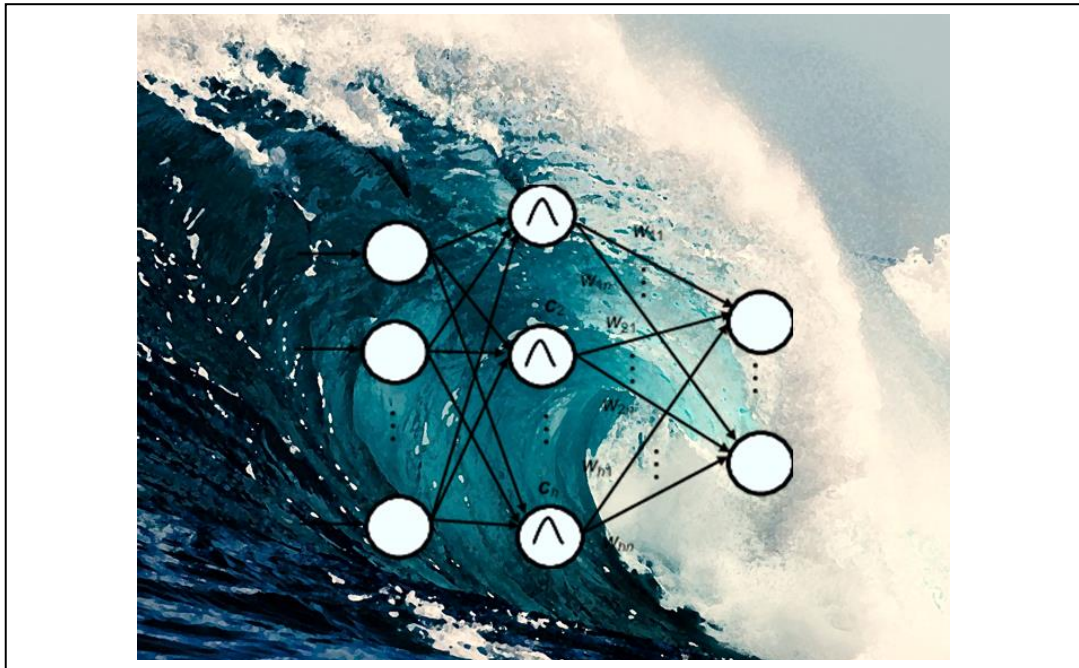


Διπλωματική Εργασία

Μοντελοποίηση δεδομένων θαλάσσιου κυματισμού με νευρωνικά δίκτυα



Φοιτητής: Αποστολάκη Ελένη
ΑΜ: 50107042

Επιβλέπων Καθηγητής

Αλέξανδρος Αλεξανδρίδης
Καθηγητής

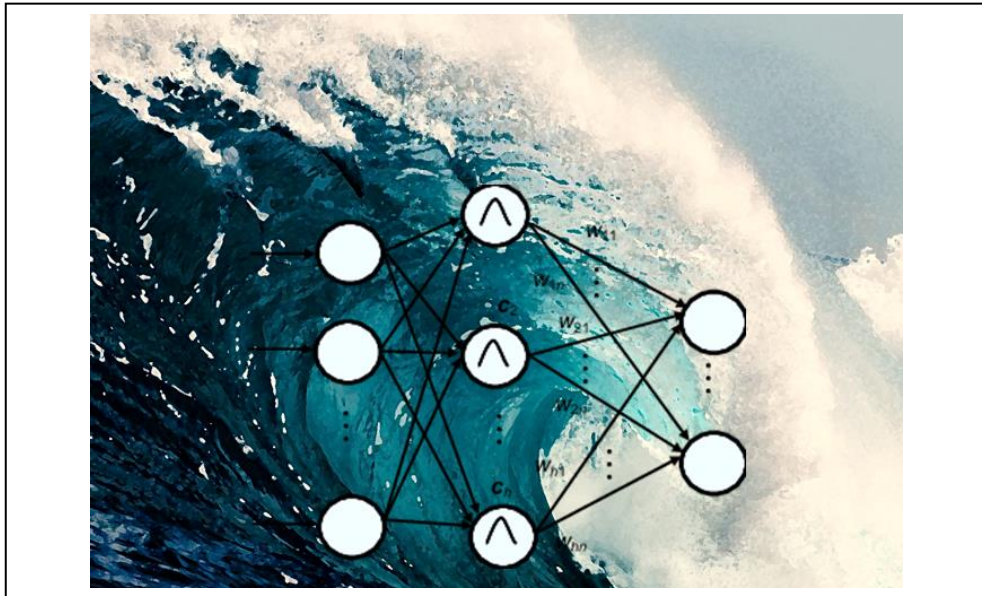
ΑΘΗΝΑ-ΑΙΓΑΛΕΩ, ΟΚΤΩΒΡΙΟΣ 2022



UNIVERSITY OF WEST ATTICA
FACULTY OF ENGINEERING
DEPARTMENT OF ELECTRICAL & ELECTRONICS
ENGINEERING

Diploma Thesis

Sea wave data modelling using neural networks



Student: Apostolaki Eleni
Registration Number: 50107042

Supervisor

Alex Alexandridis
Professor

ATHENS-EGALEO, OCTOBER 2022

Η Διπλωματική Εργασία έγινε αποδεκτή και βαθμολογήθηκε από την εξής τριμελή επιτροπή:

Αλεξανδρίδης Αλέξανδρος, Καθηγητής	Φαμέλης Ιωάννης, Καθηγητής	Ζώης Ηλίας, Αναπληρωτής Καθηγητής
(Υπογραφή)	(Υπογραφή)	(Υπογραφή)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ και Ελένη Αποστολάκη, Οκτώβριος, 2022

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν την συγγραφέα του και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις θέσεις του επιβλέποντος, της επιτροπής εξέτασης ή τις επίσημες θέσεις του Τμήματος και του Ιδρύματος.

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η κάτωθι υπογεγραμμένη Αποστολάκη Ελένη του Νικολάου, με αριθμό μητρώου 50107042 φοιτήτρια του Πανεπιστημίου Δυτικής Αττικής της Σχολής ΜΗΧΑΝΙΚΩΝ του Τμήματος ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ,

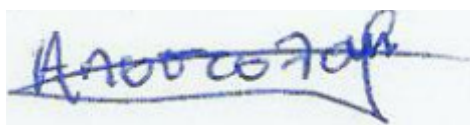
δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου.

Επιθυμώ την απαγόρευση πρόσβασης στο πλήρες κείμενο της εργασίας μου, παρά μόνο έπειτα από αίτησή μου στη Βιβλιοθήκη και έγκριση του επιβλέποντος καθηγητή.»

Η Δηλούσα
Αποστολάκη Ελένη



Ευχαριστίες

Με την παρούσα διπλωματική εργασία ολοκληρώνονται οι σπουδές μου στο προπτυχιακό πρόγραμμα σπουδών του Τμήματος Ηλεκτρολόγων & Ηλεκτρονικών του Πανεπιστημίου Δυτικής Αττικής.

Καθ' όλη την διάρκεια των σπουδών μου καθοριστικό ρόλο έπαιξε η συμβολή των καθηγητών μου στα γνωστικά αντικείμενα που παρακολούθησα και για αυτό οφείλω να τους ευχαριστήσω για τη συμβολή τους στην ολοκλήρωση των σπουδών μου.

Ιδιαίτερα θα ήθελα να ευχαριστήσω θερμά τον εισηγητή και επιβλέποντα κ. Αλέξανδρο Αλεξανδρίδη για την εμπιστοσύνη και την καθοδήγηση που μου έδωσε στην εκπόνηση της παρούσας διπλωματικής εργασίας. Επίσης, θα ήθελα να τον ευχαριστήσω για όλες τις συμβουλές και τις υποδείξεις του, καθώς και για τις γνώσεις που αποκόμισα από εκείνον καθ' όλη την διάρκεια των φοιτητικών μου χρόνων.

Επιπλέον, θα ήθελα να ευχαριστήσω την υποψήφιο Διδάκτορα, Ιωάννη Κορδάτο, για την υπομονή του και την πολύτιμη καθοδήγηση που μου παρείχε καθ' όλη τη διάρκεια της προσπάθειας αυτής μέχρι την ολοκλήρωσή της.

Ιδιαίτερες ευχαριστίες θα ήθελα να δώσω και στους Ιωάννη Φαμέλη και Ηλία Ζώη, τόσο για τον χρόνο που μου παρέθεσαν στην αξιολόγηση της διπλωματικής εργασίας, όσο και για τις γνώσεις που μου προσέφεραν.

Τέλος, οφείλω να ευχαριστήσω την οικογένειά μου και τους κοντινούς μου ανθρώπους, για τη συμπαράσταση και την υπομονή τους.

Περίληψη

Τα νευρωνικά δίκτυα είναι μαθηματικά μοντέλα που χρησιμοποιούν αλγόριθμους μάθησης εμπνευσμένους από τον ανθρώπινο εγκέφαλο για την αποθήκευση πληροφοριών. Σήμερα, ο όρος μηχανική μάθηση χρησιμοποιείται συχνά και αναφέρεται στον επιστημονικό κλάδο που ασχολείται με το σχεδιασμό και την ανάπτυξη αλγορίθμων, που επιτρέπουν στους υπολογιστές να μαθαίνουν, από δεδομένα που προέρχονται από αισθητήρες ή βάσεις δεδομένων.

Πως όμως συνδέονται τα νευρωνικά δίκτυα με το ύψος του κύματος της θάλασσας;

Η αξιόπιστη εκτίμηση του ύψους κύματος της θάλασσας μπορεί να παρέχει χρήσιμες πληροφορίες για εφαρμογές, όπως η παρακολούθηση του περιβάλλοντος και η προστασία και ανάπτυξη των λιμανιών και των παράκτιων μεταφορών. Στην παρούσα διπλωματική εργασία, αρχικά γίνεται ανασκόπηση των συμβατικών μεθόδων πρόβλεψης του ύψους κύματος, που βασίζονται σε αριθμητικά, στατιστικά και στοχαστικά μοντέλα. Ειδικότερα, μια υπάρχουσα μέθοδος πρόβλεψης είναι αυτή κατά την οποία τα μοντέλα του κυματικού φάσματος επιλύονται σε ορθογώνια δομημένα πλέγματα. Κάθε πλέγμα αποτελείται από έναν αριθμό κόμβων τα οποία ονομάζονται σημεία του πλέγματος. Ανάλογα με το αν η απόσταση των σημείων αυτών είναι μεγάλη ή μικρή, το πλέγμα θεωρείται αραιό ή πυκνό αντίστοιχα.

Στόχος της εργασίας είναι η ανάπτυξη ενός μοντέλου βασισμένο σε νευρωνικά δίκτυα, το οποίο θα δέχεται ως εισόδους τις πληροφορίες του αραιού πλέγματος και προβλέπει τι συμβαίνει στους κόμβους του πυκνού πλέγματος. Για την επίτευξή αυτού του στόχου χρησιμοποιήθηκαν τα νευρωνικά δίκτυα ακτινικής βάσης, ενώ για τον υπολογισμό των κέντρων του δικτύου χρησιμοποιήθηκε ο αλγόριθμος των ασαφών μέσων.

Κατόπιν, λοιπόν, της εκπαίδευσης του νευρωνικού δικτύου, τα αποτελέσματα των σεναρίων και η σύγκριση αυτών έδειξαν ότι υπάρχει η δυνατότητα εξαγωγής μιας πολύ καλής προσέγγισης για το μέσο και το μέγιστο ύψος κύματος του πυκνού πλέγματος τρέχοντας μόνο το αραιό, με αποτέλεσμα την σημαντική μείωση του υπολογιστικού χρόνου. Συγκεκριμένα, αν και ο χρόνος που απαιτείται για την εκπαίδευση του μοντέλου ποικίλει και μπορεί να είναι μεγάλος ανάλογα με τις εισόδους που λαμβάνει το νευρωνικό δίκτυο, ο χρόνος για την πρόβλεψη είναι μηδαμινός σε σχέση με τις συμβατικές μεθόδους πρόβλεψης.

Λέξεις – κλειδιά

Αλγόριθμος Ασαφών Μέσων, Νευρωνικά δίκτυα, Πρόβλεψη Μέσου Ύψους Κύματος, Πρόβλεψη Μέγιστου Ύψους Κύματος, Συναρτήσεις Ακτινικής Βάσης

Abstract

Neural networks are mathematical models that retain information using learning techniques derived from the human brain. Today, the phrase "machine learning" is frequently used to describe the scientific field concerned with the creation of algorithms that enable computers to learn based on information obtained from sensors or databases.

But how are neural networks connected to the height of the sea wave?

Applications like environmental monitoring, the development and protection of ports, and coastal transportation can all benefit from accurate estimation of sea wave height. This Thesis begins by reviewing the traditional wave height prediction techniques, which are based on numerical, statistical, and ensemble models. One existing prediction technique in particular resolves wave spectrum models on grids with orthogonal structures. A number of nodes called grid points make up each grid. The grid is categorized as sparse or dense depending on whether the distance between these points is large or small.

The aim of the Thesis is to create a neural network model that will take the data from the sparse grid as inputs and predict what will happen at the nodes of the dense grid. Radial basis function neural networks were used to accomplish this task, and the network's centers were determined using the Fuzzy Means algorithm.

After the training of the neural network, the results of the scenarios and their comparison revealed that it is possible to obtain a very good approximation for the significant and maximum wave height of the dense grid by running only the sparse one, resulting to a reduction in computing time. Although the time required to train the model varies and can be lengthy depending on the inputs received by the neural network, the time required for prediction is minimal when compared to conventional prediction methods.

Keywords

Fuzzy Means Algorithm, Maximum Wave Height Prediction, Neural Networks, Radial Basis Functions, Significant Wave Height Prediction

Κατάλογος Πινάκων	9
Κατάλογος Εικόνων.....	9
Αλφαβητικό Ευρετήριο	11
ΕΙΣΑΓΩΓΗ.....	12
Αντικείμενο της διπλωματικής εργασίας.....	12
Σημασία της διπλωματικής εργασίας	12
Δυσκολίες που παρουσιάζονται	12
Παραφερή έρευνα.....	13
Μεθοδολογία.....	14
Δομή	14
ΚΕΦΑΛΑΙΟ 1^ο :Μετεωρολογικές προβλέψεις-Πρόβλεψη κύματος	16
1.1 Μέθοδοι πρόγνωσης.....	16
1.1.1 Αριθμητική πρόγνωση.....	18
1.1.2 Στατιστική πρόγνωση.....	19
1.1.3 Συνοπτική πρόγνωση.....	20
1.1.4 Στοχαστική πρόγνωση.....	21
1.2 Πρόγνωση κύματος.....	21
1.2.1 Βασικά χαρακτηριστικά κύματος	21
1.2.2 Συμβατικά μοντέλα πρόβλεψης κύματος	22
1.2.3 Πρόβλεψη κύματος και χρήση νευρωνικών δικτύων	25
ΚΕΦΑΛΑΙΟ 2^ο : Νευρωνικά Δίκτυα Ακτινικής Βάσης - Αλγόριθμος Fuzzy Means.....	28
2.1 Perceptrons Πολλών Στοιβάδων	29
2.2 Νευρωνικά Δίκτυα Ακτινικής Βάσης.....	30
2.2.1 Στοιβάδα εισόδου.....	30
2.2.2 Κρυφή στοιβάδα	31
2.2.3 Στοιβάδα εξόδου	31
2.3 Συναρτήσεις ακτινικής βάσης	32
2.3.1 Παραδείγματα συναρτήσεων ακτινικής βάσης.....	33
2.4 Λειτουργία δικτύων RBF.....	34
2.5 Εκπαίδευση RBF δικτύου.....	35
2.5.1 Κανονικοποίηση δεδομένων.....	36
2.5.2 Συντεταγμένες κέντρων.....	36
2.5.3 Αλγόριθμος Fuzzy Means	39
2.5.4 Πλεονεκτήματα της χρήσης του αλγόριθμου Fuzzy Means.....	42
ΚΕΦΑΛΑΙΟ 3^ο : Πειραματικό μέρος	43
3.1 Περιγραφή δεδομένων	43
3.1.1 Dataset δεδομένων	44
3.2 Μεθοδολογία	48
3.2.1 RBF και Fuzzy Means.....	48
3.2.2 Περιγραφή Σεναρίων.....	49
3.2.3 Σενάρια με την χρήση όλων των features.....	49
3.2.4 Σενάρια με feature engineering στην ταχύτητα του ανέμου.....	51
3.2.5 Σενάρια με αφαίρεση χαρακτηριστικών.....	55
3.2.6 Σενάρια με χρήση της βαθυμετρία των διπλανών σημείων	60
3.3 Σύγκριση αποτελεσμάτων.....	63
ΚΕΦΑΛΑΙΟ 4^ο : Συμπεράσματα	66
Βιβλιογραφία – Αναφορές - Διαδικτυακές Πηγές.....	68

Κατάλογος Πινάκων

Πίνακας 3.1: Dataset Δεδομένων.

Πίνακας 3.2: Γεωγραφικό μήκος και πλάτος των nodes.

Πίνακας 3.3: Μέγιστες και ελάχιστες τιμές χαρακτηριστικών κύματος.

Πίνακας 3.4: Σενάρια χωρίς Feature Engineering.

Πίνακας 3.5: Σενάρια με Feature Engineering στην ταχύτητα του ανέμου-1.

Πίνακας 3.6: Σενάρια με Feature Engineering στην ταχύτητα του ανέμου-2.

Πίνακας 3.7: Σενάρια με αφαίρεση feature-1.

Πίνακας 3.8: Σενάρια με αφαίρεση feature-2.

Πίνακας 3.9: Σενάρια με αφαίρεση feature-3.

Πίνακας 3.10: Σενάρια με αφαίρεση feature-4.

Πίνακας 3.11: Σενάρια με χρήση της βαθυμετρία των διπλανών σημείων -1.

Πίνακας 3.12: Σενάρια με χρήση της βαθυμετρία των διπλανών σημείων-2.

Κατάλογος Εικόνων

Εικόνα 1.1: Synoptic weather map (The European Space Agency).

Εικόνα 1.2: Βασικά χαρακτηριστικά κύματος.

Εικόνα 1.3: Nested Grid Ocean Model (URI Graduate school of Oceanography Numerical Modeling Lab).

Εικόνα 2.1: Δομή MLP δικτύου με δύο κρυφές στοιβάδες.

Εικόνα 2.2: Δομή RBF δικτύου με N μεταβλητές εισόδου και L κρυφούς κόμβους.

Εικόνα 2.3: Παράδειγμα χρήσης της μεθόδου των ελαχίστων τετραγώνων.

Εικόνα 2.4: Συνάρτηση Gauss.

Εικόνα 2.5: Συνάρτηση TPS.

Εικόνα 2.6: Πολυτετραγωνική συνάρτηση.

Εικόνα 2.7: Πρόβλεψη του δικτύου RBF.

Εικόνα 2.8: Ασαφές σύνολο.

Εικόνα 2.9: Ασαφής διαχωρισμός του χώρου μιας μεταβλητής.

Εικόνα 2.10: Διαχωρισμός δισδιάστατου χώρου εισόδου με ασαφείς υποχώρους.

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Εικόνα 2.11: Βαθμός συμμετοχής του $u = [u_1 \ u_2]$.

Εικόνα 2.12: Επιλογή του κοντινότερου ασαφούς υποχώρου.

Εικόνα 3.1: Αναπαράσταση πυκνού και αραιού πλέγματος.

Εικόνα 3.2: Training, validation, testing dataset.

Εικόνα 3.3: Πρόβλεψη SWHT με όλα τα features.

Εικόνα 3.4: Πρόβλεψη MWH με όλα τα features.

Εικόνα 3.5: Ανάλυση σε WINX και WINY.

Εικόνα 3.6: Πρόβλεψη SWHT με feature engineering.

Εικόνα 3.7: Πρόβλεψη MWH με feature engineering.

Εικόνα 3.8: Πρόβλεψη SWHT με feature engineering-2.

Εικόνα 3.9: Πρόβλεψη MWH με feature engineering-2.

Εικόνα 3.10: Πρόβλεψη SWHT με SWHT_Sparse.

Εικόνα 3.11: Πρόβλεψη MWH με feature MWH_Sparse.

Εικόνα 3.12: Πρόβλεψη SWHT με feature SWHT_Sparse, WINX, WINY.

Εικόνα 3.13: Πρόβλεψη SWHT με feature SWHT_Sparse, DP.

Εικόνα 3.14: Πρόβλεψη SWHT με feature SWHT_Sparse, DP, MEANFR, WINX, WINY.

Εικόνα 3.15: Πρόβλεψη MWH με feature MWH_Sparse, DP, MEANFR, WINX, WINY.

Εικόνα 3.16: Πρόβλεψη SWHT με feature SWHT_Sparse και DP_adj.

Εικόνα 3.17: Πρόβλεψη SWHT με feature SWHT_Sparse DP, DP_adj και MEANFR.

Εικόνα 3.18: Πρόβλεψη SWHT με feature SWHT_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY.

Εικόνα 3.19: Πρόβλεψη SWHT με feature SWHT_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY.

Αλφαβητικό Ευρετήριο

ANN: Artificial Neural Networks

DP: Model Bathymetry

FM: Fuzzy Means

IEEE: The Institute for Electrical and Electronics Engineers

LMS: Least Mean Square

MAE: Mean Absolute Error

MEANFR: Mean Frequency

MLP: Multilayer Perceptron

MWDIR: Mean Wave Direction

MWH: Maximum Wave Height

NN: Neural Network

NWP: Numerical weather prediction

PKFR: Peak Frequency

RBF: Radial Basis Function

RMSE: Root Mean Square Error

R2: Correlation coefficient

SVM: Support Vector Machine

SWHT: Significant Wave Height

WINDDIR: Wind Direction

WINSF: Wind Speed

MET: Μέθοδος Ελαχίστων Τετραγώνων

ΕΙΣΑΓΩΓΗ

Στο εισαγωγικό αυτό κεφάλαιο αναλύονται το αντικείμενο και η σημασία της διπλωματικής εργασίας, οι δυσκολίες που παρουσιάζονται, η παρεμφερή έρευνα και η μεθοδολογία η οποία χρησιμοποιήθηκε.

Αντικείμενο της διπλωματικής εργασίας

Σκοπός της διπλωματικής εργασίας είναι η μοντελοποίηση μετεωρολογικών δεδομένων με την χρήση νευρωνικών δικτύων. Συγκεκριμένα, η δημιουργία ενός μοντέλου το οποίο θα επιφέρει αποτελέσματα με μεγάλη ακρίβεια και μικρό υπολογιστικό χρόνο, αξιοποιώντας τα ήδη υπάρχοντα δεδομένα τα οποία προέρχονται από αναλυτικά μοντέλα που επιλύονται σε ορθογώνια δομημένα πλέγματα. Ειδικότερα, κάθε πλέγμα αποτελείται από έναν αριθμό κόμβων τα οποία ονομάζονται σημεία του πλέγματος και ανάλογα με την απόσταση των σημείων αυτών το πλέγμα θεωρείται αραιό ή πυκνό.

Το μοντέλο μηχανικής μάθησης δέχεται ως εισόδους τα δεδομένα του αραιού πλέγματος και προβλέπει το ύψος κύματος στους κόμβους του πυκνού πλέγματος. Τα δεδομένα θαλάσσιου κυματισμού αφορούν την γεωγραφική περιοχή της Κρήτης και πιο συγκεκριμένα τον Άγιο Νικόλαο και σχετίζονται με χαρακτηριστικά όπως η βαθυμετρία, η ταχύτητα του ανέμου και η κατεύθυνση του, καθώς και η μέση συχνότητα.

Σημασία της διπλωματικής εργασίας

Τα κύματα των ωκεανών είναι πολύ σημαντικά για την πρόγνωση του καιρού και την μοντελοποίηση του κλίματος, καθώς και για τις παράκτιες κοινότητες, τις ναυτιλιακές διαδρομές και τη βιομηχανία ανοικτής θαλάσσης.

Η πρόβλεψη κύματος βοηθά τους επιστήμονες και τους ερευνητές να περιορίσουν και να μετριάσουν περιβαλλοντικές καταστροφές, όπως οι πετρελαιοκηλίδες. Ωστόσο, δεν βοηθά μόνο τους ερευνητές να προβλέψουν μια περιβαλλοντική καταστροφή, αλλά μπορεί επίσης να συμβάλει στην αποφυγή εργασίας σε μέρη όπου η κίνηση και η ενέργεια των κυμάτων καθιστούν τη γεώτρηση πολύ επικίνδυνη. Χαρακτηριστικό παραδειγμα της εφαρμογής της πρόγνωσης κύματος στον χώρο εργασίας αποτελούν οι ναυτιλιακές εταιρίες, όπου γίνεται τακτική χρήση της πρόβλεψης κύματος στις επιχειρηματικές τους δραστηριότητες με σκοπό την αποφυγή της καταστροφής φορτίων και πλοίων.

Εκτός από τον κλάδο της ναυτιλίας, λόγω του ότι η πρακτική της υδατοκαλλιέργειας ή η εκτροφή ψαριών, οστρακοειδών και φυκιών, αρχίζει να γίνεται πιο δημοφιλής σε υπεράκτια θαλάσσια περιβάλλοντα, απαιτούνται δεδομένα κύματος για την προστασία των αγροκτημάτων και τη μεγιστοποίηση της βιωσιμότητάς τους. Βέβαια, οι ιχθυοκαλλιέργειες δεν είναι η μόνη περιοχή στην οποία η πρόβλεψη κύματος μπορεί να είναι αναγκαία. Η έρευνα για την κυματική ενέργεια ως εναλλακτική ανανεώσιμη πηγή ενέργειας είναι ένα αναπτυσσόμενο πεδίο που μπορεί να επωφεληθεί από τις ακριβείς προβλέψεις κύματος.

Συμπεραίνοντας, το ύψος κύματος θεωρείται μια παράμετρος ζωτικής σημασίας για το σχεδιασμό και την κατασκευή δομών προστασίας ακτών, τη μεταφορά ιζημάτων και τον εντοπισμό και την ανάπτυξη λιμανιών, ενώ επίσης σημαντικό ρόλο παίζει στην παράκτια μηχανική, την παρακολούθηση του περιβάλλοντος και τις θαλάσσιες μεταφορές.

Δυσκολίες που παρουσιάζονται

Τα προβλήματα που παρουσιάζουν τα συμβατικά μοντέλα πρόβλεψης του ύψους κύματος σχετίζονται τόσο με την μειωμένη ακρίβεια των προβλέψεων όσο και με τους αυξημένους

υπολογιστικούς χρόνους εκτέλεσής τους. Δεδομένου του ότι η πρόβλεψη των κυμάτων είναι μια αβέβαιη και τυχαία διαδικασία, αυτό την καθιστά ιδανική για μοντελοποίηση με την χρήση νευρωνικών δικτύων. Πάρα τα πλεονεκτήματα της χρήσης των νευρωνικών δικτύων (γρήγορη ταχύτητα υπολογισμού, χαμηλό υπολογιστικό κόστος και ισχυρή μη γραμμική ικανότητα εκμάθησης) έχουν παρουσιαστεί και δυσκολίες που είναι αναγκαίο να αντιμετωπιστούν.

Αρχικά, μια από τις βασικότερες δυσκολίες συνδέεται με την επιλογή, τόσο της τοπολογίας του νευρωνικού δικτύου όσο και της μεθόδου εκπαίδευσης του. Επιπλέον, ένα κρίσιμο ζήτημα σχετίζεται με τον σχεδιασμό ενός μοντέλου που θα παράγει ακριβείς προβλέψεις, πράγμα το οποίο εξαρτάται από την επιλογή των παραμέτρων του δικτύου. Εκτός από τα παραπάνω, κατά την εκπαίδευση των νευρωνικών δικτύων παρουσιάζεται το φαινόμενο του *overfitting*. Λόγω αυτού, λοιπόν, πρέπει να δοθεί ιδιαίτερη έμφαση στον τρόπο με τον οποίο τα δεδομένα εισόδου θα χωριστούν και θα κανονικοποιηθούν πριν δοθούν σαν είσοδοι για την εκπαίδευση του νευρωνικού δικτύου.

Παρεμφερή έρευνα

Ένα παρεμφερή πεδίο στην πρόγνωση των χαρακτηριστικών του κύματος είναι η πρόβλεψη του καιρού. Από εκεί, λοιπόν, μπορούν να αντληθούν πληροφορίες για αντίστοιχες δυσκολίες που παρουσιάστηκαν σε επιστημονικές έρευνες. Ενώ υπάρχει σχεδόν καθολική συμφωνία στη βιβλιογραφία ότι ο καιρός είναι ένα δυναμικό και μη γραμμικό φαινόμενο, φαίνεται να υπάρχουν δύο σχολές σκέψης για την ταξινόμηση της φύσης μιας πρόβλεψης που γίνεται από ένα δοσμένο σύνολο δεδομένων [1].

Από τη μία πλευρά, υπάρχει η πεποίθηση από τους ερευνητές ότι εάν το σύνολο δεδομένων είναι επαρκώς μεγάλο, συνήθως εκτείνεται σε διάστημα πολλών χρόνων, τότε οι προβλέψεις που βασίζονται στο dataset είναι ντετερμινιστικές, πράγμα που σημαίνει ότι κάθε παράγοντας που απαιτείται για τον προσδιορισμό της επόμενης κατάστασης της μεταβλητής πρόβλεψης υπάρχει στο dataset και ότι μπορεί να προβλεφθεί μια διακριτή τιμή. Από την άλλη πλευρά, επικρατεί η άποψη ότι με ένα μικρότερο dataset και μερικούς εξωγενείς προγνωστικούς παράγοντες, μπορεί να πραγματοποιηθεί μια στοχαστική πρόβλεψη [2].

Υπάρχουν πειράματα στη βιβλιογραφία που ενσωματώνουν διαφορετικές τοπολογίες νευρωνικών δικτύων. Σε μια περίπτωση, τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης χρησιμοποιήθηκαν για την πρόβλεψη της ποσότητας των βροχοπτώσεων στο εύρος 1 έως 3 ωρών στην Μπανγκόκ της Ταϊλάνδης, κάτι το οποίο ισοδυναμεί με την ντετερμινιστική προσέγγιση μιας συνάρτησης παλινδρόμησης [3]. Σε μια άλλη περίπτωση, οι ερευνητές χρησιμοποίησαν ένα νευρωνικό δίκτυο ακτινικής βάσης για την πρόβλεψη των βροχερών και μη βροχερών ημερών, που ισοδυναμεί με τη χρήση του νευρωνικού δικτύου για την ντετερμινιστική προσέγγιση μιας συνάρτησης ταξινόμησης [4].

Τέλος, τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για τον υπολογισμό των πιθανοτήτων των καιρικών φαινομένων σε στοχαστικές συνθήκες. Για παράδειγμα, ερευνητές στην Αργεντινή χρησιμοποίησαν μια σειρά νευρωνικών δικτύων για να προσεγγίσουν τις αθροιστικές συναρτήσεις κατανομής για την εμφάνιση υγρών και ξηρών διαστημάτων, κάτι που ισοδυναμεί με τη χρήση ενός νευρωνικού δικτύου για τον στοχαστικό υπολογισμό των συναρτήσεων κατανομής πιθανότητας [5].

Από τις παραπάνω τοπολογίες των νευρωνικών δικτύων ιδιαίτερο ενδιαφέρον παρουσίασε αυτή των νευρωνικών δικτύων ακτινικής βάσης. Κατά την εκπαίδευση των δικτύων RBF έχουν χρησιμοποιηθεί διάφορες μέθοδοι για την εύρεση των κέντρων του δίκτυο, όπως αυτή του K-means και αυτή του Fuzzy Means. Σύμφωνα με έρευνες [6], κατά τη διάρκεια της περιόδου πλημμύρας στην πόλη Shuangyashan, παρατηρήθηκε ότι η ακρίβεια του δικτύου

RBF για την πρόβλεψη των βροχοπτώσεων βελτιώθηκε σημαντικά με την μέθοδο K-means, σε σύγκριση με το μοντέλο δικτύου MLP. Το μέσο σχετικό σφάλμα των μηνιαίων βροχοπτώσεων το 2006, το 2007 και το 2008 μειώθηκε κατά 24%. Έτσι, το μοντέλο που κατασκευάστηκε θα μπορούσε να χρησιμοποιηθεί ως μια αποτελεσματική μέθοδος για την πρόβλεψη βροχοπτώσεων ή φαινομένων που παρουσιάζουν αβεβαιότητα.

Τέλος, σε άλλους επιστημονικούς κλάδους στους οποίους χρησιμοποιήθηκε ο αλγόριθμος Fuzzy Means παρατηρήθηκαν σημαντικές βελτιώσεις στην ακρίβεια του παραγόμενου μοντέλου [7]. Επίσης, μειώθηκε σημαντικά ο υπολογιστικός χρόνος κατά την εκπαίδευση του δικτύου [8].

Μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε στην παρούσα διπλωματική εργασία χωρίζεται σε πέντε βήματα.

Το πρώτο βήμα περιλαμβάνει την μελέτη, τόσο μέσω του διαδικτύου όσο και μέσω της αντίστοιχης βιβλιογραφίας, του θεωρητικού υποβάθρου των νευρωνικών δικτύων. Κατά το στάδιο αυτό συγκεντρώθηκαν στοιχεία που αφορούσαν τα νευρωνικά δίκτυα και επιλέχθηκε η τοπολογία του δικτύου RBF. Ιδιαίτερη σημασία δόθηκε επίσης και στην θεωρία των ασαφών συνόλων, καθώς για τον υπολογισμό των κέντρων του δικτύου RBF χρησιμοποιήθηκε ο αλγόριθμος Fuzzy Means.

Κατά το δεύτερο βήμα και αφού κατανοήθηκε η λειτουργία των δικτύων ακτινικής βάσης και του αλγορίθμου των ασαφών μέσων, έγινε η συγγραφή του κώδικα. Ο προγραμματισμός πραγματοποιήθηκε μέσω του προγράμματος MATLAB. Για να τον έλεγχο της σωστής λειτουργίας του μοντέλου μηχανικής μάθησης χρησιμοποιήθηκαν μικρότερα σετ δεδομένων, για τα οποία υπήρχαν ήδη διαθέσιμα αποτελέσματα. Κατά το βήμα αυτό, πραγματοποιήθηκαν βελτιστοποιήσεις και αλλαγές στον κώδικα για την καλύτερη λειτουργία του. Τέτοιες αλλαγές ήταν η τροποποίηση των επαναληπτικών δομών, η χρήση διαφορετικών εντολών για την αντιστροφή πινάκων και η χρήση των πινάκων για την εκτέλεση πράξεων. Εφόσον διαπιστώθηκε η ορθή λειτουργία του μοντέλου, ξεκίνησε η διαδικασία μελέτης άλλων μεθόδων για την πρόβλεψη του ύψους κύματος με ή χωρίς την χρήση νευρωνικών δικτύων.

Στο τρίτο βήμα μελετήθηκε το dataset με τα δεδομένα που προέρχονται από τον Άγιο Νικόλαο της Κρήτης, καθώς και το τρέξιμο των σεναρίων και η συλλογή των αποτελεσμάτων. Για να αποφευχθεί το φαινόμενο του overfitting χρησιμοποιήθηκε περισσότερος όγκος δεδομένων στο training dataset. Με βάση τους στατιστικούς δείκτες, τον χρόνο ολοκλήρωσης της εκπαίδευσης και τον αριθμό των κέντρων, δημιουργήθηκε ένας πίνακας ο οποίος περιείχε όλα τα αποτελέσματα από τα σενάρια που υλοποιήθηκαν. Επιπλέον, κατά το βήμα αυτό, σχεδιάστηκαν διαγράμματα που στον ένα τους άξονα υπήρχαν οι τιμές των προβλέψεων και στον άλλον τους άξονα οι πραγματικές τιμές. Το καλύτερο μοντέλο ήταν αυτό που τα δεδομένα σύγκλιναν στην ευθεία με κλίση 45 μοιρών.

Στο τέταρτο βήμα, μελετήθηκε κατά πόσο η ανάλυση της κατεύθυνσης του ανέμου σε δυο κάθετες συνιστώσες X και Y θα βελτιώνει τα αποτελέσματα του μοντέλου, τόσο σε σχέση με τον υπολογιστικό χρόνο όσο και με τα στατιστικά σφάλματα.

Στο πέμπτο και τελευταίο βήμα συγκεντρώθηκαν όλα τα αποτελέσματα τα οποία προέκυψαν από όλα τα σενάρια και έγινε η σύγκριση μεταξύ τους.

Δομή

Στο πρώτο κεφάλαιο της διπλωματικής εργασίας περιγράφεται γενικά ο τρόπος με τον οποίο γίνονται οι μετεωρολογικές προβλέψεις και δίνεται ιδιαίτερη έμφαση στην πρόβλεψη του ΠΑΔΑ, Τμήμα Η&ΗΜ, Διπλωματική Εργασία, Αποστολάκη Ελένη

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

ύψους κύματος, χωρίς την χρήση νευρωνικών δικτύων. Γίνεται, δηλαδή, μια ιστορική αναδρομή στην πρόβλεψη του ύψους κύματος. Στην συνέχεια παρουσιάζεται ο τρόπος με τον οποίο τα νευρωνικά δίκτυα χρησιμοποιούνται στην πρόβλεψη του μέσου ύψους κύματος.

Στο δεύτερο κεφάλαιο αναλύεται ο τρόπος λειτουργίας των δικτύων RBF και του αλγόριθμου των ασαφών μέσων, καθώς και πως μέσω του αλγόριθμου Fuzzy Means υπολογίζονται τα κέντρα.

Στο τρίτο κεφάλαιο περιγράφονται τα δεδομένα, η ανάλυση της μεθοδολογίας που χρησιμοποιήθηκε και τυχόν βελτιστοποιήσεις οι οποίες έγιναν σε αυτή.

Τέλος, η διπλωματική εργασία ολοκληρώνεται με την παράθεση των αποτελεσμάτων, την εξαγωγή συμπερασμάτων και τις προτάσεις για μελλοντική έρευνα.

ΚΕΦΑΛΑΙΟ 1^ο :Μετεωρολογικές προβλέψεις-Πρόβλεψη κύματος

Μια από τις σημαντικότερες επιστημονικές προόδους του προηγούμενου αιώνα είναι η ικανότητά προσομοίωσης πολύπλοκων φυσικών συστημάτων χρησιμοποιώντας αριθμητικά μοντέλα, ώστε να προβλεφθεί η εξέλιξή τους. Ένα παράδειγμα τέτοιου συστήματος αποτελεί η πρόγνωση του καιρού, μέσω της εφαρμογής των αρχών της φυσικής, οι οποίες συνδυάζονται με μια ποικιλία στατιστικών αλλά και εμπειρικών τεχνικών.

Εκτός από τις προβλέψεις των ίδιων των ατμοσφαιρικών φαινομένων, η πρόγνωση του καιρού περιλαμβάνει προβλέψεις για αλλαγές στο επιφανειακό κλίμα της Γης. Αυτές οι αλλαγές προκαλούνται από ατμοσφαιρικές συνθήκες όπως το χιόνι και ο πάγος, οι παλίρροιες, οι καταιγίδες και πλημμύρες [9].

Η βάση για την πρόβλεψη του καιρού ξεκίνησε με τις θεωρίες των αρχαίων Ελλήνων φιλοσόφων και συνεχίστηκε κατά την περίοδο της Αναγέννησης. Ακολούθησε η επιστημονική επανάσταση του 17ου και 18ου αιώνα και έπειτα τα θεωρητικά μοντέλα του 20ου και 21ου αιώνα, τα οποία βοήθησαν στη βελτίωση των εφαρμογών.

Οι μετεωρολογικές προβλέψεις χρειάζονται την συστηματική συλλογή καταγραφών του καιρού διαφόρων γεωγραφικών περιοχών και την σωστή ανάλυση χρησιμοποιώντας τα δεδομένα αυτά για την πρόβλεψη. Η μετεωρολογία έχει σαν πεδίο έρευνας το πώς η ατμόσφαιρα αλλάζει με την πάροδο του χρόνου συναρτήσει διάφορων παραγόντων, έτσι σύνθετες μαθηματικές εξισώσεις μπορούν να καταστρωθούν με σκοπό την αναπαράσταση των φαινομένων. Αυτές οι εξισώσεις στην συνέχεια εξελίσσονται σε αριθμητικά μοντέλα. Η έξοδοι και οι εκτιμήσεις από αυτά τα μοντέλα μπορούν να χρησιμοποιηθούν ως βοήθημα στους μετεωρολόγους για την προετοιμασία των προβλέψεων τόσο βραχυπρόθεσμα όσο και μακροπρόθεσμα.

1.1 Μέθοδοι πρόγνωσης

Η πρόγνωση του καιρού μπορεί να οριστεί ως η πράξη πρόβλεψης μελλοντικών καιρικών συνθηκών ή ως μια προσπάθεια να υποδειχθούν οι καιρικές συνθήκες που είναι πιθανό να συμβούν.

Οι καιρικές προβλέψεις γίνονται συχνά συλλέγοντας ποσοτικά δεδομένα για την τρέχουσα κατάσταση της ατμόσφαιρας χρησιμοποιώντας την επιστημονική κατανόηση των ατμοσφαιρικών διεργασιών, για να προβάλουν την εξέλιξη της ατμόσφαιρας στο μέλλον.

Προτού γίνει εκτενέστερη επεξήγηση του ορισμού που τέθηκε παραπάνω, είναι ιδιαίτερα σημαντικό, στην προσπάθειά μας για την καλύτερη κατανόησή του, να γίνει μία συνοπτική ιστορική αναδρομή. Πιο συγκεκριμένα, στην αρχαιότητα, η πρόβλεψη βασιζόταν κυρίως στην παρατήρηση του καιρού. Η επιστημονική μελέτη αυτής και της μετεωρολογίας δεν αναπτύχθηκε έως ότου έγιναν διαθέσιμα τα πρώτα όργανα μέτρησης. Η αρχή της σηματοδοτήθηκε με την εφεύρεση του βαρόμετρου υδραργύρου από τον Evangelista Torricelli, έναν Ιταλό φυσικό-μαθηματικό, στα μέσα του 17ου αιώνα και τη σχεδόν ταυτόχρονη ανάπτυξη ενός αξιόπιστου θερμομέτρου [10]. Παράλληλα, στη μετεωρολογική έρευνα συνέβαλε σημαντικά μια σειρά αξιοσημείωτων επιτευγμάτων από χημικούς και φυσικούς του 17ου και 18ου αιώνα. Η διατύπωση των νόμων της πίεσης, της θερμοκρασίας και της πυκνότητας αερίου από τους Robert Boyle και Jacques-Alexandre-César Charles, καθώς και η ανάπτυξη του νόμου των μερικών πιέσεων μίκτων αερίων από τον John Dalton, είναι μερικές ακόμη από τις σημαντικότερες επιστημονικές ανακαλύψεις της περιόδου εκείνης

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

που κατέστησαν δυνατή τη μέτρηση και την καλύτερη κατανόηση μέχρι τότε άγνωστων πτυχών της ατμόσφαιρας και τη συμπεριφορά της [10].

Με την πάροδο των ετών, η μελέτη των καιρικών προτύπων οδήγησε σε διαφορετικές τεχνικές για την πρόγνωση του [11]. Κάποια παραδείγματά τέτοιων τεχνικών αποτελούν: η στοχαστική πρόγνωση, η αναλογική τεχνική, τα αριθμητικά μοντέλα κα. Συγκεκριμένα, για τα αριθμητικά μοντέλα, ο τρόπος της λειτουργίας τους είναι βασισμένος στον διαχωρισμό του χώρου σε μια σειρά από “κουτιά. Στη μέση κάθε πλαισίου υπάρχει ένα σημείο για το οποίο το μοντέλο υπολογίζει αναλυτικά τις μεταβλητές του καιρού και κάνει προβλέψεις. Το αποτέλεσμα αυτού του τρισδιάστατου εγκιβωτισμού της ατμόσφαιρας είναι γνωστό ως πλέγμα [12].

Επιστρέφοντας, λοιπόν, στην ανάλυση του τι σημαίνει πρόγνωση καιρού, ένα γνώρισμα της αποτελεί η ταξινόμησή της με βάση το χρόνο ή τη διάρκεια της περιόδου πρόβλεψης σε [13]:

- (1) Πρόβλεψη πολύ μικρής εμβέλειας (Very short range forecast).
- (2) Πρόβλεψη μικρής εμβέλειας (Short range forecast).
- (3) Εκτεταμένη πρόβλεψη (Extended forecast).
- (4) Πρόβλεψη μεγάλης εμβέλειας (Long range forecast).
- (5) Άμεσης πρόβλεψης (Now casting).

Ειδικότερα για τις:

Προβλέψεις πολύ μικρής εμβέλειας:

Σε αυτόν το τύπο πρόβλεψης αναφερόμαστε σε έως και 12ώρη περιγραφή των μεταβλητών του καιρού. Από αυτή μπορεί να παραχθεί ένα πλήρες σύνολο μεταβλητών όπως η θερμοκρασία του αέρα και η σχετική υγρασία, η ταχύτητα και κατεύθυνση του ανέμου, η ηλιακή ακτινοβολία, η ποσότητα και ο τύπος βροχόπτωσης με απαραίτητη προϋπόθεση να είναι η διαθεσιμότητα αποτελεσματικών συστημάτων μετάδοσης. Ένα παράδειγμα που θα μπορούσε να αναφερθεί είναι πως οι πληροφορίες για τον παγετό πρέπει να μεταδίδονται στους αγρότες που έχουν την δυνατότητα να ενεργοποιήσουν εγκαταστάσεις άρδευσης ή πυρκαγιών ή άλλα συστήματα προστασίας. Τέτοιες προβλέψεις συνήθως παρουσιάζουν μεγάλη ακρίβεια.

Προβλέψεις μικρής εμβέλειας:

Οι ημερήσιες προβλέψεις του καιρού θεωρούνται προβλέψεις μικρής εμβέλειας και βασίζονται κυρίως σε τρέχοντα καιρικά δεδομένα όπως η πίεση και οι αλλαγές θερμοκρασίας. Ισχύουν από 12 έως 72 ώρες με ακρίβεια 70 - 80%. Κάποιες ενδεικτικές εφαρμογές είναι ο προγραμματισμός άρδευσης, η προσαρμογή χρόνου αγροτικών εργασιών και η προστασία των φυτών από πιθανό παγετό.

Εκτεταμένη πρόβλεψη:

Η εκτεταμένη πρόβλεψη δίνει έμφαση στον τύπο του καιρού, τη σειρά των βροχερών ημερών, τους καιρικούς κινδύνους στη γεωργία όπως οι ισχυροί άνεμοι και πραγματοποιείται από 3 έως 10 ημέρες με ακρίβεια 60 - 70%. Είναι χρήσιμη για πολλές αγροτικές εργασίες όπως ο προγραμματισμός συγκομιδής, ο ψεκασμός, η διαχείριση αγροκτημάτων κα.

Πρόβλεψη μεγάλης εμβέλειας:

Αυτός ο τύπος πρόβλεψης είναι επωφελής για την πρόγνωση του καιρού για τους επόμενους 1-2 μήνες. Οι μακροπρόθεσμες εποχιακές προβλέψεις βοηθούν στο να επιτρέψουν την προσαρμογή της καλλιέργειας στο αναμενόμενο κλίμα.

Άμεσης πρόβλεψης:

Στην άμεση πρόβλεψη, η πρόγνωση του καιρού δίνεται 2 έως 3 ώρες νωρίτερα. Τα αποτελέσματα της είναι χρήσιμα για την αεροπορία και την πλοήγηση. Επιπλέον, μικρά και βραχυπρόθεσμα φαινόμενα όπως ο κεραυνός μπορούν να προβλεφθούν έως και 6 ώρες νωρίτερα με συστήματα ραντάρ και δορυφόρων.

Τέλος, παίζει εξίσου σημαντικό ρόλο στις μετεωρολογικές προβλέψεις, εκτός από τον χρόνο ή τη διάρκεια της περιόδου πρόβλεψης, η μέθοδος της πρόγνωσης. Οι διάφορες μέθοδοι που χρησιμοποιούνται στη σύγχρονη πρόγνωση του καιρού χωρίζονται στις εξής κατηγορίες:

- (1) Αριθμητικές μέθοδοι.
- (2) Στατιστικές μέθοδοι.
- (3) Συνοπτική πρόγνωση (Synoptic forecasting).
- (4) Στοχαστική πρόγνωση (Ensemble Forecasting).

Οι παραπάνω μέθοδοι θα αναλυθούν περαιτέρω στα παρακάτω υπό κεφάλαια:

1.1.1 Αριθμητική πρόγνωση

Η σύγχρονη πρόβλεψη του καιρού βασίζεται σε μεγάλο βαθμό στην αριθμητική πρόγνωση. Συγκεκριμένα, οι Linacre και Geerts το 1997 όρισαν την αριθμητική πρόγνωση (Numerical Weather Prediction, NWP) ως ένα απλοποιημένο σύνολο εξισώσεων που χρησιμοποιείται για τον υπολογισμό των αλλαγών των συνθηκών. Ενώ, σύμφωνα με τους Lutgens και TarBuck (1989), η λέξη «αριθμητική» είναι παραπλανητική, διότι όλοι οι τύποι μετεωρολογικών προβλέψεων βασίζονται σε ορισμένα ποσοτικά δεδομένα και επομένως θα μπορούσαν να ενταχθούν σε αυτόν τον τίτλο [12].

Σε αυτή την μέθοδο πρόγνωσης, χρησιμοποιούνται μαθηματικά μοντέλα, τόσο της ατμοσφαιράς όσο και του ωκεανού, τα οποία βασίζονται στις τρέχουσες καιρικές συνθήκες. Πιο συγκεκριμένα, η διαδικασία αποτελείται από ένα σύνολο μερικών διαφορικών εξισώσεων, οι οποίες περιγράφουν τις δυναμικές και θερμοδυναμικές διεργασίες στην ατμόσφαιρα της γης [14]. Επιβάλλουν, δηλαδή, τους βασικούς νόμους διατήρησης της μάζας, της ενέργειας και της ορμής, την αρχή διατήρησης της ποσότητας νερού και της συγκέντρωσης των ατμοσφαιρικών συστατικών, ενώ λαμβάνουν υπ' όψιν και τις εξωτερικές διαταραχές, όπως η ηλιακή θέρμανση και οι επιφανειακές έλξεις [15].

Οι προαναφερθέντες εξισώσεις μεταφράζονται σε κώδικα στον υπολογιστή χρησιμοποιώντας αριθμητικές μεθόδους και παραμετροποιήσεις άλλων φυσικών διεργασιών και συνδυάζονται με τις αρχικές και οριακές συνθήκες πριν εκτελεστούν σε μια γεωγραφική περιοχή.

Λόγω της πολυπλοκότητας της μεθόδου NWP απαιτείται η χρήση υπέρ υπολογιστών για την επίλυσή της. Στα μοντέλα αυτά δεν υπάρχει η αναλυτική λύση των εξισώσεων και για αυτό τον λόγο η εύρεση της ακριβούς αναλυτικής λύσης γίνεται χρησιμοποιώντας μια απλοποιημένη μορφή των εξισώσεων ή την αριθμητική επίλυση των βασικών εξισώσεων πάνω σε ένα πλέγμα.

Ειδικότερα, η ατμόσφαιρα αποτελείται από έναν αριθμό κόμβων τα οποία ονομάζονται σημεία του πλέγματος. Οι εξισώσεις στην περίπτωση αυτή επιλύονται προσεγγιστικά σε σημεία του πλέγματος. Τα σημεία αυτά είναι κατανομημένα με ομοιόμορφο ή ανομοιόμορφο τρόπο και καλύπτουν οριζόντια και κάθετα την περιοχή στην οποία θα πραγματοποιηθεί η πρόβλεψη. Όσο πιο μεγάλος είναι ο αριθμός των σημείων του πλέγματος τόσο πιο πολύπλοκη και χρονοβόρα καθίσταται η προσομοίωση. Σημαντική παράμετρος της μεθόδου αυτής, επομένως, είναι η απόστασή μεταξύ των σημείων του πλέγματος αλλιώς αποκαλούμενη και ως χωρική ανάλυση. Η προσομοίωση που προκύπτει από το μοντέλο, δημιουργεί τη μελλοντική κατάσταση της ατμόσφαιρας σε όλα τα σημεία του πλέγματος .

Το αριθμητικό μοντέλο χωρίζεται σε τρία στάδια τα οποία είναι:

- i. Pre-processing.
- ii. Αριθμητική επίλυση εξισώσεων.
- iii. Post processing.

Τα αριθμητικά μοντέλα χρησιμοποιούνται κυρίως σε εποχιακές προγνώσεις καιρού, σε κλιματικές προγνώσεις, σε μελέτες μετεωρολογικών φαινομένων (τοπικοί κυκλώνες, μέτωπα) και σε μελέτες μικρό κλίμακας.

Αυτό που αξίζει να τονιστεί είναι πως κάθε μοντέλο τέτοιου τύπου χαρακτηρίζεται από παραλείψεις, προσεγγίσεις και συμβιβασμούς. Τα προβλήματα που συναντώνται στην αριθμητική πρόγνωση είναι το πρόβλημα των αρχικών συνθηκών, η έλλειψη δεδομένων στον ωκεανό και την ανώτερη ατμόσφαιρα καθώς και η επίδραση τοπικών φαινομένων στις παρατηρήσεις.

Κατά την διαδικασία της αριθμητικής πρόγνωσης παρατηρούνται συχνά σφάλματα στρογγυλοποιήσεων (round-off error) καθώς ένας υπολογιστής 32-bit επιλύει αριθμούς που διαφέρουν μεταξύ τους κατά 3×10^{-8} και η οποιαδήποτε μικρότερη διαφορά χάνεται. Επιπλέον, παρατηρούνται σφάλματα αποκοπής (truncation error) και αριθμητική αστάθεια καθώς η αριθμητική λύση μπορεί να αποκλίνει από την πραγματική λύση. Επίσης, είναι πιθανό να παρουσιαστούν προβλήματα αν οι ταχύτητες είναι μεγάλες, η ανάλυση του πλέγματος μικρή και το χρονικό βήμα ολοκλήρωσης μεγάλο. Τέλος, τα σφάλματα εξαρτώνται και από την δυναμική αστάθεια.

Για να παραχθούν πιο αξιόπιστες προβλέψεις, οι μετεωρολόγοι στράφηκαν στην μελέτη πολλαπλών αριθμητικών μοντέλων πρόβλεψης του καιρού. Ένας τρόπος με τον οποίο το έχουν πετύχει αυτό είναι συγκρίνοντας πολλές προβλέψεις διαφορετικών αριθμητικών μοντέλων. Συγκεκριμένα, σύγκριναν περιφερειακές με παγκόσμιες προβλέψεις μοντέλων, προβλέψεις από τα μοντέλα που χρησιμοποιούνται σε διαφορετικά κέντρα ή λαμβάνοντας υπ'όψιν διάφορες εκδόσεις του ίδιου μοντέλου. Με τον τρόπο αυτό, εξέταζαν πώς οι νέες παρατηρήσεις άλλαζαν τα αποτελέσματα των προβλέψεων του μοντέλου με την πάροδο του χρόνου.

1.1.2 Στατιστική πρόγνωση

Η μελέτη του κλίματος είναι, σε μεγάλο βαθμό, η μελέτη των στατιστικών του καιρού. Έτσι, δεν αποτελεί έκπληξη το γεγονός ότι η στατιστική ανάλυση αποτελεί αδιάσπαστο κομμάτι των κλιματολογικών επιστημών. Η στατιστική ανάλυση βοηθά στον ποσοτικό προσδιορισμό των επιπτώσεων της αβεβαιότητας, τόσο από την άποψη της παρατήρησης και της μέτρησης όσο και από την άποψη της κατανόησης των διαδικασιών που διέπουν τη μεταβλητότητα του κλίματος. Βοηθά, επίσης, να προσδιοριστούν ποιες από τις πολλές πληροφορίες που προέρχονται από τις παρατηρήσεις του κλίματος είναι άξια ερμηνείας. Οι στατιστικές μέθοδοι απαιτούνται στην συλλογή πρωτογενών δεδομένων και στην ερμηνεία και ανάλυση των συνόλων δεδομένων υψηλού επιπέδου, ενώ συμβάλλουν στην τελική σύνθεση της γνώσης για το κλίμα [16]. Τέλος, χρησιμοποιούνται κυρίως στις μακροχρόνιες προγνώσεις του καιρού και λειτουργούν συμπληρωματικά μαζί με τις μεθόδους αριθμητικής πρόβλεψης καιρού.

Ειδικότερα, στις στατιστικές μεθόδους, λαμβάνονται υπ'όψιν τα προηγούμενα καιρικά δεδομένα, καθώς υποτίθεται ότι ο καιρός στο μέλλον θα είναι μια αναπαράσταση αυτού στο παρελθόν. Ο κύριος σκοπός της μελέτης των προηγούμενων καιρικών δεδομένων είναι να ανακαλύψει εκείνες τις πτυχές του καιρού που αποτελούν καλούς δείκτες των μελλοντικών γεγονότων. Η διαδικασία είναι να συγκεντρωθούν στατιστικά δεδομένα που σχετίζονται με τη θερμοκρασία, με την ταχύτητα και την κατεύθυνση του ανέμου, την ποσότητα συννεφιάς, την υγρασία και την εποχή του έτους. Στη συνέχεια, αυτά τα δεδομένα απεικονίζονται σε γραφήματα τα οποία παρέχουν μια εκτίμηση της μέγιστης θερμοκρασίας για την ημέρα από τα δεδομένα των τρέχων συνθηκών.

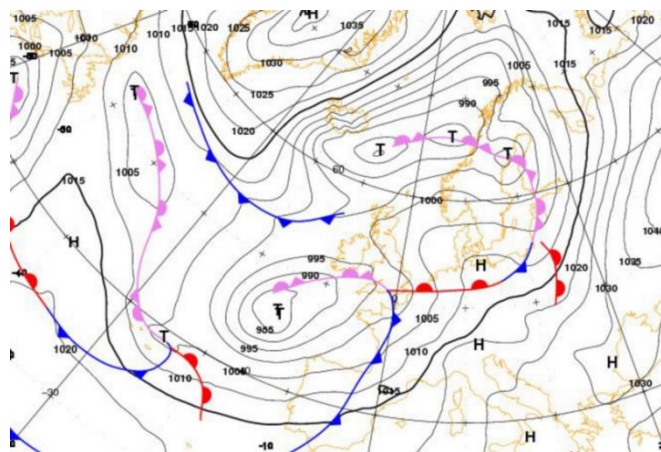
Μια στατιστική προσέγγιση για την πρόβλεψη του καιρού αποτελεί η αναλογική μέθοδος. Σε αυτή τη μέθοδο γίνεται μια προσπάθεια να εντοπιστούν παλαιότερες καιρικές καταγραφές οι οποίες είναι σχεδόν παρόμοιες με τις τρέχουσες συνθήκες. Δηλαδή, γίνεται μια προσπάθεια αντιστοίχισης του προβλεπόμενου μοτίβου σε μια ευρεία περιοχή (όπως μια ήπειρο ή ακόμα και ολόκληρο το ημισφαίριο) με παρόμοιες εποχές στο παρελθόν. Ο κατάλογος των προηγούμενων παραδειγμάτων μπορεί στη συνέχεια να συνδυαστεί και να προωθηθεί εγκαίρως για να παρουσιαστεί μια πιθανή μελλοντική εξέλιξη. Σε αυτή τη μέθοδο, όπως και στην περίπτωση της αριθμητικής μεθόδου, το κύριο πρόβλημα είναι η έλλειψη πλήρους ενημέρωσης για τις καιρικές συνθήκες.

Η αναλογική πρόβλεψη είναι μια απλή αλλά ισχυρή μέθοδος που αποτελεί ουσιαστικό μέρος της σύγχρονης πρόγνωσης του καιρού. Η Παγκόσμια Υπηρεσία Κλίματος έχει αναπτύξει μια σειρά από ευέλικτα εποχιακά αναλογικά συστήματα πρόβλεψης, τα οποία επιτρέπουν στους χρήστες να παράγουν γρήγορα και εύκολα τις δικές τους αναλογικές προβλέψεις. Όταν τα δυναμικά μοντέλα διαφωνούν μεταξύ τους και με τις σχετικές αναλογικές προβλέψεις, τότε η Παγκόσμια Υπηρεσία Κλίματος βασίζεται συχνά σε ένα ανάλογο για την παραγωγή της μηνιαίας εποχικής κλιματικής πρόβλεψης. Οι προβλέψεις υψηλής εμπιστοσύνης καθίστανται δυνατές όταν οι δυναμικές προβλέψεις και οι κατάλληλες αναλογικές προβλέψεις συμφωνούν και αλληλοσυμπληρώνονται [17].

1.1.3 Συνοπτική πρόγνωση

Η συνοπτική πρόγνωση αποτελεί την παραδοσιακή προσέγγιση στην πρόβλεψη του καιρού. Με το όρο “συνοπτική” εννοούμε ότι η παρατήρηση διαφορετικών καιρικών στοιχείων αναφέρεται σε συγκεκριμένο χρονικό διάστημα παρατήρησης. Με βάση τον ορισμό που δόθηκε, ένας “συνοπτικός” χάρτης απεικονίζει τις ατμοσφαιρικές συνθήκες σε μια δεδομένη στιγμή. Για να έχει μια σφαιρική άποψη της αλλαγής του καιρού, ένα σύγχρονο μετεωρολογικό κέντρο ετοιμάζει μια σειρά από συνοπτικά διαγράμματα κάθε μέρα. Τέτοιοι συνοπτικοί χάρτες αποτελούν την ίδια τη βάση των καιρικών προγνώσεων.

Όπως αναφέρθηκε προηγουμένως, το έργο της προετοιμασίας συνοπτικών διαγραμμάτων σε τακτική βάση περιλαμβάνει τεράστια συλλογή και ανάλυση δεδομένων παρατήρησης που λαμβάνονται από χιλιάδες μετεωρολογικούς σταθμούς. Από την προσεκτική μελέτη των καιρικών διαγραμμάτων για πολλά χρόνια διατυπώθηκαν ορισμένοι εμπειρικοί κανόνες. Αυτοί οι κανόνες βοήθησαν τους μετεωρολόγους να εκτιμήσουν τον ρυθμό και την κατεύθυνση της κίνησης των καιρικών συστημάτων. Στο σχήμα 1.1 φαίνεται ένα παράδειγμα ενός συνοπτικού καιρικού διαγράμματος:



Σχήμα 1.1: Synoptic weather map

1.1.4 Στοχαστική πρόγνωση

Σε ένα μοντέλο πρόγνωσης το οποίο προβλέπει τα χαρακτηριστικά του καιρού, τα σφάλματα πρόβλεψης αυξάνονται αναπόφευκτα με τον πάροδο του χρόνου λόγω της χαοτικής φύσης της ατμόσφαιρας και της ανακρίβειας των αρχικών παρατηρήσεων. Από ένα σημείο και μετά τα σφάλματα είναι τόσο μεγάλα με αποτέλεσμα η πρόβλεψη να μην έχει καμία συσχέτιση με την πραγματική κατάσταση της ατμόσφαιρας. Για αυτό τον λόγο, η εξέταση μιας μεμονωμένης πρόβλεψης δεν δίνει καμία ένδειξη για το πόσο πιθανό είναι αυτή η πρόβλεψη να είναι σωστή. Η στοχαστική πρόγνωση είναι ένα σχετικά νέο εργαλείο στο πεδίο της μετεωρολογικής πρόβλεψης. Το κύριο γνώρισμα της είναι ότι επιτρέπει πιο γρήγορες και επιστημονικά βασισμένες συγκρίσεις μεταξύ πολλαπλών μοντέλων. Βασίζεται, δηλαδή, στο ότι μικρά λάθη στην ανάλυση μπορούν να επηρεάσουν την πρόγνωση κατά της διάρκεια της προγνωστικής περιόδου. Αυτό σημαίνει ότι μια λίγο διαφοροποιημένη ανάλυση μπορεί να οδηγήσει σε μια εντελώς διαφορετική πρόγνωση. Λόγω του προβλήματος των αρχικών συνθηκών, λοιπόν, προστίθενται στην ανάλυση ορισμένες διαταραχές οι οποίες προκαλούν ένα εύρος από διαφορετικές αρχικές συνθήκες [18].

Επίσης, στη μέθοδο αυτή συνεπάγεται η παραγωγή πολλών προβλέψεων προκειμένου να αντικατοπτρίζεται η αβεβαιότητα στην αρχική κατάσταση της ατμόσφαιρας (λόγω σφαλμάτων στις παρατηρήσεις και ανεπαρκούς δειγματοληψίας). Η αβεβαιότητα αυτή μπορεί στη συνέχεια να εκτιμηθεί από το εύρος των διαφορετικών προβλέψεων που παράγονται. Η λογική πίσω από αυτόν τον τύπο πρόγνωσης είναι ότι δύο εκτελέσεις ενός μοντέλου δεν αρκούν για να βασιστεί μια πρόβλεψη [12].

Με τα αποτελέσματα της στοχαστική πρόγνωσης να έχουν προστεθεί στην εργαλειοθήκη της αριθμητική πρόγνωσης, οι επιστήμονες διαθέτουν ένα άλλο επίπεδο πληροφοριών που θα τους βοηθήσει να κάνουν έξυπνη χρήση της αριθμητικής πρόγνωσης, βασιζόμενοι και στα αποτελέσματα της στοχαστικής.

1.2 Πρόγνωση κύματος

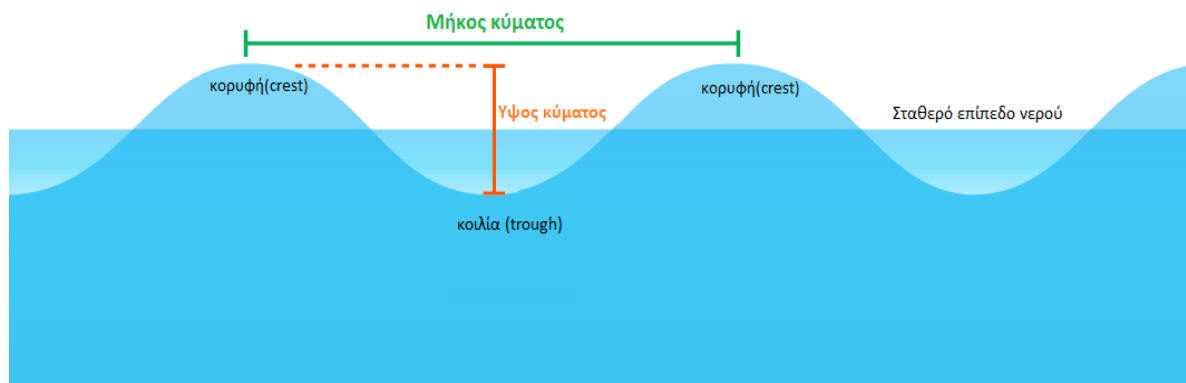
Το μεγαλύτερο μέρος του νερού της Γης είναι συγκεντρωμένο στον ωκεανό. Η ποικιλόμορφη τοπογραφία του ωκεανού τον κάνει ένα ενδιαφέρον θέμα για πολλούς επιστήμονες. Ένα από τα κοινά φαινόμενα που συμβαίνει στον ωκεανό είναι το κύμα της θάλασσας.

Τα κύματα δημιουργούνται από ενέργεια που διέρχεται από το νερό, με αποτέλεσμα να κινείται σε κύκλους. Ωστόσο, το νερό στην πραγματικότητα δεν ταξιδεύει σε κύματα. Τα κύματα μεταδίδουν ενέργεια μέσω του ωκεανού και αν δεν συναντούν εμπόδια έχουν τη δυνατότητα να ταξιδέψουν σε ολόκληρο τον ωκεανό [19]. Είναι το αποτέλεσμα της συνισταμένης των δυνάμεων που δρουν σε αυτόν και οι κυρίαρχες φυσικές δυνάμεις είναι η πίεση από την ατμόσφαιρα (ειδικά μέσω των ανέμων), οι σεισμοί, η βαρύτητα της Γης και των ουράνιων σωμάτων (Σελήνη και Ήλιος), η δύναμη Coriolis (λόγω της περιστροφής της Γης) και η επιφανειακή τάση.

1.2.1 Βασικά χαρακτηριστικά κύματος

Τα χαρακτηριστικά των κυμάτων εξαρτώνται από τις δυνάμεις ελέγχου (controlling forces). Τα παλιρροϊκά κύματα δημιουργούνται από την απόκριση στη βαρύτητα της Σελήνης και του Ήλιου και είναι κύματα μεγάλης κλίμακας. Στο άλλο άκρο της κλίμακας, τα τριχοειδή κύματα (capillary waves) κυριαρχούνται από την επιφανειακή τάση στο νερό. Υπάρχουν κύματα βαρύτητας όπου η βαρύτητα της Γης και η άνωση του νερού είναι οι κύριοι καθοριστικοί παράγοντες [20].

Η απλούστερη κυματική κίνηση μπορεί να αντιπροσωπεύεται από ένα ημιτονοειδές, προοδευτικό κύμα. Αυτό σημαίνει ότι το κύμα επαναλαμβάνεται και έχει την ομαλή μορφή της καμπύλης του ημιτόνου, όπως φαίνεται στο Σχήμα 1.2. Το κύμα αποτελείται από μια σειρά από παράλληλες κορυφές, οι οποίες είναι όλες ίσες σε ύψος και απέχουν ίση απόσταση μεταξύ τους. Η προοδευτική τους φύση φαίνεται στην κίνηση τους με σταθερή ταχύτητα σε κατεύθυνση κάθετη στις κορυφές και στο ότι δεν αλλάζουν μορφή.



Σχήμα 1.2: Βασικά χαρακτηριστικά κύματος

Σε ένα υγρό μέσο διάδοσης, όπως είναι το νερό της θάλασσας, η ενέργεια μεταφέρεται από το ένα σημείο στο άλλο με τον κυματισμό. Τα κύματα που δημιουργούνται αποτελούν μια αλληλουχία από κορυφές και κοιλιές και προσδιορίζονται από τα παρακάτω χαρακτηριστικά [21]:

- Σταθερό επίπεδο νερού (Still water line), είναι το επίπεδο της επιφάνειας της θάλασσας αν ήταν απόλυτα ήρεμη.
- Μήκος κύματος λ (Wave length), είναι η οριζόντια απόσταση μεταξύ δύο διαδοχικών κορυφών ή κοιλιών.
- Περίοδος κύματος T (Wave period), είναι το χρονικό διάστημα (σε δευτερόλεπτα) μεταξύ της διέλευσης διαδοχικών κορυφών από το ίδιο σταθερό σημείο.
- Συχνότητα f (Wave frequency), είναι ο αριθμός των κορυφών που περνούν από ένα σταθερό σημείο στην μονάδα του χρόνου. Συνήθως μετριέται σε Hertz και εναλλακτικά υπολογίζεται ως $f = 1/T$.
- Ύψος κύματος H (Wave height), είναι η κάθετη απόστασή μεταξύ της υψηλότερης κορυφής και της χαμηλότερης κοιλιάς ενός κύματος.
- Ταχύτητα u (Wave Velocity), είναι η ταχύτητα η οποία εξαρτάται από το βάθος των νερών της θάλασσας στην οποία διαδίδεται. Όσο μειώνεται το βάθος της θάλασσας η ταχύτητα των κυμάτων μειώνεται.

1.2.2 Συμβατικά μοντέλα πρόβλεψης κύματος

Στις αρχές της δεκαετίας του 1960 υπήρξε μια ταχεία ανάπτυξη της στατιστικής θεωρίας των κυμάτων των ωκεανών, με αποκορύφωμα τη βασική εξελικτική εξίσωση για το κυματικό φάσμα: την εξίσωση του ενεργειακού ισοζυγίου. Αυτή η εξίσωση εκφράζει ότι η εξέλιξη του κυματικού φάσματος καθορίζεται, αφενός από τα αδιαβατικά φαινόμενα της διάσπασης και της διάθλασης και αφετέρου από μια συνάρτηση πηγής που είναι το άθροισμα των φυσικών διεργασιών όπως η είσοδος ανέμου, η διάχυση που παρατηρείται όταν σκάει το κύμα και οι μη γραμμικές αλληλεπιδράσεις. Η εξίσωση του ενεργειακού ισοζυγίου μας δίνει τον ρυθμό μεταβολής της κατάστασης της θάλασσας που προκαλείται από αδιαβατικές διεργασίες [22].

Ωστόσο, για να ξεκινήσει η πρόβλεψη κύματος με την μέθοδο της αριθμητικής πρόγνωσης χρησιμοποιώντας την πλήρη εξίσωση του ενεργειακού ισοζυγίου, συμπεριλαμβανομένης μιας ρητής αναπαράστασης της φυσικής των μη γραμμικών αλληλεπιδράσεων, χρειάστηκαν πολλά χρόνια. Υπάρχουν διάφοροι λόγοι για τους οποίους χρειάστηκε τόσος πολύς χρόνος μέχρι να γίνει η πρόβλεψη κυμάτων ακολουθώντας αυτήν την ορθολογική προσέγγιση. Πρώτα απ' όλα, στις αρχές της δεκαετίας του '60 η σημασία των μη γραμμικών αλληλεπιδράσεων δεν αναγνωρίστηκε, με αποτέλεσμα οι ερευνητές να προσεγγίσουν τα μοντέλα κυματικού φάσματος με τέτοιο τρόπο ώστε να περιλαμβάνουν μόνο τα αποτελέσματα της κίνησης της κυματικής ενέργειας, την είσοδο ανέμου και μια στοιχειώδη μορφή διασποράς.

Τα κυματικά μοντέλα πρώτης γενιάς που αναπτύχθηκαν στη δεκαετία εκείνη και στις αρχές της δεκαετίας του 1970 αντιμετώπισαν το πρόβλημα της μοντελοποίησης του πλήρους ενεργειακού ισοζυγίου. Σε αυτά τα μοντέλα θεωρήθηκε ότι τα συστατικά του κύματος σταματούσαν ξαφνικά να αναπτύσσονται όταν έφταναν σε ένα συνολικό επίπεδο κορεσμού (Phillips 1958). Σήμερα γνωρίζουμε ότι κάτι τέτοιο δεν υφίσταται. Αν και αυτά τα λεγόμενα μοντέλα πρώτης γενιάς εφαρμόστηκαν με επιτυχία για πολλά χρόνια, υπήρχαν αμφιβολίες, από την έναρξή τους, ως προς το εάν πραγματικά αντιπροσώπευαν σωστά τη φυσική της παραγωγής κυμάτων [23].

Στη δεκαετία του 1970, τα εκτεταμένα πειράματα ανάπτυξης κυμάτων και οι μετρήσεις της παραμέτρου εισόδου του ανέμου στα κύματα (Mitsuyasu 1968, 1969, Hasselmann et al. 1973) άλλαξαν ριζικά την μορφή του φασματικού ισοζυγίου ενέργειας στο οποίο βασίστηκαν τα μοντέλα πρώτης γενιάς, οδηγώντας έτσι στην ανάπτυξη των μοντέλων κύματος δεύτερης γενιάς. Στα μοντέλα δεύτερης γενιάς οι μη γραμμικές αλληλεπιδράσεις παραμετροποιούνται προκειμένου να αναπαράχθει μια ορισμένη φασματική κατανομή ενέργειας [24].

Γενικά τα μοντέλα αυτά εφαρμόζονταν με επιτυχία και μερικά από αυτά εξακολουθούν να λειτουργούν μέχρι και σήμερα. Ωστόσο, από την αρχή έγινε σαφές ότι για έντονα ανομοιόμορφα πεδία ανέμου (π.χ. έντονοι τροπικοί κυκλώνες ή καταιγίδες μεσαίου γεωγραφικού πλάτους), οι επαγωγικοί όροι είναι σε θέση να ανταγωνιστούν την τότε γραμμική κατανομή της ενέργειας και σημαντικές αποκλίσεις από την κατά κάποιο τρόπο ισορροπημένη φασματική κατανομή. Η μελέτη Sea Wave Modeling Project (SWAMP) [25] συνέκρινε τα αποτελέσματα εννέα διαφορετικών μοντέλων πρώτης και δεύτερης γενιάς σε απλές, υποθετικές συνθήκες. Ιδιαίτερα, για ένα πεδίο ανέμου τυφώνα διαπιστώθηκε ότι τα μοντέλα συμπεριφέρονταν αρκετά διαφορετικά. Ένα παράδειγμα που μπορεί να παρουσιαστεί είναι ότι το μέγιστο ύψος κύματος κυμαινόταν μεταξύ 8 και 25 m, αντικατοπτρίζοντας έτσι, την έλλειψη γνώσης για εκείνη τη στιγμή. Συγκεκριμένα, τα αποτελέσματα των μοντέλων δεύτερης γενιάς και οι εμπειρικούς «κανόνες» δεν φαινόταν πάντα να συμφωνούν. Σε ταχέως μεταβαλλόμενες περιπτώσεις απαιτείται μια ρητή αναπαράσταση των μη γραμμικών αλληλεπιδράσεων.

Αν και τόσο τα κυματικά μοντέλα πρώτης όσο και δεύτερης γενιάς μπορούν να ρυθμιστούν ώστε να παρέχουν χρήσιμα αποτελέσματα για ορισμένες κατηγορίες αιολικών πεδίων, έχει αποδειχθεί ότι κανένα από τα υπάρχοντα μοντέλα κυμάτων δεν ήταν αξιόπιστο για ακραίες καταστάσεις για τις οποίες απαιτούνται συχνά οι προβλέψεις κύματος. Στην περίπτωση των μοντέλων δεύτερης γενιάς τα προβλήματα σχετίζονται με την αριθμητική ανάλυση και όχι την φυσική προσέγγιση [22]. Έτσι, λοιπόν, λόγω των προβλημάτων που αντιμετώπιζαν τα προηγούμενα μοντέλα, προτάθηκε η ανάπτυξη των μοντέλων τρίτης γενιάς στα οποία το κυματικό φάσμα υπολογιζόταν μόνο με την ενσωμάτωση της βασικής εξίσωσης φασματικής μεταφοράς, χωρίς τον προηγούμενο περιορισμό του φασματικού σχήματος [23]. Συγκεκριμένα, με τον όρο φαινόμενα μεταφοράς ορίζουμε, τις μη αντιστρεπτές διεργασίες της

στατιστικής φύσης που απορρέουν από την τυχαία ή τη συνεχή κίνηση των μορίων. Για παράδειγμα στα κύματα, τα φαινόμενα αυτά σχετίζονται με τη μεταφορά ορμής.

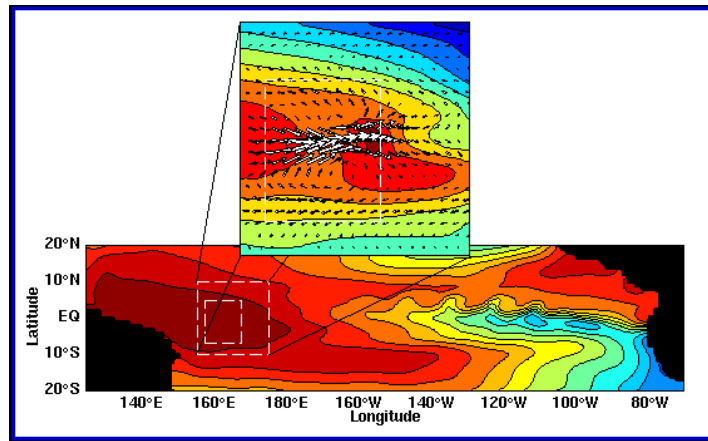
Όπως έχει ήδη αναφερθεί, τα μοντέλα πρώτης γενιάς αναπτύχθηκαν με βάση απλά αιοικά πεδία και χωρίς κυρίαρχες μη γραμμικές αλληλεπιδράσεις και απώλεια ενέργειας, ενώ τα μοντέλα δεύτερης γενιάς χρησιμοποιώντας ποίκιλα πεδία ανέμου και απλοποιημένες μη γραμμικές αλληλεπιδράσεις. Σε αυτά η επιφάνεια της θάλασσας ορίζεται ως το άθροισμα ενός μεγάλου αριθμού μεμονωμένων κυμάτων, όπου κάθε κύμα διαδίδεται με σταθερή συχνότητα σύμφωνα με τη θεωρία γραμμικών κυμάτων [26]. Από την άλλη πλευρά, η διαδικασία μοντελοποίησης βελτιώθηκε περαιτέρω στα μοντέλα τρίτης γενιάς και η διακριτοποίηση του χώρου με την βοήθεια ορθογώνιων πλεγμάτων εισήχθη στη μοντελοποίηση κυμάτων. Αυτά τα μοντέλα χρησιμοποιούν την εξίσωση του ενεργειακού ισοζυγίου για την περιγραφή της χρονικής και χωρικής εξέλιξης των φασμάτων κυμάτων.

Ειδικότερα, ο βασικός στόχος των κυματικών μοντέλων τρίτης γενιάς είναι να προσεγγίσουν το πρόβλημα της αξιολόγησης των συνθηκών των κυμάτων από φυσική άποψη, αποφεύγοντας πολλές από τις συντομεύσεις που υπήρχαν στα μοντέλα πρώτης και δεύτερης γενιάς. Στα τέλη της δεκαετίας του 1970 ήταν διαθέσιμος ένας αυξανόμενος αριθμός μοντέλων κυμάτων. Ενώ ο κοινός στόχος τους ήταν η ικανότητα σωστής εκτίμησης των συνθηκών των κυμάτων σε μια δεδομένη κατάσταση, προτάθηκαν πολλές διαφορετικές προσεγγίσεις στο πρόβλημα ανάλογα με τις συνθήκες στις οποίες αναμενόταν να λειτουργήσει κάθε μοντέλο. Σχεδόν όλα τα μοντέλα είχαν σοβαρούς περιορισμούς στη μαθηματική αναπαράσταση της φυσικής των κυμάτων ανέμου. Αυτό με τη σειρά του οδήγησε σε κακές επιδόσεις όταν τα μοντέλα έπρεπε να λειτουργούν εκτός του εύρους συνθηκών που είχαν αναπτυχθεί [27].

Τα μοντέλα κυματικού φάσματος τρίτης γενιάς επιλύονται χρησιμοποιώντας αριθμητικές μεθόδους σε ορθογώνια δομημένα πλέγματα και ακόμη και με τους σημερινούς υπολογιστές, είναι πολύ απαιτητικά υπολογιστικά [28]. Τα μοντέλα τρίτης γενιάς χρησιμοποιούνται εκτενώς για την προσομοίωση κυματικών συνθηκών και την δημιουργία προβλέψεων στον ωκεανό. Πρόσφατα, η εφαρμογή τέτοιων μοντέλων έχει επεκταθεί και στις παράκτιες περιοχές. Η κύρια εστίαση ήταν μέχρι στιγμής στη διαμόρφωση των συναρτήσεων που περιγράφουν τα διάφορα φυσικά φαινόμενα, τα οποία αφορούν τα ρηχά νερά.

Ένα παράδειγμα που μπορεί να χρησιμοποιηθεί είναι ότι, για την επίλυση των χαρακτηριστικών μεγεθών των φυσικών φαινομένων στις παράκτιες περιοχές απαιτείται η χρήση λεπτού πλέγματος. Στη ζώνη θραύσης (breaking zone) χρειάζεται ανάλυση της τάξης των 10 μέτρων. Απαιτείται επίσης υψηλή ανάλυση για την επίλυση των σύνθετων τοπογραφικών βυθού σε περιβάλλοντα με ρηχά νερά, όπως νησίδες, ύφαλοι και κανάλια. Η ανάγκη για τοπικά μοντέλα υψηλής ανάλυσης μπορεί να επιτευχθεί με τη χρήση της τεχνικής της συνεργασίας ενός τοπικού μοντέλου με λεπτό πλέγμα που ενσωματώνεται με ένα μοντέλο χονδροειδούς πλέγματος [28].

Με στόχο τη μείωση της υπολογιστικής ισχύς είναι επιθυμητό να εισαχθούν πιο ευέλικτα πλέγματα ως εναλλακτική λύση στα μοντέλα αυτά. Ο έλεγχος της κατανομής των κόμβων επιτρέπει τη βέλτιστη χρήση τους και την προσαρμογή της ανάλυσης πλέγματος στις σχετικές φυσικές κλίμακες. Τα εύκαμπτα πλέγματα μπορούν να επιτευχθούν με διάφορους τρόπους, π.χ. καμπυλόγραμμα πλέγματα πολλαπλών τμημάτων, επικαλυπτόμενα πλέγματα, βελτίωση τοπικού πλέγματος και μη δομημένα πλέγματα. Στο σχήμα 1.3 βλέπουμε ένα παράδειγμα του πλέγματος του ωκεανού.



Σχήμα 1.3: Nested Grid Ocean Model

Παρόλο που τα μοντέλα τρίτης γενιάς χρησιμοποιούν τις αριθμητικές μεθόδους που αναφέρθηκαν και τα αποτελέσματά τους είναι πολύ ακριβή απαιτούν μεγάλο υπολογιστικό χρόνο για να εκτελεστούν. Λόγω αυτού, λοιπόν, και των μειονεκτημάτων των υπόλοιπων μοντέλων, κρίθηκε αναγκαίο το να δημιουργηθεί ένας νέος τρόπος πρόβλεψης του ύψους κύματος.

1.2.3 Πρόβλεψη κύματος και χρήση νευρωνικών δικτύων

Τα τελευταία χρόνια υπάρχουν πολλές έρευνες στη βιβλιογραφία για την πρόβλεψη ύψους κύματος, οι οποίες χρησιμοποιούν μεθόδους όπως η μηχανική μάθηση, οι γενετικοί αλγόριθμοι και τα ασαφή συστήματα. Τα νευρωνικά δίκτυα μπορούν να προσεγγίσουν οποιαδήποτε περίπλοκη μη γραμμική διαδικασία χωρίς εκ των προτέρων γνώση της υποκείμενης φυσικής, και έχουν χρησιμοποιηθεί ευρέως για την επίλυση του προβλήματος της πρόβλεψης του μέσου ή του μέγιστου κύματος. Αρκετές τοπολογίες τους, όπως τα δίκτυα πρόσθιας τροφοδότησης, τα νευρωνικά δίκτυα ακτινικής βάσης και τα αναδρομικά δίκτυα, έχουν εφαρμοστεί για την πρόβλεψη του ύψους κύματος σε πληθώρα μελετών [29].

Πιο συγκεκριμένα, τα νευρωνικά δίκτυα έχουν πολλές εφαρμογές στην ωκεανογραφία, συμπεριλαμβανομένης της πρόβλεψης της στάθμης του νερού και του ύψους κύματος. Σε αυτές τις εφαρμογές, οι παράμετροι των κυμάτων προβλέπονται χρησιμοποιώντας τοπικά δεδομένα ανέμου ή κυμάτων για τη θέση πρόβλεψης, καθώς μόνο αυτά τα δεδομένα είναι διαθέσιμα. Η βελτίωση της ακρίβειας της πρόβλεψης επιχειρήθηκε με την εφαρμογή ολοένα και πιο περίπλοκων μοντέλων νευρωνικών δικτύων. Αν και οι εφαρμογές έδειξαν την ικανότητα των νευρωνικών δικτύων να προβλέπουν παραμέτρους κυμάτων με βάση δεδομένα ανέμου και κυμάτων, τα αποτελέσματά αυτών των προβλέψεων δεν ήταν πάντα ικανοποιητικά.

Σύμφωνα με έρευνες [30], έχει αποδειχθεί ότι ένα νευρωνικό δίκτυο μπορεί να κάνει εξίσου καλές προβλέψεις όπως και τα αριθμητικά μοντέλα. Ωστόσο, η ακρίβεια του μοντέλου δεν είναι εγγυημένη. Εξαρτάται κυρίως από την ποιότητα και την ποσότητα των δεδομένων, καθώς και από τον τρόπο με τον οποίο εισάγονται τα δεδομένα εισόδου στο δίκτυο. Οι εφαρμογές των νευρωνικών δικτύων στις προβλέψεις θαλάσσιων χαρακτηριστικών, που έχουν γίνει έως τώρα, προσανατολίζονται κυρίως στην εκτίμηση ή την πρόβλεψη των τιμών κάποιας συγκεκριμένης τυχαίας μεταβλητής. Οι μεταβλητές που σχετίζονται με τον ωκεανό είναι το ύψος κύματος, τα χαρακτηριστικά κατεύθυνσης των κυμάτων, τα παλιρροϊκά επίπεδα, οι χρόνοι υψηλού και χαμηλού νερού, η στάθμη της θάλασσας, οι ταχύτητες ανέμου, τα

χαρακτηριστικά εκβολών ποταμών, τα παράκτια ρεύματα και άλλες μετεωρολογικές παράμετροι των ωκεανών [31].

Από τα προαναφερθέντα χαρακτηριστικά, ωστόσο, αυτό που έχει απασχολήσει τους επιστήμονες περισσότερο είναι το ύψος κύματος. Συγκεκριμένα, η πρόβλεψη του μέσου ύψους κύματος (Significant Wave Height, SWHT) παίζει σημαντικό ρόλο σε τομείς της θαλάσσιας μηχανικής όπως η αλιεία, η παραγωγή ενέργειας και οι θαλάσσιες μεταφορές. Επιπλέον, οι τρέχουσες μέθοδοι πρόβλεψης του ύψους κύματος περιορίζονται σε μεγάλο βαθμό στην πρόβλεψη του μέσου κύματος ενός σημείου, χωρίς να λαμβάνεται υπόψη η πρόβλεψη του μέσου ύψους διπλανών σημείων. Γενικώς, τα αριθμητικά και στατιστικά μοντέλα έχουν χρησιμοποιηθεί ευρέως στην παγκόσμια πρόβλεψη της θαλάσσιας κατάστασης. Τόσο οι μέθοδοι αριθμητικών μοντέλων όσο και οι στατιστικές μέθοδοι προσπαθούν να προβλέψουν το μέσο ύψος κύματος προσεγγίζοντας το με μαθηματικά σχεσιακά μοντέλα. Κατά κανόνα η μοντελοποίηση βασίζεται σε ντετερμινιστικές εξισώσεις ή προσεγγίσεις στοχαστικού πεδίου χρόνου (stochastic time domain approaches). Ωστόσο, αυτές δεν μπορούν να εξηγήσουν πλήρως την πολυπλοκότητα και την αβεβαιότητα των κυματικών φαινομένων [32]. Λόγω της ισχυρής μη γραμμικότητας των φυσικών διεργασιών και των μηχανισμών των κυμάτων του ωκεανού, ειδικά σε ακραίες περιπτώσεις, τα μοντέλα αυτά ενδέχεται να μην έχουν υψηλή ακρίβεια και πρέπει να βελτιωθούν. Τα αριθμητικά μοντέλα απαιτούν ακριβή μετεωρολογικά και ωκεανογραφικά δεδομένα και μεγάλο όγκο υπολογιστικών εργασιών, με αποτέλεσμα ο χρόνος να εμποδίζει την ανάπτυξη ταχείας και ακριβούς πρόβλεψης του μέγιστου ύψους κύματος [33].

Με την ταχεία ανάπτυξη της τεχνητής νοημοσύνης (Artificial Intelligence), λόγω των πλεονεκτημάτων της (γρήγορη ταχύτητα υπολογισμού, χαμηλό υπολογιστικό κόστος και ισχυρή μη γραμμική ικανότητα εκμάθησης) τα τελευταία χρόνια, η μέθοδος πρόβλεψης SWHT που βασίζεται στη βαθιά μάθηση έχει εκτιμηθεί ιδιαίτερα από τους ερευνητές. Η μέθοδος βαθιάς μάθησης χρειάζεται μόνο να γνωρίζει ποιοι παράγοντες σχετίζονται με τη στοχευόμενη φυσική ποσότητα, να δημιουργήσει ένα μοντέλο πρόβλεψης εισόδου-εξόδου και να προβλέψει το SWHT για το άμεσο μέλλον. Η πρόβλεψη των κυμάτων είναι μια αβέβαιη και τυχαία διαδικασία, κάτι το οποίο την καθιστά ιδανική για μοντελοποίηση με την χρήση νευρωνικών δικτύων [34].

Σύμφωνα με τις μελέτες των Deo and Naidu το 1998, η εφαρμογή των νευρωνικών δικτύων στο πρόβλημα της πρόβλεψης κύματος έδειξε ότι ένα κατάλληλα εκπαιδευμένο δίκτυο θα μπορούσε να παρέχει ικανοποιητικά αποτελέσματα για ορισμένους τύπους προβλέψεων. Συγκεκριμένα, βρέθηκε ότι είναι χρήσιμο για τη διεξαγωγή χαρτογράφησης κυμάτων ανέμου σε ανοιχτές ευρύτερες περιοχές, σε βαθιά νερά και όταν τα διαστήματα δειγματοληψίας και πρόβλεψης είναι μεγάλα, όπως μία εβδομάδα. Ενώ, στην περίπτωση όπου οι μετρήσεις ανέμου σε μικρά διαστήματα παρουσιάζουν συνήθως διακυμάνσεις στις τιμές τους, η εκπαίδευση του δικτύου δυσκολεύει. Έτσι, στην περίπτωση των δεδομένων σύντομων διαστημάτων, η ξεχωριστή εκπαίδευση τόσο για τις περιόδους των μουσώνων όσο και για τις εποχές με καλό καιρό είναι ευεργετική. Σε αντίθεση με τα ντετερμινιστικά μοντέλα, η εισαγωγή της διάρκειας του ανέμου δεν φαίνεται να είναι απαραίτητη στα νευρωνικά δίκτυα [35].

Το 2002 οι Agrawal και Deo προέβλεψαν το μέσο ύψος κύματος και την μέση περίοδο κύματος από την είσοδο της ταχύτητας του ανέμου και από τα τρέχοντα και προηγούμενα ύψη κύματος. Βρήκαν ότι όταν πρόκειται για μικρότερα διαστήματα πρόβλεψης, το νευρωνικό δίκτυο είχε σαφή υπεροχή έναντι της στοχαστικής προσέγγισης. Ωστόσο, αυτό δεν συνέβαινε όταν μελετήθηκε μεγαλύτερο διάστημα πρόβλεψης [36]. Ένα χρόνο αργότερα, οι Deo και

Jagdale έδειξαν ότι τα νευρωνικά δίκτυα θα μπορούσαν επίσης να χρησιμοποιηθούν για την πρόβλεψη των υψών θραύσης των κυμάτων [37].

Τα τελευταία χρόνια οι προβλέψεις κυμάτων με την χρήση αριθμητικών μοντέλων που βασίζονται σε διεργασίες όπως το Wave Model (WAM) και το Simulating Waves Nearshore (SWAN), είναι πολύπλοκες και χρησιμοποιούνται ευρέως [23]. Αυτά τα μοντέλα έχουν μεγάλο υπολογιστικό κόστος σε προβλέψεις περίπλοκων κυμάτων σε μεσοκλίμακα κοντά στην ακτή καθώς και υπεράκτια. Επομένως, η εφαρμογή των μοντέλων σε μια ημερήσια πρόβλεψη κυμάτων δεν είναι κατάλληλη. Από την άλλη πλευρά, ένα μοντέλο πρόβλεψης κυμάτων που βασίζεται σε τεχνητό νευρωνικό δίκτυο έχει προταθεί για την πρόβλεψη κύματος. Το μοντέλο πρόβλεψης κύματος που βασίζεται σε νευρωνικά δίκτυα απαιτεί πολύ μειωμένο υπολογιστικό κόστος σε σχέση με το μοντέλο που βασίζεται στην αριθμητική πρόγνωση. Αυτό συμβαίνει διότι μπορεί γρήγορα να προβλέψει την κατάσταση των κυμάτων, αφού η εκπαίδευση και η επικύρωση ολοκληρωθούν επιτυχώς από πιθανές παραμέτρους εισόδου και εξόδου της πίεσης του ανέμου και της στάθμης της θάλασσας [38].

Η πραγματική δύναμη και τα πλεονεκτήματα των νευρωνικών δικτύων έγκειται στην ικανότητα αναπαράστασης τόσο γραμμικών όσο και μη γραμμικών σχέσεων απευθείας από τα δεδομένα που μοντελοποιούνται. Τα παραδοσιακά γραμμικά μοντέλα είναι ανεπαρκή όταν πρόκειται για αληθινά δεδομένα μοντελοποίησης που περιέχουν μη γραμμικά χαρακτηριστικά. Από την άλλη πλευρά, τα μειονέκτημα των μεθόδων μηχανικής μάθησης είναι ότι η ικανότητα μακροπρόθεσμης πρόβλεψης μειώνεται σημαντικά και η ικανότητα γενίκευσης του μοντέλου είναι περιορισμένη. Τέλος, τόσο τα αριθμητικά μοντέλα όσο και η μηχανική μάθηση έχουν πλεονεκτήματα και μειονεκτήματα για την πρόβλεψη ύψους κύματος και ανάλογα την εφαρμογή είναι στην κρίση του μηχανικού ποιο θα χρησιμοποιηθεί [39].

ΚΕΦΑΛΑΙΟ 2^ο : Νευρωνικά Δίκτυα Ακτινικής Βάσης - Αλγόριθμος Fuzzy Means

Ένα νευρωνικό δίκτυο είναι ένας τεράστιος παράλληλος επεξεργαστής με κατανεμημένη αρχιτεκτονική, ο οποίος αποτελείται από απλές μονάδες επεξεργασίας και έχει από την φύση του την δυνατότητα να αποθηκεύει εμπειρική γνώση και να την καθιστά διαθέσιμη για χρήση [40]. Τόσο το όνομα όσο και η δομή των νευρωνικών δικτύων είναι επηρεασμένα από τον ανθρώπινο εγκέφαλο, καθώς μιμούνται τον τρόπο που οι νευρώνες δίνουν σήμα ο ένας στον άλλον.

Η εμφάνιση των νευρωνικών δικτύων παρατηρήθηκε στη δεκαετία του 1940 με το έργο των Warren McCulloch και Walter Pitts, οι οποίοι έδειξαν ότι τα τεχνητά νευρωνικά δίκτυα μπορούσαν, κατ' αρχήν, να υπολογίσουν οποιαδήποτε αριθμητική ή λογική συνάρτηση [41]. Η έρευνα στο πεδίο της τεχνητής νοημοσύνης επιταχύνθηκε γρήγορα, με την Kunihiko Fukushima να αναπτύσσει το πρώτο αληθινό, πολυεπίπεδο νευρωνικό δίκτυο (MLP) το 1980 [42].

Η πρώτη πρακτική εφαρμογή των τεχνητών νευρωνικών δικτύων ήρθε στα τέλη της δεκαετίας του 1950, με την εφεύρεση του perceptron και του σχετικού κανόνα μάθησης από τον Frank Rosenblatt. Ο Rosenblatt και οι συνεργάτες του δημιούργησαν ένα δίκτυο perceptron και επιβεβαίωσαν την ικανότητά του να αναγνωρίζει πρότυπα [43]. Αυτή η πρόωμη επιτυχία προκάλεσε μεγάλο ενδιαφέρον για την έρευνα νευρωνικών δικτύων. Δυστυχώς, αργότερα αποδείχθηκε ότι το συγκεκριμένο δίκτυο perceptron μπορούσε να λύσει μόνο μια περιορισμένη κατηγορία προβλημάτων.

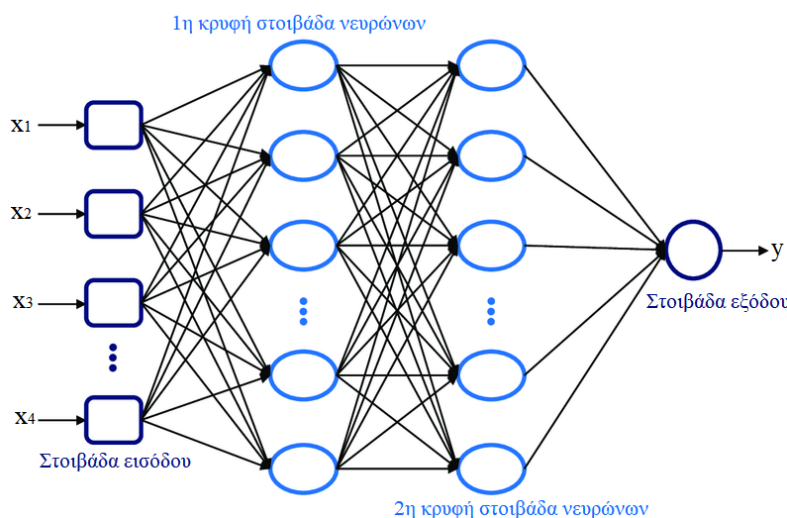
Την ίδια περίπου εποχή, ο Bernard Widrow και ο Ted Hoff εισήγαγαν [44] έναν νέο αλγόριθμο μάθησης και τον χρησιμοποίησαν για να εκπαιδεύσουν γραμμικά νευρωνικά δίκτυα, τα οποία ήταν παρόμοια σε δομή και ικανότητα με αυτά του Rosenblatt. Γενικότερα, ο αλγόριθμος μάθησης ορίζεται ως η διαδικασία κατά την οποία επιτυγχάνεται η μάθηση. Σκοπός του είναι η τροποποίηση των συναπτικών βαρών του νευρωνικού δίκτυο με κατάλληλο τρόπο ώστε να επιτευχθεί ο επιθυμητός στόχος. Ο κανόνας μάθησης Widrow - Hoff, λοιπόν, εξακολουθεί να χρησιμοποιείται έως σήμερα [44]. Κατά τη δεκαετία του 1980, δημοσιεύτηκε η αρχική εργασία σε συναρτήσεις ακτινικής βάσης από τον Powell και άλλους επιστήμονες [45]. Η πρωτοτυπία της έγκειται στο ότι οι συναρτήσεις ακτινικής βάσης χρησιμοποιήθηκαν για παρεμβολή σε έναν πολυδιάστατο χώρο. Η χρήση συναρτήσεων ακτινικής βάσης για παρεμβολή εξακολουθεί να είναι ένας σημαντικός και ενεργός τομέας έρευνας [46].

Ο αρχικός στόχος της προσέγγισης των νευρωνικών δικτύων ήταν να δημιουργηθεί ένα υπολογιστικό σύστημα που θα μπορούσε να λύσει προβλήματα όπως ο ανθρώπινος εγκέφαλος. Ωστόσο, με την πάροδο του χρόνου, οι ερευνητές μετατόπισαν το ενδιαφέρον τους στη χρήση νευρωνικών δικτύων για την αντιστοίχιση συγκεκριμένων εργασιών, οδηγώντας έτσι σε αποκλίσεις από μια αυστηρά βιολογική προσέγγιση. Έκτοτε, τα νευρωνικά δίκτυα έχουν ενταχθεί σε διάφορες εφαρμογές, όπως η υπολογιστική όραση [47], η αναγνώριση ομιλίας [48], η αυτόματη μετάφραση [49], το φιλτράρισμα κοινωνικών δικτύων, η ιατρική διάγνωση και η μοντελοποίηση εξαιρετικά ασταθών δεδομένων όπως αυτά των οικονομικών χρονοσειρών [50]. Επιπλέον, έχουν την δυνατότητα να μαθαίνουν να μοντελοποιούν τις σχέσεις μεταξύ εισόδων και εξόδων, κάνοντας γενικεύσεις και εξάγοντας συμπεράσματα. Εν κατακλείδι, τα νευρωνικά δίκτυα είναι ιδανικά στο να λύνουν πολύπλοκα προβλήματα σε πραγματικές καταστάσεις.

2.1 Perceptrons Πολλών Στοιβάδων

Το perceptron αποτελείται από έναν νευρώνα και είναι το απλούστερο είδος νευρωνικού δικτύου. Τα βασικά συστατικά ενός τεχνητού νευρώνα είναι ένα σύνολο συνάψεων, ο αθροιστής και μια συνάρτηση ενεργοποίησης.

Το Perceptron πολλών στοιβάδων (Multilayer Perceptron, MLP) είναι ο πιο γνωστός και συχνά χρησιμοποιούμενος τύπος νευρωνικού δικτύου. Αποτελείται από την στοιβάδα εισόδου και την στοιβάδα εξόδου. Όλα τα ενδιάμεσα στρώματα ονομάζονται κρυφές στοιβάδες. Τα επίπεδα είναι πλήρως διασυνδεδεμένα, πράγμα που σημαίνει ότι κάθε νευρώνας είναι συνδεδεμένος με κάθε έναν νευρώνα στο προηγούμενο και στο επόμενο επίπεδο. Το σήμα εισόδου διαδίδεται με κατεύθυνση προς τα εμπρός μέσα στο δίκτυο. Αυτή η αρχιτεκτονική ονομάζεται πρόσθιας τροφοδότησης [51]. Στο σχήμα 2.1 παρουσιάζεται η δομή ενός MLP δικτύου με δύο κρυφές στοιβάδες.



Σχήμα 2.1: Δομή MLP δικτύου με δύο κρυφές στοιβάδες.

Για την εκπαίδευση του δικτύου MLP χρησιμοποιείται ο αλγόριθμος οπισθοδιάδοσης (Back Propagation). Εισήχθη για πρώτη φορά τη δεκαετία του 1960 και σχεδόν 30 χρόνια αργότερα (1989) έγινε δημοφιλής από τους Rumelhart, Hinton και Williams [52]. Στο πρώτο βήμα του αλγορίθμου, οι συνάψεις και οι πολώσεις όλων των νευρώνων αρχικοποιούνται σε τυχαίες αρχικές τιμές. Στην συνέχεια προχωρώντας στοιβάδες από τα αριστερά προς τα δεξιά υπολογίζονται τα σήματα εξόδου όλων των νευρώνων και έπειτα τα συνολικά σήματα εξόδου και το σφάλμα σε σχέση με την έξοδο y . Τέλος, αλλάζουν οι τιμές των συνάψεων προχωρώντας στις στοιβάδες από τα δεξιά προς τα αριστερά.

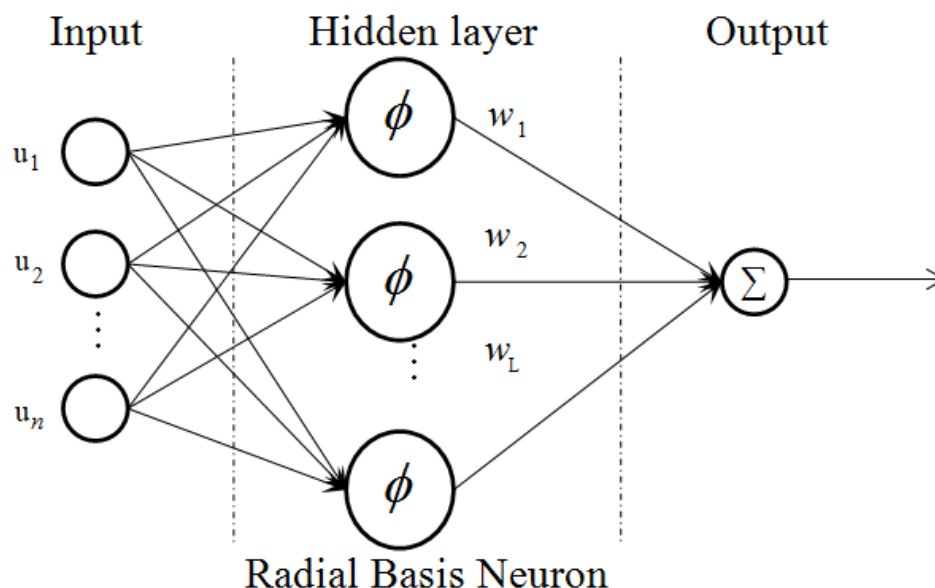
Με απλά λόγια κατά τον αλγόριθμο αυτό, μετά από κάθε διέλευση προς τα εμπρός μέσω του δικτύου, εκτελείται ένα πέρασμα (epoch) προς τα πίσω, ενώ γίνεται μια προσαρμογή των παραμέτρων του δικτύου (συνάψεις w και πολώσεις). Όταν η συνθήκη τερματισμού, η οποία μπορεί για παράδειγμα να αφορά την σύγκλιση των συνάψεων ή έναν μέγιστο αριθμό περασμάτων, γίνει αληθής, τότε ο αλγόριθμος τερματίζεται.

Σύμφωνα με τον κανόνα Δέλτα [53], ο αλγόριθμος back propagation κινεί τις συνάψεις προς τη διεύθυνση με την μέγιστη μεταβολή σφάλματος. Αυτό σημαίνει ότι η σύναψη κινείται προς την διεύθυνση που μειώνεται πιο απότομα η συνάρτηση του σφάλματος. Για αυτό τον λόγο μέθοδοι τέτοιου τύπου λέγονται μέθοδοι τάχιστης κατάβασης (steepest descent). Συνδυαζόμενη με άλλες μεθόδους βελτιστοποίησης η μέθοδος steepest descent μπορεί να αυξήσει την ταχύτητα σύγκλισης.

2.2 Νευρωνικά Δίκτυα Ακτινικής Βάσης

Τα νευρωνικά δίκτυα ακτινικής βάσης (RBF) εισήχθησαν στη βιβλιογραφία των νευρωνικών δικτύων από τους Broomhead και Lowe το 1988 [54]. Από την εμφάνισή τους μέχρι και σήμερα, τα δίκτυα RBF χρησιμοποιούνται κυρίως για προβλήματα προσέγγισης συναρτήσεων, αναγνώρισης προτύπων και πρόβλεψης χρονοσειρών. Η απλή δομή τους επιτρέπει τη μάθηση σταδιακά, μειώνει τον χρόνο εκπαίδευσης και αυτό έχει οδηγήσει στην χρήση τέτοιων δικτύων σε πολλές εφαρμογές όπως αυτές που ήδη αναφέρθηκαν [55].

Συγκεκριμένα, το δίκτυο RBF στην πιο απλή μορφή του είναι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης με τρία επίπεδα [56]. Για να χρησιμοποιηθεί ένα τέτοιο δίκτυο πρέπει να καθοριστεί η συνάρτηση ενεργοποίησης και ο αριθμός των νευρώνων της κρυφής στοιβάδας και ο αλγόριθμος εκπαίδευσης για την εύρεση των παραμέτρων του δικτύου. Όπως φαίνεται στο σχήμα 2.2, το δίκτυο αποτελείται από την στοιβάδα εισόδου, την κρυφή στοιβάδα και την στοιβάδα εξόδου η οποία έχει μόνο κόμβους άθροισης [57].



Σχήμα 2.2: Δομή RBF δικτύου με N μεταβλητές εισόδου και L κρυφούς κόμβους.

Η στοιβάδα εισόδου δεν είναι επίπεδο υπολογισμού, λαμβάνει απλώς τα δεδομένα εισόδου και τα τροφοδοτεί στην κρυφή στοιβάδα του δικτύου RBF. Ο υπολογισμός που γίνεται μέσα στην κρυφή στοιβάδα είναι πολύ διαφορετικός από τα περισσότερα νευρωνικά δίκτυα και από αυτόν προέρχεται η ισχύς του δικτύου RBF. Το επίπεδο εξόδου εκτελεί απλώς την διαδικασία της πρόβλεψης. Αξίζει επίσης να αναφερθεί ότι τα δίκτυα RBF έχουν γραμμική σχέση ανάμεσα στην κρυφή στοιβάδα και στην έξοδο [58]. Συγκεκριμένα για την κάθε στοιβάδα:

2.2.1 Στοιβάδα εισόδου

Στο επίπεδο εισόδου υπάρχει ένας νευρώνας για κάθε μια από τις μεταβλητές πρόβλεψης. Η στοιβάδα εισόδου διανέμει τις N μεταβλητές εισόδου στους L κόμβους της κρυφής στοιβάδας. Κάθε κόμβος στην κρυφή στοιβάδα συνδέεται με ένα κέντρο, ίσο σε διάσταση με τον αριθμό των διαστάσεων του προβλήματος (N). Στη στοιβάδα εισόδου δεν πραγματοποιούνται υπολογισμοί όπως συμβαίνει με τα τυπικά τεχνητά νευρωνικά δίκτυα. Οι νευρώνες εισόδου συνδέονται πλήρως με τους νευρώνες της κρυφής στοιβάδας και τροφοδοτούν την είσοδό τους προς τα εμπρός [59].

2.2.2 Κρυφή στοιβάδα

Η κρυφή στοιβάδα λαμβάνει την είσοδο, στην οποία το σύνολο των στοιχείων της μπορεί να μην είναι γραμμικά διαχωρίσιμο και το μετασχηματίζει σε ένα νέο χώρο που είναι πιο γραμμικά διαχωρίσιμος. Εκτελεί, δηλαδή, έναν μη γραμμικό μετασχηματισμό και χαρτογραφεί τον χώρο εισόδου σε έναν χώρο υψηλότερης διάστασης. Για αυτό τον λόγο και η κρυφή στοιβάδα έχει υψηλότερη διάσταση από το επίπεδο εισόδου. Αυτό βασίζεται στο θεώρημα του Cover για τη δυνατότητα διαχωρισμού των προτύπων, το οποίο δηλώνει ότι ένα μοτίβο που μετατρέπεται σε χώρο υψηλότερης διάστασης με μη γραμμικό μετασχηματισμό είναι πιο πιθανό να διαχωρίζεται γραμμικά [60]. Επομένως, ο αριθμός των νευρώνων στην κρυφή στοιβάδα θα πρέπει να είναι μεγαλύτερος από τον αριθμό των νευρώνων εισόδου.

Η δραστηριότητα $\mu_L(\mathbf{u}(k))$ του L κόμβου υπολογίζεται από την ευκλείδεια διαφορά του k διανύσματος εισόδου και του κέντρου του κόμβου. Δίνεται από την σχέση [61]:

$$\mu_L(\mathbf{u}(k)) = \|\mathbf{u}(k) - \hat{u}_l\| = \sqrt{\sum_{i=1}^N (u_i(k) - \hat{u}_{i,l})^2}, k=1, \dots, K \quad (2.1)$$

Όπου το K είναι ο συνολικός αριθμός των δεδομένων του διανύσματος εισόδου και το \hat{u}_l είναι το κέντρο του l κόμβου.

Η συνάρτηση ενεργοποίησης, ή αλλιώς αποκαλούμενη και ως συνάρτηση σύνθλιψης, χρησιμοποιείται για να περιορίσει το επιτρεπόμενο πλάτος του σήματος εξόδου ανάμεσα σε κάποιες τιμές. Στα δίκτυα RBF, η συνάρτηση ενεργοποίησης για κάθε κόμβο είναι μια συνάρτηση ακτινικής συμμετρίας και δίνεται από τον τύπο [62]:

$$g(\mu) = \mu^2 \log(\mu) \quad (2.2)$$

Η έξοδος του κρυφού νευρώνα $\mathbf{z}(k)$ ισούται με:

$$\mathbf{z}(k) = [g(\mu_1(\mathbf{u}(k))), g(\mu_2(\mathbf{u}(k))), \dots, g(\mu_L(\mathbf{u}(k)))] \quad (2.3)$$

Αυτό που αξίζει να αναφερθεί είναι ότι η διάσταση του πίνακα της εξόδου του κρυφού νευρώνα $\mathbf{z}(k)$ έχει διάσταση $N \times L$ όπου L είναι ο αριθμός των κόμβων και N ο αριθμός των μεταβλητών εισόδου.

2.2.3 Στοιβάδα εξόδου

Η στοιβάδα εξόδου χρησιμοποιεί μια γραμμική συνάρτηση ενεργοποίησης. Οι υπολογισμοί στο επίπεδο εξόδου εκτελούνται ακριβώς όπως και σε ένα τυπικό τεχνητό νευρωνικό δίκτυο το οποίο είναι ένας γραμμικός συνδυασμός μεταξύ του διανύσματος εισόδου και του διανύσματος βάρους.

Η τελική έξοδος του νευρωνικού δικτύου $\hat{y}(k)$ υπολογίζεται γραμμικά ως εξής [63]:

$$\hat{y}(k) = \mathbf{z}(k) \cdot \mathbf{w} = \sum_{l=1}^L w_l g(\mu_l(\mathbf{u}(k))) \quad (2.4)$$

Όπου w είναι ένας πίνακας που περιέχει τα συναπτικά βάρη.

Τα συναπτικά βάρη υπολογίζονται τυπικά χρησιμοποιώντας την γραμμική παλινδρόμηση των εξόδων του κρυφού στρώματος στις πραγματικές μετρούμενες εξόδους. Το πρόβλημα της παλινδρόμησης μπορεί να λυθεί χρησιμοποιώντας την μέθοδο των ελαχίστων τετραγώνων:

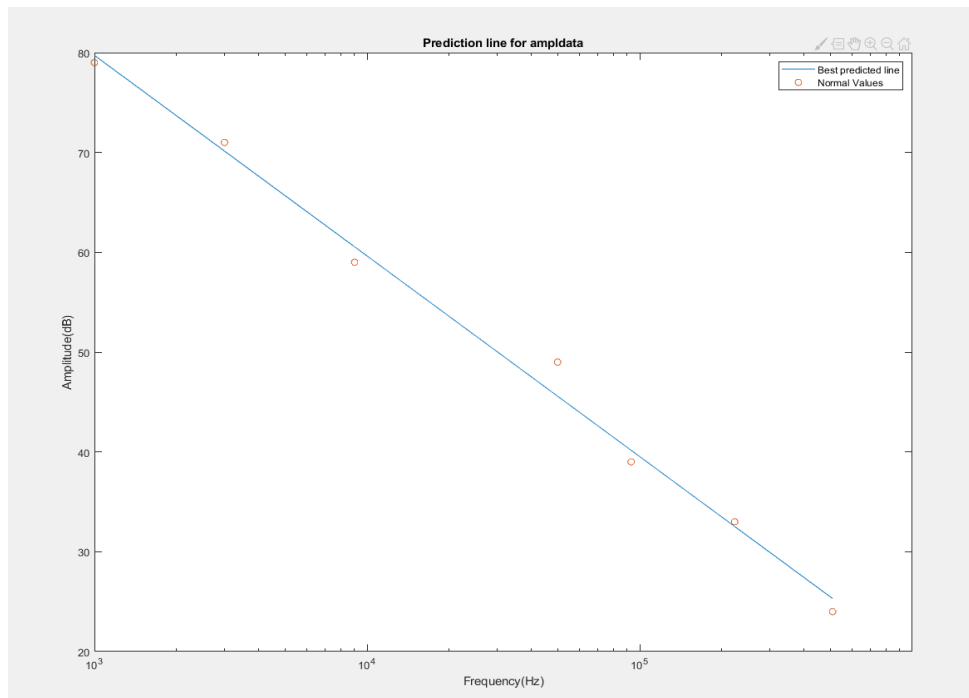
$$\mathbf{w}^T = \mathbf{Y}^T \cdot \mathbf{Z} \cdot (\mathbf{Z}^T \cdot \mathbf{Z})^{-1} \quad (2.5)$$

Μέθοδος των ελαχίστων μέσων τετραγώνων:

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Ο αλγόριθμος των ελαχίστων μέσων τετραγώνων (Least Mean Square) αποτελεί τον πρώτο γραμμικό αλγόριθμο προσαρμοστικού φιλτραρίσματος και ανακαλύφθηκε από τους Widrow και Hoff το 1960 [44]. Πιο συγκεκριμένα, είναι μια μορφή μαθηματικής ανάλυσης παλινδρόμησης που χρησιμοποιείται για τον προσδιορισμό της καλύτερης ευθείας που περνάει από ένα σύνολο δεδομένων, παρέχοντας έτσι μια οπτική επίδειξη της σχέσης μεταξύ των δεδομένων. Κάθε σημείο αντιπροσωπεύει τη σχέση μεταξύ μιας γνωστής ανεξάρτητης μεταβλητής και μιας άγνωστης εξαρτημένης μεταβλητής. Για να γίνει αντιληπτή η μέθοδος αυτή θα δοθεί το εξής απλό παράδειγμα σε δύο διαστάσεις:

Θεωρούμε ότι υπάρχει ένα σύνολο δειγμάτων $\{x_i, y_i\}$. Σκοπός είναι να βρεθεί ένα μαθηματικό μοντέλο, το οποίο να περιγράφει την σχέση αυτών των δύο μεταβλητών. Στην συγκεκριμένη περίπτωση, το μοντέλο που αναζητείται είναι η ευθεία γραμμή της μορφής $y = \alpha x + \beta$, η οποία "ταιριάζει" καλύτερα στο σύνολο των δειγμάτων. Έχοντας αυτό το μοντέλο υπάρχει η δυνατότητα πρόβλεψης των τιμών του y για νέες τιμές του x . Έστω ότι τα πειραματικά δεδομένα συμβολίζονται με τον κύκλο. Μέσω της μεθόδου των ελαχίστων τετραγώνων γίνεται μια προσπάθεια προσδιορισμού της καλύτερης ευθείας που ελαχιστοποιεί τις αποστάσεις των δεδομένων από αυτή. Στο σχήμα 2.3 φαίνεται το παράδειγμα που περιγράφηκε παραπάνω:



Σχήμα 2.3: Παράδειγμα χρήσης της μεθόδου των ελαχίστων τετραγώνων

Όπως μπορεί πλέον να γίνει αντιληπτό, η σχέση 2.5 έχει προκύψει από την λύση της εξίσωσης $\mathbf{Y} = \mathbf{Z} \cdot \mathbf{W}$. Για να συμβεί αυτό πρέπει να απομονωθεί το \mathbf{W} στην δεξιά πλευρά. Με τον πολλαπλασιασμό των δύο μελών της εξίσωσης με τον ανάστροφο του \mathbf{Z} , δημιουργείται ένας τετραγωνικός πίνακας ο οποίος είναι αντιστρέψιμος. Πολλαπλασιάζοντας, λοιπόν, και τα δύο μέλη με τον αντίστροφο του πίνακα $(\mathbf{Z}^T \cdot \mathbf{Z})$, προκύπτει ο πίνακας \mathbf{w} ο οποίος περιέχει τα συναπτικά βάρη. Ο πίνακας $(\mathbf{Z}^T \cdot \mathbf{Z})$ ονομάζεται πίνακας συνδιασποράς (covariance matrix), και περιέχει πληροφορίες για τον τρόπο που είναι κατανεμημένες οι είσοδοι του συστήματος.

2.3 Συναρτήσεις ακτινικής βάσης

Μια συνάρτηση $f(x)$ ονομάζεται ακτινικού τύπου (radial function) αν υπάρχει κάποιο κέντρο και η τιμή της συνάρτησης αυτής εξαρτάται μόνο από την απόσταση του x από το κέντρο.

Πρόκειται για μια συνάρτηση η οποία έχει τοπικό χαρακτήρα κάτι το οποίο σημαίνει πως έχει μόνο ένα μέγιστο, στο κέντρο. Εάν x είναι το διάνυσμα εισόδου και c είναι το διάνυσμα κέντρου τότε η έξοδος ορίζεται ως συνάρτηση της απόστασης του x από το κέντρο c και δίνεται από τον τύπο:

$$f(x) = f(\|x - c\|) \quad (2.6)$$

Όπως έχει ήδη αναφερθεί παραπάνω, οι συναρτήσεις αυτές χρησιμοποιούνται στους νευρώνες της κρυφής στοιβάδας. Η έξοδος των νευρώνων στην ουσία απεικονίζει την πιο κοντινή τιμή του διανύσματος σε σχέση με το διάνυσμα εισόδου u_i . Στα νευρωνικά δίκτυα ακτινικής βάσης υπάρχουν φορές που συναντάμε συναρτήσεις οι οποίες εκτός από τις τιμές των κέντρων χαρακτηρίζονται και από το πλάτος σ . Παρακάτω θα αναφερθούν ορισμένα κοινά παραδείγματα συναρτήσεων ακτινικής βάσης:

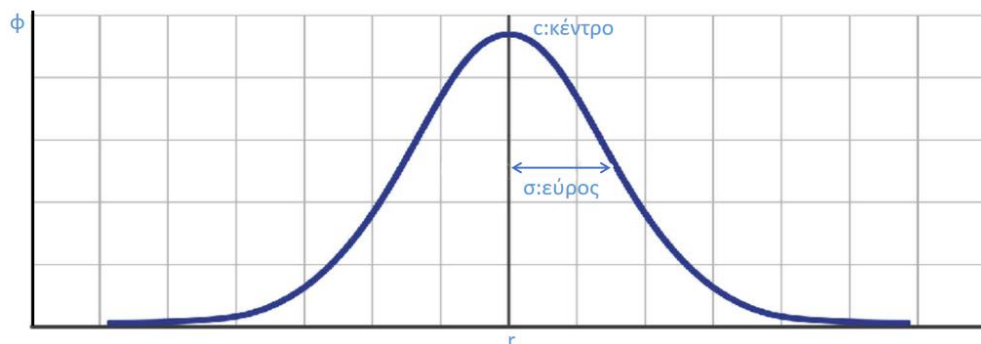
2.3.1 Παραδείγματα συναρτήσεων ακτινικής βάσης

Οι ακτινικές συναρτήσεις είναι μια ειδική κατηγορία συναρτήσεων. Το βασικό τους χαρακτηριστικό είναι ότι η απόκρισή τους μειώνεται ή αυξάνεται μονότονα με την απόσταση από ένα κεντρικό σημείο.

Η πιο κοινή συνάρτηση βάσης είναι η συνάρτηση Gauss η οποία περιγράφεται από τον τύπο:

$$\varphi(r) = e^{-\frac{r^2}{2\sigma^2}} \quad (2.7)$$

Όπου σ είναι το εύρος και c το κέντρο. Γνωρίζουμε ότι για την συνάρτησή Gauss η οποία έχει τοπικό χαρακτήρα ισχύει ότι όταν το φ τείνει στο μηδέν η απόλυτη τιμή του r τείνει στο άπειρο ($\varphi \rightarrow 0, |r| \rightarrow \infty$). Στο σχήμα 2.4 αναπαρίσταται η συνάρτηση Gauss:

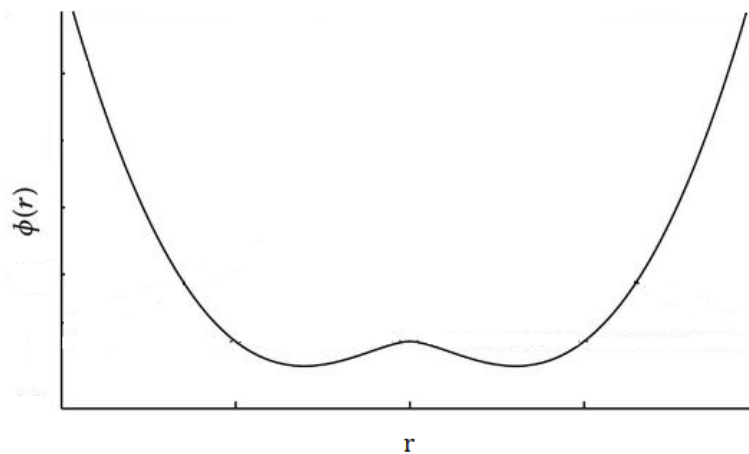


Σχήμα 2.4: Συνάρτηση Gauss

Ένα άλλο παράδειγμα ακτινικής συνάρτησης βάσης είναι η συνάρτηση TPS (Thin Plate Spline) η οποία περιγράφεται από τον τύπο:

$$\varphi(r) = r^2 \log r \quad (2.8)$$

Στο σχήμα 2.5 ενδεικτικά παρουσιάζεται η γραφική παράσταση:



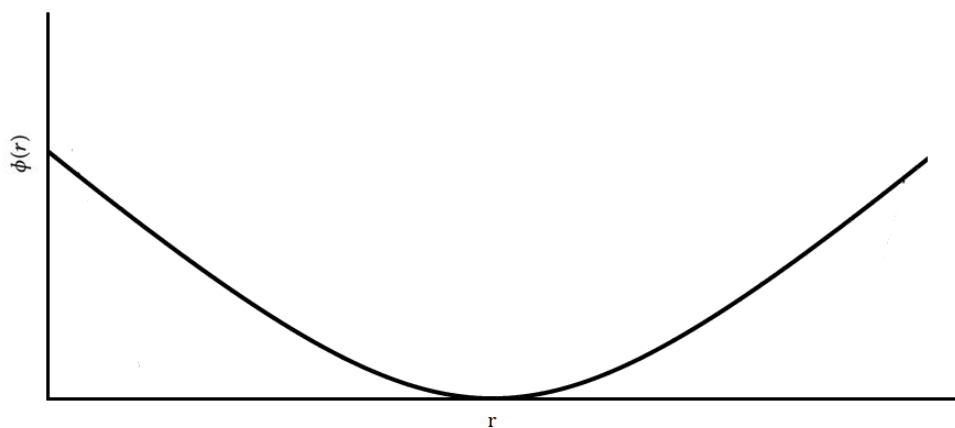
Σχήμα 2.5: Συνάρτηση TPS

Η συνάρτηση TPS χρησιμοποιείται κυρίως σε περιπτώσεις που θέλουμε να κάνουμε εξομάλυνση και παρεμβολή. Με τον όρο παρεμβολή αναφερόμαστε σε προβλέψεις δεδομένων οι οποίες είναι εντός των ορίων εκπαίδευσης. Είναι γνωστό ότι για την συνάρτησή TPS, η οποία έχει μη τοπικό χαρακτήρα, ισχύει ότι όταν το ϕ τείνει στο άπειρο η απόλυτη τιμή του r τείνει επίσης στο άπειρο ($\phi \rightarrow \infty, |r| \rightarrow \infty$).

Τέλος, υπάρχει και η πολυτετραγωνική (multiquadratic) συνάρτηση η οποία επίσης έχει μη τοπικό χαρακτήρα και περιγράφεται από την σχέση:

$$\phi(r) = \sqrt{r^2 + \sigma^2} \quad (2.9)$$

Στο σχήμα 2.6 ενδεικτικά παρουσιάζεται η γραφική παράσταση της πολυτετραγωνικής συνάρτησης:



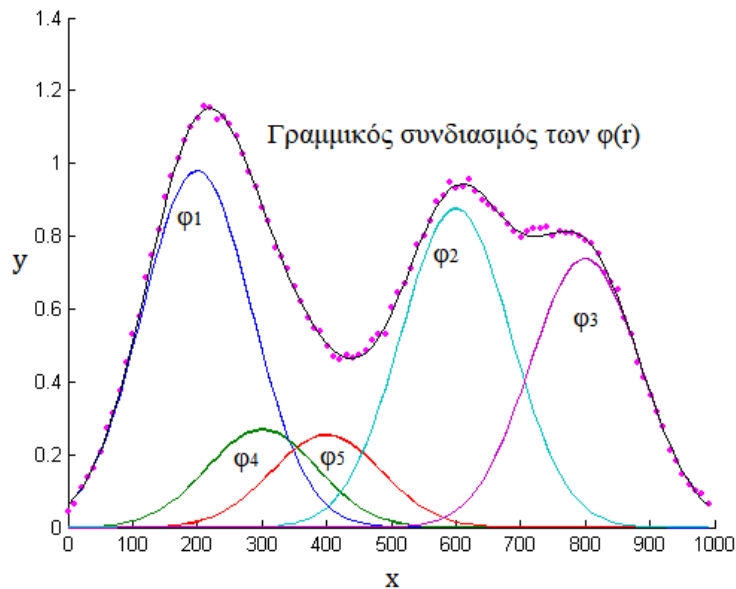
Σχήμα 2.6: Πολυτετραγωνική συνάρτηση

2.4 Λειτουργία δικτύων RBF

Η κεντρική ιδέα της λειτουργίας των δικτύων RBF βασίζεται στο ότι μια προβλεπόμενη τιμή στόχου ενός στοιχείου είναι πιθανό να είναι περίπου ίδια με άλλα στοιχεία τα οποία έχουν κοντινές τιμές σε αυτές των μεταβλητών πρόβλεψης. Τα δίκτυα RBF χρησιμοποιούν μια προσέγγιση τοπολογικού χαρακτήρα καθώς λαμβάνεται υπ'όψιν η τοπολογία του χώρου των

μεταβλητών εισόδου. Για να γίνει αντιληπτή η προσέγγισή τοπολογικού χαρακτήρα παρατίθεται το παρακάτω παράδειγμα:

Έστω ότι ένα σύστημα έχει μεταβλητές εισόδου x και μεταβλητές εξόδου y . Από την παρατήρηση των δεδομένων εκπαίδευση στο χώρο εισόδου γίνεται αντιληπτό ότι στην τοπολογία του χώρου εισόδου έχουν δημιουργηθεί πέντε ομάδες (clusters) δεδομένων. Στο κέντρο κάθε ομάδας μπορεί να τοποθετηθεί μια συναρτήσεις ακτινικής βάσης η οποία αν σταθμιστεί μεταβάλλεται το ύψος της. Το άθροισμα όλων των σταθμισμένων συναρτήσεων ακτινικής βάσης, όπως φαίνεται και στο σχήμα 2.7, δίνει την πρόβλεψη του δικτύου RBF.



Σχήμα 2.7: Πρόβλεψη του δικτύου RBF

Στο σχήμα 2.7, οι συναρτήσεις ακτινικής βάσης έχουν διαφορετικά χρώματα πράγμα το οποίο συμβολίζει ότι έχουν διαφορετικά βάρη. Όταν αυτές αθροιστούν παράγεται μια συνεχή συνάρτηση. Για να συμβεί όμως αυτό, πρέπει να είναι γνωστό το που θα τοποθετηθούν τα κέντρα των συναρτήσεων ακτινικής βάσης c και το εύρος σ τους. Για τον σκοπό αυτό χρησιμοποιούνται αλγόριθμοι όπως αυτός του k -means και του Fuzzy Means. Τέλος, αξίζει να αναφερθεί ότι στα δίκτυα RBF η γνώση δεν αποθηκεύεται μόνο στις συνάψεις, αλλά και στις παραμέτρους της συνάρτησης ενεργοποίησης [64].

2.5 Εκπαίδευση RBF δικτύου

Οι πιο συχνά χρησιμοποιούμενοι αλγόριθμοι εκπαίδευσης των δικτύων RBF έχουν δύο στάδια, τα οποία αντιμετωπίζουν τα δύο επίπεδα του δικτύου RBF χωριστά. Η εκπαίδευση ενός δικτύου ακτινικής βάσης σχετίζεται με τον υπολογισμό των ακόλουθων παραμέτρων [65]:

- i. Συντεταγμένες των κέντρων.
- ii. Συναπτικά βάρη.

Λόγω της δομής του δικτύου, ο προσδιορισμός των παραμέτρων περιλαμβάνει δύο φάσεις. Στην πρώτη λαμβάνονται τα κέντρα των κόμβων του κρυφού στρώματος, με μια μέθοδο μη εποπτευόμενης μάθησης, ενώ στην δεύτερη φάση τα συναπτικά βάρη υπολογίζονται μέσω της μεθόδου των ελαχίστων τετραγώνων.

Συγκεκριμένα, για τον καθορισμό των κέντρων της κρυφής στιβάδας λαμβάνονται υπ' όψιν μόνο τα δεδομένα εισόδου και συνήθως ο καθορισμός τους αντιμετωπίζεται ως πρόβλημα ομαδοποίησης. Κατά την δεύτερη φάση παρατηρείται ότι η έξοδος του δικτύου είναι το σταθμισμένο άθροισμα των εξόδων της κρυφής στιβάδας, πράγμα το οποίο δηλώνει ότι η

σχέση ανάμεσα στην έξοδο του δικτύου και στην έξοδο των νευρώνων της κρυφής στοιβάδας είναι γραμμική [66].

2.5.1 Κανονικοποίηση δεδομένων

Η κανονικοποίηση είναι μια τεχνική προετοιμασίας των δεδομένων που χρησιμοποιείται συχνά στη μηχανική μάθηση και ορίζεται ως η διαδικασία μετατροπής των δεδομένων στην ίδια κλίμακα. Απαιτείται μόνο όταν το εύρος των χαρακτηριστικών είναι διαφορετικό. Σε πολλές περιπτώσεις οι μεταβλητές εισόδου είναι πιθανόν να είναι σε διαφορετικές κλίμακες. Κάτι τέτοιο είναι πιθανόν να δημιουργήσει προβλήματα κατά την διάρκεια της εκπαίδευσης του νευρωνικού δικτύου. Για αυτό και κρίνεται αναγκαία η κανονικοποίηση των μεταβλητών εισόδου.

Μια διαδεδομένη μέθοδος για τον σκοπό αυτό είναι η κανονικοποίηση σε συγκεκριμένο εύρος τιμών με γραμμικό μετασχηματισμό. Αν x'_{min} είναι η ελάχιστη τιμή του ορίου και x'_{max} η μέγιστη, αυτό που επιτυγχάνεται με την μέθοδο αυτή είναι οι μεταβλητές να έρθουν ανάμεσα στα ίδια όρια. Η παρακάτω συνάρτηση δείχνει τον τρόπο κανονικοποίησης της μεταβλητή x με την συγκεκριμένη μέθοδο:

$$x' = x'_{min} + \frac{(x'_{max} - x'_{min})(x - x_{min})}{(x_{max} - x_{min})} \quad (2.10)$$

Όπου:

x' : η κανονικοποιημένη τιμή της μεταβλητής.

x'_{min} : η ελάχιστη τιμή του ορίου.

x'_{max} : η μέγιστη τιμή του ορίου.

x : η τιμή της μεταβλητής.

Ειδικά για $x'_{min} = 1$ και για $x'_{min} = -1$ η σχέση (2.10) μετατρέπεται:

$$x' = -1 + \frac{2(x - x_{min})}{(x_{max} - x_{min})} \quad (2.11)$$

Τέλος, αυτό που αξίζει να αναφερθεί είναι πως για την αποκανονικοποίηση της εξόδου θα πρέπει να έχει κρατηθεί η ελάχιστη και η μέγιστη τιμή για κάθε μεταβλητή.

2.5.2 Συντεταγμένες κέντρων

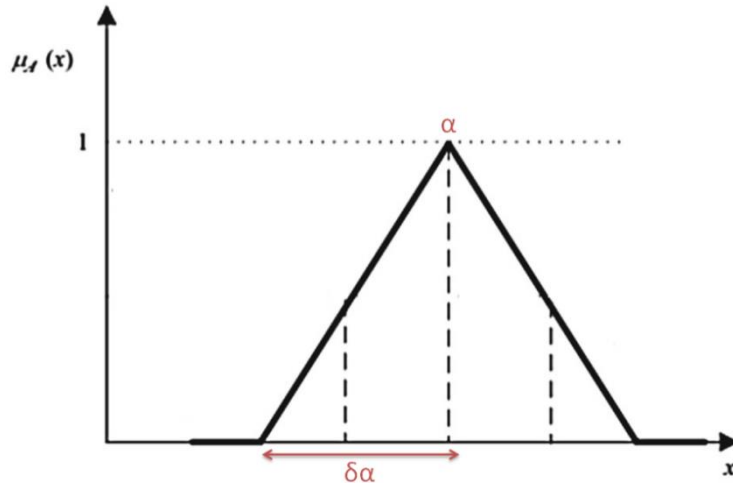
Όπως ήδη έχει αναφερθεί, κατά την άθροιση των συναρτήσεων ακτινικής βάσης πρέπει να είναι γνωστό το που θα τοποθετηθούν τα κέντρα. Για το υπολογισμό αυτό χρησιμοποιήθηκε ο αλγόριθμος των ασαφών μέσων (Fuzzy Means). Ο λόγος που εφαρμόστηκε η συγκεκριμένη μέθοδος προσδιορισμού των κέντρων, είναι διότι πρόκειται να αναλυθούν συστήματα με μεγάλη αβεβαιότητα. Όλα τα ασαφή συστήματά βασίζονται στο ότι γίνεται μια διαμέριση του χώρου των μεταβλητών εισόδου σε ένα αριθμό ασαφών συνόλων. Στην συγκεκριμένη μεθοδολογία ο χώρος κάθε μεταβλητής εισόδου u_i ($i = 1, 2, \dots, N$) χωρίζεται σε τριγωνικά ασαφή σύνολα.

Ένα ασαφές σύνολο είναι μια κατηγορία αντικειμένων με μια συνέχεια βαθμών συμμετοχής. Χαρακτηρίζεται από μια συνάρτηση συμμετοχής μ_A (membership function) η οποία εκχωρεί σε κάθε αντικείμενο έναν βαθμό ιδιότητας μέλους που κυμαίνεται μεταξύ του μηδενός και του ένα. Θεωρώντας N κανονικοποιημένες μεταβλητές εισόδου u_i , όπου $i = 1, \dots, N$, κάθε ασαφές σύνολο μπορεί να γραφτεί ως [7]:

$$A_{i,j} = \{a_{i,j}, \delta a_i\}, i = 1, \dots, N, j = 1, \dots, s_i \quad (2.11)$$

Όπου $a_{i,j}$ είναι το κέντρο του συνόλου στο οποίο η τιμή της συνάρτησης συμμετοχής είναι ίση με την μονάδα και δa το μισό του εύρους. Το δa επιλέγεται ώστε οι δυο κορυφές του τριγώνου να βρίσκονται ακριβώς στα κέντρα των δυο γειτονικών ασαφών συνόλων.

Ένα παράδειγμα ασαφούς συνόλου απεικονίζεται στο σχήμα 2.8.

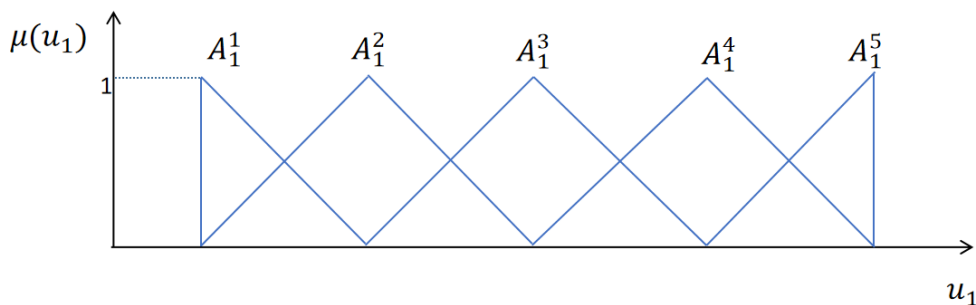


Σχήμα 2.8: Ασαφές σύνολο

Ο ασαφής διαχωρισμός (fuzzy partitioning) ορίζεται από τον διαχωρισμό του χώρου μιας μεταβλητής και περιγράφεται όπως φαίνεται παρακάτω:

$$T_i = \{A_i^1, A_i^2, \dots, A_i^s\}, 1 \leq i \leq N \quad (2.12)$$

Έστω ότι για την πρώτη μεταβλητή εισόδου u_1 γίνεται η επιλογή 5 ασαφών υποχώρων ($s = 5$). Ο ασαφής διαχωρισμός του χώρου της μεταβλητής διαμορφώνεται όπως απεικονίζεται στο σχήμα 2.9:



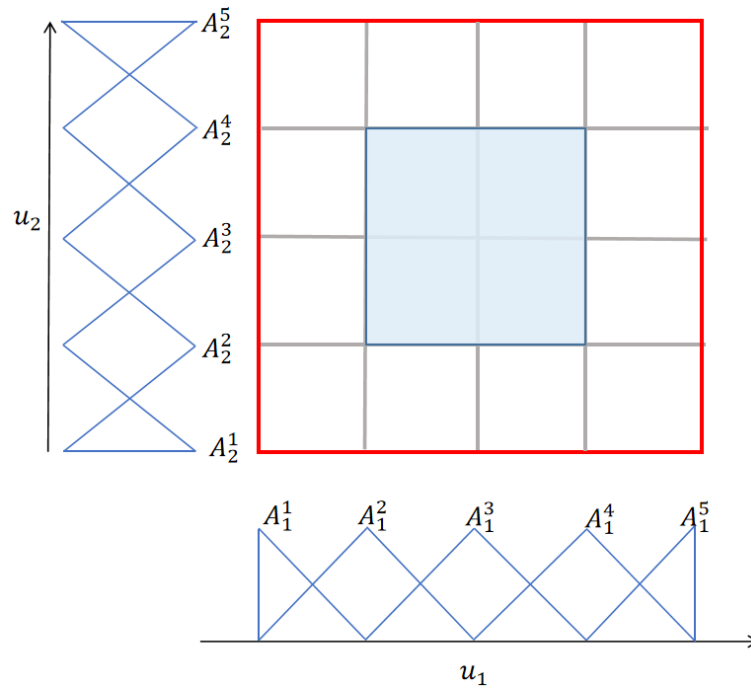
Σχήμα 2.9: Ασαφής διαχωρισμός του χώρου μιας μεταβλητής

Αυτή η τεχνική διαμέρισης δημιουργεί ένα σύνολο από S πολυδιάστατους ασαφής υποχώρους A^l (fuzzy subspaces), όπου $l=1, \dots, S$. Το σύνολο αυτό υπολογίζεται με την βοήθεια της εξίσωσης:

$$S = \prod_{i=1}^N s_i \quad (2.13)$$

Η μεταβλητή S στην ουσία απεικονίζει πως ο χώρος εισόδου έχει χωριστεί σε S ασαφής υποχώρους.

Στο σχήμα 2.10 απεικονίζεται ο διαχωρισμός ενός δισδιάστατου χώρου σε ασαφείς υποχώρους. Σε κάθε διάσταση ο χώρος της κάθε μεταβλητής είναι χωρισμένος σε 5 ασαφή σύνολα με αποτέλεσμα να έχουμε συνολικά $S=25$ ασαφείς υποχώρους.



Σχήμα 2.10: Διαχωρισμός δισδιάστατου χώρου εισόδου με ασαφείς υποχώρους

Ο στόχος του αλγορίθμου FM είναι να δημιουργήσει το κρυφό στρώμα του δικτύου RBF επιλέγοντας μόνο ένα μικρό υποσύνολο των ασαφών υποχώρων. Ο αριθμός των επιλεγμένων ασαφών υποχώρων θα πρέπει να διατηρείται σε χαμηλά επίπεδα, αλλά ταυτόχρονα τα παραγόμενα κέντρα RBF θα πρέπει να καλύπτουν επαρκώς τα διαθέσιμα δεδομένα εισόδου. Η επιλογή αυτή εκτελείται με βάση την πολυδιάστατη συνάρτηση συμμετοχής $\mu_{A^l}(\mathbf{u}(k))$ ενός διανύσματος εισόδου $\mathbf{u}(k)$ σε έναν ασαφές υποχώρο A^l . Η συνάρτηση συμμετοχής μ_{A^l} υπολογίζεται ως εξής [67]:

$$\mu_{A^l}(\mathbf{u}(k)) = \begin{cases} 1 - d_r^l(\mathbf{u}(k)), & \text{αν } d_r^l(\mathbf{u}(k)) \leq 1 \\ 0, & \text{αλλιώς} \end{cases} \quad (2.14)$$

Όπου $d_r^l(\mathbf{u}(k))$ είναι συνάρτηση της απόστασης μεταξύ του ασαφούς υποχώρου A^l και των δεδομένων εισόδου $\mathbf{u}(k)$.

Η συνάρτηση $d_r^l(\mathbf{u}(k))$ ορίζει μια υπερεπιφάνεια στον χώρο εισόδου. Αυτή η υπερεπιφάνεια σηματοδοτεί το όριο μεταξύ των διανυσμάτων εισόδου που λαμβάνουν μη μηδενικούς ή μηδενικούς βαθμούς συμμετοχής στον ασαφή υποχώρο A^l . Στην περίπτωση που έχουμε ασύμμετρη κατάτμηση του χώρου εισόδου η συνάρτηση $d_r^l(\mathbf{u}(k))$ υπολογίζεται ως εξής [68]:

$$d_r^l(\mathbf{u}(x)) = \sqrt{\sum_{i=1}^N \frac{(a_{i,j_i}^l - u_i(k))^2}{N(\delta \alpha_i)^2}} \quad (2.15)$$

Ενώ, στην περίπτωση συμμετρικής κατάτμησης του χώρου εισόδου η συνάρτηση $d_r^l(\mathbf{u}(k))$ υπολογίζεται ως εξής:

$$d_r^l(\mathbf{u}(x)) = \sqrt{\frac{\sum_{i=1}^N (a_{i,j_i}^l - u_i(k))^2}{\sqrt{N}\delta\alpha}} \quad (2.16)$$

Έχοντας ορίσει τη συνάρτηση συμμετοχής, ο αλγόριθμος προχωρά στην εύρεση του υποσυνόλου των ασαφών υποχώρων που εκχωρούν έναν μη μηδενικό πολυδιάστατο βαθμό σε όλα τα διάνυσματα εισόδου. Είναι προφανές πως ο ασαφής υποχώρος ο οποίος περιγράφει με τον καλύτερο τρόπο ένα διάνυσμα εισόδου είναι αυτός που έχει την μικρότερη ευκλείδεια απόσταση και άρα τον μεγαλύτερο βαθμό συμμετοχής. Πιο συγκεκριμένα, όσο πιο κοντά βρίσκεται το δεδομένο εισόδου σε ένα ασαφές σύνολο τόσο μεγαλύτερος θα είναι και ο βαθμός συμμετοχής. Η μέγιστη τιμή που μπορεί να πάρει ο βαθμός συμμετοχής είναι ίση με την μονάδα. Στην αντίθετη περίπτωση, όσο πιο μακριά βρίσκεται το δεδομένο από το ασαφές σύνολο τόσο πιο πολύ θα τείνει ο βαθμός συμμετοχής στην τιμή μηδέν, η οποία είναι και η ελάχιστη τιμή που μπορεί να πάρει. Ο ασαφής υποχώρος που περιγράφει με τον καλύτερο τρόπο το διάνυσμα εισόδου προκύπτει όταν ολοκληρωθούν οι παραπάνω υπολογισμοί σε όλες τις διαστάσεις.

Αυτό που πρέπει να αναφερθεί είναι πως οι ασαφείς υποχώροι που έχουν επιλεγεί θα πρέπει να καλύπτουν ικανοποιητικά το διάνυσμα εισόδου. Αν αυτό δεν συμβαίνει τότε εισάγεται νέος υποχώρος. Ο τρόπος με τον οποίο ελέγχουμε αν ο ασαφής υποχώρος καλύπτει ένα διάνυσμα εισόδου είναι μέσω της συνάρτησης $d_r^l(\mathbf{u}(k))$. Ειδικότερα, αν η ευκλείδεια απόσταση είναι μεγαλύτερη της μονάδας αυτό σημαίνει ότι ο ασαφής υποχώρος δεν καλύπτει το διάνυσμα είσοδο.

Η διαδικασία που περιγράφεται παραπάνω επιτυγχάνεται χρησιμοποιώντας έναν μη επαναλαμβανόμενο αλγόριθμο που απαιτεί μόνο ένα πέρασμα των δεδομένων εισόδου, καθιστώντας έτσι τη διαδικασία υπολογισμού των κέντρων εξαιρετικά γρήγορη. Ο αλγόριθμος αυτός παρουσιάζεται παρακάτω [69]:

2.5.3 Αλγόριθμος Fuzzy Means

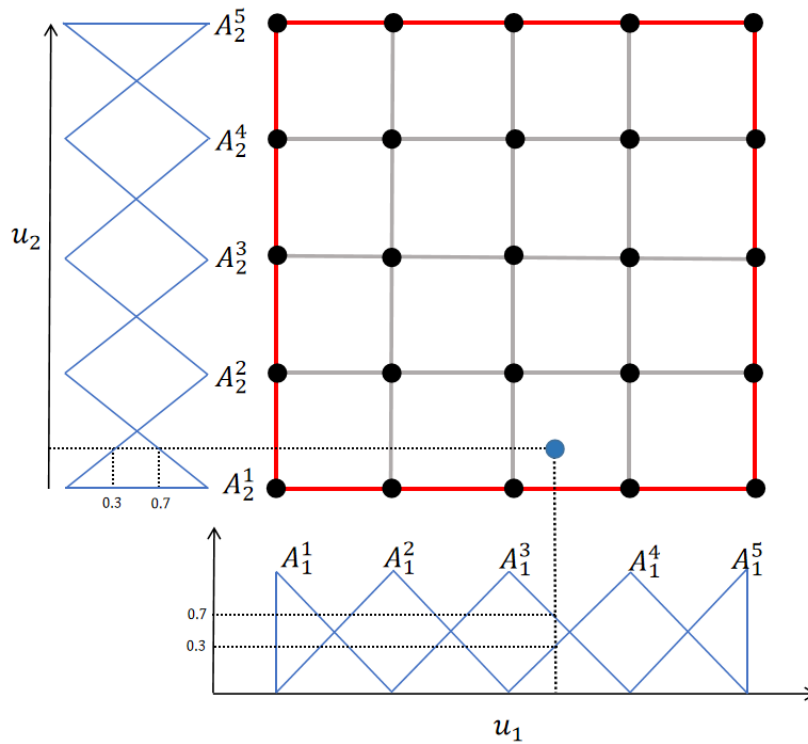
Για να κατανοήσουμε καλύτερα τον αλγόριθμο Fuzzy Means θα περιγράψουμε την διαδικασία μέσω ενός παραδείγματος. Αρχικά, είναι αναγκαίο να κανονικοποιήσουμε τις μεταβλητές εισόδου N και έπειτα να διαχωριστούν τα πεδία τιμών σε s ασαφή σύνολα. Στα σύνολα αυτά αποθηκεύονται οι τιμές του κέντρου a και το εύρος δa .

Ας πάρουμε ως παράδειγμα τον χώρο του σχήματος 2.10. Για ένα διάνυσμα εισόδου τα ασαφή σύνολα για τις αντίστοιχες μεταβλητές u_1 και u_2 είναι:

$$T_1 = \{A_1^1, A_1^2, \dots, A_1^5\} \quad (2.17)$$

$$T_2 = \{A_2^1, A_2^2, \dots, A_2^5\} \quad (2.18)$$

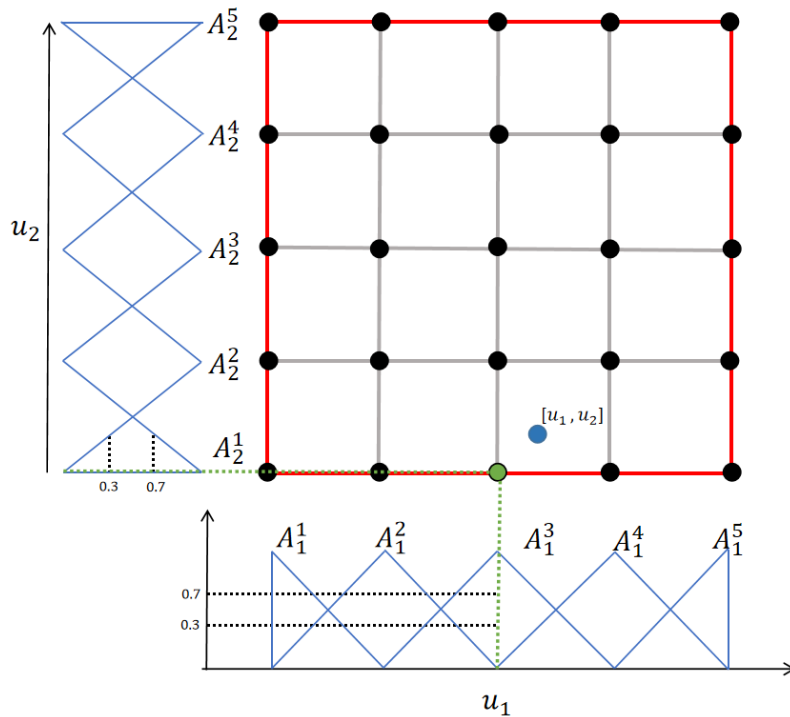
Αρχικά θα πρέπει να καθοριστούν οι βαθμοί συμμετοχής των u_1, u_2, \dots, u_N στα ασαφή σύνολα που τους αντιστοιχούν. Όπως φαίνεται στο σχήμα 2.11, ο βαθμός συμμετοχής του u_1 στο A_1^3 είναι 0.7 και στο A_1^4 είναι 0.3, ενώ στα υπόλοιπα ασαφή σύνολα είναι μηδενικός. Με την ίδια λογική ο βαθμός συμμετοχής του u_2 στο A_2^2 είναι 0.7 και στο A_2^2 είναι 0.3, ενώ στα υπόλοιπα ασαφή σύνολα είναι μηδενικός.



Σχήμα 2.11: Βαθμός συμμετοχής του $u = [u_1 \ u_2]$

Στην συνέχεια με βάση τον υπολογισμό των βαθμών συμμετοχής που έγινε παραπάνω ανατίθεται στην μεταβλητή u_1 το ασαφές σύνολο A_1^3 και στην u_2 το A_2^1 . Επιλέχθηκε, δηλαδή, το ασαφές σύνολο με τον μεγαλύτερο βαθμό συμμετοχής.

Στο σχήμα 2.12 παρουσιάζεται για το διάνυσμα εισόδου $u = [u_1 \ u_2]$ ο κοντινότερος ασαφής υποχώρος ο οποίος δίνεται ως συνδυασμός των ασαφών συνόλων τα οποία επιλέχθηκαν με βάση τον βαθμό συμμετοχής.



Σχήμα 2.12: Επιλογή του κοντινότερου ασαφούς υποχώρου

Με βάση την σχέση (2.12), ο ασαφής υποχώρος μπορεί να γραφτεί ως:

$$A_{i,j} = \{[\alpha_1 \alpha_2], [\delta\alpha_1 \delta\alpha_2]\} \quad (2.19)$$

Με τρόπο ανάλυσης που αναφέρθηκε, τόσο με την βοήθεια των σχημάτων όσο και με την βοήθεια του παραδείγματος, κατανοήθηκε ο τρόπος υπολογισμού του πρώτου κέντρου για το πρώτο δεδομένο εισόδου. Στην συνέχεια, ελέγχεται αν τα υπόλοιπα δεδομένα εισόδου βρίσκονται έξω από τον κύκλο ο οποίος ορίζεται από τα ήδη επιλεγμένα κέντρα. Αν αυτό δεν συμβαίνει με τον ίδιο τρόπο προστίθεται ένα νέο κέντρο. Αφού γίνει ένα πέρασμα όλων των δεδομένων εισόδου έχουν υπολογιστεί τα κέντρα του δικτύου RBF. Με βάση όλα τα παραπάνω, ο ψευδοκώδικας που περιγράφει τον αλγόριθμο Fuzzy Means είναι ο εξής [68]:

Αλγόριθμος Fuzzy Means

Είσοδος: Κανονικοποιημένες μεταβλητές εισόδου(Utrain) από το Training Dataset

Ακέραιος αριθμός ασαφών συνόλων $s = [s_1, s_2, \dots, s_N]$

όπου N ο αριθμός των διαστάσεων

Έξοδος: Αριθμός των επιλεγμένων κέντρων RBF L

1. Συντεταγμένες κέντρων $\hat{U} = [\hat{u}_1, \hat{u}_2, \dots, \hat{u}_L]^T$
2. Παίρνουμε το 1^ο δεδομένο: $k=1$
3. Υπολογισμός εύρους δa . Διαχωρισμός του χώρου για κάθε x_i σε s ασαφή σύνολα
4. Πρώτο κέντρο RBF
5. **Για** $i=1:N$
6. Υπολογισμός του fuzzy set με την μεγαλύτερη συνάρτηση συμμετοχής μ_A σε κάθε διάσταση i .
7. **Τέλος**
8. Δημιουργία πρώτου RBF κέντρου: $\hat{u}_1 = [\alpha_1^1, \alpha_2^1, \dots, \alpha_N^1]$
9. **Για** $k=2:K$
10. **Αν** το data point βρίσκεται έξω από τον κύκλο ο οποίος ορίζεται από τα ήδη επιλεγμένα κέντρα, δηλαδή εάν $\min_{1 \leq l \leq L} [d_r^l(\mathbf{u}(k))] > 1$
11. Προστίθεται νέο κέντρο: $L \leftarrow L+1$
12. **Για** $i=1:N$
13. Υπολογισμός του fuzzy set με την μεγαλύτερη συνάρτηση συμμετοχής μ_A σε κάθε διάσταση i .
14. **Τέλος**
15. Δημιουργία του L RBF κέντρου: $\hat{u}_L = [\alpha_1^L, \alpha_2^L, \dots, \alpha_N^L]$
16. **Τέλος**
17. **Τέλος**

Τα κέντρα RBF επιλέγονται έτσι ώστε να διασφαλίζεται ότι όλα τα σημεία των δεδομένων εκπαίδευσης καλύπτονται από τουλάχιστον μία υπερσφαίρα. Η τελική επιλογή εξαρτάται από μία μόνο παράμετρο, δηλαδή τον εκχωρημένο αριθμό ασαφών συνόλων s , που είναι κοινός για όλες τις μεταβλητές εισόδου. Γενικά παρατηρείται ότι η χρήση μεγαλύτερου αριθμού s δημιουργεί ένα πιο πυκνό πλέγμα και έχει ως αποτέλεσμα την επιλογή περισσότερων κέντρων RBF [70].

2.5.4 Πλεονεκτήματα της χρήσης του αλγόριθμου Fuzzy Means

Η μέθοδος των ασαφών μέσων παρουσιάζει πλεονεκτήματα σε σχέση με συμβατικές μεθόδους εκπαίδευσης νευρωνικών δικτύων RBF. Πιο συγκεκριμένα, πολύ σημαντικό είναι πως δεν απαιτείται εκ των προτέρων να καθοριστεί ο αριθμός των κόμβων της κρυφής στοιβάδας, αφού ο αλγόριθμος από μόνος του υπολογίζει την διάσταση της. Επιπλέον, ένα σημαντικό πλεονέκτημα της μεθόδου αυτής σχετίζεται με την μείωση του υπολογιστικού χρόνου κατά την εκπαίδευση του δικτύου. Η μείωση αυτή οφείλεται στο γεγονός του ότι ο αλγόριθμος χρειάζεται μόνο ένα πέρασμα από τα δεδομένα σε αντίθεση με άλλες μεθόδους οι οποίες χρειάζονται πολλά περάσματα έως ότου να συγκλίνουν [71]. Τέλος, ο αλγόριθμος των ασαφών μέσων έχει τη δυνατότητα να προσδιορίζει αυτόματα το μέγεθος του δικτύου.

ΚΕΦΑΛΑΙΟ 3^ο : Πειραματικό μέρος

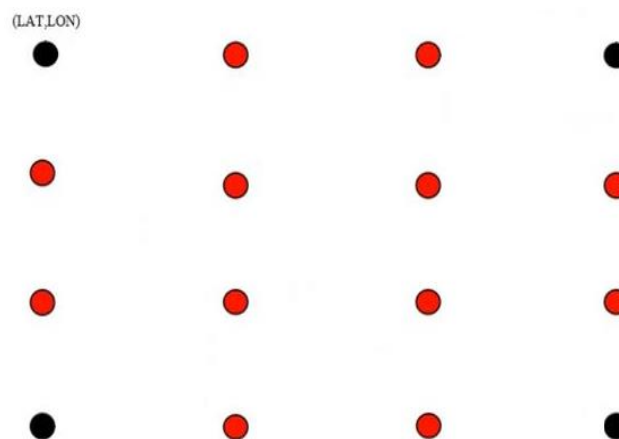
Η ανάγκη για την πρόβλεψη κύματος σε μια ποικιλία από εφαρμογές στην ωκεανογραφία ώθησε, λόγω των αδυναμιών των συμβατικών μοντέλων πρόβλεψης, στην χρήση των νευρωνικών δικτύων για την πρόβλεψη του μέσου ή του μέγιστου ύψους κύματος. Συγκεκριμένα, τα νευρωνικά δίκτυα ακτινικής βάσης λόγω της δομής τους κρίθηκαν κατάλληλα για εφαρμογές όπως αυτές που ήδη αναφέρθηκαν. Σε συνδυασμό με τον αλγόριθμο των ασαφών μέσων επιχειρήθηκε, λοιπόν, μια προσπάθεια πρόβλεψης του ύψους κύματος.

Κάνοντας, έτσι, μια προσπάθεια σύνδεσης της μέχρι τώρα θεωρίας των νευρωνικών δικτύων ακτινικής βάσης με την πρόβλεψη του ύψους κύματος, στο παρόν κεφάλαιο της διπλωματικής εργασίας περιγράφεται η προέλευση και η σύσταση των δεδομένων, καθώς και η πρακτική εφαρμογή τους ως δεδομένα εισόδου στο νευρωνικό δίκτυο RBF. Επιπλέον, παρουσιάζονται τα αποτελέσματα των σεναρίων καθώς και το feature engineering που συνέβαλε στην βελτιστοποίηση των αποτελεσμάτων. Με αυτόν τον τρόπο θα εξαχθεί το συμπέρασμα για το ποια δεδομένα εισόδου έδωσαν τους βέλτιστους στατιστικούς δείκτες.

3.1 Περιγραφή δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν στην συγκεκριμένη διπλωματική εργασία σχετίζονται με το κυματισμό της θάλασσας. Για την προσομοίωση του κυματισμού απαιτείται η λύση μερικών διαφορικών εξισώσεων. Στο dataset που μελετήθηκε τα δεδομένα αναφέρονται σε κυματικά χαρακτηριστικά δύο χρόνων και παράγονται από αναλυτικά μοντέλα με μεγάλη πολυπλοκότητα τα οποία τρέχουν σε υπέρ υπολογιστές. Παρόμοια μοντέλα χρησιμοποιούνται και από τις μετεωρολογικές υπηρεσίες για την πρόβλεψη του καιρού.

Στα μοντέλα αυτά, η γεωγραφική περιοχή μελέτης χωρίστηκε σε ένα πλέγμα (grid). Όσο πιο πυκνό είναι το πλέγμα τόσο μεγαλύτερη ακρίβεια έχει το μοντέλο. Ωστόσο, χρειάζεται περισσότερο χρόνο για να λυθεί καθώς ο υπολογιστικός χρόνος είναι ένα μέγεθος που αυξάνεται εκθετικά. Αντίθετα, η χρήση του αραιού πλέγματος παρουσιάζει το πρόβλημα της μειωμένης ακρίβειας, όμως ο υπολογιστικός χρόνος για την λύση του μοντέλου είναι μικρός. Στο σχήμα 3.1 παρουσιάζεται η αναπαράσταση των δύο πλεγμάτων:



Σχήμα 3.1: Αναπαράσταση πυκνού και αραιού πλέγματος

Τα μαύρα σημεία αντιπροσωπεύουν τα σημεία του αραιού πλέγματος, ενώ τα κόκκινα αυτά του πυκνού. Αυτό που παρουσιάζει ιδιαίτερο ενδιαφέρον για αρχή είναι οι τιμές τόσο του αραιού όσο και του πυκνού πλέγματος στα μαύρα σημεία και οι αποκλίσεις μεταξύ τους. Η διαφορά αυτή οφείλεται κυρίως στον τρόπο με τον οποίο λύνονται οι διαφορικές εξισώσεις.

Το σετ δεδομένων που μελετήθηκε και πάνω σε αυτό χρησιμοποιήθηκε η μέθοδος των νευρωνικών δικτύων, αφορά την περιοχή κοντά στον Άγιο Νικόλα της Κρήτης. Η γεωγραφική αυτή περιοχή επιλέχθηκε διότι το σφάλμα ανάμεσα στο αραιό και το πυκνό πλέγμα είναι πολύ μεγάλο. Η απόκλιση που παρατηρείται ανάμεσα στα δύο πλέγματα οφείλεται πιθανόν σε περιοχές με ιδιαίτερη γεωγραφική δομή (π.χ. μικρά νησάκια) και στην κατεύθυνση που φυσάει ο αέρας στις περιοχές αυτές. Πρέπει, λοιπόν, να δημιουργηθεί ένα μοντέλο το οποίο να δέχεται σαν εισόδους πληροφορίες από το αραιό πλέγμα και δεδομένα τα οποία ήδη έχουν μετρηθεί και παραχωρηθεί. Τα δεδομένα που χρησιμοποιούνται σαν εισόδοι στο νευρωνικό δίκτυο είναι τα εξής:

1. Η βαθυμετρία (DP) δοσμένη σε μέτρα. Με τον όρο βαθυμετρία αναφερόμαστε στις πληροφορίες που περιγράφουν το βάθος του βυθού. Η βαθυμετρία αποτελεί ουσιαστικό συστατικό για την κατανόηση της δυναμικής της θάλασσας, τόσο από την άποψη της μεταφοράς ιζημάτων όσο και στην πρόβλεψη παλίρροιας, ρευμάτων και κυμάτων [72].
2. Η μέγιστη συχνότητα (PKFR) σε Hz. Η συχνότητα των κυμάτων που αντιπροσωπεύεται από μια κορυφή στο φάσμα κύματος. Η κορυφή αυτή έχει την μέγιστη ενέργεια και είναι γνωστή και ως κυρίαρχη συχνότητα.
3. Η μέση συχνότητα (MEANFR) σε Hz.
4. Η μέση κατεύθυνση του ανέμου (MWDIR) δοσμένη σε μοίρες. Ορίζεται ως ο μέσος όρος όλων των επιμέρους κατευθύνσεων κύματος σε μια χρονοσειρά που αντιπροσωπεύει μια ορισμένη θαλάσσια κατάσταση.
5. Η ταχύτητα του ανέμου (WINSF) σε 10m με μονάδα μέτρησης m/sec. Περιγράφει το πόσο γρήγορα κινείται ο αέρας πέρα από ένα συγκεκριμένο σημείο.
6. Η κατεύθυνση του ανέμου (WINDDIR) σε μοίρες. Ορίζεται ως η πραγματική κατεύθυνση στην οποία φυσά ο άνεμος σε μια δεδομένη τοποθεσία.

Η ταχύτητα και η κατεύθυνση του ανέμου είναι σημαντικές για την παρακολούθηση και την πρόβλεψη τόσο του καιρού και όσο και του παγκόσμιου κλίματος. Οι δύο αυτές συνιστώσες του ανέμου έχουν πολλές επιπτώσεις στην επιφάνεια της θάλασσας και επηρεάζουν φαινόμενα όπως ο ρυθμός εξάτμισης, η ανάμειξη των επιφανειακών υδάτων και η ανάπτυξη καταγίδων. Κάθε μία από αυτές τις διαδικασίες έχει σημαντικές επιπτώσεις στην ποιότητα και στη στάθμη του νερού.

7. Το μέσο ύψος κύματος (SWHT) σε m. Είναι ένας όρος που χρησιμοποιείται για να δηλώσει το χαρακτηριστικό ύψος των τυχαίων κυμάτων σε μια θαλάσσια κατάσταση. Ορίζεται με τέτοιο τρόπο ώστε να αντιστοιχεί σε αυτό που παρατηρεί ένας ναυτικός όταν εκτιμά οπτικά το μέσο ύψος κύματος [73].

8. Το μέγιστο ύψος κύματος (Maximum Weight Height, MWT) σε m.

Επομένως, με βάση τα παραπάνω, το μοντέλο με το νευρωνικό δίκτυο θα δέχεται σαν είσοδο πληροφορίες από το αραιό πλέγμα και δεδομένα που είναι κοινά και στα δύο πλέγματα και αυτό με απώτερο σκοπό να προβλεφθεί το ύψος κύματος (μέσο ή μέγιστο) στα αντίστοιχα σημεία του πυκνού. Γνωρίζοντας τα αποτελέσματα του αραιού πλέγματος γίνεται μια προσπάθεια πρόβλεψης των χαρακτηριστικών του πυκνού πλέγματος.

3.1.1 Dataset δεδομένων

Όπως αναφέρθηκε, ήδη, στο dataset που μελετήθηκε περιέχονται κυματικά δεδομένα δύο χρόνων. Τα κυματικά αυτά δεδομένα φαίνονται στο πίνακα 3.1:

Πίνακας 3.1: Dataset Δεδομένων

Counter	0	1	2	3	4	5	6	7
YYYY	2020	2020	2020	2020	2020	2020	2020	2020
MM	1	1	1	1	1	1	1	1
DD	1	1	1	1	1	1	1	1
H	0	0	0	0	0	0	0	0
LON	26.30	26.35	26.35	26.15	26.20	26.30	26.35	26.15
LAT	35.15	35.15	35.20	35.25	35.25	35.25	35.25	35.30
SWHT	0.150	0.520	0.570	0.820	0.640	0.230	0.630	0.770
MWH	0.260	0.970	1.06	1.15	0.960	0.36	1.16	1.42
SWHT_Sp.	0.40	0.66	0.71	0.99	0.95	0.38	0.77	1.01
MWH_Sp.	0.74	1.23	1.330	1.84	1.78	0.67	1.44	1.89
DP_Sp.	44	192	386	88	17	22	160	128
PKFR_Sp.	0.2037	0.1975	0.1957	0.1656	0.1708	0.1733	0.1942	0.1678
MEANFR_Sp.	0.2865	0.2294	0.2279	0.1921	0.1974	0.2901	0.2204	0.1933
MWDIR_Sp.	355.72	348.69	343.86	329.39	326.81	329.52	339.46	328.36
WINSP_Sp.	5.55	5.76	6.81	5.96	6.23	7.47	7.75	6.10
WINDDIR_Sp.	319.29	321.93	315.29	305.44	304.98	308.70	311.19	305.47

Όπου η περιγραφή των δεδομένων κάθε στήλης του πίνακα 3.1 αναλύεται παρακάτω:

- Counter: Αριθμός δεδομένου.
- YYYY: Χρόνος.
- MM: Μήνας.
- DD: Ημέρα.
- H: Ωρα.
- LON: Γεωγραφικό μήκος.
- LAT: Γεωγραφικό πλάτος.
- SWHT: Μέσο ύψος κύματος στο πυκνό grid.
- MWH: Μέγιστο ύψος κύματος στο πυκνό grid.
- SWHT_Sparse: Μέσο ύψος κύματος στο αραιό grid.
- MWH_Sparse: Μέγιστο ύψος κύματος στο αραιό grid.
- DP_Sparse: Βαθυμετρία στο αραιό grid.
- PKFR_Sparse: Μέγιστη συχνότητα στο αραιό grid.
- MEANFR_Sparse: Μέση συχνότητα στο αραιό grid.
- WINSP_Sparse: Ταχύτητα του ανέμου στο αραιό grid.
- WINDDIR_Sparse: Κατεύθυνση του ανέμου στο αραιό grid.

Τα πλήθος των δεδομένων της διπλωματικής εργασίας είναι 251.120 δεδομένα. Συγκεκριμένα, προέρχονται από το χρονικό διάστημα του 2020 με 2021. Υπάρχουν καταγραφές για όλες τις ημέρες και όλους τους μήνες, ενώ κάθε μια ώρα καταγράφονται 16 μετρήσεις των χαρακτηριστικών του κύματος. Η μέγιστη τιμή του γεωγραφικού μήκους είναι 26.350 και η ελάχιστη 26.150. Αντίστοιχα, η μέγιστη του γεωγραφικού πλάτους είναι 35.350 και η ελάχιστη

35.150. Το βήμα με το οποίο αυξάνονται οι συνιστώσες αυτές είναι 0.05. Στον πίνακα 3.2 φαίνεται το γεωγραφικό μήκος και το γεωγραφικό πλάτος των nodes:

Πίνακας 3.2: Γεωγραφικό μήκος και πλάτος των nodes

LON	LAT
26.30	35.15
26.35	35.15
26.35	35.20
26.15	35.25
26.20	35.25
26.30	35.25
26.35	35.25
26.15	35.30
26.20	35.30
26.25	35.30
26.35	35.30
26.15	35.35
26.20	35.35
26.25	35.35
26.30	35.35
26.35	35.35

Στον πίνακα 3.3 παρουσιάζονται οι μέγιστες και οι ελάχιστες τιμές για τα χαρακτηριστικά κύματος που αναφέρθηκαν:

Πίνακας 3.3: Μέγιστες και ελάχιστες τιμές χαρακτηριστικών κύματος

	min	max
SWHT	0.010	5.42
MWH	0.010	10.36
SWHT_Sparse	0.010	5.68
MWH_Sparse	0.020	10.85
DP_Sparse	17	386
PKFR_Sparse	0.0827	0.5417
MEANFR_Sparse	0.0903	0.6324
MWDIR_Sparse	0	360
WINSP_Sparse	2	22.48
WINDDIR_Sparse	0	359.98

Στα μοντέλα των νευρωνικών δικτύων κρίνεται απαραίτητος ο διαχωρισμός των δεδομένων σε training, validation και testing set. Για να αποφευχθεί το φαινόμενο του overfitting, τα δεδομένα του πρώτου χρόνου χρησιμοποιήθηκαν για την εκπαίδευση του μοντέλου, ενώ αυτά του δεύτερου χρόνου χωρίστηκαν σε μισό χρόνο για την αξιολόγηση του μοντέλου στο validation set και άλλο μισό χρόνο για την επιβεβαίωση του στο testing set. Στο σχήμα 3.2 φαίνεται ο διαχωρισμός των δεδομένων, που αναφέρθηκαν παραπάνω, σε training, validation και testing set:

testing_data_set1	54896x12 double
training_data_set1	131712x12 double
validation_data_set1	64512x12 double

Σχήμα 3.2: Training, validation, testing dataset

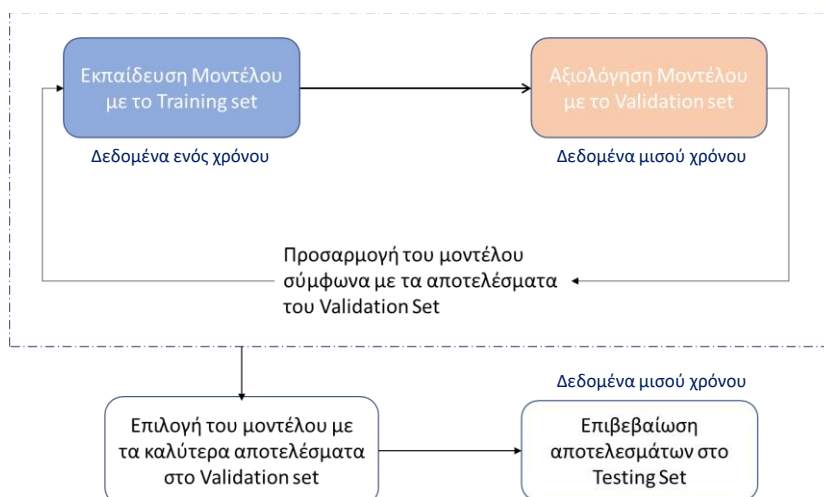
Το σετ εκπαίδευσης (training) είναι το σύνολο των δεδομένων που χρησιμοποιείται για να εκπαιδευτεί το μοντέλο και να μάθει τα κρυφά χαρακτηριστικά/μοτίβα στα δεδομένα. Το σετ αυτό θα πρέπει να έχει ένα διαφοροποιημένο σύνολο εισόδων, έτσι ώστε το μοντέλο να είναι εκπαιδευμένο σε όλα τα πιθανά σενάρια και να μπορεί να προβλέψει οποιοδήποτε δείγμα δεδομένων που μπορεί να εμφανιστεί στο μέλλον.

Το validation σετ είναι ένα σύνολο δεδομένων, ξεχωριστό από το σύνολο εκπαίδευσης, που συμβάλει στην επικύρωση της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης. Αποτελεί, δηλαδή, ένα κριτήριο που φανερώνει αν η εκπαίδευση κινείται προς τη σωστή κατεύθυνση ή όχι. Επιπλέον, από το σετ αυτό υπολογίζονται οι στατιστικούς δείκτες με σκοπό να βρεθεί το για ποια τιμή του s το μοντέλο δίνει τα καλύτερα αποτελέσματα με βάση το μικρότερο RMSE.

Το testing σετ είναι ένα σύνολο δεδομένων που αξιοποιείται για τη δοκιμή του μοντέλου μετά την ολοκλήρωση της εκπαίδευσης. Παρέχει μια αμερόληπτη μέτρηση απόδοσης του τελικού μοντέλου όσον αφορά την ακρίβεια και από το σύνολο αυτό υπολογίζονται οι δείκτες σφάλματος.

Στα μοντέλα μηχανικής μάθησης είναι γνωστό πως παρατηρείται το φαινόμενο του overfitting. Ένα στατιστικό μοντέλο λέγεται ότι κάνει overfitting οι προβλέψεις στο testing set δεν είναι ακριβείς. Κατά την διαδικασία της εκπαίδευσης και λόγω του υψηλού όγκου των δεδομένων, υπάρχει ο κίνδυνος το νευρωνικό δίκτυο να αρχίζει να μαθαίνει από τον θόρυβο και τις ανακριβείς καταχωρίσεις των δεδομένων. Όταν γίνεται ο έλεγχος με το testing dataset και παρατηρείται υψηλή διακύμανση, τότε το μοντέλο δεν κάνει σωστές προβλέψεις λόγω υπερβολικά πολλών λεπτομερειών και θορύβου. Οι αιτίες του overfitting είναι οι μη παραμετρικές και μη γραμμικές μέθοδοι και αυτό διότι αυτοί οι τύποι αλγορίθμων μηχανικής μάθησης έχουν μεγαλύτερη ελευθερία στη δημιουργία του μοντέλου και επομένως μπορούν να δημιουργήσουν μη ρεαλιστικά μοντέλα. Μια λύση για να αποφευχθεί το overfitting είναι η αύξηση των δεδομένων στο training set.

Στο σχήμα 3.3 παρουσιάζεται η διαδικασία που η περιγραφή της έγινε παραπάνω:



Σχήμα 3.3: Training Data/validation/testing

3.2 Μεθοδολογία

Απώτερος στόχος της παρούσας διπλωματικής εργασίας είναι η πρόβλεψη του ύψους κύματος του πυκνού πλέγματος γνωρίζοντας τα δεδομένα για το αραιό. Παρακάτω, περιγράφεται πως αυτό έγινε με την βοήθεια των νευρωνικών δικτύων ακτινικής βάσης και του αλγορίθμου Fuzzy Means. Επίσης, γίνεται μια αναφορά στα σενάρια που έτρεξαν και στα αποτελέσματα αυτών.

3.2.1 RBF και Fuzzy Means

Το πρώτο βήμα για την πρόβλεψη του ύψους κύματος ήταν η υλοποίηση του κώδικα Fuzzy Means ο οποίος υπολογίζει τα κέντρα του δικτύου RBF. Καταρχάς, ορίστηκε μια συνάρτηση που δέχεται σαν όρισμα τον αριθμό ασαφών συνόλων s , τις κανονικοποιημένες μεταβλητές εισόδου και επιστρέφει τον πίνακά με τις συντεταγμένες των κέντρων. Έπειτα, υπολογίστηκε η απόσταση d και με την εντολή `linspace` διαμορφώθηκε ο πίνακας $[x_1 x_2 x_3 x_4 \dots x_n]$ ο οποίος ονομάστηκε `Vector_of_fuzzysets`. Στην συνέχεια, βρέθηκαν οι συντεταγμένες του πρώτου κέντρου. Συγκεκριμένα, για το πρώτο δεδομένο μέσα σε ένα διπλό `for` υπολογίζεται σε κάθε διάσταση η απόσταση του κάθε σημείου από το αντίστοιχο σημείο του `Vector_of_fuzzysets` και αποθηκεύεται σε μια μεταβλητή η ελάχιστη απόσταση. Μετά το πέρας της διαδικασίας αυτής εκτιμήθηκαν οι συντεταγμένες του πρώτου κέντρου.

Όπως αναφέρεται στο υποκεφάλαιο 2.5.2, για να συμπληρωθεί ο πίνακας με τις συντεταγμένες των κέντρων για όλα τα δεδομένα πρέπει να ελεγχθεί εάν ο ασαφής υποχώρος καλύπτει όλα τα διανύσματα εισόδου. Ειδικότερα, εάν η ελάχιστη απόσταση $d_r^l(\mathbf{u}(k))$ είναι μεγαλύτερη από την μονάδα, τότε με τον ίδιο τρόπο, όπως και στην περίπτωση του πρώτου κέντρου, υπολογίζονται οι νέες συντεταγμένες. Ο κώδικας του Fuzzy Means δημιουργήθηκε ως μια ρουτίνα (function) η οποία καλείται μέσα στον κώδικα του RBF.

Αφού ολοκληρώθηκε το κομμάτι του αλγόριθμου για τον υπολογισμό των κέντρων, ακολούθησε η συγγραφή του κώδικα για το νευρωνικό δίκτυο RBF. Αρχικά, τα δεδομένα εισόδου διαχωρίστηκαν σε `training`, `validation` και `testing set` και κανονικοποιήθηκαν με την μέθοδο κανονικοποίησης σε συγκεκριμένο εύρος τιμών με γραμμικό μετασχηματισμό. Αξίζει επίσης να αναφερθεί, πως αφού έγινε η κανονικοποίηση του `training data set`, αποθηκεύτηκαν οι σταθερές κανονικοποίησης και έπειτα χρησιμοποιήθηκαν για την κανονικοποίηση του `validation` και `testing set`. Έπειτα, στην `for` από 4 έως 50, όπου το εύρος αυτό δηλώνει τον αριθμό ασαφών συνόλων (s), καλείται η συνάρτηση του αλγόριθμου Fuzzy Means. Το εύρος 4 έως 50 επιλέγεται διότι πάνω από αυτόν τον αριθμό των ασαφών συνόλων δεν γίνεται αισθητή η μείωση του σφάλματος και ο χρόνος για την εκπαίδευση και την πρόβλεψη του μοντέλου αυξάνεται αισθητά. Η συνάρτηση του Fuzzy Means, όπως ήδη αναφέρθηκε, επιστρέφει τον πίνακα με τις συντεταγμένες των κέντρων. Με βάση αυτές, υπολογίζεται η δραστηριότητα του κόμβου I (μ) και η έξοδος του κρυφού νευρώνα z σύμφωνα με τις εξισώσεις που αναφέρθηκαν παραπάνω (2.1, 2.3). Με την μέθοδο των ελαχίστων τετραγώνων γίνεται ο υπολογισμός των συναπτικών βαρών και στην συνέχεια της εξόδου. Η τελική της τιμή προκύπτει με την αποκανονικοποίηση των δεδομένων.

Αυτή η διαδικασία επαναλαμβάνεται τόσο στο `training` όσο και στο `validation dataset`. Αυτό συμβαίνει διότι υπολογίζοντας τους στατιστικούς δείκτες στο `validation set` και πιο συγκεκριμένα το RMSE, είναι δυνατόν να βρεθεί το για ποια τιμή του s το μοντέλο δίνει τα καλύτερα αποτελέσματα. Μετά το πέρας των 50 επαναλήψεων έχουν υπολογιστεί τα συναπτικά βάρη και οι τελικές εξοδοί για τα δυο `dataset` και τέλος υπολογίζεται η έξοδος και

οι στατιστικοί δείκτες για το testing dataset. Η έξοδος του νευρωνικού δικτύου ακτινικής βάσης αποτέλεσε και την πρόβλεψη του μέσου ή του μέγιστου ύψους κύματος.

3.2.2 Περιγραφή Σεναρίων

Αφού ολοκληρώθηκε η συγγραφή του κώδικα και η επικύρωση του μέσω άλλων benchmark datasets [7], ιδιαίτερη έμφαση δόθηκε στην περιοχή του Αγίου Νικολάου της Κρήτης. Συγκεκριμένα, τα σενάρια τα οποία υλοποιήθηκαν χωρίστηκαν στις εξής κατηγορίες:

3.2.3 Σενάρια με την χρήση όλων των features

Σε αυτή την κατηγορία σεναρίων δόθηκαν σαν είσοδοι στο νευρωνικό δίκτυο όλα τα διαθέσιμα δεδομένα του dataset. Πιο συγκεκριμένα, τα χαρακτηριστικά αυτά ήταν τα παρακάτω: SWHT_Sparse, MWH_Sparse, DP_Sparse, PKFR_Sparse, MEANFR_Sparse, WINSF_Sparse και το WINDDIR_Sparse. Η σημασία των χαρακτηριστικών αυτών περιγράφεται αναλυτικά παραπάνω. Χρησιμοποιώντας σαν είσοδο τα δεδομένα αυτά, έγινε η πρόβλεψη τόσο του SWHT όσο και του MWH του πυκνού πλέγματος.

Για την συλλογή των αποτελεσμάτων που προέκυψαν από το νευρωνικό δίκτυο δημιουργήθηκε ο πίνακας 3.4. Σε αυτόν φαίνονται χαρακτηριστικά όπως το εύρος των ασαφών υποχώρων για το οποίο έτρεξε το μοντέλο, ο βέλτιστος αριθμός αυτών ο οποίος επιλέχθηκε, ο αριθμός των κέντρων και οι στατιστικοί δείκτες για το s αυτό. Αυτό που πρέπει να τονιστεί είναι πως ο χρόνος που αναφέρεται στους πίνακες αφορά τον χρόνο της εκπαίδευσης του δικτύου διότι ο χρόνος πρόβλεψης θεωρείται αμελητέος.

Πίνακας 3.4: Σενάρια χωρίς Feature Engineering

	SWHT with all features	MWH with all features
s: fuzzy partition	4 έως 14	4 έως 14
Βέλτιστο s	14	14
Training time	1403.36 sec	1377.81 sec
Αριθμός κέντρων	1012	1012
RMSE_val	0.1395	0.2710
RMSE_te	0.1034	0.2088
R2_val	0.9498	0.9469
R2_te	0.9402	0.9450
MAE_val	0.0796	0.1560
MAE_testing	0.0703	0.1413

Οι στατιστικοί δείκτες σφάλματος οι οποίοι υπολογίστηκαν στο validation και testing set είναι το MAE, το R2 και το RMSE.

Συγκεκριμένα, το MAE ή αλλιώς το μέσο απόλυτο σφάλμα ορίζεται ως ο μέσος όρος των απόλυτων τιμών σφάλματος και υπολογίζεται ως εξής:

$$MAE = \frac{1}{n} \sum_{i=1}^N |x_i - x| \quad (3.1)$$

όπου:

n: ο αριθμός των δειγμάτων.

Σ: το σύμβολο της άθροισης.

$|x_i - x|$: το απόλυτο σφάλμα.

Το R2 είναι ένα στατιστικό μετρικό που αντιπροσωπεύει το ποσοστό της διακύμανσης για μια εξαρτημένη μεταβλητή που εξηγείται από μια ανεξάρτητη μεταβλητή ή από μεταβλητές σε ένα μοντέλο παλινδρόμησης. Συγκεκριμένα, το R2 είναι ένας δείκτης του πόσο καλύτερα μπορεί

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

να προβλεφθεί το y χρησιμοποιώντας το μοντέλο, σε σχέση με το να χρησιμοποιούνταν σαν πρόβλεψη ο μέσος όρος των μετρήσεων. Υπολογίζεται ως:

$$R2 = 1 - \frac{SSE}{SST} \quad (3.2)$$

όπου $SSE = \sum (y - y')^2$ και το $SST = \sum (y - \bar{y})^2$.

Ο όρος SSE δηλώνει το πόσο απέχουν οι προβλέψεις από τις μετρήσεις, ενώ ο όρος SST το πόσο απέχουν οι μετρήσεις από τον μέσο όρο τους.

Τέλος, το RMSE ή αλλιώς η ρίζα μέσου τετραγωνικού σφάλματος είναι ένα από τα πιο δημοφιλή στατιστικά μετρικά για την εκτίμηση της ακρίβειας των προβλεπόμενων τιμών του μοντέλου πρόβλεψης έναντι των πραγματικών τιμών κατά την εκπαίδευση των μοντέλων παλινδρόμησης. Υπολογίζεται ως:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (3.3)$$

όπου:

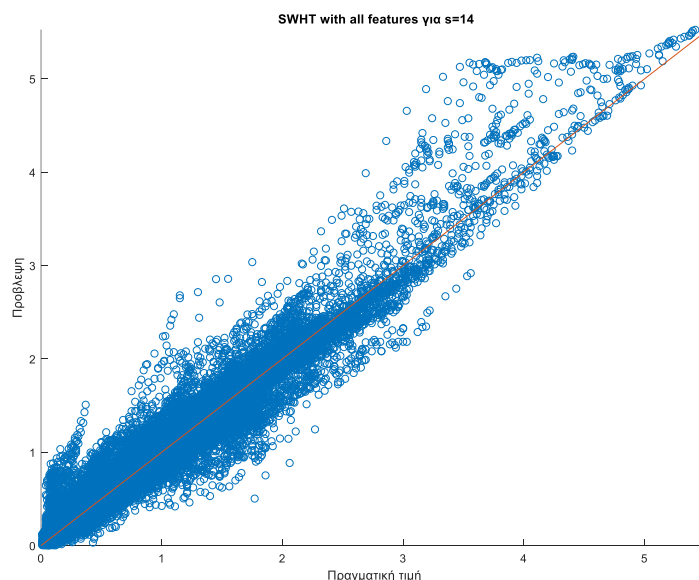
y_i : οι πραγματικές τιμές.

y'_i : οι προβλεπόμενες τιμές μιας μεταβλητής.

n : ο αριθμός των δεδομένων που είναι διαθέσιμα για ανάλυση.

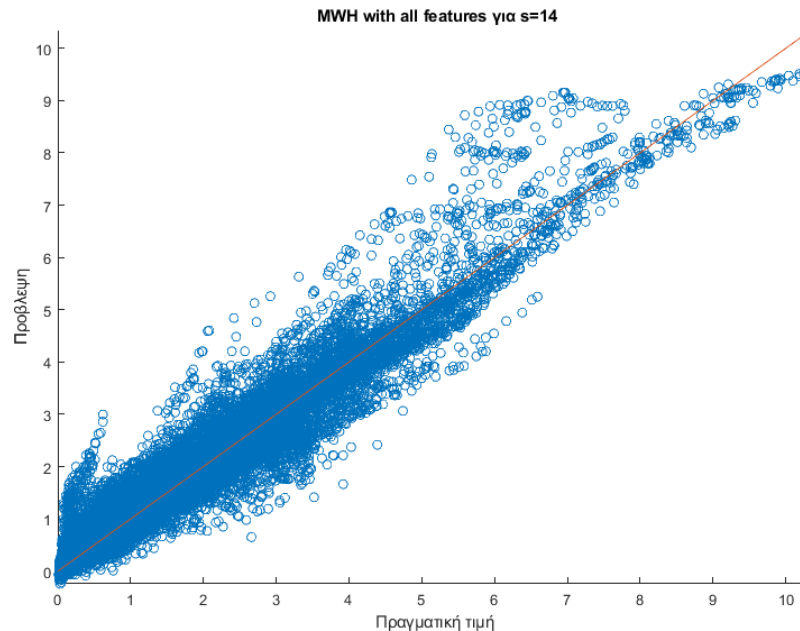
Εκτός από το τρέξιμο των σεναρίων, δημιουργήθηκαν και ορισμένα διαγράμματα. Στον άξονα x σχεδιάστηκαν οι πραγματικές τιμές ενώ στον άξονα y οι προβλέψεις. Όσο πιο πολύ συγκλίνουν τα δεδομένα στην ευθεία με κλίση 45 μοιρών τόσο πιο καλό θα μπορούσε να χαρακτηριστεί το μοντέλο. Ένας ακόμα δείκτης για την ποιότητα του, είναι οι δείκτες σφάλματος. Συγκεκριμένα, όσο πιο κοντά στην μονάδα είναι το $R2$ και όσο πιο μικρό το RMSE και το MAE, τόσο πιο καλά κάνει fit το μοντέλο.

Για το πρώτο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του SWHT με την χρήση όλων των features, αυτό που παρατηρήθηκε σύμφωνα και με τον πίνακα 3.4 ήταν πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης και της πρόβλεψης του μοντέλου ήταν 1403.36 δευτερόλεπτα. Συγκεκριμένα, για την εκπαίδευση του μοντέλου χρειάστηκαν 1400 δευτερόλεπτα, ενώ για την πρόβλεψη 3,36 δευτερόλεπτα. Από αυτή την παρατήρηση γίνεται αντιληπτό πως ο χρόνος της πρόβλεψης είναι πολύ μικρός και από εδώ και στο εξής θα θεωρείται αμελητέος. Τέλος, ο βέλτιστος ασαφής υποχώρος που επιλέχτηκε με βάση το μικρότερο RMSE στο validation set ήταν το $s = 14$. Για το s αυτό σχεδιάστηκε το



Σχήμα 3.4: Πρόβλεψη SWHT με όλα τα features

Για το δεύτερο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του MWH με την χρήση όλων των features, αυτό που παρατηρήθηκε σύμφωνα και με τον πίνακα 3.4 ήταν πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 1377.81 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος που επιλέχτηκε με βάση το μικρότερο RMSE στο validation set ήταν το $s = 14$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.5:



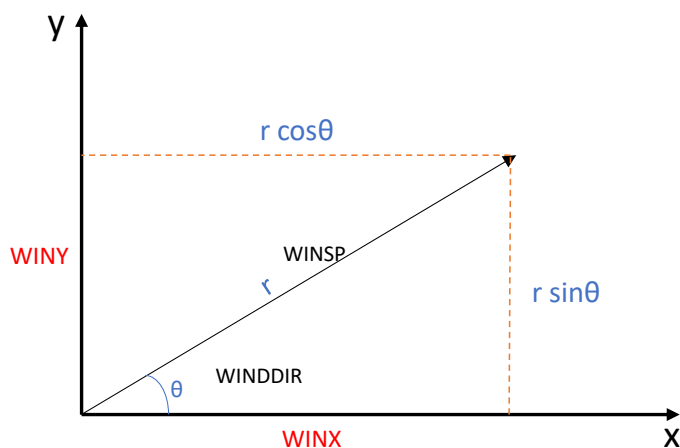
Σχήμα 3.5: Πρόβλεψη MWH με όλα τα features

3.2.4 Σενάρια με feature engineering στην ταχύτητα του ανέμου.

Το Feature engineering είναι η διαδικασία επιλογής, χειρισμού και μετατροπής ακατέργαστων δεδομένων σε χαρακτηριστικά που μπορούν να χρησιμοποιηθούν στην εποπτευόμενη ή μη εποπτευόμενη μάθηση. Προκειμένου η εκμάθηση να λειτουργήσει καλά, μπορεί να είναι απαραίτητο να σχεδιαστεί και να εκπαιδευτεί το νευρωνικό δίκτυο με διαφορετικά χαρακτηριστικά. Ένα χαρακτηριστικό (feature) είναι οποιαδήποτε μετρήσιμη είσοδος που μπορεί να χρησιμοποιηθεί σε ένα μοντέλο πρόβλεψης. Στην συγκεκριμένη περίπτωση, είναι η βαθυμετρία, η ταχύτητα και η κατεύθυνση του ανέμου, όπως αναφέρθηκε στον πίνακα 3.1. Η μηχανική χαρακτηριστικών, με απλά λόγια, είναι η πράξη μετατροπής ακατέργαστων παρατηρήσεων σε επιθυμητά χαρακτηριστικά χρησιμοποιώντας προσεγγίσεις στατιστικής ή μηχανικής μάθησης [74].

Στην ουσία με το Feature engineering αξιοποιούνται τα υπάρχοντα δεδομένα για τη δημιουργία νέων μεταβλητών που δεν περιλαμβάνονται στο training set. Με την χρήση των τεχνητών χαρακτηριστικών σκοπός είναι να βελτιωθεί η απόδοσή του μοντέλου ή με άλλα λόγια να αποκομιστούν καλύτερα αποτελέσματα.

Μελετώντας το dataset το οποίο δόθηκε, παρατηρήθηκε πως ένα χαρακτηριστικό στο οποίο θα μπορούσαμε να εφαρμοστεί η μέθοδος αυτή είναι η ταχύτητα του ανέμου. Η ταχύτητα του ανέμου στο αραιό πλέγμα μετρείται σε m/sec (WINSF_Sparse) και η κατεύθυνση του σε μοίρες (WINDDIR_Sparse). Για να επιτευχθεί το Feature engineering στην ταχύτητα του ανέμου με βάση την τριγωνομετρία έγινε ανάλυση στις δύο κάθετες συνιστώσες X και Y. Στο σχήμα 3.6 φαίνεται η ανάλυση αυτή.



Σχήμα 3.6: Ανάλυση σε WINX και WINY

Από την τριγωνομετρία είναι γνωστό ότι:

$$\sin \theta = \frac{x}{r} \text{ ή } \sin(WINDDIR) = \frac{WINX}{WINSP} \quad (3.4)$$

Επιπλέον, ισχύει ότι:

$$\cos \theta = \frac{y}{r} \text{ ή } \cos(WINDDIR) = \frac{WINY}{WINSP} \quad (3.5)$$

Με βάση τις σχέσεις (3.4), (3.5) οι συνιστώσες WINX και WINY υπολογίζονται:

$$WINX = WINSP \cdot \sin(WINDDIR) \quad (3.6)$$

$$WINY = WINSP \cdot \cos(WINDDIR) \quad (3.7)$$

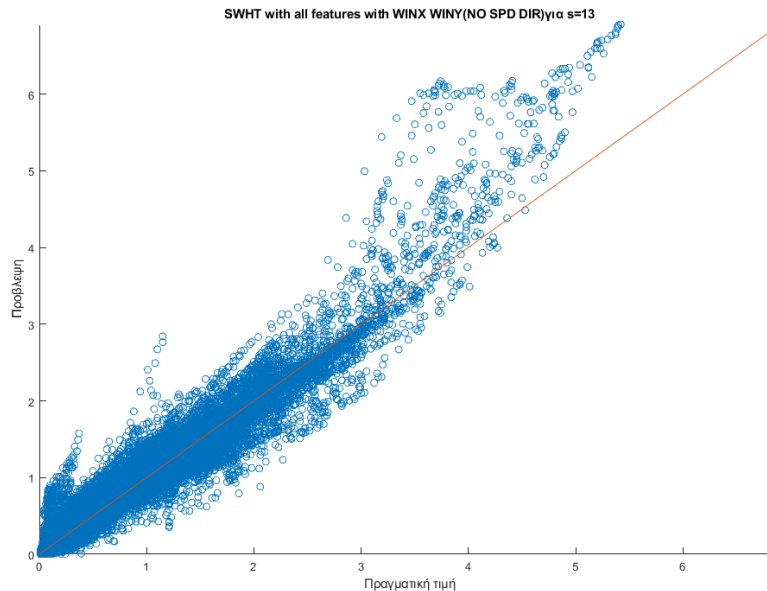
Αφού δημιουργήθηκαν οι δύο νέες συνιστώσες προστέθηκαν στο ήδη υπάρχον dataset. Έγιναν δοκιμές με το τρέξιμο νέων σεναρίων για να αποφανθεί αν η προσθήκη των συνιστωσών WINX και WINY βελτίωσε τα αποτελέσματα του μοντέλου. Στο σενάριο που εκτελέστηκε το νευρωνικό δίκτυο πήρε σαν εισόδους όλα τα features με την διαφορά του ότι αυτή την φορά χρησιμοποιήθηκαν και οι συνιστώσες WINX και WINY αντί αυτών του WINSP και WINDDIR. Προβλέφθηκε πάλι το μέσο και το μέγιστο ύψος κύματος και τα αποτελέσματα φαίνονται στον πίνακα 3.5:

Πίνακας 3.5: Σενάριο με Feature Engineering στην ταχύτητα του ανέμου-1

	SWHT with all features with WINX WINY (NO SPD DIR)	MWH with all features with WINX WINY (NO SPD/DIR)
s: fuzzy partition	4 έως 14	4 έως 14
Βέλτιστο s	13	13
Training time	936.19sec	949.14 sec
Αριθμός κέντρων	524	524
RMSE_val	0.1370	0.2746
RMSE_te	0.1062	0.2158
R2_val	0.9516	0.9455
R2_te	0.9477	0.9413
MAE_val	0.0817	0.1625
MAE_testing	0.0736	0.1495

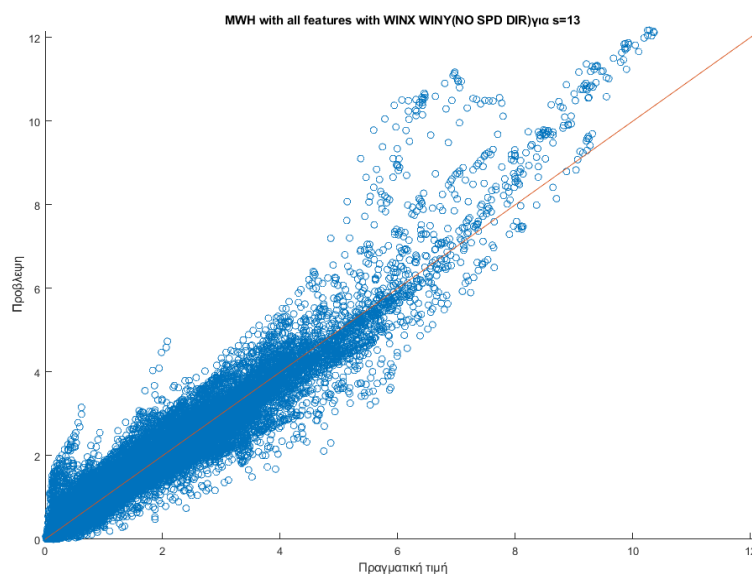
Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Για να υπάρξει και μια οπτική απεικόνιση πέρα από τους στατιστικούς δείκτες σχεδιάστηκαν και τα διαγράμματα. Για το πρώτο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του SWHT με την χρήση όλων των features και το WINX WINY, αυτό που παρατηρήθηκε, σύμφωνα και με τον πίνακα 3.5, είναι πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 936.19 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s=13$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.7:



Σχήμα 3.7: Πρόβλεψη SWHT με feature engineering

Για το δεύτερο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του MWH με την χρήση όλων των features και των WINX και WINY, αυτό που παρατηρήθηκε σύμφωνα με τον πίνακα 3.5 είναι πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 949.14 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s = 13$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.8:



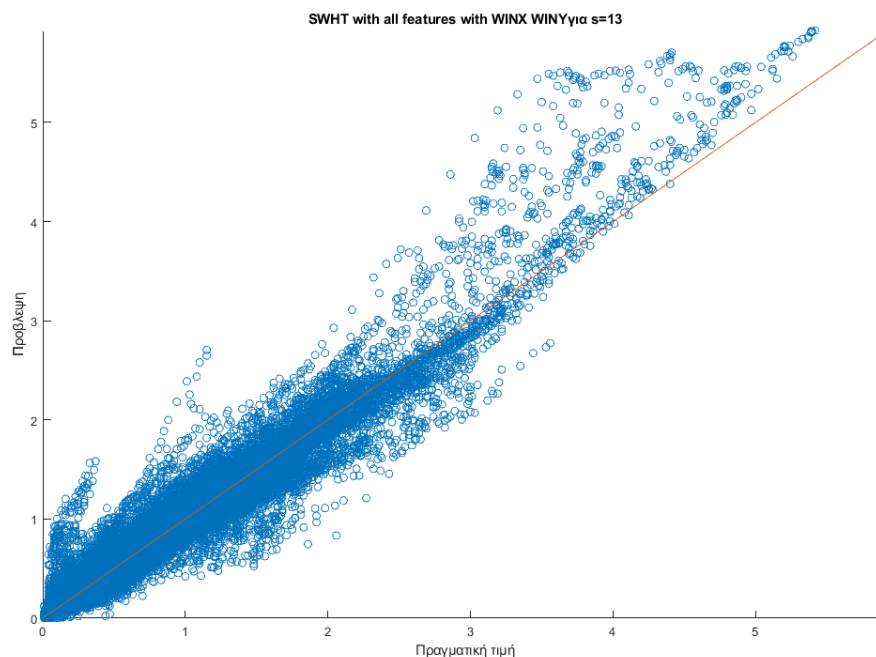
Σχήμα 3.8: Πρόβλεψη MWH με feature engineering

Μια εναλλαγή των δύο αυτών σεναρίων είναι και η προσθήκη όλων των features, της ταχύτητας και της κατεύθυνσης του ανέμου και των συνιστωσών WINX και WINY. Αφού εκτελέστηκαν τα σεσάρια αυτά, τα αποτελέσματα συγκεντρώθηκαν στον πίνακα 3.6:

Πίνακας 3.6: Σεσάρια με Feature Engineering στην ταχύτητα του ανέμου-2

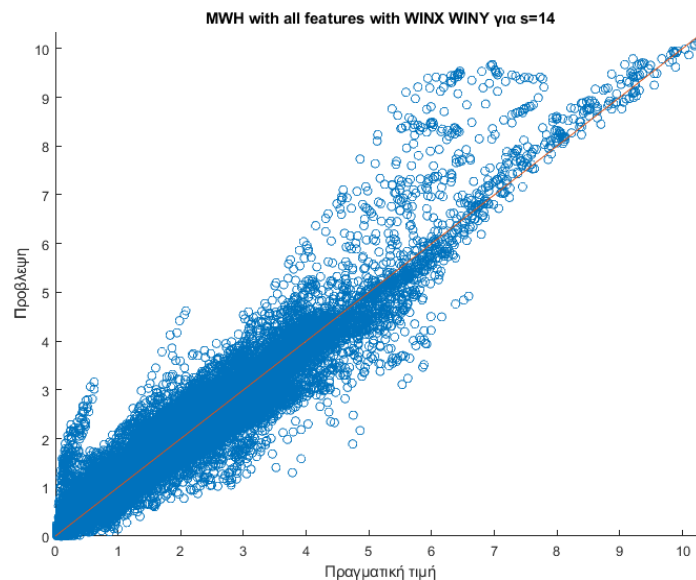
	SWHT with all features with WINX WINY	MWH with all features with WINX WINY
s: fuzzy partition	4 έως 14	4 έως 14
Βέλτιστο s	13	14
Training time	1331.67 sec	1323.34 sec
Αριθμός κέντρων	715	917
RMSE_val	0.1431	0.2858
RMSE_te	0.1068	0.2164
R2_val	0.9472	0.9410
R2_te	0.9470	0.9409
MAE_val	0.0831	0.1611
MAE_testing	0.0745	0.1458

Για το πρώτο σεσάριο παρατηρήθηκε, σύμφωνα και με τον πίνακα 3.6, πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 1331.67 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s = 13$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.9:



Σχήμα 3.9: Πρόβλεψη SWHT με feature engineering-2

Για το δεύτερο σεσάριο παρατηρήθηκε πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης ήταν 1323.34 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s = 14$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.10:



Σχήμα 3.10: Πρόβλεψη MWH με feature engineering-2

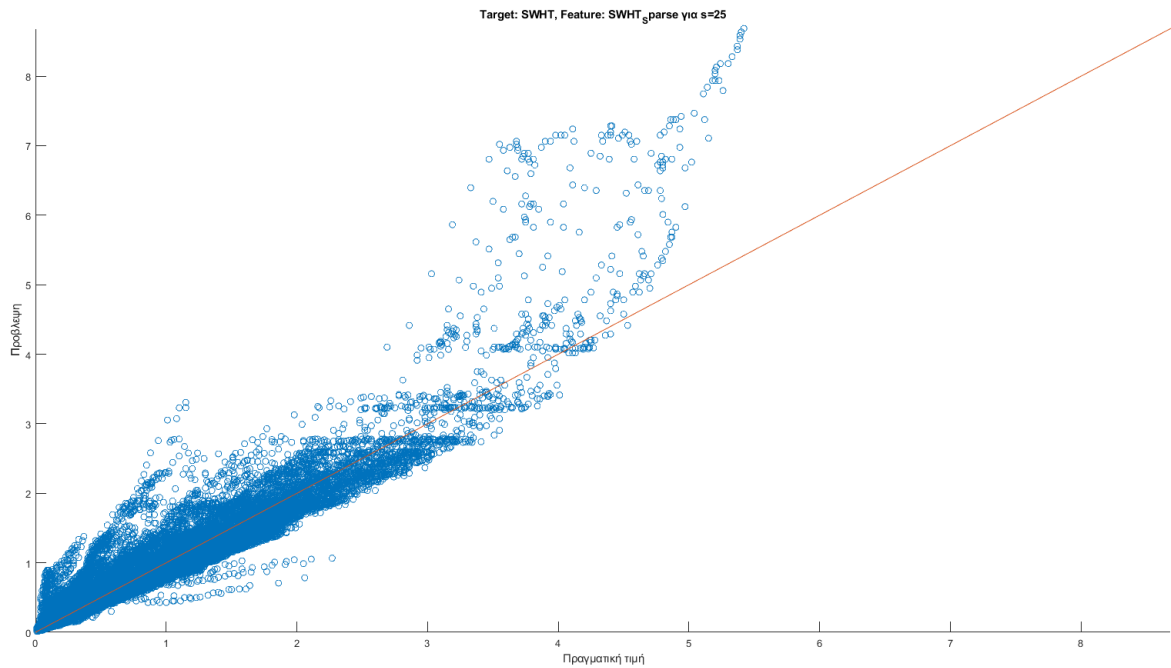
3.2.5 Σενάρια με αφαίρεση χαρακτηριστικών

Μια ακόμη ιδέα για την μείωση του χρόνου που απαιτείται για την ολοκλήρωση του κάθε τρεξίματος, είναι η αφαίρεση χαρακτηριστικών από τις εισόδους του μοντέλου. Για τον έλεγχο των επιθυμητών αποτελεσμάτων χρειάστηκε η πραγματοποίηση δοκιμών. Αρχικά, ένα σενάριο το οποίο εκτελέστηκε είτε θέλοντας να γίνει πρόβλεψη του μέγιστου είτε του μέσου ύψους κύματος, ήταν αυτό στο οποίο δόθηκαν σαν εισοδοί στο νευρωνικό δίκτυο το μέγιστο ή το μέσο ύψος κύματος του αραιού πλέγματος αντίστοιχα. Τα αποτελέσματα φαίνονται στον πίνακα 3.7:

Πίνακας 3.7: Σενάρια με αφαίρεση feature-1

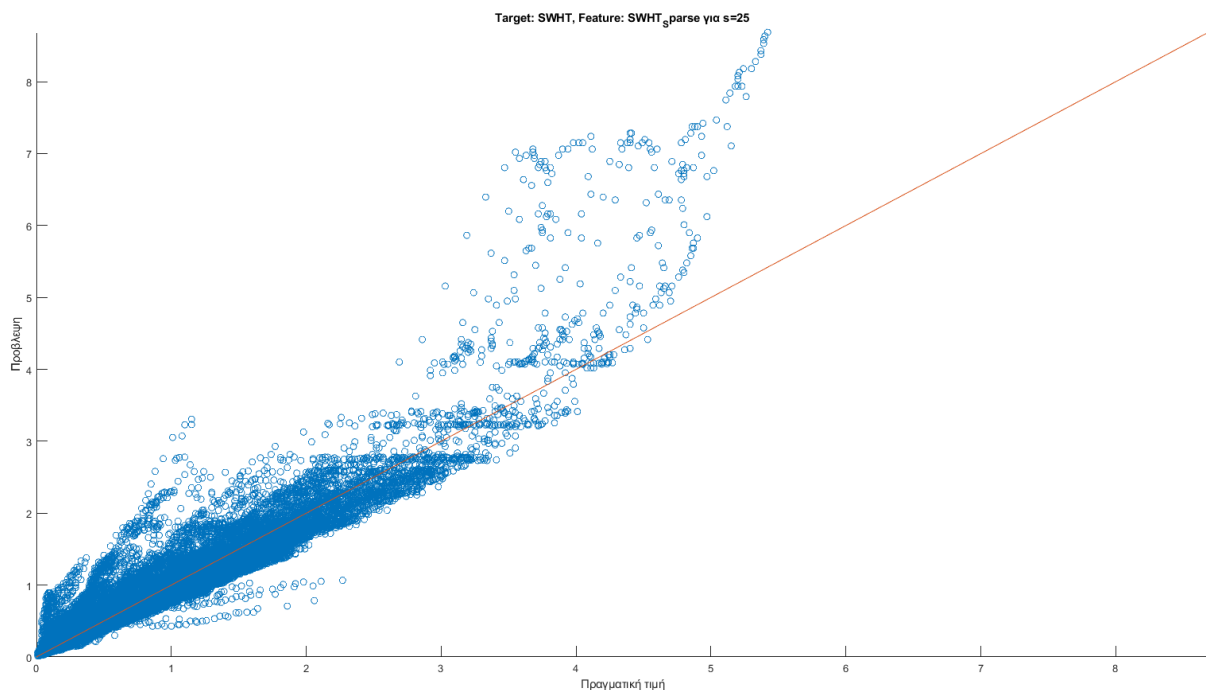
	Target: SWHT, Feature: SWHT_Sparse	Target: MWH, Feature: MWH_Sparse
s: fuzzy partition	4 έως 50	4 έως 50
Βέλτιστο s	25	25
Training time	49.3 sec	48.36 sec
Αριθμός κέντρων	23	23
RMSE_val	0.1881	0.3781
RMSE_te	0.1357	0.2881
R2_val	0.9087	0.8966
R2_te	0.9146	0.8953
MAE_val	0.1098	0.2281
MAE_testing	0.0966	0.2010

Παρατηρήθηκε, σύμφωνα και με τον πίνακα 3.4, πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης στην πρόβλεψη του SWHT με feature με SWHT_Sparse ήταν 49.3 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s = 25$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.11:



Σχήμα 3.11: Πρόβλεψη SWHT με SWHT_Sparse

Για την πρόβλεψη του MWH με feature με MWH_Sparse ο χρόνος που χρειάστηκε ήταν 48.36 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος που επιλέχτηκε με βάση το μικρότερο RMSE στο validation set ήταν το $s = 25$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.12:



Σχήμα 3.12: Πρόβλεψη MWH με feature MWH_Sparse

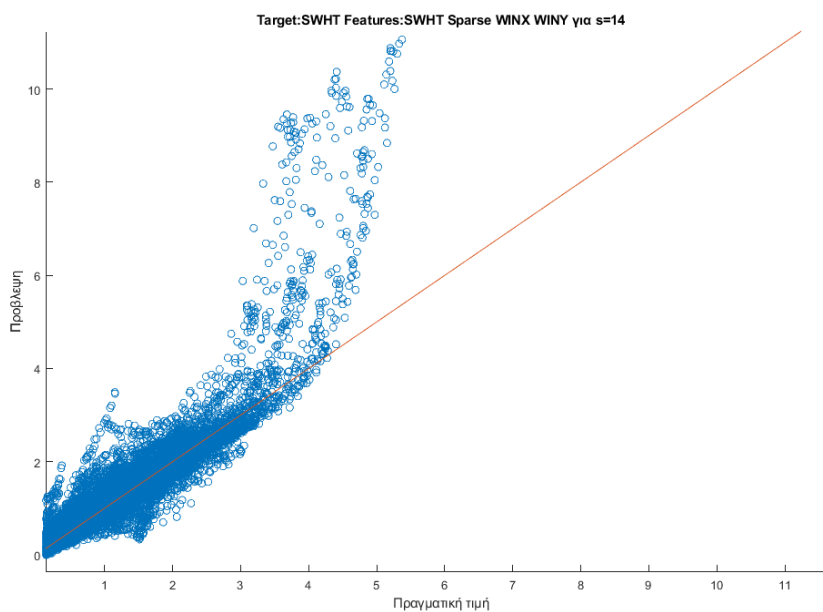
Στην συνέχεια, πραγματοποιήθηκε η δοκιμή πρόβλεψης του μέγιστου ύψους κύματος παραλλάσσοντας το ποια features έμπαιναν σαν είσοδος στο μοντέλο. Οι δοκιμές αυτές και τα αποτελέσματά τους φαίνονται στον πίνακα 3.8:

Πίνακας 3.8: Σενάρια με αφαίρεση feature-2

	Target:SWHT Feature:SWHT_Sparse WINX WINY	Target:SWHT Feature:SWHT_Sparse DP
s: fuzzy partition	4 έως 24	4 έως 50
Βέλτιστο s	14	34
Training time	841.09 sec	2078.67 sec
Αριθμός κέντρων	105	164
RMSE_val	0.1615	0.1506
RMSE_te	0.1278	0.1116
R2_val	0.9327	0.9415
R2_te	0.9242	0.9422
MAE_val	0.1022	0.0894
MAE_testing	0.00930	0.00764

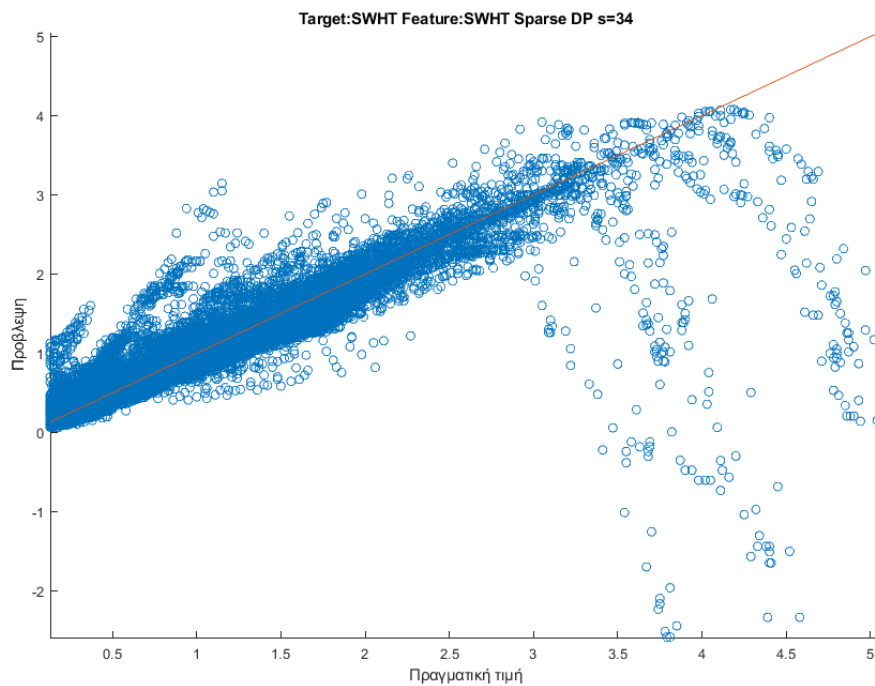
Όσο λιγότερα χαρακτηριστικά προσθέτονταν σαν είσοδοι στο μοντέλο τόσο μειωνόταν και ο υπολογιστικός χρόνος που χρειαζόταν να ολοκληρωθεί. Εκτός από αυτό όμως, μειωνόταν και η ακρίβεια του. Κάτι τέτοιο μπορεί να γίνει αντιληπτό από τα παρακάτω διαγράμματα και από τους στατιστικούς δείκτες του πίνακα 3.8.

Στο διάγραμμα 3.13 φαίνεται η πρόβλεψη του μέγιστου ύψους κύματος με features τα SWHT_Sparse, WINX και WINY για το βέλτιστο s το οποίο επιλέχτηκε:



Σχήμα 3.13: Πρόβλεψη SWHT με features SWHT_Sparse, WINX, WINY

Στο διάγραμμα 3.14 φαίνεται η πρόβλεψη του SWHT με features SWHT_Sparse και DP για το βέλτιστο s το οποίο επιλέχτηκε:



Σχήμα 3.14: Πρόβλεψη SWHT με features SWHT_Sparse ,DP

Στην προσπάθεια να βελτιωθούν τα αποτελέσματα των σεναρίων προστέθηκαν παραπάνω εισόδους στο μοντέλο. Αυτό παρουσιάζεται στον πίνακα 3.9:

Πίνακας 3.9: Σενάριο με αφαίρεση feature-3

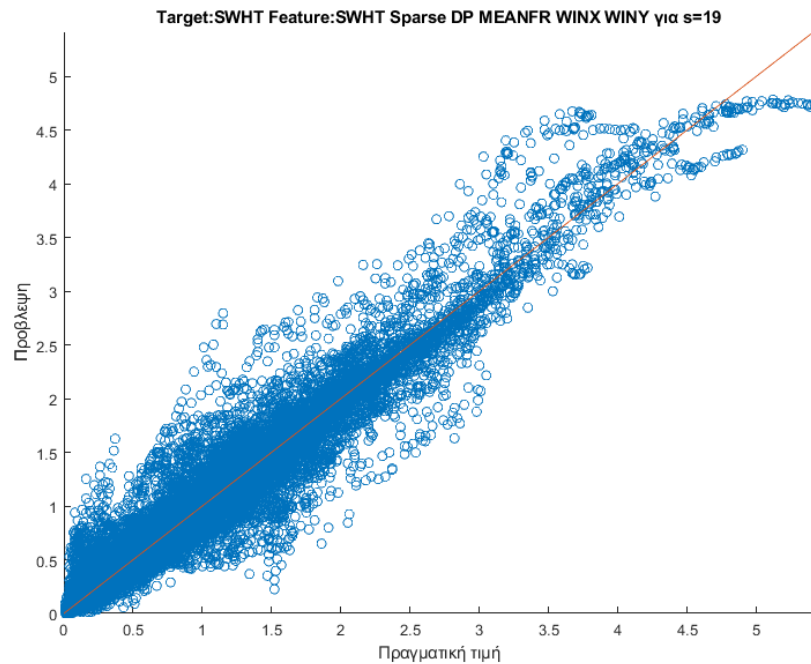
	Target:SWHT Feature:SWHT_Sparse PKFR WINX WINY	Target:SWHT Feature:SWHT_Sparse DP MEANFR	Target:SWHT Feature:SWHT_Sparse DP MEANFR WINX WINY
s: fuzzy partition	4 έως 24	4 έως 24	4 έως 20
Βέλτιστο s	12	23	19
Training time	1339.49 sec	706.91 sec	1716.22 sec
Αριθμός κέντρων	94	265	760
RMSE_val	0.1671	0.1562	0.1334
RMSE_te	0.1266	0.111	0.0999
R2_val	0.9280	0.9370	0.9541
R2_te	0.9256	0.9428	0.9538
MAE_val	0.1007	0.0912	0.0766
MAE_testing	0.0905	0.0763	0.0674

Σύμφωνα και με τον πίνακα 3.9 ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης στην πρόβλεψη του SWHT με features με SWHT_Sparse, PKFR, WINX και WINY ήταν 1339.49 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set ήταν το $s = 12$.

Στην πρόβλεψη του SWHT με features με SWHT_Sparse, DP και MEANFR η εκπαίδευση ολοκληρώθηκε σε 706.91 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν το $s=23$. Τέλος, με τα χαρακτηριστικά SWHT_Sparse, DP, MEANFR, WINX και WINY χρειάστηκαν 1716.22 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν $s = 19$.

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Από τα σενάρια αυτά σχεδιάστηκε το διάγραμμα που έδινε τα καλύτερα αποτελέσματα. Αυτό ήταν το τρίτο σενάριο, διότι σε σύγκριση με τα άλλα δύο έδινε μικρότερο RMSE και MAE, ενώ το R2 πλησίαζε πιο πολύ την μονάδα τόσο στο validation όσο και το testing dataset. Για το σενάριο αυτό σχεδιάστηκε το διάγραμμα 3.15:



Σχήμα 3.15: Πρόβλεψη SWHT με features SWHT_Sparse, DP, MEANFR, WINX, WINY

Ιδιαίτερο ενδιαφέρον παρουσίαζε το τρέξιμο των ίδιων σεναρίων με τον πίνακα 3.9 με την μόνη διάφορα να έγκειται στο ότι προβλέφθηκε το MWH και αντί του SWHT_Sparse χρησιμοποιήθηκε σαν είσοδο το MWH_Sparse. Τα αποτελέσματα φαίνονται στον πίνακα 3.10:

Πίνακας 3.10: Σενάρια με αφαίρεση feature-4

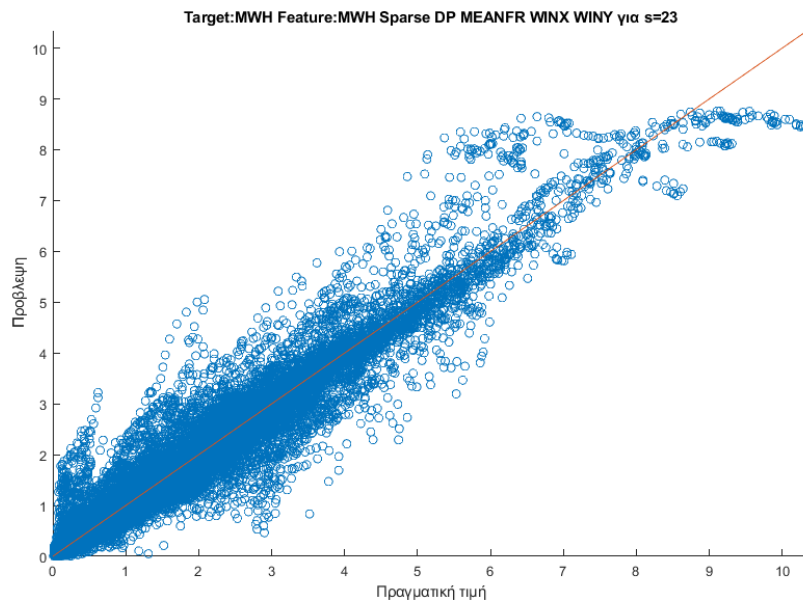
	Target:MWH Feature:MWH_Sparse PKFR WINX WINY	Target:MWH Feature:MWH_Sparse DP MEANFR	Target:MWH Feature:MWH_Sparse DP MEANFR WINX WINY
s: fuzzy partition	4 έως 24	4 έως 24	4 έως 24
Βέλτιστο s	15	23	23
Training time	1408.80 sec	739.58 sec	3309.83
Αριθμός κέντρων	178	267	1384
RMSE_val	0.3306	0.3062	0.2621
RMSE_te	0.2670	0.2260	0.1781
R2_val	0.9210	0.9322	0.9503
R2_te	0.9100	0.9355	0.9600
MAE_val	0.2023	0.1796	0.1409
MAE_testing	0.1870	0.1541	0.1133

Σύμφωνα με τον πίνακα 3.10, ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης στην πρόβλεψη του MWH με features τα MWH_Sparse, PKFR, WINX και WINY ήταν 1408.80 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος επιλέχτηκε με βάση το μικρότερο RMSE στο validation set πως ήταν το $s = 15$.

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Η πρόβλεψη του MWH με features τα MWH_Sparse, DP και MEANFR ολοκληρώθηκε σε 739.58 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν το $s = 23$. Τέλος, με τα χαρακτηριστικά MWH_Sparse, DP, MEANFR, WINX και WINY χρειάστηκαν 3309.83 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν $s = 23$.

Από τα σενάρια αυτά σχεδιάστηκε το διάγραμμα που έδινε τα καλύτερα αποτελέσματα. Αυτό ήταν το τρίτο σενάριο διότι σε σύγκριση με τα άλλα δύο έδινε μικρότερο RMSE και MAE, ενώ το R2 πλησίαζε πιο πολύ την μονάδα τόσο στο validation όσο και το testing dataset. Για το σενάριο αυτό σχεδιάστηκε το διάγραμμα 3.16:



Σχήμα 3.16: Πρόβλεψη MWH με features MWH_Sparse, DP, MEANFR, WINX, WINY

3.2.6 Σενάρια με χρήση της βαθμετρίας των διπλανών σημείων

Βασιζόμενοι στην ανάγκη βελτίωσης της ακρίβειας των προβλέψεων και στο feature engineering, γεννήθηκε η ιδέα της χρήσης της βαθμετρίας των διπλανών σημείων. Συγκεκριμένα, το νευρωνικό δίκτυο θα λάβει σαν είσοδο χαρακτηριστικά από ένα σημείο δίπλα σε αυτό που γίνεται η πρόβλεψη και αυτό θα συμβεί διότι η πληροφορία αυτή μπορεί να είναι σημαντική. Για να επιτευχθεί αυτό χρησιμοποιήθηκαν τα σημεία του dense grid, έτσι ώστε να βελτιωθεί η πρόβλεψη στα κοινά σημεία του sparse και του dense. Με άλλα λόγια το ζητούμενο ήταν η πρόβλεψη του ύψους κύματος με την η χρήση της βαθμετρίας του πυκνού πλέγματος από τα διπλανά σημεία του αραιού.

Ουσιαστικά, έπρεπε να δημιουργηθεί ένα νέο dataset το οποίο θα έχει τα features από τα κοινά σημεία του αραιού και του πυκνού πλέγματος και, την βαθμετρία του πυκνού σε σημεία που θα επιλεγθούν. Τα δεδομένα αυτά έπειτα θα χρησιμοποιηθούν στην πρόβλεψη του μέγιστου ύψους κύματος του πυκνού πλέγματος. Οι τρόποι με τους οποίους δυνητικά θα επιλεγθούν τα διπλανά σημεία ποικίλουν, καθώς οι συνδυασμοί είναι πολλοί. Για να γίνει η εξαγωγή των σημείων ακολουθήθηκε ο παρακάτω συλλογισμός:

Κάθε σημείο του πλέγματος περιγράφεται από γεωγραφικό μήκος (LON) και το γεωγραφικό πλάτος (LAT). Για να βρεθεί το κοντινότερο σημείο του dataset του πυκνού πλέγματος σε αυτό του αραιού, έπρεπε να γίνει η σύγκριση των δύο συντεταγμένων των δύο dataset. Κρατώντας ίδια την τιμή του γεωγραφικού πλάτους βρέθηκαν ποια σημεία του πυκνού πλέγματος είναι πιο κοντά στην τιμή του γεωγραφικού μήκους των σημείων του dense-sparse.

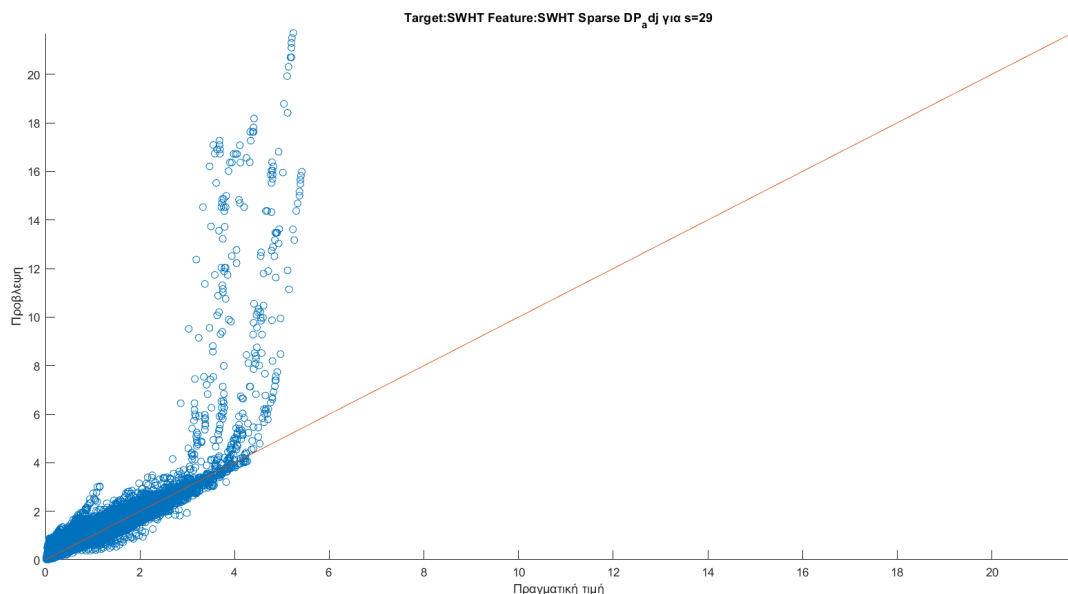
Δεδομένου ότι το σημείο του dense έχει την ίδια ημερομηνία και ώρα με αυτή του dense-sparse grid.

Αρχικά, ένα σενάριο το οποίο εκτελέστηκε, κάνοντας πρόβλεψη του μέγιστου ύψους κύματος, ήταν αυτό στο οποίο δόθηκαν σαν είσοδοι στο νευρωνικό δίκτυο το μέγιστο ύψος κύματος του αραιού πλέγματος και η βαθυμετρία των διπλανών και στην συνέχεια προστέθηκαν και τα χαρακτηριστικά της βαθυμετρία του πυκνού πλέγματος και των διπλανών σημείων και η μέση συχνότητα. Τα αποτελέσματα φαίνονται στον πίνακα 3.11:

Πίνακας 3.11: Σενάρια με χρήση της βαθυμετρία των διπλανών σημείων -1

	Target: SWHT Features: SWHT_Sparse DP_adj	Target:SWHT Features:SWHT_Sparse DP DP ADJ MEANFR
s: fuzzy partition	4 έως 50	4 έως 20
Βέλτιστο s	29	15
Training time	1303.96 sec	505.38 sec
Αριθμός κέντρων	101	149
RMSE_val	0.1476	0.1435
RMSE_te	0.1107	0.0927
R2_val	0.9438	0.9469
R2_te	0.9431	0.9601
MAE_val	0.0893	0.0818
MAE_testing	0.0752	0.0645

Για το πρώτο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του SWHT με την χρήση των features SWHT_Sparse και DP_adj, αυτό που παρατηρήθηκε σύμφωνα και με τον πίνακα 3.11 ήταν πως ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 1303.96 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος που επιλέχτηκε με βάση το μικρότερο RMSE στο validation set ήταν το $s = 29$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.17:

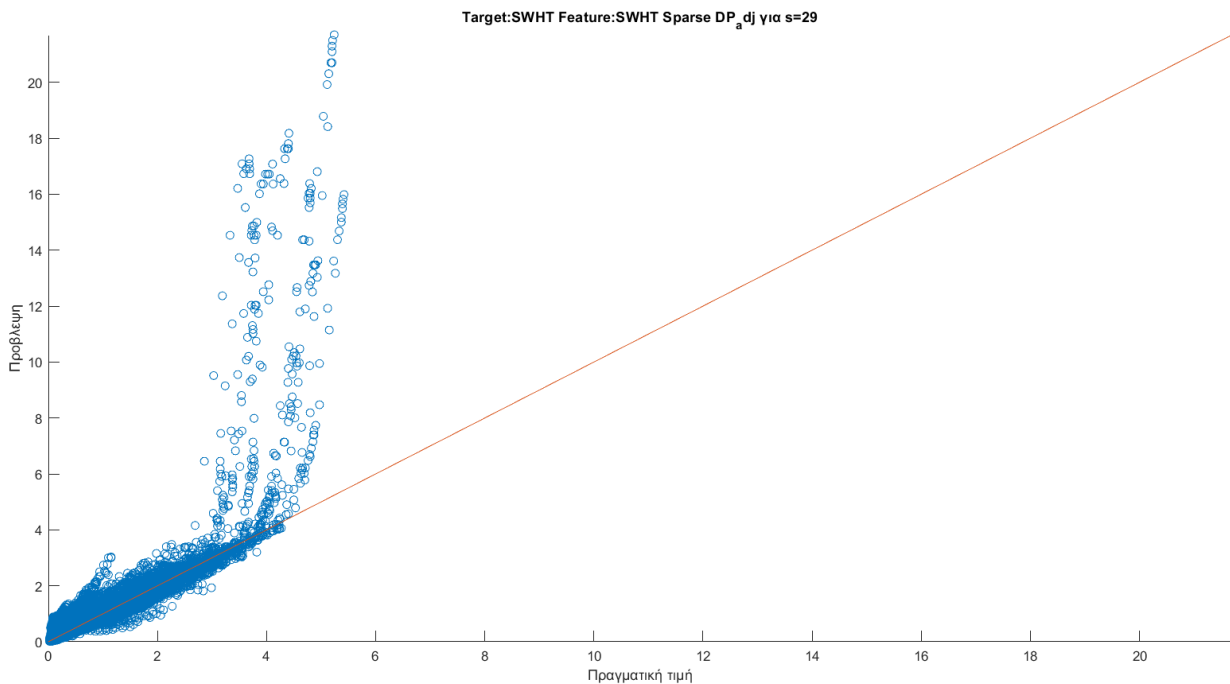


Σχήμα 3.17: Πρόβλεψη SWHT με features SWHT_Sparse και DP_adj

Για το δεύτερο σενάριο, δηλαδή το σενάριο που γίνεται πρόβλεψη του μέσου ύψους κύματος με την χρήση των features SWHT_Sparse, DP, DP_adj και MEANFR, παρατηρήθηκε, σύμφωνα με τον πίνακα 3.11, πως ο χρόνος για την ολοκλήρωση της εκπαίδευσης του μοντέλου ήταν 505.38 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος που επιλέχτηκε με ΠΑΔΑ, Τμήμα Η&ΗΜ, Διπλωματική Εργασία, Αποστολάκη Ελένη

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

βάση το μικρότερο RMSE στο validation set ήταν το $s = 15$. Για το s αυτό σχεδιάστηκε το διάγραμμα 3.18:



Σχήμα 3.18: Πρόβλεψη SWHT με features SWHT_Sparse, DP, DP_adj και MEANFR

Ιδιαίτερο ενδιαφέρον παρουσιάζει το τρέξιμο σεναρίων με την χρήση περισσότερων εισόδων καθώς, όπως είχε αποδειχτεί και παραπάνω, περισσότεροι εισοδοί αυξάνουν την ακρίβεια των προβλέψεων. Τα αποτελέσματα των σεναρίων παρουσιάζονται στον πίνακα 3.12:

Πίνακας 3.12: Σεσάρια με χρήση της βαθμετρία των διπλανών σημείων -2

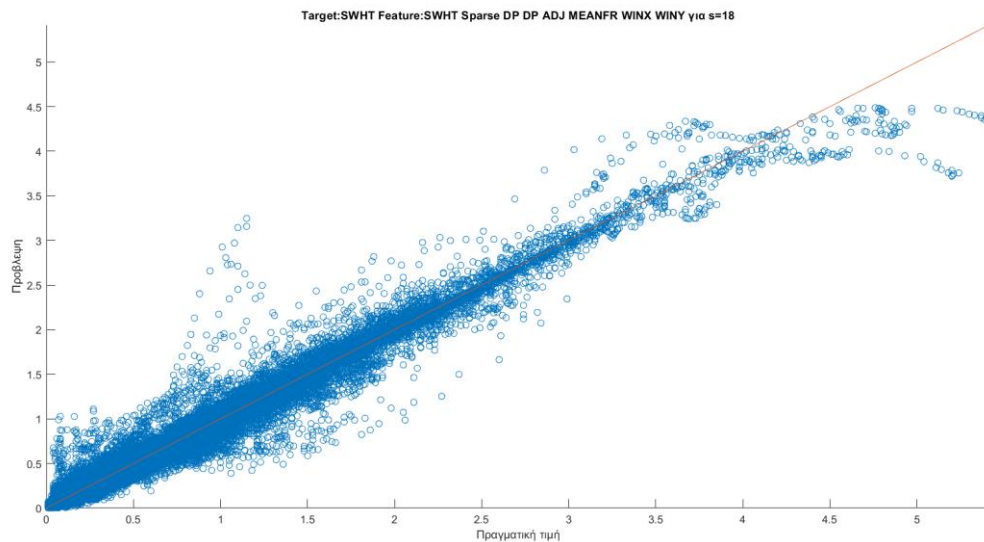
	Target:MWH Feature:MWHT Sparse DP DP ADJ MEANFR WINX WINY	Target:SWHT Feature:SWHT- MWH Sparse DP DP ADJ MEANFR WINX WINY	Target:SWHT Feature:SWHT Sparse DP DP ADJ MEANFR WINX WINY
s: fuzzy partition	4 έως 20	4 έως 20	4 έως 20
Βέλτιστο s	19	20	18
Training time	1665.63 sec	2181.53 sec	1690.98 sec
Αριθμός κέντρων	995	1081	828
RMSE_val	0.1891	0.1071	0.0997
RMSE_te	0.1318	0.0701	0.0683
R2_val	0.9741	0.9704	0.9743
R2_te	0.9781	0.9772	0.9784
MAE_val	0.1056	0.0558	0.0547
MAE_testing	0.0852	0.0437	0.0431

Σύμφωνα με τον πίνακα 3.12, ο χρόνος που χρειάστηκε για την ολοκλήρωση της εκπαίδευσης του μέγιστου ύψους κύματος με χαρακτηριστικά τα MWH_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY ήταν 1665.63 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος, που επιλέχτηκε με βάση το μικρότερο RMSE, στο validation set ήταν $s = 19$.

Μοντελοποίηση μετεωρολογικών δεδομένων με νευρωνικά δίκτυα

Η πρόβλεψη του SWHT με features με SWHT και MWH_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY ολοκληρώθηκε σε 2181.53 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν το $s=20$. Τέλος, για την πρόβλεψη του μέσου ύψους κύματος με τα χαρακτηριστικά SWHT_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY χρειάστηκαν 1690.98 δευτερόλεπτα και ο βέλτιστος ασαφής υποχώρος ήταν $s = 18$.

Από τα σενάρια αυτά σχεδιάστηκε το διάγραμμα που έδινε τα καλύτερα αποτελέσματα. Αυτό ήταν το τρίτο σενάριο διότι σε σύγκριση με τα άλλα δύο έδινε μικρότερο RMSE και MAE, ενώ το R2 πλησίαζε πιο πολύ την μονάδα τόσο στο validation όσο και το testing dataset. Το σενάριο αυτό σχεδιάστηκε το διάγραμμα 3.19:



Σχήμα 3.19: Πρόβλεψη SWHT με features SWHT_Sparse, DP, DP_ADJ, MEANFR, WINX και WINY

3.3 Σύγκριση αποτελεσμάτων

Αφού ολοκληρώθηκε το κομμάτι του τρεξίματος των σεναρίων έγινε μια προσπάθεια σύγκρισης τόσο των σεναρίων ίδιας κατηγορίας όσο και διαφορετικής. Το βέλτιστο σενάριο είναι αυτό στο οποίο το R2 πλησιάζει την μονάδα και τα RMSE και MAE είναι όσο το δυνατόν μικρότερα. Ένας ακόμα δείκτης για το πόσο καλά είναι τα αποτελέσματα του σεναρίου είναι τα διαγράμματα που στον άξονα x σχεδιάστηκαν οι πραγματικές τιμές ενώ στον άξονα y οι προβλέψεις. Συγκεκριμένα, όσο πιο πολύ συγκλίνουν τα δεδομένα στην ευθεία με κλίση 45 μοιρών τόσο καλύτερο θεωρείται το μοντέλο. Πέραν των στατιστικών δεικτών και των διαγραμμάτων, σημαντικό ρόλο παίζει επίσης και ο χρόνος ολοκλήρωσης του μοντέλου μηχανικής μάθησης. Όσο πιο μεγάλος είναι ο χρόνος αυτός τόσο αυξάνεται και το υπολογιστικό κόστος του μοντέλου, πράγμα το οποίο σημαίνει ότι μεγαλώνουν οι απαιτήσεις στην επεξεργαστική ισχύς και στο μέγεθος της μνήμης. Τέλος, ο αριθμός των κέντρων επηρεάζει τα αποτελέσματα των σεναρίων, όσο μικρότερος είναι ο αριθμός αυτός τόσο το καλύτερο. Από τους παραπάνω πίνακες έγινε φανερό πως όσο περισσότερα κέντρα προέκυψαν από τους υπολογισμούς του αλγόριθμου Fuzzy Means, τόσο αυξάνονταν και ο υπολογιστικός χρόνος που απαιτούνταν για την εκπαίδευση και την πρόβλεψη του μοντέλου.

Σχετικά με τα σενάρια στα οποία χρησιμοποιήθηκαν όλα τα features, σύμφωνα με τον πίνακα 3.4 παρατηρείται ότι το βέλτιστο σενάριο εκ των δύο είναι αυτό στο οποίο προβλέπεται το μέσο ύψος κύματος. Ειδικότερα, παρόλο που το σενάριο της πρόβλεψης του MWH χρειάζεται λιγότερο χρόνο για να ολοκληρωθεί (1377.81 seconds), ενώ αυτό του SWHT χρειάζεται περισσότερο (1403.36 seconds), το μοντέλο πρόβλεψης του SWHT παρουσιάζει καλύτερα

αποτελέσματα στους στατιστικούς δείκτες MAE και RMSE στο testing σετ. Επιπλέον, το R2 του SWHT πλησιάζει λιγότερο την μονάδα (0.9402) από το αντίστοιχο του MWH (0.9450), η διαφορά αυτή ωστόσο θεωρείται σχεδόν αμελητέα. Τέλος, φαίνεται και από τα διαγράμματα 3.4 και 3.5 πως τα δεδομένα που προέρχονται από την πρόβλεψη του SWHT συγκλίνουν περισσότερο στην ευθεία με κλίση 45 μοιρών από τα αντίστοιχα του MWH.

Η επόμενη κατηγορία σεναρίων η οποία εκτελέστηκε είναι αυτή με τα feature engineering στην ταχύτητα του ανέμου. Σύμφωνα με τον πίνακα 3.5, στα σενάρια στα οποία δεν χρησιμοποιήθηκαν τα χαρακτηριστικά της ταχύτητας και της κατεύθυνσης του ανέμου, αλλά αντί αυτών οι συνιστώσες WINX και WINY, ο χρόνος ολοκλήρωσης ήταν σαφώς μικρότερος από τα αντίστοιχα σενάρια στα οποία χρησιμοποιήθηκαν όλα τα features. Μεταξύ των σεναρίων του πίνακα 3.4 καλύτερο θεωρείται αυτό που προβλέπει το μέσο ύψος κύματος, καθώς το RMSE και το MAE είναι μικρότερα από αυτά του σεναρίου στο οποίο προβλέπεται το μέγιστο ύψος κύματος. Τέλος, το R2 στην περίπτωση αυτή είναι επίσης καλύτερο από αυτό του MWH. Μεταξύ των σεναρίων πρόβλεψης SWHT των πινάκων 3.4 και 3.5, στο σενάριο πρόβλεψης με την χρήση όλων των features τα στατιστικά σφάλματα MAE και RMSE είναι ελαφρώς καλύτερα από το σενάριο με την χρήση των features WINX και WINY.

Στο πίνακα 3.6 φαίνονται τα αποτελέσματα από το σενάριο που δέχτηκε ως είσοδο όλα τα features και τις συνιστώσες WINX και WINY. Ο χρόνος ολοκλήρωσης της εκπαίδευσης έφτασε στα επίπεδα που ήταν και στο σενάριο της πρόβλεψης όλων των features. Μεταξύ των σεναρίων του πίνακα 3.6 καλύτερο είναι το μοντέλο πρόβλεψης του SWHT λόγω του ότι το MAE και το RMSE είναι μικρότερα από τα αντίστοιχα του MWH και το R2 είναι πιο κοντά στην μονάδα από το αντίστοιχο του MWH. Συγκρίνοντας τώρα τους πίνακες 3.5 και 3.6, δηλαδή τα σενάρια που χρησιμοποιήθηκαν οι συνιστώσες WINX και WINY με και χωρίς την κατεύθυνση και την ταχύτητα του ανέμου, καλύτερο είναι αυτό που προβλέφθηκε το μέσο ύψος κύματος με την χρήση όλων των features, των συνιστωσών WINX, WINY και όχι των WINSP και WINDDIR. Αυτό αποδεικνύεται αν παρατηρηθούν οι στατιστικοί δείκτες σφάλματος των δύο πινάκων.

Σχετικά με την κατηγορία σεναρίων με αφαίρεση features παρατηρήθηκε, όπως ήταν και λογικό, πως όσο λιγότερα χαρακτηριστικά έμπαιναν σαν είσοδοι στο μοντέλο τόσο μειωνόταν ο χρόνος εκτέλεσης αλλά και η ακρίβεια των προβλέψεων. Ένας τύπος σεναρίου που παρουσίασε ιδιαίτερο ενδιαφέρον είναι αυτός του πίνακα 3.7. Ειδικότερα, έγινε μια προσπάθεια πρόβλεψης του μέγιστου και του μέσου ύψους κύματος του πυκνού grid δίνοντας μόνο σαν είσοδο στο μοντέλο το μέγιστο και το μέσο ύψος κύματος του αραιού grid. Μεταξύ των δύο καλύτερη πρόβλεψη είναι αυτή του SWHT. Στα σενάρια αυτά οι στατιστικοί δείκτες ήταν πολύ χειρότεροι από αυτούς των προηγούμενων σεναρίων των πινάκων 3.4 έως 3.5. Αυτό που έγινε αμέσως αντιληπτό είναι η σημασία των υπόλοιπων features στην πρόβλεψη του ύψους κύματος.

Συνεχίζοντας με την πρόβλεψη του μέσου ύψους κύματος στον πίνακα 3.8 φαίνονται δύο σενάρια στα οποία προστέθηκαν κάποια features σαν είσοδοι στο μοντέλο. Παρόλο που τα αποτελέσματα βελτιώθηκαν αισθητά, κανένα από τα δύο σενάρια δεν έφτασε την ακρίβεια που έφτασαν τα σενάρια στα οποία χρησιμοποιήθηκαν όλα τα features.

Επιπλέον, δοκιμάστηκαν κάποια ακόμα σενάρια για την πρόβλεψη του μέσου και του μέγιστου ύψους κύματος τα οποία παρουσιάζονται στους πίνακες 3.9 και 3.10. Τα σενάρια της πρόβλεψης του SWHT είναι πιο ακριβή από αυτά του MWH. Με βάση το πόρισμα αυτό λοιπόν, δόθηκε έμφαση στον πίνακα 3.9. Εκ των τριών σεναρίων, καλύτερα αποτελέσματα παρουσίασε το μοντέλο που δέχτηκε σαν εισόδους το SWHT_Sparse, την βαθυμετρία, την μέση συχνότητα και τις συνιστώσες WINX και WINY.

Τέλος, στα σενάρια στα οποία χρησιμοποιήθηκε σαν χαρακτηριστικό στην είσοδο του μοντέλου η βαθυμετρία των διπλανών σημείων, παρατηρήθηκε από τους πίνακες 3.10 και 3.11 ότι η χρήση περισσότερων εισόδων βελτιώνει τα αποτελέσματα του μοντέλου. Οι στατιστικοί δείκτες και τα διαγράμματα έδειξαν πως η πρόβλεψη του μέσου ύψους κύματος ήταν ικανοποιητική, ενώ τα αποτελέσματα της πρόβλεψης του μέγιστου ύψους κύματος όπως και στα υπόλοιπα σενάρια παρουσίαζε χειρότερους στατιστικούς δείκτες σφάλματος.

Συμπερασματικά, αυτό που παρατηρήθηκε ήταν πως με την χρήση όλων των εισόδων το μοντέλο έδωσε τα καλύτερα αποτελέσματα από την άποψη των στατιστικών δεικτών και των διαγραμμάτων. Ο χρόνος εκπαίδευσης του μοντέλου όμως ήταν μεγάλος. Προσθέτοντας τις συνιστώσες WINX και WINY οι στατιστικοί δείκτες σφάλματος εξακολουθούσαν να παρουσιάζουν παρόμοιες τιμές με τα προηγούμενα σενάρια, ο χρόνος εκπαίδευσης του μοντέλου μηχανικής μάθησης όμως μειώθηκε. Τέλος, στα σενάρια με την αφαίρεση χαρακτηριστικών, παρόλο που αυτά παρουσίασαν τα χειρότερα αποτελέσματα από την άποψη των στατιστικών δεικτών, οι χρόνοι εκπαίδευσης μειώθηκαν αισθητά. Ανάλογα λοιπόν την ακρίβεια η οποία είναι αναγκαίο να επιτευχθεί για την εκάστοτε εφαρμογή γίνεται χρήση της αντίστοιχης κατηγορίας σεναρίων, με αυτά της χρήσης όλων των features και αυτά με feature engineering στην ταχύτητα του ανέμου να παρουσιάζουν την πιο ακριβή πρόβλεψη.

ΚΕΦΑΛΑΙΟ 4^ο : Συμπεράσματα

Η πρόβλεψη του ύψους κύματος με τα συμβατικά μοντέλα παρουσίασε μειονεκτήματα όσον αφορά την ακρίβεια και τους υπολογιστικούς χρόνους των προβλέψεων. Συγκεκριμένα, τόσο τα κυματικά μοντέλα πρώτης όσο και δεύτερης γενιάς μπορούν να ρυθμιστούν ώστε να παρέχουν χρήσιμα αποτελέσματα για ορισμένες κατηγορίες πεδίων, ωστόσο έχει αποδειχθεί ότι κανένα από τα μοντέλα αυτά δεν είναι αξιόπιστο για ακραίες καταστάσεις για τις οποίες απαιτούνται συχνά οι προβλέψεις κύματος. Από την άλλη πλευρά, τα μοντέλα τρίτης γενιάς μπορούν να προβλέψουν με ακρίβεια τα χαρακτηριστικά κύματος ωστόσο είναι πολύ απαιτητικά υπολογιστικά ακόμα και για τους σημερινούς υπολογιστές με αποτέλεσμα να χρειάζονται πολύ χρόνο για την παραγωγή μιας πρόβλεψης. Όλα τα παραπάνω οδήγησαν στην χρήση των νευρωνικών δικτύων στην πρόβλεψη του μέσου ή του μέγιστου ύψους κύματος. Με την ανάπτυξη της τεχνητής νοημοσύνης και λόγω των πλεονεκτημάτων της (γρήγορη ταχύτητα υπολογισμού, χαμηλό υπολογιστικό κόστος και ισχυρή μη γραμμική ικανότητα εκμάθησης) η μέθοδος πρόβλεψης SWHT που βασίζεται στη μηχανική μάθηση εφαρμόστηκε σε πολλές μελέτες τα τελευταία χρόνια.

Στην παρούσα διπλωματική εργασία, χρησιμοποιήθηκαν δεδομένα τα οποία προέρχονταν από την επίλυση μοντέλων σε ορθογώνια δομημένα πλέγματα στην περιοχή του Αγίου Νικολάου στην Κρήτη. Κάθε πλέγμα αποτελείται από έναν αριθμό κόμβων τα οποία ονομάζονται σημεία του πλέγματος. Η απόσταση των σημείων αυτών καθορίζει το αραιό ή το πυκνό πλέγμα. Με βάση λοιπόν τα δεδομένα αυτά πραγματοποιήθηκε η πρόβλεψη του μέσου και του μέγιστου ύψους κύματος με την χρήση των νευρωνικών δικτύων ακτινικής βάσης. Στα δίκτυα RBF, εξαιτίας της δομής τους, ο προσδιορισμός των παραμέτρων περιλαμβάνει δύο φάσεις, πράγμα το οποίο καθιστά πιο εύκολο τον προγραμματισμό και την βελτιστοποίηση του μοντέλου μηχανικής μάθησης. Κατά την πρώτη φάση για τον υπολογισμό των κέντρων έγινε η χρήση του αλγόριθμου των ασαφών μέσων. Ο αλγόριθμος αυτός απαιτεί μόνο ένα πέρασμα των δεδομένων εισόδου, πράγμα το οποίο καθιστά την διαδικασία υπολογισμού των κέντρων εξαιρετικά γρήγορη.

Στην είσοδο του νευρωνικού δικτύου χρησιμοποιήθηκαν τα χαρακτηριστικά του αραιού πλέγματος και έγινε μια προσπάθεια πρόβλεψης του SWHT και του MWH του πυκνού πλέγματος. Ανάλογα με τον τρόπο που τα χαρακτηριστικά δόθηκαν σαν είσοδοι στο νευρωνικό δίκτυο, τα σενάρια τα οποία εκτελέστηκαν χωρίστηκαν σε τέσσερις βασικές κατηγορίες :

- i. Σενάρια με την χρήση όλων των features
- ii. Σενάρια με την χρήση feature engineering στην ταχύτητα του ανέμου
- iii. Σενάρια με αφαίρεση χαρακτηριστικών
- iv. Σενάρια με χρήση της βαθμετρίας των διπλανών σημείων

Σε όλα τα σενάρια έγινε σαφές, από τους στατιστικούς δείκτες σφάλματος και τα διαγράμματα, πως η πρόβλεψη του μέσου ύψους κύματος έδινε καλύτερα αποτελέσματα από την πρόβλεψη του μέγιστου ύψους κύματος. Αυτό δικαιολογείται από το γεγονός του ότι είναι πιο εύκολο να προβλεφθεί η μέση από μέγιστη τιμή μιας μεταβλητής.

Το μοντέλο, που έδωσε τα καλύτερα αποτελέσματα από την άποψη των στατιστικών δεικτών και των διαγραμμάτων, προερχόταν από το σενάριο στο οποίο δόθηκαν σαν είσοδοι στο νευρωνικό δίκτυο όλα τα χαρακτηριστικά. Το μειονέκτημα όμως που παρουσίασε ήταν στον χρόνο ολοκλήρωσης της εκπαίδευσης καθώς αυτός ήταν μεγάλος. Παρόμοια αποτελέσματα έδωσαν και τα σενάρια με το feature engineering στην ταχύτητα του ανέμου καθώς οι στατιστικοί δείκτες εξακολούθησαν να παρουσιάζουν παρόμοιες τιμές, ο χρόνος ολοκλήρωσης της εκπαίδευσης του μοντέλου όμως μειώθηκε. Επιπρόσθετα, στα σενάρια αφαίρεσης χαρακτηριστικών παρατηρήθηκε ότι όσο λιγότερα χαρακτηριστικά προστίθονταν σαν είσοδοι στο μοντέλο τόσο μειωνόταν και ο υπολογιστικός χρόνος που χρειαζόταν να ολοκληρωθεί η

εκπαίδευση. Εκτός από αυτό όμως, μειωνόταν και η ακρίβεια του. Τέλος, στα σενάρια με την βαθυμετρία των διπλανών σημείων έγινε αντιληπτό πως με την χρήση πολλών εισόδων τα αποτελέσματα ήταν ικανοποιητικά, καθώς σε κάποιες περιπτώσεις οι στατιστικοί δείκτες είχαν τιμές καλύτερες και από αυτές των σεναρίων με την χρήση όλων των χαρακτηριστικών. Λόγω της χιλιομετρικής απόστασης των διπλανών σημείων, ωστόσο, δεν μπορεί να προσδιοριστεί με σιγουριά κατά πόσο η πρόβλεψη του ύψους κύματος θεωρείται ορθή.

Κλείνοντας, δεν μπορεί να αποφανθεί ποιο μοντέλο από τα παραπάνω θεωρείται βέλτιστο. Εξαρτάται από την ακρίβεια η οποία είναι αναγκαίο να επιτευχθεί για την εκάστοτε εφαρμογή. Αν για παράδειγμα κάποια εφαρμογή απαιτούσε υψηλή ακρίβεια, τότε το μοντέλο στο οποίο χρησιμοποιούνται σαν εισοδοί όλα τα χαρακτηριστικά του κύματος ή το μοντέλο με το feature engineering στην ταχύτητα του ανέμου, θα θεωρούνταν ιδανικό. Αντίθετα, αν η εφαρμογή απαιτούσε μικρότερο υπολογιστικό χρόνο εκπαίδευσης και πρόβλεψης τότε το μοντέλο το οποίο θα θεωρούνταν ιδανικό θα προέρχονταν από την κατηγορία σεναρίων με αφαίρεση χαρακτηριστικών.

Τέλος, αυτό που αξίζει να αναφερθεί και αποτελεί το βασικότερο συμπέρασμα της διπλωματικής εργασίας είναι ότι χρησιμοποιώντας τα μοντέλα μηχανικής μάθησης που προτείνονται, υπάρχει η δυνατότητα εξαγωγής μιας πολύ καλής προσέγγισης για το μέσο και το μέγιστο ύψος κύματος του πυκνού πλέγματος τρέχοντας μόνο το αραιό, με αποτέλεσμα σημαντική μείωση του υπολογιστικού χρόνου για την εκπαίδευση του μοντέλου. Αν και ο χρόνος που απαιτείται για την εκπαίδευση του μοντέλου ποικίλει και μπορεί να είναι μεγάλος ανάλογα με τις εισόδους που λαμβάνει το νευρωνικό δίκτυο, ο χρόνος για την πρόβλεψη είναι μηδαμινός σε σχέση με τον αντίστοιχο των κυματικών μοντέλων τρίτης γενιάς.

Βάση των παραπάνω κάποιες προτάσεις για μελλοντική έρευνα σχετίζονται με την χρήση των μοντέλων μηχανικής μάθησης σε εφαρμογές σε μεγαλύτερες περιοχές ή σε περιοχές με διαφορετικά γεωγραφικά χαρακτηριστικά. Τέλος, μπορούν να πραγματοποιηθούν δοκιμές με περισσότερα features σαν εισοδοί στο νευρωνικό δίκτυο. Για παράδειγμα θα μπορούσε πέρα από την βαθυμετρία των διπλανών σημείων του grid να χρησιμοποιηθεί σαν είσοδος και η ταχύτητα του ανέμου των διπλανών σημείων και να αποφανθεί αν το χαρακτηριστικό αυτό βελτιώνει τις προβλέψεις ή όχι.

Βιβλιογραφία – Αναφορές - Διαδικτυακές Πηγές

- [1] A.Culclasure, “Using Neural Networks to Provide Local Weather Forecasts”, Department of Computer Sciences, Georgia Southern University, 2013
- [2] H.D. Navone and H.D. Ceccatto, “Predicting Indian monsoon rainfall: a neural network approach,” *Climate Dynamics*, vol. 10, pp. 305-312, 1994.
- [3] N.Q. Hung, M.S. Babel, S. Weesakul, and N.K. Tripathi, “An artificial neural network model for forecasting in Bangkok, Thailand,” *Hydrology and Earth System Sciences*, vol. 13, no. 8, pp. 1413-1425, August 2009.
- [4] T. Santhanam and A.C. Subhajini, “An efficient weather forecasting system using radial basis function neural network,” *Journal of Computer Science*, vol. 7, no. 7, pp. 962-966, June 2011.
- [5] J-P Boulanger, F. Martinez, and E.C. Segura, “Neural network based daily precipitation generator (nngen-p),” *Climate Dynamics*, vol. 28, pp. 307-324, September 2006.
- [6] Z. Xing, H. Guo, S. Dong, Q. Fu, J. Li, “RBF Neural Network Based on K-means Algorithm with Density Parameter and its Application to the Rainfall Forecasting”, *Computer and Computing Technologies in Agriculture*, vol.8, pp.218-225, December 2015
- [7] M. Papadimitrakis, A. Alexandridis, “Active vehicle suspension control using road preview model predictive control and radial basis function networks”, *Applied Soft Computing*, vol.120, 108646, May 2022.
- [8] A. Alexandridis, H. Sarimveis and K. Ninos, “A Radial Basis Function network training algorithm using a non-symmetric partition of the input space – Application to a Model Predictive Control configuration”, *Advances in Engineering Software*, vol.42, pp. 830-837, June 2011
- [9] A. Balasubramanian, “Weather Forecasting”, Department of studies in Earth Science, University of Mysore, Mysore India, Technical report 5, February 2016
- [10] Britannica.(2022).*History of weather forecasting*[Online].
- [11] Sumanta Hait(2021). *Basics of Weather Forecasting* [Online].
- [12] A. J Iseh and T.Y Woma ,”Weather Forecasting Models, Methods and Applications”, *International Journal of Engineering Research & Technology* , Vol. 2 Issue 12, December 2013
- [13] V. K. Didal, Brijbhooshan1, A. Todawat and K. Choudhary, “Weather Forecasting in India: A Review”, *International Journal of Current Microbiology and Applied Sciences*, vol.6, pp.577-590, November 2017
- [14] J.R Holton, *An Introduction to Dynamic Meteorology*, 4rd ed., Elsevier Academic Press, 2004
- [15] Rk Naresh, S.Kumar and E.Narwal, *Artificial Intelligence in Indian Agriculture*, Jaya Publishing House, March 2021
- [16] F.W. Zwiers and H.V. Storch ,”On the role of statistics in climate research”,*International journal of climatology*, vol.24, pp.665-680, May 2004
- [17] J. Dutton(2021),*What is Analog Forecasting?*[Online].
- [18] Ι.Πυθαρούλης, “Συνοπτική και δυναμική μετεωρολογία-Εισαγωγή” [Online].
- [19] National Ocean Service(2021,February.26),*Why does the ocean have waves?*,[Online].
- [20] World Meteorological Organization, *Guide to Wave Analysis and Forecasting*, 2018 edition, Chair, Publications Board, 2018
- [21] CDIP(2022),*Wave Measurement*,[Online].
- [22] P.A.E.M. Janssen,“Progress in ocean wave forecasting”, *Journal of Computational Physics*, 227(7), pp.3572-3594, March 2008

- [23] K. Hasselmann, E. Bauer and P. A.E.M. Janssen , “The WAM model - a third generation ocean wave prediction model”, *Journal of Computational Physics*[Online],vol.18,pp.1775-1810, January 1988
- [24] E. Caetano and Valdir Innocentini, “An improved second generation wave model” ,*Rev. bras.oceanorg*,vol.51,pp.1-21,December 2002
- [25] H. Allender, T.P. Barnett, L. Bertotti, J. Bruinsma, V.J. Cardone, L. Cavaleri, and etc, Sea wave modeling project (SWAMP).An intercomparison study of wind wave predictions models, part 1: principal results and conclusions, in: *Ocean Wave Modeling*,Plenum Press, New York,1985
- [26] G.S.Dwarakish and J.Thomas T, “Numerical wave modelling – A review”, *Aquatic Procedia*,vol.4,pp.443– 448,December 2015
- [27] L.Calaveri, L.Bertotti and P. Lionello, “Shallow Water Application of the Third-Generation WAM Wave Model”,*Journal of Geophysical Research*,vol.94,pp.8111-8124,June 1989
- [28] O.R. Sørensen, H. Kofoed-Hansen, M. Rugbjerg and L. S. Sørensen , “A Third-Generation Spectral Wave Model Using An Unstructured Finite Volume Technique”, *Coastal Engineering* 2004 ,vol.4, pp. 894-906 , April 2005
- [29] N. Krishnakumar,R.Savitha and A.A.Mamun , “Regional ocean wave height prediction using sequential learning neural networks”,*Ocean Engineering*,vol.129, pp.605-612,January 2017
- [30] S.N Londhe and M.C Deo, “Wave tranquility studies using neural networks.” ,*Ocean Engineering* ,vol.16(6),pp.419-36, March 2003
- [31] R. Kalraa,M.C.Deo, R.Kumarb,V. K. Agarwalb , “RBF network for spatial mapping of wave heights”,*Marine Structures*,vol.18,pp.289-300,March 2005
- [32] O. Makarynskyya, A.A. Pires-SilvabD. Makarynskaa and C. Ventura-Soaresc, “Artificial neural networks in wave predictions at the west coast of Portugal”, *Computers and Geosciences*, Volume 31, Issue 4, pp. 415-424, May 2005
- [33] P. Hao,S. Li ,C. Yu and G. Wu, “A prediction model of Significant Wave Height in the South China Sea Based on Attention Mechanism”, *Frontiers in Marine Science*, 9:895212,June 2022
- [34] N. Portillo Juan and V.Negro Valdecantos , “Review of the application of Artificial Neural Networks in ocean engineering”, *Ocean Engineering*, Volume 259, 111947,September 2022
- [35] M.C.Deo and C.S. Naidua, “Real time wave forecasting using neural networks”,*Ocean Engineering* ,vol.26(2),pp.191-203, August 1998
- [36] M.C. Deo, A. Jha, A.S. Chaphekar, K. Ravikant, “Neural networks for wave forecasting”, *Ocean Engineering* ,vol.28(7),pp.889-898, July 2001
- [37] M.C. Deo, and S.S.Jagdale, “Prediction of breaking waves with neural networks”, *Ocean Engineering*, vol .30 (9), pp,1163–1178, November 2002
- [38] M. Tsuda, S. Kim and Y. Matsumi, “Wave forecasts using an Artificial Neural Network for port construction works management “,*Journal of Japan Society of Civil Engineers*,vol.73,pp.151-156,October 2017
- [39] X. Zhang, Y. Li,S. Gao and P.Ren, “Ocean Wave Height Series Prediction with Numerical Long Short-Term Memory”, *Journal of Marine Science and Engineering*, 9(5):514,May 2021
- [40] S. Haykin, *Νευρωνικά Δίκτυα και Μηχανική Μάθηση*, Τρίτη έκδοση, July 2010, Εκδόσεις Παπασωτηρίου

- [41] W. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943.
- [42] K. Fukushima, “Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position”, *Biological Cybernetics*, vol.36, pp.193-202, April 1980
- [43] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, pp. 386–408, 1958.
- [44] B. Widrow, M.E. Hoff, “Adaptive switching circuits,”1960 IRE WESCON Convention Record, New York: IRE Part 4, pp. 96–104, 1960.
- [45] M.J.D. Powell, “Radial basis functions for multivariable interpolation: a review,” *Algorithms for Approximation*, pp. 143-167, Oxford, 1987.
- [46] M.T.Hagan,H. B.Demuth and M. H. Beale, *Neural Network Design*,Second edition,Martin Hagan,2014
- [47] H. Hassana, Z. Ren, H. Zhao and etc , “Review and classification of AI-enabled COVID-19 CT imaging models based on computer vision tasks,Computers in Biology and Medicine, vol.141, December 2021
- [48] A. Messaoudi, H. Haddad, C.Fourati and etc, “Tunisian Dialectal End-to-end Speech Recognition based on DeepSpeech”, *Procedia Computer Science*, vol.189, pp.183-190, May 2021
- [49] S.M.Singh and T. D. Singh, “Low resource machine translation of english–manipuri: A semi-supervised approach”, *Expert Systems with Applications*, vol.209,pp. 118-187 December 2022
- [50] O.B. Sezer, M.U. Gudelek and A.M. Ozbayoglu, “Financial Time Series Forecasting with Deep Learning : A Systematic Literature Review: 2005-2019”, *Applied Soft Computing*, vol.90, 106181, May 2020
- [51] M.C. Popesc,V.E.Balas,N.E.Mastorakis and L.P. Popescu,”Multilayer perceptron and neural networks”,*WSEAS Transactions on Circuits and Systems*, vol.8, Issue , July 2009
- [52] D.E. Rumelhart, G.E. Hinton and R.J. Williams , “Learning representations by back-propagating errors”,*Nature* ,vol.323, pp.533-536, October 1986
- [53] P. Auera, H. Burgsteinerb and W. Maassc, “A learning rule for very simple universal approximators consisting of a single layer of perceptron”, *Neural Networks*, vol.21, pp.786-795, June 2008
- [54] D.S. Broomhead and D. Lowe, “Multivariable Functional Interpolation and Adaptive Networks”,*Complex Systems*, vol.2 , pp.321-355, 1988
- [55] H.B. Celikoglu, “Application of radial basis function and generalized regression neural networks in non-linear utility function specification for travel mode choice modelling”, *Mathematical and Computer Modelling*,vol.44,pp.640-658,October 2006
- [56] N. Korovesis, D. Kandris, G. Koulouras, A. Alexandridis, “Robot Motion Control via an EEG-Based Brain–Computer Interface by Using Neural Networks and Alpha Brainwaves”, *Electronics*, vol.8, 1387, November 2019.
- [57] M. Papadimitrakis, M. Stogiannos, H. Sarimveis, A. Alexandridis, “Multi-ship control and collision avoidance using MPC and RBF-based trajectory predictions”, *Sensors*, vol.21,pp. 6959, October 2021
- [58] A.P Alexandridis, C.I. Siettos, H.K. Sarimveis, A.G. Boudouvis, G.V. Bafas, “Modeling of nonlinear process dynamics using Kohonen’s Neural Networks, Fuzzy Systems and Chebyshev Series”, *Computers and Chemical Engineering*, vol.26, pp. 479–486, May 2002 .

- [59] D. Karamichailidou, V. Kaloutsas, A. Alexandridis, “Wind Turbine Power Curve Modeling Using Radial Basis Function Neural Networks and Tabu Search”, *Renewable Energy*, vol.163, pp.2137-2152, January 2021
- [60] T.M. Cover, “Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition”, *Electronic Computers*, vol.14, pp.326-334, June 1965
- [61] D. Karamichailidou, A. Alexandridis, G. Anagnostopoulos, G. Syriopoulos, O. Sekkas, “Modeling biogas production from anaerobic wastewater treatment plants using radial basis function networks and differential evolution”, *Computers and Chemical Engineering*, vol.157C, 107629, January 2022.
- [62] A. Alexandridis, E. Chondrodima, “A medical diagnostic tool based on radial basis function classifiers and evolutionary simulated annealing”, *Journal of Biomedical Informatics*, vol.49 (2014), pp. 61-72, June 2014 .
- [63] A. Alexandridis, M. Stogiannos, N. Papaioannou, E. Zois and H. Sarimveis, “An Inverse Neural Controller Based on the Applicability Domain of RBF Network Models”, *Sensors*, vol.18, pp.315, January 2018
- [64] A. Alexandridis, M. Stogiannos, A. Kyriou and H. Sarimveis, “An offset-free neural controller based on approximating the inverse process dynamics” *Journal of Process Control*, vol.23, pp. 968–979, August 2013.
- [65] A. Alexandridis, E. Chondrodima, N. Giannopoulos and H. Sarimveis, “A Fast and Efficient Method for Training Categorical Radial Basis Function Networks”, *IEEE Transactions on Neural Networks and Learning Systems*, vol.28, pp.2831-283, November 2017
- [66] M. Stogiannos, A. Alexandridis, H. Sarimveis, “Model predictive control for systems with fast dynamics using inverse neural models”, *ISA Transactions*, vol.72, pp.161-177, January 2018
- [67] A. Alexandridis, E. Chondrodima and H. Sarimveis, “Cooperative learning for radial basis function networks using particle swarm optimization”, *Applied Soft Computing*, vol.49, pp.485-497, August 2016
- [68] A. Alexandridis, E. Chondrodima and H. Sarimveis, “Radial Basis Function Network Training Using a Nonsymmetric Partition of the Input Space and Particle Swarm Optimization”, *IEEE Transactions on Neural Networks and Learning Systems* ,vol.24, pp.219-230, February 2013
- [69] A. Alexandridis, E. Chondrodima, E. Efthimiou, G. Papadakis, F. Vallianatos and D. Triantis, “Large earthquake occurrence estimation based on radial basis function neural networks”, *IEEE Transactions on Geoscience and Remote Sensing*, vol.52, pp. 5443-5453, September 2014
- [70] A. Alexandridis, “Evolving RBF neural networks for adaptive soft-sensor design”, *International Journal of Neural Systems*, vol. 23, pp. 1350029, October 2013.
- [71] D. Karamichailidou, S. Koletsios and A. Alexandridis, “An RBF online learning scheme for non-stationary environments based on fuzzy means and Givens rotations”, *Neurocomputing*, vol.501, pp.370-386, August 2022
- [72] European Commission (2020), *European Marine Observation and Data Network (EMODnet)* [Online].
- [73] K. Thomsen, *Offshore Wind : A Comprehensive Guide to Successful Offshore Wind Farm Installation*, 2nd Edition , July 2014, Elsevier Academic Press
- [74] H. Pate (2021), *What is Feature Engineering — Importance, Tools and Techniques for Machine Learning* [Online].