

**Εθνικό Μετσόβιο Πολυτεχνείο
Τμήμα Πολιτικών Μηχανικών
Τομέας Υδατικών Πόρων**

**Διερεύνηση της δυνατότητας πρόγνωσης της εξέλιξης
μετεωρολογικών μεταβλητών με τεχνικές δυναμικών
συστημάτων**

Διπλωματική εργασία του φοιτητή Δημήτρη Πούλου.

**Επιβλέπων : Δημήτρης Κουτσογιάννης
επίκ. Καθηγητής Ε.Μ.Π.**

Ιούνιος 1998

ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή γίνεται μία προσπάθεια πρόγνωσης της εξέλιξης μετεωρολογικών μεταβλητών χρησιμοποιώντας τη θεωρία των δυναμικών συστημάτων. Ειδικότερα εφαρμόζεται η σύγχρονη μέθοδος των αναλόγων και ερευνώνται τα όρια της ακρίβειας πρόγνωσης με την εν λόγω μέθοδο.

Στην εισαγωγή γίνεται μία παρουσίαση βασικών εννοιών των δυναμικών συστημάτων. Μας απασχολούν έννοιες που θα χρησιμοποιηθούν στη συνέχεια και συγκεκριμένα θέματα ισορροπίας, αυτονομίας, fractals και διαστάσεών τους, απόκλισης τροχών, που παρουσιάζονται μέσω επιλεγμένων θεωρημάτων και παραδειγμάτων.

Στο πρώτο κεφάλαιο παρουσιάζεται η μέθοδος των αναλόγων, οι έννοιες και οι παράμετροι της και παρουσιάζεται ο αλγόριθμος και οι μεταβλητές του. Επίσης εξετάζεται το θέμα των σφαλμάτων και τίθεται το θέμα της χρονικής κλίμακας των μεγεθών. Γίνεται σχολιασμός στη δυνατότητα βελτιστοποίησης της πρόβλεψης και βελτίωσης της κατανόησης φυσικών δυναμικών συστημάτων.

Στο δεύτερο κεφάλαιο περιέχεται η εφαρμογή της μεθόδου των αναλόγων σε αρχεία μετρήσεων ατμοσφαιρικών μεγεθών, γίνεται προσπάθεια για βελτιστοποίηση της πρόγνωσης και ελέγχονται και σχολιάζονται οι περιορισμοί.

Στην εργασία περιλαμβάνεται παράτημα που περιέχει τον κώδικα του προγράμματος και το format βοηθητικών αρχείων.

SUMMARY

In this thesis we attempt to forecast the evolution of meteorological variables, using the theory on dynamical systems. More specifically, we use the method of the analogues and assess the limits of a forecast using that method.

In the introduction we present the basic concepts of dynamical systems that we use in this thesis. We present issues like stability, autonomy, fractal dimension, orbit divergence through selected theorems and examples.

In the first chapter we present the method of the analogues, its basic concepts and parameters and the algorithm we use with its variables. We also investigate the issues of prediction errors with regard to the time scale. We also discuss the feasibility of a good forecast and the improvement of the understanding of physical dynamical systems.

In the second chapter we present the application of the method of analogues on the forecast of atmospheric variables and discuss the limits of the forecasts and ways to improve them.

In addition, there is an appendix with the code of the program we used and the format of assistant computer archives.

Πρόλογος

Υπάρχει η άποψη ότι γνώση, πρόβλεψη και επιβίωση είναι πολύ κοντινές και ισχυρά αλληλοεξαρτόμενες έννοιες. Τα δυναμικά συστήματα είναι μια πηγή. Η ζωή τα χρησιμοποιεί. Τα φυτά, οι οργανισμοί, ο εγκέφαλος έχουν πολύπλοκη-fractal δομή. Αυτή τα βοηθά να προσαρμοστούν σε ανόμοια και μεταβαλλόμενα περιβάλλοντα ή αλλιώς προβλέπει ανόμοια και μεταβαλλόμενα περιβάλλοντα. Η ανθρώπινη γνώση ακολουθεί τη φύση, ανακαλύπτει τα υπάρχοντα. Εδώ τίθεται το ερώτημα: έχει η φύση βρει την τέλεια πρόβλεψη; Τίποτα δε ζει για πάντα. Υπάρχει όμως ο μετασχηματισμός, η λογική των συστημάτων. Μόνο οι ιδέες είναι αιώνιες και αναλλοίωτες. Ισως η τέλεια πρόγνωση να βρίσκεται ‘μετά τα Φυσικά’.

Θέλω να ευχαριστήσω ιδιαίτερα τον επίκουρο καθηγητή Δημήτρη Κουτσογάννη για τη βοήθεια και τη στήριξή του στην ολοκλήρωση αυτής της εργασίας. Επίσης τον Νίκο Μανδέλο για τη βοήθειά του στη συγγραφή του κώδικα και τις συζητήσεις, τους Νίκο Μαμάση και Αντώνη Χριστοφίδη για την παροχή των μετρήσεων και το χρόνο που μου διέθεσαν και επιπλέον όσον αφορά τον Νίκο Μαμάση για τις υποδείξεις του στο κείμενο αυτής της εργασίας.

ΠΕΡΙΕΧΟΜΕΝΑ	σελίδα
ΕΙΣΑΓΩΓΗ.....	5
α. Ισορροπία.....	6
β. Διατήρηση του όγκου.....	6
γ. Αυτονομία.....	7
δ. Ελκυστές και διαστάσεις.....	7
ΚΕΦΑΛΑΙΟ Α. Γενικές έννοιες από τη θεωρία χαοτικών συστημάτων.....	10
α. Τροχιές.....	10
β. Μέθοδος των αναλόγων.....	11
γ. Ο αλγόριθμος DVS.....	12
δ. Υλοποίηση του αλγόριθμου DVS.....	12
ε. Παράδειγμα.....	15
στ. Όρια πρόγνωσης.....	20
ΚΕΦΑΛΑΙΟ Β. Εφαρμογή σε χρονοσειρές μετεωρολογικών δεδομένων.....	25
α. Θερμοκρασία.....	25
β. Ατμοσφαιρική πίεση.....	32
γ. Ύγρασία.....	39
δ. Βορινή συνιστώσα ταχύτητας ανέμου.....	46
ε. Ανατολική συνιστώσα ταχύτητας ανέμου.....	52
στ Συσχέτιση προβολών της ταχύτητας του ανέμου σε άξονες.....	58
ζ Βορινή συνιστώσα ταχύτητας ανέμου (10-λεπτες μετρήσεις).....	59
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	73
ΑΝΑΦΟΡΕΣ.....	75
ΠΑΡΑΡΤΗΜΑ.....	76
α Κώδικας του αλγόριθμου DVS.....	76
β Συνεργαζόμενα αρχεία.....	88

ΕΙΣΑΓΩΓΗ

Στην εργασία αυτή ερευνάται η δυνατότητα πρόγνωσης της εξέλιξης ορισμένων μετεωρολογικών μεταβλητών χρησιμοποιώντας τη θεωρία των δυναμικών συστημάτων. Αρχικά δίδονται οι έννοιες και τα βασικά χαρακτηριστικά των δυναμικών συστημάτων. Έτσι αναπτύσσεται το θεωρητικό υπόβαθρο για την εφαρμογή της σύγχρονης μεθόδου πρόγνωσης των αναλόγων. Στη συνέχεια δίδεται ο αλγόριθμος DVS. Πρόκειται για έναν αλγόριθμο που εφαρμόζει την εν λόγω μέθοδο για ένα μεγάλο αριθμό προγνώσεων και δίνει το μέσο σφάλμα πρόγνωσης συναρτήσει των μεταβλητών του μοντέλου. Στη συνέχεια σχολιάζεται η δυνατότητα εφαρμογής αυτής της μεθόδου σε φυσικά δυναμικά συστήματα, μέσω της εξέτασης των αναμενόμενων σφαλμάτων i) σε δυναμικά συστήματα που καλύπτουν διάφορες χρονικές κλίμακες και ii) σε δυναμικά συστήματα που προτείνονται από το συγγραφέα. Αφού αναπτυχθούν τα παραπάνω θέματα, εφαρμόζεται ο αλγόριθμος σε αρχεία μετρήσεων ατμοσφαιρικών μεγεθών και ερευνάται η βελτιστοποίηση της πρόγνωσης αυτών των μεγεθών.

Ένα δυναμικό σύστημα περιγράφεται από ένα σύστημα διαφορικών εξισώσεων αν αναφέρεται σε συνεχή χρόνο, ή από ένα σύστημα εξισώσεων διαφορών αν αναφέρεται σε διακριτό χρόνο. Παράδειγμα ενός συστήματος διαφορικών εξισώσεων είναι το σύστημα του Lorenz:

$$\begin{aligned}\dot{x} &= -ax + ay \\ \dot{y} &= -xz + bx - y \\ \dot{z} &= xy - cz\end{aligned}$$

όπου τα x , y , z αποτελούν τις μεταβλητές του συστήματος, τα \dot{x} , \dot{y} , \dot{z} είναι οι παράγωγοι αυτών των μεταβλητών ως προς το χρόνο και τα a , b , c σταθερές. Ένα σύστημα εξισώσεων διαφορών είναι το:

$$\begin{aligned}x_{n+1} &= 1 - ax_n^2 + y_n \\ y_{n+1} &= bx_n\end{aligned}$$

που είναι γνωστό ως σύστημα του Hénon (Hénon map). Τα x_n , y_n είναι οι μεταβλητές του συστήματος στη θέση n της χρονοσειράς και τα a , b σταθερές.

Τα συστήματα διαφορών επιλύονται πιο εύκολα στον υπολογιστή και είναι προσφορότερα για τη μελέτη χρονοσειρών.

Τα θέματα δυναμικών συστημάτων της εισαγωγής περιγράφονται πληρέστερα από τον Tsonis (1992), από όπου και μεταφέρονται.

α. Ισορροπία

Στα σημεία στα οποία ισορροπεί ένα σύστημα (σημεία ισορροπίας) ισχύει

$$x_{n+1}^i = x_n^i$$

όπου ι είναι η διεύθυνση (δηλαδή $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$), ή για συνεχή χρόνο

$$\dot{x}^i = 0$$

Ένα γραμμικό σύστημα διαφορικών εξισώσεων

$$\dot{\vec{x}} = A\vec{x}$$

είναι ευσταθές αν όλες οι ιδιοτιμές λ του Α έχουν πραγματικό μέρος αρνητικό. Με το συμβολισμό παραγώγου διανύσματος εννοείται το διάνυσμα με συνιστώσες τις παραγώγους των συνιστωσών του αρχικού διανύσματος. Οι ιδιοτιμές λ είναι οι λύσεις της εξίσωσης:

$$DET(A - \lambda I) = 0$$

όπου $DET()$ είναι η ορίζουσα του πίνακα .

Αν το σύστημα είναι μη γραμμικό πρέπει πρώτα να προσεγγιστεί από ένα γραμμικό στη γειτονιά του σημείου ισορροπίας (με απαλοιφή των μη γραμμικών όρων) και στη συνέχεια να γίνει ο έλεγχος ευστάθειας (για μία εκτενή παρουσίαση θεμάτων ευστάθειας βλ. Χατζηδημητρίου (1988)).

β. Διατήρηση του όγκου

Ένα δυναμικό σύστημα έχει τόσες ιδιοτιμές όσους και βαθμούς ελευθερίας. Η κάθε ιδιοτιμή δείχνει την τάση γειτονικών τροχιών να συγκλίνουν ή να αποκλίνουν ως προς κάποια διεύθυνση. Αρνητική ιδιοτιμή σημαίνει σύγκλιση, θετική απόκλιση. Ένα γραμμικό σύστημα εξισώσεων διαφορών

$$\overrightarrow{x_{n+1}} = A\overrightarrow{x_n}$$

είναι διατηρητικό αν $|\det A|=1$ και μη διατηρητικό αν $|\det A|<1$. Ένα σύστημα διαφορικών εξισώσεων όμως είναι διατηρητικό αν το ίχνος του A είναι 1: $\text{trace}A=1$ και μη διατηρητικό αν $\text{trace}A<1$. Αποδεικνύεται ότι αν $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n$, τότε $\text{trace}A=\lambda_1+\lambda_2+\dots+\lambda_n$. Ετσι το άθροισμα των ιδιοτιμών καθορίζει αν το σύστημα είναι διατηρητικό.

Αν ένα σύστημα διαφορικών εξισώσεων είναι μη διατηρητικό και συγχρόνως έχει τουλάχιστον μία ιδιοτιμή θετική λέγεται χαοτικό (βλ. και Μπούντης (1988)).

γ. Αυτονομία

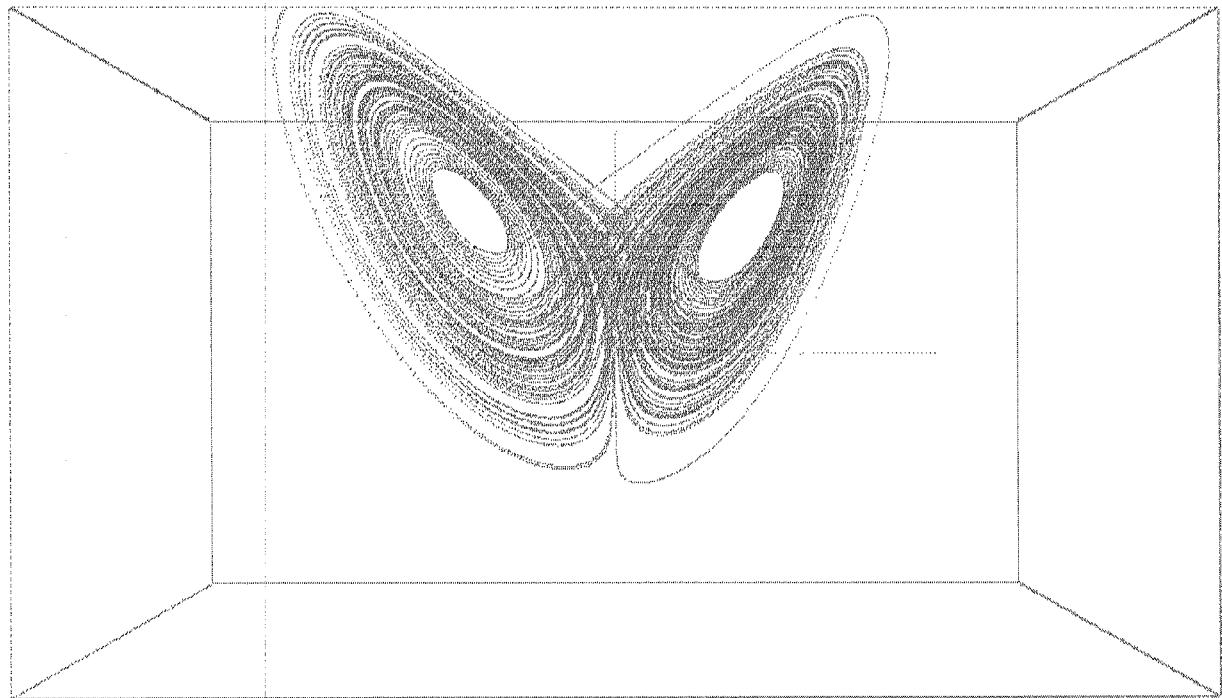
Ένα σύστημα ανεξάρτητο του χρόνου είναι αυτόνομο και δύο τροχιές του δε διασταυρώνονται ποτέ. Αν ο χρόνος αποτελεί μεταβλητή του συστήματος τότε δύο τροχιές μπορεί να διασταυρωθούν. Σε αυτήν την περίπτωση δεν μπορούμε να παρακολουθήσουμε την εξέλιξη του συστήματος ή αλλιώς έχουμε απώλεια πληροφορίας. Πρέπει τότε να προστεθεί επιπλέον η διάσταση του χρόνου στο χώρο των φάσεων (το χώρο αλληλοσυσχέτισης των μεταβλητών). Η επιπλέον εξίσωση είναι :

$$\dot{t} = 1$$

δ. Ελκυστές και διαστάσεις

Ένα μη διατηρητικό σύστημα ν διαστάσεων περιορίζεται σε ένα διαρκώς σμικρινόμενο μέρος του n -διάστατου χώρου των φάσεων. Ετσι ένα σύστημα τριών διαστάσεων διαρκώς μικραίνει τον όγκο του στο χώρο των φάσεων, ένα σύστημα δύο διαστάσεων την επιφάνειά του κ.ο.κ. Το όριο σύγκλισης καλείται ελκυστής. Ένας ελκυστής μπορεί να είναι ένα σημείο (π.χ. η θέση ηρεμίας στην οποία ισορροπεί μια φθίνουσα ταλάντωση), ένας περιοδικός κύκλος, μία 'σαμπρέλα' ή κάτι πολύ παράξενο: ένας 'παράξενος ελκυστής'. Παράξενο ελκυστή παρουσιάζουν τα χαοτικά συστήματα.

Ένας παράξενος ελκυστής έχει σχεδιαστεί στο σχήμα 0.1:



Σχήμα 0.1. Ο ελκυστής του Lorenz όπως σχεδιάσθηκε με το πρόγραμμα ‘Dynamics Numerical Explorations’ των Nusse and Yorke (1994).

Ο χώρος στον οποίο ανήκουν οι ελκυστές είναι ο χώρος των fractals. Ο χώρος των fractals είναι ένας πλήρης (complete) μετρικός χώρος. Ορίζεται για έναν χώρο n διαστάσεων και κάθε σημείο του είναι ένα συμπαγές (κλειστό και ολικά φραγμένο) υποσύνολο του n -διάστατου χώρου. Στον χώρο των fractals ο ελκυστής είναι ένα ευσταθές σημείο ισορροπίας (βλ. Barnsley (1993)).

Μία χαρακτηριστική ιδιότητα των fractals είναι ότι αποδέχονται κλασματικές διαστάσεις. Αυτή η ιδιότητα θεωρήθηκε τόσο σημαντική όσο να τους δώσει το όνομα (fraction = κλάσμα). Έχουν αναπτυχθεί μέθοδοι προσδιορισμού της διάστασης που χρησιμοποιούν αριθμητικά δεδομένα ή μετρήσεις (βλ. Grassberger(1986)). Η κλασματική διάσταση ενός ελκυστή είναι οι βαθμοί ελευθερίας του δυναμικού συστήματος μείον ο βαθμός αλληλοσυσχέτισης των βαθμών αυτών. Επομένως μπορεί να προσδιορισθεί η αλληλοσυσχέτιση των βαθμών ελευθερίας από τη διάσταση.

Έχει αποδειχθεί ότι ένας ελκυστής r διάστασης εδράζεται επαρκώς σε $2r+1$ διαστάσεις ή βαθμούς ελευθερίας (θεώρημα του Whitney). Επιπλέον έχει δειχθεί ότι η έδραση (embedding) του ελκυστή σε διάσταση $m \geq 2r+1$ διατηρεί τις τοπολογικές του ιδιότητες (θεώρημα του Takens). Η απεικόνισή του στο χώρο των φάσεων γίνεται ως συνάρτηση των $x_i, x_{i-t_1}, x_{i-t_2}, \dots, x_{i-t_n}$, χωρίς μεγάλη απόλεια πληροφορίας συνήθως. Το πλήθος των x είναι τάξης μεγέθους $2r+1$. Το σύνηθες χρονικό βήμα t (time delay) είναι μοναδιαίο. Στην εργασία αυτή θα παρουσιαστεί

ένα πρόγραμμα που, εκτός του ότι ανιχνεύει ντετερμινισμό, δίνει μία προσέγγιση της διάστασης.

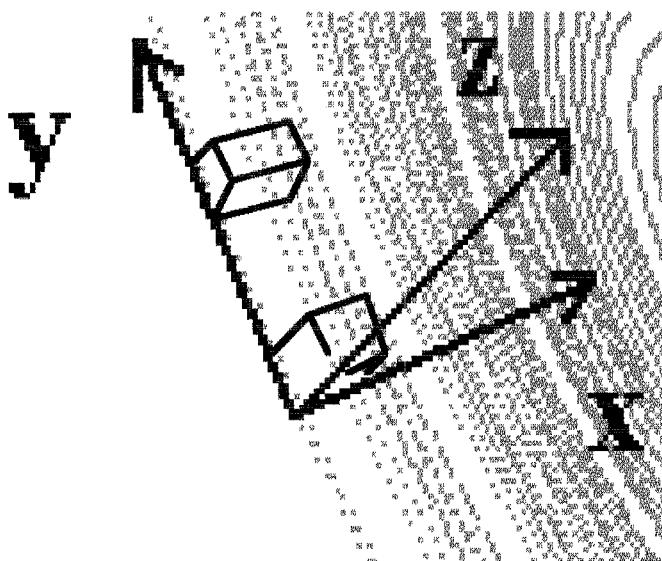
Τέλος αναφέρεται ότι ένα δυναμικό σύστημα μπορεί να έχει περισσότερους του ενός ελκυστές. Σε ποιόν ελκυστή θα ισορροπήσει το σύστημα καθορίζοντας οι αρχικές τιμές.

ΚΕΦΑΛΑΙΟ Α

Γενικές έννοιες από τη θεωρία χαοτικών συστημάτων

α. Τροχιές

Οι ιδιοτιμές ενός δυναμικού συστήματος δείχνουν τη διαστολή ή συστολή του χώρου στον οποίο εδράζεται το σύστημα. Σε κάθε ιδιοτιμή αντιστοιχεί μια διεύθυνση. Εποι αν δύο γειτονικές τροχιές στο χώρο των φάσεων συγκλίνουν ή αποκλίνουν καθορίζεται από τις ιδιοτιμές. Κατά τη διεύθυνση της αρνητικής ιδιοτιμής συγκλίνουν, της θετικής αποκλίνουν και της μηδενικής είναι παράλληλες. Αυτό φαίνεται στο τμήμα του ελκυστή του Lorenz που είναι σχεδιασμένο στο σχήμα A.1 :



Σχήμα A.1. Το γράφημα έγινε με το πρόγραμμα 'Dynamics Numerical Explorations' των Nusse and Yorke (1994).

Στο σχήμα αυτό παρατηρούμε ότι, με την προϋπόθεση ότι η φορά είναι προς το πάνω μέρος του χαρτιού, δύο σημεία του άξονα x αποκλίνουν ($\lambda_x > 0$), του y διατηρούν σταθερή απόσταση ($\lambda_y = 0$) και του z συγκλίνουν ($\lambda_z < 0$). Αυτό φαίνεται από την παραμόρφωση του στοιχειώδους όγκου. Ένα σημείο που βρίσκεται κοντά στον ελκυστή σε απόσταση z' θα διαγράψει τροχιά σύγκλισης και θα φτάσει στον ελκυστή σε άπειρο χρόνο t.

Ο παραπάνω ελκυστής εδράζεται σε τρεις διαστάσεις. Τρεις είναι και οι βαθμοί ελευθερίας του δυναμικού συστήματος. Η διάσταση του ελκυστή είναι $d \approx 1.7$ (βλ. Tsonis (1992)).

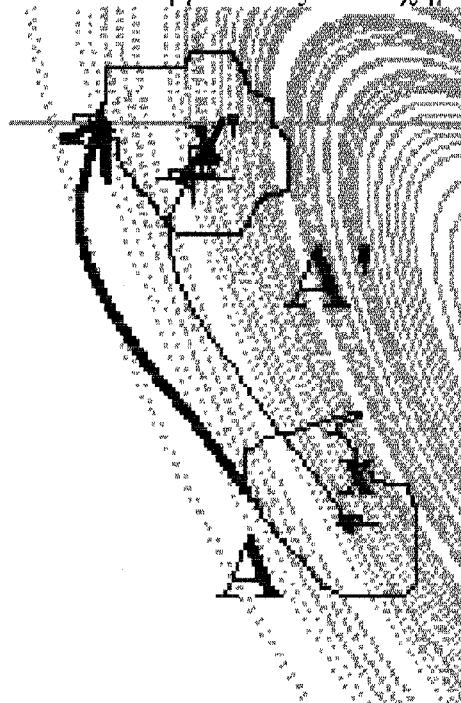
β. Μέθοδος των αναλόγων

Η μέθοδος των αναλόγων είναι μέθοδος πρόγνωσης. Στηρίζεται στην υπόθεση ότι γειτονικά σημεία στον ελκυστή παραμένουν γειτονικά ύστερα από ένα μικρό χρονικό βήμα T . Αυτή η υπόθεση στηρίζεται σε δύο έννοιες που μόνο κατά περίπτωση μπορούν να αναλυθούν: ποια σημεία μπορούν να θεωρηθούν γειτονικά και πόσο μικρό πρέπει να είναι το χρονικό βήμα. Πάντως χωρίς σε βάθος διερεύνηση αυτών των συνιστώσων η μέθοδος εργάζεται ικανοποιητικά σε μεγάλο εύρος δυναμικών συστημάτων και μεγάλο εύρος περιοχών του χώρου των φάσεών τους.

Το πρώτο βήμα της μεθόδου είναι να δημιουργηθεί ένας χώρος φάσεων από επίλυση εξισώσεων ή μετρήσεις. Στην περίπτωση που είναι γνωστό το σύστημα είναι γνωστός και ο αριθμός των διαστάσεων του χώρου (ισούται με το πλήθος των μεταβλητών του συστήματος). Στην περίπτωση όμως που πρέπει να δημιουργηθεί ο χώρος από μετρήσεις, γίνονται δοκιμές σε διάφορες διαστάσεις και επιλέγεται εκείνη που δίνει τις καλύτερες προγνώσεις. Η διάσταση αυτή λέγεται διάσταση έδρασης (embedding dimension) d και για διάσταση ελκυστή r είναι ίση περίπου με $2r+1$ (θεώρημα Takens). Όπως έχει αναφερθεί, τα δεδομένα της χρονοσειράς διατάσσονται σε d -διάστατα διανύσματα με χρονική απόσταση μεταξύ γειτονικών συνιστώσων τ (βλ. Casdagli (1991)).

Το δεύτερο βήμα είναι ο ορισμός του χρησιμοποιούμενου πλήθους γειτόνων και η εύρεσή τους.

Το τρίτο βήμα είναι η μοντελοποίηση της κίνησης των γειτόνων και η εφαρμογή του μοντέλου στην πρόβλεψη της νέας θέσης του συστήματος (βλ. Tsonis (1992)). Σχηματικά δίδεται η μέθοδος στο σχήμα A.2 :



Σχήμα A.2. Σχηματική απεικόνιση της μεθόδου των αναλόγων.

γ. Ο αλγόριθμος DVS

Ο αλγόριθμος DVS (Deterministic Versus Stochastic modeling) παρουσιάσθηκε από τους Casdagli and Weigend (1992). Πρόκειται για έναν αλγόριθμο που χρησιμοποιεί τη μέθοδο των αναλόγων και διερευνά το ντετερμινιστικό ή στοχαστικό χαρακτήρα μιας χρονοσειράς. Οι βασικές του παράμετροι είναι το πλήθος των γειτόνων (k) και η διάσταση (m) που χρησιμοποιούνται στην κατασκευή του μοντέλου. Επιπλέον ορίζονται ο χρόνος πρόβλεψης και τα χρονικά βήματα τί για την κατασκευή των διανυσμάτων του χώρου των φάσεων. Ο αλγόριθμος χωρίζει τη χρονοσειρά σε δύο μέρη: το πρώτο χρησιμοποιείται στην κατασκευή του ελκυστή και το δεύτερο στον έλεγχο των σφαλμάτων της πρόγνωσης. Κάθε πρόγνωση συγκρίνεται με την πραγματική τιμή και βρίσκεται το σφάλμα. Για έναν μεγάλο αριθμό προγνώσεων προκύπτει το μέσο σφάλμα $E(k)$ του μοντέλου για ένα συγκεκριμένο (k,m) . Με δοκιμές για διάφορα (k,m) , σχεδιάζονται οι καμπύλες $E(k)=f(k)$ για τα διάφορα m και βρίσκονται οι παράμετροι που βελτιστοποιούν το μοντέλο.

Η επικρατούσα άποψη για την ανάλυση των αποτελεσμάτων είναι ότι βελτιστοποίηση για μικρό k σημαίνει ντετερμινιστικό σύστημα ενώ για μεγάλο k στοχαστικό ή μεγάλης διάστασης ντετερμινιστικό. Ο μεγάλης διάστασης ντετερμινισμός απαιτεί μεγάλο αριθμό παραμέτρων-διαστάσεων με αποτέλεσμα το μοντέλο να υπερπροσομοιώνει τα δεδομένα και να πιάνει και το θόρυβο. Για αυτό χρειάζεται ένας μεγάλος αριθμός γειτόνων που να ανατρέι το θόρυβο. Όμως στην ουσία το πρόβλημα είναι το πλήθος των απαιτούμενων σημείων για να σχεδιαστεί ο ελκυστής, που διακρίνει ένα ντετερμινιστικό σύστημα με λίγους βαθμούς ελευθερίας από ένα με πολλούς. Αυτό το πλήθος αυξάνει εκθετικά με τη διάσταση (για το πλήθος των απαιτούμενων σημείων (βλ. Tsonis (1992)). Άλλιώς ο θόρυβος δεν έχει λόγο να επηρεάζει περισσότερο ποσοστιαία κάποιες διαστάσεις από άλλες εκτός αν αρχίσουμε να μιλάμε για συστήματα που περιέχουν ασυμβίβαστα σε χρόνο πρόβλεψης υποσυστήματα. Κάτι τέτοιο επιχειρείται στο τέλος αυτού του κεφαλαίου.

δ. Υλοποίηση του αλγόριθμου DVS

Η υλοποίηση του αλγόριθμου DVS των Casdagli and Weigend (1992) έγινε με την ανάπτυξη ενός προγράμματος υπολογιστή σε γλώσσα Pascal/Delphi. Στο πρόγραμμα έγινε πρόβλεψη α) δυνατότητας επιλογής διαφορετικών χρονικών βημάτων και β) δημιουργίας αρχείου που περιέχει τις προγνώσεις και τις πραγματικές τιμές.

Επιπλέον έχουν γίνει αλλαγές στον τρόπο εισαγωγής των δεδομένων και στη σειρά επεξεργασίας : γίνεται ξεχωριστή επεξεργασία για κάθε k .

Δεδομένης μιας χρονοσειράς πραγματικών τιμών ο αλγόριθμος επιχειρεί να την προσομοιάσει με μοντέλα της μορφής:

$$x_{i+T} \approx f(x_i, x_{i-t_1}, \dots, x_{i-t_{m-1}})$$

Στην απλούστερη περίπτωση η συνάρτηση f είναι γραμμική, οπότε πρακτικώς η παραπάνω σχέση ισοδύναμεί με μοντέλο AR. Η διαφορά από τα συνήθη μοντέλα AR της θεωρίας στοχαστικών ανελίξεων (βλ. Μψίκου (1985)), βρίσκεται στο γεγονός ότι οι συντελεστές της γραμμικής σχέσης δεν είναι σταθεροί, ούτε υπολογίζονται από τα στατιστικά χαρακτηριστικά της χρονοσειράς των x_i . Αντίθετα, για κάθε διάνυσμα $(x_i, x_{i-t_1}, \dots, x_{i-t_{m-1}})$ υπολογίζονται διαφορετικές τιμές των συντελεστών. Ο υπολογισμός αυτών των τιμών γίνεται με βάση όχι το σύνολο της διαθέσιμης χρονοσειράς, αλλά των γειτονικών ως προς τις τιμές διανυσμάτων. Πρόκειται δηλαδή για τοπικού (ως προς το χώρο των φάσεων) χαρακτήρα εκτίμηση.

Για να βρεθεί η γραμμική συνάρτηση $f : R^m \rightarrow R$ που δίνει την καλύτερη πρόβλεψη για το x_{i+T} , υπό την έννοια ότι ελαχιστοποιεί το σφάλμα, χρησιμοποιείται η μέθοδος των ελαχίστων τετραγώνων. Οι ακέραιοι T και m είναι τα εξής μεγέθη :

T : χρόνος πρόβλεψης στο μέλλον

m : διάσταση έδρασης ή διάσταση του κατασκευαζόμενου χώρου των φάσεων

Η μεταβλητή της f είναι το διάνυσμα \bar{x}_i . Τα t_i είναι τα χρονικά βήματα μεταξύ των συνιστώσων του διανύσματος.

Στη συνέχεια δίδεται ο αλγόριθμος :

1. Κανονικοποιείται η χρονοσειρά με μέση τιμή μηδέν και τυπική απόκλιση μονάδα (βλ. Κουτσογιάννης (1996)).
2. Χωρίζεται η χρονοσειρά σε δύο μέρη :

α. Το πρώτο $\{x_1, \dots, x_{Nf}\}$ χρησιμοποιείται για να γίνει ο υπολογισμός των μεταβλητών του κάθε μοντέλου.

β. Το δεύτερο $\{x_{Nf+1}, \dots, x_{Nf+Nt}\}$ για να συγκρίνουμε την πρόβλεψη με την πραγματική τιμή. Το Nf είναι το πλήθος των σημείων που θεωρούνται εξ αρχής γωστά και το Nt το πλήθος των σημείων που θα χρησιμοποιηθούν στον έλεγχο της πρόγνωσης.

3. Επιλέγεται το T .

4. Επιλέγεται το m (Τα m, T, Nt, ti, k διαβάζονται από το αρχείο input.dat).
5. Καθορίζεται το διάνυσμα \bar{x}_i για πρόβλεψη χρόνου T (i>Nf).
6. Υπολογίζονται οι αποστάσεις του διανύσματος \bar{x}_i από τα διανύσματα που βρίσκονται πριν από το x_{i-T} και επιλέγονται τα k πιο κοντινά. Ως μέτρο της απόστασης για την επιλογή των κοντινών σημείων χρησιμοποιείται η νόρμα του αθροίσματος των απολύτων διαφορών των συντεταγμένων.
7. Προσαρμόζεται ένα γραμμικό μοντέλο της μορφής :

$$x^l_{j+T} = a_0 + \sum_{n=1}^m a_n x^l_{j-t_{n-1}}$$

όπου l=1,...,k και t0=0.

Πρόκειται για ένα μοντέλο AR, τάξης m, προσαρμοσμένο στα k γειτονικότερα του \bar{x}_i σημεία (σε κάθε διάνυσμα αντιστοιχεί αμφιμονοσήμαντα ένα σημείο στον χώρο των φάσεων). Έτσι σχηματίζεται ένα σύστημα k εξισώσεων με m+1 αγνώστους (a_0, a_1, \dots, a_m) που προσεγγίζεται με τη μέθοδο ελαχίστων τετραγώνων (βλ. Μπακόπουλος και Χρυσοβέργης (1985)).

8. Γίνεται χρήση του μοντέλου για εκτίμηση μιας πρόγνωσης σε χρόνο T : x_{i+T} και υπολογίζεται το απόλυτο σφάλμα $e_{i+T}(k) = |x_{i+T} - x_{i+T}|$.
9. Επαναλαμβάνονται τα βήματα 5 ως 8 και υπολογίζεται το μέσο απόλυτο σφάλμα πρόβλεψης :

$$E_m(k) = \frac{\sum_{i=T}^{M+m} e_{i+T}(k)}{M-T}$$

10. Δημιουργούνται δύο είδη αρχείων : α) με το μέσο σφάλμα πρόβλεψης συναρτήσει των m, k (Case0?.dat), β) με την τυπική απόκλιση και τις τιμές x_{i+T} , x_{i+T}' (c_0?_0?.exp). Από τα αρχεία (α) λαμβάνονται τα δεδομένα για τα γραφήματα $E_m(k)=f(k,m)$ και από τα (β) για τα μεικτά γραφήματα πρόβλεψη-πραγματική τιμή = $f(i)$.

Στο παράρτημα δίδεται η λίστα του προγράμματος. Είναι γραμμένο σε κώδικα Pascal για Delphi.

Οι μέγιστες επιτρεπόμενες τιμές των παραμέτρων όπως καθορίζονται στον κώδικα είναι :

```
max k=5000
max m=101
max i =100,000 (πλήθος στοιχείων χρονοσειράς)
max(i-Nf)=2000
```

ε. Παράδειγμα

Θα γίνει εφαρμογή του αλγορίθμου στο σύστημα του Hènon. Το σύστημα αυτό με κατάλληλη επιλογή των παραμέτρων του γίνεται χαοτικό. Η εφαρμογή θα γίνει στο σύστημα :

$$x_{t+1} = 1.4 - x_t^2 + 0.3 y_t$$

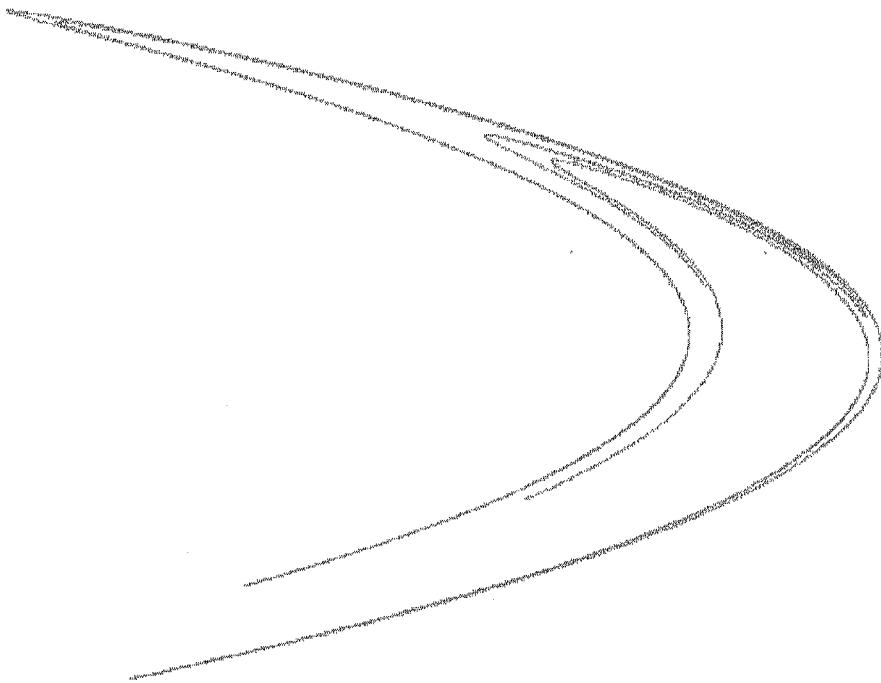
$$y_{t+1} = x_t$$

Στη συνέχεια δίδεται η λίστα ενός προγράμματος σε Pascal που δημιουργεί μια χρονοσειρά αυτού του συστήματος. Όπως φαίνεται, ο βρόγχος ανακυκλώνεται μερικές φορές πριν αρχίσει να γράφει τιμές στο αρχείο. Αυτό γίνεται για να προσεγγίσει η τροχιά των ελκυστή.

{λίστα Hènon}

```
var f3:text;
var x,y,z:real;i:integer;
begin
assign(f3,'a:\henon.dat');
rewrite(f3);
x:=0.5;y:=0.5;
for i:=1 to 300 do
begin
z:=x;
x:=1.4-x*x+0.3*y;
y:=z;
end;
for i:=1 to 3000 do
begin
z:=x;
x:=1.4-x*x+0.3*y;
writeln(f3,x);
y:=z;
end;
close(f3);
end.
```

Ο ελκυστής αυτού του συστήματος σχεδιασμένος με το πρόγραμμα ‘Dynamics Numerical Explorations’ των Nusse and Yorke (1994) δίδεται στο σχήμα A.3 :



Σχήμα A.3. Ο ελκυστής του Hénon.

Για να γίνει επεξεργασία της χρονοσειράς Hénon.dat, πρέπει αυτή να αποθηκευτεί στο αρχείο data.dat. Στη συνέχεια πρέπει να οριστούν οι μεταβλητές του αρχείου input.dat. Επειδή γνωρίζουμε το δυναμικό σύστημα που την παράγει ξέρουμε κατά προσέγγιση και τη διάσταση την οποία θα την εδράσουμε. Θα δοκιμάσουμε σε δύο και τρεις διαστάσεις και με μικρό αριθμό προγνώσεων (3000-2980=20 στοιχεία). Το αρχείο input.dat διαμορφώνεται ως εξής :

INPUT FILE

Parametric Input:

Number of Examination Cases:

2

Case I:

Amount of past time series (Nf):

2980

Prediction time (T):

1

Dimension (m):

2

T_i from T₁ to T_{m-1} (m-1 total values)

1

Amount of K_i different values:

5

K_i from K₁ to K_{final}:

5

10

50

100

200

Case II:

Amount of past time series (N_f):

2980

Prediction time (T):

1

Dimension (m):

3

T_i from T₁ to T_{m-1} (m-1 total values)

1

2

Amount of K_i different values:

5

K_i from K₁ to K_{final}:

5

10

50
100
200

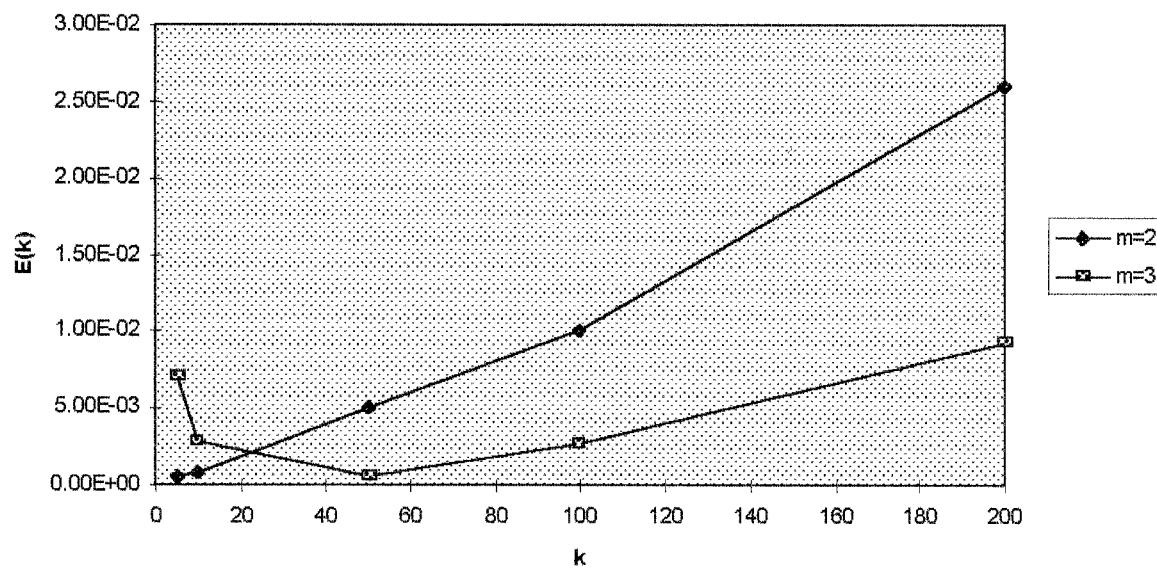
Στη συνέχεια ‘τρέχει’ το πρόγραμμα. Αυτό δημιουργεί τα αρχεία case01.dat και case02.dat. Τα αποτελέσματα δίδονται στον πίνακα A.1 :

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
2	5	5,13E-04
	10	7,93E-04
	50	5,02E-03
	100	1,03E-02
	200	2,62E-02
3	5	6,99E-03
	10	2,77E-03
	50	5,90E-04
	100	2,61E-03
	200	9,17E-03

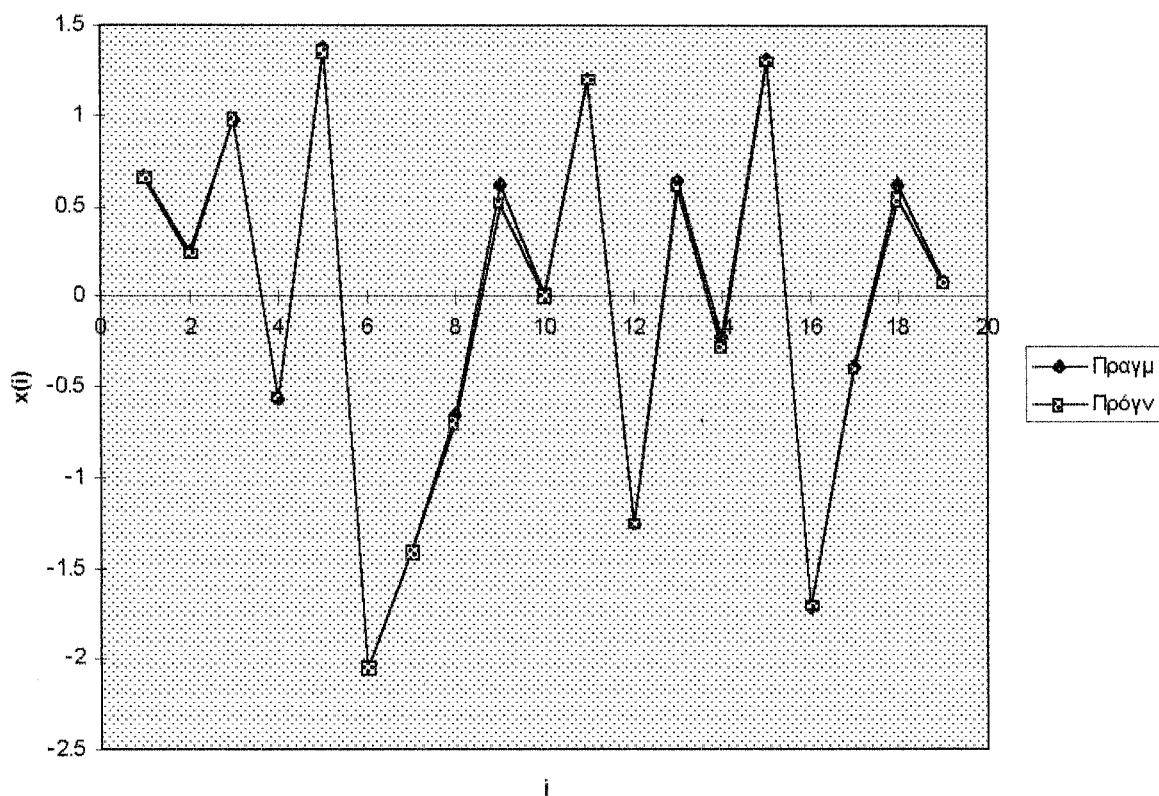
Πίνακας A.1. Αποτελέσματα πρόγνωσης στο σύστημα του Hènon.

Όλα τα σφάλματα είναι πολύ μικρά. Αυτό σημαίνει ότι η δυναμική του συστήματος αποδόθηκε από το μοντέλο. Φαίνεται επίσης ότι η πρόγνωση είναι καλύτερη για μικρό αριθμό γειτόνων. Αυτό προκύπτει από το γεγονός ότι το σύστημα είναι καθαρά ντετερμινιστικό.

Στη συνέχεια σχεδιάζονται τα γραφήματα σφαλμάτων (σχήμα A.4) και το γράφημα πραγματικών τιμών-πρόβλεψης του c_01_05.exp (σχήμα A.5).



Σχήμα A.4. Διάγραμμα $E(k)=f(k)$ για πρόγνωση στον ελκυστή του Hénon.



Σχήμα A.5. Διάγραμμα πραγματικών τιμών - προγνώσεων στα τελευταία 19 στοιχεία του Henon.dat για την χειρότερη πρόβλεψη : $m=2$, $k=200$.

στ. Όρια πρόγνωσης

Στη συνέχεια σχολιάζεται το θέμα των ορίων μιας πρόγνωσης. Γίνεται προσέγγιση μέσω δύο οδών: i) της χρονικής κλίμακας, ii) μίας νέας κατηγορίας δυναμικών συστημάτων που προτείνεται.

i) Χρονική κλίμακα

Σε ένα χαοτικό δυναμικό σύστημα μπορούμε να κάνουμε ακριβή πρόβλεψη ως κάποιο χρονικό όριο. Αν αυξήσουμε πολύ τον ορίζοντα, η πρόγνωση θα γίνει τυχαία μεταβλητή. Μεταβλητές της πρόγνωσης είναι η ανεκτή ακρίβεια πρόγνωσης, το γνωστό πλήθος των σημείων του ελκυστή και η ακρίβεια με την οποία γνωρίζουμε τις συντεταγμένες κάθε σημείου. Είδαμε στο σύστημα του Hénon ότι μπορούμε να προγνώσουμε για μοναδιαίο ορίζοντα με πολύ μεγάλη ακρίβεια. Αν όμως προσπαθήσουμε να κάνουμε πρόγνωση για ορίζοντα 50 θα δούμε ότι στην καλύτερη περίπτωση οι προγνώσεις μας θα είναι στοχαστικές.

Τι γνώση θα μπορούσαμε να έχουμε για το σύστημα του Hénon, αν το συναντούσαμε κάπου σε εξέλιξη και παίρναμε μετρήσεις κάθε 50 βήματά του; Καμία για τη δυναμική του. Κάθε δυναμικό σύστημα πρέπει να εξετάζεται στη χρονική κλίμακα στην οποία εξελίσσεται.

Αν όμως ένα δυναμικό σύστημα αποτελείται από υποσυστήματα που εξελίσσονται σε διαφορετικές χρονικές κλίμακες (ασύμβατες κλίμακες), είμαστε υποχρεωμένοι να το εξετάσουμε στη μικρότερη και να περιοριστούμε σε βραχυχρόνιες προγνώσεις.

ii) Μία νέα κατηγορία δυναμικών συστημάτων

Έστω η συνάρτηση $f(x)=\cos(1/x)$ ορισμένη στο $(-1,0)$. Η γραφική της παράσταση θυμίζει fractal. Η επεξεργασία της με το 'Chaos data analyzer', Sprott and Rowlands (1991) δίνει: α. για δείγμα 10,000 ισαπεχόντων σημείων στο διάστημα $(-0.5, -0.5E-4)$, διάσταση (correlation dimension) 1.26 ± 0.16 , β. για δείγμα 10,000 ισαπεχόντων σημείων στο διάστημα $(-0.5E-2, -0.5E-6)$, διάσταση (correlation dimension) 1.63 ± 0.12 , γ. για δείγμα 1,000 ισαπεχόντων σημείων στο διάστημα $(-0.5E-5, -0.5E-6)$, διάσταση (correlation dimension) 2.38 ± 0.08 , δ. για δείγμα 3,000 ισαπεχόντων σημείων στο διάστημα $(-1.5E-6, -0.5E-7)$, διάσταση (correlation dimension) 2.42 ± 0.03 . Στο σημείο αυτό ανακαλούμε ένα θεώρημα των fractals που λέει ότι η διάσταση ενός fractal είναι η μεγαλύτερη διάσταση των fractal υποσυνόλων του (βλ. Barnsley (1993)). Έτσι θα μπορούσαμε να δεχθούμε ως ακριβέστερα τα αποτελέσματα για διαστήματα που βρίσκονται πιο κοντά στο 0, όπου είναι όπως φαίνεται μεγαλύτερη η διάσταση. Μία κριτική αυτών των αποτελεσμάτων μας οδηγεί να υποθέσουμε ότι πιθανόν να είναι θέμα προβλήματος αντιληπτότητας, λόγω εισροής κλιμάκων αυξανόμενης ασυμβατότητας για $x \rightarrow 0$, που μας υποχρεώνει να βλέπουμε το 'fractal' αυτό σε αυξανόμενη διάσταση. Οταν η διάσταση ξεπερνάει το 2, τότε το αντικείμενο νοείται στο χώρο. Έτσι προκύπτει ότι η γραφική παράσταση της $f(x)=\cos(1/x)$

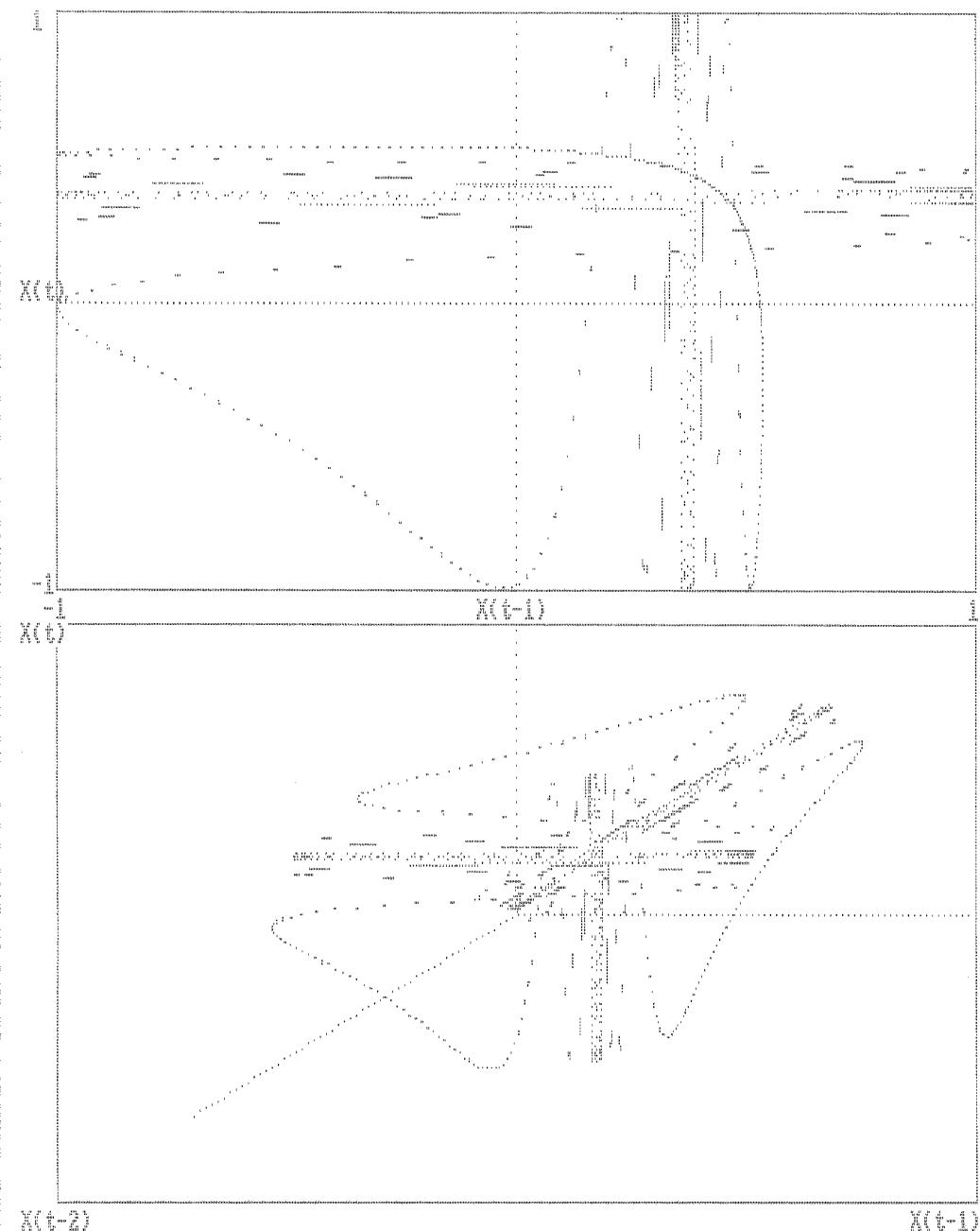
είναι η προβολή ενός αντικειμένου του χώρου, ίσως και ολόκληρου του χώρου αν η διάσταση τείνει στο 3 για διάστημα του οποίου τα όρια τείνουν στο μηδέν, στο επίπεδο. Όμως από τα παραπάνω αποτελέσματα για τη διάσταση της f , φαίνεται και ότι για απόλυτη αύξηση του αριστερού ορίου του διαστήματος στο οποίο την ορίζουμε μικραίνει η διάστασή της, πράγμα που σημαίνει ότι η διάστασή της ακολουθεί τη μικρότερη διάσταση των υποσυνόλων της. Αυτό σημαίνει ότι η συνάρτηση αυτή συμπεριφέρεται αντίστροφα από τα fractals.

Η καμπύλη της $f(x)$ αποτελείται από άπειρους κύκλους, σμικρυνόμενης περιόδου καθώς το x τείνει στο 0. Το μήκος της καμπύλης είναι άπειρο. Ένα δυναμικό σύστημα που περιέχει αυτής της μορφής συναρτήσεις έχει παράξενη συμπεριφορά. Ένα τέτοιο σύστημα είναι το (1):

$$\begin{aligned}x_{n+1} &= \cos\left(\frac{1}{\cos(t_n)}\right) \\t_{n+1} &= t_n + 1\end{aligned}$$

Στο σχήμα A.6 δίδεται ο ελκυστής του εδρασμένος σε δύο και τρεις διαστάσεις όπως σχεδιάστηκε από το πρόγραμμα ‘Chaos data analyzer’, Sprott and Rowlands (1991). Στις τρεις διαστάσεις θυμίζει τρισορθογώνιο σύστημα.

Τα αποτελέσματα της επεξεργασίας του για ορίζοντα $T=1$, σε 2, 3, 4, 5, 6 διαστάσεις δίδονται στον πίνακα A.2, και για ορίζοντες 10 και 100 σε 2 διαστάσεις στους πίνακες A.3 και A.4 αντίστοιχα.



Σχήμα Α.6. Ο ελκυστής του συστήματος (1) εδρασμένος σε δύο και τρεις διαστάσεις.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
2	10	1,134
	30	0,162
	50	0,091
	80	0,065
	120	0,070
	170	0,109
3	10	8,64
	30	0,439
	50	0,790
	80	0,143
	120	0,126
	170	0,199
4	10	0,679
	30	1,302
	50	1,861
	80	2,430
	120	0,190
	170	0,184
5	10	0,620
	30	0,799
	50	0,566
	80	0,226
	120	0,177
	170	0,233
6	10	2,250
	30	0,756
	50	1,525
	80	0,268
	120	0,609
	170	0,273

Πίνακας A.2. Αποτελέσματα πρόγνωσης με ορίζοντα 1 ώρα.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
2	10	0,496
	30	0,231
	50	0,184
	80	0,068
	120	0,092
	170	0,111

Πίνακας A.3. Αποτελέσματα πρόγνωσης με ορίζοντα 10 ώρες

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
2	10	5,856
	30	0,118
	50	0,031
	80	0,043
	120	0,058
	170	0,083

Πίνακας A.4. Αποτελέσματα πρόγνωσης με ορίζοντα 100 ώρες

Από τα αποτελέσματα της επεξεργασίας φαίνεται ότι για αυτό το σύστημα δε μπορούμε να έχουμε τέλεια πρόβλεψη. Η βελτιστοποίηση γίνεται για σχετικά μεγάλο αριθμό γειτόνων. Και προκύπτει και ότι η πρόβλεψη για μεγάλου βάθους ορίζοντα δε χάνει ακρίβεια.

Ένα πιο σύνθετο σύστημα της $\cos(1/x)$ θα απαιτούσε περισσότερους γείτονες. Για ένα δυναμικό σύστημα που θα περιείχε ένα τέτοιο υποσύστημα - αν και ντετερμινιστικό - δε θα μπορούσαμε να βελτιστοποιήσουμε την πρόβλεψη με μικρό αριθμό γειτόνων, ενώ θα υπήρχε ένα ελάχιστο συστηματικό σφάλμα στην πρόβλεψη.

ΚΕΦΑΛΑΙΟ Β

Εφαρμογή σε χρονοσειρές μετεωρολογικών δεδομένων

Θα αναλυθούν μετρήσεις ατμοσφαιρικών δεδομένων του σταθμού του Ε.Μ.Π. Τα μεγέθη που θα εξεταστούν είναι: μέση ταχύτητα ανέμου και διεύθυνσή του, θερμοκρασία, ατμοσφαιρική πίεση και υγρασία. Τα αρχεία δεδομένων είναι ωριαία του έτους 1996 και 10λεπτά της περιόδου 19/01/96 - 07/03/96.

Θα χρησιμοποιηθεί χρονικό βήμα 1 ώρας: $t_i = 1, 2, 3, \dots$

Η ανάλυση γίνεται με το πρόγραμμα για υπολογιστή που φτιάξαμε για τον αλγόριθμο DVS. Αυτό επεξεργάζεται μία χρονοσειρά μετρήσεων. Τα ωριαία αρχεία μετρήσεων που χρησιμοποιούμε έχουν 8784 στοιχεία. Έστω ότι θέτουμε $N_f = 8600$. Ας δούμε τι κάνει το πρόγραμμα για δεδομένα: διάσταση $m = 3$, αριθμός γειτόνων $k = 30$, ορίζοντας πρόγνωσης $T = 1$ ώρα: α) Κατασκευάζει ένα διάνυσμα τριών διαστάσεων με στοιχεία τις τιμές των 8600-στού, 8599-στού, 8598-στού στοιχείου της χρονοσειράς (επειδή το χρονικό βήμα το έχουμε θέσει 1 ώρα, τα στοιχεία που λαμβάνουμε στην κατασκευή του διανύσματος είναι συνεχόμενα). β) Ερευνά σε όλα τα διανύσματα από το (x_3, x_2, x_1) ως το $(x_{8599}, x_{8598}, x_{8597})$ και βρίσκει τα 30 πιο κοντινά στο $(x_{8600}, x_{8599}, x_{8598})$. γ) Βρίσκει που μεταφέρθηκε το κάθε διάνυσμα (σημείο του τρισδιάστατου χώρου), από τα 30 κοντινότερα, ύστερα από 1 ώρα (επειδή $T = 1$ ώρα) και δημιουργεί ένα γραμμικό μοντέλο της μετακίνησης με τη μέθοδο των ελαχίστων τετραγώνων. δ) Εφαρμόζει το μοντέλο του προηγούμενου βήματος στο διάνυσμα $(x_{8600}, x_{8599}, x_{8598})$ και υπολογίζει την τιμή του 8601-στού στοιχείου. Αυτή είναι μία πρόγνωση. ε) Συγκρίνει την πρόγνωση με την πραγματική τιμή του 8601-στού στοιχείου και υπολογίζει το απόλυτο σφάλμα. στ) Επαναλαμβάνει όλη τη διαδικασία για πρόβλεψη των 8602-στού, ..., 8784-στού στοιχείων και βρίσκει τα απόλυτα σφάλματα πρόγνωσης. Η μέση τιμή αυτών των σφαλμάτων είναι το μέσο απόλυτο σφάλμα πρόγνωσης.

α. Θερμοκρασία (8784 ωριαίες μετρήσεις)

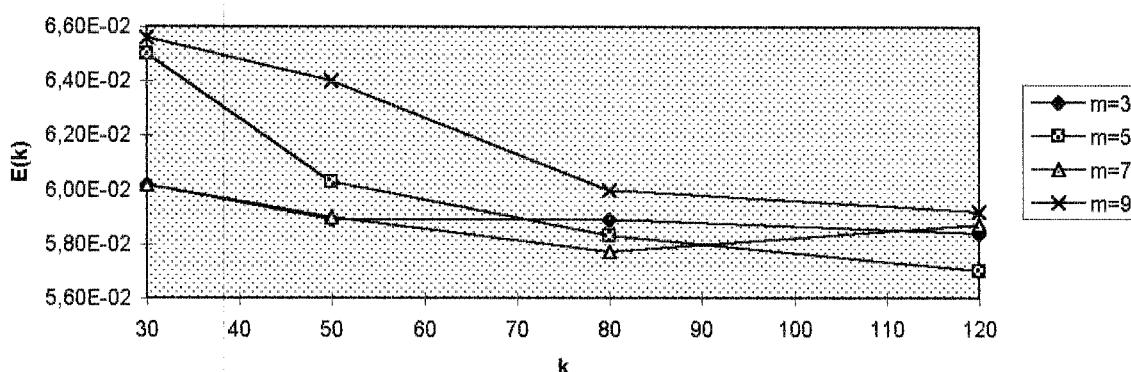
Η θερμοκρασία είναι ένα μέγεθος που χαρακτηρίζεται από μικρές μεταβολές και έναν κύκλο 24 ωρών. Θα γίνει επεξεργασία της για πρόβλεψη 1, 6, 12, 24 ωρών.

i) $T=1h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-1=183 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.1 και γραφικά στο σχήμα B.1.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	6,02E-02
	50	5,89E-02
	80	5,89E-02
	120	5,84E-02
5	30	6,50E-02
	50	6,03E-02
	80	5,83E-02
	120	5,70E-02
7	30	6,02E-02
	50	5,90E-02
	80	5,77E-02
	120	5,87E-02
9	30	6,56E-02
	50	6,40E-02
	80	6,00E-02
	120	5,92E-02

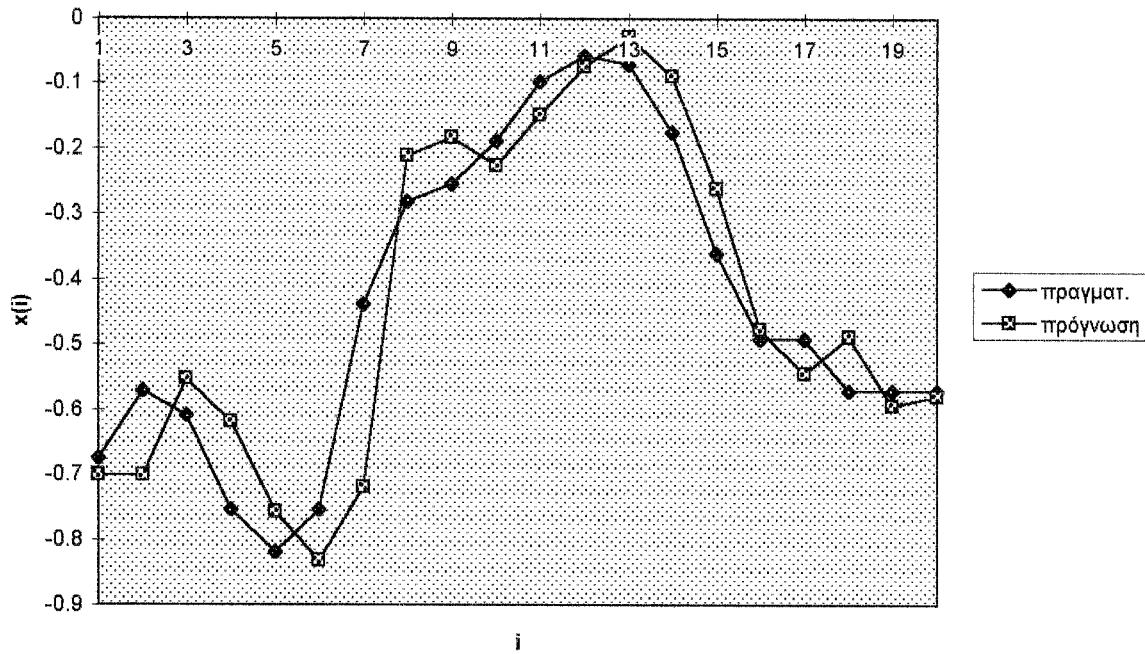
Πίνακας B.1. Αποτελέσματα πρόγνωσης της θερμοκρασίας με ορίζοντα 1 ώρα.



Σχήμα B.1. Γράφημα σφαλμάτων πρόγνωσης E(k) θερμοκρασίας για T=1h

Οι ωριαίες μεταβολές της θερμοκρασίας είναι μικρές και η πρόβλεψη είναι καλή. Το μεγάλο πλήθος των απαιτούμενων γειτόνων για βελτιστοποίηση της πρόβλεψης φανερώνει το στοχαστικό χαρακτήρα της χρονοσειράς. Ένα μικρό πλήθος βαθμών ελευθερίας (5-7) μπορεί να διαχωριστεί από το θόρυβο.

Στο σχήμα B.2 δίδεται το γράφημα πραγματικών τιμών - πρόγνωσης για τις τελευταίες 20 προγνώσεις στο δείγμα για m=5, k=120.



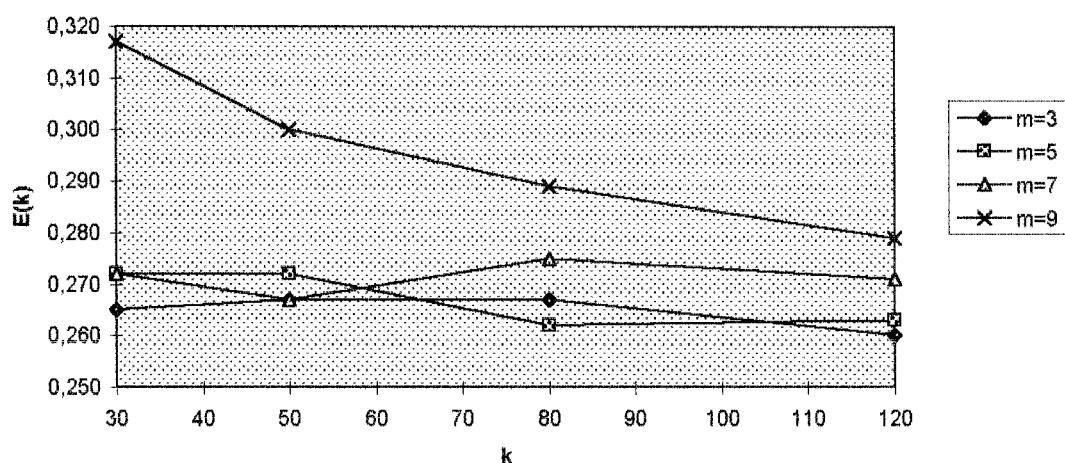
Σχήμα B.2. Γράφημα πραγματικής τιμής - πρόγνωσης για ορίζοντα 1 ώρας, των τελευταίων 20 ωρών του δείγματος ($m=5$, $k=120$).

ii) $T=6h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν $8784 - 8600 = 178$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.2 και γραφικά στο σχήμα B.3.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,265
	50	0,267
	80	0,267
	120	0,260
5	30	0,272
	50	0,272
	80	0,262
	120	0,263
7	30	0,272
	50	0,267
	80	0,275
	120	0,271
9	30	0,317
	50	0,300
	80	0,289
	120	0,279

Πίνακας B.2. Αποτελέσματα πρόγνωσης της θερμοκρασίας με ορίζοντα 6 ώρες.



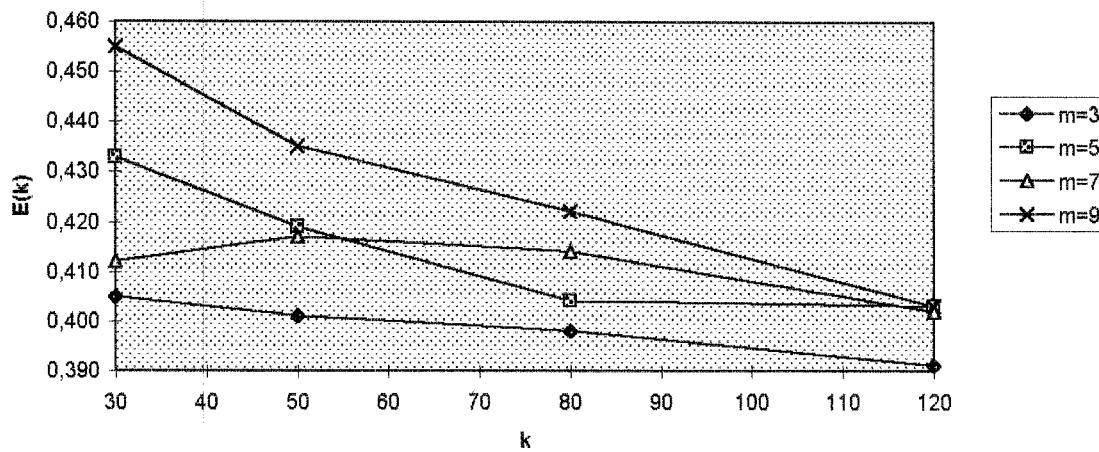
Σχήμα B.3. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ θερμοκρασίας για $T=6h$

iii) T=12h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-12=172 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.3 και γραφικά στο σχήμα B.4.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,405
	50	0,401
	80	0,398
	120	0,391
5	30	0,433
	50	0,419
	80	0,404
	120	0,403
7	30	0,412
	50	0,417
	80	0,414
	120	0,402
9	30	0,455
	50	0,435
	80	0,422
	120	0,403

Πίνακας B.3. Αποτελέσματα πρόγνωσης της θερμοκρασίας με ορίζοντα 12 ώρες.



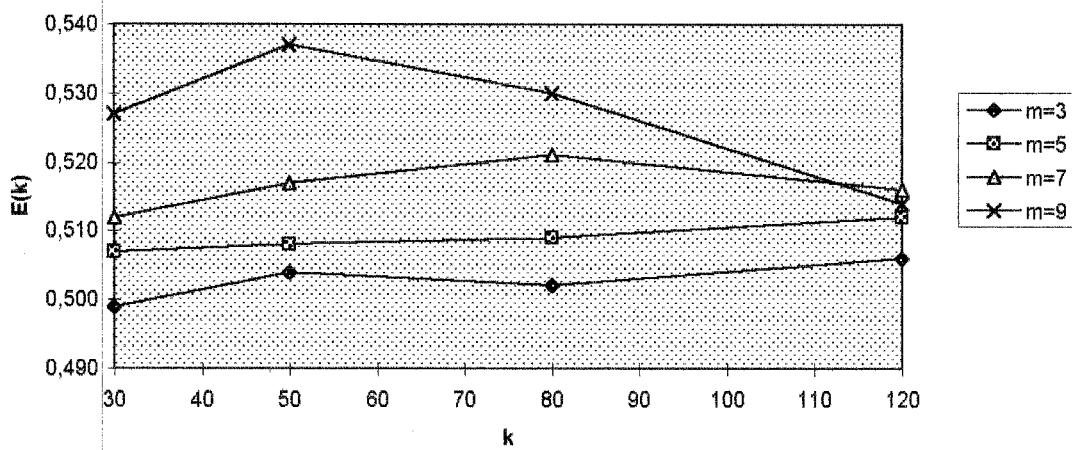
Σχήμα B.4. Γράφημα σφαλμάτων πρόγνωσης E(k) θερμοκρασίας για T=12h

iv) T=24h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-24=160 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.4 και γραφικά στο σχήμα B.5.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,499
	50	0,504
	80	0,502
	120	0,506
5	30	0,507
	50	0,508
	80	0,509
	120	0,512
7	30	0,512
	50	0,517
	80	0,521
	120	0,516
9	30	0,527
	50	0,537
	80	0,530
	120	0,514

Πίνακας B.4. Αποτελέσματα πρόγνωσης της θερμοκρασίας με ορίζοντα 24 ώρες.

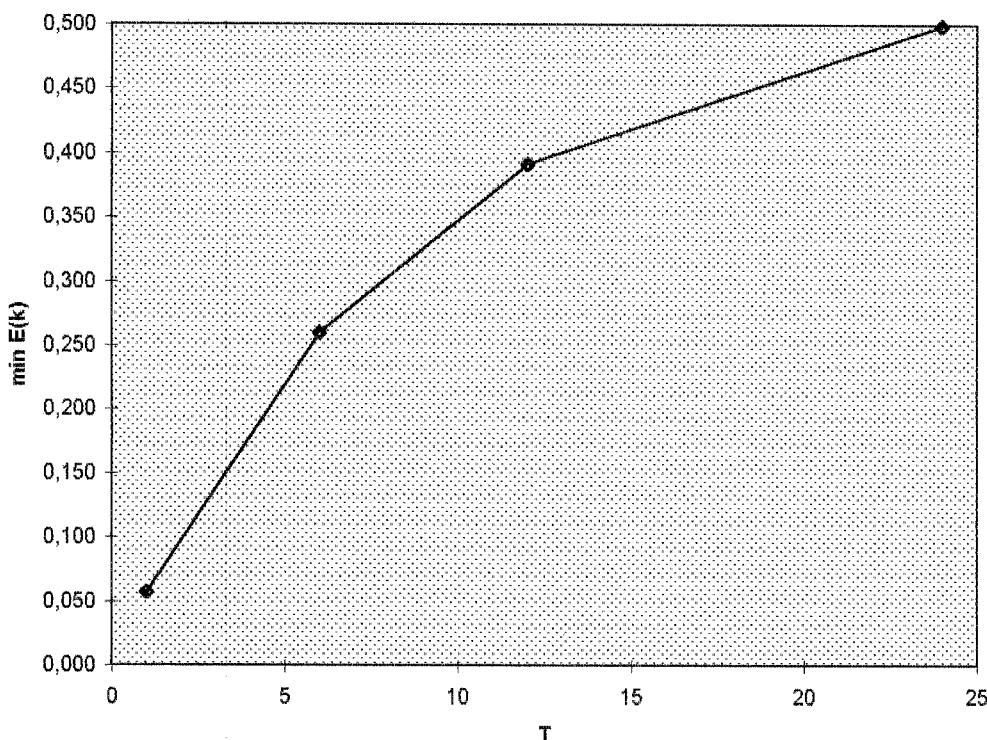


Σχήμα B.5. Γράφημα σφαλμάτων πρόγνωσης E(k) θερμοκρασίας για T=24h

v) Συσχέτιση των αποτελεσμάτων

Από τα προηγούμενα αποτελέσματα προκύπτει μία τάση μείωσης της διάστασης m , για βέλτιστοποίηση της πρόβλεψης, με την αύξηση του ορίζοντα πρόβλεψης T . Λιγότεροι ντετερμινιστικοί βαθμοί ελευθερίας ανιχνεύονται με παράλληλη αύξηση του σφάλματος πρόβλεψης. Δηλαδή ο αλγόριθμος τείνει να αναγνωρίζει ίχνη μόνο ντετερμινισμού όσο αυξάνει ο ορίζοντας πρόβλεψης και ως εκ τούτου η διαφορά στις κλίμακες ορίζοντα T - βήματος $t=1\text{ h}$ των μετρήσεων.

Στο σχήμα B.6 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m, k .



Σχήμα B.6. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για τη θερμοκρασία

Η καμπύλη του σχήματος B.6 προσεγγίζεται βέλτιστα από τη συνάρτηση :

$$\min E(k)(T) \cong 0.044 + 0.138 \ln(T)$$

με συντελεστή αυτοσυσχέτισης 0.993 .

Για $\min E(k)(T) = 0 \Rightarrow T \cong 0.727 \text{ h} \cong 44 \text{ min}$

Αν είχαμε μετρήσεις κάθε 44 λεπτά δε θα μηδενιζόταν το σφάλμα, ωστόσο αυτός είναι ένας χαρακτηριστικός χρόνος των μετρήσεων θερμοκρασίας.

β. Ατμοσφαιρική πίεση (8784 ωριαίες μετρήσεις)

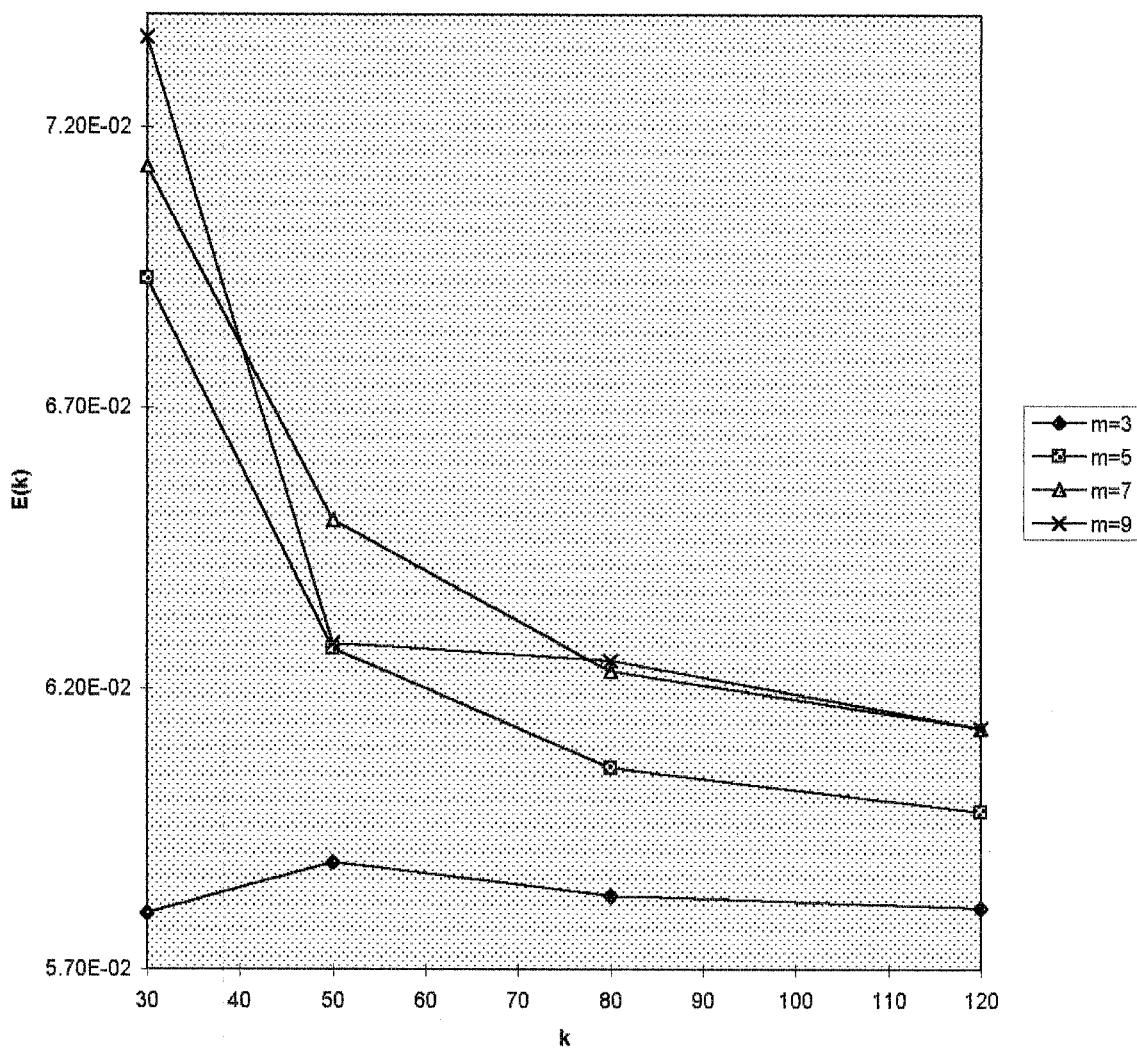
Η ατμοσφαιρική πίεση είναι ένα μέγεθος που χαρακτηρίζεται από πολύ μικρές μεταβολές. Θα γίνει επεξεργασία της για πρόβλεψη 1, 6, 12, 24 ωρών.

i) T=1h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-1=183 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.5 και γραφικά στο σχήμα B.7 .

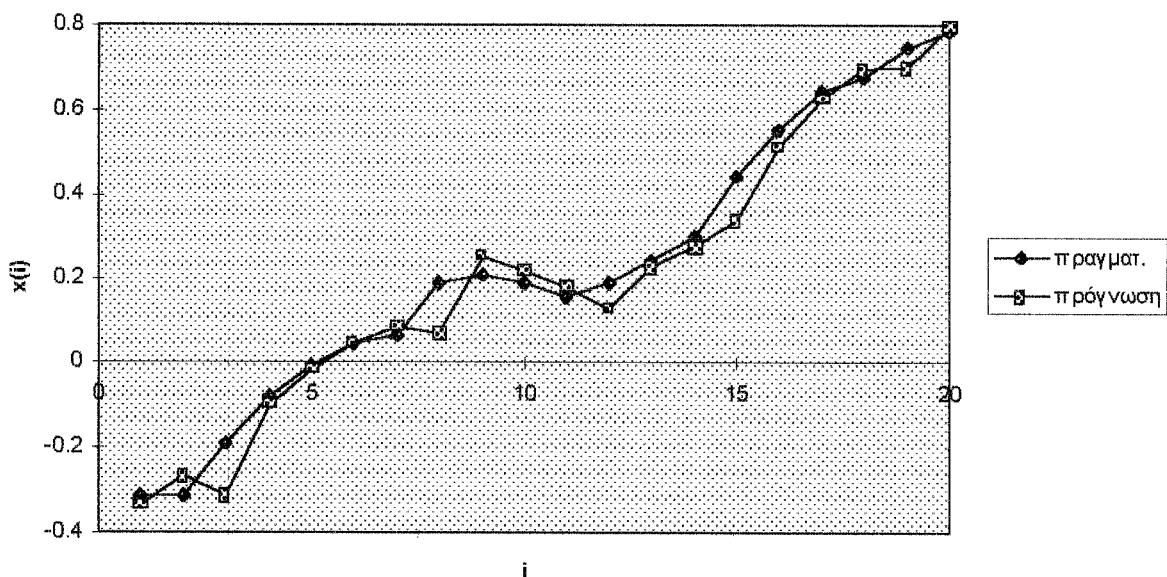
Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	5,80E-02
	50	5,89E-02
	80	5,83E-02
	120	5,81E-02
5	30	6,93E-02
	50	6,27E-02
	80	6,06E-02
	120	5,98E-02
7	30	7,13E-02
	50	6,50E-02
	80	6,23E-02
	120	6,13E-02
9	30	7,36E-02
	50	6,28E-02
	80	6,25E-02
	120	6,13E-02

Πίνακας B.5. Αποτελέσματα πρόγνωσης της ατμοσφαιρικής πίεσης με ορίζοντα 1 ώρα.



Σχήμα B.7. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ ατμ. πίεσης για $T=1h$

Στο σχήμα B.8 δίδεται το γράφημα πραγματικών τιμών - πρόγνωσης για τις τελευταίες 20 προγνώσεις στο δείγμα για $m=3$, $k=120$.



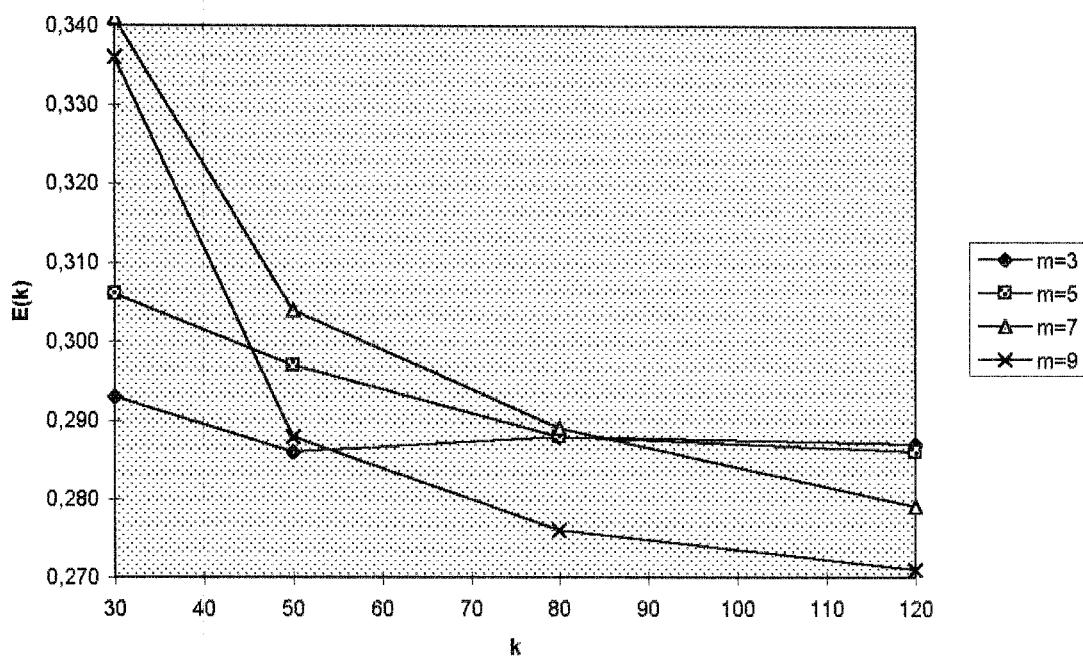
Σχήμα B.8. Γράφημα πραγματικής τιμής - πρόγνωσης για ορίζοντα 1 ώρας, των τελευταίων 20 ωρών του δείγματος τιμών ατμ. πίεσης($m=5$, $k=120$).

ii) $T=6h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 8600 (δηλαδή θα γίνουν $8784-8600-6=178$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.6 και γραφικά στο σχήμα B.9.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,293
	50	0,286
	80	0,288
	120	0,287
5	30	0,306
	50	0,297
	80	0,288
	120	0,286
7	30	0,341
	50	0,304
	80	0,289
	120	0,279
9	30	0,336
	50	0,288
	80	0,276
	120	0,271

Πίνακας B.6. Αποτελέσματα πρόγνωσης της ατμ. πίεσης με ορίζοντα 6 ώρες.



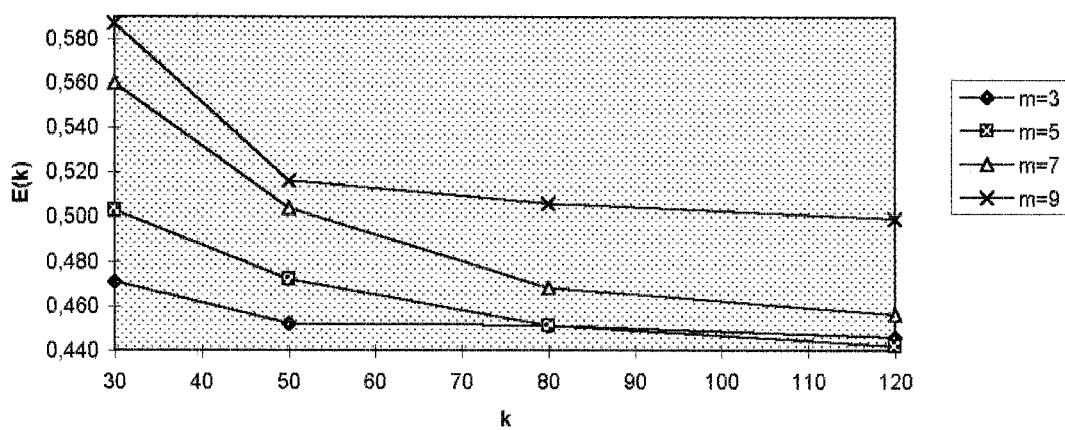
Σχήμα B.9. Γράφημα σφαλμάτων πρόγνωσης E(k) ατμ. πίεσης για T=6h

iii) T=12h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-12=172 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.7 και γραφικά στο σχήμα B.10.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,471
	50	0,452
	80	0,451
	120	0,446
5	30	0,503
	50	0,472
	80	0,451
	120	0,442
7	30	0,560
	50	0,504
	80	0,468
	120	0,456
9	30	0,587
	50	0,516
	80	0,506
	120	0,499

Πίνακας B.7. Αποτελέσματα πρόγνωσης της ατμ. πίεσης με ορίζοντα 12 ώρες.



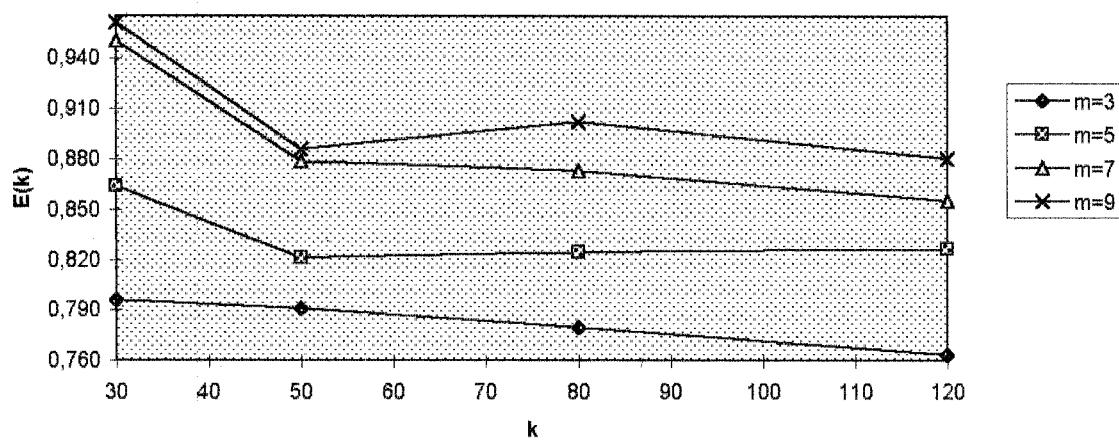
Σχήμα B.10. Γράφημα σφαλμάτων πρόγνωσης E(k) ατμ.πίεσης για T=12h

iv) T=24h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-24=160 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.8 και γραφικά στο σχήμα B.11.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,796
	50	0,791
	80	0,779
	120	0,763
5	30	0,864
	50	0,821
	80	0,824
	120	0,826
7	30	0,950
	50	0,879
	80	0,873
	120	0,855
9	30	0,961
	50	0,886
	80	0,902
	120	0,880

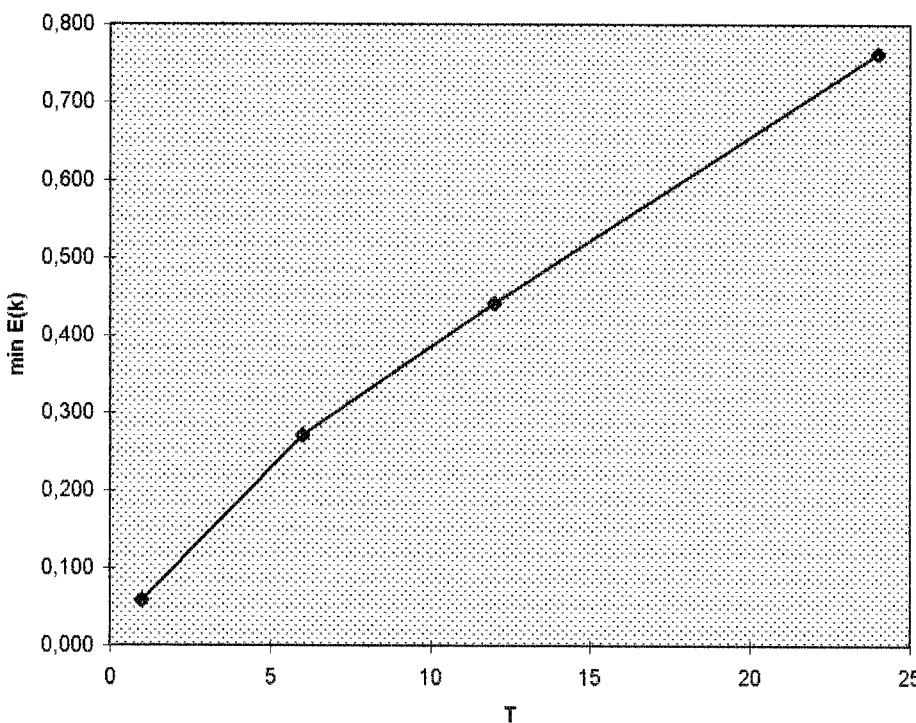
Πίνακας B.8. Αποτελέσματα πρόγνωσης της ατμ. πίεσης με ορίζοντα 24 ώρες.



Σχήμα B.11. Γράφημα σφαλμάτων πρόγνωσης E(k) ατμ. πίεσης για T=24h

v) Συσχέτιση των αποτελεσμάτων

Στο σχήμα B.12 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m , k .



Σχήμα B.12. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για την ατμ. πίεση

Η καμπύλη του σχήματος B.9 προσεγγίζεται βέλτιστα από τη συνάρτηση :

$$\min E(k)(T) \cong 0,0595 T^{0.811}$$

με συντελεστή αυτοσυσχέτισης 0.999 .

Για $\min E(k)(T) \rightarrow 0 \Rightarrow T \rightarrow 0$

Η ατμοσφαιρική πίεση είναι ένα μέγεθος δυσμετάβλητο μεν με πάρα πολλούς βαθμούς ελευθερίας δε. Ο χρόνος $T \rightarrow 0$ δεν έχει πρακτικό νόημα για σύντομη πρόβλεψη αλλά φανερώνει την αδυναμία για μακροχρόνια πρόβλεψη.

γ. Υγρασία (8784 ωριαίες μετρήσεις)

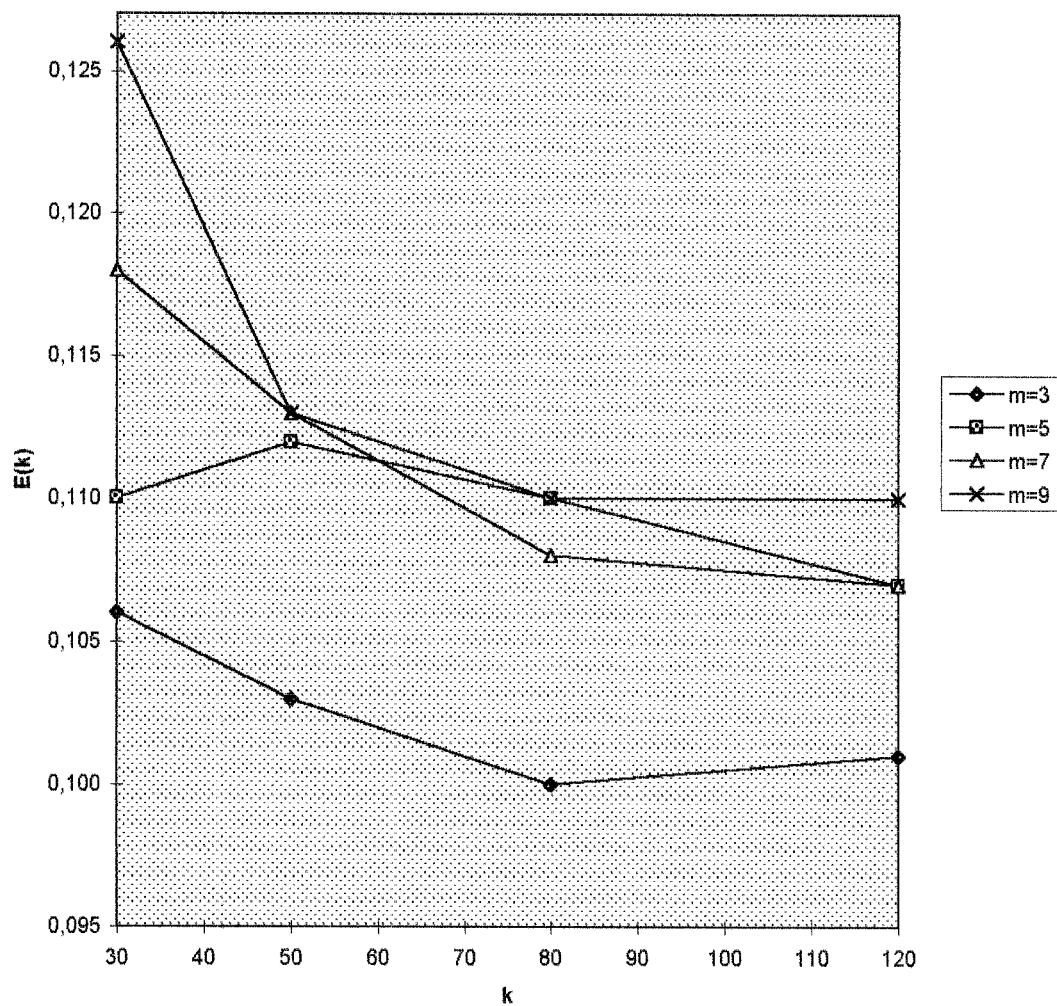
Θα γίνει επεξεργασία της υγρασίας για πρόβλεψη 1, 6, 12, 24 ωρών.

i) T=1h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-1=183 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.9 και γραφικά στο σχήμα B.13 .

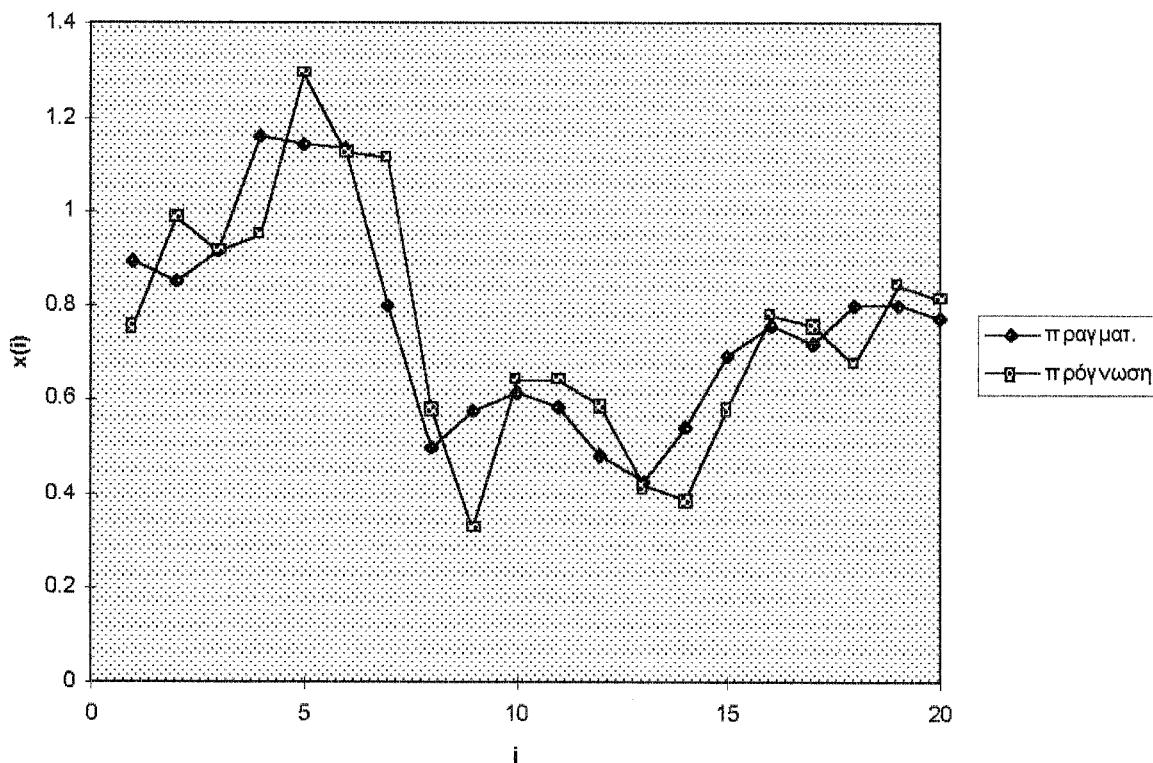
Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,106
	50	0,103
	80	0,100
	120	0,101
5	30	0,111
	50	0,112
	80	0,110
	120	0,107
7	30	0,118
	50	0,113
	80	0,108
	120	0,107
9	30	0,126
	50	0,113
	80	0,110
	120	0,110

Πίνακας B.9. Αποτελέσματα πρόγνωσης της υγρασίας με ορίζοντα 1 ώρα.



Σχήμα B.13. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ υγρασίας για $T=1h$

Στο σχήμα B.14 δίδεται το γράφημα πραγματικών τιμών - πρόγνωσης για τις τελευταίες 20 προγνώσεις στο δείγμα για $m=3$, $k=80$.



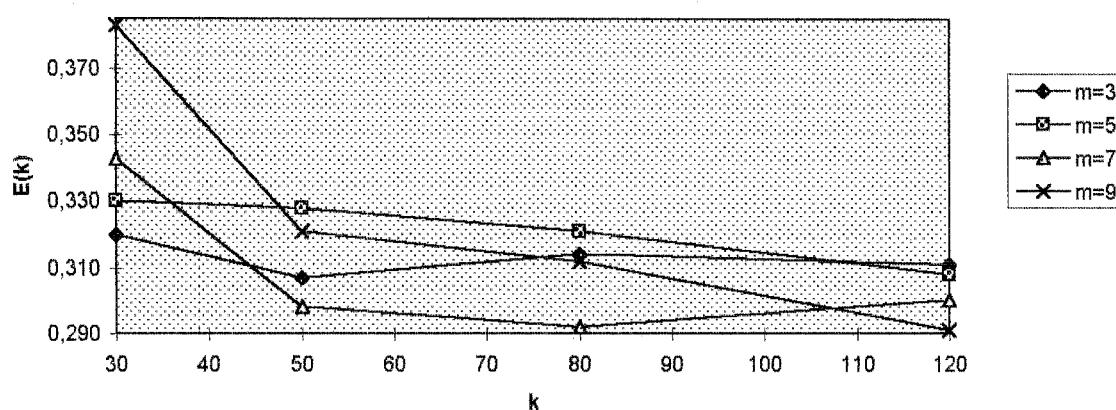
Σχήμα B.14. Γράφημα πραγματικής τιμής - πρόγνωσης για ορίζοντα 1 ώρας, των τελευταίων 20 ωρών του δείγματος τιμών υγρασίας ($m=3$, $k=80$).

ii) $T=6h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 8600 (δηλαδή θα γίνουν $8784-8600-6=178$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.10 και γραφικά στο σχήμα B.15.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,320
	50	0,307
	80	0,314
	120	0,311
5	30	0,330
	50	0,328
	80	0,321
	120	0,308
7	30	0,343
	50	0,298
	80	0,292
	120	0,300
9	30	0,383
	50	0,321
	80	0,312
	120	0,291

Πίνακας B.10. Αποτελέσματα πρόγνωσης της υγρασίας με ορίζοντα 6 ώρες.



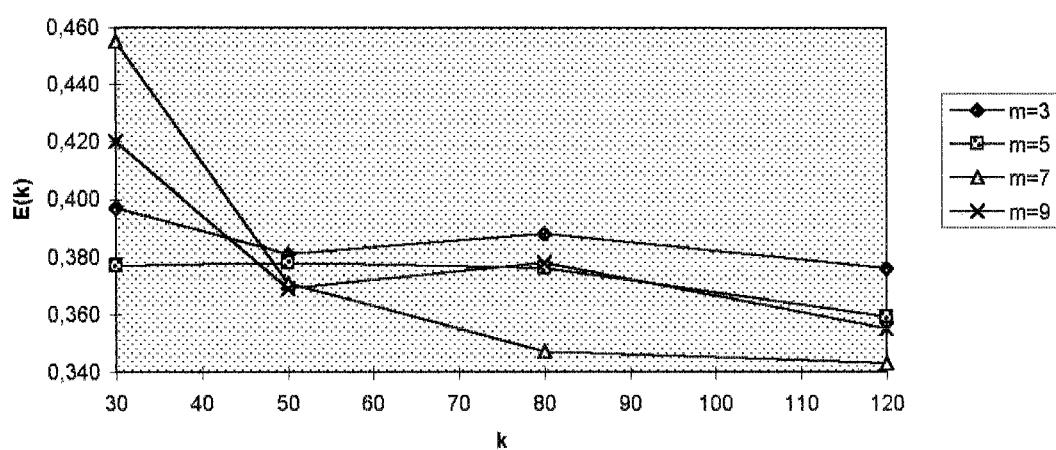
Σχήμα B.15. Γράφημα σφαλμάτων πρόγνωσης E(k) υγρασίας για T=6h

iii) T=12h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-12=172 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.11 και γραφικά στο σχήμα B.16.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,397
	50	0,381
	80	0,388
	120	0,376
5	30	0,377
	50	0,378
	80	0,376
	120	0,359
7	30	0,455
	50	0,371
	80	0,347
	120	0,343
9	30	0,420
	50	0,369
	80	0,378
	120	0,355

Πίνακας B.11. Αποτελέσματα πρόγνωσης της υγρασίας με ορίζοντα 12 ώρες.



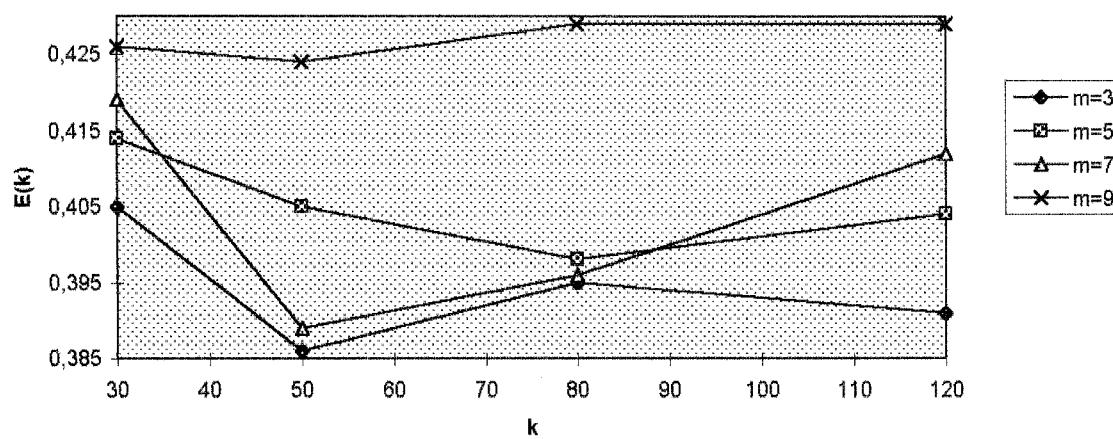
Σχήμα B.16. Γράφημα σφαλμάτων πρόγνωσης E(k) υγρασίας για T=12h

iv) T=24h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-24=160 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.12 και γραφικά στο σχήμα B.17.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,405
	50	0,386
	80	0,395
	120	0,391
5	30	0,414
	50	0,405
	80	0,398
	120	0,404
7	30	0,419
	50	0,389
	80	0,396
	120	0,412
9	30	0,426
	50	0,424
	80	0,429
	120	0,429

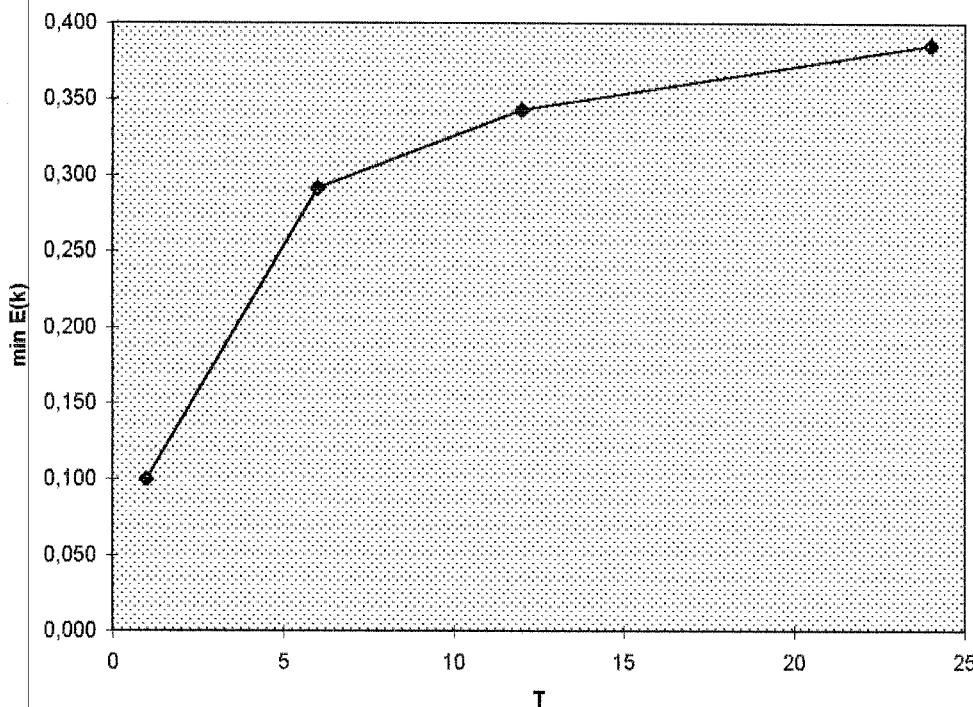
Πίνακας B.12. Αποτελέσματα πρόγνωσης της υγρασίας με ορίζοντα 24 ώρες.



Σχήμα B.17. Γράφημα σφαλμάτων πρόγνωσης E(k) υγρασίας για T=24h

v) Συσχέτιση των αποτελεσμάτων

Στο σχήμα B.18 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m , k .



Σχήμα B.18. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για την υγρασία

Η καμπύλη του σχήματος B.18 προσεγγίζεται βέλτιστα από τη συνάρτηση :

$$\min E(k)(T) \cong 0,109 + 0,0918 \ln(T)$$

με συντελεστή αυτοσυσχέτισης 0.993 .

Για $\min E(k)(T) = 0 \Rightarrow T \cong 0,305 \text{ h} \cong 18 \text{ min}$

Οι πολύ καλές μακροχρόνιες προβλέψεις σε αντίθεση με τις όχι ιδιαίτερα καλές βραχυχρόνιες (σε σύγκριση με τις αντίστοιχες της θερμοκρασίας και της ατμοσφαιρικής πίεσης) δείχνουν ύπαρξη ντετερμινισμού σε υψηλή χρονική κλίμακα.

δ. Βορινή συνιστώσα ταχύτητας ανέμου (8784 ωριαίες μετρήσεις)

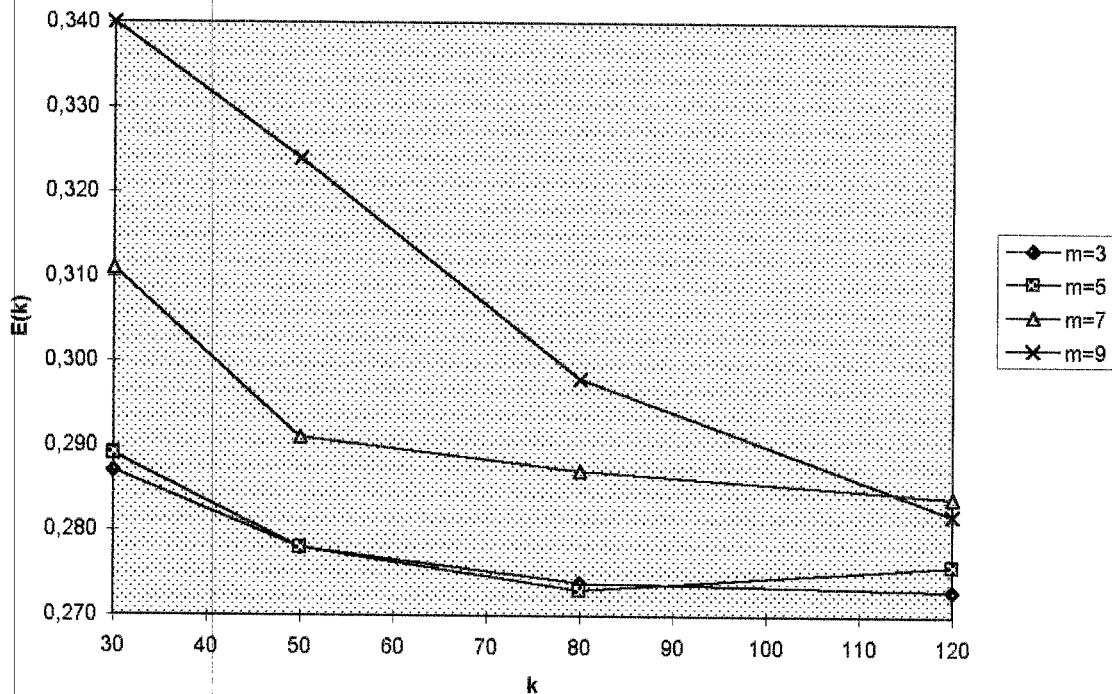
Η μέση ωριαία ταχύτητα ανέμου αναλύεται σε βορινή και ανατολική συνιστώσες. Θα γίνει επεξεργασία της βορινής συνιστώσας για πρόβλεψη 1, 2, 4, 6 ωρών.

i) T=1h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-1=183 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.13 και γραφικά στο σχήμα B.19 .

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,287
	50	0,278
	80	0,274
	120	0,273
5	30	0,289
	50	0,278
	80	0,273
	120	0,276
7	30	0,311
	50	0,291
	80	0,287
	120	0,284
9	30	0,340
	50	0,324
	80	0,298
	120	0,282

Πίνακας B.13. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας της ταχύτητας ανέμου, με ορίζοντα 1 ώρα.



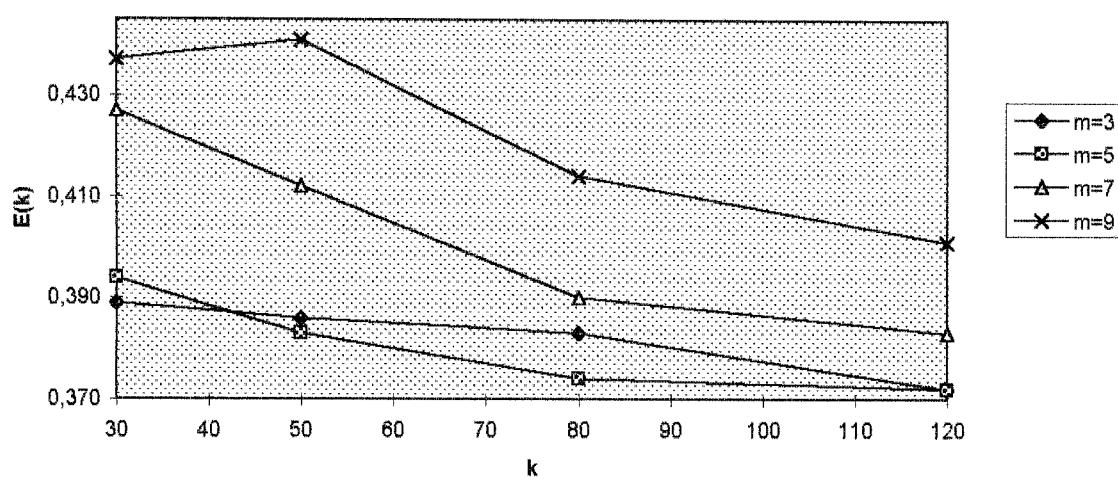
Σχήμα B.19. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας ανέμου για $T=1h$

ii) $T=2h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν $8784-8600-2=182$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.14 και γραφικά στο σχήμα B.20.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,389
	50	0,386
	80	0,383
	120	0,372
5	30	0,394
	50	0,383
	80	0,374
	120	0,372
7	30	0,427
	50	0,412
	80	0,390
	120	0,383
9	30	0,437
	50	0,441
	80	0,414
	120	0,401

Πίνακας B.14. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 2 ώρες.



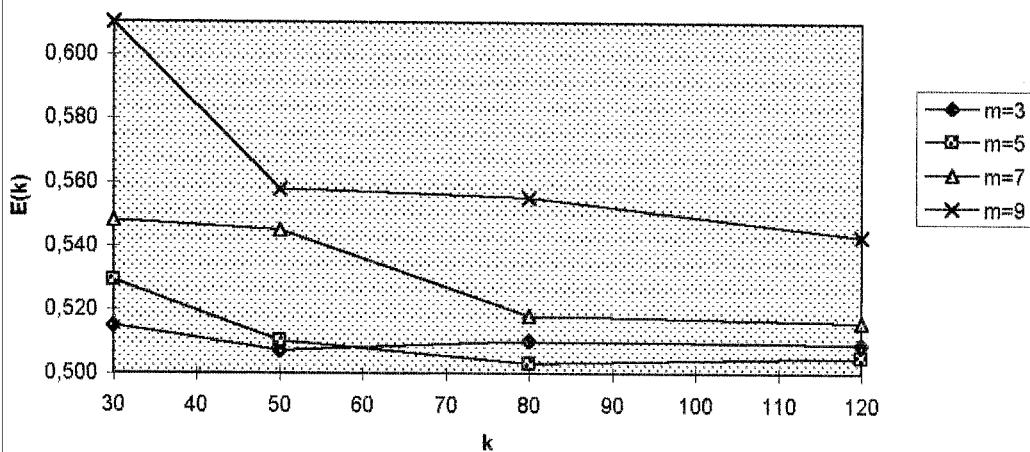
Σχήμα B.20. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ της βορινής συνιστώσας του ανέμου για $T=2h$

iii) T=4h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-4=180 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.15 και γραφικά στο σχήμα B.21.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,515
	50	0,507
	80	0,510
	120	0,509
5	30	0,529
	50	0,510
	80	0,503
	120	0,505
7	30	0,548
	50	0,545
	80	0,518
	120	0,516
9	30	0,610
	50	0,558
	80	0,555
	120	0,543

Πίνακας B.15. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 4 ώρες.



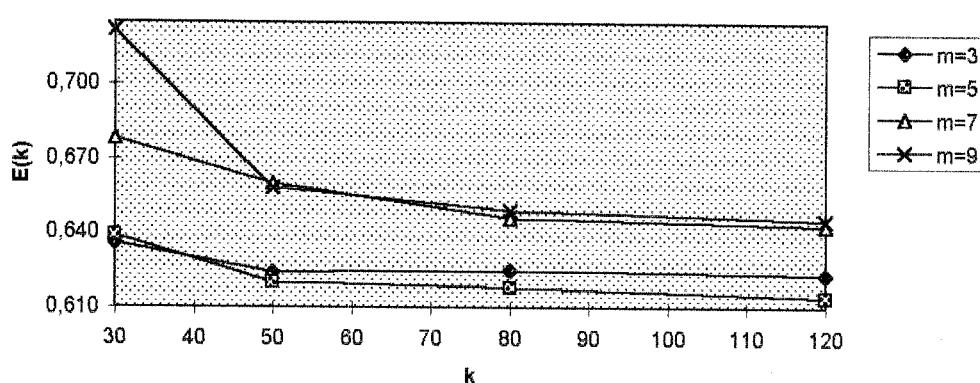
Σχήμα B.21. Γράφημα σφαλμάτων πρόγνωσης E(k) βορινής συνιστώσας του ανέμου για T=4h

iv) T=6h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-6=178 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.16 και γραφικά στο σχήμα B.22.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,636
	50	0,624
	80	0,625
	120	0,623
5	30	0,639
	50	0,620
	80	0,618
	120	0,614
7	30	0,678
	50	0,660
	80	0,646
	120	0,643
9	30	0,722
	50	0,658
	80	0,649
	120	0,645

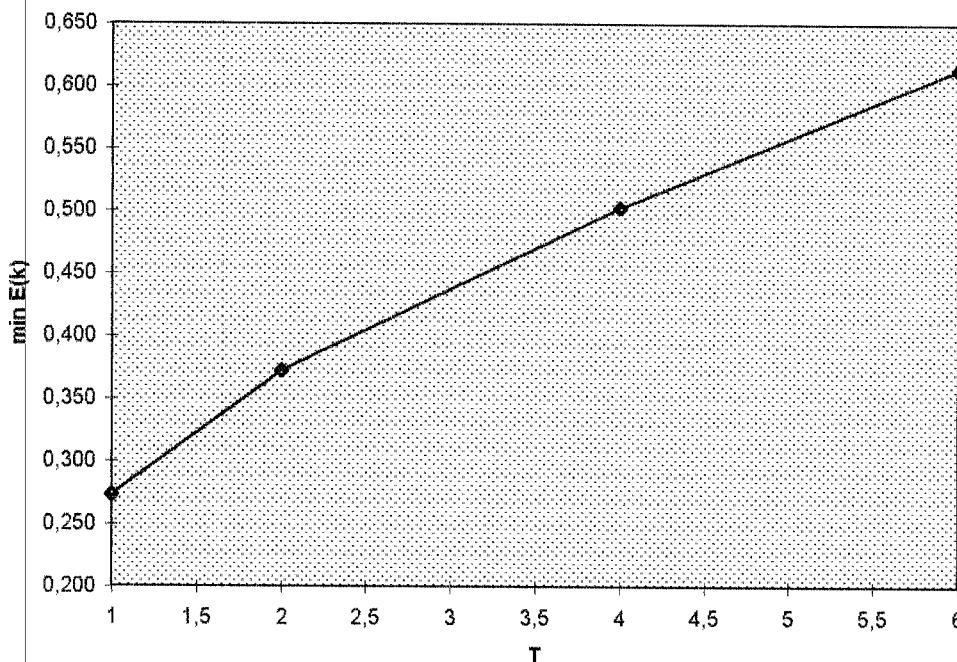
Πίνακας B.16. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 6 ώρες.



Σχήμα B.22. Γράφημα σφαλμάτων πρόγνωσης E(k) βορινής συνιστώσας του ανέμου για T=6h

v) Συσχέτιση των αποτελεσμάτων

Στο σχήμα B.23 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m , k .



Σχήμα B.23. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για τη βορινή συνιστώσα της ταχύτητας του ανέμου.

Η καμπύλη του σχήματος B.23 προσεγγίζεται βέλτιστα από τη συνάρτηση :

$$\min E(k)(T) \cong 0,272 T^{0,449}$$

με συντελεστή αυτοσυγχέτισης 0.9998 .

Για $\min E(k)(T) \rightarrow 0 \Rightarrow T \rightarrow 0$

ε. Ανατολική συνιστώσα ταχύτητας ανέμου (8784 ωριαίες μετρήσεις)

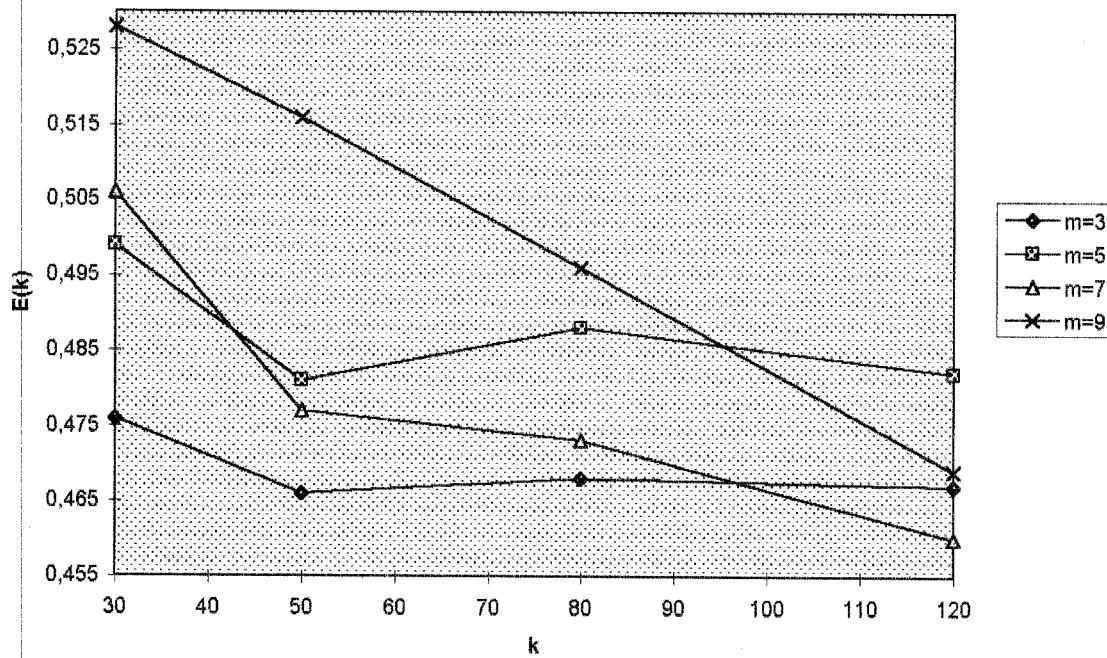
Η μέση ωριαία ταχύτητα ανέμου αναλύεται σε βορινή και ανατολική συνιστώσες. Θα γίνει επεξεργασία της ανατολικής συνιστώσας για πρόβλεψη 1, 2, 4, 6 ωρών.

i) T=1h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-1=183 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.17 και γραφικά στο σχήμα B.24 .

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,476
	50	0,466
	80	0,468
	120	0,467
5	30	0,499
	50	0,481
	80	0,488
	120	0,482
7	30	0,506
	50	0,477
	80	0,473
	120	0,460
9	30	0,528
	50	0,516
	80	0,496
	120	0,469

Πίνακας B.17. Αποτελέσματα πρόγνωσης της ανατολικής συνιστώσας της ταχύτητας ανέμου, με ορίζοντα 1 ώρα.



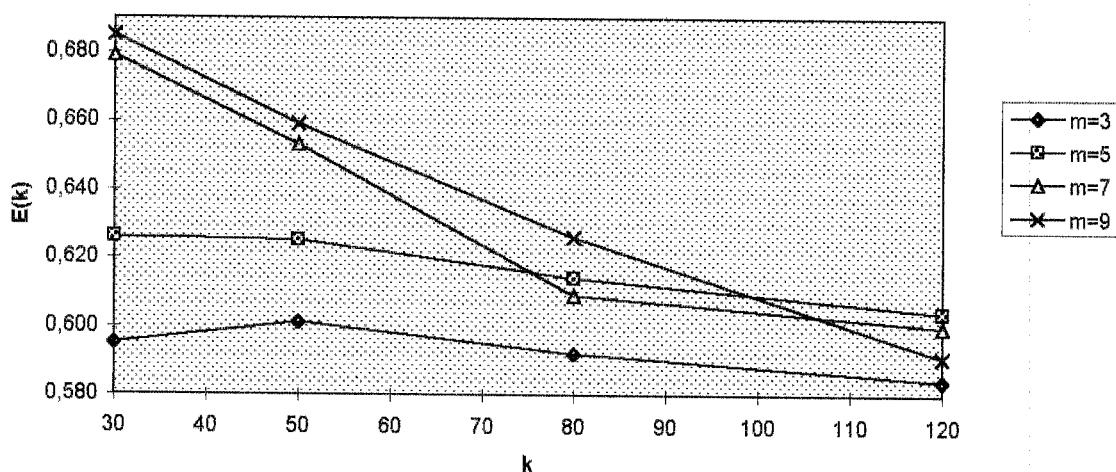
Σχήμα B.24. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ ανατολικής συνιστώσας ανέμου για $T=1h$

ii) $T=2h$

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 8600 (δηλαδή θα γίνουν $8784-8600-2=182$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.18 και γραφικά στο σχήμα B.25.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,595
	50	0,601
	80	0,592
	120	0,584
5	30	0,626
	50	0,625
	80	0,614
	120	0,604
7	30	0,679
	50	0,653
	80	0,609
	120	0,600
9	30	0,685
	50	0,659
	80	0,626
	120	0,591

Πίνακας B.18. Αποτελέσματα πρόγνωσης της ανατολικής συνιστώσας του ανέμου με ορίζοντα 2 ώρες.



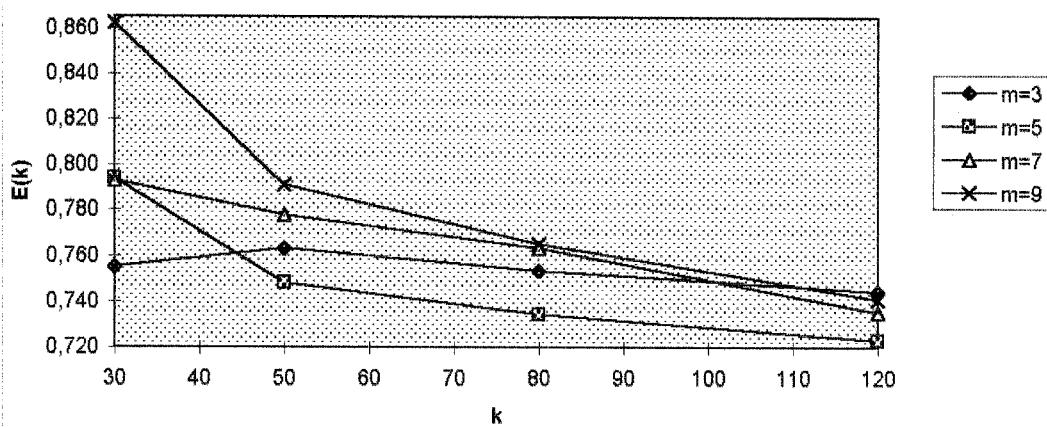
Σχήμα B.25. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ της ανατολικής συνιστώσας του ανέμου για $T=2h$

iii) T=4h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-4=180 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.19 και γραφικά στο σχήμα B.26.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,755
	50	0,763
	80	0,753
	120	0,744
5	30	0,794
	50	0,748
	80	0,734
	120	0,723
7	30	0,793
	50	0,778
	80	0,763
	120	0,735
9	30	0,862
	50	0,791
	80	0,765
	120	0,741

Πίνακας B.19. Αποτελέσματα πρόγνωσης της ανατολικής συνιστώσας του ανέμου με ορίζοντα 4 ώρες.



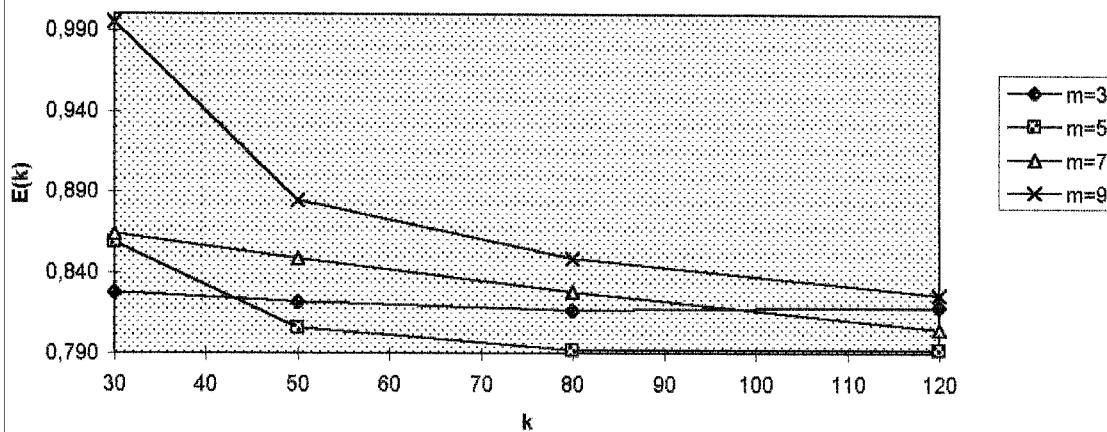
Σχήμα B.26. Γράφημα σφαλμάτων πρόγνωσης E(k) ανατολικής συνιστώσας του ανέμου για T=4h

iv) T=6h

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 8600 (δηλαδή θα γίνουν 8784-8600-6=178 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.20 και γραφικά στο σχήμα B.27.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,828
	50	0,822
	80	0,817
	120	0,819
5	30	0,859
	50	0,806
	80	0,792
	120	0,792
7	30	0,864
	50	0,849
	80	0,828
	120	0,805
9	30	0,995
	50	0,885
	80	0,849
	120	0,826

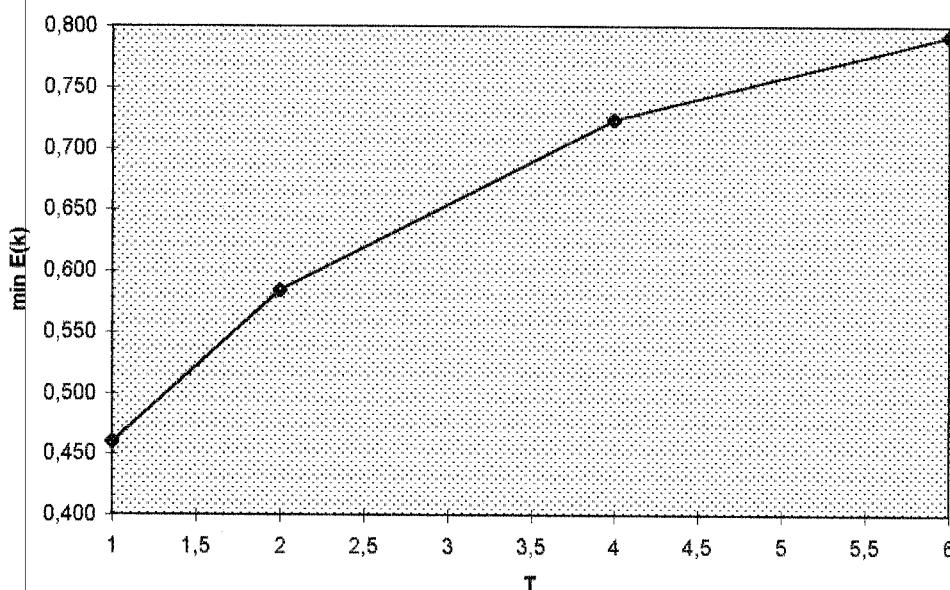
Πίνακας B.20. Αποτελέσματα πρόγνωσης της ανατολικής συνιστώσας του ανέμου με ορίζοντα 6 ώρες.



Σχήμα B.27. Γράφημα σφαλμάτων πρόγνωσης E(k) ανατολικής συνιστώσας του ανέμου για T=6h

v) Συσχέτιση των αποτελεσμάτων

Στο σχήμα B.28 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m , k .



Σχήμα B.28. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για την ανατολική συνιστώσα της ταχύτητας του ανέμου

Η ανατολική συνιστώσα της ταχύτητας του ανέμου έχει πολύ μεγαλύτερο σφάλμα πρόγνωσης από τη βορινή. Αυτό πιθανότατα συμβαίνει επειδή ο βορινός είναι ο επικρατών άνεμος.

Η καμπύλη του σχήματος B.28 προσεγγίζεται βέλτιστα από τη συνάρτηση :

$$\text{min } E(k)(T) = 0,458 + 0,187 \ln(T)$$

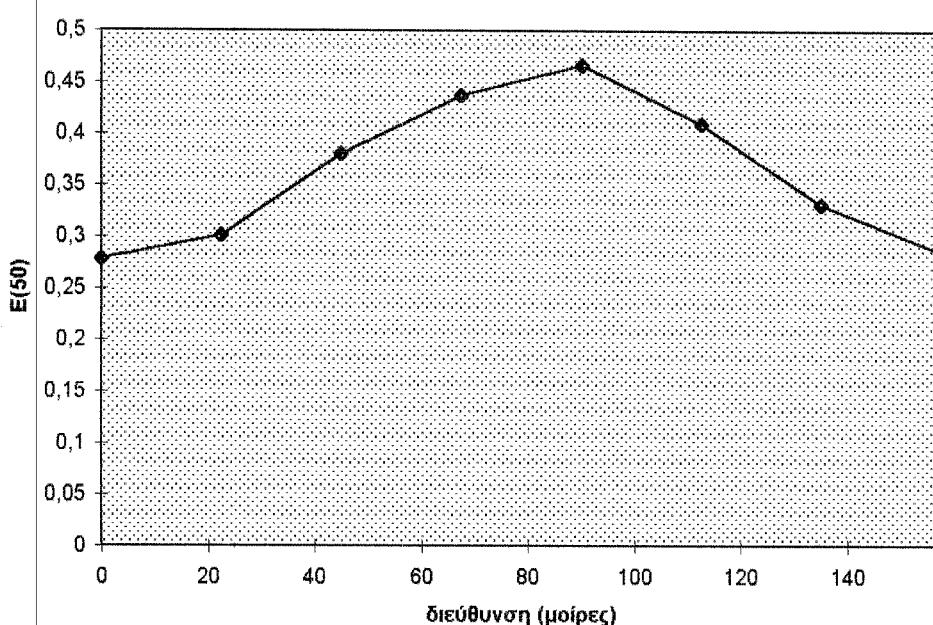
με συντελεστή αυτοσυγχέτησης 0,9996 .

**στ. Συσχέτιση προβολών της ταχύτητας του ανέμου σε άξονες γωνίας φ
(φ=0 : βοράς, φ=90 : ανατολή)**

Για ορίζοντα πρόγνωσης 1 ώρα γίνεται επεξεργασία της προβολής της ταχύτητας του ανέμου σε άξονες 0, 22,5, 45, 67,5, 90, 112,5, 135, 157,5 μοιρών. Η διεύθυνση 0 μοιρών αντιστοιχεί στο βορά και η διεύθυνση 90 μοιρών στην ανατολή. Η επεξεργασία γίνεται για : Nf=8600, m=3, k=50. Τα αποτελέσματα δίδονται στον πίνακα B.21 και το σχήμα B.29.

Γωνία προβολής	E(50)
0	0,278
22,5	0,301
45	0,381
67,5	0,437
90	0,466
112,5	0,410
135	0,331
157,5	0,285

Πίνακας B.21 Μέσο σφάλμα πρόγνωσης της προβολής της ταχύτητας του ανέμου σε άξονα, συναρτήσει της γωνίας του άξονα



Σχήμα B.29 Μέσο σφάλμα πρόγνωσης της προβολής της ταχύτητας του ανέμου σε άξονα, συναρτήσει της γωνίας του άξονα

ζ. Βορινή συνιστώσα ταχύτητας ανέμου (6897 10-λεπτες μετρήσεις)

Η μέση δεκάλεπτη ταχύτητα ανέμου αναλύεται σε βορινή και ανατολική συνιστώσες.

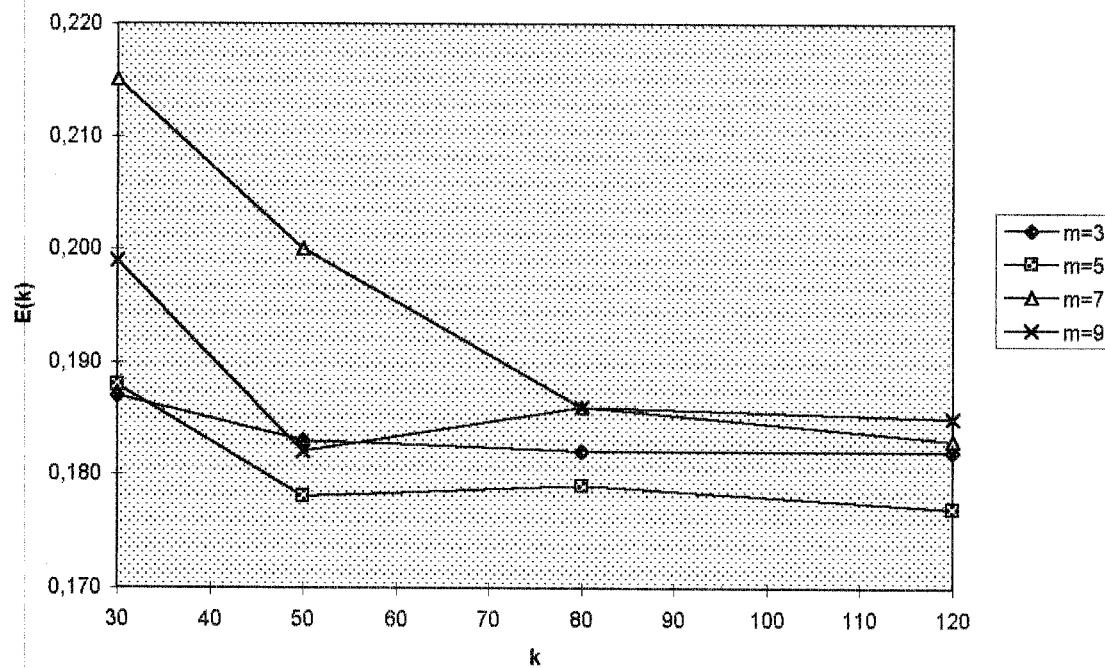
Θα γίνει επεξεργασία της βορινής συνιστώσας για πρόβλεψη 1, 2, 4, 6, 12, 24, 36 10-λέπτων.

i) $T=1\times10$ λεπτά

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 6700 (δηλαδή θα γίνουν 6897-6700-1=196 προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.22 και γραφικά στο σχήμα B.30 .

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,187
	50	0,183
	80	0,182
	120	0,182
5	30	0,188
	50	0,178
	80	0,179
	120	0,177
7	30	0,215
	50	0,200
	80	0,186
	120	0,183
9	30	0,199
	50	0,182
	80	0,186
	120	0,185

Πίνακας B.22. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας της ταχύτητας ανέμου, με ορίζοντα 1×10 λεπτά.



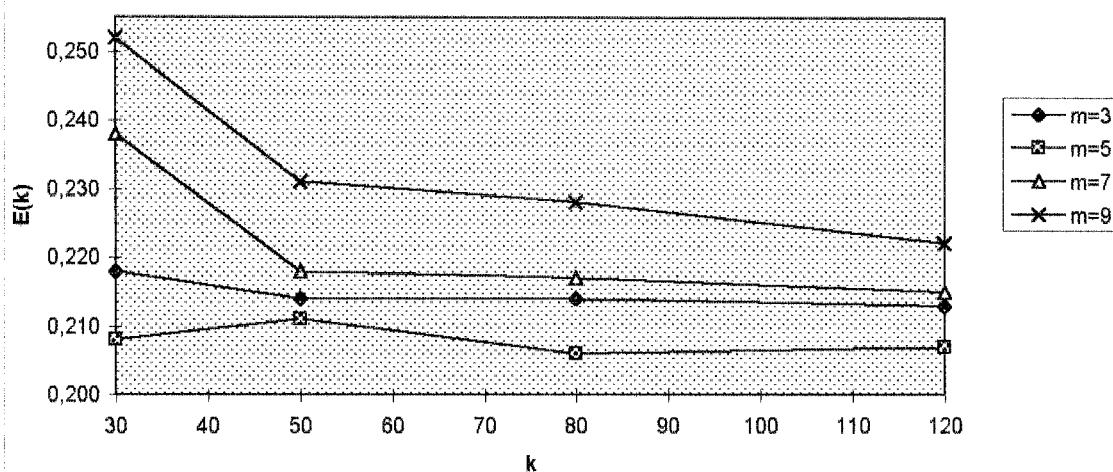
Σχήμα B.30. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας ανέμου για $T=1 \times 10$ λεπτά

ii) $T=2 \times 10$ λεπτά

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 - 2 = 195$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.23 και γραφικά στο σχήμα B.31.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,218
	50	0,214
	80	0,214
	120	0,213
5	30	0,208
	50	0,211
	80	0,206
	120	0,207
7	30	0,238
	50	0,218
	80	0,217
	120	0,215
9	30	0,252
	50	0,231
	80	0,228
	120	0,222

Πίνακας B.23. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 2×10 λεπτά.



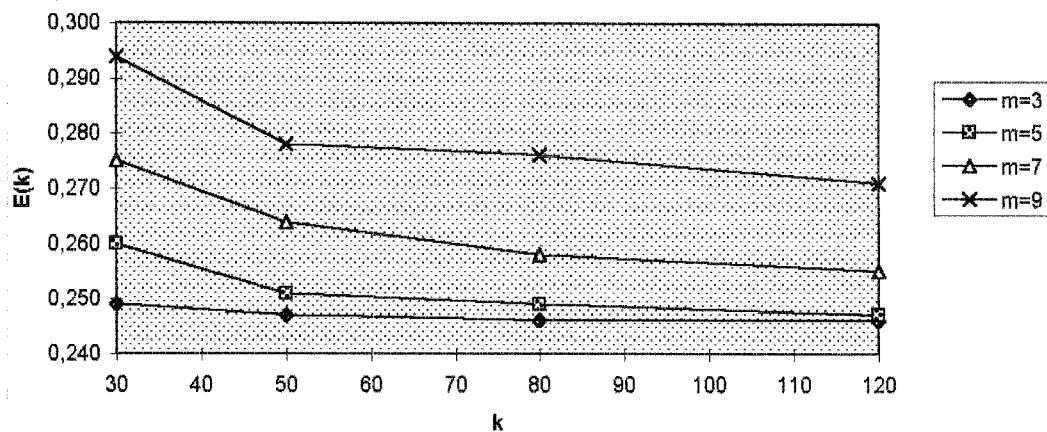
Σχήμα B.31. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ της βορινής συνιστώσας του ανέμου για $T=2 \times 10$ λεπτά

iii) $T=4 \times 10$ λεπτά

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 = 193$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.24 και γραφικά στο σχήμα B.32.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,249
	50	0,247
	80	0,246
	120	0,246
5	30	0,260
	50	0,251
	80	0,249
	120	0,247
7	30	0,275
	50	0,264
	80	0,258
	120	0,255
9	30	0,294
	50	0,278
	80	0,276
	120	0,271

Πίνακας B.24. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 4×10 λεπτά .



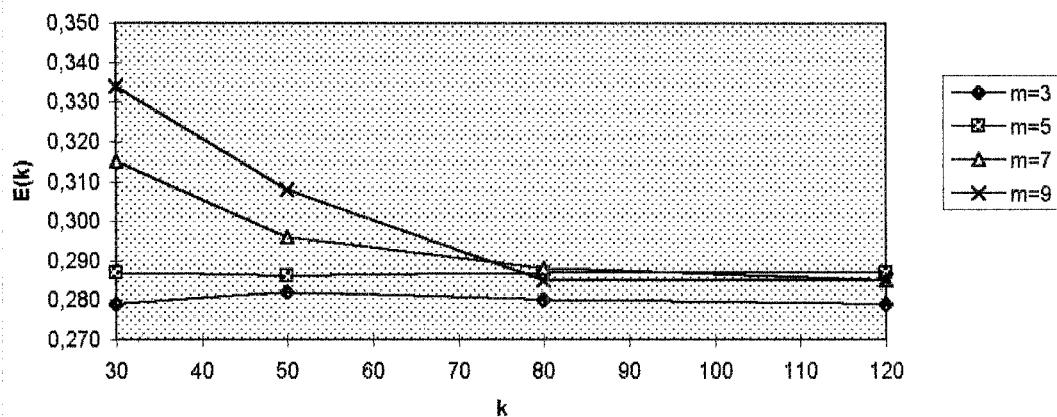
Σχήμα B.32. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας του ανέμου για $T=4 \times 10$ λεπτά

iv) $T=6 \times 10$ λεπτά (=1 ώρα)

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος κ των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 = 191$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.25 και γραφικά στο σχήμα B.33.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,279
	50	0,282
	80	0,280
	120	0,279
5	30	0,287
	50	0,286
	80	0,287
	120	0,287
7	30	0,315
	50	0,296
	80	0,288
	120	0,285
9	30	0,334
	50	0,308
	80	0,285
	120	0,285

Πίνακας B.25. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 6×10 λεπτά



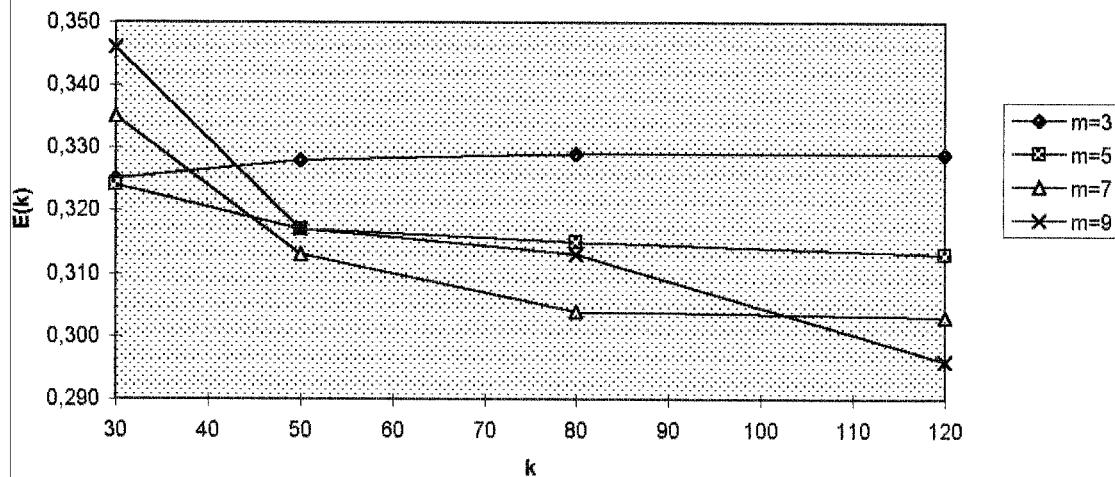
Σχήμα B.33. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας του ανέμου για $T=6 \times 10$ λεπτά

v) $T=12 \times 10$ λεπτά (=2 ώρες)

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος Nf τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 = 185$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.26 και γραφικά στο σχήμα B.34.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,325
	50	0,328
	80	0,329
	120	0,329
5	30	0,324
	50	0,317
	80	0,315
	120	0,313
7	30	0,335
	50	0,313
	80	0,304
	120	0,303
9	30	0,346
	50	0,317
	80	0,313
	120	0,296

Πίνακας B.26. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 12×10 λεπτά .



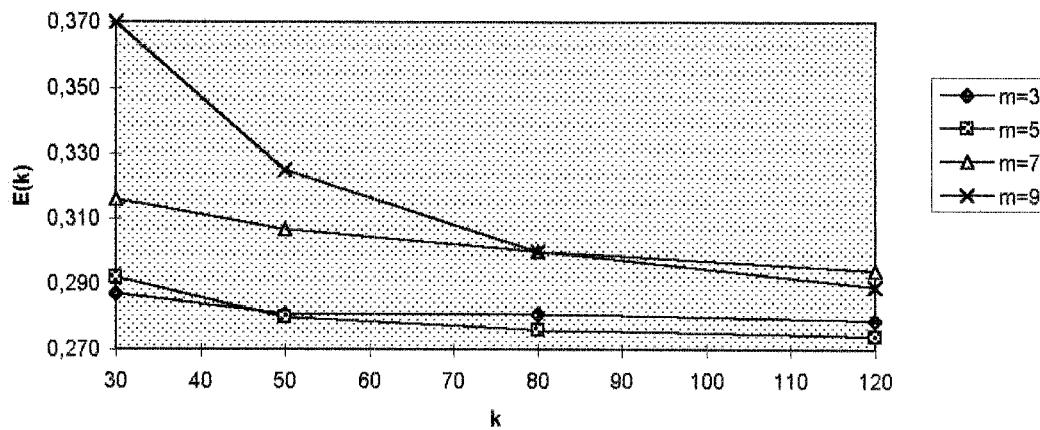
Σχήμα B.34. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ της βορινής συνιστώσας του ανέμου για $T=12 \times 10$ λεπτά

vi) $T=24 \times 10$ λεπτά (=4 ώρες)

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 = 173$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.27 και γραφικά στο σχήμα B.35.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, $E(k)$
3	30	0,287
	50	0,281
	80	0,281
	120	0,279
5	30	0,292
	50	0,280
	80	0,276
	120	0,274
7	30	0,316
	50	0,307
	80	0,300
	120	0,294
9	30	0,370
	50	0,325
	80	0,300
	120	0,289

Πίνακας B.27. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 24×10 λεπτά.



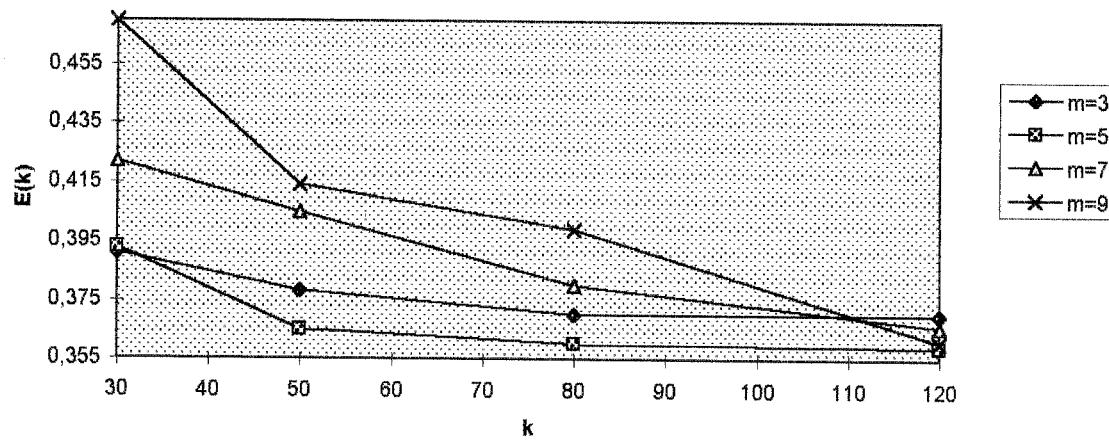
Σχήμα B.35. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας του ανέμου για $T=24 \times 10$ λεπτά

vii) $T=36 \times 10$ λεπτά (=6 ώρες)

Γίνεται έλεγχος σε 3, 5, 7, 9 διαστάσεις. Το πλήθος k των γειτόνων τίθεται διαδοχικά 30, 50, 80, 120. Το μέγεθος N_f τίθεται 6700 (δηλαδή θα γίνουν $6897 - 6700 - 36 = 161$ προβλέψεις). Τα αποτελέσματα δίδονται αριθμητικά στον πίνακα B.28 και γραφικά στο σχήμα B.36.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, $E(k)$
3	30	0,391
	50	0,378
	80	0,370
	120	0,370
5	30	0,393
	50	0,365
	80	0,360
	120	0,359
7	30	0,422
	50	0,405
	80	0,380
	120	0,366
9	30	0,470
	50	0,414
	80	0,399
	120	0,361

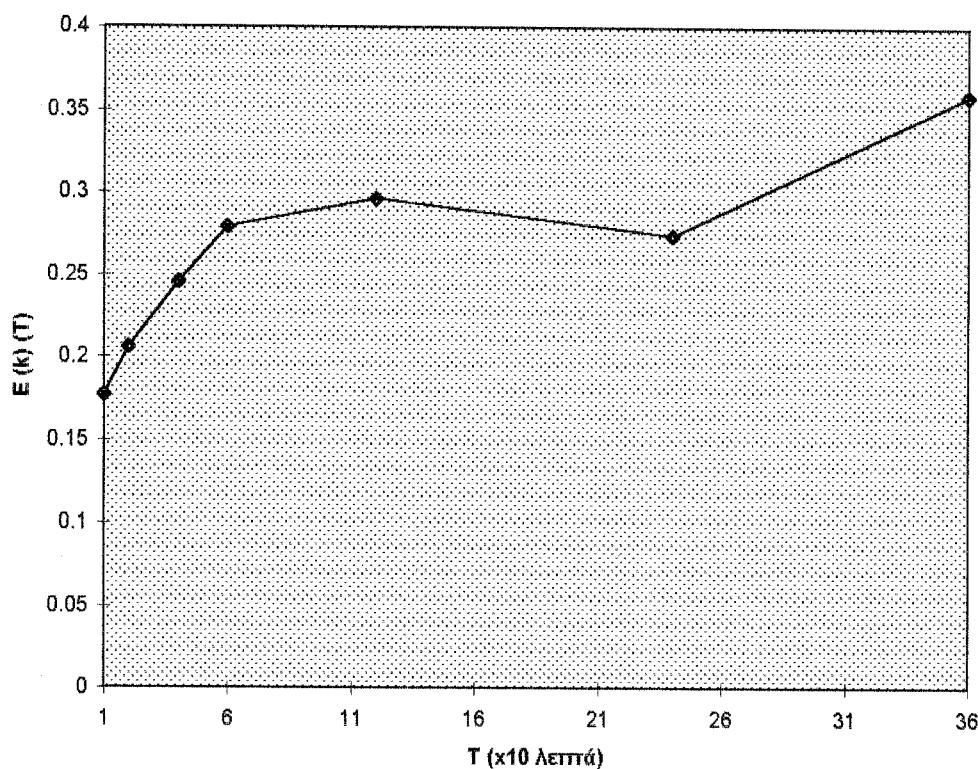
Πίνακας B.28. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 36×10 λεπτά.



Σχήμα B.36. Γράφημα σφαλμάτων πρόγνωσης $E(k)$ βορινής συνιστώσας του ανέμου για $T=36 \times 10$ λεπτά

viii) Συσχέτιση των αποτελεσμάτων

Στο σχήμα B.37 δίδεται η γραφική παράσταση του σφάλματος πρόβλεψης συναρτήσει του ορίζοντα, για τις βέλτιστες επιλογές m , k .



Σχήμα B.37. Ελάχιστο σφάλμα πρόγνωσης συναρτήσει του ορίζοντα πρόγνωσης για τη βορινή συνιστώσα της ταχύτητας του ανέμου.

Το παράδοξο αυτού του γραφήματος είναι ότι έχουμε καλύτερη πρόγνωση για ορίζοντα 4 ωρών από ότι για ορίζοντα 1 και 2 ωρών. Αυτό συμβαίνει όταν ανακαλύπτεται κάποια συχνότητα στο δείγμα. Τότε δεσμεύονται από το μοντέλο πρόσθετοι βαθμοί ελευθερίας.

Σε αυτήν την περίπτωση μπορούμε να βελτιώσουμε την πρόγνωση για ορίζοντα 1 ή 2 ωρών αν αλλάξουμε το χρονικό βήμα στην κατασκευή των διανυσμάτων του χώρου των φάσεων. Το χρονικό βήμα που χρησιμοποιήσαμε ως τώρα ήταν $1(\times 10 \text{ λεπτά})$ (ή $\times 1 \text{ ώρα}$, στην περίπτωση των δειγμάτων μέσων ωριαίων τιμών)), δηλαδή $t_i = 1, 2, 3, \dots$. Για να βελτιώσουμε την πρόβλεψη 2 ωρών τη μετατρέπουμε σε πρόβλεψη 4 ωρών. Αυτό μπορεί να γίνει δίνοντας το πρώτο βήμα $12(\times 10 \text{ λεπτά})$ ώστε να απέχει $24(\times 10 \text{ λεπτά})$ από την πρόβλεψη : $t_i = 12, 13, 14, \dots$. Αυτή είναι μια πρόβλεψη 4 ωρών, με μία επιπλέον τιμή γνωστή 2 ώρες πριν την πρόβλεψη. Για $m=3$, $t_i=12, 13, 14$, δίνει $E(80)=0,276$.

Αυτό όμως που είναι σημαντικό, είναι ότι με χρήση 10-λεπτου δείγματος φαίνεται να μπορούμε να κάνουμε πολύ καλύτερες μακροχρόνιες προγνώσεις από ότι με χρήση ωριαίου, στο συγκεκριμένο μέγεθος. Για το πλήθος των προγνώσεων που κάναμε - της τάξης των 150-200 - φαίνεται ότι τα αποτελέσματα υπερεκτιμήθηκαν. Το δείγμα ελέγχου δεν είναι μικρό σε μέγεθος αλλά σε χρονική κλίμακα. Όταν αναλύσαμε τη χρονοσειρά μέσων ωριαίων τιμών είχαμε δείγμα ελέγχου της τάξης των 150 σημείων. Αυτά αντιστοιχούσαν σε 150 ώρες. Οι 150 ώρες όμως αν αναχθούν σε 10-λεπτα είναι 900 σημεία. Αν ανξήσουμε το δείγμα ελέγχου - θέτοντας $Nf=6000$ - σε $6897-6000-24=873$ σημεία για πρόγνωση $24 \times 10 \text{ λεπτών}$ έχουμε τα αποτελέσματα του πίνακα B.29 .

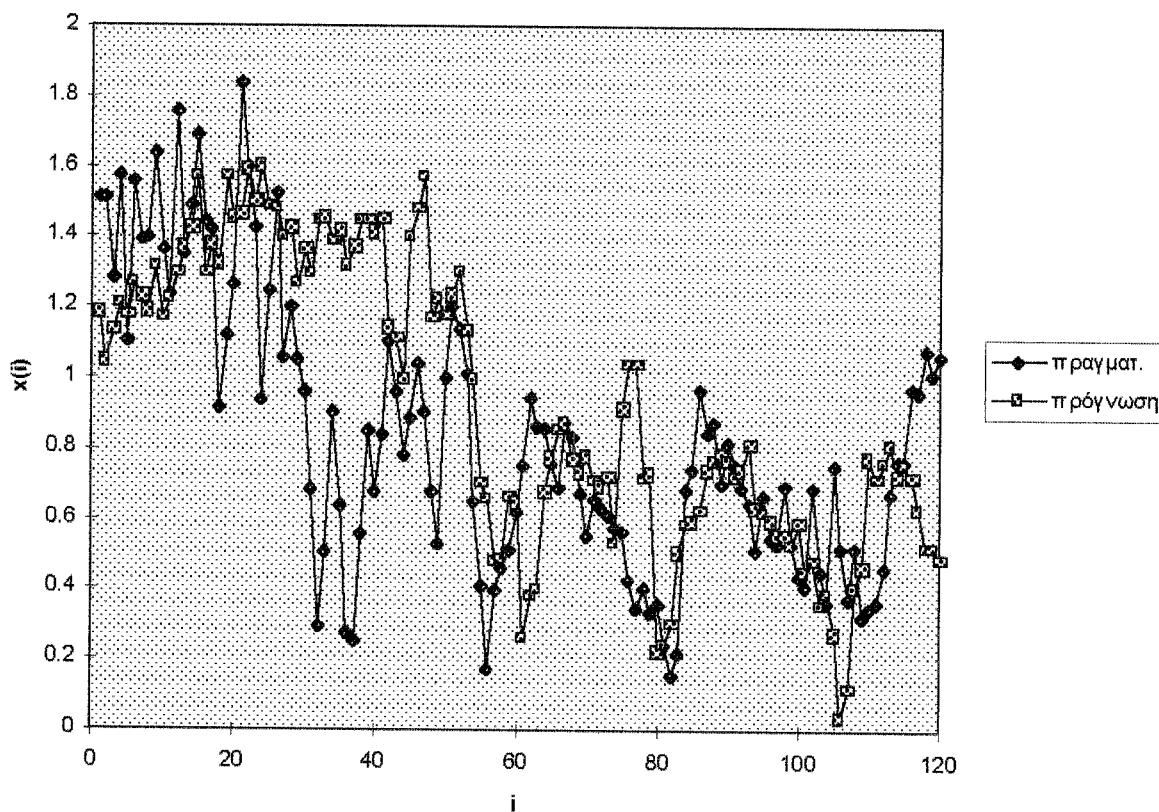
Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
3	30	0,412
	50	0,405
	80	0,404
	120	0,402
5	30	0,428
	50	0,415
	80	0,410
	120	0,407
7	30	0,463
	50	0,439
	80	0,421
	120	0,411
9	30	0,501
	50	0,456
	80	0,438
	120	0,415

Πίνακας B.29. Αποτελέσματα πρόγνωσης της βορινής συνιστώσας του ανέμου με ορίζοντα 24×10 λεπτά για $Nf=6000$.

Το μικρότερο σφάλμα πρόγνωσης είναι $E(120)=0,402$ για $m=3$. Για να είναι συγκρίσιμο με το σφάλμα πρόγνωσης 4 ωρών από αρχεία μέσων ωριαίων τιμών, πρέπει να πολλαπλασιαστεί επί $1,140=3,417/2,997$, όπου 3,417 είναι η τυπική απόκλιση του δείγματος 10-λέπτων μετρήσεων και 2,997 η τυπική απόκλιση του δείγματος ωριαίων μετρήσεων. Έτσι έχουμε σφάλμα $0,402 \times 1,140 = 0,458 < 0,503$. Δηλαδή τα αρχεία 10-λέπτων δίνουν καλύτερη πρόγνωση από τα ωριαία. Πρέπει τέλος να σημειωθεί ότι το πραγματικό μέσο σφάλμα για πρόγνωση 4 ωρών από αρχεία 10-λέπτων είναι μικρότερο του 0,458 αφού η μέση τιμή των απολύτων σφαλμάτων έξι 10-λέπτων είναι μεγαλύτερη του απόλυτου μέσου σφαλματος μιας ώρας, κατά την ανισότητα $|α| + |β| + \dots \geq |α + β + \dots|$.

Με διορθωμένα τα σφάλματα πρόβλεψης για $Nf=6000$, η βέλτιστη προσέγγιση του σφαλματος ακολουθεί την καμπύλη : $\min E(k)(T) = 0,172 T^{0,277}$.

Στο σχήμα B.38 δίδεται το γράφημα πραγματικών τιμών - πρόβλεψης για τα τελευταία 120 σημεία του αρχείου μέσων 10-λέπτων τιμών της βορινής συνιστώσας της ταχύτητας του ανέμου, για πρόβλεψη 24×10 λεπτών και $m=3$, $k=80$.



Σχήμα B.38. Γράφημα πραγματικής τιμής- πρόβλεψης για ορίζοντα 24($\times 10$ λεπτά)

Στο σχήμα B.38 φαίνεται ότι τα μεγάλα σφάλματα πρόβλεψης είναι σε περιοχές που έχουμε μεταβολή της τάσης. Το πρόβλημα είναι ότι το μοντέλο λειτουργεί σε μικρή χρονική κλίμακα (χρονικό βήμα 10 λεπτά) και δεν μπορεί να παρακολουθήσει αλλαγές μεγάλης κλίμακας. Τσως να χρειάζεται να συνεργάζονται περισσότερα του ενός μοντέλα, κάθε ένα από τα οποία θα καλύπτει διαφορετική κλίμακα, για να βελτιστοποιήσουμε μία πρόγνωση.

Τέλος αναφέρεται ότι και στην ατμοσφαιρική πίεση - της οποίας το ελάχιστο σφάλμα συναρτήσει του ορίζοντα πρόγνωσης προσεγγίζεται από καμπύλη δύναμης του ορίζοντα - η ανάλυση 10-λέπτων μετρήσεων έδωσε αποτελέσματα της τάξης μεγέθους των αποτελεσμάτων ωριαίων μετρήσεων. Αντίθετα στην υγρασία και την ανατολική συνιστώσα της ταχύτητας του ανέμου - που η καμπύλη τους είναι λογάριθμος του ορίζοντα πρόγνωσης - οι 10-λεπτες μετρήσεις δίνουν πολύ χειρότερες προγνώσεις. Στους πίνακες B.30, B.31, B.32 δίνονται τα αποτελέσματα της επεξεργασίας των μεγεθών : ατμοσφαιρική πίεση, υγρασία, ανατολική συνιστώσα της ταχύτητας του ανέμου αντίστοιχα. Η επεξεργασία έχει γίνει για $m=5$, $Nf=6000$.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
5	30	0,379
	50	0,360
	80	0,346
	120	0,340

Πίνακας B.30. Αποτελέσματα πρόγνωσης της ατμοσφαιρικής πίεσης με ορίζοντα 72×10 λεπτά για $Nf=6000$.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
5	30	0,916
	50	0,887
	80	0,873
	120	0,862

Πίνακας B.31. Αποτελέσματα πρόγνωσης της υγρασίας με ορίζοντα 72×10 λεπτά για $Nf=6000$.

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
5	30	0,569
	50	0,552
	80	0,547
	120	0,540

Πίνακας B.32. Αποτελέσματα πρόγνωσης της ανατολικής συνιστώσας του ανέμου με ορίζοντα 6×10 λεπτά για $Nf=6000$.

Τα παραπάνω σφάλματα για να είναι συγκρινόμενα με τα αντίστοιχα για ωριαίες μετρήσεις, πρέπει να πολλαπλασιαστούν με το συντελεστή : τυπ.απόκλ.10-λεπτών / τυπ.απόκλ.ωριαίων. Αυτός ο συντελεστής είναι : α. για την ατμ. πίεση $7,865/5,555=1,416$, β. για την υγρασία $7,461/11,323=0,659$, γ. για την ανατολική συνιστώσα της ταχύτητας του ανέμου $1,989/1,584=1,256$.

Έτσι τα διορθωμένα μικρότερα σφάλματα είναι :

α. Για την ατμ. πίεση : $1,416 \times 0,340 = 0,481 \approx 0,442$ των ωριαίων μετρήσεων.

β. Για την υγρασία : $0,659 \times 0,862 = 0,568 > 0,343$ των ωριαίων μετρήσεων.

γ. Για την αν. συνιστ.

ταχύτ. ανέμου : $1,256 \times 0,540 = 0,678 > 0,460$ των ωριαίων μετρήσεων.

Αν επαναλάβουμε την ίδια διαδικασία για τη θερμοκρασία θα δούμε ότι με τις 10-λεπτες μετρήσεις μπορούμε να κάνουμε καλύτερη πρόγνωση από ότι με τις ωριαίες. Τα αποτελέσματα για τις 10-λεπτες και για $m=5$, $Nf=6000$ δίδονται στον πίνακα B.33 .

Διάσταση έδρασης, m	Αριθμός γειτόνων, k	Μέσο σφάλμα πρόγνωσης, E(k)
5	30	0,743
	50	0,723
	80	0,717
	120	0,708

Πίνακας B.33. Αποτελέσματα πρόγνωσης της θερμοκρασίας με ορίζοντα 72×10 λεπτά για $Nf=6000$.

Η θερμοκρασία είναι ένα μέγεθος με έντονη εποχιακή εξάρτηση. Η τυπική απόκλιση των 10-λέπτων, που καλύπτουν ένα χρονικό διάστημα δύο μηνών, είναι 3,364 η μισή περίπου της τυπικής απόκλισης των ωριαίων μετρήσεων, που καλύπτουν χρονικό διάστημα ενός έτους, και είναι 7,618 . Έτσι το σφάλμα των 10-λέπτων γίνεται $0,708 \times 3,364 / 7,618 = 0,313 < 0,391$ των ωριαίων μετρήσεων. Αυτή πάντως η σχέση οφείλεται μάλλον στον τοπικό χαρακτήρα του δείγματος των 10-λέπτων παρά στη μεγαλύτερη ανάλυση.

Συμπεράσματα

Στην εργασία αυτή χρησιμοποιήθηκε η μέθοδος των αναλόγων, η οποία αντλείται από την επιστημονική περιοχή των δυναμικών συστημάτων, για την πρόγνωση μετεωρολογικών μεταβλητών και συγκεκριμένα: θερμοκρασίας, σχετικής υγρασίας, ατμοσφαιρικής πίεσης και ταχύτητας ανέμου. Για να την εφαρμόσουμε στα αρχεία δεδομένων που είχαμε, κατασκευάσαμε ένα πρόγραμμα υπολογιστή που βασιζόταν στον αλγόριθμο DVS των Casdagli and Weigend (1992). Ο χρόνος που χρειαζόταν για να 'τρέξει' το πρόγραμμα (σε υπολογιστή Pentium στα 100 MHz), ποίκιλε ανάλογα με τις τιμές των διάσταση, πλήθος γειτόνων, πλήθος προγνώσεων, μεταξύ 5 λεπτών και 1 ώρας, πράγμα που φανερώνει ότι ο συγκεκριμένος αλγόριθμος είναι αρκετά γρήγορος. Δοκιμάστηκαν ορίζοντες πρόγνωσης από 10 λεπτά μέχρι 24 ώρες ενώ χρησιμοποιήθηκαν ιστορικά δεδομένα με χρονική κλίμακα 10 λεπτών και 1 ώρας. Γενικώς φάνηκε ότι για μικρό ορίζοντα πρόγνωσης το σφάλμα πρόγνωσης είναι αρκετά μικρό (της τάξης του $0,1 \times$ τυπική απόκλιση για την ατμοσφαιρική πίεση και τη θερμοκρασία, του $0,15 \times$ τυπική απόκλιση για την υγρασία και του $0,25 \times$ τυπική απόκλιση για τη βορινή συνιστώσα της ταχύτητας του ανέμου), ενώ για μεγαλύτερους ορίζοντες το σφάλμα γίνεται πολύ σημαντικό (της τάξης των $0,6$ τυπικών αποκλίσεων για ορίζοντα πρόγνωσης 24 ώρες). Αυτό δείχνει ότι υπάρχει αδυναμία για μακροχρόνια πρόβλεψη αυτών των μεταβλητών με τη συγκεκριμένη μέθοδο. Τα μοντέλα δε μπορούν να προγνώσουν την αλλαγή της κλίσης της τάσης (ανοδικής ή καθοδικής).

Πιο συγκεκριμένα στα μεγέθη που εξετάσαμε και από τα αποτελέσματα που πήραμε βρήκαμε ενδείξεις ότι: α) στη θερμοκρασία ο αλγόριθμος τείνει να αναγνωρίζει ίχνη μόνο ντετερμινισμού όσο αυξάνει ο ορίζοντας πρόγνωσης, β) η ατμοσφαιρική πίεση είναι ισχυρά μη αυτόνομο μέγεθος όταν εξετάζεται με σημειακές μετρήσεις, γ) η σχετική υγρασία έχει έντονη παρουσία ντετερμινισμού σε μεγάλη χρονική κλίμακα, δ) η καμπύλη του σφάλματος πρόγνωσης συναρτήσει του ορίζοντα συσχετίζεται με την αναγκαία χρονική ανάλυση των μετρήσεων για βελτιστοποίηση της πρόγνωσης.

Στην εργασία αυτή δεν αξιοποιήθηκε το γεγονός ότι τα ατμοσφαιρικά μεγέθη αλληλοσυσχετίζονται και μία μικρή αλλαγή π.χ. της ταχύτητας ή της διεύθυνσης του ανέμου μπορεί να προκαλέσει μεγάλη αλλαγή της τιμής της υγρασίας. Μία κατεύθυνση προς την οποία μπορούμε μελλοντικά να κινηθούμε είναι να κατασκευάσουμε μοντέλο που να λαμβάνει υπόψη του όλα τα μεγέθη πριν κάνει πρόβλεψη για την εξέλιξη ενός.

Επιπλέον είδαμε ότι η σύγκριση των σφαλμάτων πρόγνωσης μπορεί να οδηγήσει σε κάποια συμπεράσματα για τα μεγέθη, όπως στην περίπτωση της ταχύτητας του

ανέμου που η πολύ καλύτερη πρόγνωση για τη βορινή από ότι για την ανατολική συνιστώσα του συνδυάζεται με το γεγονός ότι ο βορινός είναι ο επικρατών άνεμος.

Τελικά φαίνεται να διαμορφώνονται οι εξής ανάγκες για μια καλύτερη αντιμετώπιση της πρόγνωσης των ατμοσφαιρικών μεγεθών :

α. Χρησιμοποίηση ιστορικών αρχείων πολύ μεγάλου μεγέθους και με την κατάλληλη χρονική ανάλυση ώστε να μπορούν να βρεθούν αρκετές κατάλληλες σειρές αναλόγων, κοντινές με την εξεταζόμενη.

β. Δημιουργία μοντέλων που να συσχετίζουν ταυτόχρονα διαφορετικά μεγέθη.

γ. Δημιουργία μοντέλων που να καλύπτουν περισσότερες χρονικές κλίμακες π.χ. μοντέλο που να χρησιμοποιεί μέσες τιμές 10-λέπτων και μέσες τιμές ωριαίων και μέσες τιμές ημερησίων μετρήσεων. Η αρχιτεκτονική ενός τέτοιου μοντέλου μοιάζει με την αρχιτεκτονική νευρωνικού δικτύου.

δ. Ταυτόχρονη χρησιμοποίηση μετρήσεων από γειτονικές περιοχές ώστε να γίνεται συσχετιζόμενη ανάλυσή τους.

Αναφορές

- Barnsley, M., *Fractals everywhere*, second edition, Academic Press Professional, 1993.
- Casdagli, M., A Dynamical Systems Approach to Modeling Input-Output Systems, Section 3, *Nonlinear Modeling and Forecasting*, edited by M. Casdagli and S. Eubank, Addison-Wesley, 1991.
- Casdagli, M. and Weigend, A., Exploring the Continuum Between Deterministic and Stochastic Modeling, Section 3, *Time-Series Prediction : Forecasting the Future and Understanding the Past*, edited by A. Weigend and N. Gershenfeld, Addison-Wesley, 1992.
- Grassberger, P., Estimating the Fractal Dimensions and Entropies of Strange Attractors, Part 5, *Chaos*, edited by A. V. Holden, Princeton University Press, 1986.
- Nusse, H. and Yorke, J., *Dynamics Numerical Explorations*, Springer Verlag, 1994.(βιβλίο και πρόγραμμα υπολογιστή)
- Sprott, J. and Rowlands, G., *Chaos Data Analyzer*, American Institute of Physics, 1991.(πρόγραμμα υπολογιστή)
- Tsonis, A., *Chaos, from Theory to Applications*, Plenum Press, 1992.

- Κουτσογιάννης, Δ., *Στατιστική Υδρολογία*, Έκδοση 2, Τυπογραφείο Ε.Μ.Π., 1996.
- Μιμίκου, Μ., *Στοχαστική Υδρολογία*, Ε.Μ.Π., 1985.
- Μπακόπουλος, Α. και Χρυσοβέργης, Ι., *Εισαγωγή στην Αριθμητική Ανάλυση*, Εκδόσεις Συμεών, 1985.
- Μπούντης, Α., Τρόποι Μετάβασης στο Χάος, *Τάξη και Χάος στα Μη Γραμμικά Δυναμικά Συστήματα*, Τόμος Δεύτερος, Επιστημονικοί Εκδότες Τ. Μπούντης και Στ. Πνευματικός, Εκδόσεις Γ.Α.Πνευματικός, 1989.
- Χατζηδημητρίου, Ι., Συντονισμός και Αστάθεια σε Δυναμικά Συστήματα, *Τάξη και Χάος στα Μη Γραμμικά Δυναμικά Συστήματα*, Τόμος Δεύτερος, Επιστημονικοί Εκδότες Τ. Μπούντης και Στ. Πνευματικός, Εκδόσεις Γ.Α.Πνευματικός, 1989.

Παράτημα

α. Κώδικας του αλγόριθμου DVS

```

unit FINAL32;

interface

uses
  Windows, Messages, SysUtils, Classes, Graphics, Controls, Forms, Dialogs,
  StdCtrls, Buttons;

type
  TForm1 = class(TForm)
    Button1: TButton;
    BitBtn1: TBitBtn;
  procedure CVR;
    procedure Button1Click(Sender: TObject);
  private
    { Private declarations }
  public
    { Public declarations }
  end;

var
  Form1: TForm1;

implementation

{$R *.DFM}
{Constants}
CONST
  maxKCapacity = 50; {max Kvalue Array Capacity}
  maxmvCapacity = 100; {max mv Array Capacity}

  maxKvalueCapacity = 5000; {set the kvalue max capacity to all
auxilliary arrays}
  maxDimCapacity = 101; {set the dim max capacity to all dim-reference
arrays}

```

```

maxXCapacity = 100000; {set the time series array max capacity}
maxEstimCapacity = 2000; {Error Capacity}

{Global variables}
VAR

CASES      :integer; {number of different cases}
X          :array [ 1 .. maxXCapacity ] of real; {time series}
neiberhoud :array [ 1 .. maxKvalueCapacity ] of integer; {neiberhoud
indexes}
imax       :integer; {maximum number of data in data.dat file}
Nf         :integer; {amount of past time serie elements}
dim        :integer; {dimension}
kvalue     :array [ 1 .. maxKCapacity ] of integer; {k values array}
kvalue_num :integer; {amount of kvalue array elements}
mv         :array [ 1 .. maxmvCapacity ] of integer; {moves out mv steps
from the first line}
t          :integer; {prediction time}
Sigma      :real; {SIGMA}

{ARRAY DEFINATIONS}
A          :array [ 1 .. maxKvalueCapacity , 1 .. maxDimCapacity ] of
real;
AT         :array [ 1 .. maxDimCapacity , 1 .. maxKvalueCapacity ] of
real;
ATA        :array [ 1 .. maxDimCapacity , 1 .. maxDimCapacity ] of
real;
ATA_1      :array [ 1 .. maxDimCapacity , 1 .. maxDimCapacity ] of
real;
ATA_1AT    :array [ 1 .. maxDimCapacity , 1 .. maxKvalueCapacity ] of
real;
FIT        :array [ 1 .. maxDimCapacity ] of real;
B          :array [ 1 .. maxKvalueCapacity ] of real;
Unic       :array [ 1 .. maxDimCapacity , 1 .. maxDimCapacity ] of
real; {auxilliary array which
is used in ATA^-1
calculation}
Error      :array [ 1 .. maxKCapacity ] of real; {Error Diagramm}
ESTIM      :array [ 1 .. maxEstimCapacity ] of real; {Estimation
Output}

{Read data from file and working on Normalization Procedure}

```

```

Procedure Normalization;
var
  file_:text;
  mean:real;
  i,help:integer;
begin
  {open data.dat file}
  assign(file_,'data.dat');
  reset(file_);
  {Read time series and calculate mean average}
  imax:=0;
  mean:=0;
  repeat
    imax:=imax+1;
    readln(file_,X[imax]);
    mean:=mean+X[imax];
  until eof(file_)=true;
  close(file_);

  mean:=mean/imax;
  {Calculate SIGMA}
  sigma:=0;
  for i:=1 to imax do
    sigma:=sigma+sqr(X[i]-mean);
  sigma:=sqrt(sigma/(imax-1));
  {Normalization}
  for i:=1 to imax do
    X[i]:=(X[i]-mean)/sigma;
end;

```

{ Procedure which calculates distances from any (X to dist) vector in test area
 and at the same time calculates the k first neiberhoud elements}

Procedure DISTANCES(dist,kroutes:integer);

```

var
  i,j,k,l:integer;
  sum:real;
  kroutescriterio:boolean;
  neiberhoudvalue:array[1..maxKvalueCapacity]of real;
begin

```

```

for k:=1 to kroutes do
  neiberhoudvalue[k]:=1e20;

  for i:=mv[dim-1]+1 to dist-t do
    begin

      {for dist to i distance calculation start from here and stop at}
      sum:=abs( X[dist] - X[i] );
      for j:=1 to dim-1 do
        sum:= sum + abs( X[dist-mv[j]] - X[i-mv[j]] );
      {dist to i distance calculation stop here}

      {loop inwhich a kroutes-neiberhoud points sorting procedure is executed,
      starting from here}
      k:=0;kroutescriterio:=false;
      repeat
        k:=k+1;
        if sum<neiberhoudvalue[k] then
          begin
            kroutescriterio:=true;
            for l:=kroutes downto k+1 do
              begin
                neiberhoud[l]:=neiberhoud[l-1];
                neiberhoudvalue[l]:=neiberhoudvalue[l-1];
              end;
            neiberhoudvalue[k]:=sum;
            neiberhoud[k]:=i;
          end;
        until (k=kroutes) or (kroutescriterio=true);
      {sorting procedure stops here}
      end;
    end;

  {Procedure which calculates the A matrix}
  Procedure CalculateA(kroutes_:integer);
    var i,j:integer;
    begin
      for i:=1 to kroutes_ do
        for j:=1 to dim+1 do
          if j=1 then A[i,j]:=1 else
            if j=2 then A[i,j]:=X[neiberhoud[i]] else
              A[i,j]:=X[neiberhoud[i]-mv[j-2]];
    end;
  end;

```

end;

{Procedure which calculates the B matrix}

Procedure CalculateB(kroutes_:integer);
 var i:integer;
 begin
 for i:=1 to kroutes_ do
 B[i]:=X[neiberhoud[i]+t];
 end;

{Procedure which calculates the AT matrix}

Procedure CalculateAT(kroutes_:integer);
 var i,j:integer;
 begin
 for i:=1 to dim+1 do
 for j:=1 to kroutes_ do
 AT[i,j]:=A[j,i];
 end;

{Procedure which calculates the AT*A matrix}

Procedure CalculateATA(kroutes_:integer);
 var i,j,k:integer;
 sum:real;
 begin
 for i:=1 to dim+1 do
 for j:=1 to dim+1 do
 begin
 sum:=0;
 for k:=1 to kroutes_ do
 sum:=sum+AT[i,k]*A[k,j];
 ATA[i,j]:=sum;
 end;
 end;

{The solution of a linear system - Τριγωνοποίηση (phase I)}

procedure GAUSS1;

var n,i,j,k,l:integer;
 help,mji:extended;
 begin
 n:=dim+1;

```

for i:=1 to n do
for j:=1 to n do
if i<>j then Unic[i,j]:=0 else Unic[i,j]:=1;
for i:=1 to n do
begin
l:=i-1;
k:=0;
repeat
l:=l+1;
if (ATA[l,i]<>0) then k:=l;
until (k<>0) or (l=n);
if (l=n) and (ATA[k,i]=0) then begin writeln('HALT');readln;HALT;end;
if k<>i then
begin
for l:=1 to n do
begin
help:=ATA[i,l];
ATA[i,l]:=ATA[k,l];
ATA[k,l]:=help;
help:=Unic[i,l];
Unic[i,l]:=Unic[k,l];
Unic[k,l]:=help;
end;
end;

for j:=i+1 to n do
begin
mji:=ATA[j,i]/ATA[i,i];
for l:=1 to n do
begin
ATA[j,l]:=ATA[j,l]-mji*ATA[i,l];
Unic[j,l]:=Unic[j,l]-mji*Unic[i,l];
end;
end;
end;

```

{The solution of a linear system Πίσω αντικατάσταση (phase II)}

procedure GAUSS;

var n,i,j,k,l:integer;

```

sum:extended;

begin
GAUSS1;
n:=dim+1;
for l:=1 to n do
begin
if ATA[n,n]=0 then begin writeln('HALT');readln;HALT;end;
ATA_1[n,l]:=Unic[n,l]/ATA[n,n];
i:=n;
repeat
i:=i-1;
sum:=0;
for k:=i+1 to n do
sum:=ATA[i,k]*ATA_1[k,l]+sum;
ATA_1[i,l]:=(Unic[i,l]-sum)/ATA[i,i];
until i=1;
end;
end;

```

{Procedure which calculates the $(AT^*A)^{(-1)}$ matrix}

```

Procedure CalculateATA_1;
begin
    GAUSS;
end;

```

{Procedure which calculates the $(AT^*A)^{(-1)} * AT$ matrix}

```

Procedure CalculateATA_1AT(kroutes_:integer);
var i,j,k:integer;
    sum:real;
begin
    for i:=1 to dim+1 do
        for j:=1 to kroutes_ do
            begin
                sum:=0;
                for k:=1 to dim+1 do
                    sum:=sum+ATA_1[i,k]*AT[k,j];
                ATA_1AT[i,j]:=sum;
            end;
    end;

```

{Procedure which calculates the $(AT^*A)^{(-1)} * AT^*B$ matrix (FINAL MATRIX)}

```
Procedure CalculateFIT(kroutes_:integer);
var i,k:integer;
    sum:real;
begin
    for i:=1 to dim+1 do
        begin
            sum:=0;
            for k:=1 to kroutes_ do
                sum:=sum+ATA_1AT[i,k]*B[k];
            FIT[i]:=sum;
        end;
end;
```

{Procedure which calculates the $a_0, a_1, a_2, \dots, a_m$ array}

```
Procedure ELAXISTATETRAGONA(kroutes:integer);
begin
    CalculateA(kroutes);
    CalculateAT(kroutes);
    CalculateATA(kroutes);
    CalculateATA_1;
    CalculateATA_1AT(kroutes);
    CalculateB(kroutes);
    CalculateFIT(kroutes);
end;
```

{FUNCTION which calculates the estimation $a_0 + a_1 * X_1 + \dots$ }

```
FUNCTION ESTIMATION(dist:integer):real;
var sum:real;
    i:integer;
begin
    sum:=FIT[1]+FIT[2]*X[dist];
    for i:=3 to dim+1 do
        sum:=sum+FIT[i]*X[dist-mv[i-2]];
    ESTIMATION:=sum;
end;
```

{Procedure which reads the input parameters}

```
Procedure Read_Case(CASES_:integer);
var file_.text;
    i,j:integer;
```

```
begin
assign(file_,'input.dat');
reset(file_);
if CASES=0 then
begin
  for i:=1 to 4 do
    readln(file_);
  readln(file_,CASES);
end else
begin
  for i:=1 to 6 do
    readln(file_);
  for i:=1 to CASES -1 do
    begin
      for j:=1 to 10 do
        readln(file_);
      readln(file_,dim);
      for j:=1 to 2 do
        readln(file_);
      for j:=1 to dim-1 do
        readln(file_);
      for j:=1 to 2 do
        readln(file_);
      readln(file_,kvalue_num);
      for j:=1 to 2 do
        readln(file_);
      for j:=1 to kvalue_num do
        readln(file_);
      readln(file_);
    end;
    for j:=1 to 4 do
      readln(file_);
    readln(file_,Nf);
  for j:=1 to 2 do
    readln(file_);
  readln(file_,t);
  for j:=1 to 2 do
    readln(file_);
  readln(file_,dim);
  for j:=1 to 2 do
    readln(file_);
  for j:=1 to dim-1 do
```

```

      readln(file_,mv[j]);
      for j:=1 to 2 do
        readln(file_);
      readln(file_,kvalue_num);
      for j:=1 to 2 do
        readln(file_);
      for j:=1 to kvalue_num do
        readln(file_,kvalue[j]);
      readln(file_);

    end;
    close(file_);
end;

```

{Procedure which stores the results for every case}

Procedure Store(Case_:integer);

```

var
  num:string;
  file_:text;
  i:integer;
  sum:real;
begin
  str(Case_2,num);
  if Case_<10 then num[1]:='0';
  assign(file_,'Case'+num+'.dat');
  rewrite(file_);
  for i:=1 to kvalue_num do
    writeln(file_,kvalue[i],'',Error[i]);
  close(file_);
end;

```

{Procedure which stores the results for every case}

Procedure Store_VALUES(Case_,KVALUEid:integer);

```

var
  num,str_:string;
  file_:text;
  i:integer;
begin
  str(Case_2,num);
  if Case_<10 then num[1]:='0';
  str(KVALUEid:2,str_);
  if KVALUEid<10 then str_[1]:='0';

```

```

assign(file_,'C_'+num+'_'+str_+'.exp');
rewrite(file_);
writeln(file_,'SIGMA = ',SIGMA);
for i:=1 to imax-Nf-t do
writeln(file_,X[Nf+i:t]:5:3,' ',estim[i]:5:3);
close(file_);
end;

{Main Program}
Procedure TFORM1.CVR;
  var i,j,n,ll:integer;
    sum:real;
    str_:string;
begin
  Read_Case(0); {Reads all dif cases}
  for n:=1 to Cases do
  begin
    with canvas do
    begin
      str(n:3,str_);
      Textout(100,50,'CASE = '+str_);
    end;
    Read_Case(n);
    Normalization;

    for i:=1 to kvalue_num do
    begin
      with canvas do
      begin
        str(i:3,str_);
        Textout(100,100,'I examine K number: '+str_);
      end;
      sum:=0; ll:=0;
      for j:=1 to imax-Nf-t do
      begin
        Distances(Nf+j,kvalue[i]);
        ELAXISTATETRAGONA(kvalue[i]);
        ll:=ll+1;
        estim[ll]:=ESTIMATION(Nf+j);
        sum:=sum+abs(X[Nf+j:t]-ESTIMATION(Nf+j));
      end;
    end;
  end;
end;

```

```
    end;
Error[i]:=sum/(imax-Nf-t);
Store_VALUES(n,i);
end;
Store(n);
end;
MessageDlg('Calculations have Finished ...',
mtConfirmation, [mbOk], 0);
end;

procedure TForm1.Button1Click(Sender: TObject);
begin
CVR;
end;

end.
```

β. Συνεργαζόμενα αρχεία

Το πρόγραμμα συνεργάζεται με τέσσερα αρχεία. Από δύο διαβάζει δεδομένα και σε δύο καταγράφει αποτελέσματα. Πρόκειται για τα :

data.dat
 input.dat
 case???.dat (π.χ. case01.dat, case02.dat κ.λ.π.)
 c_??_?.exp (π.χ. c_01_03.exp, c_02_01.exp κ.λ.π.)

i) Αρχείο data.dat

Είναι το αρχείο που περιέχει τα στοιχεία της χρονοσειράς. Από αυτό διαβάζονται από το πρόγραμμα και σε αυτό πρέπει να έχει εγγραφεί κάθε χρονοσειρά που θέλουμε να επεξεργαστούμε. Η μορφή του πρέπει να είναι : ένα στοιχείο ανά σειρά.

ii) Αρχείο input.dat

Είναι το αρχείο που περιέχει τις τιμές των μεταβλητών Nf, m, T, ti, k.

Η μεγαλύτερή του διαίρεση είναι σε cases (περιπτώσεις). Κάθε case είναι μία ομάδα παραμέτρων. Αν υπάρχουν περισσότερες της μίας, τότε το πρόγραμμα αφού επιλύσει την πρώτη συνεχίζει με τη δεύτερη κ.ο.κ. Το μέγιστο πλήθος cases που δέχεται το πρόγραμμα είναι 100.

Η κάθε case δέχεται μία ομάδα μεταβλητών. Η μεταβλητή k επιτρέπεται να έχει μία ως πενήντα τιμές. Για περισσότερες της μίας τιμές, το πρόγραμμα αφού επεξεργαστεί τα δεδομένα με την πρώτη συνεχίζει με τη δεύτερη κ.ο.κ.

Το αρχείο αυτό έχει πολύ συγκεκριμένο format και η παραμικρή διαφοροποίησή του οδηγεί σε σφάλμα. Για αυτό υπάρχει πρότυπο του αρχείο : sample.dat που καλείται σε περίπτωση που το input.dat ‘καταστραφεί’. Αυτά που πρέπει να προσέχονται είναι να δίνονται ακριβώς όσες τιμές ζητούνται και να μένει ακριβώς μια σειρά κενή μετά από κάθε ‘ομάδα’. Στο αρχείο αυτό δίνουμε τιμές μέσω ενός editor.

Ένα υπόδειγμα του input.dat δίδεται στη συνέχεια :

INPUT FILE**Parametric Input:****Number of Examination Cases:**

3

Case I:**Amount of past time series (Nf):**

7999

Prediction time (T):

2

Dimension (m):

5

T_i from T₁ to T_{m-1} (m-1 total values)

1

2

3

4

Amount of K_i different values:

3

K_i from K₁ to K_{final}:

20

30

40

Case II:**Amount of past time series (Nf):**

7800

Prediction time (T):

1

Dimension (m):

4

T_i from T₁ to T_{m-1} (m-1 total values)

1

2

3

Amount of K_i different values:

4

K_i from K₁ to K_{final}:

10

20

30

40

Case III:

Amount of past time series (N_f):

7800

Prediction time (T):

1

Dimension (m):

5

T_i from T₁ to T_{m-1} (m-1 total values)

1

2

3

4

Amount of K_i different values:

2

K_i from K₁ to K_{final}:

10

20

iii) Αρχείο case???.dat

Στο αρχείο αυτό αποθηκεύονται τα σφάλματα E(k) και τα k στα οποία αντιστοιχούν. Δημιουργείται ένα αρχείο για κάθε case. Έτσι αν οι cases που περιέχονται στο input.dat είναι 10, δημιουργούνται τα αρχεία : case01.dat, case02.dat, ... ,case10.dat.

iv) Αρχείο c_??_???.exp

Στην πρώτη σειρά αυτού του αρχείου αποθηκεύεται η τυπική απόκλιση των δείγματος. Στη συνέχεια αποθηκεύονται σε δύο στήλες οι πραγματικές τιμές και οι προβλέψεις που είχαμε για αυτές. Δημιουργείται από ένα τέτοιο αρχείο για κάθε k από κάθε case. Π.χ. για το δεύτερο k της τρίτης case δημιουργείται το c_03_02.exp.