



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

ΠΙΘΑΝΟΤΙΚΗ ΠΡΟΒΛΕΨΗ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μάρκος Α. Κουσουνάδης - Κνούσεν

Επίβλεψη : Παύλος Σ. Γεωργιλάκης, Αναπληρωτής Καθηγητής Ε.Μ.Π.
Ιωάννης Μπαζιώνης, Υποψήφιος Διδάκτωρ Ε.Μ.Π.

Αθήνα, Ιούλιος 2021



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ

ΠΙΘΑΝΟΤΙΚΗ ΠΡΟΒΛΕΨΗ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μάρκος Α. Κουσουνάδης - Κνούσεν

Επίβλεψη : Πάυλος Σ. Γεωργιλάκης, Αναπληρωτής Καθηγητής Ε.Μ.Π.
Ιωάννης Μπαζιώνης, Υποψήφιος Διδάκτωρ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9^η Ιουλίου 2021

.....
Πάυλος Γεωργιλάκης
Αν. Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Κορρές
Καθηγητής Ε.Μ.Π.

.....
Ιωάννης Γκόνος
Αν. Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2021

.....
Μάρκος Α. Κουσουνάδης - Κνούσεν
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών
Ε.Μ.Π.

Copyright © Μάρκος Κουσουνάδης - Κνούσεν, 2021.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό 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 τον συγγραφέα.

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

ΠΕΡΙΛΗΨΗ

Σκοπός της παρούσας διπλωματικής εργασίας είναι η πιθανοτική πρόβλεψη της αιολικής παραγωγής μέσω της ανάπτυξης μοντέλων που βασίζονται σε τεχνητά νευρωνικά δίκτυα. Για την παραγωγή των διαστημάτων πρόβλεψης χρησιμοποιείται η μέθοδος εκτίμησης κάτω και άνω ορίου. Το πρώτο μοντέλο που αναπτύσσεται εκπαιδεύεται στη βάση της ελαχιστοποίησης του κριτηρίου του εύρους κάλυψης, ενώ το δεύτερο μοντέλο εκπαιδεύεται σε πολυκριτηριακά πλαίσια. Ως αλγόριθμος βελτιστοποίησης επιλέγεται η βελτιστοποίηση σμήνους σωματιδίων. Στην περίπτωση του δεύτερου μοντέλου ο αλγόριθμος τροποποιείται έτσι ώστε να είναι κατάλληλος για πολυκριτηριακή βελτιστοποίηση. Για την αρχικοποίηση των βαρών και των πολώσεων των δύο μοντέλων επιλέγεται η τεχνική αρχικοποίησης Nguyen-Widrow. Και στα δύο μοντέλα εφαρμόζεται τελεστής μετάλλαξης ο οποίος μειώνεται γραμμικά.

Για την εκπαίδευση και την αξιολόγηση των μοντέλων χρησιμοποιούνται δημόσια διαθέσιμα δεδομένα που προέρχονται από τον διαγωνισμό GEFCom 2014. Τα δεδομένα αφορούν δέκα αιολικούς σταθμούς και χωρίζονται σε εποχές ώστε να είναι δυνατή η εποχική ανάλυση.

Η ανάπτυξη του κώδικα για την κατασκευή των μοντέλων πιθανοτικής πρόβλεψης της παρούσας διπλωματικής εργασίας γίνεται στη γλώσσα προγραμματισμού Python. Για τα δεδομένα κάθε εποχής, τα μοντέλα παράγουν διαστήματα πρόβλεψης με ωριαία χρονική ανάλυση. Τα αποτελέσματα των μοντέλων συγκρίνονται μεταξύ τους αλλά και με τα αποτελέσματα άλλων μοντέλων πιθανοτικής πρόβλεψης της διαθέσιμης βιβλιογραφίας. Εξάγονται συμπεράσματα σχετικά με την απόδοση των μοντέλων αλλά και την απόδοση ανά εποχή, και προτείνονται ορισμένες πιθανές μελλοντικές επεκτάσεις της εργασίας.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Πιθανοτική πρόβλεψη αιολικής παραγωγής, διαστήματα πρόβλεψης, εκτίμηση κάτω και άνω ορίου, βελτιστοποίηση σμήνους σωματιδίων, πολυκριτηριακή βελτιστοποίηση, τεχνητά νευρωνικά δίκτυα.

ABSTRACT

The purpose of this diploma thesis is the probabilistic forecasting of wind power production through the development of forecasting models consisting of artificial neural networks. The Lower Upper Bound Estimation (LUBE) method is employed in order to generate prediction intervals. The minimization of the Coverage Width based Criterion (CWC) is used as the cost function for the training of the first model developed in this diploma thesis. The second developed model is trained in a multi objective framework. Particle Swarm Optimization (PSO) is chosen as the optimization algorithm. In the case of the second model, PSO is modified in order to suit multi objective optimization. The Nguyen-Widrow initialization technique is chosen for the initialization of the weights and biases. A linearly decreasing mutation operator is applied to both models.

For the training and the evaluation of the models, publicly available data from the Global Energy Forecasting Competition 2014 are used. The data, which are related to ten wind farms, are divided into seasons so that seasonal analysis is possible.

The code development for the construction of the probabilistic forecasting models of this diploma thesis is done in Python. For each seasonal dataset, prediction intervals are generated in an hourly time resolution. The results of the models are compared with each other and also with the results of other probabilistic forecasting models found in the available bibliography. Conclusions about the performance of the models and the performance per season are drawn. Possible future research on the topic is also proposed.

KEYWORDS

Probabilistic wind power forecasting, prediction intervals, Lower Upper Bound Estimation, Particle Swarm Optimization, multi objective optimization, Artificial Neural Networks.

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2020–2021 υπό την επίβλεψη του κ. Παύλου Γεωργιάκη, αναπληρωτή καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. στον οποίο οφείλω ιδιαίτερες ευχαριστίες για την ανάθεσή της, δίνοντάς μου την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα Γιάννη Μπαζιώνη για τη βοήθεια και την καθοδήγησή κατά την εκπόνηση της εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου Αντώνη και Σίσελ για την πολύτιμη στήριξή τους κατά τη διάρκεια των σπουδών μου, τους φίλους μου που ήταν πάντα δίπλα μου καθώς και την Κατερίνα για τη συμπαράσταση και την αγάπη της σε όλη τη διάρκεια εκπόνησης της εργασίας.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1^Ο

ΕΙΣΑΓΩΓΗ

1.1	Αντικείμενο της διπλωματικής εργασίας	1
1.2	Βιβλιογραφική ανασκόπηση	3
1.3	Δομή διπλωματικής εργασίας	4

ΚΕΦΑΛΑΙΟ 2^Ο

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1	Αρχιτεκτονική τεχνητών νευρωνικών δικτύων	5
2.2	Εκπαίδευση τεχνητών νευρωνικών δικτύων	8
2.2.1	Επιτηρούμενη εκπαίδευση	8
2.2.2	Μη επιτηρούμενη εκπαίδευση	9
2.3	Αντικειμενικές συναρτήσεις	9
2.4	Αλγόριθμοι βελτιστοποίησης	10
2.4.1	Αλγόριθμος προς τα πίσω διάδοσης του σφάλματος	10
2.4.2	Μεταερευνητικοί αλγόριθμοι	12
2.4.3	Βελτιστοποίηση σμήνους σωματιδίων	13
2.5	Πολυκριτηριακή βελτιστοποίηση	15
2.5.1	Μέτωπο Pareto	16
2.5.2	Πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων	17

ΚΕΦΑΛΑΙΟ 3^Ο

ΠΡΟΒΛΕΨΗ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ

3.1	Η σημασία της πρόβλεψης αιολικής παραγωγής	21
3.2	Πρόβλεψη αιολικής παραγωγής με χρήση μηχανικής μάθησης	23
3.2.1	Αιτιοκρατικά μοντέλα	24
3.2.2	Πιθανοτικά μοντέλα	24
3.3	Μέθοδος εκτίμησης κάτω και άνω ορίου	26

ΚΕΦΑΛΑΙΟ 4^Ο

ΑΝΑΛΥΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΑΙΟΛΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

4.1	Δεδομένα αιολικής παραγωγής	30
4.2	Δεδομένα του διαγωνισμού Global Energy Forecasting 2014	30
4.3	Επεξεργασία Δεδομένων	31
4.3.1	Μετατροπή δεδομένων	31
4.3.2	Κανονικοποίηση δεδομένων	32

ΚΕΦΑΛΑΙΟ 5^ο

ΜΟΝΤΕΛΑ ΠΙΘΑΝΟΤΙΚΗΣ ΠΡΟΒΛΕΨΗΣ ΑΙΟΛΙΚΗΣ ΙΣΧΥΟΣ

5.1	Εκτίμηση κάτω και άνω ορίου με βελτιστοποίηση σμήνους σωματιδίων του κριτηρίου του εύρους κάλυψης	35
5.1.1	Κριτήριο εύρους κάλυψης	35
5.1.2	Αρχικοποίηση βαρών και πολώσεων	37
5.1.3	Αλγόριθμος βελτιστοποίησης	39
5.1.4	Τελεστής μετάλλαξης	40
5.1.5	Μετρητικά αξιολόγησης μοντέλου	43
5.2	Εκτίμηση κάτω και άνω ορίου με πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων	44
5.2.1	Αλγόριθμος βελτιστοποίησης	45
5.2.2	Μετρητικά αξιολόγησης μοντέλου	48
5.3	Βελτιστοποίηση παραμέτρων	49
5.3.1	Βελτιστοποίηση παραμέτρων LUBE–PSO–CWC	50
5.3.2	Βελτιστοποίηση παραμέτρων LUBE–MOPSO	55
5.3.3	Βελτίωση αποτελεσμάτων	59

ΚΕΦΑΛΑΙΟ 6^ο

ΑΠΟΤΕΛΕΣΜΑΤΑ

6.1	Μεθοδολογία	61
6.1.1	Διαχωρισμός δεδομένων	61
6.1.2	Μοντέλο BELM	64
6.1.3	Αξιολόγηση μοντέλων πιθανοτικής πρόβλεψης	65
6.2	Σύγκριση αποτελεσμάτων	66
6.2.1	Σύγκριση απόδοσης μοντέλων	67
6.2.2	Εποχική ανάλυση	76

ΚΕΦΑΛΑΙΟ 7^ο

ΕΠΙΛΟΓΟΣ

7.1	Σύνοψη και συμπεράσματα	85
7.2	Μελλοντικές επεκτάσεις της εργασίας	87

ΚΕΦΑΛΑΙΟ 8^ο

ΒΙΒΛΙΟΓΡΑΦΙΑ

8.1	Βιβλιογραφία	89
-----	--------------	----

ΕΙΣΑΓΩΓΗ

1.1 ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Αντικείμενο της παρούσας διπλωματικής εργασίας αποτελεί η πρόβλεψη της αιολικής παραγωγής μέσω της ανάπτυξης μοντέλων που βασίζουν τη λειτουργία τους σε τεχνητά νευρωνικά δίκτυα. Η ανάγκη για μείωση των εκπομπών διοξειδίου του άνθρακα για τον περιορισμό των συνεπειών της κλιματικής αλλαγής καθώς και η μελλοντική πιθανή εξάντληση των αποθεμάτων των ορυκτών καυσίμων, έχουν οδηγήσει στην όλο και αυξανόμενη διείδυση των Ανανεώσιμων Πηγών Ενέργειας (ΑΠΕ) στα συστήματα ηλεκτρικής ενέργειας. Μεγάλο μερίδιο κατέχει τα τελευταία χρόνια η παραγωγή ηλεκτρικής ενέργειας από αιολικούς σταθμούς παραγωγής (αιολική ενέργεια). Η διείδυση της αιολικής ενέργειας στο δίκτυο έχει ως αποτέλεσμα την αύξηση της αβεβαιότητας και της μεταβλητότητας. Η παραγωγή ηλεκτρικής ενέργειας από αιολικούς σταθμούς εξαρτάται από τις καιρικές συνθήκες, και κυρίως από την ταχύτητα του ανέμου. Λόγω της τυχαίας φύσης του ανέμου, η παραγωγή αιολικής ισχύος είναι μόνο μερικώς ελεγχόμενη και χαρακτηρίζεται από μεγάλη στοχαστικότητα. Οι συνεχείς και μη ελεγχόμενες μεταβολές στην ταχύτητα του ανέμου σε συνδυασμό με την μορφή της καμπύλης ισχύος της ανεμογεννήτριας ως συνάρτηση της ταχύτητας του ανέμου καθώς και σε συνδυασμό με την αβεβαιότητα των βραχυπρόθεσμων μετεωρολογικών προβλέψεων καθιστούν την πρόβλεψη της αιολικής παραγωγής μία μεγάλη πρόκληση. Η ορθή πρόβλεψη της αιολικής παραγωγής έχει μεγάλη σημασία, καθώς όσο μεγαλύτερο ποσοστό καταλαμβάνει η αιολική παραγωγή στο μερίδιο της συνολικής παραγωγής ηλεκτρικής ισχύος, τόσο περισσότερο επηρεάζει τον προγραμματισμό για την κάλυψη των εθνικών ενεργειακών αναγκών. Η απόκλιση των προβλέψεων από τις πραγματικές τιμές της παραγόμενης αιολικής ισχύος οδηγούν στη μείωση της απόδοσης του συστήματος, στην εξάρτηση από συμβατικές μονάδες παραγωγής, στη σπατάλη περίσσειας παραγόμενης αιολικής ενέργειας και στην αύξηση του κόστους παραγωγής της ηλεκτρικής ενέργειας.

Για την πρόβλεψη της αιολικής παραγωγής υπάρχουν δύο είδη μοντέλων: τα φυσικά μοντέλα και τα στατιστικά μοντέλα. Στα στατιστικά μοντέλα περιλαμβάνονται και τα μοντέλα που βασίζουν τη λειτουργία τους σε τεχνητά νευρωνικά δίκτυα. Τα στατιστικά μοντέλα χωρίζονται περαιτέρω στα αιτιοκρατικά μοντέλα και στα πιθανοτικά μοντέλα. Τα αιτιοκρατικά μοντέλα παράγουν σημειακές προβλέψεις. Αντίθετα, τα πιθανοτικά μοντέλα είτε κάνουν πρόβλεψη ποσοστιμορίου είτε παράγουν διαστήματα πρόβλεψης.

Στην παρούσα διπλωματική εργασία αναπτύσσονται δύο μοντέλα με σκοπό την πιθανοτική πρόβλεψη της αιολικής παραγωγής. Τα μοντέλα παράγουν διαστήματα πρόβλεψης εντός των οποίων αναμένεται να βρεθούν οι πραγματικές τιμές της παραγόμενης αιολικής ισχύος με βάση μια πιθανότητα. Η κατασκευή των διαστημάτων πρόβλεψης από τα μοντέλα της παρούσας διπλωματικής εργασίας βασίζεται στη μέθοδο εκτίμησης κάτω και

άνω ορίου (Lower Upper Bound Estimation – LUBE) [1]. Το τεχνητό νευρωνικό δίκτυο του ενός μοντέλου εκπαιδεύεται στη βάση της βελτιστοποίησης μιας μόνο αντικειμενικής συνάρτησης, ενώ η εκπαίδευση του τεχνητού νευρωνικού δικτύου του άλλου μοντέλου έγκειται στην ταυτόχρονη βελτιστοποίηση πολλαπλών αντικειμενικών συναρτήσεων (πολυκριτηριακή βελτιστοποίηση). Οι αλγόριθμοι βελτιστοποίησης που χρησιμοποιούνται είναι η βελτιστοποίηση σμήνους σωματιδίων (Particle Swarm Optimization – PSO) [2] και η πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων (Multi Objective Particle Swarm Optimization – MOPSO) [3], οι οποίοι ανήκουν στην κατηγορία των μεταερευτικών αλγορίθμων. Τα βάρη και οι πολώσεις των τεχνητών νευρωνικών δικτύων αρχικοποιούνται με την τεχνική αρχικοποίησης Nguyen–Widrow [4], ενώ γίνεται εφαρμογή ενός τελεστή μετάλλαξης (mutation operator) με στόχο τη βελτίωση της απόδοσης των δύο μοντέλων.

Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση των μοντέλων προέρχονται από τον διεθνή διαγωνισμό Global Energy Forecasting Competition 2014 (GEFCOM 2014) [5]. Τα δεδομένα αυτά εμπεριέχουν στοιχεία για δέκα διαφορετικούς αιολικούς σταθμούς – ζώνες της Αυστραλίας. Συγκεκριμένα, εμπεριέχουν τις προβλέψεις για τις συνιστώσες του διανύσματος του ανέμου σε δύο διαφορετικά ύψη καθώς και τις πραγματικές τιμές της παραγόμενης αιολικής ισχύος κάθε αιολικού σταθμού, σε ωριαία χρονική ανάλυση. Τα δεδομένα αφορούν τα έτη 2012 και 2013. Στην παρούσα διπλωματική εργασία γίνεται διαχωρισμός των δεδομένων σε εποχές έτσι ώστε να είναι εφικτή η εποχική ανάλυση των αποτελεσμάτων.

Αφού γίνει καθορισμός των παραμέτρων των δύο μοντέλων, επιτελείται η εκπαίδευση, η παραγωγή των διαστημάτων πρόβλεψης και η αξιολόγησή τους για κάθε εποχή κάθε ζώνης. Τα αποτελέσματα των δύο μοντέλων συγκρίνονται μεταξύ τους αλλά και με τα αποτελέσματα άλλων πιθανοτικών μοντέλων της διαθέσιμης βιβλιογραφίας. Πέρα από την αξιολόγηση της απόδοσης των δύο μοντέλων της παρούσας διπλωματικής εργασίας, γίνεται εποχική ανάλυση, δηλαδή σύγκριση των αποτελεσμάτων ανά εποχή για την εξαγωγή συμπερασμάτων γύρω από την απόδοση ανά εποχή, την ύπαρξη ή όχι περιοδικών φαινομένων ανά εποχή, καθώς και για τη διερεύνηση τυχόν συσχετίσεων μεταξύ των αποτελεσμάτων των εποχών.

Βασική καινοτομία της παρούσας διπλωματικής εργασίας αποτελεί η ανάπτυξη ενός μοντέλου πιθανοτικής πρόβλεψης αιολικής παραγωγής, που βασίζει τη λειτουργία του στη μέθοδο LUBE και εκπαιδεύεται σε πολυκριτηριακά πλαίσια. Επίσης, ο τελεστής μετάλλαξης που χρησιμοποιείται έχει μεγαλύτερο αρχικό δείκτη μετάλλαξης σε σχέση με αυτόν που συναντάται συνήθως στη βιβλιογραφία ώστε να είναι κατάλληλα προσαρμοσμένος στο πρόβλημα της πρόβλεψης αιολικής παραγωγής. Άλλη συνεισφορά της παρούσας διπλωματικής εργασίας αποτελεί η εποχική ανάλυση και τα συμπεράσματα που εξάγονται από αυτήν. Τέλος, εφαρμόζεται μία μέθοδος βελτιστοποίησης παραμέτρων, ώστε οι παράμετροι των μοντέλων να καθοριστούν με βάση τα ειδικά χαρακτηριστικά του προβλήματος, ενώ η αξιολόγηση των μοντέλων γίνεται σε αποτελέσματα που παράγονται από δημόσια διαθέσιμα δεδομένα, επομένως είναι εφικτή η αξιολόγηση και η χρήση των αποτελεσμάτων των δύο μοντέλων για την άμεση σύγκριση με άλλα μοντέλα που έχουν αναπτυχθεί ή θα αναπτυχθούν στο μέλλον.

1.2 ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Την τελευταία δεκαετία οι περισσότερες έρευνες έχουν εστιάσει στην ανάπτυξη πιθανοτικών αντί αιτιοκρατικών μοντέλων για την πρόβλεψη της αιολικής παραγωγής. Τα μοντέλα αυτά κάνουν χρήση μηχανικής μάθησης, χρησιμοποιούν παραμετρικές ή μη παραμετρικές μεθόδους για την κατασκευή των διαστημάτων πρόβλεψης και εκπαιδεύονται σε μονοκριτηριακά ή πολυκριτηριακά πλαίσια. Ενδεικτικά, παρουσιάζονται ορισμένες εργασίες της διαθέσιμης βιβλιογραφίας γύρω από την ανάπτυξη πιθανοτικών μοντέλων για την πρόβλεψη της αιολικής παραγωγής.

Οι C. Wan, Z. Xu, P. Pinson, Z. Y. Dong και K. P. Wong στην εργασία [6] αναπτύσσουν ένα μοντέλο πιθανοτικής πρόβλεψης αιολικής παραγωγής που βασίζεται σε Extreme Learning Machines. Τα διαστήματα πρόβλεψης παράγονται ακολουθώντας παραμετρική μέθοδο, με την έννοια ότι πρώτα παράγονται σημειακές προβλέψεις και στη συνέχεια παράγονται τα διαστήματα πρόβλεψης, κάνοντας χρήση της μεθόδου bootstrapping ζευγαριών (pairs bootstrapping) και κάποιων στατιστικών μεθόδων.

Οι H. Wang, G. Li, G. Wang, J. Peng, H. Jiang και Y. Liu στην εργασία [7] αναπτύσσουν ένα μοντέλο που επίσης παράγει τα διαστήματα πρόβλεψης ακολουθώντας παραμετρική μέθοδο. Το μοντέλο βασίζεται στη θεωρία της βαθιάς μηχανικής μάθησης (deep learning), αποτελείται από ένα συνελκτικό τεχνητό νευρωνικό δίκτυο, ενώ στα δεδομένα ισχύος εφαρμόζεται μετασχηματισμός κυματιδίων (wavelet transformation). Αντίστοιχο μοντέλο αναπτύσσουν οι H. Wen, J. Gu, J. Ma και Z. Jin στην εργασία [8], όπου χρησιμοποιείται ένα πλήρες συνελκτικό τεχνητό νευρωνικό δίκτυο και για την παραγωγή των διαστημάτων πρόβλεψης γίνεται χρήση της θεωρίας του Μπέυζ (Bayesian theory) και της διακοπής Monte Carlo (Monte Carlo dropout).

Οι H. Quan, D. Srinivasan και A. Khosravi στην εργασία [9] κάνουν χρήση της μη παραμετρικής μεθόδου LUBE για την παραγωγή των διαστημάτων πρόβλεψης. Το πρόβλημα της βελτιστοποίησης των διαστημάτων πρόβλεψης μετατρέπεται σε πρόβλημα βελτιστοποίησης μίας αντικειμενικής συνάρτησης και γίνεται χρήση του αλγορίθμου PSO για την εκπαίδευση του τεχνητού νευρωνικού δικτύου. Αντίστοιχο μοντέλο αναπτύσσουν οι Y. K. Wu, P. E. Su, T. Y. Wu, J. S. Hong και M. Y. Hassan στην εργασία [10], όπου ως αλγόριθμος βελτιστοποίησης επιλέγεται η αναζήτηση συστήματος επιβάρυνσης (Charged System Search - CSS). Οι Z. Shi, H. Liang και V. Divanahi στην εργασία [11] αξιοποιούν τη μέθοδο LUBE σε ένα αναδρομικό τεχνητό νευρωνικό δίκτυο (recurrent neural network) επιλέγοντας ως αλγόριθμο βελτιστοποίησης τον αλγόριθμο της λιβελούλας (dragonfly algorithm).

Το μοντέλο που αναπτύσσουν οι G. Zhang, Z. Li, K. Zhang, L. Zhang, X. Hua και Y. Wang στην εργασία [12] αξιοποιεί τη συνάρτηση Copula για την παραγωγή των διαστημάτων πρόβλεψης, ενώ για την επιλογή του βέλτιστου συνδυασμού παραμέτρων κάνει χρήση του πολυκριτηριακού γενετικού αλγορίθμου μη κυριαρχούμενης ταξινόμησης II (Non dominated Sorting Genetic Algorithm II – NSGA II). Αντίστοιχο μοντέλο που εκπαιδεύεται σε πολυκριτηριακά πλαίσια αναπτύσσουν οι A. Kavousi-Fard, A. Khosravi και S. Nahavandi στην εργασία [13]. Το μοντέλο αυτό κάνει χρήση της μεθόδου LUBE και βασίζεται στη θεωρία της ασαφούς λογικής (fuzzy logic). Για την εκπαίδευση του μοντέλου προτείνεται μια τροποποιημένη εκδοχή του αλγορίθμου της νυχτερίδας (Bat Algorithm – BA).

1.3 ΔΟΜΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα διπλωματική εργασία δομείται ως εξής: στο Κεφάλαιο 2 δίνεται το θεωρητικό υπόβαθρο γύρω από τα τεχνητά νευρωνικά δίκτυα. Παρουσιάζεται το μαθηματικό μοντέλο τους, αναλύονται ο τρόπος λειτουργίας τους και τα είδη της εκπαίδευσής τους. Περιγράφεται ο ρόλος των αντικειμενικών συναρτήσεων και των αλγορίθμων βελτιστοποίησης. Γίνεται παρουσίαση του αλγορίθμου της προς τα πίσω διάδοσης του σφάλματος και του αλγορίθμου βελτιστοποίησης σμήνους σωματιδίων. Παρουσιάζεται η έννοια της πολυκριτηριακής βελτιστοποίησης και περιγράφεται η πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων.

Στο Κεφάλαιο 3 αναλύεται η σημασία της ορθής πρόβλεψης της αιολικής παραγωγής και γίνεται κατηγοριοποίηση των μοντέλων πρόβλεψης. Επίσης, παρουσιάζεται η μέθοδος κάτω και άνω ορίου, τα ειδικά χαρακτηριστικά της και τα πλεονεκτήματά της.

Στο Κεφάλαιο 4 γίνεται ανάλυση των δεδομένων που χρησιμοποιούνται στην παρούσα διπλωματική εργασία για την εκπαίδευση και αξιολόγηση των τεχνητών νευρωνικών δικτύων καθώς και περιγραφή των μεθόδων επεξεργασίας των δεδομένων.

Στο Κεφάλαιο 5 παρουσιάζεται αναλυτικά η δομή των δύο πιθανοτικών μοντέλων πρόβλεψης αιολικής παραγωγής της παρούσας διπλωματικής εργασίας, τα μετρητικά αξιολόγησης που χρησιμοποιούνται, καθώς και η διαδικασία βελτιστοποίησης των παραμέτρων των δύο μοντέλων.

Στο Κεφάλαιο 6 παρουσιάζεται η μέθοδος του διαχωρισμού των δεδομένων σε δεδομένα εκπαίδευσης και δεδομένα αξιολόγησης. Γίνεται περιγραφή του μοντέλου πιθανοτικής πρόβλεψης που χρησιμοποιείται ως σημείο αναφοράς, καθώς και της μεθόδου αξιολόγησης των αποτελεσμάτων που ακολουθείται. Καταγράφονται τα αποτελέσματα, παρουσιάζονται τα παραγόμενα διαστήματα πρόβλεψης κάθε μοντέλου και γίνεται σύγκριση της απόδοσης των μοντέλων. Για τα αποτελέσματα των δύο μοντέλων της παρούσας διπλωματικής εργασίας γίνεται εποχική ανάλυση.

Στο Κεφάλαιο 7 γίνεται η σύνοψη της παρούσας διπλωματικής εργασίας, καταγράφονται τα συνολικά συμπεράσματά της, και προτείνονται ορισμένες πιθανές επεκτάσεις της.

Το Κεφάλαιο 8 περιλαμβάνει τη βιβλιογραφία που χρησιμοποιήθηκε για την εκπόνηση της παρούσας διπλωματικής εργασίας.

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Με τον όρο τεχνητά νευρωνικά δίκτυα περιγράφονται υπολογιστικά συστήματα ικανά να απεικονίσουν οποιαδήποτε μη γραμμική συνάρτηση. Είναι ιδιαίτερα αποδοτικά σε ότι αφορά την επίλυση προβλημάτων υψηλής πολυπλοκότητας όπου η σχέση μεταξύ εισόδου και εξόδου δεν είναι σαφώς καθορισμένη ή εύκολα υπολογίσιμη [14]. Υπολογιστικά μοντέλα που βασίζονται σε τεχνητά νευρωνικά δίκτυα μπορούν να δώσουν προσεγγιστικές λύσεις σε προβλήματα για την επίλυση των οποίων τα κλασικά μαθηματικά μοντέλα είναι ανεπαρκή.

Η αρχιτεκτονική των τεχνητών νευρωνικών δικτύων είναι εμπνευσμένη από το ανθρώπινο νευρικό σύστημα [15]. Βασικό στοιχείο ενός τεχνητού νευρωνικού δικτύου αποτελεί ο νευρώνας. Το μαθηματικό μοντέλο ενός νευρώνα απεικονίζεται στο Σχήμα 2.1. Κάθε νευρώνας δέχεται εισόδους, είτε από άλλους νευρώνες (x_1, x_2, \dots, x_n), είτε από εξωτερικούς παράγοντες (πόλωση b) και παράγει μια έξοδο y . Σε κάθε είσοδο x_1, x_2, \dots, x_n αντιστοιχεί ένα βάρος w_1, w_2, \dots, w_n . Η συνολική είσοδος z του νευρώνα του Σχήματος 2.1 υπολογίζεται ως εξής:

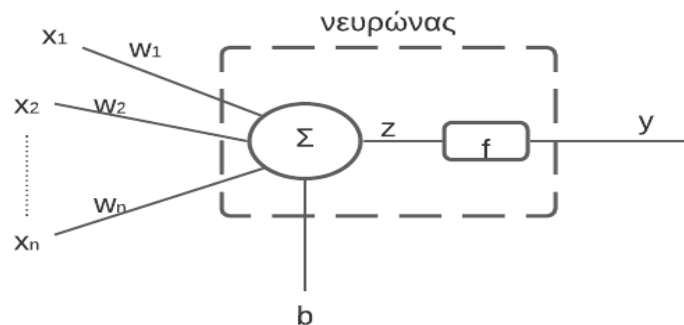
$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b \quad (2.1)$$

Η έξοδος y του νευρώνα του Σχήματος 2.1 υπολογίζεται ως εξής:

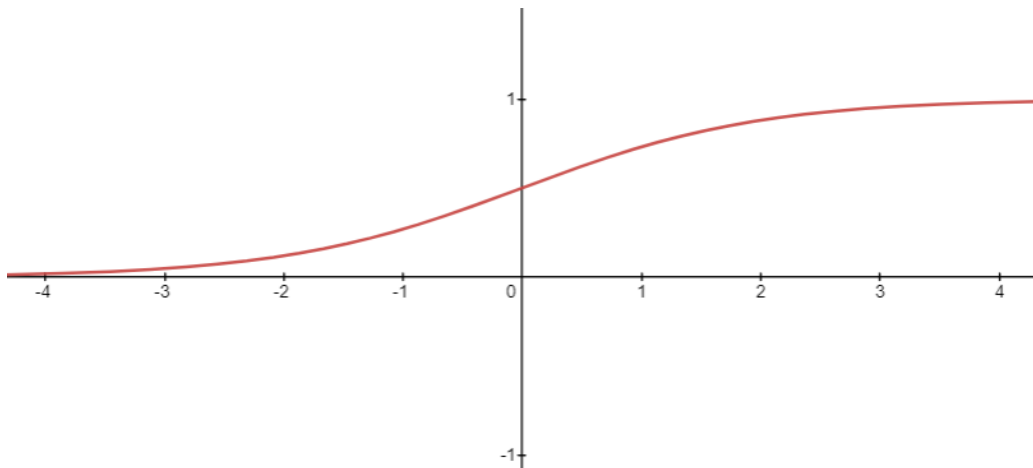
$$y = f(z) \quad (2.2)$$

όπου z η είσοδος του νευρώνα και f η συνάρτηση ενεργοποίησης. Ένα τυπικό παράδειγμα μιας συνάρτησης ενεργοποίησης είναι η σιγμοειδής (Σχήμα 2.2):

$$f(z) = \frac{1}{1 + e^{-z}} \quad (2.3)$$



Σχήμα 2.1: Το μαθηματικό μοντέλο ενός νευρώνα τεχνητού νευρωνικού δικτύου



Σχήμα 2.2: Η σιγμοειδής συνάρτηση ενεργοποίησης

Άλλες συχνά χρησιμοποιούμενες συναρτήσεις ενεργοποίησης είναι η διπολική σιγμοειδής, η ReLU και η υπερβολική εφαπτομένη, οι οποίες υπολογίζονται σύμφωνα με τις εξισώσεις (2.4), (2.5) και (2.6), αντίστοιχα:

$$f_{b_sigmoid}(z) = \frac{1 - e^{-z}}{1 + e^{-z}} \quad (2.4)$$

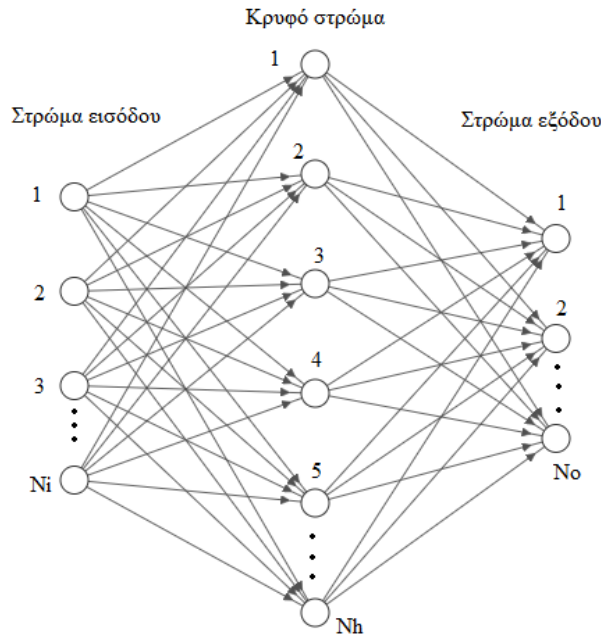
$$f_{ReLU}(z) = \max(0, z) \quad (2.5)$$

$$f(z)_{tanh} = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (2.6)$$

Ένα τεχνητό νευρωνικό δίκτυο αποτελείται από νευρώνες συνδεδεμένους παράλληλα και σειριακά μεταξύ τους μέσω συνάψεων. Ένας ή περισσότεροι νευρώνες συνδεδεμένοι παράλληλα μεταξύ τους συγκροτούν ένα στρώμα. Υπάρχουν τρία είδη στρωμάτων: το στρώμα εισόδου, το στρώμα εξόδου και τα κρυφά στρώματα [14]. Το στρώμα εισόδου αποτελείται από παθητικούς νευρώνες, δηλαδή νευρώνες που δεν επιτελούν κάποια υπολογιστική διαδικασία, απλώς περιέχουν πληροφορία η οποία μεταφέρεται μέσω των συνάψεων στους νευρώνες του κρυφού στρώματος. Αντίθετα, τα κρυφά στρώματα και το στρώμα εξόδου αποτελούνται από νευρώνες με μαθηματικό μοντέλο σαν αυτό του Σχήματος 2.1.

Ανάλογα με την κατεύθυνση της ροής της πληροφορίας, τα τεχνητά νευρωνικά δίκτυα χωρίζονται σε δύο κατηγορίες [16]:

- Τα νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης. Η ροή της πληροφορίας ακολουθεί μία κατεύθυνση, από το στρώμα εισόδου προς τα κρυφά στρώματα, και από εκεί προς το στρώμα εξόδου. Στο Σχήμα 2.3 απεικονίζεται ένα νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης, με ένα κρυφό στρώμα.
- Τα αναδρομικά νευρωνικά δίκτυα. Σε αντίθεση με τα νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης, η πληροφορία στα αναδρομικά νευρωνικά δίκτυα μπορεί να μεταφέρεται μεταξύ νευρώνων του ίδιου στρώματος, καθώς και προς τα πίσω (σε νευρώνα προηγούμενου στρώματος).



Σχήμα 2.3: Νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης ενός κρυφού στρώματος

Για το τεχνητό νευρωνικό δίκτυο του Σχήματος 2.3, η σχέση μεταξύ εισόδου και εξόδου είναι η εξής:

$$y_i = f_1 \left(\sum_{j=1}^{N_h} w_{ij} f_2 \left(\sum_{k=1}^{N_i} v_{jk} x_k + b_{vj} \right) + b_{wi} \right) \quad i = 0, 1, \dots, N_o \quad (2.7)$$

όπου y_i η έξοδος του i -οστού νευρώνα του στρώματος εξόδου, f_1 η συνάρτηση ενεργοποίησης του στρώματος εξόδου, f_2 η συνάρτηση ενεργοποίησης του κρυφού στρώματος, w_{ij} το βάρος της σύναψης μεταξύ του i -οστού νευρώνα του στρώματος εξόδου και του j -οστού νευρώνα του κρυφού στρώματος, b_{wi} η πόλωση του i -οστού νευρώνα του

στρώματος εξόδου, v_{jk} το βάρος της σύναψης μεταξύ του k -οστού νευρώνα του στρώματος εισόδου και του j -οστού νευρώνα του κρυφού στρώματος, b_{vj} η πόλωση του j -οστού νευρώνα του κρυφού στρώματος και x_k η τιμή του k -οστού νευρώνα του στρώματος εισόδου. Τα N_i , N_h , N_o αντιστοιχούν στο πλήθος νευρώνων του στρώματος εισόδου, του κρυφού στρώματος και του στρώματος εξόδου.

Το πλήθος των νευρώνων σε κάθε στρώμα, καθώς και το πλήθος των κρυφών στρωμάτων εξαρτάται από τα χαρακτηριστικά του εκάστοτε προβλήματος. Πολύ μικρός αριθμός νευρώνων ενδεχομένως να αδυνατεί να επιλύσει κάποιο πρόβλημα υψηλής πολυπλοκότητας (υποπροσαρμογή), ενώ αντίθετα πολύ μεγάλος αριθμός νευρώνων ενδεχομένως να μειώσει την ικανότητα γενίκευσης του μοντέλου (υπερπροσαρμογή).

Τα τεχνητά νευρωνικά δίκτυα έχουν πλήθος πλεονεκτημάτων. Είναι μη γραμμικά, επομένως μπορούν να προσεγγίσουν καλύτερα τις λύσεις προβλημάτων υψηλής πολυπλοκότητας σε σχέση με τα γραμμικά μοντέλα. Μπορούν να εξάγουν πληροφορίες και άγνωστες σχέσεις μεταξύ των δεδομένων, ενώ επίσης έχουν τη δυνατότητα να προσαρμοστούν σε άγνωστα δεδομένα (γενίκευση). Τέλος, η παράλληλη σύνδεση μεταξύ των νευρώνων δίνει τη δυνατότητα παράλληλων (ταυτόχρονων) υπολογισμών, καθιστώντας έτσι

τα τεχνητά νευρωνικά δίκτυα ταχύτερα σε σχέση με άλλα υπολογιστικά συστήματα. Τα παραπάνω πλεονεκτήματα έχουν οδηγήσει στην όλο και αυξανόμενη χρήση των τεχνητών νευρωνικών δικτύων, κυρίως σε εφαρμογές που αφορούν εκτίμηση συναρτήσεων, ταξινόμηση και πρόβλεψη [17].

2.2 ΕΚΠΑΙΔΕΥΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Η ικανότητα των τεχνητών νευρωνικών δικτύων να βρίσκουν μοτίβα και συσχετίσεις μεταξύ των δεδομένων έτσι ώστε στη συνέχεια να μπορούν να επιτελέσουν λειτουργίες όπως ταξινόμηση και πρόβλεψη, βασίζεται στην εκπαίδευσή τους. Η εκπαίδευση των τεχνητών νευρωνικών δικτύων έγκειται στην επαναλαμβανόμενη τροφοδότησή τους με ένα σύνολο δεδομένων και στην αναπροσαρμογή των βάρων και των πολώσεων έτσι ώστε για κάθε δεδομένο εισόδου να παράγεται η επιθυμητή έξοδος [14]. Το σύνολο αυτό των δεδομένων ονομάζεται σύνολο εκπαίδευσης, και είναι υποσύνολο του συνολικού διαθέσιμου συνόλου δεδομένων.

Υπάρχουν δύο μέθοδοι εκπαίδευσης των τεχνητών νευρωνικών δικτύων: (i) η επιτηρούμενη εκπαίδευση και (ii) η μη επιτηρούμενη εκπαίδευση.

2.2.1 Επιτηρούμενη εκπαίδευση

Στην επιτηρούμενη εκπαίδευση, το σύνολο εκπαίδευσης αποτελείται από δεδομένα εισόδου και δεδομένα εξόδου. Αρχικά, τα βάρη και οι πολώσεις αρχικοποιούνται είτε με τυχαίο τρόπο είτε με κάποια μέθοδο αρχικοποίησης. Στη συνέχεια το τεχνητό νευρωνικό δίκτυο τροφοδοτείται επαναλαμβανόμενα με το σύνολο δεδομένων εκπαίδευσης, αναπροσαρμόζοντας σε κάθε επανάληψη τα βάρη και τις πολώσεις, έως ότου για κάθε δεδομένο εισόδου να προσεγγίζεται επαρκώς η επιθυμητή έξοδος. Η αναπροσαρμογή των βάρων και των πολώσεων γίνεται με κατάλληλο τρόπο ώστε να μειώνεται η απόκλιση μεταξύ των πραγματικών και των επιθυμητών εξόδων. Όταν η απόκλιση αυτή έχει φτάσει στα επιθυμητά επίπεδα, οι τιμές των βάρων και των πολώσεων «κλειδώνουν» και η διαδικασία της εκπαίδευσης ολοκληρώνεται. Τέλος, έχοντας συγκλίνει στα κατάλληλα βάρη και στις κατάλληλες πολώσεις, το τεχνητό νευρωνικό δίκτυο αξιολογείται τροφοδοτώντας το με τα υπόλοιπα δεδομένα, τα οποία αποτελούν το σύνολο αξιολόγησης. Σε ορισμένες εφαρμογές, η διαδικασία της εκπαίδευσης συνεχίζεται και κατά τη διάρκεια της αξιολόγησης, αυξάνοντας την προσαρμοστικότητα, μειώνοντας ωστόσο την ταχύτητα του μοντέλου .

Η σύγκλιση ενός τεχνητού νευρωνικού δικτύου στα κατάλληλα βάρη και στις κατάλληλες πολώσεις δεν είναι δεδομένη. Η ανικανότητα ενός τεχνητού νευρωνικού δικτύου να εκπαιδευτεί μπορεί να σχετίζεται με διάφορους παράγοντες, όπως με την αρχιτεκτονική του, τον αλγόριθμο με βάση τον οποίο αναπροσαρμόζονται τα βάρη και οι πολώσεις καθώς και με την ποιότητα και το πλήθος των δεδομένων εκπαίδευσης [15].

Η συντριπτική πλειονότητα των τεχνητών νευρωνικών δικτύων χρησιμοποιούν την επιτηρούμενη εκπαίδευση, κυρίως λόγω της αποτελεσματικότητας και της ακρίβειάς της [15].

2.2.2 Μη επιτηρούμενη εκπαίδευση

Σε αντίθεση με την επιτηρούμενη εκπαίδευση, στη μη επιτηρούμενη εκπαίδευση το τεχνητό νευρωνικό δίκτυο τροφοδοτείται μόνο με δεδομένα εισόδου, επομένως θα πρέπει μόνο του να αναγνωρίσει μοτίβα ή να βρει τα χαρακτηριστικά με βάση τα οποία θα ταξινομήσει τα δεδομένα εισόδου, χωρίς εξωτερική βοήθεια. Ένα δεδομένο εξόδου δηλαδή, εκπαιδεύεται ώστε να αναπτύξει μια μη προκαθορισμένη ανταπόκριση σε ομάδες δεδομένων εισόδου που παρουσιάζουν κάποια στατιστική συσχέτιση. Η διαδικασία αυτή ονομάζεται προσαρμοστικότητα [15]. Φυσικά, το τεχνητό νευρωνικό δίκτυο χρειάζεται πληροφόρηση με κάποιο τρόπο, ώστε να μπορεί να προσαρμοστεί. Η πληροφορία αυτή δίνεται μέσω της αρχιτεκτονικής του τεχνητού νευρωνικού δικτύου, καθώς και μέσω κάποιων κανόνων μάθησης [15].

Η μη επιτηρούμενη εκπαίδευση είναι ταχύτερη από την επιτηρούμενη, ωστόσο έχει μικρότερη ακρίβεια. Χρησιμοποιείται κυρίως για να γίνει κάποια αρχική κατηγοριοποίηση των δεδομένων εισόδου [16]. Λόγω της σχετικά μικρής κατανόησης που υπάρχει γύρω από την μη επιτηρούμενη εκπαίδευση, οι εφαρμογές στις οποίες χρησιμοποιείται είναι λίγες, ωστόσο αποτελεί πεδίο συστηματικής ακαδημαϊκής έρευνας. Από εδώ και πέρα, όπου γίνεται στην εργασία αυτή αναφορά σε εκπαίδευση τεχνητών νευρωνικών δικτύων, θα εννοείται επιτηρούμενη εκπαίδευση.

2.3 ΑΝΤΙΚΕΙΜΕΝΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ

Σε κάθε επανάληψη κατά την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου είναι απαραίτητο να γίνει μια σύγκριση μεταξύ των παραγόμενων εξόδων του τεχνητού νευρωνικού δικτύου και των επιθυμητών εξόδων του συνόλου εκπαίδευσης. Στη βάση αυτής της σύγκρισης στη συνέχεια γίνεται η αναπροσαρμογή των βαρών και των πολώσεων. Πιο συγκεκριμένα, για κάθε δεδομένο υπολογίζεται το σφάλμα μεταξύ παραγόμενης και επιθυμητής εξόδου, και στη συνέχεια υπολογίζεται η μέση τιμή αυτών των σφαλμάτων για όλο το σύνολο εκπαίδευσης.

Μία αντικειμενική συνάρτηση ενός τεχνητού νευρωνικού δικτύου μπορεί να είναι η ελαχιστοποίηση της μέσης τιμής των σφαλμάτων του συνόλου εκπαίδευσης [18]. Το αν πρέπει να γίνει ελαχιστοποίηση ή μεγιστοποίηση της αντικειμενικής συνάρτησης εξαρτάται από την αντικειμενική συνάρτηση που επιλέγεται καθώς και από τη φύση του προβλήματος. Η τιμή της αντικειμενικής συνάρτησης εξαρτάται από τα βάρη, τις πολώσεις και το σύνολο εκπαίδευσης του τεχνητού νευρωνικού δικτύου. Σκοπός της αντικειμενικής συνάρτησης είναι η ποσοτικοποίηση της απόκλισης μεταξύ παραγόμενων και επιθυμητών εξόδων συνολικά, και η ελαχιστοποίηση αυτής της απόκλισης. Ορισμένες πολύ συχνά χρησιμοποιούμενες αντικειμενικές συναρτήσεις είναι:

- Ελαχιστοποίηση του μέσου απόλυτου σφάλματος (Mean Absolute Error – MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.8)$$

- Ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (Mean Squared Error – MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.9)$$

- Ελαχιστοποίηση της ρίζας του μέσου τετραγωνικού σφάλματος (Root Mean Squared Error – RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.10)$$

όπου n το πλήθος των δεδομένων εκπαίδευσης, y_i η παραγόμενη έξοδος για το δείγμα i και \hat{y}_i η επιθυμητή έξοδος για το δείγμα i .

2.4 ΑΛΓΟΡΙΘΜΟΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Σε κάθε επανάληψη κατά την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου τα βάρη και οι πολώσεις αναπροσαρμόζονται κατάλληλα ώστε να ικανοποιείται η αντικειμενική συνάρτηση. Η μέθοδος με την οποία γίνεται αυτή η αναπροσαρμογή καθορίζεται από τον αλγόριθμο βελτιστοποίησης.

2.4.1 Αλγόριθμος προς τα πίσω διάδοσης του σφάλματος

Ο πιο διαδεδομένος αλγόριθμος βελτιστοποίησης για την εκπαίδευση τεχνητών νευρωνικών δικτύων εμπρόσθιας τροφοδότησης είναι ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος (back – propagation algorithm) ο οποίος επινοήθηκε από τους Minsky και Papert το 1969 [19], και αφορούσε τεχνητά νευρωνικά δίκτυα χωρίς κρυφά στρώματα. Μία επιτυχημένη γενίκευση του αλγορίθμου ώστε να αφορά οποιοδήποτε τεχνητό νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης προτάθηκε από τους Rumelhart, Hinton και Williams το 1986 [20].

Η αναπροσαρμογή των βαρών και των πολώσεων βασίζεται στην μέθοδο βελτιστοποίησης «μείωση της κλίσης» (gradient descent) η οποία εφαρμόζεται στην αντικειμενική συνάρτηση που έχει επιλεγεί [14]. Τα βάρη και οι πολώσεις μεταβάλλονται σε κάθε επανάληψη σύμφωνα με τις εξισώσεις (2.11) και (2.12):

$$\Delta w_{pj,qk}(n+1) = lr \cdot \delta_{qk} \cdot y_{pj} + m \cdot \Delta w_{pj,qk}(n) \quad (2.11)$$

$$w_{pj,qk}(n+1) = w_{pj,qk}(n) + \Delta w_{pj,qk}(n+1) \quad (2.12)$$

όπου n ο αριθμός της τρέχουσας επανάληψης, $w_{pj,qk}$ η τιμή του βάρους που συνδέει τον νευρώνα p του στρώματος j με τον νευρώνα q του στρώματος k , δ_{qk} η τιμή του συντελεστή δ για τον νευρώνα q του στρώματος k , y_{pj} η έξοδος του νευρώνα p του στρώματος j , lr ο ρυθμός μάθησης και m η ορμή (momentum).

Για τον υπολογισμό του συντελεστή δ γίνεται χρήση του γενικευμένου κανόνα δ . Ο γενικευμένος κανόνας δ περιλαμβάνει δύο φάσεις [16]. Στην πρώτη φάση κάθε είσοδος x προωθείται «προς τα μπροστά» έτσι ώστε να υπολογιστούν οι έξοδοι των νευρώνων του στρώματος εξόδου. Οι έξοδοι αυτές συγκρίνονται με τις επιθυμητές εξόδους και υπολογίζεται η αντικειμενική συνάρτηση (μετρητικό σφάλματος) που έχει επιλεγεί. Στη δεύτερη φάση, το σφάλμα αυτό μεταδίδεται προς τα πίσω, γίνεται ο υπολογισμός του δ για κάθε νευρώνα του τεχνητού νευρωνικού δικτύου και στη συνέχεια γίνεται η αναπροσαρμογή των βαρών και των πολώσεων.

Ο συντελεστής δ υπολογίζεται με διαφορετικό τρόπο ανάλογα τη θέση του βάρους που πρέπει να αναπροσαρμοστεί. Εάν το βάρος αφορά τη σύνδεση νευρώνα κρυφού στρώματος με νευρώνα του στρώματος εξόδου, ο συντελεστής δ υπολογίζεται ως εξής:

$$\delta_q^o = (\hat{y}_q - y_q) \frac{\partial f^o}{\partial z_q} \quad (2.13)$$

όπου δ_q^o ο συντελεστής δ του νευρώνα q του στρώματος εξόδου, \hat{y}_q η επιθυμητή έξοδος του νευρώνα q του στρώματος εξόδου, y_q η παραγόμενη έξοδος του νευρώνα q του στρώματος εξόδου, f^o η συνάρτηση ενεργοποίησης του στρώματος εξόδου και z_q η είσοδος του νευρώνα q του στρώματος εξόδου. Για νευρώνες που δεν ανήκουν στο στρώμα εξόδου δεν υπάρχει επιθυμητή έξοδος, επομένως εάν το βάρος αφορά τη σύνδεση μεταξύ νευρώνων κρυφών στρωμάτων ή νευρώνα του στρώματος εισόδου με νευρώνα κρυφού στρώματος, ο συντελεστής δ υπολογίζεται ως εξής:

$$\delta_{qk} = \frac{\partial f^k}{\partial z_{qk}} \sum_{p=1}^{N_o} \delta_p^j w_{qk,pj} \quad (2.14)$$

όπου δ_{qk} ο συντελεστής δ του νευρώνα q του στρώματος k , f^k η συνάρτηση ενεργοποίησης του στρώματος k , z_{qk} η είσοδος του νευρώνα q του στρώματος k , N_o το πλήθος των νευρώνων του στρώματος j , δ_p^j ο συντελεστής δ του νευρώνα p του στρώματος j , $w_{qk,pj}$ το βάρος που συνδέει τον νευρώνα q του στρώματος k με τον νευρώνα p του στρώματος j . Τέλος, ο ρυθμός μάθησης lr και η ορμή m λαμβάνουν τιμές μεταξύ του 0.01 και του 1. Η ορμή m συνεισφέρει στην αποφυγή παγίδευσης του αλγορίθμου σε κάποιο τοπικό βέλτιστο, ενώ ο ρυθμός μάθησης lr σχετίζεται με την ταχύτητα σύγκλισης του αλγορίθμου σε κάποιο βέλτιστο [14].

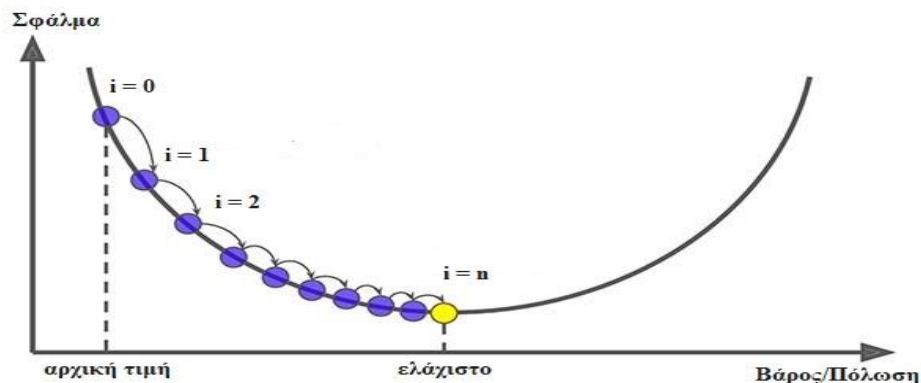
Ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος με τη μέθοδο βελτιστοποίησης «μείωση της κλίσης» λόγω της απλότητας στη χρήση του αλλά και της ταχύτητάς του, αποτελεί τον πιο διαδεδομένο αλγόριθμο βελτιστοποίησης με πληθώρα εφαρμογών. Η αναπροσαρμογή των βαρών και των πολώσεων σε κάθε επανάληψη γίνεται με τέτοιο τρόπο ώστε να οδηγεί στη μείωση της παραγώγου της μέσης τιμής του μετρητικού σφάλματος, προκειμένου να τείνει προς το μηδέν και άρα σε κάποιο ελάχιστο, τοπικό ή ολικό (Σχήμα 2.4). Ωστόσο, παρουσιάζει και κάποια μειονεκτήματα. Το μεγαλύτερο μειονέκτημά του είναι η ευκολία με την οποία μπορεί να παγιδευτεί σε κάποιο τοπικό βέλτιστο, ειδικά όταν το πρόβλημα παρουσιάζει μεγάλη πολυπλοκότητα [20]. Ταυτόχρονα, αφορά μόνο προβλήματα στα οποία το μετρητικό σφάλματος και κατά επέκταση η αντικειμενική συνάρτηση μπορούν να παραγωγιστούν [1].

2.4.2 Μεταερευνητικοί αλγόριθμοι

Εάν η αντικειμενική συνάρτηση είναι μη παραγωγίσιμη, δηλαδή παρουσιάζει κάποια ασυνέχεια, κλασικοί αλγόριθμοι βελτιστοποίησης που βασίζονται στη μέθοδο βελτιστοποίησης «μείωση της κλίσης», όπως ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος, είναι ακατάλληλοι. Οι στοχαστικοί αλγόριθμοι βελτιστοποίησης από την άλλη, οι οποίοι διακρίνονται σε δύο κατηγορίες, τους ευρετικούς και τους μεταερευνητικούς αλγορίθμους, δεν κάνουν χρήση παραγώγου [21].

Οι ευρετικοί αλγόριθμοι βασίζονται στην μέθοδο της δοκιμής και σφάλματος (trial and error). Οι ευρετικοί αλγόριθμοι κάνουν ελάχιστες έως καθόλου υποθέσεις για το πρόβλημα που πρόκειται να βελτιστοποιηθεί. Παρόλο που μπορούν να δώσουν ποιοτικές λύσεις σε εύλογα χρονικά διαστήματα, δεν υπάρχει εγγύηση ότι οι λύσεις αυτές θα είναι βέλτιστες, καθώς δεν είναι δυνατόν να εξερευνηθεί ολόκληρο το πεδίο τιμών. Οι ευρετικοί αλγόριθμοι χρησιμοποιούνται όταν δεν είναι αναγκαίο να βρεθούν οι βέλτιστες λύσεις, αλλά κάποιες λύσεις που υπολογίζονται γρήγορα και είναι επαρκώς ποιοτικές [21].

Οι μεταερευνητικοί αλγόριθμοι αποτελούν εξέλιξη των ευρετικών αλγορίθμων. Η μέθοδος αναζήτησης ενός βέλτιστου είναι παρόμοια με αυτή των ευρετικών αλγορίθμων. Ωστόσο, σε αντίθεση με τους ευρετικούς αλγορίθμους, οι μεταερευνητικοί αλγόριθμοι



Σχήμα 2.4: Τα βάρη και οι πολώσεις αναπροσαρμόζονται σε κάθε επανάληψη έτσι ώστε να μειώνεται η κλίση της μέσης τιμής του μετρητικού σφάλματος (μέθοδος gradient descent).

εμπλέκουν μεθόδους τυχαιοποίησης ώστε να αποφεύγεται η παγίδευση σε κάποιο τοπικό βέλτιστο, αλλά και μεθόδους τοπικής αναζήτησης ώστε να εξασφαλίζεται η εύρεση μιας βέλτιστης λύσης. Κατά συνέπεια, οι μεταερευτικοί αλγόριθμοι είναι κατάλληλοι για προβλήματα καθολικής βελτιστοποίησης [21].

Οι μεταερευτικοί αλγόριθμοι μπορούν να ταξινομηθούν με πολλούς τρόπους. Ένας τρόπος είναι η ταξινόμησή τους στις εξής δύο κατηγορίες: (i) στους αλγόριθμους που στηρίζονται σε πληθυσμούς (πολλαπλές λύσεις σε κάθε επανάληψη) και (ii) στους αλγόριθμους που ακολουθούν την τροχιά μιας λύσης [21]. Ορισμένοι συχνά χρησιμοποιούμενοι μεταερευτικοί αλγόριθμοι είναι: ο αλγόριθμος προσομοιωμένης απόπτωσης (simulated annealing) [22], ο γενετικός αλγόριθμος (genetic algorithm) [23], η βελτιστοποίηση αποικίας μυρμηγκιών (ant colony optimization) [24], ο αλγόριθμος νυχτερίδας (bat algorithm) [25], και η βελτιστοποίηση σμήνους σωματιδίων (particle swarm optimization) [2]. Από αυτούς, ο αλγόριθμος προσομοιωμένης απόπτωσης ακολουθεί την τροχιά μιας λύσης, ενώ οι υπόλοιποι αλγόριθμοι στηρίζονται σε πληθυσμούς.

2.4.3 Βελτιστοποίηση Σμήνους Σωματιδίων

Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (particle swarm optimization) είναι ένας μεταερευτικός αλγόριθμος που στηρίζεται σε πληθυσμούς και προτάθηκε από τους Kennedy και Eberhart [2]. Ανήκει στην κατηγορία των αλγορίθμων που αντλούν την έμπνευσή τους από κάποιο φυσικό φαινόμενο. Συγκεκριμένα, προσομοιάζει την κοινωνική συμπεριφορά και κίνηση των πτηνών που συγκροτούν ένα σμήνος για την εύρεση τροφής.

Όπως λοιπόν τα πτηνά κινούνται μέσα σε ένα σμήνος ακολουθώντας κάποιους κανόνες ώστε να βρουν τροφή, έτσι και τα σωματίδια του αλγορίθμου κινούνται στο πεδίο ορισμού του προβλήματος ώστε να αναζητήσουν τη βέλτιστη λύση. Κάθε σωματίδιο αποτελεί μια υποψήφια λύση, και προσδιορίζεται από τρεις παραμέτρους: (i) την τρέχουσα θέση του \vec{x} , (ii) την ταχύτητά του \vec{v} και (iii) την καλύτερη θέση στην οποία έχει βρεθεί (Pbest). Η θέση ενός σωματιδίου αντιστοιχεί σε μια υποψήφια λύση. Στην περίπτωση που η βελτιστοποίηση σμήνους σωματιδίων χρησιμοποιείται ως αλγόριθμος βελτιστοποίησης τεχνητού νευρωνικού δικτύου, η θέση ενός σωματιδίου είναι ένα διάνυσμα τιμών που ταυτίζεται με τα βάρη και τις πολώσεις της τρέχουσας επανάληψης. Η ταχύτητα των σωματιδίων σχετίζεται με την κατεύθυνση προς την οποία θα κινηθούν, δηλαδή με το διάνυσμα μεταβολής των βαρών και των πολώσεων. Η αξιολόγηση της θέσης ενός σωματιδίου γίνεται κάνοντας χρήση της αντικειμενικής συνάρτησης. Κάθε σωματίδιο φέρει ως πληροφορία την καλύτερη θέση στην οποία έχει βρεθεί, δηλαδή τη θέση εκείνη για την οποία ικανοποιείται περισσότερο η αντικειμενική συνάρτηση. Τέλος, σε επίπεδο σμήνους, αποθηκεύεται και η συνολικά καλύτερη θέση (Gbest), δηλαδή η καλύτερη θέση που έχει καταγραφεί μεταξύ όλων των σωματιδίων του σμήνους έως την τρέχουσα επανάληψη.

Σε κάθε επανάληψη, η κατεύθυνση και το μέτρο της κίνησης ενός σωματιδίου είναι συνάρτηση τριών παραγόντων (Σχήμα 2.5): (i) του διανύσματος της ταχύτητας του σωματιδίου κατά τη διάρκεια της προηγούμενης επανάληψης, (ii) της ατομικά καλύτερης του θέσης (Pbest), και (iii) της συνολικά καλύτερης θέσης (Gbest). Οι εξισώσεις με βάση τις

οποίες ανανεώνονται η ταχύτητα και η θέση του κάθε σωματιδίου σε κάθε επανάληψη είναι οι εξής [26]:

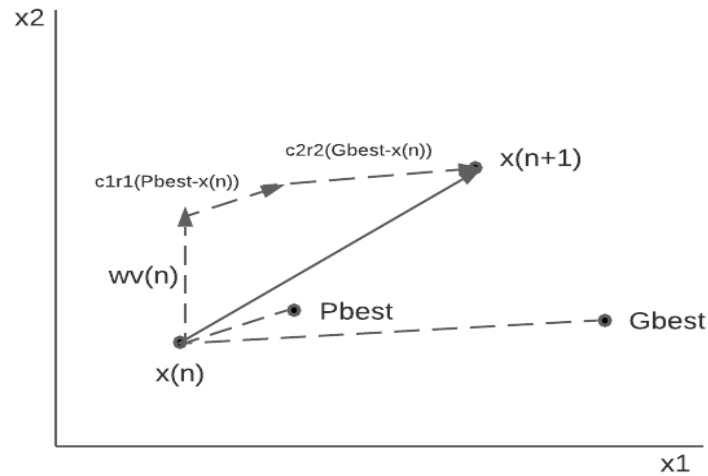
$$\vec{v}(n+1) = w\vec{v}(n) + c_1 r_1 (\vec{P}best - \vec{x}(n)) + c_2 r_2 (\vec{G}best - \vec{x}(n)) \quad (2.15)$$

$$\vec{x}(n+1) = \vec{x}(n) + \vec{v}(n+1) \quad (2.16)$$

Οι εξισώσεις (2.15) και (2.16) αφορούν μία διάσταση. Αν το πρόβλημα έχει περισσότερες διαστάσεις, τότε οι εξισώσεις (2.15) και (2.16) εφαρμόζονται για κάθε διάσταση. Οι συντεταγμένες της θέσης x αντιστοιχούν στα βάρη και στις πολώσεις, επομένως αποτελούν μια υποψήφια λύση του προβλήματος. Το w ονομάζεται βάρος αδράνειας (inertia weight) και παίρνει τιμές στο διάστημα $[0.1, 1.2]$. Σχετίζεται με τη δυνατότητα ενός σωματιδίου να αλλάζει κατεύθυνση. Όσο μικρότερο είναι το w , τόσο γρηγορότερα συγκλίνουν τα σωματίδια, ενώ όσο μεγαλώνει, τα σωματίδια δίνουν περισσότερο βάρος στην τρέχουσα κατεύθυνσή τους παρά στις καλύτερες καταγεγραμμένες θέσεις με αποτέλεσμα βραδύτερη σύγκλιση. Οι συντελεστές c_1, c_2 παίρνουν τιμές στο διάστημα $[1, 2]$ και είναι στην ουσία συντελεστές βαρύτητας. Το c_1 αφορά την ατομικά καλύτερη θέση του εκάστοτε σωματιδίου ενώ το c_2 την καλύτερη θέση συνολικά μέσα στο σμήνος. Τα r_1, r_2 αποτελούν τυχαίες μεταβλητές στο διάστημα $[0, 1]$ που παράγονται σε κάθε επανάληψη για κάθε σωματίδιο. Τέλος, το n είναι ο αριθμός της τρέχουσας επανάληψης.

Η θέση και η ταχύτητα κάθε σωματιδίου αρχικοποιείται τυχαία ή με κάποια μέθοδο αρχικοποίησης. Για κάθε σωματίδιο, η αρχική του θέση ορίζεται ως η ατομικά καλύτερή του θέση ($Pbest$). Στη συνέχεια υπολογίζεται η συνολικά καλύτερη θέση μέσα στο σμήνος ($Gbest$). Η ταχύτητα και η θέση κάθε σωματιδίου ανανεώνονται σύμφωνα με τις εξισώσεις (2.12) και (2.13) και γίνεται αξιολόγηση της νέας τους θέσης. Εφόσον η ταχύτητα και η θέση βγουν έξω από το πεδίο ορισμού του προβλήματος, επαναφέρονται εντός των ορίων τους. Εάν είναι απαραίτητο, ανανεώνονται τα $Pbest$ κάθε σωματιδίου, καθώς και το $Gbest$. Η διαδικασία επαναλαμβάνεται έως ότου ικανοποιηθεί κάποιο κριτήριο τερματισμού του αλγορίθμου, όπως η συμπλήρωση του μέγιστου αριθμού επαναλήψεων ή η συμπλήρωση πολλών επαναλήψεων χωρίς μεταβολή του $Gbest$. Η τελική συνολικά καλύτερη θέση ($Gbest$) αποτελεί και τη λύση του προβλήματος.

Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων χρησιμοποιείται όλο και συχνότερα τα τελευταία χρόνια, κυρίως σε προβλήματα αριθμητικής βελτιστοποίησης, λόγω των πολλών πλεονεκτημάτων του [26]. Έχει πληθώρα πρακτικών και ερευνητικών εφαρμογών και είναι ιδανικός για την επίλυση προβλημάτων βελτιστοποίησης παραμέτρων [27]. Επίσης, καθώς ανήκει στην κατηγορία των μεταερευνητικών αλγορίθμων, δεν απαιτεί να είναι παραγωγίσιμη η αντικειμενική συνάρτηση. Είναι αποδοτικότερος, ταχύτερος και λιγότερο πολύπλοκος σε σχέση με άλλους μεταερευνητικούς αλγορίθμους. Οι λίγες υπερπαραμέτροι που έχει (πλήθος σωματιδίων, c_1, c_2, w), τον καθιστούν ευέλικτο και πρακτικό. Τέλος, σε ότι αφορά τα μειονεκτημάτά του, λόγω της μεγάλης ταχύτητας σύγκλισης που έχει είναι αρκετά πιθανό τα σωματίδια να παγιδευτούν σε κάποιο μη ολικό βέλτιστο.



Σχήμα 2.5: Η μεταβολή της θέσης ενός σωματιδίου για ένα πρόβλημα δύο διαστάσεων

2.5 ΠΟΛΥΚΡΙΤΗΡΙΑΚΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ

Ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος και η βελτιστοποίηση σμήνους σωματιδίων που περιγράφηκαν στις Ενότητες 2.4.1 και 2.4.3, αντίστοιχα, αφορούν τη βελτιστοποίηση μίας μόνο αντικειμενικής συνάρτησης. Τα περισσότερα προβλήματα ωστόσο του πραγματικού κόσμου αφορούν την ταυτόχρονη βελτιστοποίηση πολλών, συχνά αντικρουόμενων, κριτηρίων. Η ταυτόχρονη βελτιστοποίηση πολλών αντικειμενικών συναρτήσεων ονομάζεται πολυκριτηριακή βελτιστοποίηση (multi – objective optimization).

Ο γενικός τύπος ενός προβλήματος πολυκριτηριακής βελτιστοποίησης περιγράφεται από την σχέση 2.14 [28]:

$$\begin{array}{ll}
 \text{Μεγιστοποίηση ή Ελαχιστοποίηση} & f_m(x), & m = 1, 2, \dots, M \\
 \text{περιορισμοί} & g_j(x) \geq 0, & j = 1, 2, \dots, J \\
 & h_k(x) = 0, & k = 1, 2, \dots, K \\
 & x_i^L \leq x_i \leq x_i^U, & i = 1, 2, \dots, I
 \end{array} \quad (2.17)$$

όπου $f(x) = (f_1(x), f_2(x), \dots, f_M(x))^T$ οι αντικειμενικές συναρτήσεις του προβλήματος, $g_j(x)$ και $h_k(x)$ οι περιορισμοί του προβλήματος, $x = (x_1, x_2, \dots, x_I)^T$ μία λύση του προβλήματος, x_i^L και x_i^U το κάτω και το άνω όριο της διάστασης i της λύσης x .

Ένα κλασικό παράδειγμα δύο αντικρουόμενων αντικειμενικών συναρτήσεων είναι η σχέση κόστους και ποιότητας όπου η μία αντικειμενική συνάρτηση έγκειται στην ελαχιστοποίηση του κόστους ενώ η άλλη στη μεγιστοποίηση της ποιότητας. Ωστόσο, συνήθως η μείωση του κόστους συνεπάγεται και μείωση της ποιότητας, ενώ αντίστοιχα η αύξηση της ποιότητας συνεπάγεται αύξηση του κόστους.

2.5.1 Μέτωπο Pareto

Ένας τρόπος επίλυσης ενός προβλήματος πολλών αντικρουόμενων αντικειμενικών συναρτήσεων είναι η μετατροπή του σε πρόβλημα μίας αντικειμενικής συνάρτησης και, στη συνέχεια, η ελαχιστοποίηση ή μεγιστοποίηση αυτής της μίας αντικειμενικής συνάρτησης. Αυτό μπορεί να γίνει είτε με τον συνδυασμό των αντικειμενικών συναρτήσεων σε μία ενιαία αντικειμενική συνάρτηση, ή με την ιεράρχηση των αντικειμενικών συναρτήσεων ανάλογα με την σπουδαιότητά τους και τον καθορισμό αντίστοιχων βαρών για την κάθε συνάρτηση. Ωστόσο, η ενοποίηση των αντικειμενικών συναρτήσεων σε μία συνάρτηση πολλές φορές δεν είναι εφικτή, ενώ η ιεράρχησή τους μπορεί να είναι πολύπλοκη και απαιτεί περαιτέρω πληροφορίες για το πρόβλημα, που ενδέχεται να μην είναι διαθέσιμες.

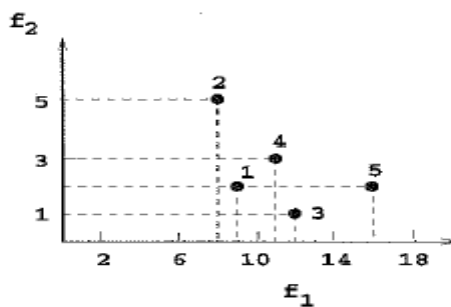
Ένα πρόβλημα πολυκριτηριακής βελτιστοποίησης δεν έχει μία και μοναδική λύση. Μια λύση μπορεί να υπερισχύει σε κάποιες αντικειμενικές συναρτήσεις, ενώ μια άλλη λύση μπορεί να υπερισχύει σε κάποιες άλλες. Επομένως, σε αντίθεση με τα προβλήματα μιας αντικειμενικής συνάρτησης, τα προβλήματα πολυκριτηριακής βελτιστοποίησης περιλαμβάνουν ένα σύνολο λύσεων. Η εύρεση αυτών των λύσεων βασίζεται στην κατά Pareto υπερίσχυση (Pareto dominance). Σύμφωνα με την κατά Pareto υπερίσχυση, μία λύση x_1 υπερισχύει μιας λύσης x_2 εάν [28]:

- Η λύση x_1 είναι καλύτερη από τη x_2 στη βελτιστοποίηση τουλάχιστον μιας αντικειμενικής συνάρτησης.
- Η λύση x_1 δεν είναι χειρότερη από τη x_2 στη βελτιστοποίηση καμιάς αντικειμενικής συνάρτησης.

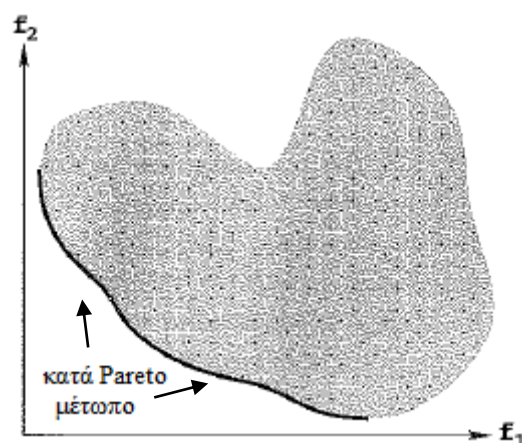
Επομένως, μεταξύ δύο λύσεων x_1, x_2 μπορούν να ισχύουν οι ακόλουθες τρεις σχέσεις [28]:

- i. Η λύση x_1 να υπερισχύει της λύσης x_2
- ii. Η λύση x_2 να υπερισχύει της λύσης x_1
- iii. Να μην υπερισχύει καμία λύση από τις δύο (μη κυριαρχούμενες λύσεις)

Στο Σχήμα 2.6 απεικονίζονται πέντε λύσεις ενός προβλήματος βελτιστοποίησης δύο αντικειμενικών συναρτήσεων, η μία έγκειται στη μεγιστοποίηση της f_1 , και η άλλη στην ελαχιστοποίηση της f_2 . Η λύση 1 υπερισχύει της λύσης 2, καθώς είναι καλύτερη και στις δύο αντικειμενικές συναρτήσεις. Η λύση 5 από την άλλη υπερισχύει της λύσης 1 καθώς είναι καλύτερη στην αντικειμενική συνάρτηση f_1 και όχι χειρότερη στην αντικειμενική συνάρτηση f_2 . Τέλος, οι λύσεις 3 και 5 είναι μη κυριαρχούμενες μεταξύ τους, καθώς η 3 είναι καλύτερη



Σχήμα 2.6: Πέντε λύσεις ενός προβλήματος βελτιστοποίησης δύο αντικειμενικών συναρτήσεων



Σχήμα 2.7: Το μέτωπο Pareto ενός προβλήματος ελαχιστοποίησης δύο αντικειμενικών συναρτήσεων.

στην αντικειμενική συνάρτηση f_2 , ενώ η 5 είναι καλύτερη στην αντικειμενική συνάρτηση f_1 .

Το σύνολο όλων των μη κυριαρχούμενων λύσεων στο πεδίο ορισμού του προβλήματος ονομάζεται «μέτωπο Pareto» (Pareto front) [28]. Στο Σχήμα 2.7 το γραμμοσκιασμένο τμήμα αποτελεί το σύνολο των πιθανών λύσεων ενός προβλήματος βελτιστοποίησης δύο αντικειμενικών συναρτήσεων (ελαχιστοποίηση της f_1 και ελαχιστοποίηση της f_2), ενώ η μαύρη έντονη γραμμή αποτελεί το μέτωπο Pareto.

Η πολυκριτηριακή βελτιστοποίηση έχει δύο στόχους ισάξιας σπουδαιότητας. Πρώτον, να βρεθεί ένα σύνολο λύσεων όσο πιο κοντά γίνεται στο μέτωπο Pareto. Δεύτερον, το σύνολο λύσεων που θα βρεθεί να διατηρεί όσο το δυνατόν μεγαλύτερη ποικιλομορφία [28].

Μέσω της πολυκριτηριακής βελτιστοποίησης είναι δυνατό να μοντελοποιηθεί οποιοδήποτε πρόβλημα, ανεξάρτητα από το πόσο πολύπλοκο είναι και από το αν είναι αντικρουόμενες οι αντικειμενικές του συναρτήσεις. Η μοντελοποίηση είναι εύκολη, και δεν απαιτείται η εκ των προτέρων γνώση της σπουδαιότητας κάθε αντικειμενικής συνάρτησης. Η επίλυση ενός πολυκριτηριακού προβλήματος περιλαμβάνει ένα σύνολο μη κυριαρχούμενων λύσεων αντί για μία και μοναδική λύση, δίνοντας έτσι τη δυνατότητα της επιλογής, κάτι χρήσιμο καθώς για το ίδιο πρόβλημα, ανάλογα τις συνθήκες, μπορεί να είναι προτιμότερη διαφορετική λύση κάθε φορά.

2.5.2 Πολυκριτηριακή Βελτιστοποίηση Σμήνους Σωματιδίων

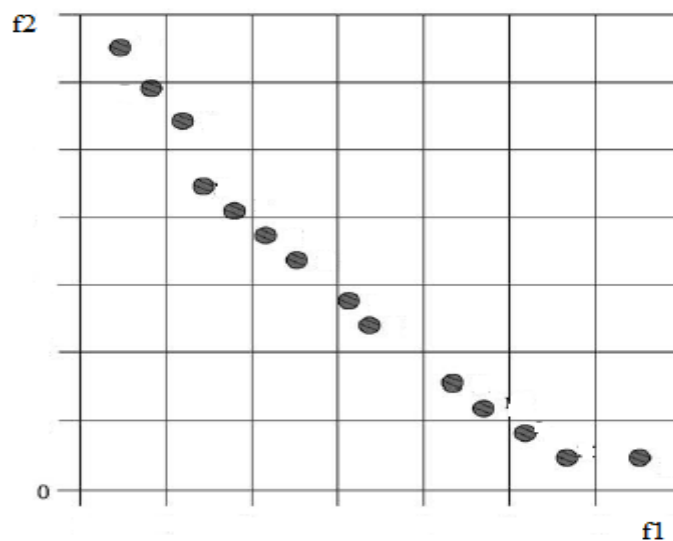
Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων για την επίλυση προβλημάτων πολλαπλών αντικειμενικών συναρτήσεων (Multi – Objective Particle Swarm Optimization) προτάθηκε από τους Coello και Lechuga το 2002 [3]. Η βασική ιδέα του αλγορίθμου είναι ίδια με αυτήν που περιγράφηκε στην Ενότητα 2.4.3. Για την προσαρμογή του αλγορίθμου στην πολυκριτηριακή βελτιστοποίηση γίνεται χρήση δύο επιπλέον μηχανισμών, της κατά Pareto υπερίσχυσης και της γεωγραφικής αναπαράστασης των μη κυριαρχούμενων λύσεων.

Αφού αρχικοποιηθούν οι θέσεις, οι ταχύτητες και οι ατομικά καλύτερες θέσεις (Pbest) των σωματιδίων, γίνεται σύγκριση μεταξύ των λύσεων που φέρει κάθε σωματίδιο με βάση την κατά Pareto κυριαρχία. Οι μη κυριαρχούμενες λύσεις αποθηκεύονται σε ένα αρχείο αποθήκευσης (repository), ώστε στη συνέχεια να επιλεγεί από εκεί η συνολικά καλύτερη θέση (Gbest), που θα επιτελέσει τον ρόλο του οδηγού για την ανανέωση της ταχύτητας και της θέσης κάθε σωματιδίου.

Αφού βρεθούν και αποθηκευτούν όλες οι μη κυριαρχούμενες λύσεις της τρέχουσας επανάληψης, ο χώρος ο οποίος εξερευνήθηκε από τα σωματίδια χωρίζεται σε υπερκύβους (hyper cubes) έτσι ώστε να δημιουργηθεί ένα πλέγμα. Οι άξονες του πλέγματος αντιστοιχούν στις αντικειμενικές συναρτήσεις του προβλήματος, επομένως κάθε μη κυριαρχούμενη λύση του αρχείου αποθήκευσης αναπαρίσταται γεωγραφικά πάνω στο πλέγμα με βάση την απόδοσή της σε κάθε αντικειμενική συνάρτηση. Στο Σχήμα 2.8 φαίνεται το πλέγμα που δημιουργήθηκε για ένα πρόβλημα ελαχιστοποίησης δύο αντικειμενικών συναρτήσεων f_1 και f_2 . Κάθε υπερκύβος λαμβάνει μία τιμή h , που προκύπτει από την διαίρεση ενός οποιοδήποτε αριθμού $x > 1$ με το πλήθος των μη κυριαρχούμενων λύσεων που βρίσκονται εντός του υπερκύβου. Οι λύσεις εντός ενός υπερκύβου λαμβάνουν την τιμή h του υπερκύβου. Η τιμή h εκφράζει τον συνωστισμό λύσεων εντός του υπερκύβου. Όσες περισσότερες λύσεις υπάρχουν εντός ενός υπερκύβου, τόσο μικρότερο θα είναι το h .

Η συνολικά καλύτερη θέση (Gbest) για την ανανέωση της ταχύτητας κάθε σωματιδίου επιλέγεται ως εξής: αφού γίνει η γεωγραφική αναπαράσταση των μη κυριαρχούμενων λύσεων και η ανάθεση της τιμής h σε κάθε λύση, η συνολικά καλύτερη θέση (Gbest) επιλέγεται τυχαία μεταξύ των λύσεων. Η πιθανότητα να επιλεγεί μια λύση είναι ανάλογη με την τιμή h που έχει η συγκεκριμένη λύση. Επομένως είναι πιο πιθανό να επιλεγεί μια λύση από έναν υπερκύβο με μικρό συνωστισμό λύσεων παρά μια λύση από έναν υπερκύβο με μεγάλο συνωστισμό. Με αυτόν τον τρόπο εξασφαλίζεται η απαραίτητη ποικιλομορφία μεταξύ των λύσεων.

Η ατομικά καλύτερη θέση (Pbest) κάθε σωματιδίου επιλέγεται με βάση την κατά



Σχήμα 2.8: Γεωγραφική αναπαράσταση στο δημιουργημένο πλέγμα των μη κυριαρχούμενων λύσεων ενός προβλήματος ελαχιστοποίησης δύο αντικειμενικών συναρτήσεων

Pareto υπερίσχυση. Εάν δύο λύσεις είναι μη κυριαρχούμενες μεταξύ τους, επιλέγεται τυχαία μία ως Pbest. Οι ταχύτητες και οι θέσεις των σωματιδίων στη συνέχεια ανανεώνονται με βάση τις εξισώσεις (2.15) και (2.16). Στο τέλος κάθε επανάληψης, εάν το αρχείο αποθήκευσης παρουσιάσει υπερχείλιση, διαγράφονται οι μη κυριαρχούμενες λύσεις με τη μικρότερη τιμή h , ώστε να εξασφαλίζεται η μεγαλύτερη δυνατή ποικιλομορφία. Ο αλγόριθμος επαναλαμβάνεται έως ότου ικανοποιηθούν τα κριτήρια τερματισμού. Οι τελικές λύσεις που είναι αποθηκευμένες στο αρχείο αποθήκευσης, είναι και οι λύσεις του προβλήματος.

Ο αλγόριθμος της πολυκριτηριακής βελτιστοποίησης σμήνους σωματιδίων είναι αποδοτικός και άμεσα συγκρίσιμος με άλλους πετυχημένους μεταερευτικούς αλγορίθμους πολυκριτηριακής βελτιστοποίησης, όπως ο γενετικός αλγόριθμος μη κυριαρχούμενης ταξινόμησης II (non dominated sorting genetic algorithm II – NSGA II) [29]. Διατηρεί όλα τα πλεονεκτήματα της πολυκριτηριακής βελτιστοποίησης, σε συνδυασμό με τα πλεονεκτήματα του αλγορίθμου βελτιστοποίησης σμήνους σωματιδίων. Παράγει αντίστοιχα αποτελέσματα με αυτά του γενετικού αλγορίθμου μη κυριαρχούμενης ταξινόμησης II, έχοντας μεγαλύτερη ταχύτητα σύγκλισης και μικρότερο υπολογιστικό φορτίο [3].

ΠΡΟΒΛΕΨΗ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ

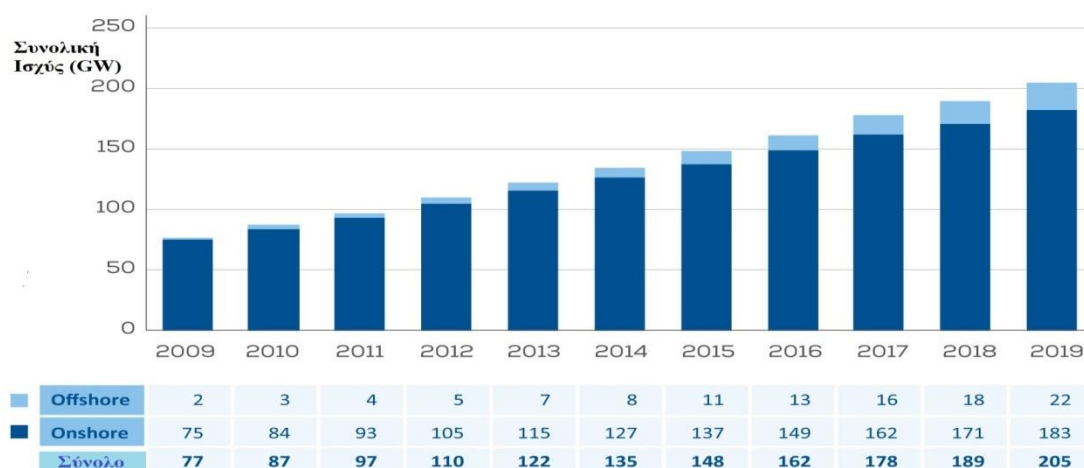
3.1 Η ΣΗΜΑΣΙΑ ΤΗΣ ΠΡΟΒΛΕΨΗΣ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ

Η κλιματική αλλαγή και οι συνεπαγόμενες επιπτώσεις που έχει στον πλανήτη, έχουν πυροδοτήσει τις τελευταίες δεκαετίες έντονες συζητήσεις και δράσεις σχετικά με την αντιμετώπισή τους. Η κλιματική κρίση αποτελεί κατά πολλούς τη μεγαλύτερη κρίση που θα αντιμετωπίσει η ανθρωπότητα τον 21^ο αιώνα. Σε ότι αφορά τον περιορισμό της κλιματικής αλλαγής, κοινή συνισταμένη μεταξύ των ειδικών αποτελεί η μείωση των εκπομπών αερίων του θερμοκηπίου, όπως του διοξειδίου του άνθρακα (CO₂). Πρόσφατες έρευνες υποδεικνύουν ότι πρέπει να μειωθούν κατά 45% οι εκπομπές CO₂ που οφείλονται σε ανθρώπινη δραστηριότητα έως το 2030, σε σχέση με το 2010, ώστε η αύξηση της θερμοκρασίας να περιοριστεί στους 1.5° C [30].

Μεγάλο μερίδιο στις εκπομπές CO₂ έχει η βιομηχανία παραγωγής ηλεκτρικής ενέργειας, λόγω της μεγάλης εξάρτησής της από τα ορυκτά καύσιμα. Ταυτόχρονα, πολλοί ισχυρίζονται ότι ο πλανήτης βρίσκεται στα πρόθυρα μιας παγκόσμιας ενεργειακής κρίσης λόγω της εξάντλησης των αποθεμάτων των ορυκτών καυσίμων. Τα παραπάνω έχουν οδηγήσει στην όλο και αυξανόμενη διεξόδυση της ηλεκτρικής ενέργειας που προέρχεται από Ανανεώσιμες Πηγές Ενέργειας (ΑΠΕ) στο ηλεκτρικό δίκτυο. Η παραγωγή ηλεκτρικής ενέργειας από ΑΠΕ βασίζεται σε φυσικά φαινόμενα και διαδικασίες, επομένως σε σχέση με τις συμβατικές μονάδες παραγωγής έχουν λιγότερες εκπομπές αερίων του θερμοκηπίου και μικρότερο περιβαλλοντικό αποτύπωμα. Σύμφωνα με το Ευρωπαϊκό Κοινοβούλιο, ως ΑΠΕ θεωρούνται η αιολική, η ηλιακή, η υδροηλεκτρική, η γεωθερμική, η κυματική ενέργεια των ωκεανών, η ενέργεια που εκλύεται από τα αέρια σε χώρους υγειονομικής ταφής, η ενέργεια από αέρια μονάδων επεξεργασίας λυμάτων, και η ενέργεια από βιοαέρια [31]. Η Ευρωπαϊκή Επιτροπή εκτιμάει ότι το μερίδιο της ηλεκτρικής ενέργειας προερχόμενης από ΑΠΕ στη συνολική παραγόμενη ηλεκτρική ενέργεια, θα πρέπει να αυξηθεί από 21% που ήταν το 2015 σε τουλάχιστον 45% μέχρι το 2030 [32].

Την τελευταία δεκαετία αναπτύσσεται ραγδαία η παραγωγή ηλεκτρικής ενέργειας από αιολικούς σταθμούς παραγωγής (αιολική ενέργεια). Στο Σχήμα 3.1 φαίνεται η εξέλιξη της συνολικής εγκατεστημένης αιολικής ισχύος στην Ευρωπαϊκή Ένωση από το 2009 έως το 2019. Στην Ελλάδα, η συνολική εγκατεστημένη αιολική ισχύς το 2020 έφτανε το 15% της συνολικής παραγωγής ηλεκτρικής ισχύος [33]. Η αυξανόμενη αυτή διεξόδυση της αιολικής ενέργειας φέρει όμως ταυτόχρονα αρκετά προβλήματα και προκλήσεις.

Η πιο σημαντική πρόκληση των αιολικών σταθμών παραγωγής είναι η μεγάλη αύξηση της μεταβλητότητας και της αβεβαιότητας που προκαλεί στο σύστημα ηλεκτρικής ενέργειας η μεταβλητή και μη ελεγχόμενη παραγωγή των αιολικών σταθμών. Σε αντίθεση με τις συμβατικές μονάδες παραγωγής αλλά και κάποιες μονάδες παραγωγής που στηρίζονται σε

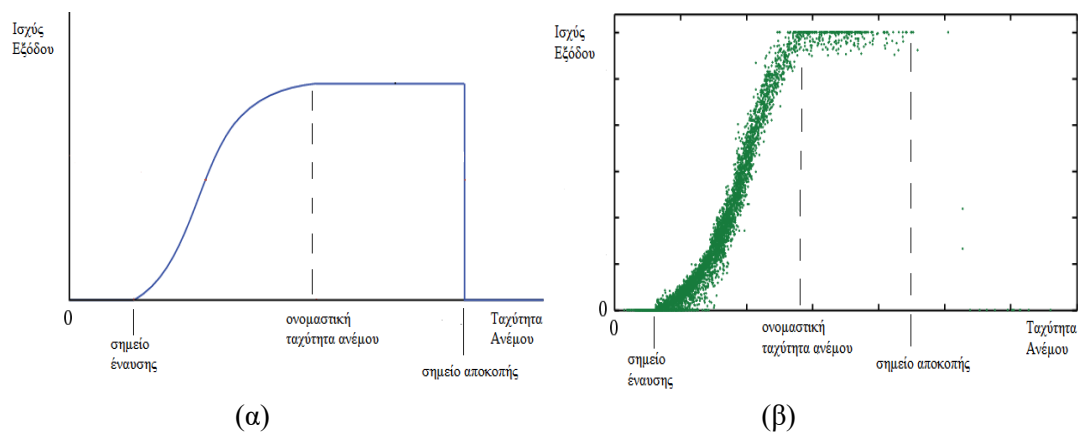


Σχήμα 3.1: Η συνολική εγκατεστημένη αιολική ισχύς στην Ευρωπαϊκή Ένωση

ΑΠΕ (π.χ. υδροηλεκτρικά), η παραγωγή αιολικής ισχύος είναι μόνο μερικώς ελεγχόμενη και χαρακτηρίζεται από μεγάλη στοχαστικότητα [32]. Η παραγωγή αιολικής ισχύος είναι κατά βάση συνάρτηση της ταχύτητας του ανέμου, αλλά και άλλων καιρικών και χωροταξικών παραγόντων, όπως η κατεύθυνση του ανέμου, η υγρασία και το φαινόμενο της σκίασης. Η τυχαία φύση του ανέμου και γενικότερα της ατμόσφαιρας προκαλεί συνεχείς αλλαγές στην παραγωγή της αιολικής ισχύος. Η αβεβαιότητα της βραχυπρόθεσμης καιρικής πρόβλεψης κάνει ακόμα δυσκολότερο τον έλεγχο της αιολικής παραγωγής. Στο Σχήμα 3.2(α), από την θεωρητική καμπύλη ισχύος μιας ανεμογεννήτριας, φαίνεται πως για ταχύτητες ανέμου μικρότερες από την ονομαστική, μικρές μεταβολές της ταχύτητας του ανέμου αντιστοιχούν σε μεγάλες μεταβολές της παραγωγής ισχύος.

Ακόμα και αν οι βραχυπρόθεσμες καιρικές προβλέψεις είναι άψογες, η πρόβλεψη αιολικής παραγωγής αποτελεί πρόκληση. Η αιολική παραγωγή, ακριβώς επειδή είναι συνάρτηση πολλών τυχαίων παραγόντων, αποτελεί μια στοχαστική διαδικασία. Ως στοχαστική διαδικασία ονομάζεται η διαδικασία που ακόμα και αν η αρχική της κατάσταση και οι εξισώσεις που την διέπουν είναι γνωστές, δεν είναι δυνατόν να προβλεφθεί η ακριβής πορεία της. Έτσι, ακόμα και αν δύο διαφορετικές στιγμές επικρατούν για την ίδια ανεμογεννήτρια οι ίδιες ακριβώς συνθήκες, δεν είναι σίγουρο ότι θα παραχθεί η ίδια ποσότητα αιολικής ισχύος. Στο Σχήμα 3.2(β) φαίνεται η πραγματική σχέση αιολικής ισχύος και ταχύτητας του ανέμου [34].

Η γνώση της παραγόμενης ισχύος κάθε στιγμή έχει μεγάλη σημασία για την κάλυψη των εθνικών ενεργειακών αναγκών και την αποφυγή σβέσης του συστήματος και άλλων ανεπιθύμητων συμβάντων. Οι παραγωγοί θα πρέπει να έχουν τη δυνατότητα να προβλέπουν σωστά την παραγωγή αιολικής ισχύος, ώστε να μπορούν έγκαιρα και αξιόπιστα να καταθέτουν τις προσφορές τους και να γίνεται με ακρίβεια ο κεντρικός σχεδιασμός. Όταν οι προβλέψεις της αιολικής παραγωγής είναι χαμηλότερες από την πραγματική παραγωγή των αιολικών σταθμών, οι συμβατικές μονάδες παραγωγής αναγκάζονται να λειτουργήσουν σε χαμηλότερη ισχύ, μειώνοντας την αποδοτικότητα του συστήματος. Ακόμη, εάν τα σφάλματα στην πρόβλεψη της αιολικής παραγωγής είναι αρκετά μεγάλα, για παράδειγμα, αν η πραγματική παραγωγή των αιολικών σταθμών είναι πολύ μεγαλύτερη από τις προβλέψεις,



Σχήμα 3.2: (α) η θεωρητική καμπύλη αιολικής ισχύος – ταχύτητας ανέμου, (β) η πραγματική καμπύλη αιολικής ισχύος – ταχύτητας ανέμου

είναι πιθανό η περίσσεια παραγόμενη αιολική ισχύς να μείνει ανεκμετάλλευτη. Όταν οι προβλέψεις της αιολικής παραγωγής είναι μεγαλύτερες από την πραγματική παραγωγή των αιολικών σταθμών, ενεργοποιούνται οι συμβατικές μονάδες παραγωγής ώστε να καλύψουν τη διαφορά, αυξάνοντας την παραγωγή αέριων ρύπων. Και στις δύο περιπτώσεις, αυξάνεται το κόστος λειτουργίας του συστήματος, ενώ λόγω της μείωσης της απόδοσης γίνεται σπατάλη παραγόμενης ενέργειας.

Η πρόβλεψη αιολικής παραγωγής έχει επομένως τεράστια σημασία και αποτελεί πεδίο συστηματικής ακαδημαϊκής έρευνας. Σε αντίθεση με την πρόβλεψη φορτίου που είναι πλέον μια ώριμη επιστήμη, η πρόβλεψη αιολικής παραγωγής εμφανίζει πολύ μεγαλύτερα σφάλματα που είναι της τάξης του 15–20% της ονομαστικής ισχύος του αιολικού σταθμού [35]. Προκειμένου να αντιμετωπιστούν οι παραπάνω προκλήσεις, έχουν αναπτυχθεί διάφορα μοντέλα πρόβλεψης αιολικής παραγωγής, άλλα απλά και άλλα πολύπλοκα, που βασίζουν τη λειτουργία τους σε μία σειρά από διαφορετικές επιστήμες.

3.2 ΠΡΟΒΛΕΨΗ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΜΕ ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Τα μοντέλα πρόβλεψης αιολικής παραγωγής μπορούν να διαχωριστούν σε δύο βασικές κατηγορίες με βάση τον τρόπο επεξεργασίας των μετεωρολογικών προβλέψεων: τα φυσικά μοντέλα και τα στατιστικά μοντέλα [36]. Στα φυσικά μοντέλα, αρχικά γίνεται η κατάλληλη επεξεργασία των μετεωρολογικών προβλέψεων, ώστε να προσαρμοστούν στην τοποθεσία του αιολικού σταθμού από την αρχική τοποθεσία της μετεωρολογικής πρόβλεψης. Στη συνέχεια, τα προσαρμοσμένα δεδομένα πρόβλεψης εφαρμόζονται στη θεωρητική καμπύλη αιολικής ισχύος κάθε ανεμογεννήτριας ώστε να παραχθούν οι τιμές αιολικής ισχύος.

Στα στατιστικά μοντέλα γίνεται χρήση ιστορικών μετεωρολογικών προβλέψεων και των αντίστοιχων ιστορικών δεδομένων παραγωγής ισχύος, τα οποία τροφοδοτούνται στο μοντέλο ώστε αυτό να προσεγγίσει τη μεταξύ τους σχέση. Τα στατιστικά μοντέλα χωρίζονται σε δύο περαιτέρω κατηγορίες, τα δομικά μοντέλα και τα «μαύρα κουτιά», ανάλογα με το αν χρειάζεται ο χρήστης να έχει κάποια τεχνογνωσία. Στα «μαύρα κουτιά», ο χρήστης δεν

χρειάζεται να έχει ιδιαίτερη τεχνογνωσία καθώς όλες οι διεργασίες επιτελούνται από το ίδιο το μοντέλο. Στην κατηγορία αυτή ανήκουν τα μοντέλα που βασίζονται στην τεχνητή νοημοσύνη, επομένως και τα μοντέλα που βασίζονται στα τεχνητά νευρωνικά δίκτυα [37]. Τα τεχνητά νευρωνικά δίκτυα έχουν μεγάλη ικανότητα γενίκευσης και προσέγγισης της σχέσης μεταξύ δεδομένων εισόδου και εξόδου. Σε αντίθεση με τα φυσικά μοντέλα που απαιτούν εξειδίκευση σε κάθε αιολικό σταθμό και ανεμογεννήτρια, τα μοντέλα που βασίζονται σε τεχνητά νευρωνικά δίκτυα αποδίδουν ανεξάρτητα από τα ειδικά χαρακτηριστικά που έχει κάθε αιολικός σταθμός. Τα στατιστικά μοντέλα που κάνουν χρήση μηχανικής μάθησης αλλά ταυτόχρονα αξιοποιούν κάποια μορφή τεχνογνωσίας του χρήστη, ανήκουν στα λεγόμενα «γκρι κουτιά» [37].

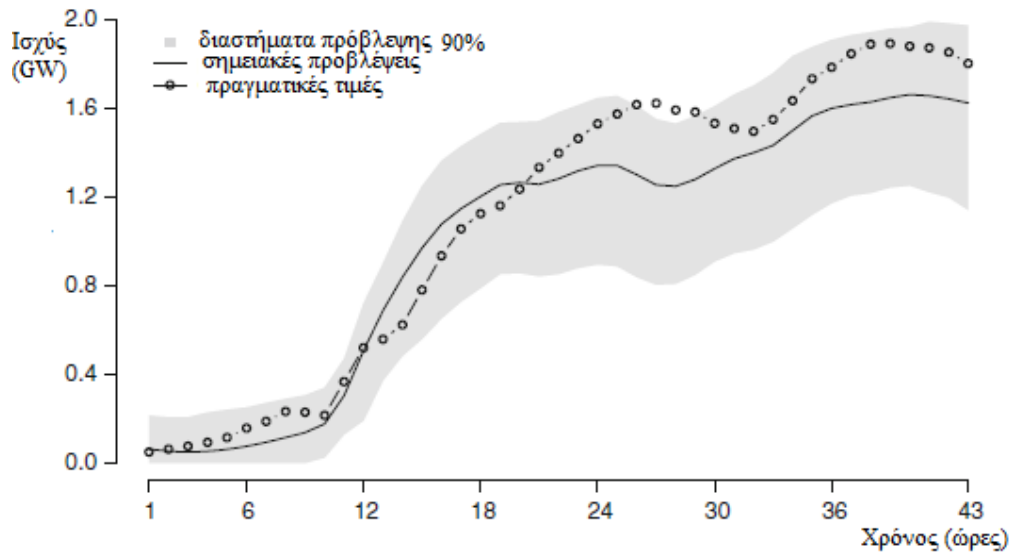
3.2.1 Αιτιοκρατικά μοντέλα

Ως αιτιοκρατικά μοντέλα (deterministic models) χαρακτηρίζονται τα μοντέλα που τη χρονική στιγμή t παράγουν μια σημειακή πρόβλεψη για τη χρονική στιγμή $t+k$ [38]. Από τη φύση της η σημειακή πρόβλεψη δίνει την αίσθηση ενός αιτιοκρατικού πλαισίου, δηλαδή μιας αρκετά μεγάλης σιγουριάς για την πρόβλεψη η οποία γίνεται. Ωστόσο τα αιτιοκρατικά μοντέλα χρησιμοποιούνται και σε στοχαστικά πλαίσια, όπου η πρόβλεψη μετατρέπεται σε εκτίμηση, με την έννοια ότι εισάγεται η έννοια της αβεβαιότητας στη σημειακή πρόβλεψη, ακόμα και αν δεν ποσοτικοποιείται με κάποιον τρόπο. Στην ουσία δηλαδή η πρόβλεψη που παράγεται είναι ο μέσος όρος όλων των συμβάντων που μπορούν να πραγματοποιηθούν τη χρονική στιγμή $t+k$ με βάση τη γνώση που υπάρχει έως τη χρονική στιγμή t [38].

Για πολλά χρόνια χρησιμοποιούνταν ως επί το πλείστον αιτιοκρατικά μοντέλα για την πρόβλεψη αιολικής παραγωγής. Η πρόβλεψη αιολικής παραγωγής ωστόσο αποτελεί στοχαστική διαδικασία, επομένως περιέχει μεγάλη αβεβαιότητα. Στο Σχήμα 3.3 φαίνονται οι σημειακές προβλέψεις αιολικής παραγωγής για όλους τους χειρσαίους αιολικούς σταθμούς της δυτικής Δανίας στις 4 Απριλίου 2007 και ώρα 00:00, με ωραία χρονική ανάλυση και χρονικό ορίζοντα 43 ωρών [38]. Από το Σχήμα 3.3 φαίνεται ότι είναι σχεδόν αδύνατον να γίνει σημειακή πρόβλεψη χωρίς σημαντικό σφάλμα, επομένως μπορεί να εξαχθεί μόνο μια μέση τιμή αιολικής παραγωγής για ένα ορισμένο χρονικό ορίζοντα. Η μη ποσοτικοποίηση της αβεβαιότητας στα αιτιοκρατικά μοντέλα, σε συνδυασμό με τον κορεσμό που έχει επέλθει στην ανάπτυξη και βελτίωσή τους, έχει οδηγήσει στην όλο και αυξανόμενη έρευνα γύρω από την ανάπτυξη πιθανοτικών μοντέλων πρόβλεψης.

3.2.2 Πιθανοτικά μοντέλα

Σε αντίθεση με τα αιτιοκρατικά μοντέλα, τα πιθανοτικά μοντέλα πρόβλεψης έχουν ως στόχο την πλήρη πληροφόρηση του χρήστη σχετικά με τα πιθανά μελλοντικά συμβάντα. Η πιθανοτική πρόβλεψη περιλαμβάνει τρεις κατηγορίες: (α) την πρόβλεψη ποσοστμοριών



Σχήμα 3.3: Σημειακές προβλέψεις και διαστήματα πρόβλεψης παραγωγής για όλους τους χειρσαίους αιολικούς σταθμούς της δυτικής Δανίας στις 4 Απριλίου 2007 και ώρα 00:00 [38].

(β) την παραγωγή διαστημάτων πρόβλεψης και (γ) την πρόβλεψη της πυκνότητας πιθανότητας [38].

Η πρόβλεψη ενός ποσοστιμορίου $\hat{q}_{t+k|t}^{(a)}$ με ονομαστικό επίπεδο εμπιστοσύνης a , είναι μια εκτίμηση που γίνεται τη χρονική στιγμή t για τη χρονική στιγμή $t+k$, του ποσοστιμορίου $q_{t+k|t}^{(a)}$ για την τυχαία μεταβλητή Y_{t+k} , δεδομένου ενός μοντέλου g , των εκτιμώμενων παραμέτρων $\hat{\theta}_t$ και του συνόλου πληροφοριών Ω_t [38]:

$$P \left[Y_{t+k} \leq \hat{q}_{t+k|t}^{(a)} \mid g, \hat{\theta}_t, \Omega_t \right] = a \quad (3.1)$$

Η σχέση (3.1) εκφράζει ότι τη χρονική στιγμή t υπάρχει πιθανότητα a η παραγόμενη αιολική ισχύς τη χρονική στιγμή $t+k$ να είναι μικρότερη ή ίση της τιμής $\hat{q}_{t+k|t}^{(a)}$.

Παρόλο που η πρόβλεψη ποσοστιμορίου είναι μια πιθανοτική πρόβλεψη, δεν μπορεί να ποσοτικοποιήσει αισθητά την αβεβαιότητα της πρόβλεψης. Για αυτό τον λόγο έχουν αναπτυχθεί μοντέλα που παράγουν διαστήματα πρόβλεψης. Ένα διάστημα πρόβλεψης που παράγεται τη χρονική στιγμή t για τη χρονική στιγμή $t+k$ ορίζει ένα εύρος πιθανών τιμών για την τυχαία μεταβλητή Y_{t+k} για ένα συγκεκριμένο επίπεδο εμπιστοσύνης $(1-\beta)$, όπου $\beta \in [0,1]$ είναι ο ονομαστικός δείκτης κάλυψης [38]:

$$P \left[Y_{t+k} \in \hat{I}_{t+k|t}^{(\beta)} \mid g, \hat{\theta}_t, \Omega_t \right] = 1 - \beta \quad (3.2)$$

Το διάστημα πρόβλεψης $\hat{I}_{t+k|t}^{(\beta)}$ προσδιορίζεται από το κάτω όριο του ($\hat{q}_{t+k|t}^{(\underline{a})}$) και από το άνω όριο του ($\hat{q}_{t+k|t}^{(\bar{a})}$), τα οποία όρια είναι προβλέψεις ποσοστιμορίου με επίπεδα εμπιστοσύνης \underline{a} και \bar{a} , αντίστοιχα, για τα οποία ισχύει η σχέση (3.3):

$$\bar{\alpha} - \underline{\alpha} = 1 - \beta \quad (3.3)$$

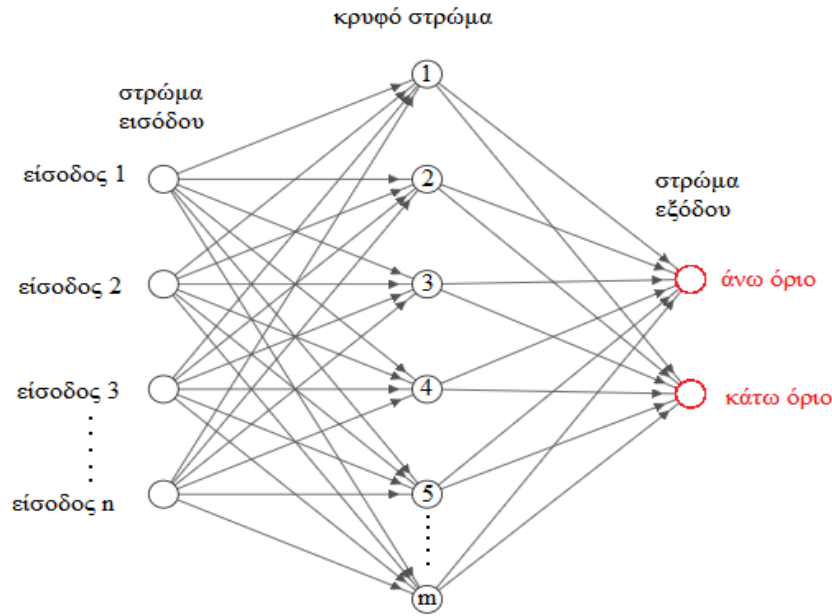
Ένα διάστημα πρόβλεψης εκφράζει ότι τη χρονική στιγμή t υπάρχει η πιθανότητα $1 - \beta$ η παραγόμενη αιολική ισχύς τη χρονική στιγμή $t+1$ να ανήκει στο διάστημα πρόβλεψης $\hat{I}_{t+k|t}^{(\beta)}$. Στο Σχήμα 3.3, μαζί με τις σημειακές προβλέψεις και τις πραγματικές τιμές, φαίνονται και τα αντίστοιχα διαστήματα πρόβλεψης με ονομαστικό επίπεδο εμπιστοσύνης 0.9 (90%).

Τα διαστήματα πρόβλεψης, αν και ποσοτικοποιούν πληρέστερα την αβεβαιότητα της πρόβλεψης σε σχέση με τα ποσοστιμώρια, δεν δίνουν το σύνολο των πληροφοριών σχετικά με την μελλοντική παραγωγή αιολικής ισχύος. Μια πλήρης πληροφόρηση μπορεί να δοθεί μόνο μέσω των προβλέψεων πυκνότητας πιθανότητας. Ως πρόβλεψη πυκνότητας πιθανότητας τη χρονική στιγμή t για τη χρονική στιγμή $t+1$ ορίζεται η πλήρης περιγραφή της συνάρτησης πυκνότητας πιθανότητας ή της κατανομής πυκνότητας πιθανότητας μιας τυχαίας μεταβλητής Y_{t+k} δεδομένων του μοντέλου g , των εκτιμώμενων παραμέτρων $\hat{\theta}_t$ και του συνόλου πληροφοριών Ω_t [38]. Ωστόσο, η ανάπτυξη μοντέλων για την πρόβλεψη της πυκνότητας πιθανότητας μιας τυχαίας μεταβλητής, όπως είναι η αιολική παραγωγή, είναι αρκετά πιο περίπλοκη σε σχέση με την παραγωγή διαστημάτων πρόβλεψης, απαιτεί μεγάλο όγκο δεδομένων υψηλής ποιότητας, ενώ σε ορισμένες περιπτώσεις είναι αδύνατον να επιτύχει υψηλά ποσοστά ακρίβειας.

3.3 ΜΕΘΟΔΟΣ ΕΚΤΙΜΗΣΗΣ ΚΑΤΩ ΚΑΙ ΑΝΩ ΟΡΙΟΥ

Οι περισσότερες μέθοδοι για την παραγωγή διαστημάτων πρόβλεψης με χρήση τεχνητών νευρωνικών δικτύων ακολουθούν δύο βασικά βήματα. Αρχικά μέσω των τεχνητών νευρωνικών δικτύων γίνεται σημειακή πρόβλεψη για το μελλοντικό συμβάν, και στη συνέχεια αξιοποιώντας τη σημειακή πρόβλεψη παράγονται μέσω στατιστικών μεθόδων τα κάτω και άνω όρια των διαστημάτων πρόβλεψης. Τα πιθανοτικά μοντέλα που ακολουθούν αυτές τις μεθόδους, αν και είναι καλύτερα σε σχέση με τα αιτιοκρατικά, έχουν ορισμένα μειονεκτήματα, όπως μεγάλη υπολογιστική πολυπλοκότητα, δυσκολία στην υλοποίηση, ενώ, επίσης, κάνουν ορισμένες ειδικές υποθέσεις για την κατανομή των δεδομένων [27].

Τα μειονεκτήματα αυτά ξεπερνιούνται με τη μέθοδο «εκτίμησης κάτω και άνω ορίου» (Lower Upper Bound Estimation – LUBE) που προτάθηκε από τους A. Khosravi, S. Nahavandi, D. Creighton και A. Atiya το 2011 [1]. Η μέθοδος LUBE είναι μία μέθοδος πιθανοτικής πρόβλεψης και συγκεκριμένα μια μέθοδος για παραγωγή διαστημάτων πρόβλεψης. Ένα μοντέλο πιθανοτικής πρόβλεψης που κάνει χρήση της μεθόδου LUBE, αποτελείται από ένα τεχνητό νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης με δύο νευρώνες στο στρώμα εξόδου που αντιπροσωπεύουν το κάτω και το άνω όριο του διαστήματος πρόβλεψης (Σχήμα 3.4). Για κάθε δεδομένο εισόδου δηλαδή, το τεχνητό νευρωνικό δίκτυο παράγει δύο εξόδους, η μικρότερη έξοδος αντιπροσωπεύει το κάτω όριο, ενώ η μεγαλύτερη έξοδος το άνω όριο. Σε αντίθεση με τις περισσότερες μεθόδους πιθανοτικής πρόβλεψης, η μέθοδος LUBE παράγει απευθείας τα διαστήματα πρόβλεψης μέσω του τεχνητού νευρωνικού δικτύου, χωρίς κάποια επιπρόσθετη στατιστική διαδικασία.



Σχήμα 3.4: Τεχνητό νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης με δύο νευρώνες στο στρώμα εξόδου: το κάτω και άνω όριο του διαστήματος πρόβλεψης

Οι αντικειμενικές συναρτήσεις που περιγράφονται στις εξισώσεις (2.8), (2.9) και (2.10) του Κεφαλαίου 2 αφορούν σημειακές προβλέψεις, επομένως δεν μπορούν να χρησιμοποιηθούν για την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου όπως αυτό του Σχήματος 3.4. Αντίθετα, είναι απαραίτητο να χρησιμοποιηθούν μετρητικά σφάλματος και αντικειμενικές συναρτήσεις που αφορούν διαστήματα πρόβλεψης. Ένα διάστημα πρόβλεψης έχει δύο χαρακτηριστικά: το εύρος του και την πιθανότητα κάλυψης. Δύο μετρητικά σφάλματος που ανταποκρίνονται σε αυτά τα χαρακτηριστικά είναι τα παρακάτω [27]:

(α) Πιθανότητα Κάλυψης Διαστήματος Πρόβλεψης (Prediction Interval Coverage Probability – PICP):

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i, \quad \begin{cases} c_i = 1 & \text{αν } y_i \in [L_i, U_i] \\ c_i = 0 & \text{αλλιώς} \end{cases} \quad (3.4)$$

όπου n το πλήθος των δεδομένων, y_i η πραγματική τιμή εξόδου, L_i και U_i το κάτω και άνω όριο του διαστήματος πρόβλεψης για το δεδομένο i . Το PICP εκφράζει το ποσοστό των δεδομένων που βρίσκονται εντός του αντίστοιχου διαστήματος πρόβλεψης. Για να είναι αξιόπιστα τα διαστήματα πρόβλεψης, θα πρέπει το PICP να είναι μεγαλύτερο ή ίσο του ονομαστικού επιπέδου εμπιστοσύνης. Στην ιδανική περίπτωση που όλα τα δεδομένα βρίσκονται εντός των διαστημάτων πρόβλεψής τους, το PICP ισούται με 1.

(β) Κανονικοποιημένο Μέσο Εύρος Διαστήματος Πρόβλεψης (Prediction Interval Normalized Average Width – PINAW):

$$PINAW = \frac{1}{nR} \sum_{i=1}^n (U_i - L_i) \quad (3.5)$$

όπου R το εύρος των δεδομένων εξόδου. Το PINAW εκφράζει το εύρος των διαστημάτων πρόβλεψης. Όσο αυξάνεται το εύρος, αυξάνεται και το PICP, ωστόσο μειώνεται η σπουδαιότητα της πληροφορίας που φέρουν τα διαστήματα πρόβλεψης.

Τα διαστήματα πρόβλεψης πρέπει να έχουν όσο το δυνατόν μικρότερο εύρος και ταυτόχρονα όσο το δυνατόν μεγαλύτερη πιθανότητα κάλυψης. Επομένως, η βελτιστοποίηση των διαστημάτων πρόβλεψης έχει δύο στόχους, την μεγιστοποίηση του PICP και την ελαχιστοποίηση του PINAW [27]. Ωστόσο, οι δύο αυτοί στόχοι είναι αντικρουόμενοι. Η βελτιστοποίηση των διαστημάτων πρόβλεψης της μεθόδου LUBE είναι επομένως από τη φύση της ένα πρόβλημα βελτιστοποίησης πολλαπλών αντικειμενικών συναρτήσεων και συγκεκριμένα δύο αντικειμενικών συναρτήσεων: της μεγιστοποίησης του PICP και της ελαχιστοποίησης του PINAW.

Καθώς η βελτιστοποίηση των διαστημάτων πρόβλεψης αποτελεί πρόβλημα πολυκριτηριακής βελτιστοποίησης, υπάρχουν δύο προσεγγίσεις για την εκπαίδευση του τεχνητού νευρωνικού δικτύου. Η πρώτη προσέγγιση είναι η διατήρηση του πολυκριτηριακού χαρακτήρα, κάνοντας χρήση κάποιου πολυκριτηριακού μεταερευτικού αλγορίθμου βελτιστοποίησης, όπως η πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων. Η δεύτερη προσέγγιση είναι η μετατροπή του προβλήματος σε πρόβλημα βελτιστοποίησης μίας αντικειμενικής συνάρτησης μέσω της ενοποίησης των δύο αντικειμενικών συναρτήσεων σε μία. Ωστόσο, η ενιαία αντικειμενική συνάρτηση που προκύπτει συνήθως είναι μη παραγωγίσιμη, μη γραμμική και πολύπλοκη, επομένως δεν είναι αποδοτικοί οι κλασικοί αλγόριθμοι βελτιστοποίησης που βασίζονται σε τεχνικές μείωσης της κλίσης, όπως ο αλγόριθμος της προς τα πίσω διάδοσης του σφάλματος (Ενότητα 2.4.1). Για τον λόγο αυτό, κατάλληλοι αλγόριθμοι για την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου της μεθόδου LUBE είναι οι μεταερευτικοί αλγόριθμοι (Ενότητα 2.4.2).

Η μέθοδος LUBE παρουσιάζει πολλά πλεονεκτήματα σε σχέση με άλλες μεθόδους πιθανοτικής πρόβλεψης. Τα περισσότερα πλεονεκτήματα πηγάζουν από το γεγονός ότι τα διαστήματα πρόβλεψης παράγονται απευθείας από το τεχνητό νευρωνικό δίκτυο. Η έλλειψη περαιτέρω μαθηματικών υπολογισμών και στατιστικών μεθόδων για την παραγωγή των διαστημάτων πρόβλεψης καθιστούν τη μέθοδο LUBE ταχύτερη και λιγότερο πολύπλοκη στην υλοποίησή της. Επίσης, το γεγονός ότι δεν γίνονται υπολογισμοί παραγώγων κατά τη διάρκεια της εκπαίδευσης του τεχνητού νευρωνικού δικτύου μειώνει την πιθανότητα η μέθοδος LUBE να παγιδευτεί σε κάποιο τοπικό βέλτιστο, καθιστώντας τη μέθοδο LUBE πιο αξιόπιστη σε σχέση με άλλες πιθανοτικές μεθόδους [27]. Η μέθοδος LUBE είναι μια μη παραμετρική μέθοδος, που σημαίνει ότι δεν κάνει καμία υπόθεση για την κατανομή των δεδομένων και των σφαλμάτων πρόβλεψης [1].

Λόγω των παραπάνω, η μέθοδος LUBE χρησιμοποιείται όλο και περισσότερο για την κατασκευή μοντέλων πιθανοτικής πρόβλεψης. Τα μοντέλα πρόβλεψης που βασίζονται στη μέθοδο LUBE έχουν εφαρμογές σε πολλά πεδία, μεταξύ των οποίων και η πρόβλεψη αιολικής παραγωγής. Ταυτόχρονα, γίνεται συστηματική ακαδημαϊκή έρευνα γύρω από τη βελτίωσή της, τον εμπλουτισμό της, τον συνδυασμό της με άλλες μεθόδους και εργαλεία, καθώς και την εφαρμογή της σε διαφορετικά προβλήματα πρόβλεψης.

ΑΝΑΛΥΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΑΙΟΛΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

4.1 ΔΕΔΟΜΕΝΑ ΑΙΟΛΙΚΗΣ ΠΑΡΑΓΩΓΗΣ

Η παραγωγή ηλεκτρικής ισχύος από αιολικούς σταθμούς παραγωγής (αιολική ισχύς) εξαρτάται από διάφορους παράγοντες, όπως την ταχύτητα και τη διεύθυνση του ανέμου, την υγρασία, την πυκνότητα του ανέμου, την ατμοσφαιρική πίεση, τη θερμοκρασία, το μήκος των πτερυγίων και το ύψος της ανεμογεννήτριας, τη διάταξη των ανεμογεννητριών μέσα σε έναν αιολικό σταθμό και τη φθορά [39]. Από αυτούς τους παράγοντες κάποιοι είναι σταθεροί στη διάρκεια του χρόνου (κυρίως οι παράγοντες που σχετίζονται με τα κατασκευαστικά χαρακτηριστικά των ανεμογεννητριών) ενώ άλλοι παράγοντες μεταβάλλονται στη διάρκεια του χρόνου, όπως οι καιρικοί παράγοντες. Αν και η αιολική ισχύς ενός αιολικού σταθμού εξαρτάται από όλους τους παραπάνω παράγοντες, η μεταβολή της αιολικής ισχύος εξαρτάται μόνο από τους παράγοντες που μεταβάλλονται με τον χρόνο, καθώς η συνεισφορά των υπόλοιπων παραγόντων είναι σταθερή. Επομένως, για την πρόβλεψη της αιολικής παραγωγής με χρήση τεχνητών νευρωνικών δικτύων γίνεται χρήση μόνο των παραγόντων που μεταβάλλονται με τον χρόνο.

Ως δεδομένα αιολικής παραγωγής ορίζονται οι μεταβλητοί παράγοντες που επηρεάζουν την αιολική παραγωγή, όπως η ταχύτητα και η διεύθυνση του ανέμου, η υγρασία, η πυκνότητα του ανέμου, η θερμοκρασία, η φθορά καθώς και η ίδια η αιολική παραγωγή, δηλαδή η ποσότητα ηλεκτρικής ισχύος που παράγεται. Στην περίπτωση της πρόβλεψης αιολικής παραγωγής με χρήση μηχανικής μάθησης, τα δεδομένα της αιολικής παραγωγής χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση των τεχνητών νευρωνικών δικτύων. Συγκεκριμένα, χρησιμοποιούνται ιστορικά αιολικά δεδομένα που αφορούν τον αιολικό σταθμό για τον οποίο θα γίνει η πρόβλεψη, ώστε αρχικά να εκπαιδευτεί και στη συνέχεια να αξιολογηθεί το τεχνητό νευρωνικό δίκτυο. Τα δεδομένα που επηρεάζουν την αιολική παραγωγή αποτελούν τα δεδομένα εισόδου, ενώ οι αντίστοιχες τιμές της παραγόμενης αιολικής ισχύος αποτελούν τα δεδομένα εξόδου. Αφού εκπαιδευτεί και αξιολογηθεί το τεχνητό νευρωνικό δίκτυο, τροφοδοτείται στην είσοδο με προβλέψεις των μελλοντικών τιμών των παραγόντων που χρησιμοποιήθηκαν ως δεδομένα εισόδου ώστε τελικά να παραχθούν οι προβλέψεις που αφορούν την αιολική ισχύ.

Τα δεδομένα αιολικής παραγωγής μπορούν να έχουν διαφορετικές χρονικές και χωρικές αναλύσεις. Ως χρονική ανάλυση ορίζεται η χρονική απόσταση μεταξύ δύο διαδοχικών παρατηρήσεων στον χρόνο. Σε ότι αφορά τα αιολικά δεδομένα, η χρονική ανάλυση συνήθως είναι από 5 λεπτά έως μερικές ώρες. Η χωρική ανάλυση αφορά τη σχετική χωρική απόσταση μεταξύ των παρατηρήσεων. Για παράδειγμα, για την πρόβλεψη της παραγωγής ενός αιολικού σταθμού μπορούν να χρησιμοποιηθούν αιολικά δεδομένα από

γειτονικούς αιολικούς σταθμούς, εάν παρατηρηθεί ότι υπάρχει συσχέτιση μεταξύ των τοποθεσιών τους.

Η πρόβλεψη των καιρικών παραγόντων έχει μεγάλη σημασία για την αξιοπιστία της πρόβλεψης της αιολικής παραγωγής. Μια μέθοδος πρόβλεψης των καιρικών παραγόντων αποτελεί η αριθμητική πρόβλεψη καιρού (Numerical Weather Prediction). Η αριθμητική πρόβλεψη καιρού μοντελοποιεί μαθηματικά την ατμόσφαιρα και παράγει προβλέψεις έχοντας ως βάση τις παρούσες καιρικές συνθήκες. Ένα μοντέλο αριθμητικής πρόβλεψης καιρού αρχικά συλλέγει ατμοσφαιρικές παρατηρήσεις. Οι παρατηρήσεις αυτές συνήθως εμπεριέχουν μεγάλα σφάλματα, λόγω της τυχαίας φύσης της ατμόσφαιρας, των σφαλμάτων των καταγραφικών οργάνων, αλλά και της αραιής τοποθέτησης των καταγραφικών οργάνων σε αραιοκατοικημένες και δύσβατες περιοχές. Για τη μείωση των σφαλμάτων, οι παρατηρήσεις περνούν από τη διαδικασία της αφομοίωσης (data assimilation). Η διαδικασία της αφομοίωσης έγκειται κατά κύριο λόγο στον συνδυασμό των παρατηρήσεων με προβλέψεις που έγιναν στο παρελθόν και αφορούν την τρέχουσα στιγμή. Στη συνέχεια, τα επεξεργασμένα δεδομένα χρησιμοποιούνται για την επίλυση επτά εξισώσεων που περιγράφουν μαθηματικά τις μεταβολές της ατμόσφαιρας: τον δεύτερο νόμο του Νεύτωνα (τρεις εξισώσεις, μία για κάθε συνιστώσα της ταχύτητας), την αρχή διατήρησης της μάζας, την αρχή διατήρησης της συγκέντρωσης των ατμοσφαιρικών συστατικών, τον πρώτο θερμοδυναμικό νόμο και την αρχή διατήρησης της ποσότητας του νερού. Η επίδραση ορισμένων ατμοσφαιρικών διεργασιών που δεν μπορούν να μοντελοποιηθούν λαμβάνεται υπόψη μέσω της ανάπτυξης σχημάτων παραμετροποίησης. Γνωρίζοντας τις αρχικές συνθήκες (επεξεργασμένες ατμοσφαιρικές παρατηρήσεις), τις εξισώσεις μοντελοποίησης της ατμόσφαιρας και τα σχήματα παραμετροποίησης, μπορούν να παραχθούν μελλοντικές καιρικές προβλέψεις, κατάλληλες για τροφοδότηση σε τεχνητά νευρωνικά δίκτυα με σκοπό την πρόβλεψη της αιολικής παραγωγής [40].

4.2 ΔΕΔΟΜΕΝΑ ΔΙΑΓΩΝΙΣΜΟΥ GLOBAL ENERGY FORECASTING COMPETITION 2014

Τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία για την εκπαίδευση και αξιολόγηση των τεχνητών νευρωνικών δικτύων είναι τα δεδομένα που δόθηκαν στους συμμετέχοντες του διαγωνισμού Global Energy Forecasting Competition 2014 (GEFCom 2014) [5]. Ο διαγωνισμός αυτός είχε ως αντικείμενο την ανάπτυξη μοντέλων πιθανοτικής πρόβλεψης για τέσσερις διαφορετικές εφαρμογές: πρόβλεψη αιολικής παραγωγής, πρόβλεψη ηλιακής παραγωγής, πρόβλεψη ηλεκτρικού φορτίου, και πρόβλεψη τιμής ηλεκτρικού ρεύματος.

Τα δεδομένα αφορούν δέκα αιολικούς σταθμούς παραγωγής στην Αυστραλία. Σε ότι αφορά την πρόβλεψη αιολικής παραγωγής, τα δεδομένα αποτελούνται από την πρόβλεψη του διανύσματος του ανέμου και από την παραγόμενη αιολική ισχύ. Το διάνυσμα του ανέμου αποτελείται από τις καρτεσιανές συνιστώσες u (οριζόντιος άξονας, Δύση–Ανατολή) και v (κατακόρυφος άξονας, Νότος–Βορράς). Οι προβλέψεις του διανύσματος του ανέμου δίνονται για δύο υψόμετρα: 10 και 100 μέτρα από το έδαφος. Η παραγόμενη αιολική ισχύς κάθε αιολικού σταθμού δίνεται κανονικοποιημένη ως προς την ονομαστική εγκατεστημένη ισχύ

του αιολικού σταθμού, στο διάστημα $[0, 1]$. Επομένως, για κάθε έναν από τους δέκα αιολικούς σταθμούς υπάρχουν πέντε κατηγορίες δεδομένων: οι προβλέψεις των συνιστωσών του διανύσματος του ανέμου u και v για υψόμετρα 10 και 100 μέτρα ($u_{10}, v_{10}, u_{100}, v_{100}$) και οι αντίστοιχες κανονικοποιημένες τιμές παραγόμενης αιολικής ισχύος (P_o). Οι προβλέψεις των συνιστωσών του διανύσματος του ανέμου αποτελούν τα δεδομένα εισόδου των τεχνητών νευρωνικών δικτύων, ενώ οι τιμές της παραγόμενης αιολικής ισχύος αποτελούν τα δεδομένα εξόδου.

Οι προβλέψεις του διανύσματος του ανέμου έχουν παραχθεί με χρονικό ορίζοντα πρόβλεψης 24 ωρών. Η χρονική ανάλυση των προβλέψεων του διανύσματος του ανέμου καθώς και των τιμών παραγόμενης αιολικής ισχύος είναι μία (1) ώρα, επομένως για κάθε ημέρα υπάρχουν 24 σύνολα δεδομένων. Τα δεδομένα αφορούν περίοδο 23 μηνών, από την 1^η Ιανουαρίου του 2012, έως και την 30^η Νοεμβρίου του 2013.

4.3 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

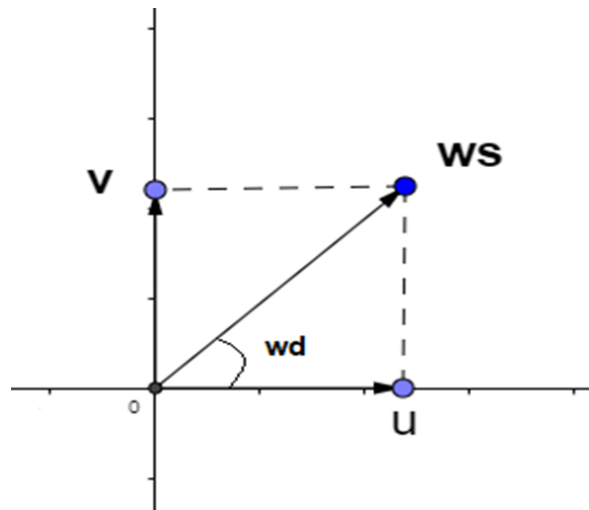
4.3.1 Μετατροπή δεδομένων

Ως δεδομένα εισόδου δίνονται οι προβλέψεις του διανύσματος του ανέμου, υπό τη μορφή των συνιστωσών u και v , στο καρτεσιανό σύστημα συντεταγμένων. Η παραγωγή αιολικής ισχύος εξαρτάται από την ταχύτητα του ανέμου, αλλά και από τη διεύθυνσή του. Ωστόσο, η επίδραση της διεύθυνσης του ανέμου είναι πολύ μικρότερη, καθώς οι ανεμογεννήτριες έχουν τη δυνατότητα να περιστρέφονται έτσι ώστε η ωφέλιμη επιφάνεια των πτερυγίων τους να είναι κάθετη στη διεύθυνση του ανέμου. Επομένως, όταν η ταχύτητα του ανέμου παραμένει σταθερή, η μεταβολή της παραγόμενης αιολικής ισχύος είναι αμελητέα για αλλαγές στη διεύθυνση του ανέμου [41].

Η άνιση αυτή επίδραση μεταξύ της ταχύτητας και της διεύθυνσης του ανέμου στην παραγόμενη αιολική ισχύ δεν μπορεί να αναδειχθεί όταν οι προβλέψεις ανέμου δίνονται υπό τη μορφή των συνιστωσών u και v . Η μετατροπή των συνιστωσών u και v σε ταχύτητα ανέμου (ws) και διεύθυνση ανέμου (wd) θα δώσει τη δυνατότητα στο τεχνητό νευρωνικό δίκτυο να αναγνωρίσει κατά την εκπαίδευσή του ότι η ταχύτητα ανέμου επηρεάζει κυρίαρχα την παραγωγή αιολικής ισχύος, επομένως θα δώσει μεγαλύτερο βάρος στα δεδομένα εισόδου που αφορούν την ταχύτητα ανέμου, αυξάνοντας έτσι την ακρίβεια των προβλέψεων.

Η μετατροπή αυτή επιτυγχάνεται με τη μεταφορά από το καρτεσιανό σύστημα συντεταγμένων στο πολικό. Το μέτρο της συνισταμένης των συνιστωσών u και v ισοδυναμεί με την ταχύτητα του ανέμου, ενώ η γωνία της συνισταμένης με τον οριζόντιο άξονα αντιστοιχεί στη διεύθυνση του ανέμου (Σχήμα 4.1). Η ταχύτητα του ανέμου υπολογίζεται από τη σχέση:

$$ws = \sqrt{u^2 + v^2} \quad (4.1)$$



Σχήμα 4.1: Η σχέση της ταχύτητας και της διεύθυνσης του ανέμου σε σχέση με τις συνιστώσες u και v .

Η διεύθυνση του ανέμου υπολογίζεται με τη βοήθεια της συνάρτησης «υπερεφαπτομένη 2» ($\arctan2$):

$$wd = \arctan2(v, u) = \begin{cases} \arctan\left(\frac{v}{u}\right) & \text{αν } u > 0 \\ \arctan\left(\frac{v}{u}\right) + \pi & \text{αν } u < 0 \text{ και } v \geq 0 \\ \arctan\left(\frac{v}{u}\right) + \pi & \text{αν } u < 0 \text{ και } v < 0 \\ \frac{\pi}{2} & \text{αν } u = 0 \text{ και } v > 0 \\ -\frac{\pi}{2} & \text{αν } u = 0 \text{ και } v < 0 \\ \text{απροσδιόριστο} & \text{αν } u = 0 \text{ και } v = 0 \end{cases} \quad (4.2)$$

όπου $\arctan\left(\frac{v}{u}\right)$ η υπερεφαπτομένη των συνιστωσών u και v .

4.3.2 Κανονικοποίηση δεδομένων

Η κανονικοποίηση των δεδομένων εισόδου και εξόδου αποτελεί μια πολύ συχνή διαδικασία προ-επεξεργασίας των δεδομένων. Ως κανονικοποίηση στον τομέα της μηχανικής μάθησης ορίζεται η διαδικασία επεξεργασίας των δεδομένων έτσι ώστε να αποκτήσουν ένα καινούριο εύρος τιμών, διατηρώντας παράλληλα τη μορφή της κατανομής τους. Στόχος της κανονικοποίησης είναι η εξασφάλιση ότι δεν θα υπάρχουν μεγάλες αποκλίσεις στην τάξη μεγέθους των δεδομένων. Κατά αυτόν τον τρόπο, μειώνεται η επιρροή των ακραίων τιμών και των δεδομένων με πολύ μεγαλύτερη τάξη μεγέθους στην εκπαίδευση του τεχνητού νευρωνικού δικτύου. Μέσω της κανονικοποίησης ρυθμίζεται επίσης η αναλογία μεταξύ θετικών και αρνητικών τιμών των δεδομένων. Για ορισμένα προβλήματα, τα τεχνητά νευρωνικά δίκτυα εκπαιδεύονται καλύτερα αν υπάρχουν και θετικές και αρνητικές τιμές. Ωστόσο πολλές φορές τα δεδομένα περιλαμβάνουν μόνο θετικές τιμές, δυσκολεύοντας έτσι

τη διαδικασία της εκπαίδευσης. Σε αυτές τις περιπτώσεις η κανονικοποίηση σε ένα εύρος τιμών με κέντρο το 0 εξασφαλίζει την ύπαρξη αρνητικών τιμών, επιταχύνοντας έτσι την εκπαίδευση του τεχνητού νευρωνικού δικτύου, πετυχαίνοντας ταυτόχρονα καλύτερα αποτελέσματα [42].

Στον τομέα της μηχανικής μάθησης επιλέγεται συνήθως η κανονικοποίηση στο διάστημα $[-1, 1]$ ή στο διάστημα $[0, 1]$. Στα δεδομένα του διαγωνισμού GEFCom 2014, τα δεδομένα εξόδου, δηλαδή οι τιμές της αιολικής ισχύος είναι ήδη κανονικοποιημένες στο διάστημα $[0, 1]$. Επομένως, πρέπει να γίνει κανονικοποίηση στα δεδομένα εισόδου.

Υπάρχουν διάφορες μέθοδοι κανονικοποίησης. Στην παρούσα διπλωματική εργασία επιλέγεται η κανονικοποίηση μεγίστου – ελαχίστου. Η κανονικοποίηση μεγίστου – ελαχίστου γίνεται με τη βοήθεια της σχέσης [43]:

$$x'_i = (b - a) \frac{x_i - x_{min}}{x_{max} - x_{min}} + a \quad (4.3)$$

όπου x'_i η κανονικοποιημένη τιμή του δεδομένου i , x_i η τιμή του δεδομένου i πριν την κανονικοποίηση, x_{min} και x_{max} η ελάχιστη και η μέγιστη τιμή της κατηγορίας δεδομένων στην οποία ανήκει το δεδομένο i και $[a, b]$ το διάστημα στο οποίο επιλέγεται να γίνει η κανονικοποίηση. Τα δεδομένα εισόδου του διαγωνισμού GEFCom 2014 περιλαμβάνουν και αρνητικές τιμές επομένως επιλέγεται η κανονικοποίηση να γίνει στο διάστημα $[-1, 1]$. Η σχέση (4.3) για $a = -1$ και $b = 1$ διαμορφώνεται ως εξής:

$$x'_i = 2 \cdot \frac{x_i - x_{min}}{x_{max} - x_{min}} - 1 \quad (4.4)$$

ΜΟΝΤΕΛΑ ΠΙΘΑΝΟΤΙΚΗΣ ΠΡΟΒΛΕΨΗΣ ΑΙΟΛΙΚΗΣ ΙΣΧΥΟΣ

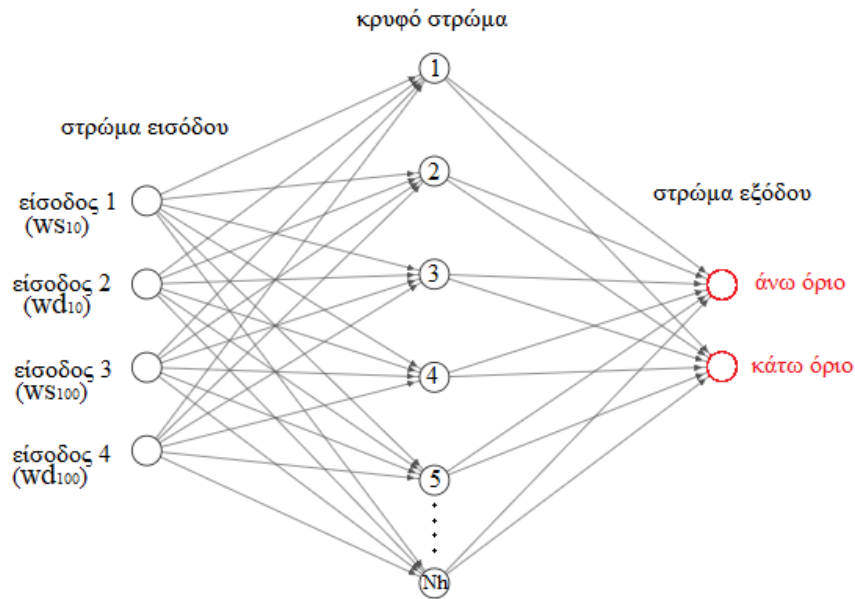
5.1 ΕΚΤΙΜΗΣΗ ΚΑΤΩ ΚΑΙ ΑΝΩ ΟΡΙΟΥ ΜΕ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΣΜΗΝΟΥΣ ΣΩΜΑΤΙΔΙΩΝ ΤΟΥ ΚΡΙΤΗΡΙΟΥ ΤΟΥ ΕΥΡΟΥΣ ΚΑΛΥΨΗΣ

Το μοντέλο εκτίμησης κάτω και άνω ορίου με βελτιστοποίηση σμήνους σωματιδίων του κριτηρίου του εύρους κάλυψης (Lower Upper Bound Estimation with Particle Swarm Optimization and Coverage Width Criterion, LUBE–PSO–CWC) [27] αποτελεί ένα μοντέλο πιθανοτικής πρόβλεψης που βασίζεται στη μέθοδο LUBE. Ένα μοντέλο που βασίζεται στη μέθοδο LUBE αποτελείται από ένα τεχνητό νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης με δύο νευρώνες στο στρώμα εξόδου, που αντιστοιχούν στο κάτω και άνω όριο του παραγόμενου διαστήματος πρόβλεψης [1]. Τα δεδομένα εισόδου, που προέρχονται από τον διαγωνισμό GEFCom 2014 [5], χωρίζονται σε τέσσερις κατηγορίες, την ταχύτητα και τη διεύθυνση του ανέμου στα δέκα και στα εκατό μέτρα ($ws_{10}, wd_{10}, ws_{100}, wd_{100}$). Επομένως το στρώμα εισόδου του τεχνητού νευρωνικού δικτύου αποτελείται από τέσσερις νευρώνες, έναν για κάθε κατηγορία δεδομένων. Το τεχνητό νευρωνικό δίκτυο επιλέγεται να έχει ένα κρυφό στρώμα, του οποίου το πλήθος νευρώνων N_h καθορίζεται μέσω μιας τεχνικής βελτιστοποίησης παραμέτρων που παρουσιάζεται στην Ενότητα 5.3. Το τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE–PSO–CWC απεικονίζεται στο Σχήμα 5.1. Ο καθορισμός των συναρτήσεων ενεργοποίησης του κρυφού στρώματος και του στρώματος εξόδου γίνεται στην Ενότητα 5.3.

5.1.1 Κριτήριο εύρους κάλυψης

Η βελτιστοποίηση των διαστημάτων πρόβλεψης της μεθόδου LUBE αποτελεί μια διαδικασία που από τη φύση της είναι πολυκριτηριακή. Το μοντέλο LUBE–PSO–CWC ακολουθεί την προσέγγιση της μετατροπής του προβλήματος πολυκριτηριακής βελτιστοποίησης σε πρόβλημα βελτιστοποίησης μίας αντικειμενικής συνάρτησης μέσω της ενοποίησης των δύο αντικειμενικών συναρτήσεων, της μεγιστοποίησης του PICP (σχέση 3.4) και της ελαχιστοποίησης του PINAW (σχέση 3.5). Η ενιαία αντικειμενική συνάρτηση που προκύπτει ονομάζεται κριτήριο εύρους κάλυψης (Coverage Width based Criterion – CWC) και έγκειται στην ελαχιστοποίηση της τιμής του CWC [1]:

$$CWC = PINAW(1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \quad (5.1)$$



Σχήμα 5.1: Το τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE – PSO – CWC

όπου $\gamma(PICP) = 1$ και μ το ονομαστικό επίπεδο εμπιστοσύνης. Όταν το PICP είναι μεγαλύτερο ή ίσο του ονομαστικού επιπέδου εμπιστοσύνης, ο εκθέτης $-\eta(PICP - \mu)$ γίνεται αρνητικός, επομένως ο όρος $e^{-\eta(PICP - \mu)}$ αποκτάει αμελητέα τιμή. Σε αυτήν την περίπτωση αποκτάει μεγαλύτερο βάρος το PINAW, δηλαδή το εύρος του διαστήματος πρόβλεψης, στον υπολογισμό του CWC. Όταν το PICP είναι μικρότερο του ονομαστικού επιπέδου εμπιστοσύνης, ο εκθέτης $-\eta(PICP - \mu)$ γίνεται θετικός, επομένως ο όρος $e^{-\eta(PICP - \mu)}$ δεν είναι αμελητέος. Ο ρόλος του η είναι να μεγεθύνει τον εκθετικό όρο, έτσι ώστε ακόμα και αν το PICP είναι ελάχιστα μικρότερο του ονομαστικού επιπέδου εμπιστοσύνης, η τιμή του CWC να παρουσιάζει απότομη εκθετική αύξηση. Για αυτόν τον λόγο το η παίρνει μεγάλες τιμές, συνήθως μεταξύ του 50 και του 100 [1].

Βασικός στόχος στην παραγωγή των διαστημάτων πρόβλεψης είναι το PICP να είναι μεγαλύτερο ή ίσο του ονομαστικού επιπέδου εμπιστοσύνης. Εάν ικανοποιείται αυτή η συνθήκη, η τιμή του CWC εξαρτάται κυρίως από το PINAW, επομένως απαραίτητη προϋπόθεση για να μειωθεί το CWC είναι να μειωθεί το μέσο εύρος των διαστημάτων πρόβλεψης εξασφαλίζοντας ταυτόχρονα ότι $PICP \geq \mu$. Στην περίπτωση που το PICP είναι μικρότερο του ονομαστικού επιπέδου εμπιστοσύνης, η τιμή του CWC παρουσιάζει απότομη εκθετική αύξηση.

Αν και το PINAW εκφράζει ικανοποιητικά το μέσο εύρος των διαστημάτων πρόβλεψης, υπάρχουν μετρητικά εύρους που αποδίδουν καλύτερα κατά την εκπαίδευση του τεχνητού νευρωνικού δικτύου. Ένα τέτοιο μετρητικό είναι η κανονικοποιημένη ρίζα της μέσης τιμής του τετραγώνου του εύρους των διαστημάτων πρόβλεψης (Prediction Interval Normalized Root mean square Width – PINRW) που περιγράφεται από την εξίσωση (5.2) [27]:

$$PINRW = \frac{1}{R} \sqrt{\frac{\sum_{i=1}^n (U_i - L_i)^2}{n}} \quad (5.2)$$

Το PINRW δεν είναι ευθέως ανάλογο του μέσου εύρους των διαστημάτων πρόβλεψης, όπως είναι το PINAW, αλλά έχει την ιδιότητα να «μεγεθύνει» τα μεγαλύτερα διαστήματα πρόβλεψης κατά τον υπολογισμό του εύρους τους. Για αυτόν τον λόγο επιλέγεται το PINRW έναντι του PINAW για την εκπαίδευση του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE-PSO-CWC. Συνεπώς, η εξίσωση (5.1) διαμορφώνεται ως εξής [27]:

$$CWC = PINRW(1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \quad (5.3)$$

Κατά την εκπαίδευση του τεχνητού νευρωνικού δικτύου, υπάρχει το ενδεχόμενο όλα τα παραγόμενα διαστήματα πρόβλεψης, κατά τη διάρκεια μιας επανάληψης, να έχουν μηδενικό εύρος, δηλαδή $PINAW = PINRW = 0$. Σύμφωνα με τις εξισώσεις (5.1) και (5.3) αυτό θα οδηγήσει σε μηδενικό CWC. Καθώς η αντικειμενική συνάρτηση της εκπαίδευσης του τεχνητού νευρωνικού δικτύου έγκειται στην ελαχιστοποίηση του CWC, το τεχνητό νευρωνικό δίκτυο θα θεωρήσει το μηδενικό CWC ως βέλτιστο, ενώ στην πραγματικότητα το PICP θα είναι πολύ μικρότερο του ονομαστικού επιπέδου εμπιστοσύνης. Για την αποφυγή αυτού του ενδεχομένου, το γινόμενο στις εξισώσεις (5.1) και (5.3) μετατρέπεται σε άθροισμα. Επομένως, η τελική αντικειμενική συνάρτηση που επιλέγεται για την εκπαίδευση του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE-PSO-CWC έγκειται στην ελαχιστοποίηση της τιμής του CWC, η οποία υπολογίζεται σύμφωνα με την εξίσωση (5.4) [9]:

$$CWC = PINRW + \gamma(PICP)e^{-\eta(PICP-\mu)} \quad (5.4)$$

Ως ονομαστικό επίπεδο εμπιστοσύνης επιλέγεται $\mu = 0.9$. Το PICP των διαστημάτων πρόβλεψης θα πρέπει να είναι μεγαλύτερο ή ίσο του 0.9, δηλαδή τουλάχιστον το 90% των πραγματικών τιμών της αιολικής ισχύος θα πρέπει να βρίσκεται εντός των αντίστοιχων διαστημάτων πρόβλεψής τους. Σε ότι αφορά το η , επιλέγεται $\eta = 80$.

5.1.2 Αρχικοποίηση βαρών και πολώσεων

Η αρχικοποίηση των βαρών και των πολώσεων ενός τεχνητού νευρωνικού δικτύου επηρεάζει σε μεγάλο βαθμό την ταχύτητα σύγκλισής του κατά τη διάρκεια της εκπαίδευσής του. Οι βέλτιστες αρχικές τιμές των βαρών και των πολώσεων είναι αδύνατον να καθοριστούν καθώς εξαρτώνται από το σύνολο των δεδομένων εκπαίδευσης [44]. Για αυτόν τον λόγο έχουν αναπτυχθεί διάφορες τεχνικές αρχικοποίησης. Η επιλογή της κατάλληλης τεχνικής αρχικοποίησης των βαρών και των πολώσεων γίνεται έχοντας ως κριτήρια το είδος των δεδομένων, τη φύση του προβλήματος και την αρχιτεκτονική του τεχνητού νευρωνικού δικτύου.

Η πιο συχνή τεχνική αρχικοποίησης είναι η τυχαία αρχικοποίηση λόγω της απλότητάς της, αλλά και της ικανότητάς της να παράγει λύσεις με μεγάλη ποικιλομορφία, καθώς τα βάρη και οι πολώσεις μπορούν να συγκλίνουν σε διαφορετικές λύσεις ανάλογα με τις αρχικές τους τιμές [44]. Για αυτόν τον λόγο επιλέγεται η τυχαία αρχικοποίηση των βαρών και των πολώσεων του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE-PSO-CWC.

Συγκεκριμένα επιλέγεται τυχαία και ομοιόμορφη αρχικοποίηση, συμμετρική γύρω από το μηδέν, με μοναδιαίο εύρος διαστήματος. Όλα τα βάρη και οι πολώσεις αρχικοποιούνται, λαμβάνοντας τυχαίες τιμές στο διάστημα $[-0.5, 0.5]$.

Ωστόσο, λόγω της μεγάλης πολυπλοκότητας που έχει η πρόβλεψη της αιολικής παραγωγής, γίνεται χρήση και της τεχνικής αρχικοποίησης Nguyen–Widrow [4], ώστε να εξασφαλίζεται η σύγκλιση του μοντέλου σε ένα βέλτιστο, εντός μικρού χρονικού διαστήματος. Η αρχικοποίηση Nguyen–Widrow εφαρμόζεται για τα βάρη που ενώνουν τους νευρώνες του στρώματος εισόδου με τους νευρώνες του κρυφού στρώματος ($w_{in \rightarrow hidden}$) καθώς και για τις πολώσεις των νευρώνων του κρυφού στρώματος (b_{hidden}). Κάθε βάρος $w_{in \rightarrow hidden}$, με την εφαρμογή της αρχικοποίησης Nguyen–Widrow, αρχικοποιείται ως εξής:

$$\hat{w}_{in \rightarrow hidden}^{ij} = \beta \frac{w_{in \rightarrow hidden}^{ij}}{|w_{hidden}^j|} \quad (5.5)$$

όπου $w_{in \rightarrow hidden}^{ij}$ η αρχική τυχαία τιμή στο διάστημα $[-0.5, 0.5]$ του βάρους που συνδέει τον νευρώνα i του στρώματος εισόδου με τον νευρώνα j του κρυφού στρώματος, $|w_{hidden}^j|$ η νόρμα του νευρώνα j του κρυφού στρώματος και β μια παράμετρος που υπολογίζεται σύμφωνα με τη σχέση (5.6):

$$\beta = 0.7 \sqrt{\frac{N_{in}}{N_h}} \quad (5.6)$$

όπου N_{in} το πλήθος των νευρώνων του στρώματος εισόδου και N_h το πλήθος των νευρώνων του κρυφού στρώματος. Η νόρμα του νευρώνα j του κρυφού στρώματος υπολογίζεται σύμφωνα με τη σχέση (5.7):

$$|w_{hidden}^j| = \sqrt{\sum_{i=1}^{N_{in}} (w_{ij}^2) + b_j^2} \quad (5.7)$$

όπου w_{ij} το βάρος που συνδέει τον νευρώνα i του στρώματος εισόδου με τον νευρώνα j του κρυφού στρώματος και b_j η πόλωση του νευρώνα j του κρυφού στρώματος. Στη συνέχεια οι πολώσεις του κρυφού στρώματος λαμβάνουν τυχαίες τιμές στο διάστημα $[-\beta, \beta]$:

$$\hat{b}_{hidden}^j = random([- \beta, \beta]) \quad (5.8)$$

Με την τεχνική αρχικοποίησης Nguyen–Widrow ανατίθεται σε κάθε νευρώνα του κρυφού στρώματος μια περιοχή μέσα σε ένα εύρος τιμών που καθορίζεται από την παράμετρο β . Το εύρος τιμών αυτό χωρίζεται σε N_h ίσα τμήματα, ένα για κάθε νευρώνα του κρυφού στρώματος. Οι αρχικές θέσεις των νευρώνων μέσα στην περιοχή τους, καθορίζονται από τις τιμές των πολώσεών τους, οι οποίες είναι τυχαίες. Κατά αυτόν τον τρόπο, παρόλο που διατηρείται μια μορφή τυχειότητας, οι αρχικές τιμές των βαρών μεταξύ στρώματος εισόδου

και κρυφού στρώματος είναι ομοιόμορφα κατανεμημένες, αυξάνοντας την ταχύτητα σύγκλισης του τεχνητού νευρωνικού δικτύου.

Η τεχνική αρχικοποίηση Nguyen–Widrow είναι κατάλληλη για δεδομένα εισόδου τα οποία είναι συμμετρικά ως προς το μηδέν [27], όπως τα δεδομένα εισόδου που χρησιμοποιούνται στην παρούσα διπλωματική εργασία, τα οποία είναι κανονικοποιημένα στο διάστημα $[-1, 1]$. Επίσης, η τεχνική αρχικοποίηση Nguyen–Widrow επιλέγεται για το μοντέλο LUBE–PSO–CWC λόγω της αποτελεσματικότητάς της και της ευρείας εφαρμογής της σε προβλήματα διαφορετικής φύσης, ανεξαρτήτως πολυπλοκότητας.

5.1.3 Αλγόριθμος Βελτιστοποίησης

Ως αλγόριθμος βελτιστοποίησης για την εκπαίδευση του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE–PSO–CWC επιλέγεται ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων [2], καθώς τα πλεονεκτήματά του τον καθιστούν κατάλληλο για ένα πρόβλημα υψηλής πολυπλοκότητας, όπως αυτό της πιθανοτικής πρόβλεψης της αιολικής παραγωγής.

Κάθε σωματίδιο του σμήνους αντιπροσωπεύει μία διαφορετική λύση, επομένως σε κάθε σωματίδιο αντιστοιχεί ένα τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE–PSO–CWC. Η θέση και η ταχύτητα κάθε σωματιδίου αποτελούνται από τόσες διαστάσεις όσες είναι και οι διαστάσεις του μοντέλου. Οι διαστάσεις του μοντέλου είναι ίσες με το πλήθος των βαρών και των πολώσεων του τεχνητού νευρωνικού δικτύου, δηλαδή:

$$n_{dim} = N_{in} \cdot N_h + N_h + 2 \cdot N_h + 2 \quad (5.9)$$

Η σχέση (5.9) ισχύει για τεχνητό νευρωνικό δίκτυο με δύο νευρώνες στο στρώμα εξόδου, όπως αυτό του μοντέλου LUBE–PSO–CWC. Η θέση κάθε σωματιδίου ταυτίζεται με τα βάρη και τις πολώσεις του τεχνητού νευρωνικού δικτύου που αντιπροσωπεύει, επομένως οι θέσεις των σωματιδίων είναι ήδη αρχικοποιημένες. Η ταχύτητα κάθε σωματιδίου αρχικοποιείται παρόμοια με τα βάρη και τις πολώσεις, δηλαδή τυχαία με ομοιόμορφη κατανομή στο διάστημα $[-0.5, 0.5]$. Ως αρχικά καλύτερη θέση κάθε σωματιδίου θεωρείται η αρχική του θέση.

Για κάθε σωματίδιο παράγονται τα διαστήματα πρόβλεψης και γίνεται η αξιολόγησή τους έχοντας ως κριτήριο την ελαχιστοποίηση του CWC. Στη συνέχεια προσδιορίζεται η συνολικά καλύτερη θέση μέσα στο σμήνος. Η ταχύτητα και η θέση κάθε σωματιδίου ανανεώνονται σύμφωνα με τις εξισώσεις (2.15) και (2.16). Οι εξισώσεις αυτές εφαρμόζονται για κάθε διάσταση του μοντέλου. Εάν κάποια διάσταση της ταχύτητας (ή της θέσης) ενός σωματιδίου βγει εκτός του καθορισμένου εύρους τιμών της $[-v_{max}, v_{max}]$ (ή $[-x_{max}, x_{max}]$), περιορίζεται εντός αυτού. Συγκεκριμένα, αν για κάποια διάσταση i της ταχύτητας ενός σωματιδίου προκύψει $v_i > v_{max}$, η διάσταση της ταχύτητας αυτής λαμβάνει την τιμή v_{max} . Αν για κάποια διάσταση i της ταχύτητας ενός σωματιδίου προκύψει $v_i < -v_{max}$, η διάσταση της ταχύτητας αυτής λαμβάνει την τιμή $-v_{max}$. Όμοια προσέγγιση ακολουθείται για κάθε διάσταση της θέσης κάθε σωματιδίου. Αφού ανανεωθούν οι ταχύτητες και οι θέσεις, παράγονται τα νέα διαστήματα πρόβλεψης, γίνεται εκ νέου η αξιολόγηση της θέσης κάθε σωματιδίου με βάση την αντικειμενική συνάρτηση ελαχιστοποίησης του CWC

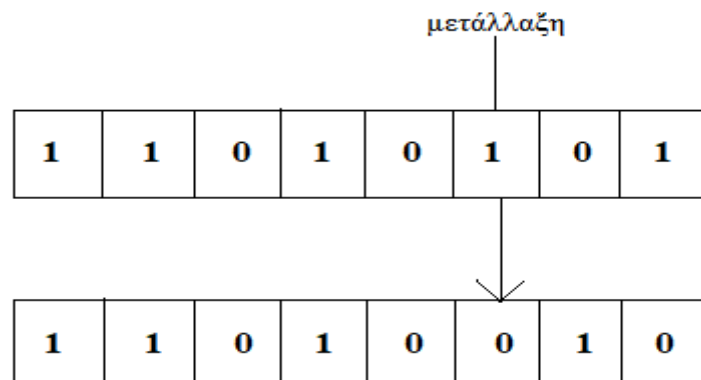
και, αν είναι απαραίτητο, ανανεώνεται η καλύτερη θέση κάθε σωματιδίου και η συνολικά καλύτερη θέση μέσα στο σμήνος. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να συμπληρωθεί ο καθορισμένος αριθμός επαναλήψεων του αλγορίθμου.

Το πλήθος των επαναλήψεων πρέπει να είναι αρκετά μεγάλο, ώστε τα σωματίδια να έχουν χρόνο να συγκλίνουν στο ολικό βέλτιστο. Ωστόσο, από ένα σημείο και έπειτα οι μεταβολές στα τελικά βάρη και στις τελικές πολώσεις είναι αμελητέες, άρα η αύξηση του αριθμού των επαναλήψεων πέρα από αυτό το σημείο αυξάνει τον χρόνο επίλυσης του προβλήματος χωρίς να βελτιώνει ιδιαίτερα τη λύση του. Ως οριακή τιμή για τις θέσεις των σωματιδίων επιλέγεται $x_{max} = 4$, επομένως κάθε διάσταση των θέσεων των σωματιδίων περιορίζεται εντός του διαστήματος $[-4, 4]$. Η μέγιστη ταχύτητα των σωματιδίων επιλέγεται έτσι ώστε οι μεταβολές των θέσεών τους να μην είναι αμελητέες, αλλά ταυτόχρονα ούτε πολύ μεγάλες με αποτέλεσμα την εξώθηση των σωματιδίων εκτός των ορίων τους. Ως μέγιστη ταχύτητα επιλέγεται $v_{max} = 1$, επομένως κάθε διάσταση των ταχυτήτων των σωματιδίων περιορίζεται εντός του διαστήματος $[-1, 1]$. Ο καθορισμός των παραμέτρων c_1 , c_2 , w των εξισώσεων (2.15) και (2.16) καθώς και του πλήθους των σωματιδίων και του πλήθους των επαναλήψεων για το μοντέλο LUBE-PSO-CWC γίνεται στην Ενότητα 5.3.1.

5.1.4 Τελεστής μετάλλαξης

Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων έχει μεγάλη ταχύτητα σύγκλισης. Λόγω της μεγάλης ταχύτητας σύγκλισης, είναι πιθανόν τα σωματίδια να συγκλίνουν πρόωρα σε κάποιο μη ολικό βέλτιστο. Για την αποφυγή της πρόωρης σύγκλισης σε μη ολικό βέλτιστο, γίνεται εφαρμογή του τελεστή μετάλλαξης (mutation operator) [27].

Ο τελεστής μετάλλαξης, ως μέθοδος, προέρχεται από τον γενετικό αλγόριθμο. Στον γενετικό αλγόριθμο κάθε γονίδιο απεικονίζεται συνήθως ως μια αλληλουχία δυαδικών ψηφίων. Η πιθανότητα να μεταλλαχθεί ένα γονίδιο καθορίζεται από την πιθανότητα μετάλλαξης m . Η μετάλλαξη έγκειται στη δυαδική αντικατάσταση ενός τυχαίου ψηφίου του γονιδίου, όπως φαίνεται στο Σχήμα 5.2. Σκοπός του τελεστή μετάλλαξης στον γενετικό αλγόριθμο είναι να αποκαταστήσει χαμένο ή μη εξερευνημένο γενετικό υλικό στον πληθυσμό, ώστε να αποτραπεί η πρόωρη σύγκλιση σε κάποιο μη ολικό βέλτιστο [45].



Σχήμα 5.2: Παράδειγμα μετάλλαξης σε μια ακολουθία δυαδικών ψηφίων

Ο τελεστής μετάλλαξης μπορεί να επεκταθεί και σε αλγορίθμους των οποίων οι πληθυσμοί απεικονίζονται με πραγματικούς αριθμούς και όχι με αλληλουχίες δυαδικών ψηφίων. Όπως συμβαίνει και στις ακολουθίες δυαδικών ψηφίων, έτσι και στους πραγματικούς αριθμούς, η πιθανότητα να μεταλλαχθεί ένα σωματίδιο του πληθυσμού καθορίζεται από την πιθανότητα μετάλλαξης. Η διαφορά έγκειται στην ίδια τη μετάλλαξη. Συνήθως τα σωματίδια μεταλλάσσονται ακολουθώντας μια προκαθορισμένη εξίσωση, της οποίας τουλάχιστον ένας όρος παράγεται τυχαία, ακολουθώντας μια τυποποιημένη κατανομή.

Η πιθανότητα μετάλλαξης καθορίζεται από τον δείκτη μετάλλαξης. Ως δείκτης μετάλλαξης ορίζεται ο αριθμός των μεταλλάξεων που γίνεται στο σύνολο του πληθυσμού σε μία επανάληψη. Ο δείκτης μετάλλαξης καθορίζεται από δύο χαρακτηριστικά: την αρχική του τιμή και τη χρονική μεταβολή του. Συνήθως η αρχική τιμή του δείκτη μετάλλαξης επιλέγεται ίση με τη μονάδα [46]. Όσον αφορά τη χρονική μεταβολή του, ο δείκτης μετάλλαξης συνήθως μειώνεται όσο προχωρούν οι επαναλήψεις. Αυτό συμβαίνει διότι στις αρχικές επαναλήψεις είναι επιθυμητές οι μεταλλάξεις καθώς αυξάνεται η τυχαιότητα και άρα ο χώρος εξερεύνησης, μειώνοντας την πιθανότητα να συγκλίνει ο αλγόριθμος σε κάποιο μη ολικό βέλτιστο. Ωστόσο, όσο προχωρούν οι επαναλήψεις, είναι επιθυμητή η μείωση της τυχαιότητας, ώστε να μπορέσει να συγκλίνει ο αλγόριθμος. Σε κάποιες περιπτώσεις προτιμάται ο δείκτης μετάλλαξης να παραμένει σταθερός σε όλη τη διάρκεια των επαναλήψεων. Η επιλογή της αρχικής τιμής και της χρονικής μεταβολής του δείκτη μετάλλαξης εξαρτάται από τον αλγόριθμο βελτιστοποίησης και τη φύση του προβλήματος.

Ο τελεστής μετάλλαξης προσαρμόστηκε με επιτυχία στη βελτιστοποίηση σμήνους σωματιδίων από τους N. Higashi και H. Iba [47]. Η μετάλλαξη στη βελτιστοποίηση σμήνους σωματιδίων αφορά τη θέση των σωματιδίων. Σε κάθε επανάληψη, υπάρχει μια πιθανότητα με βάση την οποία κάθε διάσταση της θέσης κάθε σωματιδίου υπόκειται σε μετάλλαξη. Η μετάλλαξη που επιλέχθηκε από τους N. Higashi και H. Iba είναι γκαουσιανή, με την έννοια ότι ο τυχαίος όρος στην εξίσωση μετάλλαξης της θέσης των σωματιδίων προκύπτει ακολουθώντας κανονική κατανομή. Συγκεκριμένα, η μετάλλαξη της θέσης των σωματιδίων ακολουθεί την εξίσωση (5.10):

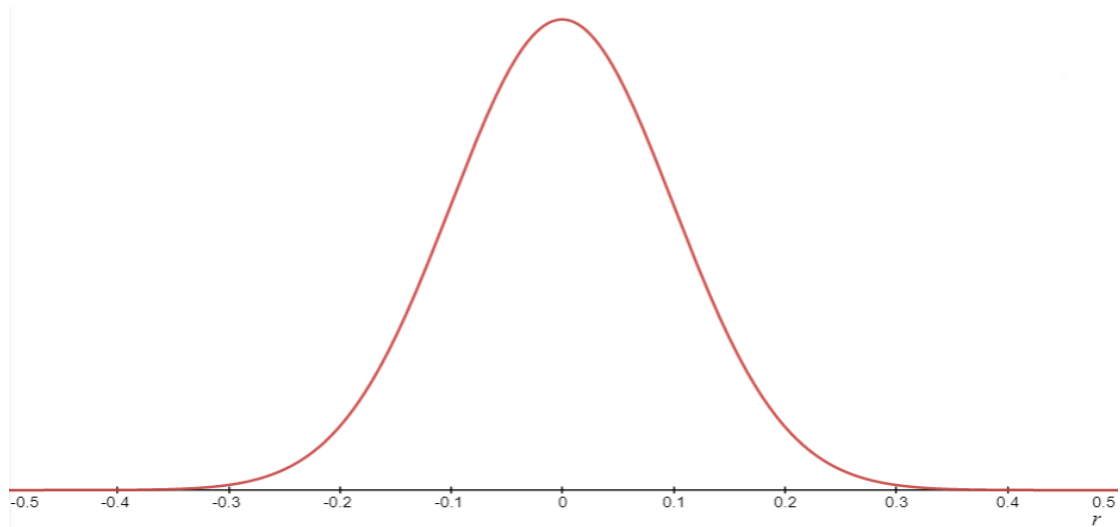
$$x' = x \left(1 + r \left(gaussian(\mu = 0, \sigma = 0.1) \right) \right) \quad (5.10)$$

όπου $r \left(gaussian(\mu = 0, \sigma = 0.1) \right)$ τυχαίος αριθμός που ακολουθεί κανονική κατανομή με μέση τιμή $\mu = 0$ και τυπική απόκλιση $\sigma = 0.1$. Η αρχική τιμή του δείκτη μετάλλαξης είναι τέτοια ώστε κατά την πρώτη επανάληψη να συμβεί μία μετάλλαξη στο σύνολο των σωματιδίων. Ο δείκτης μετάλλαξης μειώνεται γραμμικά έτσι ώστε στην τελευταία επανάληψη να έχει τιμή ίση με μηδέν.

Στην παρούσα διπλωματική εργασία ο τελεστής μετάλλαξης εκφράζεται από την εξίσωση (5.11) [46]:

$$x_i' = x_i + r \left(gaussian(\mu = 0, \sigma = 0.1) \right) \quad (5.11)$$

Η μετάλλαξη προκύπτει από το άθροισμα της τιμής της διάστασης i της θέσης ενός σωματιδίου με μια τυχαία τιμή r που ακολουθεί κανονική κατανομή με μέση τιμή $\mu = 0$ και τυπική απόκλιση $\sigma = 0.1$ (Σχήμα 5.3). Επομένως, η μετάλλαξη είναι γκαουσιανή και



Σχήμα 5.3: Η κανονική κατανομή με μέση τιμή $\mu=0$ και τυπική απόκλιση $\sigma=0.1$ που ακολουθεί η τυχαία τιμή r της εξίσωσης μετάλλαξης (5.11)

εφαρμόζεται σε κάθε διάσταση των θέσεων των σωματιδίων σύμφωνα με την πιθανότητα μετάλλαξης m .

Το πρόβλημα της πιθανοτικής πρόβλεψης αιολικής παραγωγής είναι ιδιαίτερα πολύπλοκο, με πολλά τοπικά βέλτιστα. Το γεγονός αυτό σε συνδυασμό με την υψηλή ταχύτητα σύγκλισης του αλγορίθμου βελτιστοποίησης σμήνους σωματιδίων καθιστά αρκετά πιθανό το ενδεχόμενο της πρόωρης σύγκλισης των σωματιδίων σε κάποιο μη ολικό βέλτιστο. Για την αποφυγή του φαινομένου αυτού, επιλέγεται μια πιθανότητα μετάλλαξης αρκετά μεγαλύτερη από αυτή που συναντάται συνήθως στη βιβλιογραφία. Συγκεκριμένα, επιλέγεται πιθανότητα μετάλλαξης τέτοια ώστε να γίνονται είκοσι μεταλλάξεις κατά μέσο όρο στο σύνολο των διαστάσεων κάθε σωματιδίου. Κατά αυτόν τον τρόπο αυξάνεται η τυχαιότητα των αρχικών θέσεων των σωματιδίων και συνεπώς το εύρος τιμών που εξερευνούν τα σωματίδια. Ταυτόχρονα αυξάνεται η πιθανότητα ένα σωματίδιο που έχει παγιδευτεί σε κάποιο μη ολικό βέλτιστο να απεμπλακεί από αυτό. Σε ότι αφορά τη χρονική μεταβολή της πιθανότητας μετάλλαξης επιλέγεται γραμμική μείωση, ώστε όσο προχωρούν οι επαναλήψεις να υπερισχύει ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων έναντι των μεταλλάξεων και να δοθεί η δυνατότητα στα σωματίδια να συγκλίνουν. Στις περισσότερες περιπτώσεις γραμμικής μείωσης της πιθανότητας μετάλλαξης της βιβλιογραφίας, η πιθανότητα μετάλλαξης μηδενίζεται με το τέλος των επαναλήψεων. Ωστόσο, επειδή στην παρούσα διπλωματική εργασία επιλέγεται αρκετά μεγαλύτερη αρχική πιθανότητα μετάλλαξης από αυτή που συνηθίζεται, η γραμμική μείωση προσαρμόζεται έτσι ώστε να είναι πιο απότομη, και η πιθανότητα μετάλλαξης να μηδενίζεται στα 2/3 των επαναλήψεων. Συνεπώς, για το τελευταίο 1/3 των επαναλήψεων η πιθανότητα μετάλλαξης είναι μηδενική, δίνοντας περιθώριο στα σωματίδια του αλγορίθμου βελτιστοποίησης σμήνους σωματιδίων να συγκλίνουν. Η χρονική μεταβολή της πιθανότητας μετάλλαξης που επιλέγεται εκφράζεται από την εξίσωση (5.12):

$$m = \frac{-30 \cdot \text{iter} + 20}{n_{dim}} \quad (5.12)$$

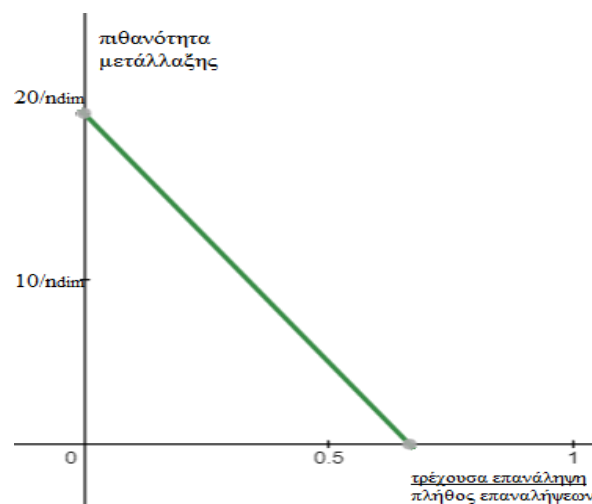
όπου n_{dim} το πλήθος των διαστάσεων του προβλήματος και $iter$ ο αριθμός της τρέχουσας επανάληψης. Η χρονική μεταβολή της πιθανότητας μετάλλαξης αναπαρίσταται γραφικά στο Σχήμα 5.4.

5.1.5 Μετρητικά αξιολόγησης μοντέλου

Αφού ολοκληρωθεί η εκπαίδευση του τεχνητού νευρωνικού δικτύου, τροφοδοτούνται σε αυτό το σύνολο των δεδομένων αξιολόγησης, ώστε να κριθεί εάν ήταν επιτυχής η εκπαίδευση. Η τελική συνολικά καλύτερη καταγεγραμμένη θέση μέσα στο σμήνος (G_{best}) αντιστοιχεί στα τελικά βάρη του τεχνητού νευρωνικού δικτύου. Κάθε διάσταση της θέσης G_{best} αντιστοιχεί σε ένα βάρος ή μια πόλωση. Με την τροφοδότηση των δεδομένων εισόδου του συνόλου αξιολόγησης, το τεχνητό νευρωνικό δίκτυο παράγει τα διαστήματα πρόβλεψης. Στη συνέχεια, τα διαστήματα πρόβλεψης συγκρίνονται με τα δεδομένα εξόδου του συνόλου αξιολόγησης μέσω κάποιων μετρητικών αξιολόγησης, και αξιολογείται η απόδοση του μοντέλου.

Τα μετρητικά αξιολόγησης που επιλέγονται για το μοντέλο LUBE-PSO-CWC είναι το PICP (σχέση (3.4)), το PINAW (σχέση (3.5)) και το CWC (σχέση (5.1)). Το PINAW επιλέγεται αντί του PINRW καθώς κατά την αξιολόγηση του μοντέλου δεν είναι απαραίτητο να «μεγεθύνονται» τα μεγαλύτερα διαστήματα πρόβλεψης. Αντίθετα, είναι προτιμότερο το μετρητικό αξιολόγησης να είναι ευθέως ανάλογο της μέσης τιμής του εύρους των διαστημάτων πρόβλεψης, ώστε να δίνεται μια ακριβέστερη εικόνα για το εύρος των διαστημάτων πρόβλεψης. Ταυτόχρονα το PINAW είναι λιγότερο πολύπλοκο υπολογιστικά σε σχέση με το PINRW. Σε ότι αφορά το CWC, όταν επιλέγεται ως μετρητικό αξιολόγησης, το $\gamma(PICP)$ δεν είναι πάντα ίσο με 1. Αντίθετα, ισχύει η σχέση (5.13):

$$\gamma(PICP) = \begin{cases} 0 & PICP \geq \mu \\ 1 & PICP < \mu \end{cases} \quad (5.13)$$



Σχήμα 5.4: Γραφική αναπαράσταση της μεταβολής της πιθανότητας μετάλλαξης σε σχέση με τις επαναλήψεις

Επομένως, όταν το P_{ICP} είναι μεγαλύτερο του ονομαστικού επιπέδου εμπιστοσύνης, το CWC ισούται με το P_{INAW}. Όταν το CWC χρησιμοποιείται για την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου, το $\gamma(PICP)$ ισούται πάντα με 1 ώστε να εξασφαλίζεται ότι τα παραγόμενα διαστήματα πρόβλεψης έχουν λίγο μεγαλύτερο P_{ICP} από το ονομαστικό επίπεδο εμπιστοσύνης και όχι οριακά ίσο. Το περιθώριο μεταξύ P_{ICP} και ονομαστικού επιπέδου εμπιστοσύνης κατά την εκπαίδευση, αυξάνει την πιθανότητα να ικανοποιείται το ονομαστικό επίπεδο εμπιστοσύνης στη μετέπειτα χρήση του μοντέλου [1]. Από τα τρία μετρητικά αξιολόγησης, το CWC αποτελεί το κυριότερο, καθώς εμπεριέχει πληροφορία και για τον δείκτη κάλυψης και για το μέσο εύρος των διαστημάτων.

5.2 ΕΚΤΙΜΗΣΗ ΚΑΤΩ ΚΑΙ ΑΝΩ ΟΡΙΟΥ ΜΕ ΠΟΛΥΚΡΙΤΗΡΙΑΚΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΣΜΗΝΟΥΣ ΣΩΜΑΤΙΔΙΩΝ

Το μοντέλο εκτίμησης κάτω και άνω ορίου με πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων (Lower Upper Bound Estimation with Multi Objective Particle Swarm Optimization ή LUBE–MOPSO) αποτελεί ένα μοντέλο πιθανοτικής πρόβλεψης που βασίζεται στη μέθοδο LUBE. Το τεχνητό νευρωνικό δίκτυο που χρησιμοποιείται για την παραγωγή των διαστημάτων πρόβλεψης είναι αντίστοιχο του μοντέλου LUBE–PSO–CWC (Σχήμα 5.1), δηλαδή εμπρόσθιας τροφοδότησης με ένα κρυφό στρώμα, τέσσερις νευρώνες στο στρώμα εισόδου και δύο νευρώνες στο στρώμα εξόδου. Το πλήθος των νευρώνων του κρυφού στρώματος και οι συναρτήσεις ενεργοποίησης καθορίζονται στην Ενότητα 5.3.

Το μοντέλο LUBE–MOPSO ακολουθεί την προσέγγιση της διατήρησης του πολυκριτηριακού χαρακτήρα που έχει από τη φύση της η βελτιστοποίηση των διαστημάτων πρόβλεψης. Η βελτιστοποίηση των διαστημάτων πρόβλεψης μπορεί να γίνει με βάση δύο αντικειμενικές συναρτήσεις, τη μεγιστοποίηση του P_{ICP} (σχέση (3.4)) και την ελαχιστοποίηση του P_{INRW} (σχέση (5.2)). Ο τύπος που περιγράφει το πρόβλημα της βελτιστοποίησης των διαστημάτων πρόβλεψης είναι:

$$\text{βελτιστοποίηση διαστημάτων πρόβλεψης} \rightarrow \begin{cases} \text{μεγιστοποίηση } P_{ICP} \\ \text{ελαχιστοποίηση } P_{INRW} \\ P_{ICP} \geq \mu \end{cases} \quad (5.14)$$

όπου μ το ονομαστικό επίπεδο εμπιστοσύνης. Ως ονομαστικό επίπεδο εμπιστοσύνης επιλέγεται $\mu=0.9$. Μέχρι στιγμής στη διαθέσιμη βιβλιογραφία, δεν υπάρχει παράδειγμα μοντέλου που να αξιοποιεί τη μέθοδο LUBE σε πολυκριτηριακά πλαίσια, δηλαδή με τη χρήση πολυκριτηριακού αλγορίθμου βελτιστοποίησης, με σκοπό την πιθανοτική πρόβλεψη της αιολικής παραγωγής.

Καθώς το μοντέλο LUBE–MOPSO προορίζεται για πιθανοτική πρόβλεψη της αιολικής παραγωγής και τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγησή του είναι ίδια με αυτά που χρησιμοποιούνται για το μοντέλο LUBE–PSO–CWC, η μέθοδος αρχικοποίησης των βαρών και των πολώσεων που επιλέγεται είναι ίδια με αυτή των βαρών και των πολώσεων του μοντέλου LUBE–PSO–CWC. Όλα τα βάρη και οι πολώσεις του τεχνητού νευρωνικού δικτύου αρχικοποιούνται τυχαία, ομοιόμορφα και συμμετρικά γύρω από το μηδέν με μοναδιαίο εύρος διαστήματος. Στη συνέχεια, για τα βάρη

μεταξύ του στρώματος εισόδου και του κρυφού στρώματος, και τις πολώσεις του κρυφού στρώματος, εφαρμόζεται η μέθοδος αρχικοποίησης Nguyen–Widrow που περιγράφεται από τις σχέσεις (5.5) και (5.8).

5.2.1 Αλγόριθμος Βελτιστοποίησης

Ο αλγόριθμος βελτιστοποίησης που επιλέγεται για την εκπαίδευση του τεχνητού νευρωνικού δικτύου βασίζεται στον αλγόριθμο πολυκριτηριακής βελτιστοποίησης σμήνους σωματιδίων που προτάθηκε από τους Coello και Lechuga [3]. Κάθε σωματίδιο του σμήνους αντιπροσωπεύει μία διαφορετική λύση, επομένως σε κάθε σωματίδιο αντιστοιχεί ένα τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE–MOPSO. Οι διαστάσεις των θέσεων και των ταχυτήτων των σωματιδίων υπολογίζονται σύμφωνα με την εξίσωση (5.9).

Οι θέσεις των σωματιδίων έχουν ήδη αρχικοποιηθεί καθώς ταυτίζονται με τα βάρη και τις πολώσεις των τεχνητών νευρωνικών δικτύων που αντιπροσωπεύουν. Οι ταχύτητες των σωματιδίων αρχικοποιούνται τυχαία, ομοιόμορφα και συμμετρικά γύρω από το μηδέν με μοναδιαίο εύρος διαστήματος. Ως καλύτερη καταγεγραμμένη θέση κάθε σωματιδίου επιλέγεται η αρχική του θέση.

Για κάθε σωματίδιο παράγονται τα διαστήματα πρόβλεψης και υπολογίζονται τα αντίστοιχα PICP και PINRW. Στη συνέχεια, γίνεται μία προς μία σύγκριση των λύσεων που φέρουν τα σωματίδια ώστε να βρεθούν οι καλύτερες έγκυρες λύσεις, οι οποίες καταχωρούνται στο αρχείο αποθήκευσης (repository). Μια λύση θεωρείται έγκυρη όταν δεν παραβιάζει κάποιον από τους περιορισμούς του προβλήματος. Στο μοντέλο LUBE–MOPSO έγκυρες λύσεις θεωρούνται εκείνες με PICP μεγαλύτερο ή ίσο του ονομαστικού επιπέδου εμπιστοσύνης. Κατά τη σύγκριση δύο λύσεων x_1 και x_2 μπορούν να προκύψουν οι παρακάτω περιπτώσεις:

- i. Αμφότερες οι x_1 και x_2 να μην είναι έγκυρες. Σε αυτή την περίπτωση επιλέγεται η λύση που παραβιάζει λιγότερο τους περιορισμούς του προβλήματος.
- ii. Η λύση x_1 να είναι έγκυρη ενώ η x_2 όχι. Σε αυτή την περίπτωση επιλέγεται η x_1 .
- iii. Η λύση x_2 να είναι έγκυρη ενώ η x_1 όχι. Σε αυτή την περίπτωση επιλέγεται η x_2 .
- iv. Αμφότερες οι x_1 και x_2 να είναι έγκυρες. Σε αυτή την περίπτωση η επιλογή βασίζεται στην κατά Pareto υπερίσχυση [28]. Εάν οι δύο λύσεις είναι μη κυριαρχούμενες μεταξύ τους, επιλέγονται και οι δύο.

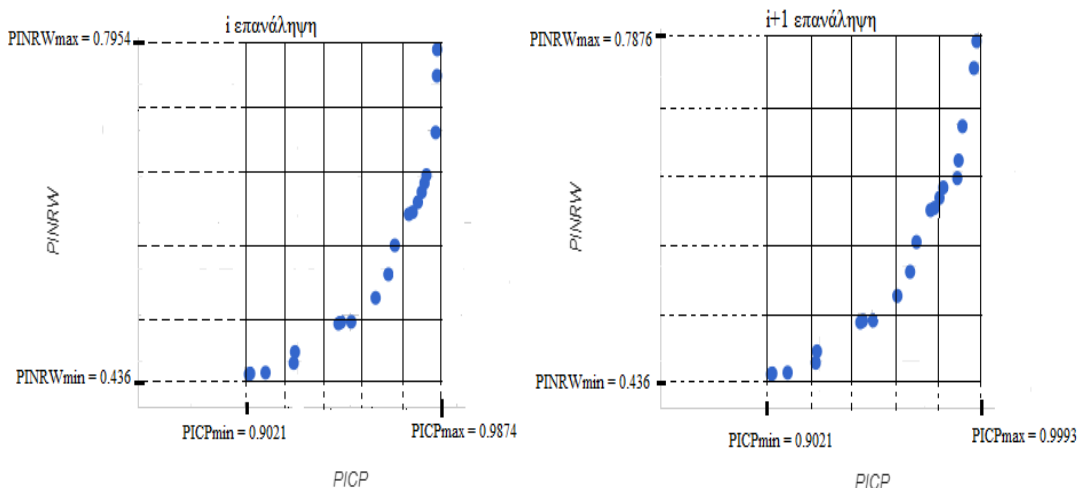
Σκοπός του αλγορίθμου είναι οι τελικές καταχωρημένες λύσεις του αρχείου αποθήκευσης να προσεγγίζουν όσο καλύτερα γίνεται το μέτωπο Pareto [28] και να διατηρούν μεγάλη ποικιλομορφία μεταξύ τους. Για την εξασφάλιση της ποικιλομορφίας των λύσεων ακολουθείται η μέθοδος της γεωγραφικής αναπαράστασης των λύσεων του αρχείου αποθήκευσης και ο διαχωρισμός του χώρου σε υπερκύβους ώστε να δημιουργηθεί ένα πλέγμα (Ενότητα 2.5.2). Στην παρούσα διπλωματική εργασία γίνεται χρήση της έννοιας του προσαρμοστικού πλέγματος [48]. Οι διαστάσεις του προσαρμοστικού πλέγματος δεν είναι σταθερές αλλά εξαρτώνται από τις λύσεις που εμπεριέχονται στο αρχείο αποθήκευσης.

Καθώς η βελτιστοποίηση των διαστημάτων πρόβλεψης του μοντέλου LUBE–MOPSO είναι πρόβλημα δύο αντικειμενικών συναρτήσεων, το πλέγμα έχει δύο διαστάσεις και η κάθε λύση φέρει δύο συντεταγμένες, που αντιστοιχούν στο PICP και το PINRW. Σε κάθε επανάληψη, εάν είναι απαραίτητο, υπολογίζονται οι μέγιστες και οι ελάχιστες τιμές των PICP και PINRW ($PICP_{max}, PINRW_{max}, PICP_{min}, PINRW_{min}$) μεταξύ των λύσεων του αρχείου αποθήκευσης. Οι τέσσερις αυτές τιμές αποτελούν και τα όρια του προσαρμοστικού πλέγματος ώστε κάθε λύση να εμπεριέχεται μέσα στο πλέγμα. Το πλήθος των υπερκύβων που συγκροτούν το πλέγμα παραμένει σταθερό, επομένως ανάλογα με το εμβαδόν του πλέγματος προσαρμόζονται και οι διαστάσεις των υπερκύβων. Οι διαστάσεις των υπερκύβων σε κάθε επανάληψη προκύπτουν από τις σχέσεις (5.15) και (5.16):

$$x_{PICP} = \frac{PICP_{max} - PICP_{min}}{n_{hypercubes}} \quad (5.15)$$

$$y_{PINRW} = \frac{PINRW_{max} - PINRW_{min}}{n_{hypercubes}} \quad (5.16)$$

όπου x_{PICP} το μήκος των υπερκύβων, y_{PINRW} το ύψος των υπερκύβων και $n_{hypercubes}$ το πλήθος των υπερκύβων σε κάθε διάσταση. Το PICP αντιστοιχεί στον άξονα x'x και το PINRW στον άξονα y'y. Ένα παράδειγμα γεωγραφικής αναπαράστασης είκοσι λύσεων του μοντέλου LUBE–MOPSO με χρήση του προσαρμοστικού πλέγματος και πέντε υπερκύβων ανά διάσταση για δύο διαδοχικές επαναλήψεις φαίνεται στο Σχήμα 5.5. Κατά την επανάληψη $i+1$ το μέγιστο PICP και το μέγιστο PINRW έχουν μεταβληθεί σε σχέση με την επανάληψη i , επομένως μεταβάλλονται και τα όρια του προσαρμοστικού πλέγματος. Οι υπερκύβοι κατά την επανάληψη $i+1$ έχουν ελάχιστα μεγαλύτερες διαστάσεις σε σχέση με την επανάληψη i , ώστε να καλύψουν τη διαφορά στα όρια του πλέγματος.



Σχήμα 5.5: Η γεωγραφική απεικόνιση είκοσι λύσεων του αρχείου αποθήκευσης, το προσαρμοστικό πλέγμα και οι οριακές τιμές των PICP, PINRW κατά τη διάρκεια δύο διαδοχικών επαναλήψεων της εκπαίδευσης του μοντέλου LUBE–MOPSO.

Σε κάθε κύβο ανατίθεται μια τιμή h η οποία ισούται με το πλήθος των λύσεων που βρίσκεται εντός του. Εάν σε κάποια επανάληψη υπάρξει υπερχειλίση του αρχείου αποθήκευσης, διαγράφονται οι λύσεις που περισσεύουν. Η διαγραφή γίνεται με βάση την τιμή h . Η διαγραφή της πρώτης λύσης γίνεται επιλέγοντας τυχαία μια λύση από τον υπερκύβο με την υψηλότερη τιμή h . Η τιμή h αυτού του υπερκύβου μειώνεται κατά μία μονάδα και η διαδικασία επαναλαμβάνεται έως ότου το πλήθος των λύσεων εντός του αρχείου αποθήκευσης γίνει ίσο με το μέγεθος του αρχείου αποθήκευσης. Κατά αυτόν τον τρόπο εξασφαλίζεται ότι θα διατηρούνται οι λύσεις με τη μεγαλύτερη ποικιλομορφία.

Το προσαρμοστικό πλέγμα επιλέγεται έναντι του σταθερού πλέγματος διότι αυξάνει την ποικιλομορφία μεταξύ των τελικών λύσεων εντός του αρχείου αποθήκευσης. Η προσαρμογή των ορίων του και κατά επέκταση των διαστάσεων των υπερκύβων δίνει τη δυνατότητα απεικόνισης του συνωστισμού λύσεων σε υψηλότερη ανάλυση.

Η επιλογή της συνολικά καλύτερης θέσης μέσα στο σμήνος (Gbest) για την εξίσωση ανανέωσης των ταχυτήτων των σωματιδίων (εξίσωση (2.15)) βασίζεται στη μέθοδο που προτείνεται από τους B. Xue, M. Zhang και W. N. Browne [49]. Από το σύνολο των λύσεων του αρχείου αποθήκευσης επιλέγονται τυχαία δύο λύσεις. Όλες οι λύσεις έχουν την ίδια πιθανότητα να επιλεγούν ανεξαρτήτως της τιμής h του υπερκύβου στον οποίο βρίσκονται. Από τις δύο επιλεγμένες λύσεις, ως συνολικά καλύτερη θέση επιλέγεται η λύση που βρίσκεται εντός του υπερκύβου με τη μικρότερη τιμή h . Αν η τιμή h των δύο υπερκύβων είναι ίδια, επιλέγεται τυχαία μία από τις δύο λύσεις. Σε σχέση με τη μέθοδο επιλογής της συνολικά καλύτερης θέσης μέσα στο σμήνος που προτείνεται από τους Coello και Lechuga [3] (Ενότητα 2.5.2), η μέθοδος που επιλέγεται για το μοντέλο LUBE-MOPSO της παρούσας διπλωματικής εργασίας είναι υπολογιστικά λιγότερο πολύπλοκη. Ταυτόχρονα μειώνει την πιθανότητα να εγκλωβιστούν τα σωματίδια σε ένα τοπικό βέλτιστο κατά τη διάρκεια των πρώτων επαναλήψεων [49].

Οι ταχύτητες και οι θέσεις των σωματιδίων ανανεώνονται σύμφωνα με τις εξισώσεις (2.15) και (2.16). Οι εξισώσεις (2.15) και (2.16) εφαρμόζονται για κάθε διάσταση του μοντέλου. Εάν κάποια διάσταση της ταχύτητας (ή της θέσης) ενός σωματιδίου βγει εκτός του καθορισμένου εύρους τιμών της $[-v_{max}, v_{max}]$ (ή $[-x_{max}, x_{max}]$), περιορίζεται εντός αυτού. Συγκεκριμένα, αν για κάποια διάσταση i της ταχύτητας ενός σωματιδίου προκύψει $v_i > v_{max}$, η ταχύτητα της διάστασης αυτής λαμβάνει την τιμή v_{max} . Αν για κάποια διάσταση i της ταχύτητας ενός σωματιδίου προκύψει $v_i < -v_{max}$, η ταχύτητα της διάστασης αυτής λαμβάνει την τιμή $-v_{max}$. Ομοια διαδικασία ακολουθείται για κάθε διάσταση των θέσεων των σωματιδίων. Στις θέσεις των σωματιδίων εφαρμόζεται τελεστής μετάλλαξης που εκφράζεται από τη σχέση (5.11). Η πιθανότητα μετάλλαξης μεταβάλλεται σύμφωνα με την εξίσωση (5.12). Εάν είναι απαραίτητο, οι θέσεις των σωματιδίων που υπέστησαν μετάλλαξη περιορίζονται εντός του εύρους τιμών τους. Με βάση τις καινούριες θέσεις των σωματιδίων παράγονται τα νέα διαστήματα πρόβλεψης και υπολογίζονται εκ νέου τα PICP και PINRW για κάθε σωματίδιο. Εάν είναι απαραίτητο, αντικαθίσταται η καλύτερη καταγεγραμμένη θέση (Pbest) κάθε σωματιδίου. Το κριτήριο αντικατάστασης της καλύτερης καταγεγραμμένης θέσης κάθε σωματιδίου είναι η κατά Pareto υπερίσχυση. Εάν οι δύο λύσεις είναι μη κυριαρχούμενες μεταξύ τους, επιλέγεται με τυχαίο τρόπο ως Pbest μία από τις δύο. Οι λύσεις που φέρουν τα σωματίδια συγκρίνονται μεταξύ τους αλλά και με τις λύσεις που είναι αποθηκευμένες στο αρχείο αποθήκευσης. Οι καλύτερες έγκυρες λύσεις αποθηκεύονται στο

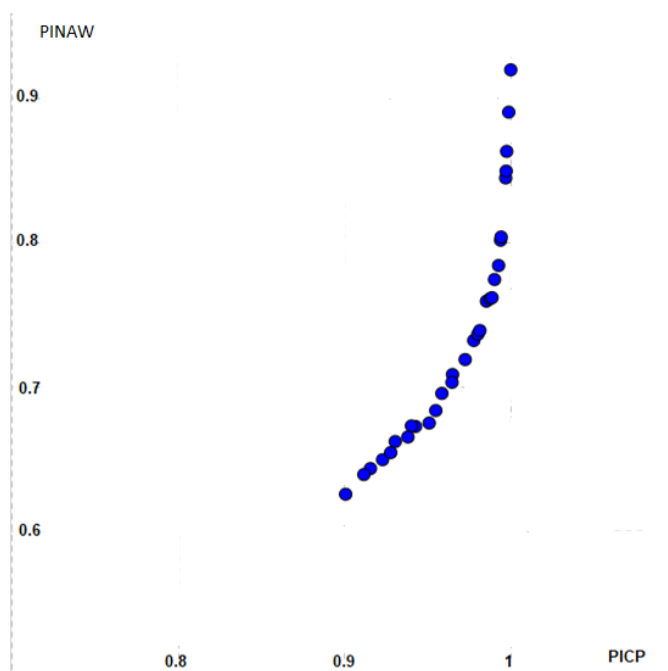
αρχείο αποθήκευσης με βάση τα κριτήρια σύγκρισης που περιγράφηκαν παραπάνω. Εφόσον χρειάζεται, ανανεώνονται οι διαστάσεις του προσαρμοστικού πλέγματος και των υπερκύβων, και διαγράφονται οι λύσεις με τον μεγαλύτερο συνωστισμό στην περίπτωση υπερχειλίσης του αρχείου αποθήκευσης. Επιλέγεται η συνολικά καλύτερη καταγεγραμμένη θέση μέσα στο σμήνος και ανανεώνονται εκ νέου οι θέσεις και οι ταχύτητες των σωματιδίων. Η διαδικασία επαναλαμβάνεται μέχρι να συμπληρωθεί ο καθορισμένος αριθμός επαναλήψεων.

Το πλήθος των επαναλήψεων πρέπει να είναι αρκετά μεγάλο ώστε να προλάβουν τα σωματίδια να συγκλίνουν στο ολικό βέλτιστο, ωστόσο από ένα σημείο και έπειτα οι μεταβολές στα τελικά βάρη και στις τελικές πολώσεις είναι αμελητέες. Η αύξηση του αριθμού των επαναλήψεων πέρα από αυτό το σημείο αυξάνει τον χρόνο επίλυσης του προβλήματος χωρίς να βελτιώνει ιδιαίτερα τη λύση του. Το μέγεθος του αρχείου αποθήκευσης επιλέγεται ίσο με το πλήθος των σωματιδίων ($n_{particles}$). Κάθε διάσταση του πλέγματος καθορίζεται να έχει $\frac{n_{particles}}{4}$ υπερκύβους, επομένως το πλέγμα συνολικά αποτελείται από $\frac{n_{particles}}{4} \cdot \frac{n_{particles}}{4} = \frac{n_{particles}^2}{16}$ υπερκύβους. Ως οριακή τιμή για τις θέσεις των σωματιδίων επιλέγεται $x_{max} = 4$ επομένως κάθε διάσταση των θέσεων των σωματιδίων περιορίζεται εντός του διαστήματος $[-4, 4]$. Ως μέγιστη ταχύτητα επιλέγεται $v_{max} = 1$, επομένως κάθε διάσταση των ταχυτήτων των σωματιδίων περιορίζεται εντός του διαστήματος $[-1, 1]$. Ο καθορισμός των παραμέτρων c_1, c_2, w των εξισώσεων (2.15) και (2.16) καθώς και του πλήθους των σωματιδίων και του πλήθους των επαναλήψεων για το μοντέλο LUBE-MOPSO γίνεται στην Ενότητα 5.3.2.

5.2.2 Μετρητικά αξιολόγησης μοντέλου

Οι λύσεις που είναι αποθηκευμένες στο αρχείο αποθήκευσης κατά το τέλος της εκπαίδευσης του τεχνητού νευρωνικού δικτύου αποτελούν τις λύσεις του προβλήματος. Κάθε διάσταση μιας λύσης αντιστοιχεί σε ένα βάρος ή μια πόλωση, επομένως υπάρχουν διαφορετικοί συνδυασμοί βαρών και πολώσεων, ίσοι με το πλήθος των λύσεων που βρίσκονται στο αρχείο αποθήκευσης. Σε αντίθεση με το μοντέλο LUBE-PSO-CWC, στο μοντέλο LUBE-MOPSO παράγονται πολλαπλές λύσεις κατά την εκπαίδευσή του. Το τεχνητό νευρωνικό δίκτυο τροφοδοτείται με τα δεδομένα εισόδου του συνόλου αξιολόγησης. Η διαδικασία της αξιολόγησης συντελείται ξεχωριστά για κάθε λύση του αρχείου αποθήκευσης. Επομένως παράγονται διαστήματα πρόβλεψης για κάθε συνδυασμό βαρών και πολώσεων που προκύπτουν από τις λύσεις του αρχείου αποθήκευσης. Τα διαστήματα πρόβλεψης συγκρίνονται με τα δεδομένα εξόδου του συνόλου αξιολόγησης μέσω κάποιων μετρητικών αξιολόγησης, και αξιολογείται η απόδοση του μοντέλου.

Τα μετρητικά αξιολόγησης που επιλέγονται για το μοντέλο LUBE-MOPSO είναι το PICP (σχέση (3.4)), το PINAW (σχέση (3.5)) και το CWC (σχέση (5.1)). Το PINAW επιλέγεται αντί του PINRW για τους ίδιους λόγους που επιλέχθηκε για την αξιολόγηση του μοντέλου LUBE-PSO-CWC (Ενότητα 5.1.5). Καθώς το CWC επιλέγεται ως μετρητικό αξιολόγησης, η τιμή του γ (PICP) καθορίζεται από τη σχέση (5.13). Για το η επιλέγεται $\eta=80$. Από τα τρία μετρητικά αξιολόγησης, το CWC αποτελεί το κυριότερο, καθώς εμπεριέχει πληροφορία και για τον δείκτη κάλυψης και για το μέσο εύρος των διαστημάτων.



Σχήμα 5.6: Παράδειγμα απεικόνισης της αξιολόγησης λύσεων που παράχθηκαν κατά την εκπαίδευση του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE–MOPSO

Στο Σχήμα 5.6 φαίνεται ένα παράδειγμα γεωγραφικής απεικόνισης της αξιολόγησης τριάντα λύσεων που παράχθηκαν κατά την εκπαίδευση του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE–MOPSO. Πέρα από την ευχέρεια που προσφέρει το μοντέλο LUBE–MOPSO λόγω της δυνατότητας επιλογής μεταξύ διαφορετικών λύσεων, η οπτική απεικόνιση των λύσεων με βάση την αξιολόγησή τους διευκολύνει την κατανόηση της συσχέτισης μεταξύ των αντικειμενικών συναρτήσεων του προβλήματος.

5.3 ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

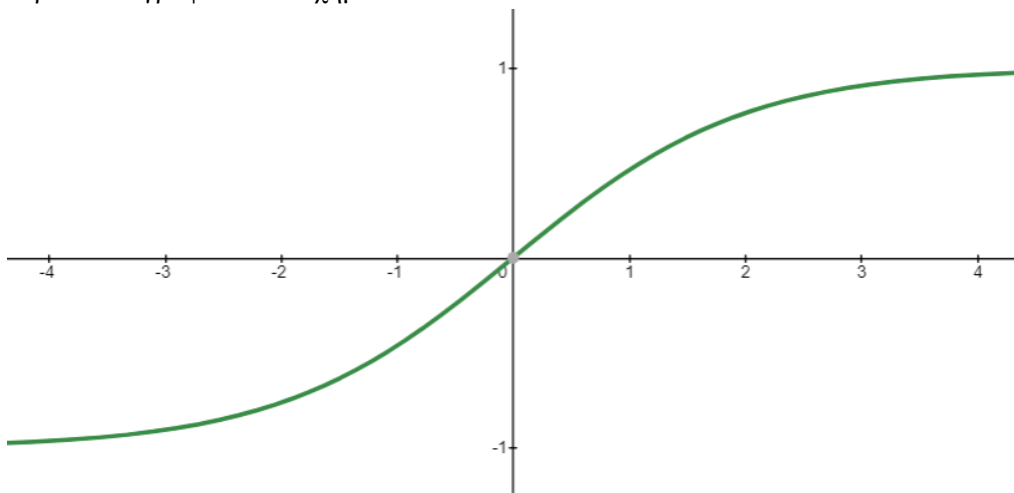
Ο καθορισμός των παραμέτρων στα στατιστικά μοντέλα πρόβλεψης που βασίζουν τη λειτουργία τους σε τεχνητά νευρωνικά δίκτυα επηρεάζει σε μεγάλο βαθμό την αποδοτικότητα των μοντέλων. Εάν ένα τεχνητό νευρωνικό δίκτυο αποτελείται από λιγότερους νευρώνες από το απαιτούμενο, δεν θα μπορέσει να μοντελοποιήσει ικανοποιητικά το πρόβλημα. Το φαινόμενο αυτό ονομάζεται υποπροσαρμογή. Εάν το πλήθος των νευρώνων και των κρυφών στρωμάτων αυξηθούν υπερβολικά, το τεχνητό νευρωνικό δίκτυο θα μοντελοποιήσει και τον θόρυβο που εμπεριέχεται στα δεδομένα εκπαίδευσης, χάνοντας έτσι τη δυνατότητα γενίκευσης. Το φαινόμενο αυτό ονομάζεται υπερπροσαρμογή. Σε ότι αφορά τις παραμέτρους των αλγορίθμων βελτιστοποίησης, η επιλογή τους εξαρτάται από τη φύση του προβλήματος και από τη δομή του τεχνητού νευρωνικού δικτύου. Οι τιμές των παραμέτρων των μοντέλων LUBE–PSO–CWC και LUBE–MOPSO της παρούσας διπλωματικής εργασίας δεν λαμβάνονται έτοιμες από κάποια άλλη εργασία που αφορά ανάπτυξη αντίστοιχων μοντέλων, αλλά προσδιορίζονται από την αρχή μέσω διαφόρων μεθόδων βελτιστοποίησης παραμέτρων. Η βελτιστοποίηση παραμέτρων γίνεται με μία διαδικασία επαναλαμβανόμενων δοκιμών με

διαφορετικούς συνδυασμούς τιμών των παραμέτρων. Η βελτιστοποίηση παραμέτρων στην παρούσα διπλωματική εργασία γίνεται για να υπάρξει προσαρμογή των μοντέλων στα δεδομένα που χρησιμοποιούνται και για να επιτευχθεί η μέγιστη δυνατή βελτίωση της αποδοτικότητάς τους. Οι μέθοδοι βελτιστοποίησης και οι συνδυασμοί παραμέτρων που επιλέγονται για δοκιμή στην παρούσα διπλωματική εργασία προκύπτουν από παραδείγματα που υπάρχουν στη διαθέσιμη βιβλιογραφία, από τη θεωρία των αλγορίθμων βελτιστοποίησης, αλλά και εμπειρικά μέσα από δοκιμές και πειραματισμούς.

5.3.1 Βελτιστοποίηση παραμέτρων LUBE–PSO–CWC

Τα δεδομένα εξόδου του διαγωνισμού GEFCom 2014 είναι ήδη κανονικοποιημένα στο διάστημα $[0, 1]$. Επομένως η συνάρτηση ενεργοποίησης του στρώματος εξόδου του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE–PSO–CWC θα πρέπει να έχει σύνολο τιμών στο διάστημα $[0, 1]$. Τέτοιες συναρτήσεις ενεργοποίησης είναι η σιγμοειδής και η ReLU που περιγράφονται από τις σχέσεις (2.3) και (2.5), αντίστοιχα. Πειραματικά προκύπτει ότι σταθερά καλύτερα αποτελέσματα παράγει η σιγμοειδής συνάρτηση ενεργοποίησης. Η σιγμοειδής συνάρτηση ενεργοποίησης, λόγω της μορφής της, έχει καλύτερη απόκριση από τη ReLU στη μοντελοποίηση μη γραμμικών προβλημάτων, επομένως είναι ιδανική για πρόβλεψη της αιολικής παραγωγής. Επομένως, ως συνάρτηση ενεργοποίησης του στρώματος εξόδου επιλέγεται η σιγμοειδής (Σχήμα 2.2).

Τα δεδομένα εισόδου κανονικοποιούνται στο διάστημα $[-1, 1]$. Στην περίπτωση αυτή προτιμάται η συνάρτηση ενεργοποίησης του κρυφού στρώματος να έχει σύνολο τιμών στο διάστημα $[-1, 1]$ ώστε να μπορεί να προσαρμοστεί καλύτερα στα δεδομένα εισόδου [27]. Τέτοιες συναρτήσεις ενεργοποίησης είναι η διπολική σιγμοειδής και η υπερβολική εφαπτομένη που περιγράφονται από τις σχέσεις (2.4) και (2.6), αντίστοιχα. Πειραματικά προκύπτει ότι η διπολική σιγμοειδής έχει πιο σταθερή απόκριση και παράγει καλύτερα αποτελέσματα σε σχέση με την υπερβολική εφαπτομένη. Επομένως, ως συνάρτηση ενεργοποίησης του κρυφού στρώματος επιλέγεται η διπολική σιγμοειδής, η οποία αναπαρίσταται γραφικά στο Σχήμα 5.7.



Σχήμα 5.7: Η διπολική σιγμοειδής συνάρτηση ενεργοποίησης

Σε ότι αφορά τη βελτιστοποίηση σμήνους σωματιδίων, θα πρέπει να γίνει καθορισμός των παραμέτρων w, c_1, c_2 της εξίσωσης (2.15). Το βάρος αδράνειας w παίρνει τιμές στο διάστημα $[0.1, 1.2]$. Όσο μεγαλύτερες τιμές παίρνει το w τόσο μεγαλύτερο βάρος δίνουν τα σωματίδια στη δική τους τρέχουσα κατεύθυνση. Όταν το w επιλέγεται μικρό, τα σωματίδια δίνουν περισσότερο βάρος στην καλύτερη θέση που έχουν καταγράψει το καθένα ξεχωριστά, αλλά και στη συνολικά καλύτερη θέση μέσα στο σμήνος για τον καθορισμό της κατεύθυνσής τους. Συνήθως επιλέγεται γραμμική μείωση του w όσο προχωρούν οι επαναλήψεις, ώστε στις πρώτες επαναλήψεις να γίνεται η μέγιστη δυνατή εξερεύνηση του πεδίου τιμών, ενώ όσο πλησιάζει το τέλος των επαναλήψεων να ευνοείται η σύγκλιση των σωματιδίων. Σε ορισμένα προβλήματα, το w επιλέγεται να λαμβάνει μια τυχαία τιμή σε κάθε επανάληψη [49].

Στην παρούσα διπλωματική εργασία για το μοντέλο LUBE-PSO-CWC γίνονται δοκιμές για γραμμική μείωση του w αλλά και για παραγωγή τυχαίας τιμής σε κάθε επανάληψη. Τα διαστήματα που επιλέγονται για δοκιμή είναι τα $[0.4, 0.7]$ και $[0.7, 0.9]$. Επομένως, γίνονται δοκιμές για τέσσερις διαφορετικούς συνδυασμούς που προκύπτουν από τους δύο τρόπους μεταβολής του w και από τα δύο διαφορετικά πεδία τιμών. Οι δοκιμές γίνονται για $c_1 = 1.22, c_2 = 1.47$ [27], 80 σωματίδια, 13 νευρώνες κρυφού στρώματος και 100 επαναλήψεις. Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση του τεχνητού νευρωνικού δικτύου αφορούν τον χειμώνα του 2013 (Δεκέμβριος του 2012 έως Φεβρουάριο του 2013) της 1^{ης} ζώνης. Για κάθε συνδυασμό γίνεται τρεις φορές εκπαίδευση και αξιολόγηση και υπολογίζεται η μέση τιμή των μετρητικών αξιολόγησης. Ως δεδομένα εκπαίδευσης επιλέγονται τα δεδομένα που αφορούν τους δύο πρώτους μήνες, ενώ για την αξιολόγηση χρησιμοποιούνται τα δεδομένα του τρίτου μήνα. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 5.1. Κριτήριο για την επιλογή του βέλτιστου συνδυασμού αποτελεί η μέση τιμή του CWC, καθώς εμπεριέχει πληροφορία ταυτόχρονα και για τον βαθμό κάλυψης και για το μέσο εύρος των διαστημάτων πρόβλεψης. Επομένως, σύμφωνα με τα αποτελέσματα του Πίνακα 5.1, επιλέγεται γραμμική μείωση του w με μέγιστη τιμή $w_{max} = 0.7$ και ελάχιστη τιμή $w_{min} = 0.4$, επειδή αυτή η επιλογή δίνει την ελάχιστη τιμή του δείκτη CWC.

Οι παράμετροι c_1 και c_2 παίρνουν τιμές στο διάστημα $[1, 2]$. Συνήθως παραμένουν σταθερές σε όλη τη διάρκεια του αλγορίθμου. Στην παρούσα διπλωματική εργασία για το μοντέλο LUBE-PSO-CWC γίνονται δοκιμές για c_1 και c_2 στο διάστημα $[1, 2]$ με βήμα 0.1. Επομένως, τα c_1 και c_2 παίρνουν 11 διαφορετικές τιμές το καθένα, και γίνονται δοκιμές 121 διαφορετικών συνδυασμών. Οι δοκιμές γίνονται για γραμμική μείωση του w στο διάστημα $[0.4, 0.7]$, 80 σωματίδια, 13 νευρώνες κρυφού στρώματος και 100 επαναλήψεις. Τα δεδομένα αφορούν τον χειμώνα του 2013 της 1^{ης} ζώνης. Για κάθε συνδυασμό γίνεται τρεις φορές εκπαίδευση και αξιολόγηση και υπολογίζεται η μέση τιμή των μετρητικών αξιολόγησης. Ως δεδομένα εκπαίδευσης επιλέγονται τα δεδομένα που αφορούν τους δύο πρώτους μήνες, ενώ για την αξιολόγηση χρησιμοποιούνται τα δεδομένα του τρίτου μήνα. Κριτήριο για την επιλογή του βέλτιστου συνδυασμού αποτελεί η μέση τιμή του CWC. Στον Πίνακα 5.2 παρουσιάζεται μέρος των αποτελεσμάτων που περιλαμβάνει και τον βέλτιστο συνδυασμό c_1 και c_2 . Ο βέλτιστος συνδυασμός προκύπτει για $c_1 = 1.2$ και $c_2 = 1.3$.

Στη συνέχεια καθορίζεται το πλήθος των νευρώνων του κρυφού στρώματος του τεχνητού νευρωνικού δικτύου και το πλήθος των σωματιδίων της βελτιστοποίησης σμήνους

Πίνακας 5.1: Καθορισμός βέλτιστου βάρους αδράνειας w για το μοντέλο LUBE-PSO-CWC.

w	Μέση τιμή PICP	Μέση τιμή PINAW	Μέση τιμή CWC
Τυχαίο [0.7, 0.9]	0.9327	0.5593	0.5593
Τυχαίο [0.4, 0.7]	0.9241	0.5563	0.5563
Γραμμική μείωση [0.7, 0.9]	0.9358	0.5709	0.5709
Γραμμική μείωση [0.4, 0.7]	0.9253	0.5548	0.5548

σωματιδίων. Ο καθορισμός του πλήθους των νευρώνων του κρυφού στρώματος γίνεται με χρήση της μεθόδου διασταυρωμένης επικύρωσης K τμημάτων (K-Fold Cross Validation) [50]. Η μέθοδος διασταυρωμένης επικύρωσης K τμημάτων έγκειται στον διαχωρισμό του συνόλου δεδομένων εκπαίδευσης σε K ίσα υποσύνολα και στην επανάληψη της εκπαίδευσης και της αξιολόγησης του τεχνητού νευρωνικού δικτύου K φορές. Ως δεδομένα εκπαίδευσης χρησιμοποιούνται $K-1$ υποσύνολα του αρχικού συνόλου των δεδομένων εκπαίδευσης, ενώ το υποσύνολο που περισσεύει χρησιμοποιείται ως σύνολο δεδομένων αξιολόγησης. Η διαδικασία επαναλαμβάνεται K φορές ώστε κάθε ένα από τα υποσύνολα του αρχικού συνόλου των δεδομένων εκπαίδευσης να έχει χρησιμοποιηθεί μία φορά ως σύνολο δεδομένων αξιολόγησης.

Η διασταυρωμένη επικύρωση K τμημάτων είναι μια μέθοδος που χρησιμοποιείται μεταξύ άλλων για βελτιστοποίηση παραμέτρων. Το σύνολο των δεδομένων εκπαίδευσης σε ορισμένες περιπτώσεις παρουσιάζει εποχικότητα ή κάποιο άλλο είδος μεροληψίας, επομένως είναι πιθανό οι βέλτιστες τιμές των παραμέτρων που θα προκύψουν να είναι επίσης μεροληπτικές. Η εφαρμογή της διασταυρωμένης επικύρωσης K τμημάτων στη διαδικασία βελτιστοποίησης των παραμέτρων εξασφαλίζει ότι εξαλείφεται οποιαδήποτε εποχικότητα ή μεροληψία του συνόλου των δεδομένων εκπαίδευσης. Για κάθε συνδυασμό παραμέτρων εφαρμόζεται εκ νέου η διασταυρωμένη επικύρωση K τμημάτων και λαμβάνεται η μέση τιμή ή η διάμεσος των τιμών του μετρητικού αξιολόγησης που προκύπτουν από τις K επαναλήψεις της διαδικασίας. Ο διαχωρισμός του αρχικού συνόλου των δεδομένων εκπαίδευσης σε K διαφορετικά υποσύνολα και η επανάληψη της εκπαίδευσης του τεχνητού νευρωνικού δικτύου με διαφορετικά δεδομένα εκπαίδευσης κάθε φορά, εξασφαλίζει ότι οι τιμές των παραμέτρων που θα προκύψουν θα είναι πραγματικά βέλτιστες.

Για τον καθορισμό του βέλτιστου πλήθους των νευρώνων του κρυφού στρώματος του τεχνητού νευρωνικού δικτύου του μοντέλου LUBE-PSO-CWC επιλέγεται διασταυρωμένη επικύρωση 5 τμημάτων (Σχήμα 5.8). Ως σύνολο δεδομένων εκπαίδευσης χρησιμοποιούνται τα δεδομένα που αφορούν τους δύο πρώτους μήνες του χειμώνα του 2013 της 1^{ης} ζώνης. Τα δεδομένα αυτά αποτελούνται από 1488 ωριαίες παρατηρήσεις. Επομένως, το σύνολο εκπαίδευσης σε κάθε επανάληψη της διασταυρωμένης επικύρωσης 5 τμημάτων αποτελείται από 1190 παρατηρήσεις, ενώ το σύνολο αξιολόγησης από τις υπόλοιπες 298 ωριαίες παρατηρήσεις. Το πλήθος των νευρώνων του κρυφού στρώματος μεταβάλλεται από 1 έως 20 με μοναδιαίο βήμα. Για κάθε αριθμό νευρώνων γίνονται 5 επαναλήψεις με διαφορετικά σύνολα δεδομένων εκπαίδευσης και αξιολόγησης, και λαμβάνεται η διάμεσος των μετρητικών αξιολόγησης για την τελική επιλογή του βέλτιστου αριθμού νευρώνων. Οι

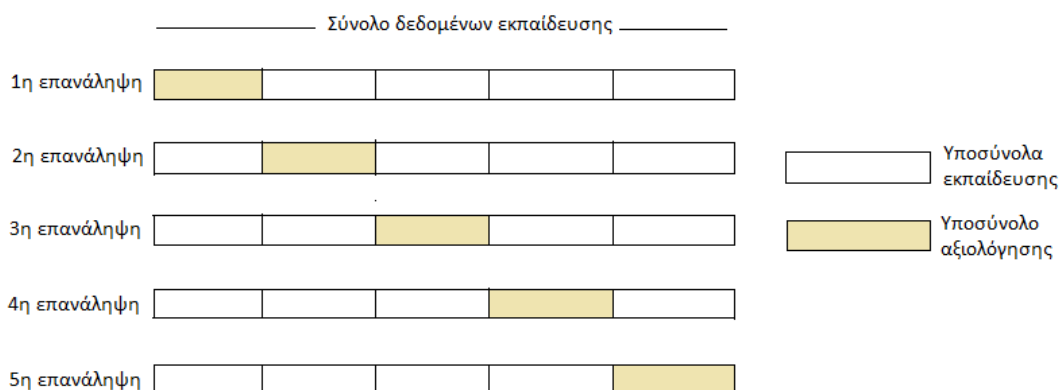
Πίνακας 5.2: Ενδεικτικά αποτελέσματα της βελτιστοποίησης των παραμέτρων c_1 και c_2 για το μοντέλο LUBE-PSO-CWC.

c_1	c_2	Μέση τιμή PICP	Μέση τιμή PINAW	Μέση τιμή CWC
1.2	1	0.9333	0.5679	0.5679
1.2	1.1	0.9370	0.5688	0.5688
1.2	1.2	0.9302	0.5616	0.5616
1.2	1.3	0.9228	0.5464	0.5464
1.2	1.4	0.934	0.5634	0.5634
1.2	1.5	0.9401	0.5684	0.5684
1.2	1.6	0.942	0.5663	0.5663

δοκιμές γίνονται για γραμμική μείωση του w στο διάστημα $[0.4, 0.7]$, $c_1 = 1.2$, $c_2 = 1.3$ και 100 επαναλήψεις του αλγορίθμου βελτιστοποίησης.

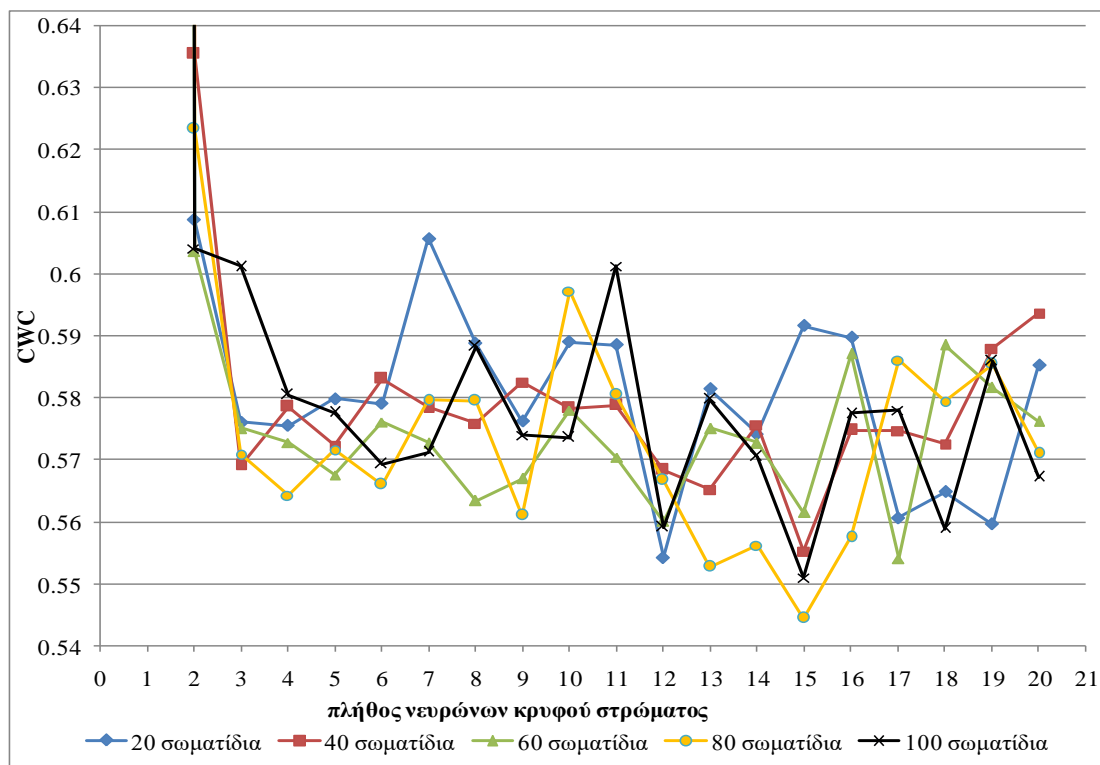
Η διασταυρωμένη επικύρωση 5 τμημάτων εφαρμόζεται πέντε φορές, για διαφορετικό πλήθος σωματιδίων κάθε φορά. Για το πλήθος σωματιδίων γίνονται δοκιμές για 20, 40, 60, 80 και 100 σωματίδια. Ο συνδυασμός πλήθους νευρώνων κρυφού στρώματος και πλήθους σωματιδίων με τη μικρότερη διάμεσο CWC επιλέγεται ως ο βέλτιστος συνδυασμός. Στο Σχήμα 5.9 παρουσιάζονται τα αποτελέσματα για τις πέντε διασταυρωμένες επικυρώσεις 5 τμημάτων. Η ελάχιστη τιμή της διαμέσου του CWC παρατηρείται για 80 σωματίδια, 15 νευρώνες κρυφού στρώματος και ισούται με $CWC = 0.5445$.

Σε ότι αφορά το πλήθος των επαναλήψεων του αλγορίθμου βελτιστοποίησης, πειραματικά προκύπτει ότι τα σωματίδια συγκλίνουν ικανοποιητικά στις 100 επαναλήψεις. Κάτω από τις 100 επαναλήψεις, τα σωματίδια δεν έχουν πάντα αρκετό χρόνο για να συγκλίνουν, ενώ πάνω από τις 100 επαναλήψεις η διαφορά στα αποτελέσματα είναι αμελητέα. Επομένως επιλέγονται 100 επαναλήψεις, ώστε να εξασφαλίζεται η σύγκλιση των σωματιδίων κρατώντας ταυτόχρονα τον χρόνο εκπαίδευσης του τεχνητού νευρωνικού δικτύου σε χαμηλά επίπεδα. Οι τελικές τιμές των παραμέτρων του μοντέλου LUBE-PSO-CWC παρουσιάζονται συγκεντρωτικά στον Πίνακα 5.3.

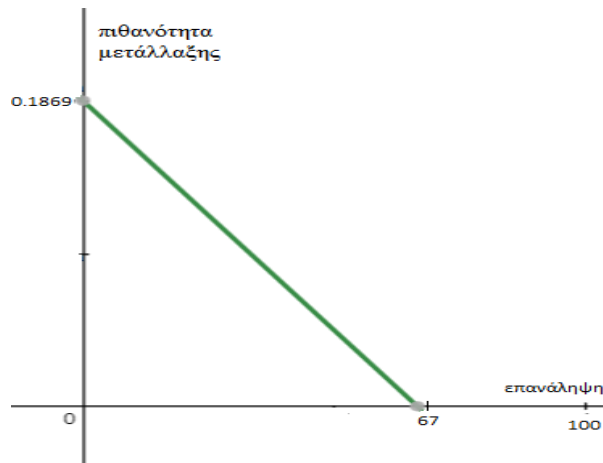
**Σχήμα 5.8:** Ο διαχωρισμός των δεδομένων εκπαίδευσης κατά τη διασταυρωμένη επικύρωση 5 τμημάτων

Πίνακας 5.3: Οι βέλτιστες τιμές των παραμέτρων του μοντέλου LUBE-PSO-CWC.

Παράμετρος	Βέλτιστη τιμή
Συνάρτηση ενεργοποίησης κρυφού στρώματος	Διπολική σιγμοειδής
Συνάρτηση ενεργοποίησης στρώματος εξόδου	Σιγμοειδής
Πλήθος νευρώνων στρώματος εισόδου (N_i)	4
Πλήθος νευρώνων κρυφού στρώματος (N_h)	15
Ονομαστικό επίπεδο εμπιστοσύνης (μ)	0.9
h	80
Πλήθος σωματιδίων ($n_{particles}$)	80
w	Γραμμική μείωση, $w_{max} = 0.7$, $w_{min} = 0.4$
c_1	1.2
c_2	1.3
Πλήθος επαναλήψεων (n_{iter})	100
x_{max}	4
v_{max}	1



Σχήμα 5.9: Αποτελέσματα διασταυρωμένης επικύρωσης 5 τμημάτων για πέντε διαφορετικά πλήθη σωματιδίων του μοντέλου LUBE-PSO-CWC.



Σχήμα 5.10: Η γραφική αναπαράσταση της πιθανότητας μετάλλαξης κάθε διάστασης της θέσης ενός σωματιδίου του μοντέλου LUBE-PSO-CWC σε συνάρτηση με την τρέχουσα επανάληψη.

Αντικαθιστώντας τις τιμές για το πλήθος των νευρώνων του στρώματος εισόδου N_i και το πλήθος των νευρώνων του κρυφού στρώματος N_h στη σχέση (5.6) για την παράμετρο β της αρχικοποίησης Nguyen-Widrow, προκύπτει $\beta \approx 1.968$. Από τη σχέση (5.9) προκύπτει ότι οι διαστάσεις του μοντέλου είναι $n_{dim} = 107$.

Η γραφική αναπαράσταση της μεταβολής της πιθανότητας μετάλλαξης σε συνάρτηση με την τρέχουσα επανάληψη του αλγορίθμου φαίνεται στο Σχήμα 5.10. Η πιθανότητα μετάλλαξης μηδενίζεται μετά την 66^η επανάληψη του αλγορίθμου. Η αρχική τιμή της πιθανότητας μετάλλαξης ισούται με $m = 0.1869$ (σχέση 5.12). Κάθε διάσταση της θέσης ενός σωματιδίου έχει δηλαδή πιθανότητα ίση με 18.69% να υποστεί μετάλλαξη κατά τη διάρκεια της πρώτης επανάληψης.

5.3.2 Βελτιστοποίηση παραμέτρων LUBE-MOPSO

Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση του μοντέλου LUBE-MOPSO είναι τα ίδια με αυτά που χρησιμοποιούνται για την εκπαίδευση του μοντέλου LUBE-PSO-CWC. Το τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE-MOPSO έχει ένα κρυφό στρώμα, δύο νευρώνες στο στρώμα εξόδου και τέσσερις νευρώνες στο στρώμα εισόδου, όπως ακριβώς το τεχνητό νευρωνικό δίκτυο του μοντέλου LUBE-PSO-CWC. Επομένως, επιλέγονται οι ίδιες συναρτήσεις ενεργοποίησης με αυτές του μοντέλου LUBE-PSO-CWC, δηλαδή η διπολική σιγμοειδής (σχέση (2.4)) για το κρυφό στρώμα και η σιγμοειδής (σχέση (2.3)) για το στρώμα εξόδου.

Ο καθορισμός των παραμέτρων w, c_1, c_2 της εξίσωσης (2.15) της πολυκριτηριακής βελτιστοποίησης σμήνους σωματιδίων για το μοντέλο LUBE-MOPSO γίνεται με τον ίδιο τρόπο που γίνεται για τη βελτιστοποίηση σμήνους σωματιδίων του μοντέλου LUBE-PSO-CWC (Ενότητα 5.3.1). Επειδή το μοντέλο LUBE-MOPSO παράγει πολλαπλές λύσεις, είναι αναγκαίο να επιλέγεται μία λύση κάθε φορά, με βάση κάποιο κριτήριο, ώστε με βάση την

Πίνακας 5.4: Καθορισμός βέλτιστου βάρους αδράνειας w για το μοντέλο LUBE–MOPSO.

w	Μέση τιμή PICP	Μέση τιμή PINAW	Μέση τιμή CWC
Τυχαίο [0.7, 0.9]	0.9043	0.5226	0.5226
Τυχαίο [0.4, 0.7]	0.9018	0.5206	0.5206
Γραμμική μείωση [0.7, 0.9]	0.9093	0.5325	0.5325
Γραμμική μείωση [0.4, 0.7]	0.9074	0.5233	0.5233

αξιολόγηση της λύσης αυτής να γίνεται η βελτιστοποίηση. Το κριτήριο που επιλέγεται είναι η τιμή του CWC. Εάν μια λύση έχει μικρό CWC, έχει PICP μεγαλύτερο του ονομαστικού επιπέδου εμπιστοσύνης. Από εκεί και πέρα, όσο μικρότερο είναι το CWC, τόσο μικρότερο είναι και το μέσο εύρος των διαστημάτων πρόβλεψης. Επομένως, κατά το τέλος της αξιολόγησης των λύσεων που παράχθηκαν κατά την εκπαίδευση του μοντέλου LUBE–MOPSO, επιλέγεται κάθε φορά η λύση με το μικρότερο CWC για σύγκριση, ώστε τα διαστήματα πρόβλεψης να έχουν όσο το δυνατόν μικρότερο εύρος ικανοποιώντας ταυτόχρονα το ονομαστικό επίπεδο εμπιστοσύνης. Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση του τεχνητού νευρωνικού δικτύου αφορούν τον χειμώνα του 2013 (Δεκέμβριος του 2012 έως Φεβρουάριο του 2013) της 1^{ης} ζώνης. Για κάθε συνδυασμό γίνεται τρεις φορές εκπαίδευση και αξιολόγηση και υπολογίζεται η μέση τιμή των μετρητικών αξιολόγησης. Ως δεδομένα εκπαίδευσης επιλέγονται τα δεδομένα που αφορούν τους δύο πρώτους μήνες, ενώ για την αξιολόγηση χρησιμοποιούνται τα δεδομένα του τρίτου μήνα. Κριτήριο για την επιλογή του βέλτιστου συνδυασμού αποτελεί η μέση τιμή του CWC, καθώς εμπεριέχει πληροφορία ταυτόχρονα και για τον βαθμό κάλυψης και για το μέσο εύρος των διαστημάτων πρόβλεψης.

Οι δοκιμές για τον καθορισμό του w γίνονται για $c_1 = 1.22, c_2 = 1.47$ [27], 80 σωματίδια, 13 νευρώνες κρυφού στρώματος και 100 επαναλήψεις. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 5.4. Η ελάχιστη μέση τιμή του CWC παρατηρείται για παραγωγή τυχαίας τιμής του w σε κάθε επανάληψη στο διάστημα [0.4, 0.7].

Οι δοκιμές για τον καθορισμό των c_1 και c_2 γίνονται για παραγωγή τυχαίας τιμής του w σε κάθε επανάληψη στο διάστημα [0.4, 0.7], 80 σωματίδια, 13 νευρώνες κρυφού στρώματος και 100 επαναλήψεις. Στον Πίνακα 5.5 παρουσιάζεται μέρος των αποτελεσμάτων που περιλαμβάνει και τον βέλτιστο συνδυασμό c_1 και c_2 . Ο βέλτιστος συνδυασμός προκύπτει για $c_1 = 1.5$ και $c_2 = 1.6$.

Το βέλτιστο πλήθος των νευρώνων του κρυφού στρώματος καθορίζεται μέσω της μεθόδου διασταυρωμένης επικύρωσης K τμημάτων (Ενότητα 5.3.1). Συγκεκριμένα επιλέγεται διασταυρωμένη επικύρωση 5 τμημάτων (Σχήμα 5.8). Το πλήθος των νευρώνων του κρυφού στρώματος μεταβάλλεται από 1 έως 20 με μοναδιαίο βήμα. Η διαδικασία επαναλαμβάνεται πέντε φορές, για 20, 40, 60, 80 και 100 σωματίδια. Ως σύνολο δεδομένων εκπαίδευσης χρησιμοποιούνται τα δεδομένα που αφορούν τους δύο πρώτους μήνες του χειμώνα του 2013 της 1^{ης} ζώνης. Τα δεδομένα αυτά αποτελούνται από 1488 ωριαίες παρατηρήσεις. Επομένως το σύνολο των δεδομένων εκπαίδευσης σε κάθε επανάληψη της διασταυρωμένης επικύρωσης 5 τμημάτων αποτελείται από 1190 παρατηρήσεις, ενώ το σύνολο των δεδομένων αξιολόγησης αποτελείται από τις υπόλοιπες 298 ωριαίες παρατηρήσεις. Οι δοκιμές γίνονται

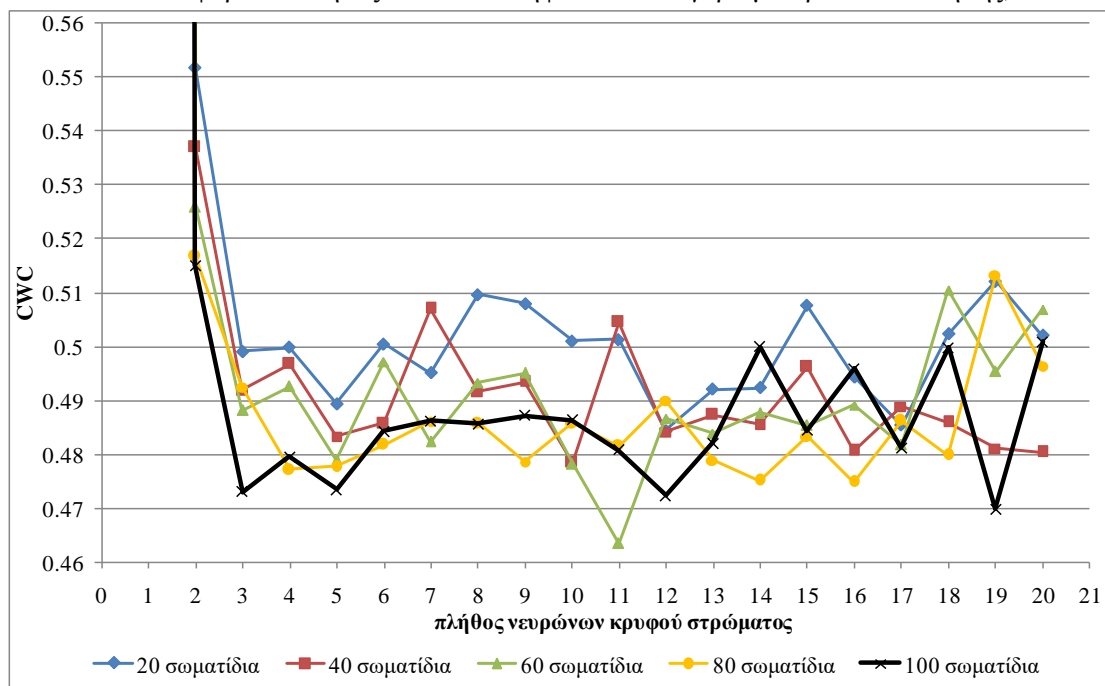
Πίνακας 5.5: Ενδεικτικά αποτελέσματα της βελτιστοποίησης των παραμέτρων c_1 και c_2 για το μοντέλο LUBE-MOPSO.

c_1	c_2	Μέση τιμή PICIP	Μέση τιμή PINAW	Μέση τιμή CWC
1.5	1.3	0.9012	0.5191	0.5191
1.5	1.4	0.9031	0.5179	0.5179
1.5	1.5	0.9025	0.5130	0.5130
1.5	1.6	0.9024	0.5106	0.5106
1.5	1.7	0.9012	0.5127	0.5127
1.5	1.8	0.9025	0.519	0.519
1.5	1.9	0.9037	0.5167	0.5167

για παραγωγή τυχαίας τιμής του w σε κάθε επανάληψη στο διάστημα $[0.4, 0.7]$, $c_1 = 1.5$, $c_2 = 1.6$ και 100 επαναλήψεις του αλγορίθμου βελτιστοποίησης. Ο συνδυασμός πλήθους νευρώνων κρυφού στρώματος και πλήθους σωματιδίων με τη μικρότερη διάμεσο CWC επιλέγεται ως ο βέλτιστος συνδυασμός. Στο Σχήμα 5.11 παρουσιάζονται τα αποτελέσματα για τις πέντε διασταυρωμένες επικυρώσεις 5 τμημάτων. Η ελάχιστη τιμή της διαμέσου του CWC παρατηρείται για 60 σωματίδια, 11 νευρώνες κρυφού στρώματος και ισούται με $CWC = 0.4634$.

Το μέγεθος του αρχείου αποθήκευσης επιλέγεται να είναι ίσο με το βέλτιστο πλήθος σωματιδίων. Επομένως, το αρχείο αποθήκευσης έχει χωρητικότητα 60 λύσεων. Κάθε διάσταση του προσαρμοστικού πλέγματος περιλαμβάνει $\frac{n_{particles}}{4} = 15$ υπερκύβους (Ενότητα 5.2.1). Επομένως το προσαρμοστικό πλέγμα αποτελείται από $15 \cdot 15 = 225$ υπερκύβους.

Σε ότι αφορά το πλήθος των επαναλήψεων του αλγορίθμου βελτιστοποίησης,



Σχήμα 5.11: Αποτελέσματα διασταυρωμένης επικύρωσης 5 τμημάτων για πέντε διαφορετικά πλήθη σωματιδίων του μοντέλου LUBE-MOPSO.

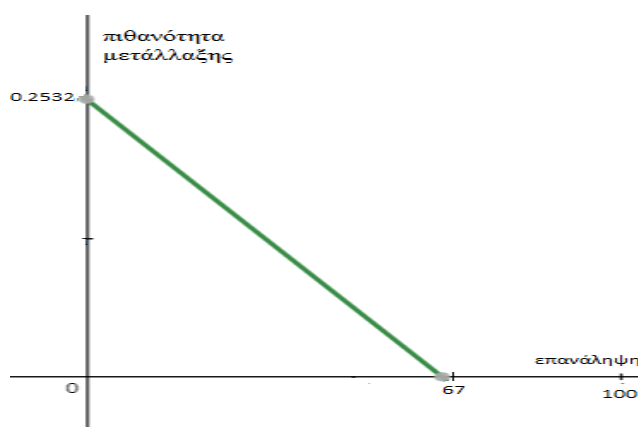
Πίνακας 5.6: Οι τελικές τιμές των παραμέτρων του μοντέλου LUBE–MOPSO.

Παράμετρος	Βέλτιστη τιμή
Συνάρτηση ενεργοποίησης κρυφού στρώματος	Διπολική σιγμοειδής
Συνάρτηση ενεργοποίησης στρώματος εξόδου	Σιγμοειδής
Πλήθος νευρώνων στρώματος εισόδου (N_i)	4
Πλήθος νευρώνων κρυφού στρώματος (N_h)	11
Ονομαστικό επίπεδο εμπιστοσύνης (μ)	0.9
Μέγεθος αρχείου αποθήκευσης	60
x_{max}	4
Πλήθος σωματιδίων ($n_{particles}$)	60
w	Τυχαία τιμή σε κάθε επανάληψη, $w_{max} = 0.7, w_{min} = 0.4$
c_1	1.5
c_2	1.6
Πλήθος επαναλήψεων (n_{iter})	100
Πλήθος υπερκύβων του προσαρμοστικού πλέγματος	225
v_{max}	1

επιλέγονται 100 επαναλήψεις, ώστε να εξασφαλίζεται η σύγκλιση των σωματιδίων κρατώντας ταυτόχρονα τον χρόνο εκπαίδευσης του τεχνητού νευρωνικού δικτύου σε χαμηλά επίπεδα. Οι τελικές τιμές των παραμέτρων του μοντέλου LUBE–MOPSO παρουσιάζονται συγκεντρωτικά στον Πίνακα 5.6.

Αντικαθιστώντας τις τιμές για το πλήθος νευρώνων του στρώματος εισόδου N_i και το πλήθος νευρώνων του κρυφού στρώματος N_h στη σχέση (5.6) για την παράμετρο β της αρχικοποίησης Nguyen – Widrow, προκύπτει $\beta \approx 1.8212$. Από τη σχέση (5.9) προκύπτει ότι οι διαστάσεις του μοντέλου είναι $n_{dim} = 79$.

Η γραφική αναπαράσταση της μεταβολής της πιθανότητας μετάλλαξης σε συνάρτηση



Σχήμα 5.12: Η γραφική αναπαράσταση της πιθανότητας μετάλλαξης κάθε διάστασης της θέσης ενός σωματιδίου του μοντέλου LUBE–MOPSO σε συνάρτηση με την τρέχουσα επανάληψη.

Πίνακας 5.7: Σύγκριση αποτελεσμάτων πριν και μετά τη βελτιστοποίηση παραμέτρων

Μοντέλο	Διάμεσος CWC
LUBE-PSO-CWC (πριν)	0.5709
LUBE-PSO-CWC (μετά)	0.5445
Διαφορά LUBE-PSO-CWC	-4.62%
LUBE-MOPSO (πριν)	0.5325
LUBE-MOPSO (μετά)	0.4634
Διαφορά LUBE-MOPSO	-12.98%

με την τρέχουσα επανάληψη του αλγορίθμου φαίνεται στο Σχήμα 5.12. Η πιθανότητα μετάλλαξης μηδενίζεται μετά την 66^η επανάληψη του αλγορίθμου. Η αρχική τιμή της πιθανότητας μετάλλαξης ισούται με $m = 0.2532$ (σχέση 5.12). Κάθε διάσταση της θέσης ενός σωματιδίου έχει δηλαδή πιθανότητα ίση με 25.32% να υποστεί μετάλλαξη κατά τη διάρκεια της πρώτης επανάληψης.

5.3.3 Βελτίωση αποτελεσμάτων

Από τους Πίνακες 5.3 και 5.6 φαίνεται ότι τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO έχουν μεγάλο πλήθος παραμέτρων. Το εύρος τιμών κάποιων παραμέτρων είναι επίσης μεγάλο, ενώ οι περισσότερες από αυτές είναι δυνατόν να μεταβάλλονται σε συνάρτηση με τον χρόνο. Επομένως, είναι εξαιρετικά δύσκολο και χρονοβόρο να δοκιμαστούν όλοι οι δυνατοί συνδυασμοί παραμέτρων των δύο μοντέλων.

Ο καθορισμός κάποιων παραμέτρων γίνεται πριν ξεκινήσει η διαδικασία βελτιστοποίησης των υπόλοιπων παραμέτρων. Έτσι, από το μεγάλο πλήθος διαθέσιμων συναρτήσεων ενεργοποίησης δοκιμάζονται μόνο οι ευρέως χρησιμοποιούμενες. Τα αποτελέσματα κρίνονται περισσότερο ποιοτικά παρά ποσοτικά. Τα όρια της θέσης των σωματιδίων επιλέγονται με βάση τον τρόπο αρχικοποίησής τους. Τα όρια της ταχύτητας των σωματιδίων ετεροκαθορίζονται από τα όρια της θέσης τους. Όμοια, το μέγεθος του αρχείου αποθήκευσης και του πλήθους των υπερκύβων επιλέγονται σε συνάρτηση με το πλήθος των σωματιδίων.

Για τη βελτιστοποίηση των υπόλοιπων παραμέτρων επιλέγονται κάποιες τιμές ως βάση, οι οποίες μένουν σταθερές κατά τη διάρκεια της βελτιστοποίησης μέχρι να αντικατασταθούν από τις βέλτιστες τιμές τους. Οι αρχικές τιμές για τα w , c_1 και c_2 είναι ίδιες με αυτές που προτείνονται στο [27]. Το αρχικό πλήθος των νευρώνων του κρυφού στρώματος, το αρχικό πλήθος σωματιδίων και το πλήθος επαναλήψεων επιλέγονται αυθαίρετα, μέσω δοκιμών και παρατήρησης των αποτελεσμάτων. Οι τιμές για τις οποίες δοκιμάζεται κάθε παράμετρος καθώς και ο τρόπος της χρονικής μεταβολής της, εάν αυτή είναι εφικτή, καθορίζεται πειραματικά μέσω δοκιμών, λαμβάνοντας υπόψη τη θεωρία των τεχνητών νευρωνικών δικτύων και της βελτιστοποίησης σμήνους σωματιδίων.

Παρόλο που η βελτιστοποίηση παραμέτρων της παρούσας διπλωματικής εργασίας παρουσιάζει ατέλειες, η βελτίωση που επιφέρει στην απόδοση των μοντέλων είναι

σημαντική. Στον Πίνακα 5.7 παρουσιάζονται συγκριτικά τα αποτελέσματα για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO πριν και μετά τη βελτιστοποίηση των αποτελεσμάτων. Τα αποτελέσματα των δύο μοντέλων πριν τη βελτιστοποίηση παράχθηκαν για γραμμική μείωση του w στο διάστημα $[0.7, 0.9]$, $c_1 = 1.22$, $c_2 = 1.47$, 80 σωματίδια, 13 νευρώνες κρυφού στρώματος και 100 επαναλήψεις του αλγορίθμου βελτιστοποίησης. Το μοντέλο LUBE-PSO-CWC παρουσιάζει μείωση (βελτίωση) της διαμέσου του CWC κατά 4.62% μετά τη βελτιστοποίηση των παραμέτρων του, ενώ το μοντέλο LUBE-MOPSO παρουσιάζει μείωση (βελτίωση) της διαμέσου του CWC κατά 12.98%.

ΑΠΟΤΕΛΕΣΜΑΤΑ

6.1 ΜΕΘΟΔΟΛΟΓΙΑ

Η μεθοδολογία που ακολουθείται για την εκπαίδευση και την αξιολόγηση των μοντέλων LUBE-PSO-CWC [27] και LUBE-MOPSO είναι κοινή. Τα δύο μοντέλα τροφοδοτούνται με τα ίδια δεδομένα έτσι ώστε να μπορεί να συγκριθεί, με αξιόπιστο τρόπο, η αποδοτικότητά τους. Η επεξεργασία των δεδομένων (Ενότητα 4.3) καθώς και ο διαχωρισμός τους σε δεδομένα εκπαίδευσης και δεδομένα αξιολόγησης γίνεται ακολουθώντας την ίδια ακριβώς μέθοδο. Αφού ολοκληρωθεί η διαδικασία της εκπαίδευσης των μοντέλων, τροφοδοτούνται με το σύνολο δεδομένων αξιολόγησης ώστε να αξιολογηθεί η απόδοσή τους (Κεφάλαιο 5). Για να εξασφαλιστεί η αμεροληψία σε ότι αφορά τη σχετική απόδοση των μοντέλων, η διαδικασία της εκπαίδευσης και της αξιολόγησης επαναλαμβάνονται αρκετές φορές. Τα αποτελέσματα των μοντέλων LUBE-PSO-CWC και LUBE-MOPSO συγκρίνονται επίσης με τα αποτελέσματα άλλων μοντέλων που υπάρχουν στη διαθέσιμη βιβλιογραφία.

6.1.1 Διαχωρισμός δεδομένων

Τα δεδομένα που χρησιμοποιούνται προέρχονται από τον διαγωνισμό GEFCom 2014 [5]. Στην παρούσα διπλωματική εργασία επιλέγονται τα δεδομένα που αφορούν την περίοδο από την 1^η Ιουνίου 2012 και ώρα 00:00 έως την 31^η Μαΐου 2013 και ώρα 23:00. Τα δεδομένα αυτά επιλέγονται και για τις δέκα διαθέσιμες ζώνες. Για να μπορέσει να γίνει εποχική ανάλυση, τα δεδομένα κάθε ζώνης χωρίζονται σε τέσσερις ομάδες που αντιστοιχούν στις τέσσερις εποχές:

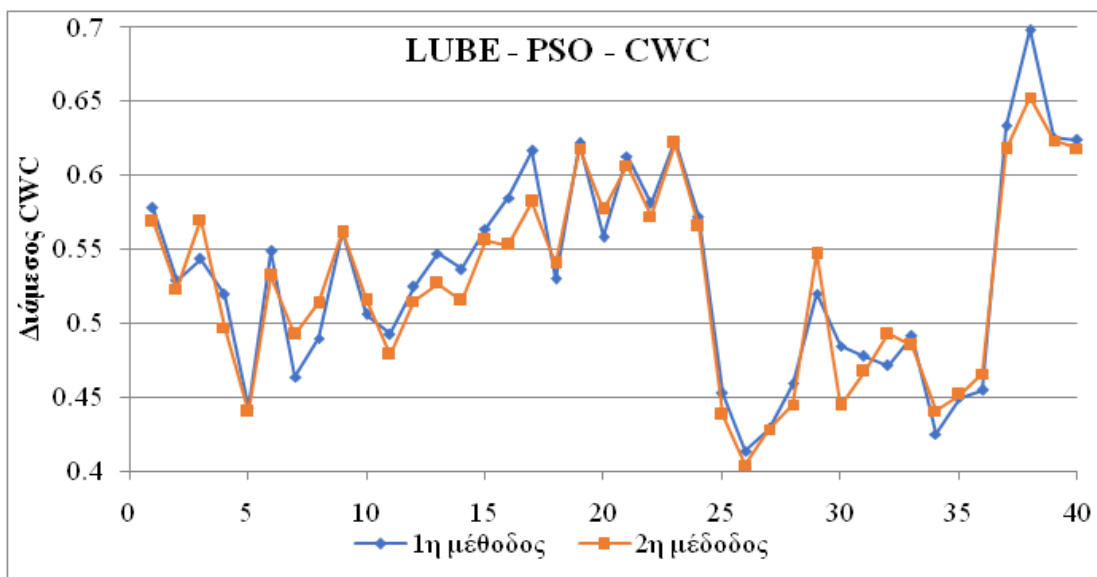
1. Καλοκαίρι (1^η Ιουνίου 2012 έως και 31^η Αυγούστου 2012): 2208 ωριαία σύνολα παρατηρήσεων.
2. Φθινόπωρο (1^η Σεπτεμβρίου 2012 έως και 30^η Νοεμβρίου 2012): 2184 ωριαία σύνολα παρατηρήσεων.
3. Χειμώνας (1^η Δεκεμβρίου 2012 έως και 28^η Φεβρουαρίου 2013): 2160 ωριαία σύνολα παρατηρήσεων.
4. Άνοιξη (1^η Μαρτίου 2013 έως και 31^η Μαΐου 2013): 2208 ωριαία σύνολα παρατηρήσεων.

Επομένως, διαμορφώνονται 40 σύνολα δεδομένων, ένα για κάθε εποχή σε κάθε ζώνη. Τα δεδομένα εισόδου μεταφέρονται, από το καρτεσιανό, στο πολικό σύστημα συντεταγμένων και κανονικοποιούνται με τη μέθοδο μεγίστου-ελαχίστου στο διάστημα $[-1, 1]$ (Ενότητα 4.3).

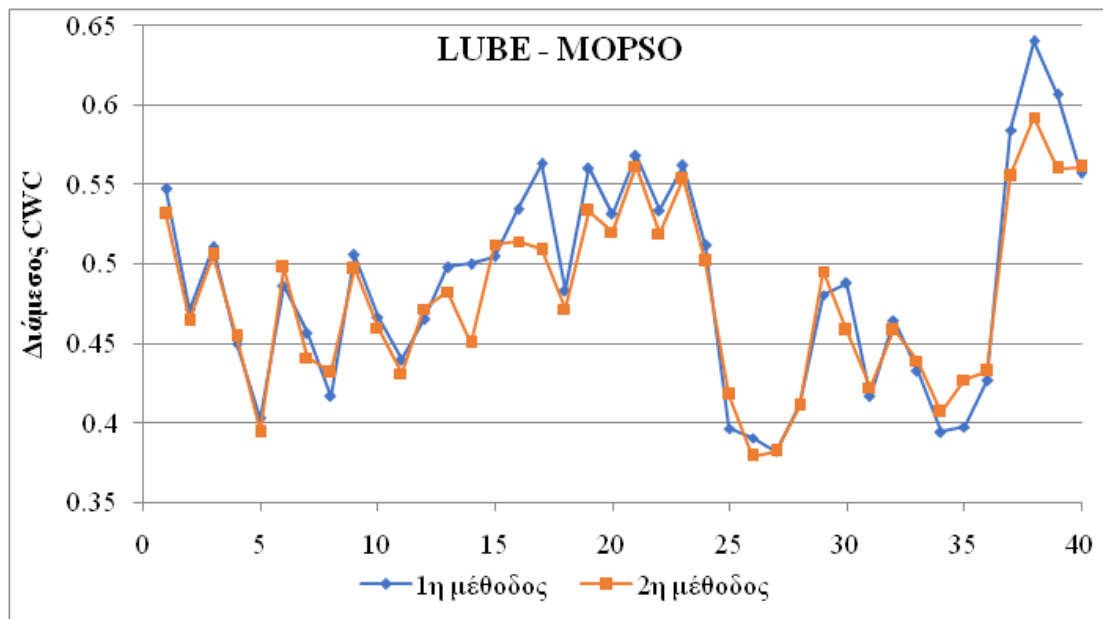
Για τον διαχωρισμό των δεδομένων σε δεδομένα εκπαίδευσης και δεδομένα αξιολόγησης, δοκιμάζονται δύο μέθοδοι. Στην πρώτη μέθοδο, ως δεδομένα εισόδου ορίζονται τα δεδομένα των δύο πρώτων μηνών κάθε εποχής, ενώ τα δεδομένα του τρίτου μήνα ορίζονται ως δεδομένα αξιολόγησης. Στη δεύτερη μέθοδο, τα δεδομένα χωρίζονται σε ημέρες, δηλαδή σε υποσύνολα 24 ωριαίων παρατηρήσεων. Κάθε τρία υποσύνολα ορίζονται ως δεδομένα εκπαίδευσης ενώ τα δεδομένα του τέταρτου υποσυνόλου ορίζονται ως δεδομένα αξιολόγησης. Η διαδικασία διαχωρισμού των δεδομένων της δεύτερης μεθόδου περιγράφεται από την εξίσωση (6.1):

$$\begin{aligned} \text{δεδομένα εκπαίδευσης: } & (4k + 1)^n, (4k + 2)^n, (4k + 3)^n \text{ παρατήρηση} \\ \text{δεδομένα αξιολόγησης: } & (4k + 4)^n \text{ παρατήρηση} \\ k = & 0, 1, \dots, \frac{n_{data}}{4} - 4 \end{aligned} \quad (6.1)$$

όπου n_{data} το πλήθος των ωριαίων παρατηρήσεων κάθε εποχής. Για τη σύγκριση των δύο μεθόδων διαχωρισμού των δεδομένων, τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO εκπαιδεύονται και αξιολογούνται 5 φορές με καθένα από τα 40 σύνολα δεδομένων. Τα μετρητικά αξιολόγησης που χρησιμοποιούνται είναι το PICP (σχέση (3.4)), το PINAW (σχέση (3.5)) και το CWC (σχέση (5.1)). Για κάθε σύνολο δεδομένων, υπολογίζεται η διάμεσος κάθε μετρητικού αξιολόγησης για τις πέντε επαναλήψεις της διαδικασίας της εκπαίδευσης και της αξιολόγησης. Στο Σχήμα 6.1 παρουσιάζονται τα αποτελέσματα του μοντέλου LUBE-PSO-CWC για τις δύο μεθόδους. Στο Σχήμα 6.2 παρουσιάζονται τα αντίστοιχα αποτελέσματα για το μοντέλο LUBE-MOPSO. Ως βασικό κριτήριο σύγκρισης επιλέγεται η διάμεσος του CWC, καθώς το CWC εμπεριέχει πληροφορία ταυτόχρονα για τον δείκτη κάλυψης και για το μέσο εύρος των διαστημάτων πρόβλεψης. Το μοντέλο LUBE-MOPSO κατά την ολοκλήρωση της εκπαίδευσης παράγει πολλαπλές λύσεις. Η λύση που επιλέγεται σε κάθε επανάληψη της διαδικασίας είναι εκείνη με τη μικρότερη τιμή CWC,



Σχήμα 6.1: Η διάμεσος του CWC κάθε συνόλου δεδομένων του μοντέλου LUBE-PSO-CWC για τις δύο μεθόδους διαχωρισμού των δεδομένων.



Σχήμα 6.2: Η διάμεσος του CWC κάθε συνόλου δεδομένων του μοντέλου LUBE–MOPSO για τις δύο μεθόδους διαχωρισμού των δεδομένων.

καθώς εξασφαλίζει ότι το PICP είναι μεγαλύτερο ή ίσο του ονομαστικού επιπέδου εμπιστοσύνης, με το μικρότερο δυνατό μέσο εύρος διαστημάτων πρόβλεψης.

Το μοντέλο LUBE–PSO–CWC παρουσιάζει καλύτερα αποτελέσματα, δηλαδή μικρότερη διάμεσο CWC, με τη δεύτερη μέθοδο διαχωρισμού δεδομένων σε 29 από τα 40 σύνολα δεδομένων. Το μοντέλο LUBE–MOPSO παρουσιάζει καλύτερα αποτελέσματα με τη δεύτερη μέθοδο σε 26 από τα 40 σύνολα δεδομένων. Στη συνέχεια υπολογίζονται οι διάμεσοι των 40 τιμών των διαμέσων των PICP, PINAW και CWC για τις δύο μεθόδους διαχωρισμού των δεδομένων. Τα αποτελέσματα και για τα δύο μοντέλα παρουσιάζονται στον Πίνακα 6.1.

Από τα Σχήματα 6.1 και 6.2 και τον Πίνακα 6.1 φαίνεται ότι η μέθοδος διαχωρισμού των δεδομένων που περιγράφεται από τη σχέση (6.1) βελτιώνει τα διαστήματα πρόβλεψης και των δύο μοντέλων σε ότι αφορά τον δείκτη κάλυψής τους και το μέσο εύρος τους. Συγκεκριμένα, για το μοντέλο LUBE–PSO–CWC η διάμεσος του PICP βελτιώνεται κατά 0.35% ενώ η διάμεσος του PINAW βελτιώνεται κατά 0.82%. Για το μοντέλο LUBE–MOPSO η διάμεσος του PICP βελτιώνεται κατά 0.27% ενώ η διάμεσος του PINAW βελτιώνεται κατά 2.7%. Επομένως, για τον διαχωρισμό των δεδομένων σε δεδομένα εκπαίδευσης και δεδομένα αξιολόγησης επιλέγεται η μέθοδος που περιγράφεται από τη σχέση (6.1). Η βελτίωση των αποτελεσμάτων με τη δεύτερη μέθοδο διαχωρισμού ήταν αναμενόμενη, καθώς στα δεδομένα εκπαίδευσης περιλαμβάνονται παρατηρήσεις από όλους τους μήνες κάθε συνόλου, αυξάνοντας την ικανότητα γενίκευσης των μοντέλων. Αντίθετα, με την πρώτη μέθοδο διαχωρισμού τα μοντέλα καλούνται να δημιουργήσουν διαστήματα πρόβλεψης για έναν μήνα για τον οποίο δεν έλαβαν κανένα δεδομένο κατά τη διάρκεια της εκπαίδευσης. Η συσχέτιση μεταξύ των δεδομένων του τρίτου μήνα κάθε εποχής με τα δεδομένα των υπόλοιπων δύο μηνών είναι μικρή. Αντίθετα, όπως είναι λογικό, τα δεδομένα μιας ημέρας παρουσιάζουν μεγαλύτερη συσχέτιση με τα δεδομένα των τριών ημερών που προηγήθηκαν.

Πίνακας 6.1: Οι διάμεσοι των 40 τιμών των διαμέσων των PICP, PINAW και CWC για τις δύο μεθόδους διαχωρισμού των δεδομένων των μοντέλων LUBE–PSO–CWC και LUBE–MOPSO.

Μοντέλο	Μέθοδος Διαχωρισμού Δεδομένων	Διάμεσος PICP	Διάμεσος PINAW	Διάμεσος CWC
LUBE–PSO–CWC	1 ^η μέθοδος	0.936594	0.529172	0.529172
	2 ^η μέθοδος	0.939889	0.524825	0.524825
	Διαφορά (%)	0.35	-0.82	-0.82
LUBE–MOPSO	1 ^η μέθοδος	0.905556	0.484805	0.484805
	2 ^η μέθοδος	0.908017	0.471693	0.471693
	Διαφορά (%)	0.27	-2.7	-2.7

6.1.2 Μοντέλο BELM

Για να την πλήρη και αξιόπιστη αξιολόγηση της απόδοσης των μοντέλων LUBE–PSO–CWC και LUBE–MOPSO δεν είναι αρκετή η σύγκριση μεταξύ τους αλλά είναι ανάγκη να συγκριθούν με την απόδοση άλλων ήδη δοκιμασμένων μοντέλων πιθανοτικής πρόβλεψης που συναντώνται στη διαθέσιμη βιβλιογραφία. Στην παρούσα διπλωματική εργασία, το μοντέλο που επιλέγεται ως σημείο αναφοράς είναι το Bootstrap Extreme Learning Machine (BELM) [51].

Το μοντέλο BELM αποτελεί ένα μοντέλο πιθανοτικής πρόβλεψης. Σε αντίθεση με τα μοντέλα LUBE–PSO–CWC και LUBE–MOPSO όπου τα όρια των διαστημάτων πρόβλεψης παράγονται απευθείας από το τεχνητό νευρωνικό δίκτυο, το τεχνητό νευρωνικό δίκτυο του μοντέλου BELM παράγει σημειακές προβλέψεις και στη συνέχεια με την κατάλληλη στατιστική επεξεργασία παράγονται τα όρια των διαστημάτων πρόβλεψης.

Το τεχνητό νευρωνικό δίκτυο του μοντέλου BELM είναι ένα Extreme Learning Machine (ELM) [52]. Τα ELM είναι τεχνητά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης ενός κρυφού στρώματος. Η διαφορά τους με τα συμβατικά τεχνητά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης είναι στον τρόπο εκπαίδευσής τους. Η εκπαίδευση των ELM δεν αποτελεί επαναληπτική διαδικασία, αντίθετα έγκειται στον υπολογισμό ενός γενικευμένου αντίστροφου πίνακα Moore–Penrose και σε έναν πολλαπλασιασμό πινάκων. Από αυτήν την ιδιότητα προκύπτει και το βασικό πλεονέκτημα των ELM, που είναι η μεγάλη ταχύτητα εκπαίδευσής τους.

Αφού παραχθούν οι σημειακές προβλέψεις από το ELM, τα διαστήματα πρόβλεψης κατασκευάζονται κάνοντας χρήση της μεθόδου bootstrapping [53] και συγκεκριμένα της μεθόδου bootstrapping ζευγαριών (pairs bootstrapping). Η μέθοδος bootstrapping ζευγαριών βασίζεται στη δημιουργία ενός νέου συνόλου δεδομένων εισόδου–εξόδου κάνοντας δειγματοληψία με επανατοποθέτηση στα αρχικά δεδομένα χρησιμοποιώντας ομοιόμορφη κατανομή. Για B διαφοροποιημένα υποσύνολα εκπαίδευσης παράγονται B σύνολα σημειακών

προβλέψεων. Το κάτω και άνω όριο των διαστημάτων πρόβλεψης υπολογίζονται μέσω των εξισώσεων (6.2) και (6.3) [6]:

$$L(x_i) = \hat{y}(x_i) - z_{1-\alpha/2} \sqrt{\sigma_{\hat{y}}^2(x_i)} \quad (6.2)$$

$$U(x_i) = \hat{y}(x_i) + z_{1-\alpha/2} \sqrt{\sigma_{\hat{y}}^2(x_i)} \quad (6.3)$$

όπου $L(x_i)$ και $U(x_i)$ το κάτω και άνω όριο, αντίστοιχα, για το σύνολο δεδομένων εισόδου του δείγματος i , $\hat{y}(x_i)$ η μέση τιμή των παραγόμενων προβλέψεων για το δείγμα i , $\sqrt{\sigma_{\hat{y}}^2(x_i)}$ η τυπική απόκλιση των προβλέψεων για το δείγμα i και $z_{1-\alpha/2}$ η αθροιστική πιθανότητα της κατανομής της μεταβλητής z η οποία ακολουθεί τυποποιημένη κανονική κατανομή $N(0,1)$. Για το α ισχύει η σχέση $\alpha = 1 - \mu$ όπου μ το ονομαστικό επίπεδο εμπιστοσύνης.

6.1.3 Αξιολόγηση μοντέλων πιθανοτικής πρόβλεψης

Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγηση του μοντέλου BELM είναι ίδια με εκείνα που χρησιμοποιούνται για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO. Ο διαχωρισμός των δεδομένων γίνεται ακολουθώντας τη μέθοδο που περιγράφεται από τη σχέση (6.1). Κατά αυτόν τον τρόπο το μοντέλο BELM μπορεί να χρησιμοποιηθεί αξιόπιστα ως σημείο αναφοράς.

Για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO έχει γίνει ήδη βελτιστοποίηση παραμέτρων (Ενότητα 5.3). Για το μοντέλο BELM γίνεται βελτιστοποίηση παραμέτρων ξεχωριστά για καθένα από τα 40 διαφορετικά σύνολα δεδομένων. Οι πολλαπλές βελτιστοποιήσεις παραμέτρων είναι απαραίτητες καθώς το μοντέλο BELM παρουσιάζει μεγαλύτερη αστάθεια σε σχέση με τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO όταν το πλήθος των δεδομένων είναι σχετικά μικρό. Επομένως, ακόμα και αν το πλήθος και το είδος των δεδομένων παραμένει σταθερό, μικρές αλλαγές στις τιμές των δεδομένων μπορούν να οδηγήσουν σε μεγάλες αποκλίσεις σε ότι αφορά τις βέλτιστες παραμέτρους του μοντέλου BELM. Ως παράμετροι προς βελτιστοποίηση για το μοντέλο BELM επιλέγονται το πλήθος των νευρώνων του κρυφού στρώματος και το πλήθος B των διαφοροποιημένων υποσυνόλων εκπαίδευσης της μεθόδου bootstrapping. Για κάθε σύνολο δεδομένων δοκιμάζονται τιμές στο διάστημα $[80, 100]$ για το πλήθος των νευρώνων του κρυφού στρώματος και τιμές στο διάστημα $[300, 700]$ για το B . Τα διαστήματα αυτά προκύπτουν πειραματικά μέσα από δοκιμές. Επομένως, τα αποτελέσματα που παρουσιάζονται και αφορούν το μοντέλο BELM, είναι για διαφορετικά πλήθη νευρώνων κρυφού στρώματος και διαφορετικές τιμές του B , που προκύπτουν έπειτα από διαδικασία βελτιστοποίησης.

Και για τα τρία μοντέλα γίνονται πέντε επαναλήψεις της εκπαίδευσης και της αξιολόγησης για κάθε εποχή κάθε ζώνης. Σε ότι αφορά το μοντέλο LUBE-MOPSO που παράγει πολλαπλές λύσεις έπειτα από κάθε εκπαίδευση, επιλέγεται η λύση με τη μικρότερη τιμή CWC. Τα μετρητικά αξιολόγησης που επιλέγονται είναι το PICP, το PINAW και το CWC. Για τη στατιστική αξιολόγηση των πέντε επαναλήψεων για κάθε εποχή και κάθε ζώνη,

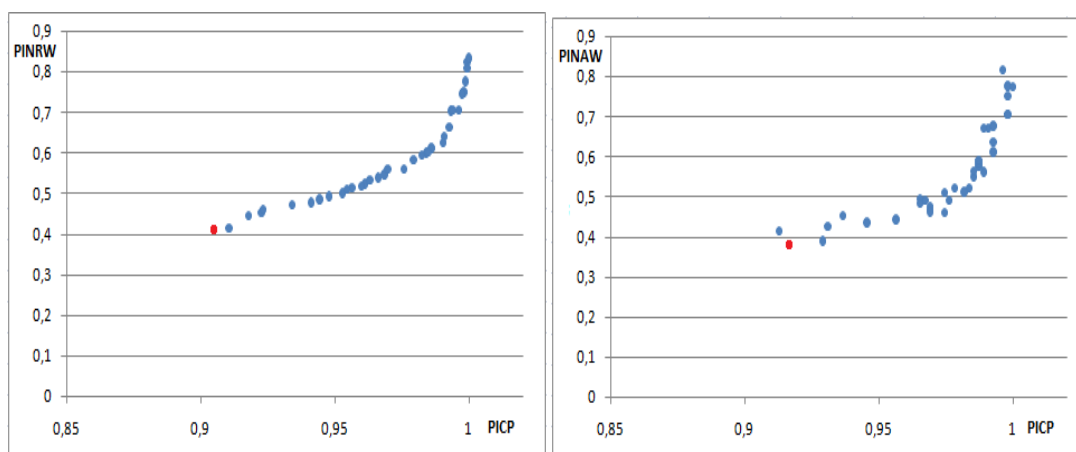
Πίνακας 6.2: Αποτελέσματα του μοντέλου LUBE–MOPSO για τα δεδομένα του καλοκαιριού της 2^{ης} ζώνης.

LUBE – MOPSO, 2 ^η Ζώνη – Καλοκαίρι			
Επανάληψη	PICP	PINAW	CWC
1	0.900362	0.38735	0.38735
2	0.903986	0.443597	0.443597
3	0.911232	0.401426	0.401426
4	0.907609	0.390717	0.390717
5	0.927536	0.394474	0.394474
Διάμεσος	0.907609	0.394474	0.394474

υπολογίζεται η διάμεσος των μετρητικών αξιολόγησης κάθε επανάληψης. Στο Σχήμα 6.3(α) φαίνονται ως παράδειγμα τα αποτελέσματα της πρώτης επανάληψης της εκπαίδευσης του μοντέλου LUBE–MOPSO για τα δεδομένα που αφορούν το καλοκαίρι της 2^{ης} ζώνης. Το πλήθος λύσεων του αρχείου αποθήκευσης κατά το τέλος της εκπαίδευσης ισούται με 36. Στο

Σχήμα 6.3(β) φαίνονται τα αποτελέσματα της πρώτης αξιολόγησης του μοντέλου LUBE–MOPSO για τα δεδομένα που αφορούν το καλοκαίρι της 2^{ης} ζώνης. Με κόκκινο χρώμα απεικονίζεται η λύση με την ελάχιστη τιμή CWC. Στον Πίνακα 6.2 φαίνονται τα αποτελέσματα της αξιολόγησης του μοντέλου LUBE–MOPSO έπειτα από τις πέντε επαναλήψεις της εκπαίδευσης. Τα αποτελέσματα αυτά αντιστοιχούν στις λύσεις με την ελάχιστη τιμή CWC σε κάθε επανάληψη.

6.2 ΣΥΓΚΡΙΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

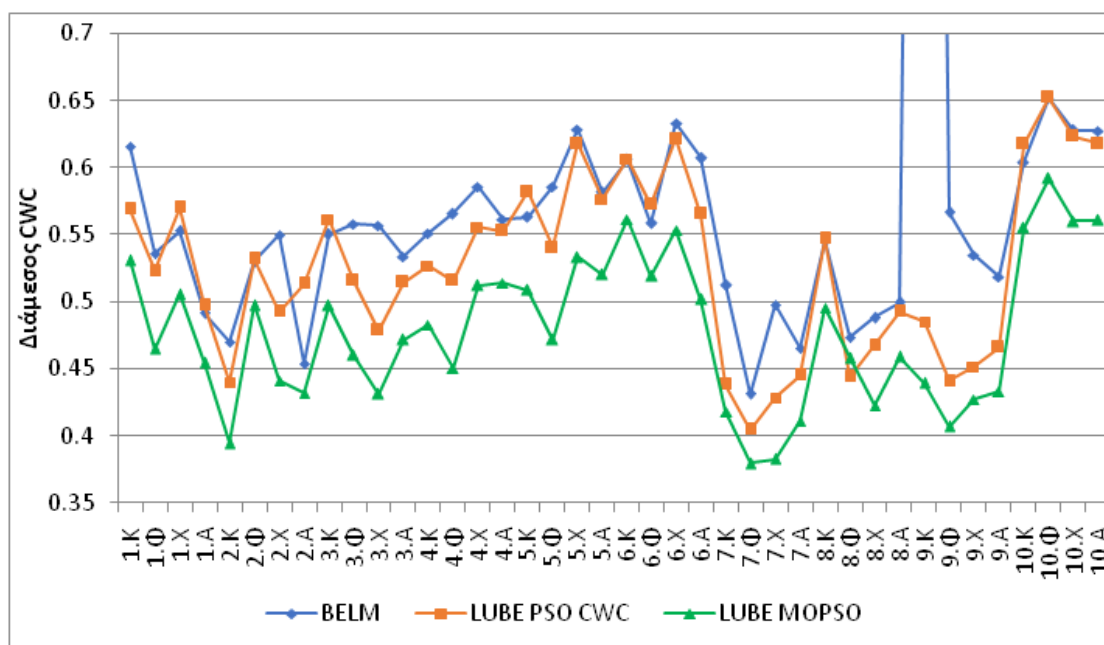


Σχήμα 6.3: (α) Οι λύσεις της πρώτης επανάληψης της εκπαίδευσης του μοντέλου LUBE–MOPSO για τα δεδομένα που αφορούν το καλοκαίρι της 2^{ης} ζώνης. (β) τα αποτελέσματα της αξιολόγησης των λύσεων της πρώτης επανάληψης της εκπαίδευσης του μοντέλου LUBE–MOPSO για τα δεδομένα που αφορούν το καλοκαίρι της 2^{ης} ζώνης.

6.2.1 Σύγκριση απόδοσης μοντέλων

Στο Σχήμα 6.4 φαίνονται τα αποτελέσματα των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM για κάθε εποχή κάθε ζώνης. Οι διάμεσες τιμές του CWC προκύπτουν από τις αξιολογήσεις των πέντε επαναλήψεων της εκπαίδευσης των μοντέλων για κάθε περίπτωση. Στους Πίνακες 6.3, 6.4 και 6.5 φαίνονται ενδεικτικά τα αποτελέσματα της αξιολόγησης των πέντε επαναλήψεων της εκπαίδευσης των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM, αντίστοιχα, για τα δεδομένα όλων των εποχών της 7^{ης} ζώνης. Στα Σχήματα 6.5, 6.6 και 6.7 φαίνονται τα παραγόμενα διαστήματα πρόβλεψης της πρώτης επανάληψης της εκπαίδευσης και αξιολόγησης των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM, αντίστοιχα, για τα δεδομένα του φθινοπώρου της 7^{ης} ζώνης. Με κόκκινο φαίνονται οι πραγματικές τιμές της παραγόμενης αιολικής ισχύος ενώ οι μπλε γραμμές αντιστοιχούν στα παραγόμενα διαστήματα πρόβλεψης. Στα Σχήματα 6.5, 6.6 και 6.7 τα πάνω διαστήματα πρόβλεψης αντιστοιχούν στις πρώτες 48 προβλέψεις, δηλαδή στις προβλέψεις με χρονικό ορίζοντα δύο ημερών, ενώ τα κάτω διαστήματα πρόβλεψης αντιστοιχούν στο σύνολο των δεδομένων αξιολόγησης, δηλαδή στις 546 παρατηρήσεις του συνόλου αξιολόγησης.

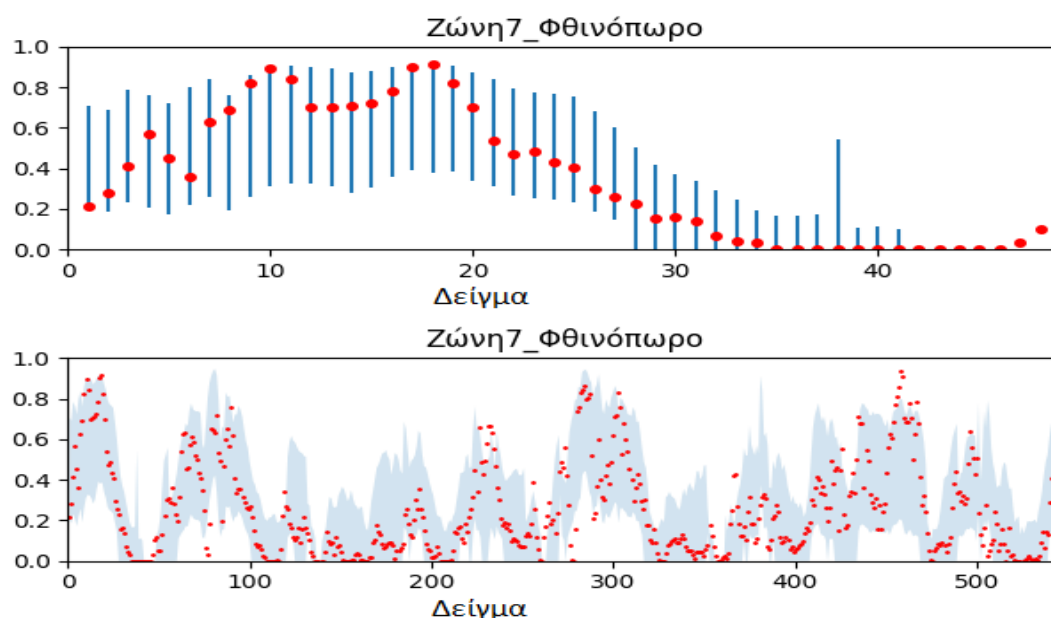
Παρατηρώντας το Σχήμα 6.4, φαίνεται ότι το μοντέλο LUBE-MOPSO υπερέχει σε ότι αφορά τη διάμεση τιμή του CWC στη συντριπτική πλειοψηφία των περιπτώσεων. Το μοντέλο LUBE-MOPSO παράγει δηλαδή τα διαστήματα πρόβλεψης με το μικρότερο μέσο εύρος, διατηρώντας παράλληλα τον δείκτη κάλυψης μεγαλύτερο ή ίσο από το ονομαστικό επίπεδο εμπιστοσύνης. Το μοντέλο LUBE-PSO-CWC υπερέχει έναντι του μοντέλου BELM



Σχήμα 6.4: Οι διάμεσες τιμές του CWC των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM για κάθε εποχή κάθε ζώνης. Στον οριζόντιο άξονα το νούμερο αντιστοιχεί στον αριθμό της ζώνης ενώ το γράμμα αντιστοιχεί στο αρχικό γράμμα κάθε εποχής (K: Καλοκαίρι, Φ: Φθινόπωρο, X: Χειμώνας, A: Άνοιξη).

Πίνακας 6.3: Τα αποτελέσματα της αξιολόγησης των πέντε επαναλήψεων της εκπαίδευσης του μοντέλου LUBE-PSO-CWC για τα δεδομένα όλων των εποχών της 7^{ης} ζώνης.

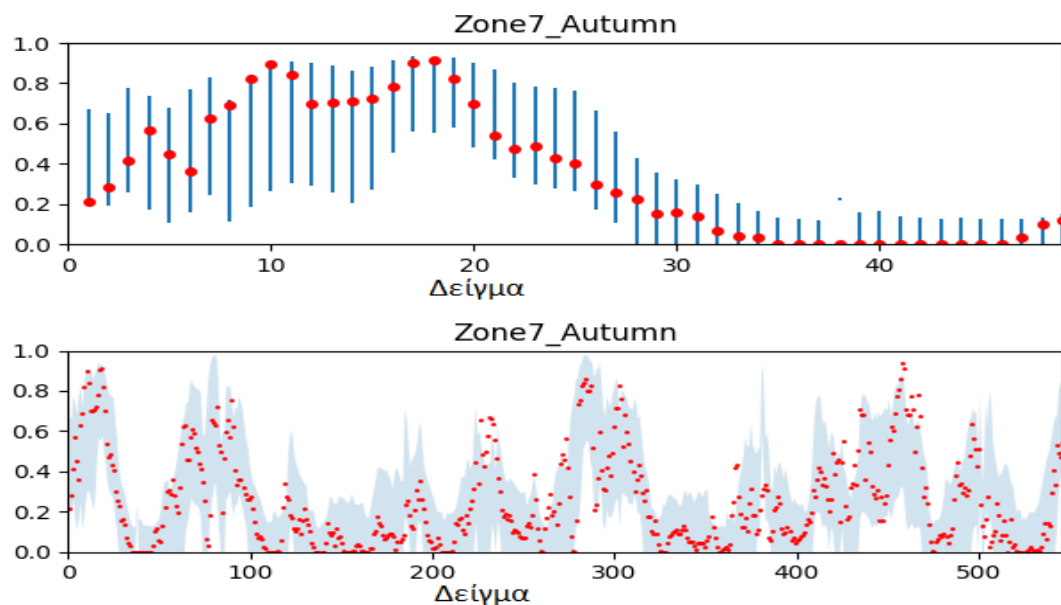
LUBE-PSO-CWC						
7 ^η Ζώνη – Καλοκαίρι						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.942029	0.951087	0.947464	0.923913	0.934783	0.942029
PINAW	0.437874	0.459689	0.426278	0.433482	0.444034	0.437874
CWC	0.437874	0.459689	0.426278	0.433482	0.444034	0.437874
7 ^η Ζώνη – Φθινόπωρο						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.945055	0.948718	0.935897	0.948718	0.93956	0.945055
PINAW	0.404243	0.418443	0.396628	0.399027	0.427136	0.404243
CWC	0.404243	0.418443	0.396628	0.399027	0.427136	0.404243
7 ^η Ζώνη – Χειμώνας						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.95	0.92037	0.944444	0.938889	0.948148	0.944444
PINAW	0.431178	0.427708	0.428225	0.47275	0.424555	0.428225
CWC	0.431178	0.427708	0.428225	0.47275	0.424555	0.428225
7 ^η Ζώνη – Άνοιξη						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.932971	0.951087	0.940217	0.943841	0.95471	0.943841
PINAW	0.418477	0.442807	0.452236	0.445033	0.485029	0.445033
CWC	0.418477	0.442807	0.452236	0.445033	0.485029	0.445033



Σχήμα 6.5: Τα παραγόμενα διαστήματα πρόβλεψης του συνόλου δεδομένων αξιολόγησης έπειτα από την πρώτη εκπαίδευση του μοντέλου LUBE-PSO-CWC, για τα δεδομένα του φθινοπώρου της 7^{ης} ζώνης.

Πίνακας 6.4: Τα αποτελέσματα της αξιολόγησης των πέντε επαναλήψεων της εκπαίδευσης του μοντέλου LUBE-MOPSO για τα δεδομένα όλων των εποχών της 7^{ης} ζώνης.

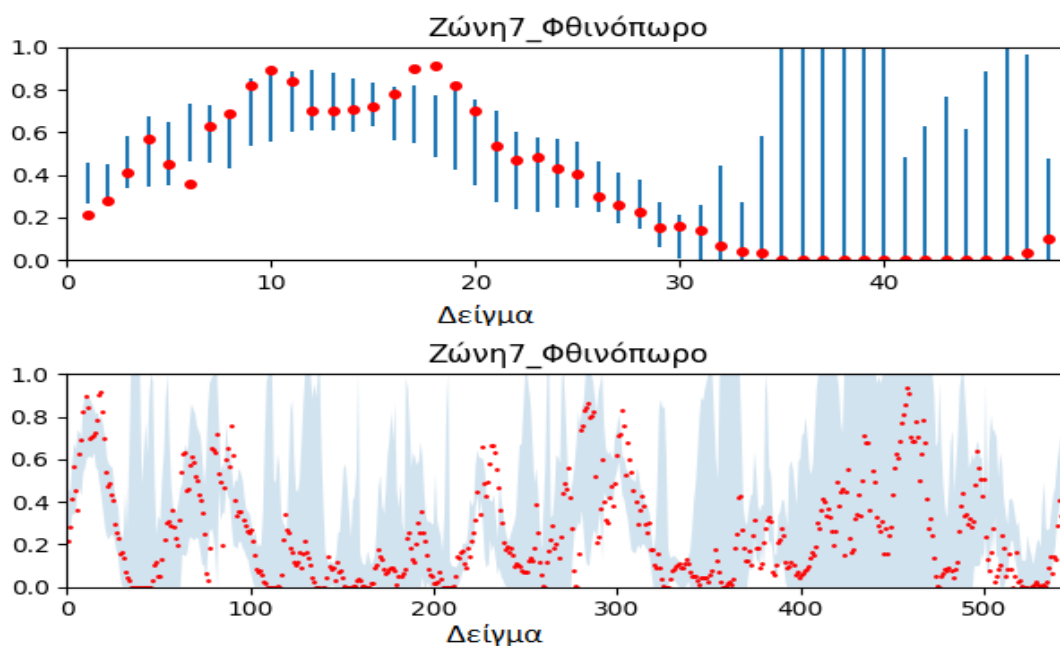
LUBE-MOPSO						
7 ^η Ζώνη – Καλοκαίρι						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.900362	0.913043	0.932971	0.900362	0.914855	0.913043
PINAW	0.41788	0.40189	0.434617	0.414362	0.443209	0.41788
CWC	0.41788	0.40189	0.434617	0.414362	0.443209	0.41788
7 ^η Ζώνη – Φθινόπωρο						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.912088	0.912088	0.915751	0.901099	0.912088	0.912088
PINAW	0.38482	0.406861	0.352516	0.373941	0.379367	0.379367
CWC	0.38482	0.406861	0.352516	0.373941	0.379367	0.379367
7 ^η Ζώνη – Χειμώνας						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.903704	0.905556	0.911111	0.905556	0.92037	0.905556
PINAW	0.374572	0.382472	0.438422	0.377537	0.391885	0.382472
CWC	0.374572	0.382472	0.438422	0.377537	0.391885	0.382472
7 ^η Ζώνη – Άνοιξη						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.902174	0.916667	0.905797	0.903986	0.902174	0.903986
PINAW	0.411764	0.398783	0.452534	0.410969	0.400778	0.410969
CWC	0.411764	0.398783	0.452534	0.410969	0.400778	0.410969



Σχήμα 6.6: Τα παραγόμενα διαστήματα πρόβλεψης του συνόλου δεδομένων αξιολόγησης έπειτα από την πρώτη εκπαίδευση του μοντέλου LUBE-MOPSO, για τα δεδομένα του φθινοπώρου της 7^{ης} ζώνης.

Πίνακας 6.5: Τα αποτελέσματα της αξιολόγησης των πέντε επαναλήψεων της εκπαίδευσης του μοντέλου BELM για τα δεδομένα όλων των εποχών της 7^{ης} ζώνης.

BELM						
7 ^η Ζώνη – Καλοκαίρι						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.927536	0.907609	0.902174	0.900362	0.889493	0.902174
PINAW	0.520722	0.5125	0.495	0.50293	0.511255	0.511255
CWC	0.520722	0.5125	0.495	0.50293	2.828965	0.5125
7 ^η Ζώνη – Φθινόπωρο						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.873626	0.891941	0.90293	0.904762	0.917582	0.90293
PINAW	0.522943	0.464129	0.399998	0.419534	0.431348	0.431348
CWC	8.770278	2.319523	0.399998	0.419534	0.431348	0.431348
7 ^η Ζώνη – Χειμώνας						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.894444	0.933333	0.890741	0.916667	0.905556	0.905556
PINAW	0.45616	0.4851	0.478075	0.497856	0.463056	0.478075
CWC	2.015783	0.4851	2.575564	0.497856	0.463056	0.497856
7 ^η Ζώνη – Άνοιξη						
επανάληψη	1	2	3	4	5	Διάμεσος
PICP	0.923913	0.916667	0.92029	0.925725	0.925725	0.923913
PINAW	0.468896	0.455027	0.465497	0.464865	0.468675	0.465497
CWC	0.468896	0.455027	0.465497	0.464865	0.468675	0.465497

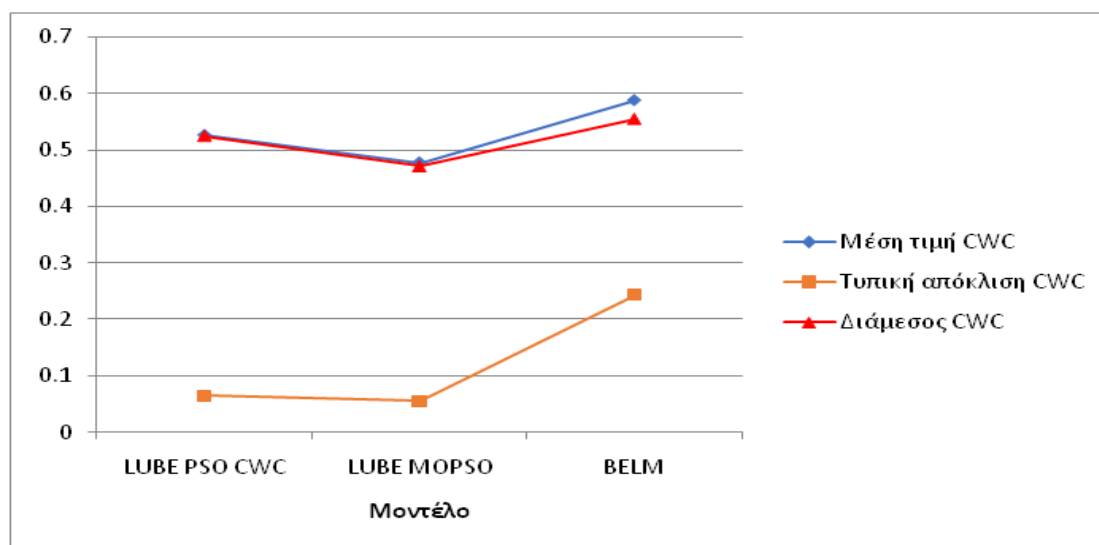


Σχήμα 6.7: Τα παραγόμενα διαστήματα πρόβλεψης του συνόλου δεδομένων αξιολόγησης έπειτα από την πρώτη εκπαίδευση του μοντέλου BELM, για τα δεδομένα του φθινοπώρου της 7^{ης} ζώνης.

σε ότι αφορά τη διάμεση τιμή του CWC στις περισσότερες περιπτώσεις. Για τα δεδομένα του καλοκαιριού της 9^{ης} ζώνης, η διάμεση τιμή του CWC για το μοντέλο BELM παρουσιάζει πολύ μεγάλη αύξηση. Συγκεκριμένα λαμβάνει τιμή ίση με $CWC = 2.0538$. Αυτό σημαίνει ότι στην πλειοψηφία των πέντε επαναλήψεων της εκπαίδευσης και της αξιολόγησης του μοντέλου BELM, τα παραγόμενα διαστήματα πρόβλεψης είχαν δείκτη κάλυψης μικρότερο του ονομαστικού επιπέδου εμπιστοσύνης. Η διάμεση τιμή του PICP του μοντέλου BELM για τα δεδομένα του καλοκαιριού της 9^{ης} ζώνης είναι $PICP = 0.8949 < 0.9 = \mu$.

Τα συμπεράσματα που εξάγονται από το Σχήμα 6.4 επιβεβαιώνονται και από τους Πίνακες 6.3, 6.4 και 6.5 που αφορούν τα δεδομένα όλων των εποχών για την 7^η ζώνη. Σε όλες τις εποχές, το μοντέλο LUBE-MOPSO παρουσιάζει τη μικρότερη διάμεσο CWC ενώ το μοντέλο LUBE-PSO-CWC παρουσιάζει μικρότερη διάμεσο CWC έναντι του μοντέλου BELM. Σε όλες τις εποχές της 7^{ης} ζώνης τα διαστήματα πρόβλεψης του μοντέλου LUBE-PSO-CWC έχουν το μεγαλύτερο μέσο δείκτη κάλυψης, δηλαδή τη μεγαλύτερη διάμεση τιμή του PICP. Επομένως, είναι περισσότερες οι πραγματικές τιμές αιολικής ισχύος που βρίσκονται εντός των παραγόμενων διαστημάτων πρόβλεψης. Από τον Πίνακα 6.5 φαίνεται ότι σε πέντε περιπτώσεις το μοντέλο BELM παράγει διαστήματα πρόβλεψης με τιμή PICP μικρότερη του ονομαστικού επιπέδου εμπιστοσύνης. Σε σύγκριση με τα μοντέλα LUBE-MOPSO και LUBE-PSO-CWC, το μοντέλο BELM παρουσιάζει μεγαλύτερη αστάθεια σε ότι αφορά τον βασικό στόχο της βελτιστοποίησης των διαστημάτων πρόβλεψης, δηλαδή το να υπερτερεί ο δείκτης κάλυψης έναντι του ονομαστικού επιπέδου εμπιστοσύνης. Στα Σχήματα 6.5, 6.6 και 6.7 φαίνεται ποιοτικά η διαφορά των διαστημάτων πρόβλεψης κάθε μοντέλου. Τα διαστήματα πρόβλεψης του μοντέλου LUBE-MOPSO έχουν κατά μέσο όρο μικρότερο εύρος, διατηρώντας ταυτόχρονα τον δείκτη κάλυψης μεγαλύτερο του ονομαστικού επιπέδου εμπιστοσύνης.

Στο Σχήμα 6.8 αναπαρίστανται γραφικά η μέση τιμή, η διάμεσος και η τυπική απόκλιση των CWC για κάθε εποχή σε κάθε ζώνη των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM. Πρόκειται ουσιαστικά για τις μέσες τιμές, τις διαμέσους και τις τυπικές αποκλίσεις των τιμών του CWC που αναπαρίστανται γραφικά στο



Σχήμα 6.8: Η διάμεσος, η μέση τιμή και η τυπική απόκλιση του CWC των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM για το σύνολο των εποχών και των ζωνών.

Πίνακας 6.6: Σύγκριση της απόδοσης των μοντέλων LUBE–PSO–CWC, LUBE–MOPSO και BELM συνολικά για όλες τις εποχές και όλες τις ζώνες.

	Διάμεσος PICP	Διάμεσος PINAW	Διάμεσος CWC	Χρόνος Εκτέλεσης
BELM	0.909838	0.550045	0.554695	0' 27"
LUBE–PSO–CWC	0.939889	0.524825	0.524825	3' 37"
LUBE–MOPSO	0.908017	0.471693	0.471693	2' 42"
Διαφορά BELM και LUBE– PSO–CWC (%)	3.3	–4.59	–5.38	703.7
Διαφορά BELM και MOPSO– LUBE (%)	–0.2	–14.24	–14.96	500
Διαφορά LUBE–PSO–CWC και MOPSO–LUBE (%)	–3.39	–10.12	–10.12	–25.35

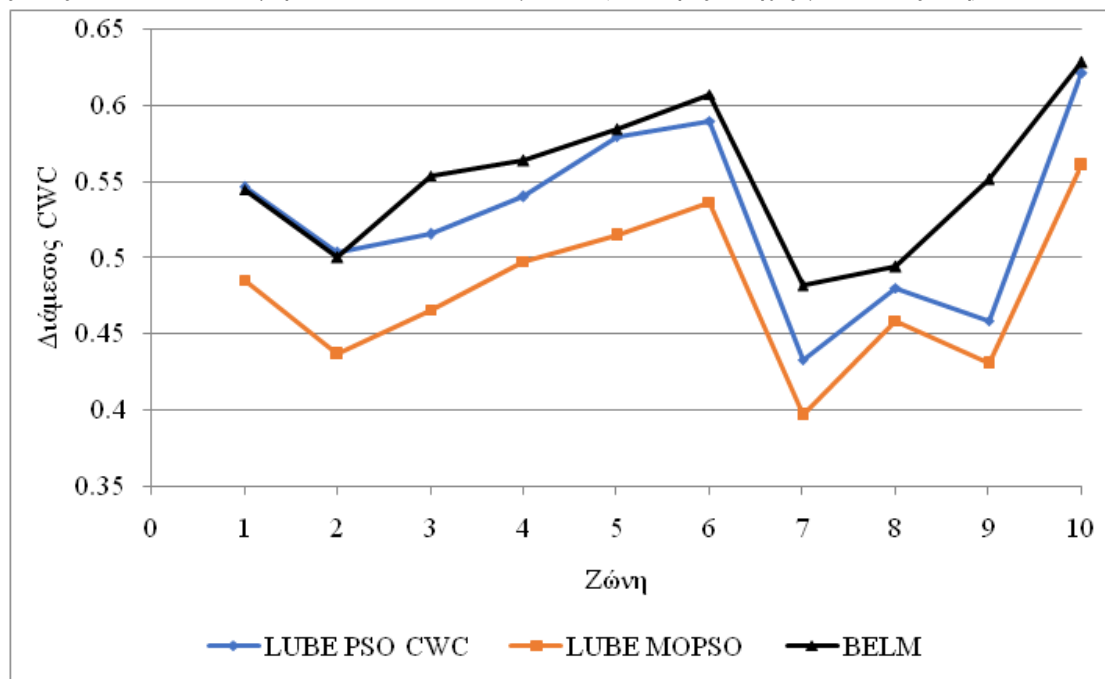
Σχήμα 6.4. Από τις μέσες τιμές και τις διαμέσους του CWC φαίνεται η διαφορά στην ποιότητα των παραγόμενων διαστημάτων πρόβλεψης συνολικά. Το μοντέλο LUBE–MOPSO υπερέρχει έναντι των υπόλοιπων δύο μοντέλων, με μέση τιμή CWC ίση με 0.4778 και διάμεση τιμή CWC ίση με 0.4717. Το μοντέλο LUBE–PSO–CWC, με μέση τιμή CWC ίση με 0.5265 και διάμεση τιμή CWC ίση με 0.5248, υπερέρχει έναντι του μοντέλου BELM του οποίου η μέση τιμή του CWC ισούται με 0.5882 και η διάμεση τιμή του CWC ισούται με 0.5547. Η τυπική απόκλιση των διαμέσων τιμών του CWC του μοντέλου LUBE–PSO–CWC ισούται με 0.0647, ενώ η αντίστοιχη τιμή του μοντέλου LUBE–MOPSO ισούται με 0.0548. Το μοντέλο LUBE–MOPSO παρουσιάζει δηλαδή ελαφρώς μικρότερη διακύμανση τιμών σε ότι αφορά το CWC έναντι του μοντέλου LUBE–PSO–CWC. Επομένως, το μοντέλο LUBE–MOPSO πέρα από καλύτερα αποτελέσματα σε ότι αφορά το μέσο εύρος των διαστημάτων πρόβλεψης, παρουσιάζει και μεγαλύτερη σταθερότητα έναντι των υπόλοιπων μοντέλων. Η τυπική απόκλιση των διαμέσων τιμών του CWC του μοντέλου BELM είναι αρκετά μεγαλύτερη σε σύγκριση με τις αντίστοιχες τυπικές αποκλίσεις των υπόλοιπων μοντέλων. Συγκεκριμένα ισούται με 0.2437. Η σχετικά μεγάλη τιμή της τυπικής απόκλισης του CWC για το μοντέλο BELM φανερώνει την αστάθεια του μοντέλου, δηλαδή την ανικανότητά του σε πολλές περιπτώσεις να ικανοποιήσει το ονομαστικό επίπεδο εμπιστοσύνης.

Στον Πίνακα 6.6 παρουσιάζονται συγκεντρωτικά οι διάμεσες τιμές των PICP, PINAW και CWC για όλες τις εποχές και όλες τις ζώνες καθώς και οι μέσοι χρόνοι εκτέλεσης των μοντέλων LUBE–PSO–CWC, LUBE–MOPSO και BELM. Ως χρόνος εκτέλεσης ορίζεται ο χρόνος από τη φόρτωση και επεξεργασία των δεδομένων έως και την παραγωγή των διαστημάτων πρόβλεψης για το σύνολο των δεδομένων αξιολόγησης. Με μαύρο έντονο χρώμα καταγράφονται τα καλύτερα αποτελέσματα κάθε κατηγορίας. Επίσης, παρουσιάζονται οι ποσοστιαίες διαφορές μεταξύ των τιμών των μοντέλων. Έχοντας ως κριτήριο σύγκρισης τη διάμεσο των τιμών του CWC, τα διαστήματα πρόβλεψης του μοντέλου LUBE–MOPSO παρουσιάζουν βελτίωση κατά 14.96% έναντι αυτών του μοντέλου BELM και κατά 10.12% έναντι αυτών του LUBE–PSO–CWC. Η διάμεση τιμή του PICP του μοντέλου LUBE–PSO–CWC ισούται με 0.9399 και είναι κατά 3% περίπου μεγαλύτερη έναντι των υπόλοιπων μοντέλων. Όπως είναι αναμενόμενο, το μοντέλο BELM παρουσιάζει πολύ μικρότερο μέσο χρόνο εκτέλεσης, μόλις 27 δευτερόλεπτα, καθώς βασίζει τη λειτουργία

του σε Extreme Learning Machines. Τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO βασίζουν τη λειτουργία τους σε συμβατικά τεχνητά νευρωνικά δίκτυα που εκπαιδεύονται μέσω μεταερευτικών αλγορίθμων, επομένως ο χρόνος εκτέλεσής τους είναι αρκετά μεγαλύτερος. Συγκεκριμένα, το μοντέλο LUBE-PSO-CWC παρουσιάζει μέσο χρόνο εκτέλεσης ίσο με 3 λεπτά και 37 δευτερόλεπτα, ενώ το μοντέλο LUBE-MOPSO έχει μέσο χρόνο εκτέλεσης ίσο με 2 λεπτά και 42 δευτερόλεπτα. Παρόλο που το μοντέλο LUBE-MOPSO εκπαιδεύεται σε πολυκριτηριακά πλαίσια, επιτελώντας περαιτέρω λειτουργίες, όπως τη σύγκριση και αποθήκευση πολλαπλών λύσεων ή τη δημιουργία του προσαρμοστικού πλέγματος, είναι κατά 25.35% ταχύτερο έναντι του μοντέλου LUBE-PSO-CWC. Αυτό συμβαίνει λόγω της διαφοράς των δύο μοντέλων στις βέλτιστες τιμές των παραμέτρων τους, και συγκεκριμένα στο μικρότερο πλήθος σωματιδίων και νευρώνων κρυφού στρώματος που έχει το μοντέλο LUBE-MOPSO σε σχέση με το μοντέλο LUBE-PSO-CWC.

Στο Σχήμα 6.9 παρουσιάζονται οι διάμεσες τιμές CWC των τεσσάρων εποχών κάθε ζώνης για τα μοντέλα LUBE-PSO-CWC, LUBE-MOPSO και BELM. Στην ουσία συγκρίνεται η απόδοση των μοντέλων συνολικά μέσα στο έτος σε κάθε διαφορετική ζώνη. Όπως φαίνεται από το Σχήμα 6.9, το μοντέλο LUBE-MOPSO παρουσιάζει τα καλύτερα αποτελέσματα σε όλες τις ζώνες. Με εξαίρεση την 1^η και τη 2^η ζώνη, το μοντέλο LUBE-PSO-CWC παρουσιάζει καλύτερα αποτελέσματα έναντι του μοντέλου BELM.

Στον Πίνακα 6.7 συνοψίζονται όλα τα παραπάνω αποτελέσματα της σύγκρισης της απόδοσης των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM. Με μαύρο έντονο χρώμα καταγράφονται τα καλύτερα αποτελέσματα κάθε κατηγορίας. Το μοντέλο LUBE-MOPSO παρουσιάζει την καλύτερη απόδοση, καταγράφοντας τις χαμηλότερες τιμές στη διάμεσο και την τυπική απόκλιση του CWC συνολικά, αλλά και σε κάθε μία από τις δέκα ζώνες. Από τα 40 διαφορετικά σύνολα δεδομένων (τέσσερις εποχές για κάθε ζώνη), στα 39



Σχήμα 6.9: Οι διάμεσες τιμές του CWC ανά ζώνη για τα μοντέλα LUBE-PSO-CWC, LUBE-MOPSO και BELM.

Πίνακας 6.7: Σύνοψη των αποτελεσμάτων της σύγκρισης της απόδοσης των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM.

Παράμετρος σύγκρισης	LUBE-MOPSO	LUBE-PSO-CWC	BELM
Διάμεσος CWC	0.471693	0.524825	0.554695
Τυπική απόκλιση CWC	0.0548	0.0647	0.2437
Διάμεσος PICP	0.908017	0.939889	0.909838
Μέσος χρόνος εκτέλεσης	2' 42"	3' 37"	0' 27"
Πλήθος ζωνών με την καλύτερη απόδοση (10 ζώνες)	10	0	0
Πλήθος ζωνών με τη δεύτερη καλύτερη απόδοση (10 ζώνες)	0	8	2
Πλήθος περιπτώσεων με την καλύτερη απόδοση (4 εποχές για κάθε ζώνη → 40 περιπτώσεις)	39	1	0
Πλήθος περιπτώσεων με τη δεύτερη καλύτερη απόδοση (4 εποχές για κάθε ζώνη → 40 περιπτώσεις)	1	28	11

το μοντέλο LUBE-MOPSO παράγει καλύτερα διαστήματα πρόβλεψης σε σχέση με αυτά των μοντέλων LUBE-PSO-CWC και BELM. Το μοντέλο LUBE-PSO-CWC παράγει διαστήματα πρόβλεψης με τη μεγαλύτερη διάμεση τιμή του δείκτη κάλυψης. Το μέσο εύρος των διαστημάτων πρόβλεψης του είναι μεγαλύτερο από αυτό των διαστημάτων πρόβλεψης του μοντέλου LUBE-MOPSO. Ταυτόχρονα παρουσιάζει ελαφρώς μεγαλύτερη αστάθεια. Σε 8 από τις 10 ζώνες παρουσιάζει τη δεύτερη καλύτερη απόδοση, ενώ από τα 40 διαφορετικά σύνολα δεδομένων, στο ένα παρουσιάζει την καλύτερη απόδοση και στα 28 τη δεύτερη καλύτερη απόδοση. Σε ότι αφορά τον χρόνο εκτέλεσης, είναι το μοντέλο που χρειάζεται τον περισσότερο χρόνο για την εκπαίδευσή του και την παραγωγή των διαστημάτων πρόβλεψης για το σύνολο των δεδομένων αξιολόγησης. Το μοντέλο BELM είναι το μοντέλο με τον μικρότερο μέσο χρόνο εκτέλεσης. Ωστόσο καταγράφει τη χειρότερη απόδοση σε 8 από τις 10 ζώνες και σε 29 από τα 40 διαφορετικά σύνολα δεδομένων. Είναι το μοντέλο που παρουσιάζει τη μεγαλύτερη αστάθεια, δηλαδή τη μεγαλύτερη διακύμανση τιμών του CWC. Σε ένα από τα 40 διαφορετικά σύνολα δεδομένων η διάμεσος του PICP είναι μικρότερη του ονομαστικού επιπέδου εμπιστοσύνης, ενώ στο 23.5% των συνολικών επαναλήψεων της εκπαίδευσης και της αξιολόγησης για όλα τα διαφορετικά σύνολα δεδομένων (200 επαναλήψεις) η τιμή του PICP είναι μικρότερη του ονομαστικού επιπέδου εμπιστοσύνης. Καθώς η εκπαίδευση ενός Extreme Learning Machine δεν βασίζεται σε επαναληπτικές διαδικασίες, το πλήθος των δεδομένων επηρεάζει σε μεγαλύτερο βαθμό την επιτυχημένη ή μη έκβαση της εκπαίδευσης σε σχέση με τα συμβατικά τεχνητά νευρωνικά δίκτυα. Η αστάθεια του μοντέλου BELM πιθανώς οφείλεται στο μικρό σχετικά πλήθος των δεδομένων εκπαίδευσης που χρησιμοποιείται στην παρούσα διπλωματική εργασία.

Για την πληρέστερη αξιολόγηση της απόδοσης των μοντέλων LUBE-PSO-CWC και LUBE-MOPSO της παρούσας διπλωματικής εργασίας, γίνεται σύγκριση με τα

Πίνακας 6.8: Σύγκριση αποτελεσμάτων των μοντέλων FCNN, LUBE-PSO-CWC, LUBE-MOPSO και BELM για τα δεδομένα της 1^{ης} και της 7^{ης} ζώνης.

Μοντέλο	1 ^η Ζώνη		7 ^η Ζώνη	
	PICP	PINAW	PICP	PINAW
FCNN	0.857	0.494	0.904	0.346
BELM	0.854	0.4688	0.9096	0.4446
LUBE-PSO-CWC	0.8873	0.4844	0.9235	0.4049
LUBE-MOPSO	0.9013	0.4536	0.9249	0.385

αποτελέσματα του πιθανοτικού μοντέλου πρόβλεψης που αναπτύχθηκε από τους H. Wen, J. Gu, J. Ma και Z. Jin [8]. Το μοντέλο της εργασίας [8] αποτελείται από ένα πλήρες συνελκτικό τεχνητό νευρωνικό δίκτυο (Fully Convolutional Neural Network – FCNN) [54]. Το πλήρες συνελκτικό τεχνητό νευρωνικό δίκτυο εκπαιδεύεται σύμφωνα με τη θεωρία του Μπέυζ (Bayesian theory) και τη διακοπή Monte Carlo (Monte Carlo dropout) [55].

Για να είναι αξιόπιστη η σύγκριση της απόδοσης των μοντέλων, πρέπει τα δεδομένα που θα χρησιμοποιηθούν για την εκπαίδευση και την αξιολόγηση να είναι ίδια. Το μοντέλο της εργασίας [8] εκπαιδεύεται και αξιολογείται με τα δεδομένα της 1^{ης} και της 7^{ης} ζώνης του διαγωνισμού GEFCom 2014. Ως δεδομένα εκπαίδευσης θεωρούνται τα δεδομένα από την 1^η Ιανουαρίου του 2012 έως και την 30^η Σεπτεμβρίου του 2012 για κάθε μία από τις δύο ζώνες, ενώ ως δεδομένα αξιολόγησης θεωρούνται τα δεδομένα του Νοεμβρίου του 2012. Τα δεδομένα του Οκτωβρίου του 2012 χρησιμοποιούνται ως σύνολο δεδομένων επαλήθευσης ώστε να κριθεί εάν είναι επιτυχημένη η εκπαίδευση πριν την τελική αξιολόγηση του μοντέλου. Επομένως τα μοντέλα LUBE-PSO-CWC, LUBE-MOPSO και BELM εκπαιδεύονται εκ νέου, κατά τον ίδιο τρόπο που εκπαιδεύεται το μοντέλο της εργασίας [8]. Τα μετρητικά αξιολόγησης που χρησιμοποιούνται είναι το PICP και το PINAW. Για να είναι άμεσα συγκρίσιμα τα αποτελέσματα μεταξύ των μοντέλων, ως ονομαστικό επίπεδο εμπιστοσύνης για την εκπαίδευση των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM ορίζεται η στρογγυλοποιημένη προς τα κάτω τιμή του PICP του μοντέλου FCNN για κάθε ζώνη. Επομένως, για την 1^η ζώνη ορίζεται $\mu = 0.85$ ενώ για τη 7^η ζώνη ορίζεται $\mu = 0.9$.

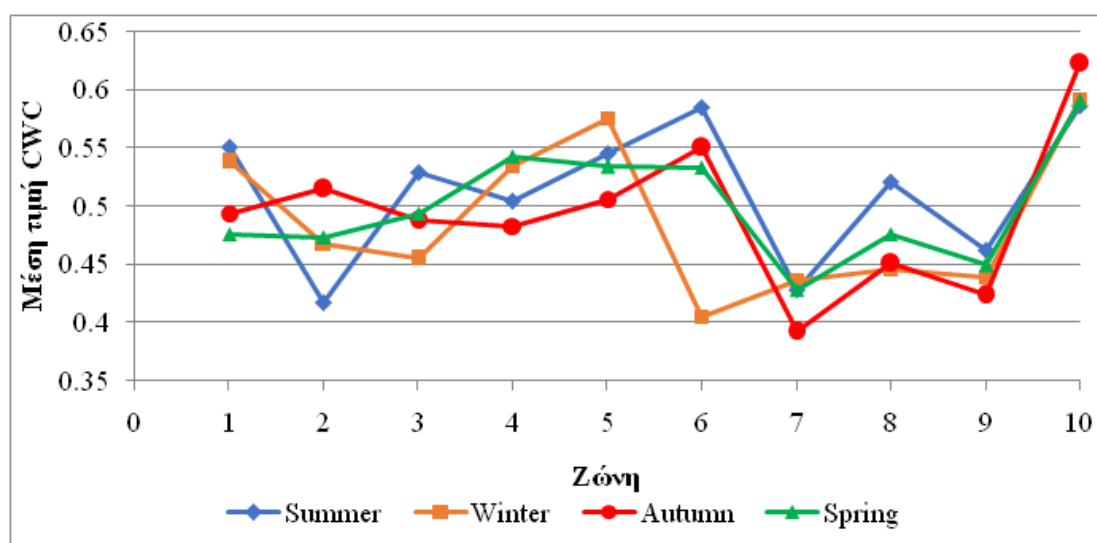
Στον Πίνακα 6.8 παρουσιάζονται τα αποτελέσματα των μοντέλων FCNN, LUBE-PSO-CWC, LUBE-MOPSO και BELM για τα δεδομένα της 1^{ης} και της 7^{ης} ζώνης. Με μαύρο έντονο χρώμα καταγράφονται τα καλύτερα αποτελέσματα κάθε κατηγορίας. Για τα δεδομένα της 1^{ης} ζώνης, το μοντέλο LUBE-MOPSO παρουσιάζει την καλύτερη επίδοση καθώς πετυχαίνει τον μεγαλύτερο βαθμό κάλυψης με το μικρότερο μέσο εύρος διαστημάτων πρόβλεψης. Το μοντέλο LUBE-PSO-CWC παρουσιάζει καλύτερα αποτελέσματα σε σχέση με το μοντέλο FCNN και στον βαθμό κάλυψης και στο μέσο εύρος των διαστημάτων πρόβλεψης. Σε σχέση με το μοντέλο BELM, το μοντέλο LUBE-PSO-CWC υπερτερεί στον βαθμό κάλυψης, ωστόσο υστερεί στο μέσο εύρος των διαστημάτων πρόβλεψης. Για τα δεδομένα της 7^{ης} ζώνης τα αποτελέσματα των μοντέλων LUBE-PSO-CWC και LUBE-MOPSO έχουν ελάχιστη απόκλιση. Τα δύο μοντέλα πετυχαίνουν τα καλύτερα αποτελέσματα σε ότι αφορά τον βαθμό κάλυψης, ωστόσο το μοντέλο FCNN παρουσιάζει το μικρότερο μέσο εύρος των διαστημάτων πρόβλεψης. Σε σχέση με το μοντέλο BELM, τα μοντέλα LUBE-

PSO-CWC και LUBE-MOPSO παρουσιάζουν καλύτερα αποτελέσματα τόσο στον βαθμό κάλυψης όσο και στο μέσο εύρος των διαστημάτων πρόβλεψης.

Συμπερασματικά, μεταξύ των μοντέλων LUBE-PSO-CWC, LUBE-MOPSO και BELM, το μοντέλο LUBE-MOPSO παρουσιάζει σταθερά την καλύτερη επίδοση. Οι επιδόσεις των μοντέλων LUBE-PSO-CWC και BELM είναι παραπλήσιες, με το μοντέλο LUBE-PSO-CWC να υπερτερεί οριακά. Λαμβάνοντας υπόψη τα αποτελέσματα του Πίνακα 6.8, αμφότερα τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO παρουσιάζουν αποτελέσματα που είναι άμεσα συγκρίσιμα με επιτυχημένα πιθανοτικά μοντέλα της διαθέσιμης βιβλιογραφίας, όπως το μοντέλο FCNN της εργασίας [8].

6.2.2 Εποχική ανάλυση

Ο διαχωρισμός των δεδομένων του διαγωνισμού GEFCom 2014 σε εποχές συντελείται με σκοπό την εποχική ανάλυση. Ως εποχική ανάλυση ορίζεται η εκπαίδευση και αξιολόγηση ενός μοντέλου ανά εποχή, με στόχο τη διερεύνηση για τυχόν συσχετίσεις μεταξύ των δεδομένων κάθε εποχής, ανάδειξη περιοδικών φαινομένων και σύγκριση της αποδοτικότητας ανά εποχή, ώστε να εξαχθούν ορισμένα συμπεράσματα τα οποία θα μπορούν να ληφθούν υπόψη κατά τη σχεδίαση ενός μοντέλου με σκοπό τη βελτίωση της αποδοτικότητάς του. Η παραγωγή ισχύος ενός αιολικού σταθμού εξαρτάται σε μεγάλο βαθμό από τις καιρικές συνθήκες οι οποίες παρουσιάζουν αυξημένη ομοιότητα ανά εποχή. Η επιτυχημένη ενσωμάτωση των συμπερασμάτων της εποχικής ανάλυσης σε ένα μοντέλο πιθανοτικής πρόβλεψης της αιολικής παραγωγής το καθιστά πιο ευέλικτο και αποδοτικό, καθώς ανάλογα την εποχή και τα χαρακτηριστικά της θα προσαρμόζονται η δομή και οι τιμές των παραμέτρων του μοντέλου. Τα μοντέλα που χρησιμοποιούνται για την εποχική ανάλυση είναι το LUBE-PSO-CWC και το LUBE-MOPSO. Για την εποχική ανάλυση χρησιμοποιούνται τα αποτελέσματα που παρουσιάστηκαν στην Ενότητα 6.2.1.

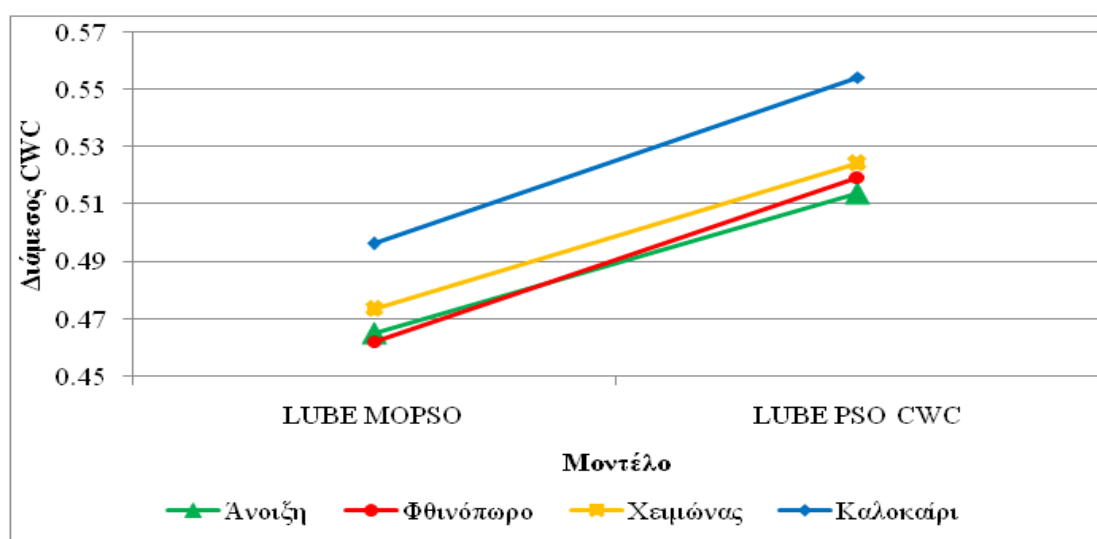


Σχήμα 6.10: Η μέση τιμή του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO σε κάθε ζώνη.

Στο Σχήμα 6.10 απεικονίζεται η μέση τιμή του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO σε κάθε ζώνη. Η μέση τιμή του CWC προκύπτει δηλαδή από την αξιολόγηση των πέντε επαναλήψεων της εκπαίδευσης των δύο μοντέλων για κάθε εποχή σε κάθε ζώνη. Ο λόγος για τον οποίο λαμβάνεται η μέση τιμή των αποτελεσμάτων και των δύο μοντέλων, και όχι κάθε μοντέλου ξεχωριστά, είναι γιατί οι διαφορές στις σχετικές μεταβολές των αποτελεσμάτων του κάθε μοντέλου μεταξύ των εποχών σε κάθε ζώνη είναι αμελητέες. Παρατηρώντας το Σχήμα 6.10 δεν φαίνεται να υπάρχει κάποια προφανής σχέση μεταξύ της αποδοτικότητας των μοντέλων και της κάθε εποχής. Ανάλογα τη ζώνη, η σχετική απόδοση των μοντέλων ανά εποχή μεταβάλλεται. Ενδεχομένως αυτό συμβαίνει διότι κάθε ζώνη αντιστοιχεί σε ένα διαφορετικό αιολικό σταθμό στην Αυστραλία, δηλαδή σε διαφορετική γεωγραφική τοποθεσία. Οι τοποθεσίες αυτές πιθανόν να διαφέρουν πολύ ως προς το υψόμετρο και το κλίμα. Επομένως, αν σε κάθε ζώνη επικρατούν διαφορετικές καιρικές ιδιομορφίες, η εξαγωγή συμπερασμάτων από την εποχική ανάλυση δυσκολεύει.

Ωστόσο ορισμένα πρώτα συμπεράσματα μπορούν να εξαχθούν από το Σχήμα 6.10. Το φθινόπωρο, τα δύο μοντέλα καταγράφουν την καλύτερη απόδοση σε σχέση με τις υπόλοιπες εποχές σε 4 από τις 10 ζώνες, ενώ μόλις σε 2 από τις 10 ζώνες η φθινοπωρινή απόδοση των δύο μοντέλων είναι η χειρότερη. Αντίστοιχα, το καλοκαίρι τα δύο μοντέλα καταγράφουν τη χειρότερη απόδοση σε σχέση με τις υπόλοιπες εποχές σε 4 από τις 10 ζώνες, ενώ μόλις σε 2 από τις 10 ζώνες η καλοκαιρινή απόδοση των δύο μοντέλων είναι η καλύτερη.

Στο Σχήμα 6.11 απεικονίζονται οι διάμεσες τιμές του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO. Σε κάθε εποχή για κάθε μοντέλο υπάρχουν δέκα διάμεσες τιμές του CWC, μία για κάθε ζώνη. Οι τιμές του Σχήματος 6.11 προκύπτουν υπολογίζοντας τη διάμεσο των δέκα τιμών αυτών για κάθε μοντέλο. Το συμπέρασμα που εξάγεται από το Σχήμα 6.10 σχετικά με την απόδοση των μοντέλων το καλοκαίρι επιβεβαιώνεται και από το Σχήμα 6.11. Ενώ και για τα δύο μοντέλα οι διάμεσες τιμές του CWC για τις τρεις υπόλοιπες εποχές έχουν μικρή απόκλιση, οι διάμεσες τιμές του CWC που



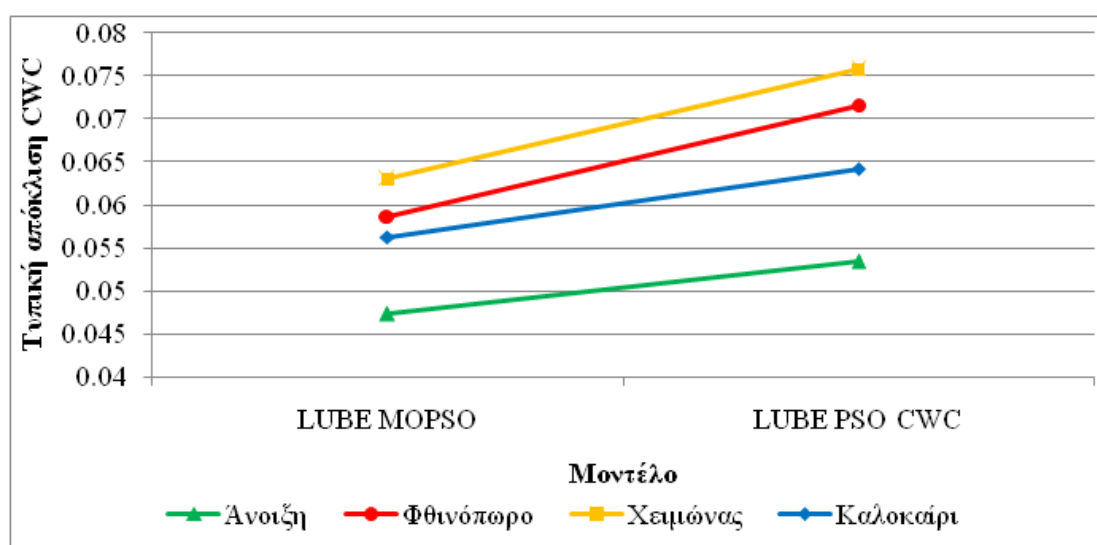
Σχήμα 6.11: Η διάμεση τιμή του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO.

Πίνακας 6.9: Οι διάμεσες τιμές του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO.

Εποχή	LUBE-PSO-CWC	LUBE-MOPSO
Καλοκαίρι	0.554099	0.496232
Φθινόπωρο	0.519249	0.462274
Χειμώνας	0.524093	0.473474
Άνοιξη	0.514076	0.465108

αντιστοιχούν στο καλοκαίρι έχουν σχετικά μεγάλη απόκλιση. Οι ακριβείς τιμές του Σχήματος 6.11 παρουσιάζονται στον Πίνακα 6.9. Με έντονο μαύρο χρώμα καταγράφονται οι τιμές της εποχής που παρουσιάζει τη μικρότερη διάμεση τιμή του CWC για κάθε μοντέλο. Οι ποσοστιαίες διαφορές μεταξύ των διαμέσων τιμών του CWC για το φθινόπωρο, την άνοιξη και τον χειμώνα κυμαίνονται στο 1.5 ~ 2%, ενώ οι αντίστοιχες ποσοστιαίες διαφορές μεταξύ των διαμέσων τιμών του CWC του καλοκαιριού και των υπόλοιπων εποχών κυμαίνονται στο 4.8 ~ 7.8%. Από τον Πίνακα 6.9 φαίνεται επίσης ότι τα αποτελέσματα για την άνοιξη και το φθινόπωρο παρουσιάζουν μια οριακή βελτίωση σε σχέση με τον χειμώνα και για τα δύο μοντέλα.

Στο Σχήμα 6.12 απεικονίζεται η τυπική απόκλιση του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO. Για τον υπολογισμό της τυπικής απόκλισης του CWC κάθε μοντέλου, λαμβάνονται υπόψη οι διάμεσες τιμές CWC ανά εποχή για κάθε ζώνη. Η τυπική απόκλιση του CWC εκφράζει τη διακύμανση των διαμέσων τιμών του CWC ανά ζώνη. Όσο μεγαλύτερη είναι η τυπική απόκλιση του CWC μιας εποχής, τόσο μεγαλύτερη διακύμανση παρουσιάζουν οι τιμές του CWC στις δέκα ζώνες. Η μεγάλη διακύμανση τιμών υποδεικνύει ότι η απόδοση των μοντέλων έχει μεγαλύτερη εξάρτηση από τη γεωγραφική τοποθεσία σε ορισμένες εποχές. Από το Σχήμα 6.12 φαίνεται πως και για τα δύο μοντέλα η μικρότερη τυπική απόκλιση των τιμών του CWC καταγράφεται την άνοιξη, ενώ η μεγαλύτερη τυπική απόκλιση των τιμών του CWC καταγράφεται τον χειμώνα. Η δεύτερη



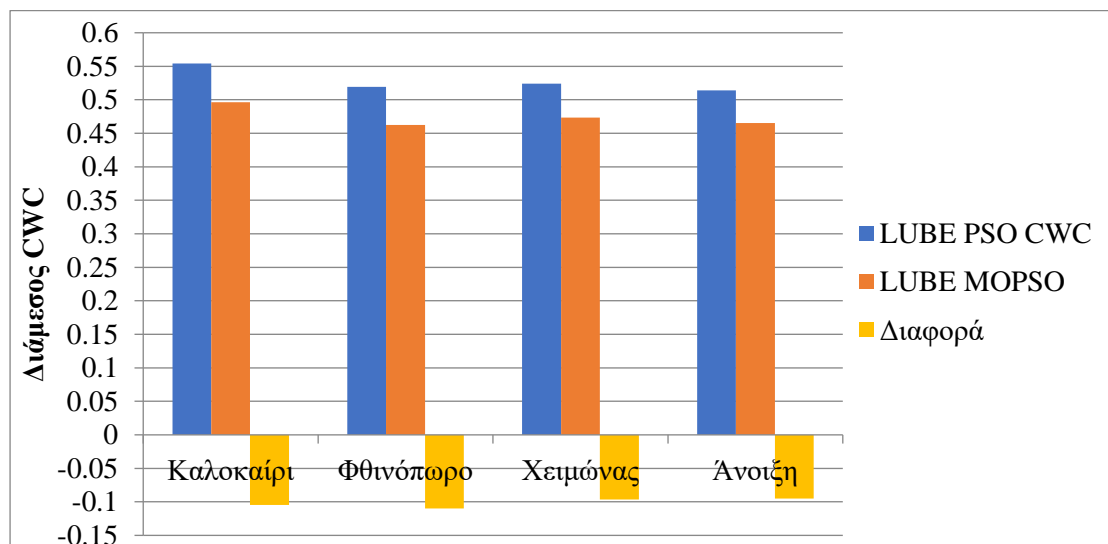
Σχήμα 6.12: Η τυπική απόκλιση του CWC ανά εποχή για τα μοντέλα LUBE-PSO-CWC και LUBE-MOPSO.

Πίνακας 6.10: Η τυπική απόκλιση των τιμών του CWC ανά εποχή για τα μοντέλα LUBE–PSO–CWC και LUBE–MOPSO.

Εποχή	LUBE–PSO–CWC	LUBE–MOPSO
Καλοκαίρι	0.0642	0.0562
Φθινόπωρο	0.0715	0.0586
Χειμώνας	0.0757	0.063
Άνοιξη	0.0534	0.0473

μεγαλύτερη τυπική απόκλιση των τιμών του CWC καταγράφεται το φθινόπωρο, ενώ η δεύτερη μικρότερη τυπική απόκλιση των τιμών του CWC καταγράφεται το καλοκαίρι. Από το Σχήμα 6.12 φαίνεται επίσης πως η τυπική απόκλιση των τιμών του CWC την άνοιξη είναι αρκετά μικρότερη σε σχέση με τις τυπικές αποκλίσεις των τιμών του CWC των υπόλοιπων εποχών. Οι ακριβείς τιμές του Σχήματος 6.12 παρουσιάζονται στον Πίνακα 6.10. Με έντονο μαύρο χρώμα καταγράφεται η τιμή της εποχής που παρουσιάζει τη μικρότερη τυπική απόκλιση του CWC για κάθε μοντέλο. Ενώ οι ποσοστιαίες διαφορές των τυπικών αποκλίσεων μεταξύ του χειμώνα, του καλοκαιριού και του φθινοπώρου για τα δύο μοντέλα κυμαίνονται στο 4 ~ 15% περίπου, οι ποσοστιαίες διαφορές των τυπικών αποκλίσεων μεταξύ της άνοιξης και των υπόλοιπων εποχών κυμαίνονται στο 15 ~ 29.5% περίπου. Τα αποτελέσματα της άνοιξης έχουν μικρότερη διακύμανση τιμών, επομένως έχουν μικρότερη εξάρτηση από τη γεωγραφική τοποθεσία του αιολικού σταθμού. Αντίθετα, τα αποτελέσματα του φθινοπώρου και του χειμώνα παρουσιάζουν μεγαλύτερη διακύμανση τιμών, δηλαδή έχουν μεγαλύτερη εξάρτηση από την τοποθεσία του αιολικού σταθμού.

Στο Σχήμα 6.13 συγκρίνονται οι διάμεσες τιμές του CWC του μοντέλου LUBE–PSO–CWC για κάθε εποχή που προκύπτουν από τις διάμεσες τιμές του CWC των δέκα



Σχήμα 6.13: Συγκριτικά η διάμεσος του CWC ανά εποχή για τα μοντέλα LUBE–PSO–CWC και LUBE–MOPSO, καθώς και η ποσοστιαία διαφορά μεταξύ των αποτελεσμάτων των δύο μοντέλων.

Πίνακας 6.11: Η διάμεσος του CWC ανά εποχή για τα μοντέλα LUBE–PSO–CWC και LUBE–MOPSO και η ποσοστιαία διαφορά μεταξύ των αποτελεσμάτων των δύο μοντέλων

Εποχή	LUBE–PSO–CWC	LUBE–MOPSO	Διαφορά (%)
Καλοκαίρι	0.5540994	0.4962319	-10.4435
Φθινόπωρο	0.5192486	0.4622746	-10.9724
Χειμώνας	0.524093	0.4734749	-9.6582
Άνοιξη	0.5140755	0.4651082	-9.5253

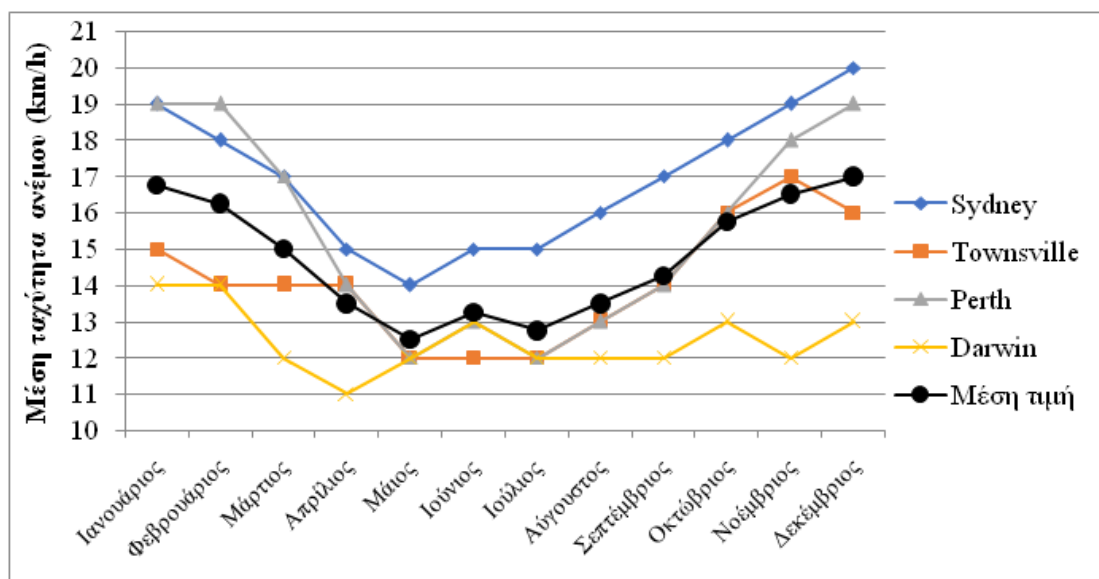
ζωνών, με τις αντίστοιχες διάμεσες τιμές του CWC του μοντέλου LUBE–MOPSO. Με κίτρινο χρώμα απεικονίζεται η ποσοστιαία μείωση που επιφέρει στη διάμεση τιμή του CWC κάθε εποχής το μοντέλο LUBE–MOPSO έναντι του μοντέλου LUBE–PSO–CWC. Οι κίτρινες μπάρες ποσοτικοποιούν δηλαδή την ποσοστιαία βελτίωση που επιφέρει στα αποτελέσματα κάθε εποχής το μοντέλο LUBE–MOPSO. Στον Πίνακα 6.11 φαίνονται οι ακριβείς τιμές του Σχήματος 6.13. Με έντονο μαύρο χρώμα φαίνεται η μεγαλύτερη ποσοστιαία διαφορά που καταγράφεται μεταξύ των δύο μοντέλων, η οποία παρατηρείται το φθινόπωρο και ισούται με -10.9724%. Ωστόσο, παρατηρώντας το Σχήμα 6.13 και τις τιμές του Πίνακα 6.11, φαίνεται πως οι ποσοστιαίες βελτιώσεις που επιφέρει το μοντέλο LUBE–MOPSO ανά εποχή έχουν αμελητέα διαφορά μεταξύ τους, δηλαδή οι βελτιώσεις δεν παρουσιάζουν κάποια εμφανή εξάρτηση από την εποχή. Επομένως, δεν μπορεί να εξαχθεί κάποιο ουσιαστικό συμπέρασμα γύρω από την επίδραση των εποχών στη βελτίωση των αποτελεσμάτων που επιφέρει το μοντέλο LUBE–MOPSO.

Στους Πίνακες 6.12 και 6.13 συνοψίζονται όλα τα αποτελέσματα που προκύπτουν από την εποχική ανάλυση για τα μοντέλα LUBE–PSO–CWC και LUBE–MOPSO, αντίστοιχα. Με έντονο μαύρο χρώμα καταγράφονται οι βέλτιστες τιμές κάθε κατηγορίας. Η μικρότερη διάμεσος CWC παρατηρείται την άνοιξη για το μοντέλο LUBE–PSO–CWC και το φθινόπωρο για το μοντέλο LUBE–MOPSO. Ωστόσο, οι διαφορές στις διάμεσους του CWC μεταξύ του χειμώνα, του φθινοπώρου και της άνοιξης και για τα δύο μοντέλα είναι μικρές. Και για τα δύο μοντέλα, η μεγαλύτερη διάμεσος CWC παρατηρείται το καλοκαίρι. Η τιμή της διαμέσου του CWC το καλοκαίρι παρουσιάζει μεγάλη απόκλιση σε σχέση με τις διαμέσους του CWC των υπόλοιπων εποχών. Σε ότι αφορά τις κατατάξεις της απόδοσης των εποχών ανά ζώνη, υπάρχουν συνολικά 20 περιπτώσεις, που προκύπτουν από τα αποτελέσματα των δύο μοντέλων ξεχωριστά για τις δέκα ζώνες. Σε 9 από τις 20 περιπτώσεις το φθινόπωρο παρουσιάζει την καλύτερη απόδοση, ενώ σε 15 από τις 20 παρουσιάζει την καλύτερη ή τη δεύτερη καλύτερη απόδοση. Το καλοκαίρι από την άλλη παρουσιάζει τη χειρότερη απόδοση σε 9 από τις 20 περιπτώσεις. Λαμβάνοντας υπόψη τις κατατάξεις της απόδοσης των εποχών ανά ζώνη και τις διάμεσες τιμές του CWC ανά εποχή για όλες τις ζώνες, προκύπτει το γενικό συμπέρασμα ότι το καλοκαίρι καταγράφονται τα χειρότερα αποτελέσματα, δηλαδή τα διαστήματα πρόβλεψης που παράγονται τους καλοκαιρινούς μήνες έχουν μεγαλύτερο μέσο εύρος διαστημάτων πρόβλεψης, ενώ το φθινόπωρο καταγράφονται τα καλύτερα αποτελέσματα, δηλαδή τα διαστήματα πρόβλεψης που παράγονται τους φθινοπωρινούς μήνες έχουν το μικρότερο μέσο εύρος διαστημάτων πρόβλεψης. Τα αποτελέσματα αυτά ενδεχομένως σχετίζονται με τη μέση ταχύτητα ανέμου ανά εποχή. Στο Σχήμα 6.14 παρουσιάζονται οι μέσες τιμές της ταχύτητας του ανέμου ανά μήνα για τέσσερις

Πίνακας 6.12: Σύνοψη των αποτελεσμάτων της εποχικής ανάλυσης για το μοντέλο LUBE–PSO–CWC.

Μοντέλο	LUBE–PSO–CWC			
	Καλοκαίρι	Φθινόπωρο	Χειμώνα	Άνοιξη
Διάμεσος CWC για όλες τις ζώνες	0.5540994	0.5192486	0.524093	0.5140755
Τυπική απόκλιση CWC για όλες τις ζώνες	0.0642	0.0715	0.0757	0.0534
Πλήθος ζωνών με την καλύτερη απόδοση	2	5	1	2
Πλήθος ζωνών με τη δεύτερη καλύτερη απόδοση	1	2	4	3
Πλήθος ζωνών με τη χειρότερη απόδοση	3	2	4	1
Πλήθος ζωνών με τη δεύτερη χειρότερη απόδοση	4	1	1	4

πόλεις της Αυστραλίας καθώς και ο μέσος όρος των τιμών αυτών [56]. Οι τέσσερις πόλεις επιλέχθηκαν έτσι ώστε να καλύπτεται κάθε άκρο της Αυστραλίας (Sydney–ΝΑ Αυστραλία, Townsville–ΒΑ Αυστραλία, Perth–ΝΔ Αυστραλία, Darwin–ΒΔ Αυστραλία). Κατά αυτόν τον τρόπο είναι δυνατόν να δοθεί μια γενική εικόνα για τη μέση ταχύτητα του ανέμου ανά μήνα συνολικά στην επικράτεια της Αυστραλίας. Αυτό είναι αναγκαίο καθώς η γεωγραφική τοποθεσία των αιολικών σταθμών που συμβολίζουν οι ζώνες δεν είναι γνωστή. Στο Σχήμα 6.15 παρουσιάζονται οι μέσες τιμές της ταχύτητας του ανέμου ανά εποχή για τις τέσσερις πόλεις της Αυστραλίας καθώς και ο μέσος όρος των τιμών αυτών. Το καλοκαίρι παρατηρούνται κατά μέσο όρο οι χαμηλότερες ταχύτητες ανέμου, ενώ το φθινόπωρο και τον χειμώνα παρατηρούνται κατά μέσο όρο οι υψηλότερες ταχύτητες ανέμου. Υψηλότερες κατά μέσο όρο ταχύτητες ανέμου αντιστοιχούν σε μεγαλύτερα χρονικά διαστήματα λειτουργίας μιας ανεμογεννήτριας στην ονομαστική περιοχή λειτουργίας, επομένως παρατηρείται μεγαλύτερη σταθερότητα στα αποτελέσματα και αποδοτικότερα διαστήματα πρόβλεψης.

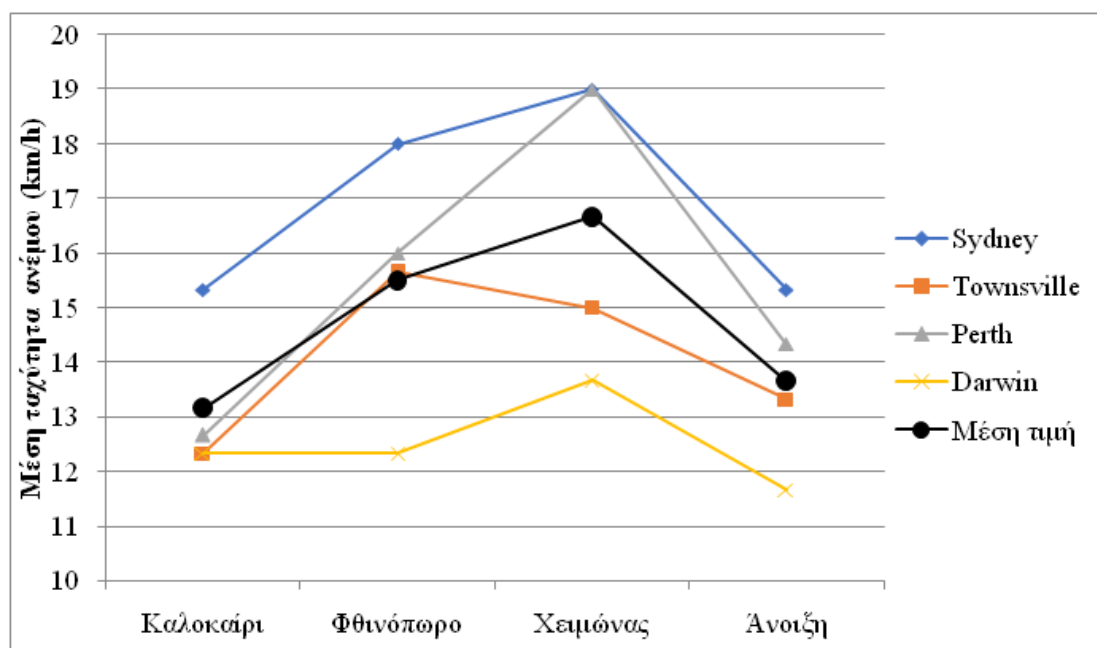
**Σχήμα 6.14:** Η μέση ταχύτητα ανέμου ανά μήνα για τέσσερις πόλεις της Αυστραλίας.

Πίνακας 6.13: Σύνοψη των αποτελεσμάτων της εποχικής ανάλυσης για το μοντέλο LUBE–MOPSO.

Μοντέλο	LUBE – MOPSO			
	Καλοκαίρι	Φθινόπωρο	Χειμώνας	Άνοιξη
Διάμεσος CWC για όλες τις ζώνες	0.4962319	0.4622746	0.4734749	0.4651082
Τυπική απόκλιση CWC για όλες τις ζώνες	0.0562	0.0586	0.063	0.0473
Πλήθος ζωνών με την καλύτερη απόδοση	2	4	2	2
Πλήθος ζωνών με τη δεύτερη καλύτερη απόδοση	2	4	3	1
Πλήθος ζωνών με τη χειρότερη απόδοση	6	2	1	1
Πλήθος ζωνών με τη δεύτερη χειρότερη απόδοση	0	0	4	6

Χαμηλότερες κατά μέσο όρο ταχύτητες ανέμου αντιστοιχούν σε μεγαλύτερα χρονικά διαστήματα μη λειτουργίας μιας ανεμογεννήτριας ή λειτουργίας στην περιοχή της μη γραμμικής σχέσης ισχύος και ταχύτητας ανέμου, επομένως παρατηρείται μεγαλύτερη αστάθεια στα αποτελέσματα, αύξηση της αβεβαιότητας και παραγωγή χειρότερων διαστημάτων πρόβλεψης. Έτσι εξηγείται εν μέρει η χειρότερη απόδοση των μοντέλων για δεδομένα του καλοκαιριού και η καλύτερη απόδοσή τους για δεδομένα του φθινοπώρου.

Σε ότι αφορά την τυπική απόκλιση των διαμέσων τιμών του CWC ανά ζώνη για κάθε εποχή, τα καλύτερα αποτελέσματα και για τα δύο μοντέλα παρατηρούνται την άνοιξη, και δευτερευόντως το καλοκαίρι. Επομένως, την άνοιξη και το καλοκαίρι η απόδοση των



Σχήμα 6.14: Η μέση ταχύτητα ανέμου ανά εποχή για τέσσερις πόλεις της Αυστραλίας. μοντέλων εξαρτάται λιγότερο από τη γεωγραφική τοποθεσία του αιολικού σταθμού. Αντίθετα, το φθινόπωρο και τον χειμώνα παρατηρείται μεγαλύτερη διακύμανση τιμών από

ζώνη σε ζώνη. Οι μεγαλύτερες ταχύτητες ανέμου που επικρατούν κατά μέσο όρο τον χειμώνα και το φθινόπωρο σε κάποιες γεωγραφικές περιοχές εκδηλώνονται έντονα, ενώ σε άλλες δεν εκδηλώνονται καθόλου. Επομένως, το φθινόπωρο και τον χειμώνα η απόδοση των μοντέλων παρουσιάζει μεγαλύτερη εξάρτηση από τη γεωγραφική τοποθεσία του αιολικού σταθμού.

Τα συμπεράσματα από την εποχική ανάλυση της παρούσας διπλωματικής εργασίας είναι χαμηλού επιπέδου, με την έννοια ότι αφορούν μόνο τη σύγκριση της αποδοτικότητας μεταξύ των εποχών και ότι δεν μπορούν να γενικευτούν με βεβαιότητα. Η αδυναμία γενίκευσης έχει να κάνει κυρίως με το γεγονός ότι τα δεδομένα εκπαίδευσης και αξιολόγησης αφορούν μόλις ένα έτος. Εάν υπήρχαν διαθέσιμα δεδομένα και για άλλα έτη για τους ίδιους αιολικούς σταθμούς, τα συμπεράσματα θα είχαν μεγαλύτερη αξιοπιστία και θα μπορούσαν να γενικευτούν με μεγαλύτερη βεβαιότητα. Η εξαγωγή συμπερασμάτων για τυχόν συσχετίσεις μεταξύ των εποχών και την παρουσία περιοδικών φαινομένων εντός των δεδομένων κάθε εποχής απαιτεί περισσότερες πληροφορίες για τον εκάστοτε αιολικό σταθμό καθώς και χρήση στατιστικών εργαλείων ανώτερου επιπέδου. Στην παρούσα διπλωματική εργασία, τα στατιστικά εργαλεία που χρησιμοποιούνται για την εποχική ανάλυση είναι απλά. Ταυτόχρονα τα δεδομένα που χρησιμοποιούνται αφορούν μόνο τις συνιστώσες του διανύσματος του ανέμου σε δύο διαφορετικά ύψη, ενώ πολλές πληροφορίες όπως η ακριβής γεωγραφική τοποθεσία των αιολικών σταθμών είναι άγνωστες.

ΕΠΙΛΟΓΟΣ

7.1 ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα διπλωματική εργασία αναπτύχθηκαν δύο μοντέλα που βασίζουν τη λειτουργία τους σε τεχνητά νευρωνικά δίκτυα με σκοπό την πρόβλεψη αιολικής παραγωγής. Η πρόβλεψη είναι πιθανοτική, με την έννοια ότι αντί για σημειακές προβλέψεις παράγονται διαστήματα πρόβλεψης. Τα μοντέλα της παρούσας διπλωματικής εργασίας βασίζονται στη μέθοδο εκτίμησης κάτω και άνω ορίου (Lower Upper Bound Estimation – LUBE), δηλαδή οι οριακές τιμές των παραγόμενων διαστημάτων πρόβλεψης προκύπτουν απευθείας από το τεχνητό νευρωνικό δίκτυο, χωρίς να μεσολαβεί κάποια επιπρόσθετη στατιστική διαδικασία.

Κατά την εκπαίδευση των μοντέλων τα βάρη και οι πολώσεις των τεχνητών νευρωνικών δικτύων αναπροσαρμόζονται κατάλληλα με στόχο τη βελτιστοποίηση των παραγόμενων διαστημάτων πρόβλεψης. Η εκπαίδευση του πρώτου μοντέλου γίνεται στη βάση μίας μόνο αντικειμενικής συνάρτησης, της ελαχιστοποίησης του κριτηρίου του εύρους κάλυψης (Coverage Width based Criterion – CWC). Ως αλγόριθμος βελτιστοποίησης επιλέγεται ο μεταερευτικός αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (Particle Swarm Optimization – PSO). Το πρώτο μοντέλο ονομάζεται LUBE–PSO–CWC. Η εκπαίδευση του δεύτερου μοντέλου γίνεται σε πολυκριτηριακά πλαίσια. Οι αντικειμενικές συναρτήσεις που επιλέγονται είναι δύο, η μεγιστοποίηση της πιθανότητας κάλυψης του διαστήματος πρόβλεψης (Prediction Interval Coverage Probability – PICP) και η ελαχιστοποίηση της κανονικοποιημένης ρίζας της μέσης τιμής του τετραγώνου του εύρους των διαστημάτων πρόβλεψης (Prediction Interval Normalized Root mean square Width – PINRW). Ως αλγόριθμος βελτιστοποίησης επιλέγεται η πολυκριτηριακή βελτιστοποίηση σμήνους σωματιδίων (Multi Objective Particle Swarm Optimization – MOPSO). Το δεύτερο μοντέλο ονομάζεται LUBE–MOPSO. Μέχρι σήμερα, η μέθοδος LUBE σε πολυκριτηριακά πλαίσια δεν έχει χρησιμοποιηθεί με σκοπό την πρόβλεψη αιολικής παραγωγής. Για την αρχικοποίηση των βαρών και των πολώσεων των δύο μοντέλων επιλέγεται η τεχνική αρχικοποίησης Nguyen–Widrow. Και στα δύο μοντέλα εφαρμόζεται γκαουσιανός τελεστής μετάλλαξης, με σχετικά μεγάλη αρχική πιθανότητα μετάλλαξης που μειώνεται γραμμικά όσο προχωρούν οι επαναλήψεις. Ο καθορισμός των παραμέτρων των δύο μοντέλων γίνεται με βάση τα αποτελέσματα της μεθόδου βελτιστοποίησης παραμέτρων που επιλέγεται.

Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και αξιολόγηση των μοντέλων LUBE–PSO–CWC και LUBE–MOPSO προέρχονται από τον διεθνή διαγωνισμό Global Energy Forecasting Competition 2014 (GEFCom 2014) και είναι δημόσια, επομένως δεν υπάρχουν περιορισμοί σε ότι αφορά την πρόσβαση σε αυτά. Τα δεδομένα αφορούν δέκα διαφορετικούς αιολικούς σταθμούς. Για κάθε αιολικό σταθμό γίνεται διαχωρισμός των

δεδομένων σε εποχές (καλοκαίρι, φθινόπωρο, χειμώνας, άνοιξη) ώστε να μπορεί να συντελεστεί εποχική ανάλυση.

Το μοντέλο που χρησιμοποιείται ως σημείο αναφοράς είναι το Bootstrap Extreme Learning Machine (BELM). Η απόδοση των μοντέλων LUBE-PSO-CWC και LUBE-MOPSO συγκρίνεται με την απόδοση του μοντέλου BELM. Στη συνέχεια γίνεται εποχική ανάλυση στα αποτελέσματα των μοντέλων LUBE-PSO-CWC και LUBE-MOPSO και εξάγονται συμπεράσματα για κάθε εποχή.

Το μοντέλο LUBE-MOPSO παρουσιάζει συνολικά καλύτερη απόδοση σε σχέση με τα μοντέλα LUBE-PSO-CWC και BELM. Τα διαστήματα πρόβλεψης που παράγει έχουν το μικρότερο μέσο εύρος, ικανοποιώντας ταυτόχρονα το ονομαστικό επίπεδο εμπιστοσύνης σε ότι αφορά την πιθανότητα κάλυψης. Τα αποτελέσματα αυτά ισχύουν ανεξαρτήτως αιολικού σταθμού και εποχής, επομένως το μοντέλο LUBE-MOPSO παρουσιάζει και τη μεγαλύτερη σταθερότητα. Το μοντέλο LUBE-MOPSO είναι ταχύτερο σε σχέση με το μοντέλο LUBE-PSO-CWC. Ταυτόχρονα δίνει μεγαλύτερη ευελιξία στον χρήστη καθώς παράγει πολλαπλές λύσεις στο τέλος κάθε εκπαίδευσης. Το μοντέλο LUBE-PSO-CWC παρουσιάζει καλύτερη απόδοση σε σχέση με το μοντέλο BELM. Τα διαστήματα πρόβλεψης που παράγει έχουν τη μεγαλύτερη πιθανότητα κάλυψης, ωστόσο είναι το μοντέλο με το μεγαλύτερο χρόνο εκτέλεσης. Το μοντέλο BELM που παρουσιάζει τα χειρότερα αποτελέσματα, είναι πολύ ταχύτερο σε σχέση με τα μοντέλα LUBE-MOPSO και LUBE-PSO-CWC της παρούσας διπλωματικής εργασίας.

Κατά την εποχική ανάλυση συγκρίνονται τα αποτελέσματα των μοντέλων LUBE-MOPSO και LUBE-PSO-CWC ανά εποχή, για όλους τους αιολικούς σταθμούς. Τα χειρότερα αποτελέσματα παρατηρούνται το καλοκαίρι, ενώ τα καλύτερα το φθινόπωρο. Ωστόσο, τα αποτελέσματα μεταξύ φθινοπώρου, χειμώνα και άνοιξης έχουν μικρή απόκλιση. Η σχέση αυτή μεταξύ των αποτελεσμάτων κάθε εποχής πιθανώς σχετίζεται με την μέση ταχύτητα του ανέμου ανά εποχή. Το καλοκαίρι που η μέση ταχύτητα του ανέμου είναι μικρότερη σε σχέση με το φθινόπωρο και τον χειμώνα, μια ανεμογεννήτρια βρίσκεται περισσότερο χρόνο στη μη γραμμική περιοχή λειτουργίας, επομένως η παραγωγή ισχύος παρουσιάζει μεγαλύτερη μεταβλητότητα. Το φθινόπωρο και τον χειμώνα τα αποτελέσματα παρουσιάζουν μεγαλύτερη αστάθεια σε σχέση με την άνοιξη και το καλοκαίρι. Το γεγονός αυτό οφείλεται πιθανώς στο ότι το καλοκαίρι και την άνοιξη οι καιρικές συνθήκες παρουσιάζουν μεγαλύτερη ομοιογένεια ανεξαρτήτως της γεωγραφικής τοποθεσίας των δέκα αιολικών σταθμών.

Συμπερασματικά, τα μοντέλα LUBE-MOPSO και LUBE-PSO-CWC παρουσιάζουν απόδοση η οποία είναι άμεσα συγκρίσιμη με την απόδοση άλλων πιθανοτικών μοντέλων της διαθέσιμης βιβλιογραφίας. Το μοντέλο LUBE-MOPSO και γενικότερα η πιθανοτική πρόβλεψη αιολικής παραγωγής με χρήση της μεθόδου LUBE σε πολυκριτηριακά πλαίσια, παρουσιάζει μεγάλη δυναμική με πολλές προοπτικές επέκτασης και βελτίωσης. Ωστόσο η επιτυχημένη πρόβλεψη της αιολικής παραγωγής και η βέλτιστη ερμηνεία και αξιοποίηση των διαστημάτων πρόβλεψης συνεχίζουν να αποτελούν πρόκληση. Το πρόβλημα της μεγάλης μεταβλητότητας της αιολικής παραγωγής λόγω της εξάρτησης από την τυχαία φύση του ανέμου, αλλά και από πληθώρα άλλων παραγόντων, αναδεικνύει τα όρια που υπάρχουν σε ότι αφορά τη βελτίωση της πρόβλεψης της αιολικής παραγωγής. Σε ότι αφορά την εποχική ανάλυση εξάγονται ορισμένα συμπεράσματα γύρω από την απόδοση ανά εποχή, ωστόσο τα

συμπεράσματα αυτά είναι δύσκολο να γενικευτούν με βεβαιότητα. Η δυσκολία αυτή σχετίζεται με την ποσότητα των διαθέσιμων δεδομένων. Η πρόσβαση σε περισσότερα δεδομένα, που αφορούν πολλά έτη και εμπεριέχουν περισσότερα μεγέθη που σχετίζονται με την αιολική παραγωγή θα αυξήσει την ικανότητα γενίκευσης των συμπερασμάτων και θα επιτρέψει την εποχική ανάλυση υψηλότερου επιπέδου. Τα συμπεράσματα της παρούσας διπλωματικής εργασίας από την εποχική ανάλυση μπορούν ωστόσο να αποτελέσουν τη βάση για περαιτέρω μελλοντική μελέτη γύρω από τη συσχέτιση εποχής και αιολικής παραγωγής.

7.2 ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Μία πιθανή επέκταση της παρούσας διπλωματικής εργασίας αποτελεί η κατάλληλη τροποποίηση του αλγορίθμου πολυκριτηριακής βελτιστοποίησης σμήνους σωματιδίων (Multi Objective Particle Swarm Optimization - MOPSO). Οι πολυκριτηριακοί αλγόριθμοι έχουν χρησιμοποιηθεί ελάχιστα για την εκπαίδευση μοντέλων πιθανοτικής πρόβλεψης αιολικής παραγωγής. Ο αλγόριθμος MOPSO που αναπτύσσεται στην παρούσα διπλωματική εργασία αντλεί τα χαρακτηριστικά του από διάφορες παραλλαγές του αλγορίθμου πολυκριτηριακής βελτιστοποίησης σμήνους σωματιδίων που συναντώνται στη διαθέσιμη βιβλιογραφία. Ωστόσο, οι παραλλαγές αυτές αφορούν άλλα αντικείμενα, που δεν έχουν σχέση με την πρόβλεψη αιολικής παραγωγής. Επομένως, με την κατάλληλη τροποποίηση του αλγορίθμου και τις κατάλληλες προσθήκες, όπως η εισαγωγή ενός τελεστή διασταύρωσης (crossover operator), ο αλγόριθμος μπορεί να προσαρμοστεί με τον βέλτιστο τρόπο στο πρόβλημα της πρόβλεψης της αιολικής παραγωγής, βελτιώνοντας τα αποτελέσματα.

Η επέκταση της παρούσας διπλωματικής εργασίας θα μπορούσε να αφορά και την επεξεργασία των αιολικών δεδομένων. Με την προσθήκη μιας παρελθοντικής χρονοσειράς των τιμών της παραγόμενης αιολικής ισχύος στα δεδομένα εισόδου, είναι δυνατή η εφαρμογή του μετασχηματισμού κυματιδίων (wavelet transformation). Η προσθήκη των παρελθοντικών τιμών παραγόμενης αιολικής ισχύος στα δεδομένα εισόδου μπορεί να οδηγήσει σε αύξηση της απόδοσης του μοντέλου, ανάλογα με τη συσχέτιση που παρουσιάζουν οι παρελθοντικές και οι μελλοντικές τιμές αιολικής ισχύος. Η εφαρμογή του μετασχηματισμού κυματιδίων αναλύει τη χρονοσειρά της αιολικής ισχύος σε συνιστώσες, καθαρίζοντας τον θόρυβο που εμπεριέχεται σε αυτή, επομένως διευκολύνει τη διαδικασία αναγνώρισης της συσχέτισης μεταξύ παρελθοντικών και μελλοντικών τιμών της αιολικής ισχύος από το τεχνητό νευρωνικό δίκτυο.

Μία ακόμα πιθανή επέκταση της παρούσας διπλωματικής εργασίας είναι η εισαγωγή της έννοιας της χωροχρονικής συσχέτισης (spatio-temporal correlation). Η χωροχρονική συσχέτιση έγκειται στη διερεύνηση της συσχέτισης των αιολικών δεδομένων δύο ή περισσότερων γειτονικών αιολικών σταθμών σε συνάρτηση με τον χρόνο. Αφού διερευνηθεί η ύπαρξη και το είδος της συσχέτισης μεταξύ των υπό εξέταση αιολικών σταθμών, για κάθε αιολικό σταθμό γίνεται τροφοδότηση αιολικών δεδομένων που αφορούν γειτονικούς αιολικούς σταθμούς στο τεχνητό νευρωνικό δίκτυο, σε συνδυασμό με τα αιολικά δεδομένα που αφορούν τον υπό εξέταση αιολικό σταθμό. Κατά αυτόν τον τρόπο αυξάνονται τα δεδομένα εισόδου από τα οποία εξαρτάται η παραγόμενη αιολική ισχύς ενός αιολικού σταθμού, που κατά πάσα πιθανότητα θα οδηγήσει και σε βελτίωση των παραγόμενων

διαστημάτων πρόβλεψης. Τέλος, με τη χωροχρονική συσχέτιση είναι πιθανό να εξαχθούν περαιτέρω συμπεράσματα σχετικά με την εποχική ανάλυση, με μεγαλύτερη ικανότητα γενίκευσης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

7.1 ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, “Lower upper bound estimation method for construction of neural network-based prediction intervals,” *IEEE Trans. Neural Networks*, vol. 22, no. 3, pp. 337–346, 2011, doi: 10.1109/TNN.2010.2096824.
- [2] J. Kennedy and R. Eberhart, “Particle swarm optimization,” *Proc. ICNN’95 - Int. Conf. Neural Networks*, doi: 10.1109/icnn.1995.488968.
- [3] C. A. Coello Coello and M. S. Lechuga, “MOPSO: A proposal for multiple objective particle swarm optimization,” *Proc. 2002 Congr. Evol. Comput. CEC 2002*, vol. 2, pp. 1051–1056, 2002, doi: 10.1109/CEC.2002.1004388.
- [4] D. Nguyen and B. Widrow, “Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights,” *IJCNN. Int. Jt. Conf. Neural Networks*, pp. 21–26, 1990, doi: 10.1109/ijcnn.1990.137819.
- [5] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, “Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 896–913, 2016, doi: 10.1016/j.ijforecast.2016.02.001.
- [6] C. Wan, Z. Xu, P. Pinson, Z. Yang Dong, and K. P. Wong, “Probabilistic Forecasting of Wind Power Generation Using Extreme Learning Machine,” *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044.
- [7] H. zhi Wang, G. qiang Li, G. bing Wang, J. chun Peng, H. Jiang, and Y. tao Liu, “Deep learning based ensemble approach for probabilistic wind power forecasting,” *Appl. Energy*, vol. 188, pp. 56–70, 2017, doi: 10.1016/j.apenergy.2016.11.111.
- [8] H. Wen, J. Gu, J. Ma, and Z. Jin, “Probabilistic wind power forecasting via bayesian deep learning based prediction intervals,” *IEEE Int. Conf. Ind. Informatics*, vol. 2019-July, pp. 1091–1096, 2019, doi: 10.1109/INDIN41052.2019.8972125.
- [9] N. A. Shrivastava, A. Khosravi, and B. K. Panigrahi, “Prediction Interval Estimation of Electricity Prices Using PSO-Tuned Support Vector Machines,” *IEEE Trans. Ind. Informatics*, vol. 11, no. 2, pp. 322–331, 2015, doi: 10.1109/TII.2015.2389625.
- [10] Y. Wu, P. Su, T. Wu, J. Hong, and M. Y. Hassan, “Probabilistic Wind-Power Forecasting Using,” vol. 54, no. 6, pp. 5609–5620, 2018.
- [11] Z. Shi, H. Liang, and V. Dinavahi, “Direct Interval Forecast of Uncertain Wind Power Based on Recurrent Neural Networks,” *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1177–1187, 2018, doi: 10.1109/TSTE.2017.2774195.
- [12] G. Zhang, Z. Li, K. Zhang, L. Zhang, X. Hua, and Y. Wang, “Multi-objective interval prediction of wind power based on conditional copula function,” *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 4, pp. 802–812, 2019, doi: 10.1007/s40565-018-0495-0.
- [13] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, “A new fuzzy-based combined prediction interval for wind power forecasting,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 18–26, 2016, doi: 10.1109/TPWRS.2015.2393880.
- [14] P. S. Georgilakis, *Spotlight on Modern Transformer Design (Power Systems)*, vol. 38. Springer, 2009.

- [15] M. G. Anderson D, “Artificial Neural Networks Technology,” *Kaman Sci. Corp.*, vol. 258, no. 6, pp. 1–83, 1992.
- [16] G. Montavon, “Introduction to Neural Networks,” *Lect. Notes Phys.*, vol. 968, no. November, pp. 37–62, 2020, doi: 10.1007/978-3-030-40245-7_4.
- [17] Π. Σ. Γεωργιλάκης, *Σύγχρονα Συστήματα Μεταφοράς και Διανομής Ηλεκτρικής Ενέργειας*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΕΑΒ), 2015.
- [18] Mohammed Zeeshan Mulla, “Cost, Activation, Loss Function|| Neural Network|| Deep Learning. What are these?” [Online]. Available: <https://medium.com/@zeeshanmulla/cost-activation-loss-function-neural-network-deep-learning-what-are-these-91167825a4de>.
- [19] M. L. Minsky and S. Papert, *Perceptrons*. Amsterdam University Press, 1972.
- [20] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0.
- [21] X. S. Yang, *Nature-inspired Metaheuristic Algorithms*. Amsterdam University Press, 2010.
- [22] E. Aarts and J. Korst, “Chapter 2 Simulated annealing 2.1 Introduction of the algorithm,” *Simulated Annealing Theory Appl.*, p. 7, 1987.
- [23] D. Whitley, “A genetic algorithm tutorial,” *Stat. Comput.*, vol. 4, no. 2, pp. 65–85, 1994, doi: 10.1007/BF00175354.
- [24] S. Fidanova, “Ant Colony Optimization,” *Stud. Comput. Intell.*, vol. 947, no. November, pp. 3–8, 2021, doi: 10.1007/978-3-030-67380-2_2.
- [25] X. S. Yang, “A new metaheuristic Bat-inspired Algorithm,” *Stud. Comput. Intell.*, vol. 284, pp. 65–74, 2010, doi: 10.1007/978-3-642-12538-6_6.
- [26] G. T. Pulido and C. A. Coello Coello, “A constraint-handling mechanism for particle swarm optimization,” *Proc. 2004 Congr. Evol. Comput. CEC2004*, vol. 2, no. 4, pp. 1396–1403, 2004, doi: 10.1109/cec.2004.1331060.
- [27] H. Quan, D. Srinivasan, and A. Khosravi, “Particle swarm optimization for construction of neural network-based prediction intervals,” *Neurocomputing*, vol. 127, pp. 172–180, 2014, doi: 10.1016/j.neucom.2013.08.020.
- [28] K. Deb, *Multi-objective optimization using evolutionary algorithms*. Wiley, 2005.
- [29] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002, doi: 10.1109/4235.996017.
- [30] C. Mac Domhnaill and L. Ryan, “Towards renewable electricity in Europe: Revisiting the determinants of renewable electricity in the European Union,” *Renew. Energy*, vol. 154, pp. 955–965, 2020, doi: 10.1016/j.renene.2020.03.084.
- [31] *OECD Factbook*. Organisation for Economic Co-operation and Development, 2016.
- [32] I. González-Aparicio and A. Zucker, “Impact of wind power uncertainty forecasting on the market integration of wind energy in Spain,” *Appl. Energy*, vol. 159, pp. 334–349, 2015, doi: 10.1016/j.apenergy.2015.08.104.
- [33] “Wind energy in Europe 2020 Statistics and the outlook for 2021–2025,” 2021. [Online]. Available: <https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-in-2020-trends-and-statistics/>.
- [34] V. Sohoni, S. C. Gupta, and R. K. Nema, “A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems,” *J. Energy*, vol. 2016, no. region 4, pp. 1–18, 2016, doi: 10.1155/2016/8519785.
- [35] Κ. Παπαθανασίου, “Πρόβλεψη αιολικής και φωτοβολταϊκής παραγωγής με χρήση τεχνικών μηχανικής μάθησης,” Εθνικό Μετσόβιο Πολυτεχνείο, 2021.
- [36] G. Boyle, *Renewable Electricity and the Grid*. Earthscan, 2007.
- [37] P. Pinson, “Estimation of the uncertainty in wind power forecasting,” *Paristechbibrilkcom*, p. 248, 2006.
- [38] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating*

- Renewables in Electricity Markets: Operational Problems (International Series in Operations Research & Management Science, 205)*, 2014th ed. Springer, 2013.
- [39] M. H. El-Ahmar, A. H. M. El-Sayed, and A. M. Hemeida, "Evaluation of factors affecting wind turbine output power," *2017 19th Int. Middle-East Power Syst. Conf. MEPCON 2017 - Proc.*, vol. 2018-Febru, no. December, pp. 1471–1476, 2018, doi: 10.1109/MEPCON.2017.8301377.
- [40] Z. Pu and E. Kalnay, *Numerical Weather Prediction Basics: Models, Numerical Methods, and Data Assimilation*. 2018.
- [41] S. Li, D. C. Wunsch, S. Member, E. A. O. Hair, M. G. Giesselmann, and S. Member, "Power generation," *Consult. Specif. Eng.*, vol. 29, no. 6, p. 221, 2001, doi: 10.1007/978-981-32-9938-2_2.
- [42] Timo Stöttner, "Why Data should be Normalized before Training a Neural Network," 2019. [Online]. Available: <https://towardsdatascience.com/why-data-should-be-normalized-before-training-a-neural-network-c626b7f66c7d>.
- [43] S. G. K. Patro and K. K. sahu, "Normalization: A Preprocessing Stage," *Iarjset*, pp. 20–22, 2015, doi: 10.17148/iarjset.2015.2305.
- [44] G. Thimm and E. Fiesler, "Neural network initialization," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 930, pp. 535–542, 1995, doi: 10.1007/3-540-59497-3_220.
- [45] M. Srinivas and L. M. Patnaik, "Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms," *IEEE Trans. Syst. Man Cybern.*, vol. 24, no. 4, pp. 656–667, 1994, doi: 10.1109/21.286385.
- [46] P. S. Andrews, "An investigation into mutation operators for particle swarm optimization," *2006 IEEE Congr. Evol. Comput. CEC 2006*, pp. 1044–1051, 2006, doi: 10.1109/cec.2006.1688424.
- [47] N. Higashi and H. Iba, "Particle swarm optimization with Gaussian mutation," *2003 IEEE Swarm Intell. Symp. SIS 2003 - Proc.*, pp. 72–79, 2003, doi: 10.1109/SIS.2003.1202250.
- [48] C. A. Coello Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 256–279, 2004, doi: 10.1109/TEVC.2004.826067.
- [49] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimization for feature selection in classification: A multi-objective approach," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1656–1671, 2013, doi: 10.1109/TSMCB.2012.2227469.
- [50] Stone M., "Cross-Validatory Choice and Assessment of Statistical Predictions," *J. R. Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 111–147, 1974.
- [51] Γ. Δ. Σκεύης, "Πιθανοτική Πρόβλεψη Αιολικής Παραγωγής με τη Χρήση Extreme Learning Machines," Εθνικό Μετσόβιο Πολυτεχνείο, 2020.
- [52] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, pp. 489–501, 2006.
- [53] D. A. Freedman, "Bootstrapping Regression Models," *Ann. Stat.*, vol. 9, no. 6, pp. 1218–1228, 1981.
- [54] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "Transfer learning for time series classification," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, pp. 1367–1376, 2019, doi: 10.1109/BigData.2018.8621990.
- [55] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," *33rd Int. Conf. Mach. Learn. ICML 2016*, vol. 3, pp. 1651–1660, 2016.
- [56] "timeanddate.com." [Online]. Available: timeanddate.com/weather/australia.