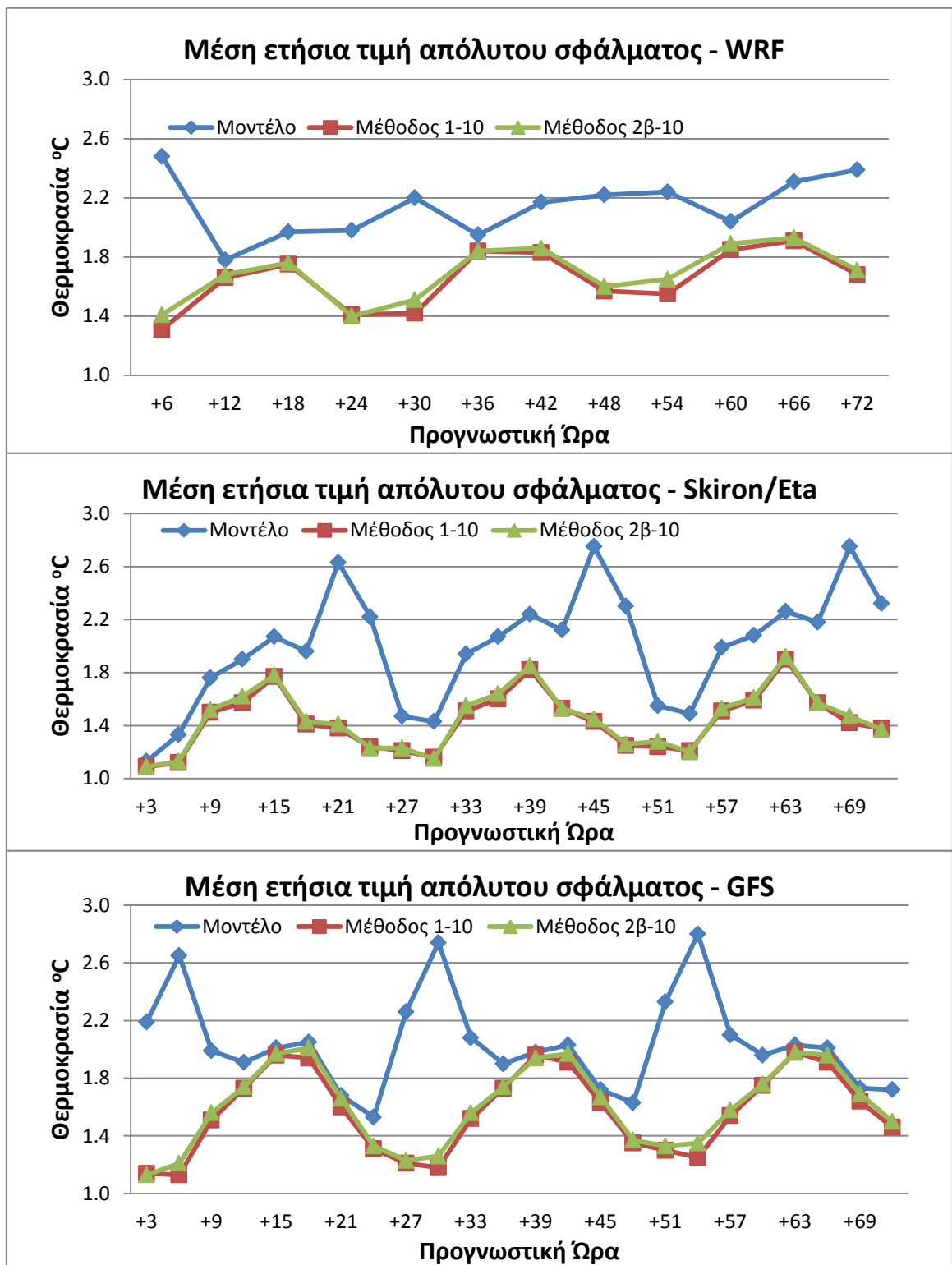
Παρατηρώντας το Σχήμα 18 στο οποίο φαίνεται το μέσο ετήσιο απόλυτο σφάλμα ανά προγνωστική ώρα, για κάθε μοντέλο, για τις μεθόδους 1-10 και 2β-10 μπορεί να δει κάποιος ότι σε όλες τις προγνωστικές ώρες και στα τρία μοντέλα υπάρχει μείωση της τιμής αυτής. Ελαφρώς μεγαλύτερη είναι η μείωση του σφάλματος όταν χρησιμοποιείται η μέθοδος 1-10, αλλά και πάλι η διαφορά της από την μέθοδο 2β-10 είναι μικρή. Σε κάποιες περιπτώσεις η μείωση της μέσης ετήσιας τιμής του απόλυτου σφάλματος είναι εντυπωσιακή, ενώ σε άλλες πολύ μικρή. Σε κανένα μοντέλο και σε καμία προγνωστική ώρα ωστόσο δεν παρατηρήθηκε αύξηση της μέσης ετήσιας

τιμής του απόλυτου σφάλματος. Το σφάλμα της 21^{ης}, της 45^{ης} και της 69^{ης} προγνωστικής ώρας από το μοντέλο Skiron/Eta ήταν 2.6°C πριν τη χρήση του φίλτρου και 1.4°C μετά, μειώθηκε δηλαδή κατά 46%. Αντίθετα η τιμή του σφάλματος της 15^{ης}, της 39^{ης} και της 63^{ης} προγνωστικής ώρας μειώθηκε ελάχιστα. Το απόλυτο σφάλμα της πρόγνωσης για το μοντέλο WRF είναι σχεδόν πάντα μεγαλύτερο από 2.0°C πριν τη χρήση του φίλτρου, αλλά μετά είναι πάντα μικρότερο από 2.0°C. Παρόλα αυτά, η μείωση και σε αυτή την περίπτωση δεν είναι πάντα η ίδια, αλλά κυμαίνεται από 0.1°C ως 1.1°C ανάλογα με την προγνωστική ώρα. Όμοια και για το μοντέλο GFS, η μέση τιμή του απόλυτου σφάλματος την 54^η προγνωστική ώρα μειώνεται από τους 2.7°C στους 1.4°C, αντίθετα την 63^η προγνωστική ώρα η μείωση είναι σχεδόν μηδενική ενώ τις υπόλοιπες προγνωστικές ώρες η μείωση παίρνει ενδιάμεσες τιμές. Η μείωση λοιπόν του μέσου ετήσιου απόλυτου σφάλματος είναι δεδομένη σε κάθε περίπτωση με το φίλτρο, χωρίς όμως ποτέ να προκαλεί ελάττωση του σφάλματος σε τιμή μικρότερη από τον 1.0°C.

Οι προγνωστικές ώρες στις οποίες παρατηρείται μικρή ως και μηδενική μείωση δεν είναι τυχαίες σε κανένα από τα τρία μοντέλα. Αντίθετα συνδέονται στενά με την τιμή του μέσου ετήσιου σφάλματος που παρουσιάζουν. Όταν αυτή η μέση τιμή βρίσκεται στο διάστημα (-0.6°C , 0.6°C) τότε το αντίστοιχο μέσο ετήσιο απόλυτο σφάλμα παραμένει σχεδόν αμετάβλητο, ενώ όταν η τιμή αυτή βρίσκεται έξω από αυτό το διάστημα, τότε το φίλτρο λειτουργεί και μειώνει το σφάλμα που εντοπίζει. Μεγάλες μειώσεις παρατηρούνται όταν το σφάλμα είναι μεγαλύτερο από 2.0°C. Δηλαδή η μεγαλύτερη μείωση παρατηρείται όταν το σφάλμα είναι μεγάλο, πράγμα που είναι απόλυτα λογικό.



Σχήμα 17: Μέση ετήσια τιμή σφάλματος ανά προγνωστική ώρα για κάθε μοντέλο και για τις μεθόδους 1-10 και 2β-10.

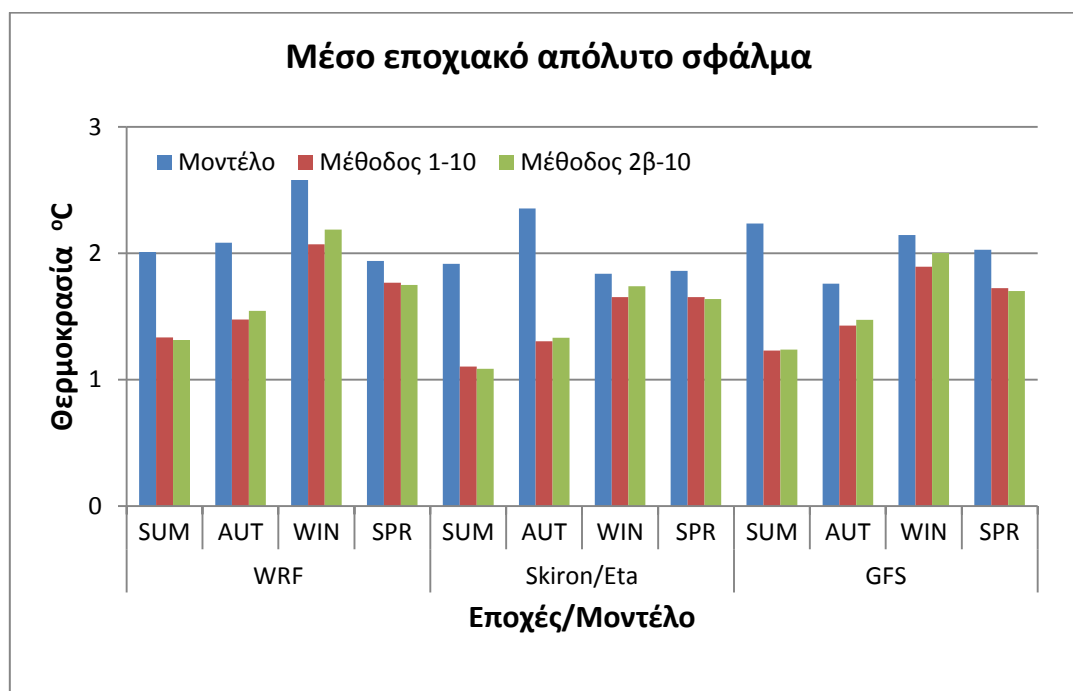


Σχήμα 18: Μέση ετήσια τιμή απόλυτου σφάλματος ανά προγνωστική ώρα για κάθε μοντέλο και για τις μεθόδους 1-10 και 2β-10.

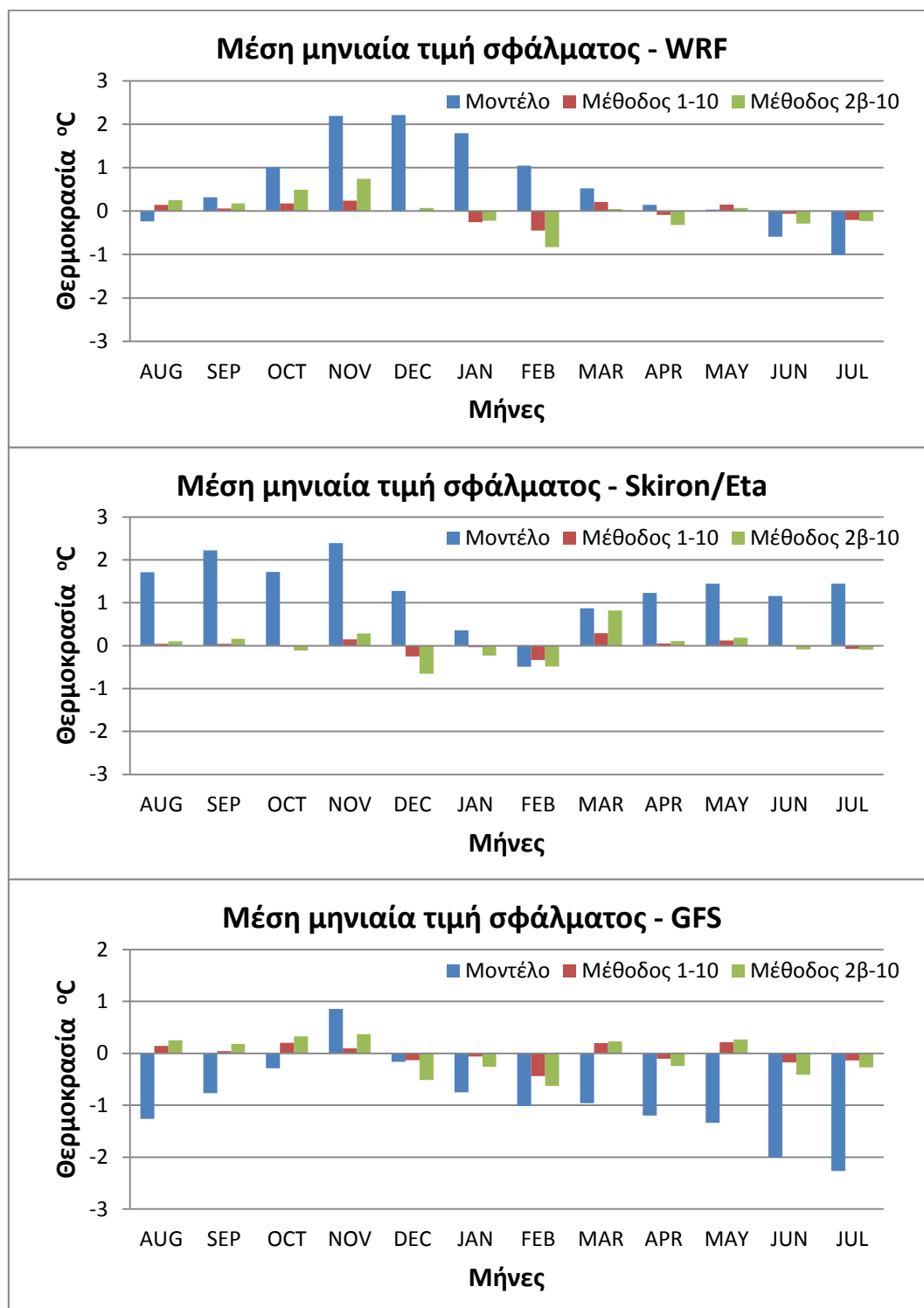
Στα Σχήματα 20 και 21 παρουσιάζονται η μέση μηνιαία τιμή σφάλματος και το μέσο μηνιαίο απόλυτο σφάλμα της πρόγνωσης της θερμοκρασίας κάθε μοντέλου πριν και μετά τη χρήση του φίλτρου Kalman. Η μέση μηνιαία τιμή σφάλματος της μεθόδου 1-10 είναι γενικά πιο κοντά στην τιμή μηδέν από αυτή της μεθόδου 2β-10. Για τη μέθοδο 1-10 η τιμή αυτή είναι πάντα μικρότερη από 0.3°C και μεγαλύτερη από -0.5°C . Ο μήνας Φεβρουάριος παρουσιάζει το μεγαλύτερο κατά απόλυτη τιμή μέσο μηνιαίο σφάλμα και στα τρία μοντέλα. Στις περισσότερες περιπτώσεις η χρήση του φίλτρου ελαττώνει, ή ακόμα και εξαφανίζει, το σφάλμα και ειδικά στο μοντέλο Skiron/Eta οι εννιά (9) από τους δώδεκα (12) μήνες παρουσιάζουν σφάλμα μετά τη χρήση του φίλτρου μικρότερο από 0.2°C (κατά απόλυτη τιμή). Εξαίρεση αποτελεί ο μήνας Μάιος από το μοντέλο WRF, όπου η μέση τιμή σφάλματος πριν τη χρήση του φίλτρου ήταν 0.0°C και μετά αυξήθηκε στους 0.1°C , αλλά μια τέτοια αύξηση είναι αναμενόμενη. Παράλληλα, παρατηρείται ότι υπάρχει μείωση του μέσου απόλυτου σφάλματος σε κάθε μήνα και στα τρία μοντέλα εκτός από τον μήνα Μάρτιο στο μοντέλο Skiron/Eta, όπου η τιμή μένει αμετάβλητη. Γενικά η μείωση του απόλυτου σφάλματος είναι μεγαλύτερη στη μέθοδο 1-10 από ότι στη μέθοδο 2β-10, αλλά η διαφορά ανάμεσα στις δύο είναι πάντα μικρότερη από 0.3°C . Για το μοντέλο WRF η μεγαλύτερη μείωση παρατηρείται τους μήνες Νοέμβριο, Δεκέμβριο και Ιούλιο όπου το απόλυτο σφάλμα μειώνεται περισσότερο από 0.8°C και η μικρότερη μείωση του μήνες Μάρτιο και Απρίλιο, όπου η τιμή αυτή μειώνεται μόνο κατά 0.1°C με 0.2°C . Πριν τη χρήση του φίλτρου πέντε (5) μήνες παρουσίαζαν μέσο απόλυτο σφάλμα μεγαλύτερο από 2.1°C , ενώ μετά μόλις ένας. Όμοια για το μοντέλο Skiron/Eta οι μήνες Αύγουστος, Σεπτέμβριος και Δεκέμβριος παρουσιάζουν μείωση της απόλυτης τιμής του σφάλματος μεγαλύτερη από 1.0°C . Μετά τη χρήση του φίλτρου δεν υπάρχει κανένας μήνας με μέσο απόλυτο σφάλμα μεγαλύτερο από 2.0°C , ενώ προηγουμένως υπήρχαν έξι (6) μήνες.

Τη μεγαλύτερη μείωση απόλυτου σφάλματος για το μοντέλο GFS παρουσιάζει ο Ιούλιος με 1.3°C και ακολουθεί ο Ιούνιος με 0.9°C , ενώ τη μικρότερη ο Μάρτιος (0.05°C). Έξι μήνες παρουσιάζουν μέσο απόλυτο σφάλμα μεγαλύτερο από 2.0°C πριν τη χρήση του φίλτρου, αλλά μόνο ένας μετά. Και στα τρία μοντέλα το μέγιστο μέσο μηνιαίο απόλυτο σφάλμα μειώνεται και μετακινείται, από τον Νοέμβριο στον Ιανουάριο στο WRF, από τον Νοέμβριο στον Μάρτιο για το Skiron/Eta και από τον Ιούλιο στον Μάρτιο για το GFS. Αξιοσημείωτο είναι το γεγονός ότι ο

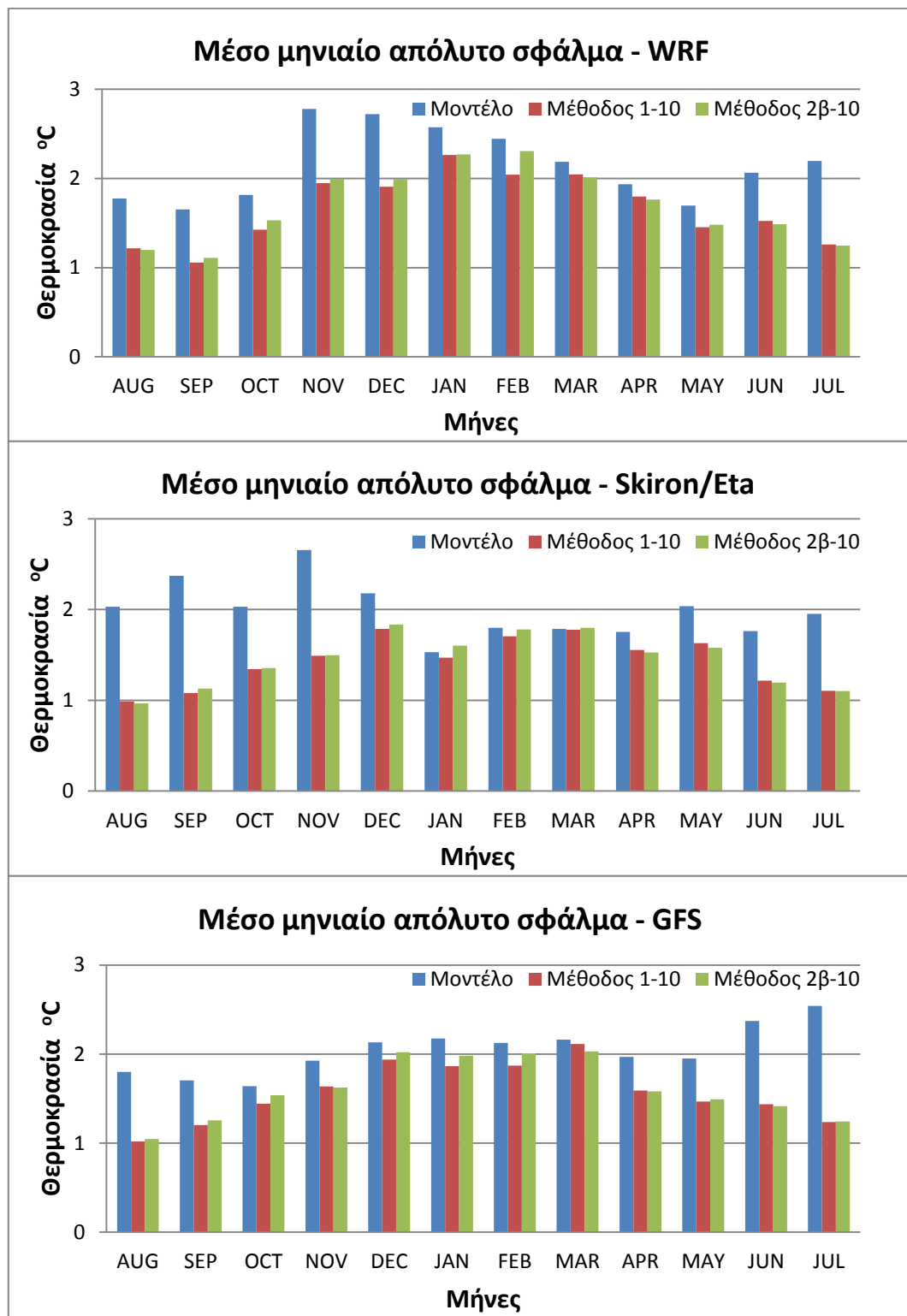
Ιούλιος για το GFS παρουσιάζει το μεγαλύτερο απόλυτο σφάλμα πριν τη χρήση του φίλτρου, αλλά μόλις το τρίτο μικρότερο μετά τη χρήση του φίλτρου. Η ετήσια κύμανση του απόλυτου σφάλματος αλλάζει και στα τρία μοντέλα και τείνει να δώσει μια κοινή εικόνα σύμφωνα με την οποία παρατηρείται μια απλή κύμανση με μέγιστη τιμή τον χειμώνα και ελάχιστη το καλοκαίρι. Σε κάθε μοντέλο οι φθινοπωρινοί μήνες παρουσιάζουν μικρότερο απόλυτο σφάλμα από τους εαρινούς, όπως φαίνεται και στο Σχήμα 19.



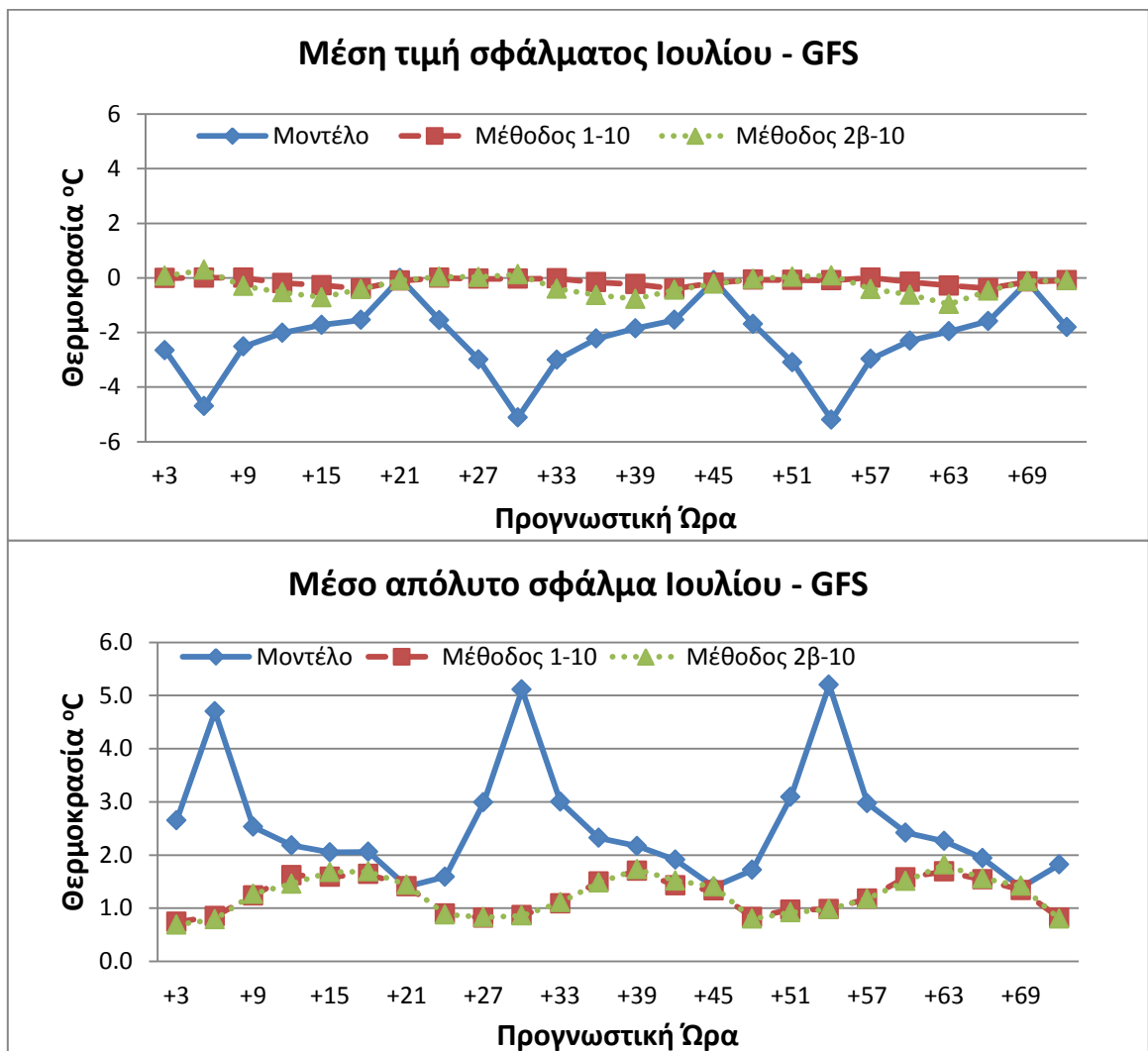
Σχήμα 19: Μέσο εποχιακό απόλυτο σφάλμα της πρόγνωσης της θερμοκρασίας κάθε μοντέλου πριν και μετά τη χρήση του φίλτρου Kalman.



Σχήμα 20: Μέση μηνιαία τιμή σφάλματος της πρόγνωσης της θερμοκρασίας κάθε μοντέλου πριν και μετά τη χρήση του φίλτρου Kalman.



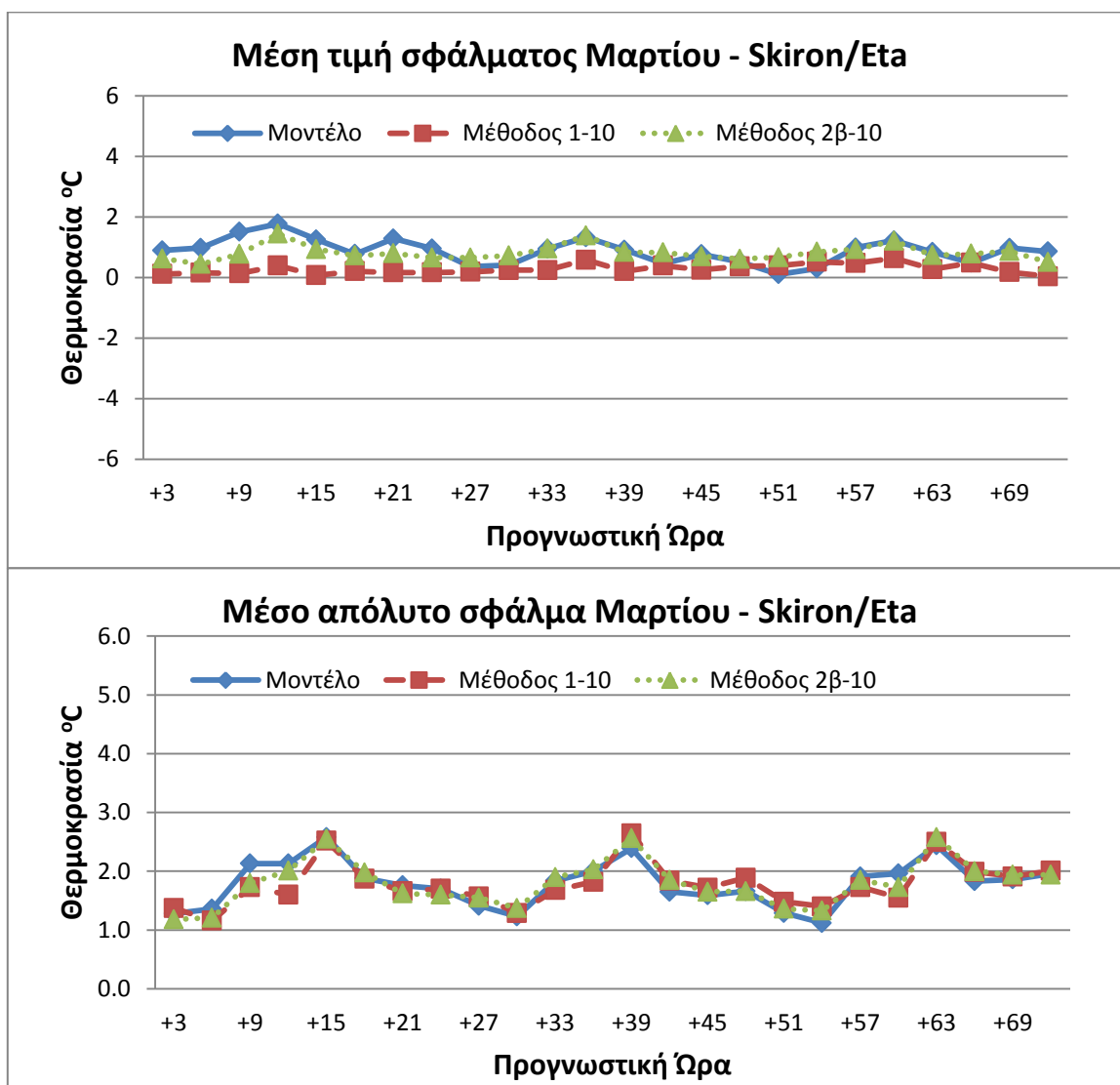
Σχήμα 21: Μέσο μηνιαίο απόλυτο σφάλμα της πρόγνωσης της θερμοκρασίας κάθε μοντέλου πριν και μετά τη χρήση του φίλτρου Kalman.



Σχήμα 22: Μέση τιμή σφάλματος και μέσο απόλυτο σφάλμα για το μήνα Ιούλιο ανά προγνωστική ώρα από το μοντέλο GFS.

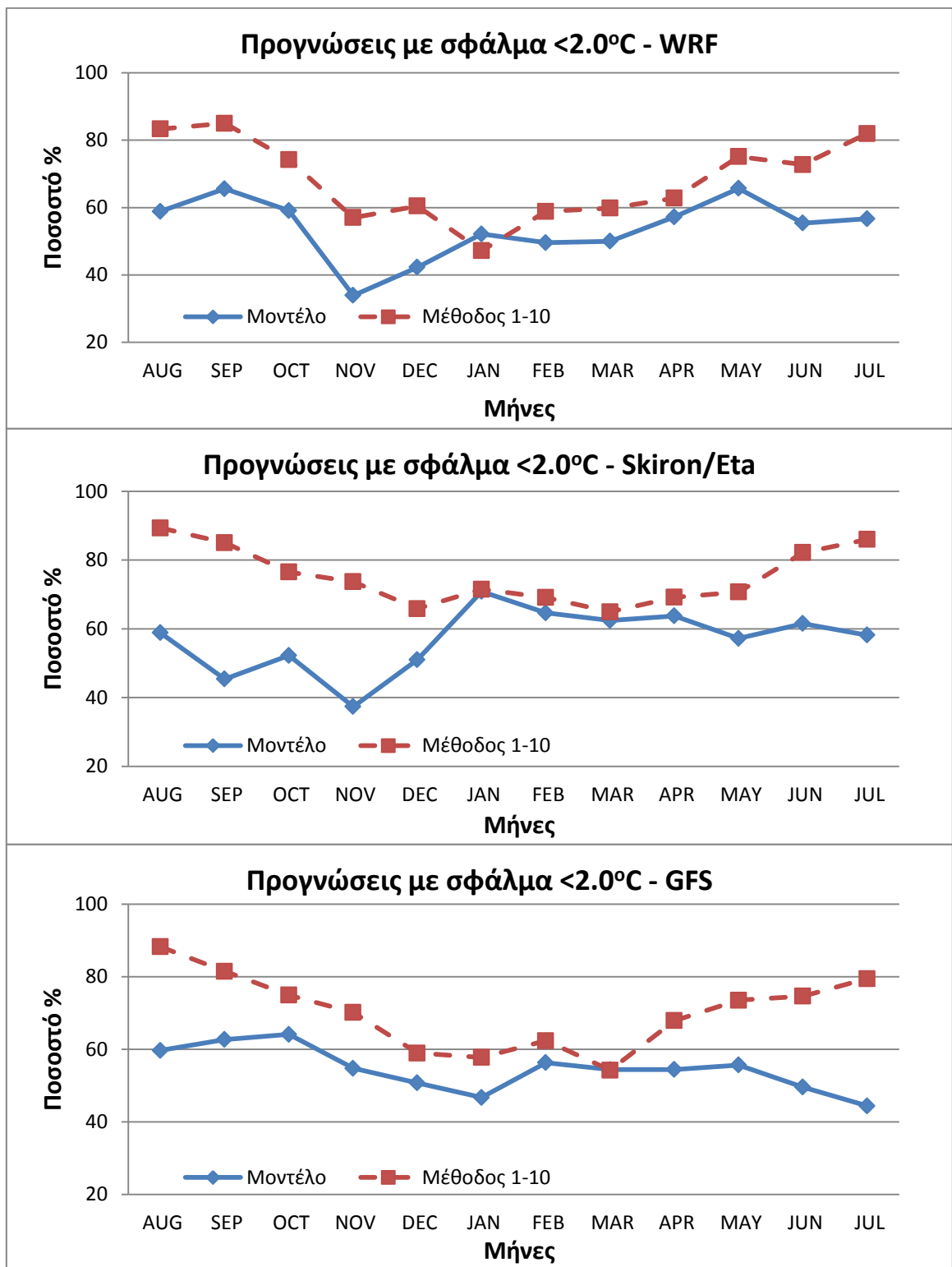
Η μεγαλύτερη μηνιαία μείωση του απόλυτου σφάλματος παρατηρείται το μήνα Ιούλιο στο μοντέλο GFS. Κατά τη διάρκεια αυτού του μήνα το απόλυτο σφάλμα που παρουσίαζε το μοντέλο ήταν 2.5°C και αυτό μειώθηκε στους 1.2°C μετά τη χρήση του φίλτρου με τη μέθοδο 1-10. Ωστόσο, αυτή η μείωση όπως φαίνεται και από το Σχήμα 22 δεν είναι ίδια για όλες τις προγνωστικές ώρες. Αντίθετα σε κάποιες από αυτές η μείωση πλησιάζει ή και ξεπερνά τους 4.0°C (6^η, 30^η και 54^η) και σε άλλες είναι σχεδόν μηδενική (21^η, 45^η και 69^η). Το μέγεθος της μείωσης εξαρτάται άμεσα από το μέγεθος του σφάλματος που υπάρχει. Όταν η μέση τιμή του

σφάλματος είναι κοντά στο μηδέν, τότε δεν παρατηρείται μεγάλη μείωση του απόλυτου σφάλματος. Στο Σχήμα 23 φαίνονται οι αντίστοιχες τιμές για το μήνα Μάρτιο του μοντέλου Skiron/Eta το οποίο παρουσίασε σχεδόν μηδενική μείωση του απόλυτου σφάλματος. Όπως φαίνεται, όλες οι προγνωστικές ώρες παρουσιάζουν μέσο σφάλμα μικρότερο από 2.0°C και οι περισσότερες μικρότερο από 1.0°C. Έτσι το αντίστοιχο μέσο απόλυτο σφάλμα παρουσιάζει μικρή μεταβολή, με μικρή μείωση σε κάποιες προγνωστικές ώρες και μικρή αύξηση σε κάποιες άλλες, διατηρώντας την τιμή του αμετάβλητη στις περισσότερες.



Σχήμα 23: Μέση τιμή σφάλματος και μέσο απόλυτο σφάλμα για το μήνα Μάρτιο ανά προγνωστική ώρα από το μοντέλο Skiron/Eta.

Στο Σχήμα 24 φαίνεται για κάθε μοντέλο το ποσοστό προγνώσεων θερμοκρασίας με μέσο απόλυτο σφάλμα μικρότερο από 2.0°C για κάθε μοντέλο πριν και μετά την εφαρμογή του φίλτρου Kalman με τη μέθοδο 1-10. Η τιμή των 2.0°C επιλέχθηκε γιατί όταν σε μία πρόγνωση θερμοκρασίας συμβαίνει το απόλυτο σφάλμα να είναι μικρότερο από αυτή την τιμή, τότε η πρόγνωση θεωρείται επιτυχημένη σύμφωνα με το Ευρωπαϊκό Κέντρο Μεσοπρόθεσμων Προγνώσεων (ECMWF) (Galanis, 2002). Σε όλους τους μήνες σε κάθε μοντέλο παρατηρήθηκε αύξηση του ποσοστού αυτού εκτός από τον Ιανουάριο στο μοντέλο WRF. Σε αυτόν τον μήνα υπάρχει μείωση από το 52% στο 47%. Ωστόσο για το WRF πριν την εφαρμογή του φίλτρου υπήρχαν μόνο δύο μήνες με ποσοστό μεγαλύτερο από 60% (Σεπτέμβριος και Μάιος και οι δύο με 65%), ενώ μετά από αυτή μόλις τρεις έχουν ποσοστό μικρότερο από 60%, με το Σεπτέμβριο να παρουσιάζει και πάλι το μεγαλύτερο ποσοστό (85%). Η εφαρμογή του φίλτρου αύξησε την μέση τιμή του ποσοστού των δώδεκα αυτών μηνών από το 54% στο 68%. Για το μοντέλο Skiron/Eta το ποσοστό αυτό παρουσίαζε μέγιστο τον Ιανουάριο με τιμή 71% και ελάχιστη το Νοέμβριο με τιμή 38%. Η μικρότερη αύξηση παρατηρήθηκε τον Ιανουάριο με το ποσοστό να φτάνει το 72%, ενώ η μεγαλύτερη το Σεπτέμβριο, όπου από 45% έφτασε το 85%. Πριν την εφαρμογή του φίλτρου τα μεγαλύτερα ποσοστά παρατηρούνταν τους μήνες Ιανουάριο, Φεβρουάριο και Μάρτιο, ενώ μετά από αυτή σε αυτούς τους μήνες παρατηρούνται τα μικρότερα και τα μεγαλύτερα δίνονται πλέον από τους θερινούς μήνες. Αυτό το αποτέλεσμα είναι σύμφωνο με τη μεγαλύτερη μείωση του απόλυτου σφάλματος τους θερινούς μήνες και την παρατήρηση ότι μετά την εφαρμογή του φίλτρου Kalman τις μεγαλύτερες τιμές μέσου μηνιαίου απόλυτου σφάλματος παρουσιάζουν να οι χειμερινοί μήνες. Η μέση τιμή του ποσοστού για το δωδεκάμηνο αυξάνεται από 57% σε 75% και είναι η μεγαλύτερη αύξηση που παρατηρείται ανάμεσα στα τρία μοντέλα. Τέλος για το μοντέλο GFS η μεγαλύτερη αύξηση παρατηρείται τον Ιούλιο από 44% σε 79% που είναι και ο μήνας με τη μεγαλύτερη μείωση του απόλυτου σφάλματος ενώ αντίθετα παραμένει ίδιο το ποσοστό για το μήνα Μάρτιο (54%), ο οποίος είναι ο μήνας που παρουσιάζει το μέγιστο απόλυτο σφάλμα μετά την εφαρμογή του φίλτρου. Μόλις δύο μήνες παρουσιάζαν ποσοστό μεγαλύτερο από 60% πριν την εφαρμογή του φίλτρου, αλλά μόλις τρεις παρουσιάζουν μικρότερο από 60% μετά από αυτή. Για το έτος που εξετάζεται η συνολική αύξηση είναι από το 54% στο 70%.



Σχήμα 24: Ποσοστό προγνώσεων θερμοκρασίας με μέσο απόλυτο σφάλμα μικρότερο από 2.0°C για κάθε μοντέλο πριν και μετά την εφαρμογή του φίλτρου Kalman.

5.3 Συμπεράσματα

Με τη βοήθεια του φίλτρου Kalman επιχειρήθηκε μείωση του σφάλματος της πρόγνωσης της θερμοκρασίας ενός αριθμητικού μοντέλου. Εφαρμόστηκαν δύο παρόμοιες μορφές του φίλτρου. Στην πρώτη περίπτωση χρησιμοποιήθηκε μόνο μία παράμετρος πράγμα που οδήγησε σε μια αρκετά απλή μορφή του φίλτρου. Στην δεύτερη περίπτωση έγινε η υπόθεση ότι το σφάλμα της πρόγνωσης σε μία συγκεκριμένη ημέρα μπορεί να περιγραφεί σαν ένα πολυώνυμο και χρησιμοποιήθηκαν ένα μηδενικού βαθμού, ένα πρωτοβάθμιο και ένα δευτεροβάθμιο πολυώνυμο. Τα τρία αριθμητικά μοντέλα πρόγνωσης καιρού που χρησιμοποιήθηκαν, ήταν το Weather Research and Forecasting - WRF, το Skiron/Eta και το Global Forecasting System – GFS, ενώ παράλληλα χρησιμοποιήθηκαν πραγματικές τιμές θερμοκρασίας, από παρατηρήσεις, από τον μετεωρολογικό σταθμό του αεροδρομίου της Θεσσαλονίκης. Τα δεδομένα καλύπτουν την περίοδο από τον Ιούνιο 2011 ως τον Ιούλιο του 2012, ενώ η στατιστική μελέτη των σφαλμάτων έγινε για το τελευταίο δωδεκάμηνο ώστε να μην υπάρχει καμία εξάρτηση από τις αρχικές συνθήκες.

Η εφαρμογή του φίλτρου Kalman σε προγνωστικές τιμές θερμοκρασίας έδειξε για άλλη μια φορά ότι μπορεί να βελτιώσει την πρόγνωση της μεταβλητής αυτής. Για τη θερμοκρασία σε απόσταση 2m από την επιφάνεια του εδάφους φαίνεται ότι αρκεί μια απλή εκτίμηση του επόμενου σφάλματος, βασισμένη σε δεδομένα από τις προηγούμενες 10 ημέρες. Τα αποτελέσματα δείχνουν να είναι παρόμοια, ανεξάρτητα από το μοντέλο στο οποίο εφαρμόζεται η μέθοδος και ανεξάρτητα από το αν η αρχική πρόγνωση του μοντέλου υπερεκτιμά ή υποεκτιμά την πραγματική τιμή.

Ο πρώτος στόχος που είναι η μείωση της μέσης τιμή του σφάλματος επιτυγχάνεται απόλυτα, αφού σε κάθε προγνωστική - και στα τρία μοντέλα - αυτή η τιμή βρίσκεται πολύ κοντά στο μηδέν. Μείωση της απόλυτης τιμής του σφάλματος παρατηρείται σε ένα ποσοστό των προγνώσεων από 61% ως 64% με τη εφαρμογή του φίλτρου. Υπάρχουν δηλαδή περιπτώσεις στις οποίες το σφάλμα του μοντέλου ήταν μικρότερο πριν την εφαρμογή του φίλτρου. Έτσι ο

δεύτερος στόχος που είναι η μείωση του απόλυτου σφάλματος επιτυγχάνεται όπως φαίνεται σε ένα μεγάλο ποσοστό, όχι όμως απόλυτα.

Σε περιόδους που η μέση τιμή του σφάλματος είναι μεγάλη, τότε παρατηρείται μείωση του απόλυτου σφάλματος που κάποιες φορές ξεπερνά τον 1.0°C . Όταν όμως η μέση τιμή του σφάλματος είναι κοντά στο μηδέν πριν την εφαρμογή του φίλτρου, τότε φαίνεται ότι η μείωση του απόλυτου σφάλματος είναι περιορισμένη. Σε ετήσια βάση φαίνεται ότι όταν σε μια προγνωστική ώρα η μέση τιμή του σφάλματος βρίσκεται στο διάστημα $(-0.6^{\circ}\text{C}, 0.6^{\circ}\text{C})$, τότε δεν παρατηρείται παρά μικρή μείωση του απόλυτου σφάλματος. Ακόμη και σε μήνες όπου υπάρχει μεγάλη μείωση του σφάλματος, αυτή δεν είναι ίδια σε όλες τις προγνωστικές ώρες, αλλά και εδώ όταν το σφάλμα είναι μεγάλο παρατηρείται μείωση αυτού, αλλά όταν είναι μικρό μένει σχεδόν αμετάβλητο.

Η εφαρμογή του φίλτρου προκαλεί αλλαγή της ετήσιας κύμανσης του σφάλματος μετακινώντας το μέγιστο σε μήνες που προηγουμένως είχαν μικρές τιμές και το ελάχιστο σε μήνες που είχαν μεγάλες τιμές. Οι χειμερινοί μήνες είναι αυτοί με το μέγιστο απόλυτο σφάλμα, πράγμα που πριν την εφαρμογή του φίλτρου δεν συνέβαινε πάντα. Αντίθετα οι θερινοί μήνες είναι αυτοί με το ελάχιστο σφάλμα πλέον, ενώ ανάμεσα στους φθινοπωρινούς και τους εαρινούς οι πρώτοι παρουσιάζουν μικρότερο σφάλμα.

Η πιθανότητα μια πρόγνωση να παρουσιάσει σφάλμα μικρότερο από 2.0°C κατά απόλυτη τιμή αυξάνεται με την εφαρμογή του φίλτρου από περίπου 55% στο 70% ή και περισσότερο, με τους θερινούς μήνες να παρουσιάζουν τη μεγαλύτερη αύξηση.

Από τα παραπάνω διαφαίνεται ότι η εφαρμογή του φίλτρου με τη βοήθεια μιας μικρής χρονοσειράς δεδομένων είναι ικανή να δώσει καλύτερες προγνώσεις θερμοκρασίας, αφού ένα τμήμα του σφάλματος που παρατηρείται συστηματικά εντοπίζεται και αφαιρείται από την τιμή της πρόγνωσης. Το φίλτρο είναι έτσι ορισμένο ώστε η εκτίμηση που εξάγει να είναι αυτή με τη μεγαλύτερη πιθανότητα εμφάνισης, με αποτέλεσμα σε ένα βάθος χρόνου να καταφέρνει να εκμηδενίσει τη μέση τιμή του σφάλματος.

Τα πλεονεκτήματα της μεθόδου είναι το γεγονός ότι αποτελεί μια σχετικά απλή μέθοδο, η οποία δεν απαιτεί μεγάλο όγκο δεδομένων, προσαρμόζεται εύκολα σε αλλαγές του μοντέλου και δεν επιβαρύνει την πρόγνωση με υπολογιστικό κόστος. Δίνει έτσι μια εύκολη αλλά ουσιαστική και χρήσιμη βοήθεια στο πρόβλημα της σωστής πρόγνωσης της τιμής μιας μεταβλητής.

ΠΕΡΙΛΗΨΗ

Ένα αριθμητικό μοντέλο πρόγνωσης καιρού παρουσιάζει σφάλματα κατά την διαδικασία προσομοίωσης των φαινομένων που λαμβάνουν χώρα στην ατμόσφαιρα. Για αυτόν τον λόγο έχουν αναπτυχθεί μέθοδοι με τις οποίες σε ένα μεταγενέστερο στάδιο επιτυγχάνεται μείωση των σφαλμάτων και κατά συνέπεια βελτίωση της πρόγνωσης. Σε αυτήν την εργασία η μέθοδος που επιλέχθηκε για το σκοπό αυτό είναι το φίλτρο Kalman. Ανάμεσα στα πλεονεκτήματα που διαθέτει η μέθοδος αυτή είναι το γεγονός ότι αποτελεί μια σχετικά απλή μέθοδος, η οποία δεν απαιτεί μεγάλο όγκο δεδομένων, προσαρμόζεται εύκολα σε αλλαγές του μοντέλου και δεν επιβαρύνει την πρόγνωση με υπολογιστικό κόστος. Η εφαρμογή της μεθόδου έγινε σε προγνώσεις τιμών θερμοκρασίας 2m από την επιφάνεια του εδάφους από τρία αριθμητικά μοντέλα πρόγνωσης καιρού το Weather Research and Forecasting - WRF, το Skiron/Eta και το Global Forecasting System - GFS ενώ χρησιμοποιήθηκαν πραγματικές τιμές θερμοκρασίας από παρατηρήσεις από τον μετεωρολογικό σταθμό του αεροδρομίου της Θεσσαλονίκης. Τα δεδομένα καλύπτουν την περίοδο από τον Ιούνιο 2011 ως τον Ιούλιο του 2012, ενώ η στατιστική μελέτη των σφαλμάτων έγινε για το τελευταίο δωδεκάμηνο ώστε να μην υπάρχει καμία εξάρτηση από τις αρχικές συνθήκες. Τα αποτελέσματα έδειξαν ότι είναι αρκετή μία απλή εκτίμηση του σφάλματος της επόμενης πρόγνωσης, με χρήση δεδομένων από τις προηγούμενες 10 προγνώσεις ώστε να παρατηρηθεί μια σημαντική μείωση του σφάλματος. Αξιοσημείωτο είναι το γεγονός ότι μετά την εφαρμογή του φίλτρου άλλαξε η ετήσια κύμανση του σφάλματος της πρόγνωσης και στα τρία μοντέλα δίνοντας μια παρόμοια εικόνα σύμφωνα με την οποία οι χειμερινοί μήνες διατηρούσαν υψηλό μέσο απόλυτο σφάλμα, ενώ το αντίστοιχο σφάλμα των θερινών μειωνόταν σε ένα μεγάλο βαθμό, δείχνοντας ότι το καλοκαίρι το σφάλμα των μοντέλων είναι κατά το μεγαλύτερο μέρος του συστηματικό.

ABSTRACT

Numerical weather prediction models are known to have systematic errors when they simulate the motion of the atmosphere especially in surface variables such as temperature at 2m. In order to make predictions more accurate and reduce these errors a bulk of methods have been developed. One of them is Kalman filtering designed by R.E. Kalman back in 1960 as a solution to the state space Wiener problem. This method has some very useful advantages. First of all Kalman filtering is a simple method which require a very small database. In addition it can adapt very fast to any necessary changes of the model and above all it is a computationally cheap method. Kalman filtering applied to temperature forecasts (2m) obtained by three numerical weather prediction models (Weather Research and Forecasting - WRF, Skiron/Eta and Global Forecasting System - GFS). Observations of temperature at 2m made available from the weather station located at the airport of Thessaloniki, Greece. The data refers to the period from June 2011 to July 2012 but statistical analysis made only for the last twelve months of this period so as to be independent from the initial values that are used. It is shown that a simple estimation of the forthcoming predicted value made using data from the 10 past predictions is sufficient to significantly reduce the mean error. Moreover the annual fluctuation of the error changed when the filter applied in all three models. Winter's months maintained the original error while on the other hand summer's months reduced it, so that max values transferred from summer to winter, obviously showing, that despite its magnitude the error of the forecasts during summer remained systematic. In any case the method showed great skill when the error was biased making this unbiased and reducing the mean absolute error but when the model's error was unbiased then filter's performance was much poorer.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Anadranistakis, M., Lagouvardos, K., Kotroni, V., Skouras, K., 2002. Combination of Kalman filter and an empirical method for the correction of near-surface temperature forecasts: application over Greece. *Geophysical Research Letters*, 29 (16), 1776–1779.

Anadranistakis, M., Lagouvardos, K., Kotroni, V., Elefteriadis, H., 2004. Correcting temperature and humidity forecasts using Kalman Filtering: potential for agricultural protection in Northern Greece. *Atmospheric Research*, 71, 115–125.

Betts, A. K., 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quarterly Journal of the Royal Meteorological Society*, 112, 677–691.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, and arctic air-mass data sets. *Quarterly Journal of the Royal Meteorological Society*, 112, 693–709.

Charney, J.G., Fjortoff, R., von Neumann, J., 1950. Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, 2, 237-254.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land-surface/ hydrology model with the Penn State/ NCAR MM5 modeling system. Part I: Model description and implementation. *Monthly Weather Review*, 129, 569-585.

Crochet, P., 2004. Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. *Meteorological Applications*, 11, 173–187.

Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., Stull, R., 2006. Ozone ensemble forecasts: 2. A Kalman-filter predictor bias correction. *Journal of Geophysical Research*, 111, D05308, doi:10.1029/2005JD006311.

Delle Monache L., Grell, G. A., McKeen, S., Wilczak, J., Pagowski, M. O., Peckham, S., Stull R., McHenry, J., McQueen, J., 2008. A Kalman-filter bias correction method applied to deterministic, ensemble averaged and probabilistic forecasts of surface ozone. *Tellus*, 60B, 238-249.

Ferrier, B. S., Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. *Preprints, 15th Conference on Numerical Weather Prediction*, San Antonio, TX, American Meteorology Society, 280-283.

Fels, S. B. and M. D. Schwarzkopf, 1975: The Simplified Exchange Approximation: A New Method for Radiative Transfer Calculations, *Journal of Atmospheric Sciences*, 32, 1475-1488.

Galanis, G., Anadranistakis, M., 2002. A one-dimensional Kalman filter for the correction of near surface temperature forecasts. *Meteorological Applications*, 9, 437–441.

Galanis, G., Louka, P., Katsafados, P., Kallos, G., Pytharoulis, I., 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. *Annales Geophysicae*, 24, 2451–2460.

Glahn, H. R., Lowry, D. A., 1972. The use of Model Output Statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, 11, 1203–1211.

Hartmann, D. L., 2013. ATM S 552 - Objective Analysis. *University of Washington*. pp 265.

Homleid, M., 1995. Diurnal corrections of short-term surface temperature forecasts using the Kalman filter. *Weather Forecasting*, 10, 689–707.

Homleid, M., 2004: Weather dependent statistical adaptation of 2 meter temperature forecasts using regression methods and Kalman filter. Report No.6/2004. *Norwegian Meteorological Institute*. ISSN 1503-8025.

Houtekamer, P. L., Herschel L. Mitchell, 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, 126, 796–811.

Janjic, Z. I., 1990: The step-mountain coordinate: physical package, *Monthly Weather Review*, 118, 1429-1443.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review*, 122, 927–945.

Janjic, Z. I., 1996: The surface layer in the NCEP Eta Model. *Eleventh Conference on Numerical Weather Prediction*, Norfolk, VA, 19-23 August, American Meteorology Society, Boston, MA, 354-355.

Janjic, Z. I., 2000: Comments on “Development and Evaluation of a Convection Scheme for Use in Climate Models”, *Journal of Atmospheric Sciences*, 57, p. 3686.

Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso model, *NCEP Office Note*, No. 437, pp 61.

Jingying, F., Dong, J., Yaohuan, H., Dafang, Z., Yong W., 2013. A Kalman Filter-Based Method for Reconstructing GMS-5 Global Solar Radiation by Introduction of In Situ Data Energies. *Institute of Geographic Sciences and Natural Resources Research, CAS*, 6, 2804-2818.

Kallos, G., Nickovic, S., Sakellaridis, G., Anadranistakis, E. et al., 1997. The regional weather forecasting system Skiron: An overview. *Proc. of the Symposium on Regional Weather Prediction on Parallel Computer Environments*, Athens, Greece, pp 9.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82, 34–45.

Kalman, R.E. and Bucy R.S. (1961). New results in linear filtering and prediction problems, Transactions of the ASME. Series D, *Journal of Basic Engineering*, 83, 95–108.

Καρακώστας, Θ., 2012. Σημειώσεις Γενικής και Δυναμικής Μετεωρολογίας. Θεσσαλονίκη, σελ 53.

Katsafados, P., Papadopoulos, A., Mavromatidis, E., Gikas, N., 2011. Quantitative verification statistics of WRF predictions over the Mediterranean region. *12th Annual WRF Users' Event*, 20-24 June 2011, Boulder, CO.

Klein, W. H., Lewis, F., 1970. Computer forecasts of maximum and minimum temperatures. *Journal of Applied Meteorology*, 9, 350–359.

Kotroni, V., Lagouvardos, K., 2004. Evaluation of MM5 High-Resolution Real-Time Forecasts over the Urban Area of Athens, Greece. *Journal of Applied Meteorology*, 43, 1666-1678.

Lacanette, K., 1991. A Basic Introduction to Filters Active, Passive, and Switched-Capacitor. *National Semiconductor Application*, Note 779. pp 22.

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. *Journal of Wind Engineering and Industrial Aerodynamics*, 96, 2348– 2362.

Lynch, P., 1993. Richardson's forecast factory: the \$64,000 question, *The Meteorological Magazine*, Vol 122, 69–70.

Lynch, P., 1999. Richardson's Marvelous Forecast, pp 61-73 in *The Life Cycles of Extratropical Cyclones*, M.A. Shapiro and S. Gronas, Eds., *American Meteorological Society*, Boston, 355pp.

Mackenzie D., 2003. Ensemble Kalman Filters Bring Weather Models Up to Date. *SIAM News*, 36, Number 8, pp 4.

Maybeck, P. S., 1979. Stochastic Models, Estimation, and Control, Vol 1, *Academic Press*, pp 411.

Meinhold, R., Singpurwall, N., 1983. Understanding the Kalman Filter. *The American Statistician*, 37 (2), 123-127.

Mellor, G.L., Yamada, T., 1982. Development of a turbulence closure model for geophysical fluids problems. *Review of Geophysics and Space Physics*, 20, 851-875.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long-wave. *Journal of Geophysical Research*, 102 (D14), 16663-16682.

Monin, A.S. and A.M. Obukhov, 1954: Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contributions of the Geophysical Institute of the Slovak Academy of Sciences, USSR*, (151), 163-187.

Μουστακίδης, Γ., 2003. Βασικές Τεχνικές Ψηφιακής Επεξεργασίας Σημάτων. *Εκδ. Τζιόλας*, σελ. 160.

Naifang Beia, Fuqing Zhanga, 2007. Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the Mei-Yu front of China. *Quarterly Journal of the Royal Meteorological Society*, 133, 83–99.

Πυθαρούλης, Ι., 2009. Υψηλής Ανάλυσης Επιχειρησιακές Αριθμητικές Προγνώσεις Καιρού για Αγρομετεωρολογικές Εφαρμογές. 6^ο Πανελλήνιο Συνέδριο Εταιρείας Γεωργικών Μηχανικών Ελλάδος "Η Γεωργική Μηχανική και η Μηχανική βιοσυστημάτων στην εποχή των βιοκαυσίμων και των κλιματικών αλλαγών", Θεσσαλονίκη 8-10 Οκτωβρίου 2009, σελ. 75-82.

Persson, A., 1991. Kalman filtering – a new approach to adaptive statistical interpretation of numerical meteorological forecasts. *Lectures and papers presented at the WMO Training Workshop on the Interpretation of NWP Products in Terms of Local Weather Phenomena and Their Verification*, Wageningen, the Netherlands. WMO PSMP Report Series no. 34, XX-27-XX-32.

Reid, I., 2001 (a). Estimation I. *Hilary Term*, pp 36.

Reid, I., 2001 (b). Estimation II. *Hilary Term*, pp 44.

Richardson, L. F., 1922. Weather Prediction by Numerical Process. *Cambridge University Press*, 236 pp.

Schwarzkopf, M. D., and S. B. Fels, 1991: The simplified exchange method revisited – An accurate, rapid method for computation of infrared cooling rates and fluxes. *Journal of Geophysical Research*, 96 (D5), 9075-9096.

Simonsen, C., 1991. Self adaptive model output statistics based on Kalman filtering. *Lectures and papers presented at the WMO Training Workshop on the Interpretation of NWP Products in Terms of Local Weather Phenomena and Their Verification*, Wageningen, the Netherlands. WMO PSMP Report Series no. 34, XX-33-XX-37.

Smith S. W., 1999. The Scientist and Engineer's Guide to Digital Signal Processing. *California Technical Publishing*, pp 640.

Τίγκελης, Ι., Φραντζεσκάκης, Δ., 2008. Σημειώσεις του μαθήματος Εισαγωγή στα Συστήματα τηλεπικοινωνιών. *Πανεπιστήμιο Αθηνών*, σελ. 205.

Welch G. and Bishop G., 2004. An Introduction to the Kalman Filter. *University of North Carolina, Department of Computer Science, SIGGRAPH 2001*, TR 95-041, pp. 1-16.

Wilks, D.S., 1995. Statistical Methods in Atmospheric Sciences, *Academic Press*, 464 pp.

Zielonka, J.R., Gaudet, B.J., Seaman, N.L., Stauffer, D.R., Deng, A., Hunter, G.K., 2008. *WRF and MM5 realtime system statistical comparisons using the model evaluation toolkit (MET)*. 9th WRF Users Workshop, June 23 - 27, 2008, Boulder, CO, 7 pp.

Zhao, Q., and Carr, F. H., 1997. A prognostic cloud scheme for operational NWP models. *Monthly Weather Review*, 125, 1931–1953.

ZhiXin, H., Zheng, J.Y., GuoFeng, W., XueZhen, Z., QuanSheng, G., 2010. 1876–1878 severe drought in North China: Facts, impacts and climatic background. *Chinese Science Bulletin*, Vol. 55, Issue 26, 3001-3007.