



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Χημικών Μηχανικών
Τομέας II – Ανάλυσης, Σχεδιασμού και Ανάπτυξης Διεργασιών
και Συστημάτων

Ανάπτυξη ρεαλιστικών μοντέλων μειωμένης τάξης για μη γραμμικές διεργασίες χημικής απόθεσης από ατμό

Διπλωματική Εργασία
της
Ελισάβετ Άγγλου

Επιβλέπων Καθηγητής
Ανδρέας Μπουντουβής

Αθήνα, 2017

Ευχαριστίες

Με την ολοκλήρωση της εργασίας αυτής θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή Ανδρέα Μπουντουβή για την ευκαιρία που έδωσε να εκπονήσω τη διπλωματική μου εργασία υπό την επίβλεψη του ιδίου και της ερευνητικής του ομάδας. Ακόμα, θα ήθελα να τον ευχαριστήσω για την εμπιστοσύνη που μου έδειξε αναθέτοντας μου το συγκεκριμένο θέμα τον περασμένο Ιανουάριο, καθώς και για τις εκτενείς διορθώσεις του στην παρούσα εργασία.

Δεν θα μπορούσα να μην ευχαριστήσω τη Δρα Ελένη Κορωνάκη για τη συνεχή και πολύτιμη καθοδήγηση της από την αρχή ως το τέλος της εργασίας αυτής, η οποία είναι αποτέλεσμα της άρτιας συνεργασίας μου με εκείνη και την ευχαριστώ για όλα. Θα ήθελα να την ευχαριστήσω για τη βοήθεια της πάνω στο κομμάτι της υπολογιστικής ρευστοδυναμικής και στο σχεδιασμό του μειωμένης τάξης μοντέλου καθώς και για την επιμονή της στην επίτευξη του στόχου, την εκπόνηση της διπλωματικής εργασίας.

Επιπροσθέτως, θα ήθελα να ευχαριστήσω τον Θοδωρή Παπαπέτρου για τις γνώσεις που μου έδωσε μέσω των γραπτών του κατά την εφαρμογή της ανάλυσης κύριων συνιστωσών και στην εκπαίδευση των νευρωνικών δικτύων. Επίσης, χωρίς τα υπολογιστικά του δεδομένα καθώς και τις διερευνήσεις που πραγματοποίησε, η παρούσα εργασία δεν θα μπορούσε να ολοκληρωθεί. Η διπλωματική μου εργασία αποτελεί συνέχεια της δικής του δουλειάς, και δεν μπορώ παρά να τον ευχαριστήσω.

Επίσης, ευχαριστώ τον υποψήφιο διδάκτορα Πάυλο Γκίνη, για τη συνεχή παρακολούθηση της εργασίας, καθώς και για τις γνώσεις τις οποίες μου παρείχε στο κομμάτι των χημικών αντιδράσεων παρόλο που τελικά δεν καταφέραμε να τις χρησιμοποιήσουμε στο έπακρο.

Τέλος, το μεγαλύτερο ευχαριστώ οφείλω στους δικούς μου ανθρώπους που ήταν συνεχώς δίπλα μου από το πρώτο έτος έως το πέρας της φοιτητικής μου ζωής στο ΕΜΠ, τους γονείς μου, τις αδερφές μου και το φίλο μου Άλκη που με βοήθησαν να είμαι προσηλωμένη στους στόχους μου. Συγχρόνως, ένα μεγάλο ευχαριστώ σε όλους τους φίλους που έκαναν αυτήν την 5ετή πορεία πιο ευχάριστη και ενδιαφέρουσα.

Περιεχόμενα

1.	Εισαγωγή.....	8
2.	Φυσικό Σύστημα	10
2.1.	Χημική απόθεση από ατμό	10
2.2.	Μαθηματική Μοντελοποίηση Προβλήματος	12
2.3.	Επίλυση προβλήματος	16
2.4.	Πολλαπλότητα Χώρου Λύσεων.....	17
3.	Μείωση Τάξης Μοντέλου	21
3.1	Εισαγωγή	21
3.2	Μείωση Τάξης Δυναμικών Συστημάτων	22
3.3	Μέθοδος Galerkin	23
3.4	Ανάλυση Κύριων Συνιστωσών.....	24
3.4.1	Κατασκευή Βάσης PCA	24
3.4.2	Επιλογή Διάστασης Βάσης	25
3.5	Μέθοδος των Στιγμιούπων.....	26
3.5.1	Κανονικοποίηση Δεδομένων	28
3.6	Αξιολόγηση της μείωσης τάξης – Σφάλματα	29
4.	Σχεδιασμός Νευρωνικού Δικτύου.....	32
4.1	Εισαγωγή	32
4.2	Βιολογικά Νευρωνικά Δίκτυα	32
4.3	Τεχνητά Νευρωνικά Δίκτυα.....	34
4.4	Μαθηματική Μοντελοποίηση Τεχνητού Νευρωνικού Δικτύου	34
4.5	Αρχιτεκτονική Τεχνητού Νευρωνικού Δικτύου.....	36
4.5.1	Δίκτυο Πρόσθιας Τροφοδότησης.....	37
4.5.2	Δίκτυο με Ανατροφοδότηση - Δίκτυο NARX	38
4.6	Εκπαίδευση Νευρωνικού Δικτύου	39
4.6.1	Μέθοδοι εκπαίδευσης	39
4.7	Εφαρμογή.....	41
4.8	Αξιολόγηση Νευρωνικού Δικτύου – Σφάλματα	42
4.9	Συμπεράσματα	43
5.	Χημικές Αντιδράσεις	45
5.1	Χημικό Σύστημα	45
5.2	Εφαρμογή.....	47
6.	Αποτελέσματα.....	49
6.1	Συλλογή Αποτελεσμάτων	49
6.2	Μείωση Τάξης.....	50

6.3	Σχεδιασμός Ορθοκανονικής Βάσης.....	50
6.3.1	Διερεύνηση 1: Τυχαία επιλογή βηματικών επιβολών από όλο το εύρος λύσεων για την κατασκευή της βάσης.....	51
6.3.2	Διερεύνηση 2: Επιλογή βηματικών επιβολών από τους 2 διαφορετικούς κλάδους πολλαπλότητας για την κατασκευή της Βάσης	56
6.3.3	Διερεύνηση 3: Εμπλουτισμός των Δεδομένων με Επιβολές Εναλλαγής μεταξύ των Κλάδων Πολλαπλότητας.....	59
6.4	Σχεδιασμός Νευρωνικού Δικτύου	62
6.4.1	Εκπαίδευση δικτύου	62
6.4.2	Αξιολόγηση Νευρωνικών Δικτύων.....	64
6.5	Μελέτη Χημικού Συστήματος	72
7.	Συμπεράσματα	77
8.	Προτάσεις για Περαιτέρω Έρευνα	79
9.	Βιβλιογραφία	81

Περίληψη

Στο πλαίσιο της παρούσας εργασίας μελετήθηκε η δυναμική συμπεριφορά ενός αντιδραστήρα χημικής απόθεσης από ατμό (ΧΑΑ) για την παραγωγή λεπτών υμενίων. Το σύστημα μη γραμμικών εξισώσεων που περιγράφουν τα φαινόμενα τα οποία επικρατούν επιλύεται με χρήση του εμπορικού κώδικα υπολογιστικής ρευστοδυναμικής ANSYS Fluent 13. Οι εξισώσεις εμφανίζουν πολλαπλότητα λύσεων για τις ίδιες συνθήκες λειτουργίας του αντιδραστήρα, η οποία οφείλεται στην ύπαρξη και ανταγωνισμό διαφορετικών μηχανισμών ροής. Ειδικότερα, για παροχή εισόδου η οποία κυμαίνεται μεταξύ $2.3 \cdot 10^{-5} - 5 \cdot 10^{-5}$ kg/s μπορούν να υπολογισθούν τρεις καταστάσεις ισορροπίας, δύο ευσταθείς και μία ασταθής. Στον ένα ευσταθή κλάδο κυριαρχεί η εξαναγκασμένη ροή και στον άλλο ευσταθή κλάδο η φυσική κυκλοφορία λόγω άνωσης. Οι δύο κλάδοι ευσταθών λύσεων συνδέονται με έναν κλάδο ασταθών λύσεων. Εξαιτίας αυτής της ιδιαιτερότητας μελετήθηκε η απόκριση του συστήματος σε βηματικές επιβολές της μαζικής παροχής, οι οποίες πραγματοποιήθηκαν στην περιοχή πολλαπλότητας. Τα υπολογιστικά αυτά δεδομένα χρησιμοποιήθηκαν για το σχεδιασμό ενός μοντέλου μειωμένης τάξης, με χρήση της μεθόδου ανάλυσης κύριων συνιστωσών και με την εκπαίδευση τεχνητών νευρωνικών δικτύων. Το μοντέλο αυτό σχεδιάστηκε ούτως ώστε να έχει τη δυνατότητα προσομοίωσης της δυναμικής συμπεριφοράς του κάθε κλάδου ξεχωριστά καθώς και εναλλαγές μεταξύ αυτών. Από τη μελέτη αυτή, προκύπτει ένα μοντέλο μειωμένης τάξης το οποίο, αφού οικοδομηθεί και προκειμένου να χρησιμοποιηθεί, δεν απαιτεί καμία πρότερη γνώση των εξισώσεων και των μηχανισμών που κυριαρχούν στον αντιδραστήρα, για την περιγραφή της δυναμικής απόκρισής του μετά από την επιβολή μιας τυχαίας διαταραχής. Σε πρώτη φάση διερευνήθηκε το πλήθος των απαιτούμενων υπολογιστικών δεδομένων για το σχεδιασμό του μοντέλου μειωμένης τάξης, ούτως ώστε αυτό να μπορεί να περιγράψει με μεγάλη σχετική ακρίβεια την πραγματική κατάσταση στον αντιδραστήρα. Η επιλογή του καλύτερου μοντέλου πραγματοποιήθηκε με γνώμονα την ποιότητα της προσέγγισης σε συνδυασμό με τον «όγκο» της απαιτούμενης πληροφορίας. Με αυτό τον τρόπο επιλέχθηκε το μοντέλο το οποίο χρησιμοποιεί την ορθοκανονική βάση η οποία αποτελείται από τρεις ιδιοκαταστάσεις σε συνδυασμό με ένα νευρωνικό δίκτυο το οποίο περιέχει 12 κρυφά επίπεδα. Σημειώνεται ότι η ορθοκανονική βάση σχεδιάστηκε με χρήση 13 βηματικών επιβολών, ενώ το νευρωνικό δίκτυο εκπαιδεύτηκε με χρήση 26. Οι συγκεκριμένες επιβολές επιλέγονται κάθε φορά ούτως ώστε να περιγράφουν όλο το εύρος πολλαπλότητας, το οποίο μελετάται. Σε δεύτερη φάση, διερευνήθηκε η επάρκεια του μοντέλου μειωμένης τάξης που έχει σχεδιαστεί σε ότι αφορά στην πρόβλεψη του ρυθμού απόθεσης υμενίων πυριτίου (Si) από σιλάνιο (SiH₄). Προέκυψε, ότι η προσεγγιστική λύση του ρυθμού απόθεσης ελάχιστα διαφέρει από την ακριβή λύση, γεγονός το οποίο καθιστά το συγκεκριμένο μοντέλο ένα πολύ χρήσιμο εργαλείο για την πρόβλεψη των χαρακτηριστικών των παραγόμενων υμενίων.

Development of realistic reduced-order models for nonlinear chemical vapor deposition processes

Diploma thesis

by

Elisabeth Anglou

Abstract

This thesis is about the study of the dynamic behavior of a chemical vapor deposition reactor (CVD) used for the production of thin films. The equations describing the phenomena that occur in the reactor are solved using the ANSYS Fluent 13 computational fluid dynamics (CFD) code. The equations exhibit a multiplicity of solutions for the same operating conditions of the reactor due to the different competing flow mechanisms. In particular, three steady states, two stable and one unstable can be calculated for mass inlet flow between $2.3 \cdot 10^{-5}$ and $5 \cdot 10^{-5}$ kg / s. In one branch of stable solutions the dominant mechanism is forced convection, whereas in another branch of stable solutions free convection dominates through buoyancy. The two stable branches are connected with an unstable solutions branch. Data of the dynamic response of the detailed system were used to design a reduced order model using the principal component analysis (PCA) method combined with properly trained artificial neural networks (ANN). This model has been designed so that it can simulate the dynamic behavior of each branch separately as well as switches between the two. Once built, the reduced order model requires no prior knowledge of the equations and mechanisms that dominate in the reactor in order to describe its dynamic response after a step change of the mass flow. The number of computational data required for the design of the reduced order model was investigated so that it could accurately describe the actual state in the reactor. The choice of the best model was based on the quality of the approximation and the size of the required information. In this context, the model that was chosen includes a PCA basis with three eigenvectors combined with a neural network containing 12 hidden layers. It is noted that the PCA basis was designed using 13 step changes, while the neural network was trained using 26. The step changes are selected in order to describe the entire range of multiplicity being studied. In the last part of the work, the adequacy of the designed model in predicting the rate of deposition of silicon (Si) films from silane (SiH_4) was investigated. It was found that the approximate solution of the deposition rate is almost equal with the exact solution, which makes this model a very useful tool for predicting the characteristics of the produced films.

Εισαγωγή

Στο πλαίσιο της παρούσας εργασίας μελετάται η διεργασία της χημικής απόθεσης από ατμό, που παρουσιάζει μεγάλο ενδιαφέρον τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο, αφού βρίσκει ευρεία εφαρμογή στη βιομηχανία για την παραγωγή λεπτών υμενίων. Η διεργασία χαρακτηρίζεται από πολύπλοκα φαινόμενα μεταφοράς και σύνθετο δίκτυο χημικών αντιδράσεων, γεγονός το οποίο καθιστά απαραίτητη τη μελέτη του συστήματος για τη σωστή κατανόησή του, και συνεπώς την παραγωγή προϊόντος (υμενίων) επιθυμητών προδιαγραφών. Σύμφωνα με προηγούμενη έρευνα [1] [2], στο συγκεκριμένο σύστημα μπορούν να υπολογισθούν τρεις καταστάσεις ισορροπίας, δύο ευσταθείς και μία ασταθής για τις ίδιες συνθήκες λειτουργίας.

Η ρύθμιση του συστήματος αυτού είναι μια αρκετά δύσκολη έως αδύνατη διαδικασία, η οποία συντελείται σχεδόν εμπειρικά ή ακόμη και διαισθητικά από τον εκάστοτε χειριστή στη βιομηχανία. Η δυναμική μελέτη του συστήματος καθίσταται απαραίτητη για τη ρύθμιση του αντιδραστήρα και την παραγωγή του επιθυμητού προϊόντος.

Στόχος της συγκεκριμένης εργασίας είναι κατασκευή ενός μοντέλου μειωμένης τάξης το οποίο θα περιγράφει τη δυναμική συμπεριφορά του συστήματος για τους δύο ευσταθείς κλάδους μόνιμων καταστάσεων, καθώς και εναλλαγές μεταξύ αυτών. Το συγκεκριμένο πρόβλημα επιλύεται με χρήση ενός εμπορικού λογισμικού υπολογιστικής ρευστοδυναμικής (ANSYS Fluent 13) [3]. Τα λογισμικά αυτά, επιλύουν το εκάστοτε πρόβλημα με διακριτοποίηση του χωρίου λύσης, σε μικρές περιοχές, όπου επιλύονται διαδοχικά οι εξισώσεις που περιγράφουν το φυσικό σύστημα.

Η αναζητούμενη λύση, η οποία τελικά προκύπτει από τον κώδικα υπολογιστικής ρευστοδυναμικής, εξαρτάται εν γένει από το χώρο και το χρόνο. Στην πρώτη φάση κατασκευής του μοντέλου μειωμένης τάξης, χρησιμοποιείται η μέθοδος της ανάλυσης των κύριων συνιστωσών, και πιο συγκεκριμένα η μέθοδος στιγμιοτύπων, με στόχο τη μείωση της τάξης του μοντέλου, με υπολογισμό των ιδιοδιανυσμάτων τα οποία περιέχουν μεγάλο ποσοστό πληροφορίας. Από τα ιδιανύσματα αυτά, με κατάλληλους συντελεστές προκύπτει ως γραμμικός συνδυασμός, μια ικανοποιητική προσέγγιση της λύσης. Στη δεύτερη φάση, χρησιμοποιούνται τα νευρωνικά δίκτυα μέσω των οποίων υπολογίζονται οι χρονικά μεταβαλλόμενοι συντελεστές των ιδιανυσμάτων που εξασφαλίζουν μία προσέγγιση της δυναμικής συμπεριφοράς του μελετώμενου συστήματος. Αξίζει να σημειωθεί ότι η κατασκευή των παραπάνω μοντέλων πραγματοποιείται με γνώση μόνο των αποτελεσμάτων των υπολογιστικών προσομοιώσεων, δεν απαιτεί δηλαδή καμία γνώση των εξισώσεων που διέπουν το φυσικό σύστημα.

Στο τελικό στάδιο της εργασίας, υπολογίζεται η προσέγγιση των χαρακτηριστικών του παραγόμενου προϊόντος όπως προκύπτει από την εφαρμογή του μοντέλου μειωμένης τάξης. Πιο συγκεκριμένα, διερευνάται πόσο επιδρά στην πρόβλεψη του ρυθμού απόθεσης, ο προσεγγιστικός υπολογισμός του πεδίου ροής και θερμοκρασίας μέσα στον αντιδραστήρα. Το παράδειγμα που παρουσιάζεται αφορά στην απόθεση υμενίου πυριτίου (Si) από σιλάνιο (SiH_4).

Στο πρώτο κεφάλαιο της εργασίας περιγράφεται με λεπτομέρεια το φυσικό σύστημα και ο τρόπος επίλυσης του προβλήματος. Στη συνέχεια, αναπτύσσεται το θεωρητικό υπόβαθρο της μεθοδολογίας ανάλυσης των κύριων συνιστωσών, των νευρωνικών δικτύων καθώς και των χημικών αντιδράσεων που χρησιμοποιούνται. Τέλος, παρουσιάζονται τα αποτελέσματα των υπολογισμών που πραγματοποιήθηκαν, τα συμπεράσματα που προέκυψαν, καθώς και οι προτάσεις για περαιτέρω έρευνα.

Φυσικό Σύστημα

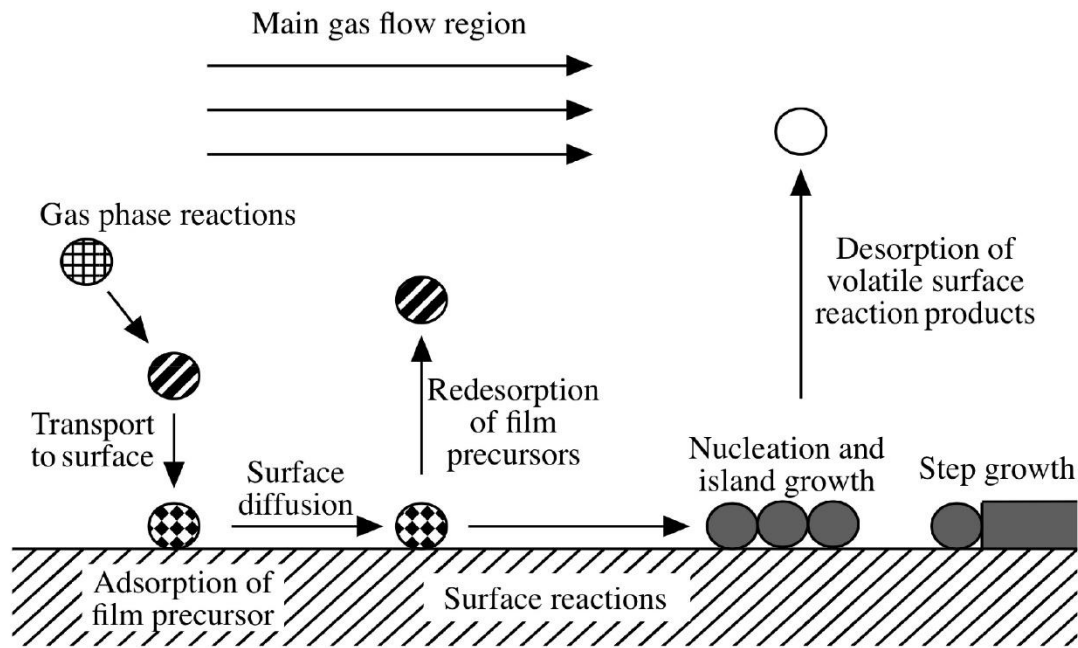
2.1. Χημική απόθεση από ατμό

Χημική απόθεση από ατμό (Chemical Vapor Deposition, CVD) ονομάζεται η διεργασία κατά την οποία αποτίθενται λεπτά στερεά υμένια (films) πάνω σε ειδικές επιφάνειες μέσω χημικών αντιδράσεων. Η διεργασία αυτή, βρίσκει ευρεία εφαρμογή στη βιομηχανία για την παραγωγή μικρο ή νάνο-ηλεκτρονικών συσκευών εξαιτίας της εύκολης προσαρμογής και μεταβολής των συνθηκών λειτουργίας του αντιδραστήρα στις εκάστοτε απαιτήσεις του προϊόντος. [2]

Η συνολική διεργασία περιλαμβάνει σύνθετους μηχανισμούς χημικής κινητικής καθώς και περίπλοκα φαινόμενα μεταφοράς ορμής, μάζας και θερμότητας. Τα αντιδρώντα, τα οποία περιέχονται στο πρόδρομο υλικό, μεταφέρονται στον αντιδραστήρα μέσω ενός φέροντος αερίου. Στη συνέχεια, πραγματοποιείται απόθεση σε στερεά δισκία (wafers) τα οποία στηρίζονται σε ειδικές επιφάνειες (substrates) και μέσω χημικών αντιδράσεων σχηματίζεται το υμένιο. Επιπλέον, αντιδράσεις μπορεί να λαμβάνουν χώρα μακριά από το υπόστρωμα, στην αέρια φάση στον κύριο όγκο του αντιδραστήρα πριν αυτά φτάσουν στην επιφάνεια του δισκίου όπου αντιδρούν. [2], [4]

Η διεργασία της χημικής απόθεσης από ατμό παρουσιάζεται σχηματικά στην Εικόνα 2-1 και συνοψίζεται στα εξής:

- Μεταφορά μάζας των αέριων αντιδρώντων
- Αντιδράσεις στην αέρια φάση
- Διάχυση των αέριων αντιδρώντων στην επιφάνεια απόθεσης
- Ρόφηση στην επιφάνεια
- Επιφανειακή διάχυση και αντιδράσεις
- Σχηματισμός του υμενίου
- Εκρόφηση των παραπροϊόντων από την επιφάνεια
- Διάχυση των παραπροϊόντων στην αέρια φάση
- Απομάκρυνση των παραπροϊόντων από την επιφάνεια απόθεσης



Εικόνα 2-1: Διεργασία Χημικής απόθεσης από ατμό

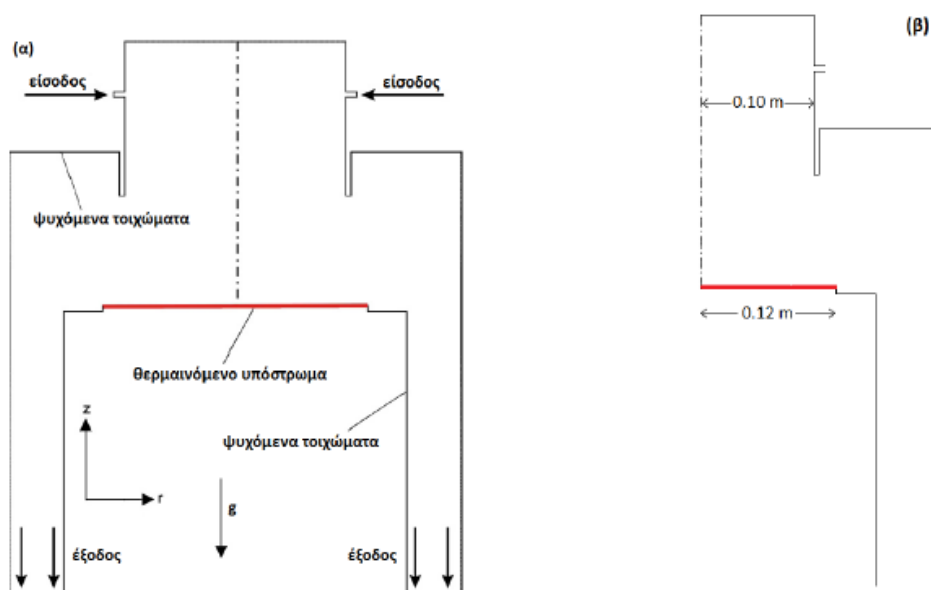
Οι διεργασίες χημικής απόθεσης από ατμό κατηγοριοποιούνται από τον τύπο του παραγόμενου υμενίου και από τις συνθήκες λειτουργίας του αντιδραστήρα. Όσον αφορά τις λειτουργικές παραμέτρους που επηρεάζουν τη διεργασία σημαντικές είναι η πίεση, όπου διακρίνονται αυτή της χαμηλής πίεσης (0,001 atm) CVD και της ατμοσφαιρικής πίεσης CVD, καθώς και ο τρόπος παροχής ενέργειας ενεργοποίησης στο υπόστρωμα, όπου διακρίνονται η απλή CVD, όπου η ενέργεια παρέχεται μέσω άμεσης θέρμανσης, η φωτονιακά βοηθούμενη CVD, όπου η ενέργεια παρέχεται μέσω laser ή υπεριώδους ακτινοβολίας και η υποβοηθούμενη από πλάσμα CVD. Σχετικά με τον τύπο του παραγόμενου προϊόντος, αυτό μπορεί να είναι Si από $\text{Si}_{4-x}\text{Cl}_x$ ή από SiH_4 , GaAs από $\text{Ga}(\text{CH}_3)_3$, InP από $\text{In}(\text{CH}_3)_3$ και PH_3 κ.α. [2]

Τα υμένια τα οποία παράγονται κάτω από συγκεκριμένες συνθήκες λειτουργίας θα πρέπει να έχουν ελεγχόμενες ιδιότητες και να μπορούν να αναπαράγονται. Οι βασικές ιδιότητες οι οποίες καθορίζουν την ποιότητα των παραχθέντων υμενίων είναι το πάχος, η ομοιομορφία, η μορφολογία της επιφάνειας, η σύσταση και η καθαρότητα. Το εύρος των αποδεκτών ορίων των παραπάνω ιδιοτήτων εξαρτάται από το είδος της διεργασίας, το υλικό καθώς και τη χρήση για την οποία προορίζεται το υμένιο. [2]

Οι αντιδραστήρες που χρησιμοποιούνται στη διεργασία της χημικής απόθεσης από ατμό απαρτίζονται από τέσσερα κύρια μέρη: το σύστημα εισόδου, τον κύριο αντιδραστήρα, την πηγή θέρμανσης και το σύστημα εξαγωγής αερίων. Έχουν αναπτυχθεί πολλά είδη αντιδραστήρων για την παραγωγή διαφορετικών ιδιοτήτων υμενίων, οι οποίοι διαφοροποιούνται ως προς το μέγεθος, σχήμα, προσανατολισμό του αντιδραστήρα και του δισκίου απόθεσης, το πλήθος των υποστρωμάτων, τη δυνατότητα περιστροφής του δισκίου και τέλος τη θέρμανση ή τη ψύξη των τοιχωμάτων, η οποία αποτρέπει την απόθεση του εκάστοτε προϊόντος στα τοιχώματα του αντιδραστήρα. [2]

Στο πλαίσιο της παρούσας εργασίας χρησιμοποιήθηκε ο αντιδραστήρας που παρουσιάζεται στην Εικόνα 2-2. Πρόκειται για έναν κυλινδρικό αντιδραστήρα ο οποίος

προσομοιώνεται σε διδιάστατη (2D), αξονική συμμετρία με θερμαινόμενο υπόστρωμα στους 700 K και ψυχόμενα τοιχώματα στους 300 K. Πιο συγκεκριμένα, το αέριο μίγμα εισέρχεται στον αντιδραστήρα από την πάνω πλευρά του σωλήνα έγχυσης διαμέτρου 0,20m, και στη συνέχεια οδηγείται στο δισκίο απόθεσης διαμέτρου 0,24 m. Εκεί, σχηματίζεται το υμένιο, όπως παρουσιάστηκε παραπάνω, ενώ τα παραπροϊόντα και τα αδρανή οδηγούνται στην έξοδο του αντιδραστήρα. Η πίεση λειτουργίας διατηρείται σταθερή στα 1300 Pa και έτσι μπορεί να θεωρηθεί ότι τα αέρια στο εσωτερικό του αντιδραστήρα συμπεριφέρονται ως ιδανικά και η ροή είναι στρωτή. Αξίζει να σημειωθεί ότι τα τοιχώματα του αντιδραστήρα ψύχονται, ούτως ώστε οι αντιδράσεις απόθεσης να μην ενεργοποιούνται και συνεπώς να μην αποτίθεται προϊόν και στα τοιχώματα [2], [5].



Εικόνα 2-2: (α) Τομή αντιδραστήρα CVD που χρησιμοποιήθηκε στην παρούσα εργασία
(β) Γεωμετρία του αντιδραστήρα σε αξονική μορφή

2.2. Μαθηματική Μοντελοποίηση Προβλήματος

Τα μαθηματικά μοντέλα που αναπτύσσονται βοηθούν στην κατανόηση της εκάστοτε διεργασίας, καθώς και στη βελτιστοποίηση αυτής όσον αφορά το κόστος, την ποιότητα του παραγόμενου προϊόντος κ.α. Τα φαινόμενα που εξελίσσονται σε μία διεργασία χημικής απόθεσης από ατμό είναι φαινόμενα μεταφοράς – ορμής, μάζας και θερμότητας – καθώς και χημικές αντιδράσεις και οι μηχανισμοί τους που λαμβάνουν χώρα στην αέρια φάση και στην επιφάνεια του δισκίου [2].

Στην παρούσα εργασία μελετήθηκε, σε πρώτη φάση, η ροή του αδρανούς αερίου N_2 και η μεταφορά ενέργειας στον αντιδραστήρα, χωρίς να λαμβάνονται υπόψη οι χημικές αντιδράσεις. Η απλοποίηση αυτή βασίζεται στο γεγονός ότι το πεδίο ροής και θερμοκρασίας που αναπτύσσεται στο εσωτερικό του αντιδραστήρα ΧΑΑ διαφέρει ελάχιστα από τα αντίστοιχα πεδία που αναπτύσσονται εάν ληφθούν υπόψη οι χημικές αντιδράσεις της διεργασίας, υπό την προϋπόθεση ότι η συγκέντρωση της πρόδρομης ένωσης στο φέρον αέριο είναι πολύ μικρή [2]. Στη δεύτερη φάση της εργασίας, θα προστεθούν χημικές

αντιδράσεις, οπότε χρήσιμο είναι να αναλυθούν σε αυτή την ενότητα το ισοζύγιο μάζας των χημικών συστατικών. Οι εξισώσεις που χρησιμοποιήθηκαν για τη μοντελοποίηση του προβλήματος σε μεταβατική κατάσταση παρουσιάζονται στη συνέχεια και περιλαμβάνουν τους μηχανισμούς των φαινομένων μεταφοράς μάζας, ορμής και ενέργειας [2].

i. Εξίσωση Συνέχειας

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.1)$$

Όπου t ο χρόνος, ρ η πυκνότητα και \mathbf{u} το διάνυσμα της ταχύτητας

ii. Εξίσωση Ορμής

$$\frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{g} \quad (2.2)$$

Όπου p η πίεση λειτουργίας, \mathbf{g} η επιτάχυνση της βαρύτητας και $\boldsymbol{\tau}$ ο τανυστής τάσης ο οποίος ορίζεται ως εξής:

$$\boldsymbol{\tau} = \mu \left((\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I} \right) \quad (2.3)$$

Όπου μ το ιξώδες και \mathbf{I} ο μοναδιαίος δυαδικός.

iii. Εξίσωση Ενέργειας

$$C_p \frac{\partial(\rho T)}{\partial t} + C_p \nabla \cdot (\rho \mathbf{u} T) = \nabla(\lambda \nabla T) - \sum_{i=1}^N \left(\mathbf{j}_i \cdot \frac{\nabla H_i}{M_i} \right) - \sum_{i=1}^N \sum_{k=1}^{N_g} H_i r_k^g \quad (2.4)$$

Όπου C_p η ειδική θερμοχωρητικότητα σε σταθερή πίεση, λ η θερμική αγωγιμότητα, \mathbf{j}_i ο ρυθμός διάχυσης του χημικού συστατικού i , H_i η ενθαλπία σχηματισμού του συστατικού i , M_i το μοριακό βάρος του συστατικού i , N το σύνολο των συστατικών του μίγματος, N_g το πλήθος των αέριων συστατικών που λαμβάνουν μέρος στις ογκομετρικές αντιδράσεις και r_k^g ο καθαρός ρυθμός της ογκομετρικής αντίδρασης k .

iv. Εξίσωση Ιδανικών Αερίων

$$p_i = c_i RT \quad (2.5)$$

όπου p_i η πίεση του i συστατικού, c_i η συγκέντρωση του i συστατικού και R η σταθερά των ιδανικών αερίων.

Οι τέσσερις παραπάνω εξισώσεις επιλύονται ταυτόχρονα και προκύπτει έτσι το πεδίο ροής και η κατανομή της θερμοκρασίας στο εσωτερικό του αντιδραστήρα ΧΑΑ.

Στο τελευταίο μέρος της εργασίας θα μελετηθεί η απόθεση πυριτίου (Si) από σιλένιο (SiH₄), συνεπώς σκόπιμο είναι να αναφερθούν οι εξισώσεις μεταφοράς μάζας των χημικών στοιχείων που αποτίθεται στο δισκίο για τη δημιουργία υμενίου. Η μελέτη της κατανομής της συγκέντρωσης των συστατικών σε όλες τις περιοχές του αντιδραστήρα είναι ιδιαίτερως σημαντική προκειμένου να υπολογιστούν χρήσιμα μεγέθη όπως ο ρυθμός απόθεσης στο υπόστρωμα ή η ομοιομορφίας και η καθαρότητας του υμενίου.

v. Εξίσωση Μεταφοράς Χημικών Στοιχείων

$$\frac{\partial(\rho\omega_i)}{\partial t} + \nabla \cdot (\rho\mathbf{u}\omega_i) = -\nabla \cdot \mathbf{j}_i + M_i \sum_{k=1}^{N_g} r_k^g \quad (2.6)$$

Όπου ω_i το κλάσμα μάζας του συστατικού i και το άθροισμα των κλασμάτων μάζας ισούται με τη μονάδα δηλαδή ισχύει $\sum_{i=1}^N \omega_i = 1$ συνεπώς η παραπάνω εξίσωση επιλύεται για $n - 1$ συστατικά, για όλα τα συστατικά δηλαδή πλην του φέροντος αερίου.

Στις διεργασίες χημικής απόθεσης από ατμό χρησιμοποιούνται συνήθως πολύ-συστατικά μίγματα στα οποία ισχύουν οι εξισώσεις Stephan – Maxwell, αντί του νόμου του Fick, για τον υπολογισμό του ρυθμού διάχυσης.

$$\sum_{j=1}^n \frac{f_i f_j}{D_{ji}} \left(\frac{\mathbf{j}_j}{\rho_j} - \frac{\mathbf{j}_i}{\rho_i} \right) = \nabla f_i - \frac{\nabla T}{T} \sum_{k=1}^n \frac{f_i f_j}{D_{ji}} \left(\frac{D_{T,j}}{\rho_j} - \frac{D_{T,i}}{\rho_i} \right) \quad (2.7)$$

Όπου f_i, f_j τα μοριακά κλάσματα των συστατικών i και j , D_{ji} ο δυαδικός συντελεστής διάχυσης Stephan – Maxwell, $D_{T,i}$ και $D_{T,j}$ οι συντελεστές θερμικής διάχυσης των συστατικών i και j και $\mathbf{j}_i, \mathbf{j}_j$ οι ρυθμοί διάχυσης οι οποίοι δίνονται από τη σχέση:

$$\mathbf{j}_i = - \sum_{j=1}^{n-1} \rho D_{ij} \nabla \omega_i - D_{T,i} \frac{\nabla T}{T} \quad (2.8)$$

Η εξίσωση (2.8) αποτελεί έκφραση του νόμου του Fick λαμβάνοντας υπόψη την επίδραση Soret, και έχει ισχύ μόνο όταν η σύσταση του μίγματος δεν μεταβάλλεται, ή όταν ο συντελεστής D_{ij} είναι ανεξάρτητος από αυτή. Ο υπολογισμός του συντελεστή αυτού γίνεται μέσω της εξίσωσης Chapman - Enskog:

$$D_{ji} = 0.0018 \frac{[T^3(\frac{1}{M_j} + \frac{1}{M_i})]^{0.5}}{P_{abs} \sigma_{ji}^2 \Omega_D} \quad (2.9)$$

Όπου P_{abs} η απόλυτη πίεση, σ_{ij} η ενεργή διατομή των συγκρούσεων και Ω_D το ολοκλήρωμα σύγκρουσης, το οποίο εκφράζει το μέτρο της αλληλεπίδρασης μεταξύ των μορίων του συστήματος και είναι συνάρτηση της ποσότητας:

$$T_D^* = \frac{T}{(\mathcal{E}/k_B)_{ji}} \quad (2.10)$$

Όπου k_B η σταθερά Boltzmann ίση με $1.3807 \cdot 10^{-23}$ και ο όρος $(\mathcal{E}/k_B)_{ji}$ η ενεργειακή παράμετρος που υπολογίζεται από το γεωμετρικό μέσο ως εξής:

$$(\mathcal{E}/k_B)_{ji} = \sqrt{(\mathcal{E}/k_B)_j * (\mathcal{E}/k_B)_i} \quad (2.11)$$

Η ενεργή διατομή των συγκρούσεων σ_{ij} υπολογίζεται από τον αριθμητικό μέσο των ατομικών σ ως εξής:

$$\sigma_{ji} = \frac{1}{2} (\sigma_i + \sigma_j) \quad (2.12)$$

Ο συντελεστής θερμικής διάχυσης D_T υπολογίζεται από τον Fluent μέσω της εμπειρικής σχέσης η οποία αποτελεί έκφραση του φαινομένου Soret:

$$D_{T,i} = -2,59 * 10^{-7} T^{0,659} \left[\frac{M_i^{0,511} f_i}{\sum_{i=1}^n M_i^{0,511} f_i} - \omega_i \right] \left[\frac{\sum_{i=1}^n M_i^{0,511} f_i}{\sum_{i=1}^n M_i^{0,489} f_i} \right] \quad (2.13)$$

Επισημαίνεται ότι, ο Fluent έχει βιβλιοθήκες και βάσεις δεδομένων οι οποίες χρησιμοποιούνται για τον υπολογισμό του ιξώδους, της ειδικής θερμοχωρητικότητας και των υπόλοιπων απαραίτητων χαρακτηριστικών μεγεθών για κάθε συστατικό τα οποία περιέχονται στις παραπάνω εξισώσεις.

Οι συνοριακές συνθήκες που απαιτούνται για την επίλυση του προβλήματος συνοψίζονται ως εξής: [2], [5]

- i. Στην είσοδο του αντιδραστήρα, το αέριο μίγμα εισέρχεται σε σταθερή θερμοκρασία ίση με 298 K και με ομοιόμορφη ταχύτητα κάθετη στην είσοδο. Η ταχύτητα εισόδου εξαρτάται άμεσα από τη μαζική παροχή εισόδου, η οποία στο συγκεκριμένο πρόβλημα κυμαίνεται μεταξύ $1,5 \cdot 10^{-5}$ kg/s και $6,0 \cdot 10^{-5}$ kg/s.
- ii. Στην έξοδο του αντιδραστήρα, επιβάλλεται ταχύτητα κάθετη στην έξοδο, μηδενική βαθμίδα θερμοκρασίας και ροής μάζας, καθώς και τιμή αναφοράς για την πίεση.

- iii. Στα τοιχώματα του αντιδραστήρα ισχύει η συνθήκη της μη-ολίσθησης για την ταχύτητα και η θερμοκρασία διατηρείται σταθερή και ίση με 300 K στα ψυχόμενα τοιχώματα και με 700 K για το θερμαινόμενο υπόστρωμα.
- iv. Στον άξονα συμμετρίας, η ταχύτητα έχει μηδενική συνιστώσα κάθετη στον άξονα και η βαθμίδα κάθε μεταβλητής είναι μηδενική στη διεύθυνση κάθετη στον άξονα.

Τέλος, απαραίτητη για την επίλυση του προβλήματος είναι η ύπαρξη αρχικών συνθηκών, οι οποίες είναι οι γνωστές κατανομές στο χώρο των πεδίων $P(r, 0)$, $\mathbf{u}(r, 0)$ και $T(r, 0)$, στο χρόνο $t=0$. Οι αρχικές συνθήκες έχουν προκύψει από επίλυση του προβλήματος στη μόνιμη κατάσταση για κάποια τιμή της παροχής εισόδου.

Η λύση του προβλήματος είναι οι κατανομές στο χώρο και στο χρόνο των πεδίων $P(r, t)$, $\mathbf{u}(r, t)$ και $T(r, t)$.

2.3. Επίλυση προβλήματος

Το σύστημα εξισώσεων που παρουσιάστηκε στην προηγούμενη ενότητα επιλύεται σε σύστημα κυλινδρικών συντεταγμένων με τον εμπορικό κώδικα υπολογιστικής μηχανικής ANSYS Fluent 13.0, με χρήση της μεθόδου των πεπερασμένων όγκων. Ο αντιδραστήρας διαμερίζεται σε 15066 κελιά στα οποία υπολογίζεται για κάθε χρονική στιγμή η πίεση, η θερμοκρασία και το διάνυσμα της ταχύτητας (η αξονική και η ακτινική συνιστώσα) μέσω των εξισώσεων που διέπουν το φυσικό σύστημα όπως αναλύθηκε προηγουμένως, δημιουργώντας έτσι ένα διάνυσμα κατάστασης $\mathbf{x}(t)$ με 60264 στοιχεία [6], [2].

Κατά τη διάρκεια της δυναμικής μελέτης του αντιδραστήρα ΧΑΑ, το σύστημα εξισώσεων δεν αφήνεται να συγκλίνει σε κάθε χρονικό βήμα Δt της μεταβατικής κατάστασης, αλλά αντίθετα καλείται εξωτερικά από περιβάλλον Matlab [7], οπότε για την k επανάληψη ως αρχικές συνθήκες λαμβάνονται τα στοιχεία του διανύσματος κατάστασης της $\mathbf{x}(t_{k-1})$ με παροχή εισόδου $\dot{m}(t_k)$. Σε κάθε χρονικό βήμα, πραγματοποιούνται 200 βήματα της επαναληπτικής μεθόδου. Ο αριθμός των επαναλήψεων σε κάθε χρονικό βήμα δεν είναι απαραίτητο να εξασφαλίζει σύγκλιση αλλά σε κάθε περίπτωση είναι αρκετά μεγάλος ώστε η νόρμα της διαφοράς δύο διαδοχικών επαναλήψεων να είναι για την εξίσωση συνέχειας (2.2) περίπου στο 10^{-5} ενώ για τις άλλες τρεις εξισώσεις περίπου στο 10^{-7} . Σημειώνεται ότι ο χρόνος διακριτοποιείται σε διαστήματα $\Delta t=0,1$ s και ισχύει ότι:

$$t_k = t_{k-1} + \Delta t \quad (2.14)$$

Επίσης, η διαδικασία επίλυσης θεωρείται ότι ξεκινάει τη χρονική στιγμή $t_0 = 0$ s. Η μαζική παροχή είναι η μόνη ανεξάρτητη μεταβλητή η οποία εξαρτάται από το χρόνο και συνεπώς η τιμή της $\dot{m}(t)$ δρα ως μεταβλητή εισόδου και το διάνυσμα κατάστασης $\mathbf{x}(t)$ ως μεταβλητή εξόδου.

Στην παρούσα εργασία υπολογίστηκε η δυναμική συμπεριφορά του συστήματος από μια μόνιμη κατάσταση σε μία άλλη, δηλαδή επιβλήθηκαν στο σύστημα βηματικές επιβολές. Συγκεκριμένα, σε μια επιβολή με n χρονικά βήματα υπολογίζονται διαδοχικά n το πλήθος διανύσματα κατάστασης $\mathbf{x}(t)$, καθένα από τα οποία αποτελεί ένα στιγμιότυπο (snapshot) του συστήματος, και όλα μαζί δημιουργούν την τροχιά (trajectory) του συστήματος. Έχει δημιουργηθεί, δηλαδή, ένα πίνακας ο οποίος καταγράφει τη δυναμική συμπεριφορά του συστήματος από τη μία κατάσταση ισορροπίας σε μία άλλη.

$$Y = [x(t_0) \quad x(t_1) \quad \dots \quad x(t_{n-1})] \quad (2.15)$$

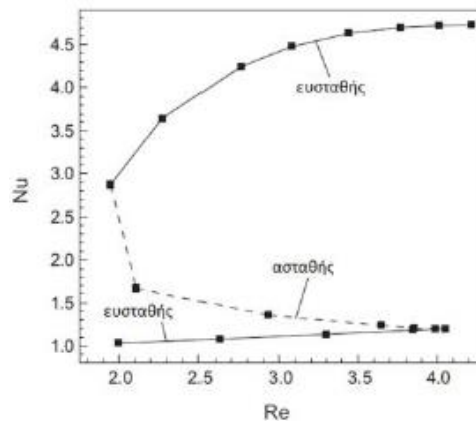
Η επίλυση του προβλήματος πραγματοποιείται μέχρι σύγκλισης στη νέα μόνιμη κατάσταση, η οποία ελέγχεται εξωτερικά από το περιβάλλον Matlab, μέχρις ότου η νόρμα της διαφοράς δύο διαδοχικών διανυσμάτων κατάστασης να είναι μικρότερη από ένα συγκεκριμένο όριο ανοχής (tolerance) [6].

$$\frac{\|x(t_{k-1}) - x(t_k)\|}{\|x(t_k)\|} < 10^{-6} \quad (2.16)$$

Η ανοχή 10^{-6} στην επίλυση του συστήματος εξισώσεων υποδηλώνει ότι το σύστημα έχει πρακτικά συγκλίνει στη νέα μόνιμη κατάσταση. Στο πλαίσιο της επιβεβαίωσης της συγκεκριμένης παραδοχής, συγκρίθηκαν οι μόνιμες καταστάσεις, οι οποίες έχουν επιλυθεί με τη βοήθεια του Fluent, με τις καταστάσεις $x(t_{n-1})$ του πίνακα Y . Το σχετικό σφάλμα των δύο καταστάσεων τείνει στο μηδέν και άρα η κατάσταση του αντιδραστήρα έχει συγκλίνει στη νέα μόνιμη κατάσταση, συνεπώς η παραδοχή της συγκεκριμένης ανοχής κρίνεται αποδεκτή.

2.4. Πολλαπλότητα Χώρου Λύσεων

Χαρακτηριστικό του συγκεκριμένου συστήματος είναι η πολλαπλότητα λύσεων που παρουσιάζει στις καταστάσεις ισορροπίας, η οποία οφείλεται στους μη γραμμικούς όρους των εξισώσεων (2.1)-(2.5). Ειδικότερα, για κάθε τιμή της μαζικής παροχής $\dot{m}(t)$ που κυμαίνεται μεταξύ $2.3 \cdot 10^{-5}$ kg/s και $5.0 \cdot 10^{-5}$ kg/s, το σύστημα μπορεί να συγκλίνει σε τρεις μόνιμες καταστάσεις, δύο ευσταθείς και μία ασταθής, ανάλογα με την αρχική κατάσταση ισορροπίας στην οποία βρίσκεται το σύστημα στην εκάστοτε χρονική στιγμή. Στην Εικόνα 2-3 παρουσιάζεται διάγραμμα των τριών κλάδων λύσεων συναρτήσει του αριθμού Reynolds, ο οποίος είναι ανάλογος της παροχής εισόδου, και του αριθμού Nusselt, ο οποίος χαρακτηρίζει τη μόνιμη κατάσταση του συστήματος [1], [2].

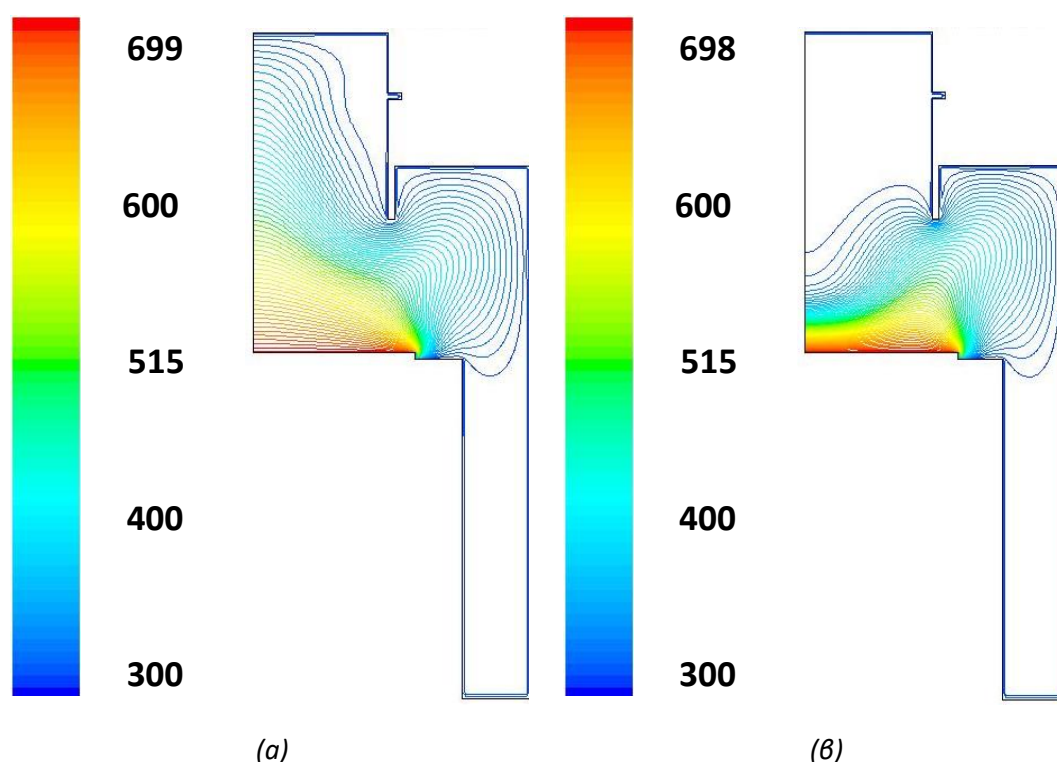


Εικόνα 2-3: Πολλαπλότητα χώρου λύσεων του μελετώμενου συστήματος

Η φυσική σημασία της πολλαπλότητας λύσεων που παρουσιάστηκε παραπάνω έγκειται στο γεγονός ότι στο φυσικό σύστημα ανταγωνίζονται δύο μηχανισμοί, η φυσική κυκλοφορία όπου κυριαρχεί η άνωση λόγω βαθμίδας θερμοκρασίας μεταξύ του θερμαινόμενου υποστρώματος και του κύριου όγκου του αντιδραστήρα, και η εξαναγκασμένη κυκλοφορία όπου κυριαρχεί η ορμή του εισερχόμενου ρευστού. Πιο συγκεκριμένα, στον πάνω ευσταθή κλάδο παρατηρείται η εξαναγκασμένη συναγωγή και

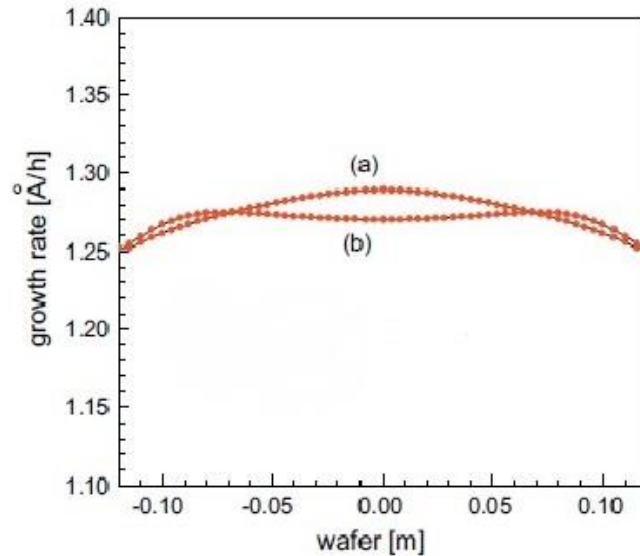
στον κάτω ευσταθή κλάδο η φυσική κυκλοφορία. Για κάθε παροχή $\dot{m}(t) \geq 5.0 \cdot 10^{-5}$ kg/s στο σύστημα κυριαρχεί μόνο ένας μηχανισμός, αυτός της εξαναγκασμένης κυκλοφορίας, ενώ για κάθε παροχή $\dot{m}(t) \leq 2.3 \cdot 10^{-5}$ kg/s κυριαρχεί αυτός της φυσικής κυκλοφορίας. Συνεπώς, ο πάνω κλάδος εκτείνεται για κάθε $\dot{m}(t) \geq 2.3 \cdot 10^{-5}$ kg/s, ενώ ο κάτω ευσταθής κλάδος για κάθε $\dot{m}(t) \leq 5.0 \cdot 10^{-5}$ kg/s [1], [2].

Στην Εικόνα 2-4 παρουσιάζεται συγκριτικά η κατανομή της θερμοκρασίας που επικρατεί στον αντιδραστήρα των δύο μηχανισμών ροής για την ίδια τιμή παροχής εισόδου, η οποία προφανώς βρίσκεται μέσα στην περιοχή πολλαπλότητας. Επιπροσθέτως, αν επιβληθεί στο σύστημα μία βηματική επιβολή η οποία ξεκινά από μία μόνιμη κατάσταση για παροχή εκτός πολλαπλότητας λύσεων στον πάνω κλάδο και καταλήγει σε μία παροχή η οποία βρίσκεται μέσα στο χώρο πολλαπλότητας λύσεων, ο Fluent θα συγκλίνει συνήθως στη λύση η οποία βρίσκεται πιο κοντά, δηλαδή στη μόνιμη κατάσταση στον πάνω κλάδο. Τέλος, ο χρόνος σύγκλισης αυξάνει σημαντικά καθώς το σύστημα προσπαθεί να συγκλίνει σε σημείο όπου πραγματοποιείται στροφή από τον ένα ευσταθή κλάδο στον άλλο, όσο πλησιάζει δηλαδή στην περιοχή αστάθειας [1], [2].



Εικόνα 2-4: Κατανομή θερμοκρασίας στον αντιδραστήρα για $m=4.0 \cdot 10^{-5}$ kg/s (α) στον κάτω ευσταθή κλάδο και (β) στον πάνω ευσταθή κλάδο

Όπως έχει προαναφερθεί, τα πεδία ροής και θερμοκρασίας δεν επηρεάζονται από την ύπαρξη άλλων συστατικών στο μίγμα πέραν του φέροντος αερίου, τη διάχυση αυτών και τις μεταξύ τους χημικές αντιδράσεις. Υπενθυμίζεται ότι τα παραπάνω ισχύουν στην περίπτωση όπου τα αντιδρώντα βρίσκονται σε ίχνη μέσα στον αντιδραστήρα [2]. Αντίθετα, ο ρυθμός απόθεσης καθώς και η ομοιομορφία του σχηματιζόμενου υμενίου εξαρτώνται σημαντικά από το μηχανισμό ροής που επικρατεί στον αντιδραστήρα, όπως παρουσιάζεται στην Εικόνα 2-5 [2].



Εικόνα 2-5: Ο ρυθμός απόθεσης του σχηματιζόμενου υμενίου συναρτήσει της θέσης στο υπόστρωμα για σταθερή παροχή στον (α) κάτω ευσταθή κλάδο και (β) στον πάνω ευσταθή κλάδο

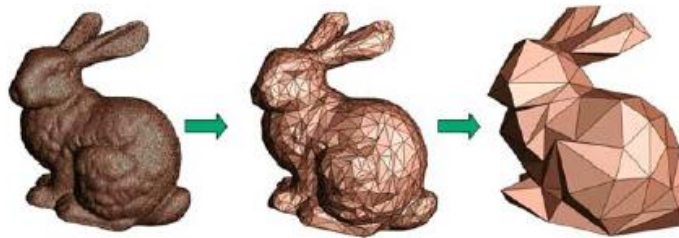
Στην Εικόνα 2-5 παρουσιάζονται οι διαφορές στην ομοιομορφία του παραγόμενου υμενίου όταν στον αντιδραστήρα επικρατούν συνθήκες φυσικής κυκλοφορίας σε σχέση με την εξαναγκασμένη ροή. Όπως έχει προαναφερθεί, επιδιώκεται από τη βιομηχανία η όσο το δυνατόν μεγαλύτερη ομοιομορφία, η οποία όμως είναι δύσκολο να ελεγχθεί. Τα υμένια τα οποία παράγονται σε συνθήκες εξαναγκασμένης συναγωγής χαρακτηρίζονται από μεγαλύτερη ομοιομορφία [2], [1].

Συμπεραίνεται, λοιπόν, ότι ο ρυθμός απόθεσης και η ομοιομορφία του υμενίου δεν καθορίζεται πλήρως από την παροχή εισόδου του ρευστού. Αυτό καθιστά τη λεπτομερή μελέτη της διεργασίας απαραίτητη τόσο σε θεωρητικό επίπεδο όσο και σε πρακτικό για την παραγωγή υμενίων που να ικανοποιούν τις προδιαγραφές της εκάστοτε εφαρμογής, αφού μια μικρή μεταβολή στην παροχή εισόδου μπορεί να αλλάξει ριζικά τα χαρακτηριστικά τους. Στο πλαίσιο της παρούσας εργασίας θα σχεδιαστεί ένα μοντέλο μειωμένης τάξης το οποίο θα μπορεί να προβλέψει τις συνέπειες των συγκεκριμένων αλλαγών με ικανοποιητική ακρίβεια και σε σύντομο χρονικό διάστημα.

Μείωση Τάξης Μοντέλου

3.1 Εισαγωγή

Η μείωση τάξης μοντέλου (Model Order Reduction, MOR) αναπτύχθηκε στο πεδίο της θεωρίας συστημάτων και ρύθμισης ως μια μέθοδος με κύριο στόχο τη μείωση της πολυπλοκότητας δυναμικών συστημάτων διατηρώντας παράλληλα τις βασικές ιδιότητες του πλήρους τάξης μοντέλου (Full Order Model), ως προς τις μεταβλητές εισόδου και εξόδου. Η πολυπλοκότητα των δυναμικών συστημάτων έγκειται στο γεγονός ότι απαιτούνται για την περιγραφή τους μεγάλο αριθμό εξισώσεων και μεταβλητών. Τα μοντέλα μειωμένης τάξης μπορούν να πραγματοποιούν προσομοιώσεις σε πολύ σύντομο χρονικό διάστημα χρησιμοποιώντας μικρό ποσοστό υπολογιστικής μνήμης και συγχρόνως να παράγουν ακούρτως ικανοποιητικά αποτελέσματα, να περιγράφουν δηλαδή το σύστημα με ικανοποιητική ακρίβεια. Στην Εικόνα 3-1 παρουσιάζεται η βασική ιδέα της μείωσης τάξης μοντέλου: η μικρή απαίτηση σε πληροφορία για την περιγραφή ενός συστήματος [8].



Εικόνα 3-1: Γραφική Αναπαράσταση της Μείωσης Τάξης Μοντέλου

Οι απαιτήσεις της βιομηχανίας ως προς την ποιότητα των παραγόμενων προϊόντων, συνεχώς αυξάνονται, απαιτώντας πιο πολύπλοκα συστήματα και διεργασίες για την επίτευξη τους. Τα συστήματα αυτά, μπορεί να περιλαμβάνουν σύνθετα φαινόμενα ροής ρευστών, χημικές αντιδράσεις κ.λ.π. Είναι προφανές ότι τα μαθηματικά μοντέλα τα οποία περιγράφουν τα συστήματα αυτά είναι αρκετά περίπλοκα καθιστώντας έτσι δύσκολη την ανάλυση, τη μελέτη, την προσομοίωση και το σχεδιασμό τους. Για τα συγκεκριμένα συστήματα, μοντέλα μειωμένης πολυπλοκότητας, τα οποία διατηρούν συγχρόνως τα κυριότερα χαρακτηριστικά τους είναι ιδιαίτερα ελκυστικά και βοηθούν σημαντικά την κατανόηση, τη βελτίωση καθώς και τη ρύθμιση της όλης διεργασίας [9].

Στη σύγχρονη εποχή, οι ηλεκτρονικοί υπολογιστές έχουν τη δυνατότητα να επιλύουν τα περίπλοκα μοντέλα τα οποία περιγράφουν σύνθετα συστήματα μηχανικής, γρήγορα, χωρίς απλοποιήσεις και με μεγάλη ακρίβεια. Εύλογα εγείρεται το ερώτημα κατά πόσο είναι απαραίτητη η μείωση της τάξης των μοντέλων, όταν υπολογισμοί περατώνονται από μία μηχανή. Ωστόσο, η επίλυση του μοντέλου πλήρους τάξης για τις σύνθετες διεργασίες είναι μια χρονοβόρα διαδικασία και απαιτεί μεγάλη υπολογιστική μνήμη, ιδιαίτερα όσο αυξάνεται η απαιτούμενη ακρίβεια και λεπτομέρεια. Συνεπώς, το μοντέλο πλήρους τάξης είναι δύσκολο έως αδύνατο να χρησιμοποιηθεί για την αυτόματη ρύθμιση του συστήματος. Για αυτό το λόγο, κρίνεται απαραίτητη η ύπαρξη μοντέλων τα οποία μπορούν να προσεγγίζουν την κατάσταση του συστήματος σε μικρό χρονικό διάστημα [8].

3.2 Μείωση Τάξης Δυναμικών Συστημάτων

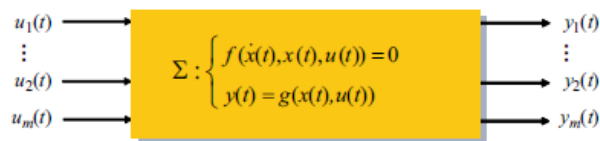
Τα αναπτυσσόμενα μοντέλα στην υπολογιστική μηχανική αποτελούνται συνήθως από συστήματα μερικών ή/και συνήθων διαφορικών εξισώσεων σε συνδυασμό με τις συνοριακές και αρχικές συνθήκες, όπως οι εξισώσεις Navier-Stokes στη ρευστοδυναμική και οι εξισώσεις Maxwell στον ηλεκτρομαγνητισμό. Για δυναμικά συστήματα που περιγράφονται από μερικές διαφορικές εξισώσεις με ανεξάρτητες μεταβλητές το χώρο και το χρόνο, μετά από διακριτοποίηση στο χώρο προκύπτει ένα σύστημα συνήθων διαφορικών εξισώσεων. Έστω ότι το σύστημα περιγράφεται είναι το εξής: [8]

$$\frac{dx}{dt} = f(x, u) \quad (3.1)$$

$$y = g(x, u) \quad (3.2)$$

Όπου u είναι η είσοδος του συστήματος, y η έξοδος και x η μεταβλητή κατάστασης. Η πολυπλοκότητα του συστήματος χαρακτηρίζεται από το πλήθος των μεταβλητών κατάστασης, δηλαδή τη διάσταση n του διανύσματος x . Ο τελεστής f μπορεί να περιλαμβάνει γραμμικούς και μη γραμμικούς όρους που στην περίπτωση των διεργασιών ΧΑΑ αφορούν τη διάχυση, τη συναγωγή και την αντίδραση. Ο τελεστής g συνδέει την είσοδο με την έξοδο του συστήματος [8].

Το δυναμικό σύστημα παρουσιάζεται γραφικά στην Εικόνα 3-2.



Εικόνα 3-2: Γραφική Αναπαράσταση Δυναμικού Συστήματος

Η μείωση της τάξης ενός μοντέλου αφορά τη μείωση της διάστασης του διανύσματος κατάστασης $x(t)$, διατηρώντας παράλληλα τη σχέση μεταξύ της εισόδου και της εξόδου του συστήματος. Το νέο δυναμικό σύστημα ορίζεται ως εξής:

$$\frac{d\hat{x}}{dt} = \hat{f}(\hat{x}, u) \quad (3.3)$$

$$y = \hat{g}(\hat{x}, u) \quad (3.4)$$

Όπου \hat{x} είναι το νέο διάνυσμα κατάστασης διάστασης d πολύ μικρότερη του N και \hat{f} , \hat{g} οι προσεγγίσεις των συναρτήσεων f , g αντίστοιχα. Για τη δημιουργία καλής προσέγγισης του αρχικού συστήματος πρέπει να ικανοποιούνται οι παρακάτω προϋποθέσεις: [8]

- Μικρό σφάλμα προσέγγισης
- Διατήρηση ιδιοτήτων του αρχικού συστήματος όπως π.χ. η ευστάθεια
- Υπολογιστικά αποτελεσματική διαδικασία μείωσης

Το σύστημα εξισώσεων που θα μελετηθεί στο πλαίσιο της παρούσας διπλωματικής εργασίας μπορεί να απλοποιηθεί στις δύο παρακάτω εξισώσεις:

$$\mathbf{y} = \mathbf{x} \quad (3.5)$$

$$\mathbf{u} = \dot{\mathbf{m}} \quad (3.6)$$

όπου $\dot{\mathbf{m}}$ είναι η μαζική παροχή εισόδου στον αντιδραστήρα.

3.3 Μέθοδος Galerkin

Η μέθοδος Galerkin είναι μια μέθοδος διακριτοποίησης και μείωσης τάξης μερικών διαφορικών εξισώσεων που βασίζεται στο χωρισμό των μεταβλητών και την προβολή της πλήρους λύσης σε συναρτήσεις βάσης. Έστω ότι το σύστημα περιγράφεται από τις εξής μερικές διαφορικές εξισώσεις, με τις απαραίτητες αρχικές και συνοριακές συνθήκες: [9]

$$\frac{\partial x}{\partial t} = D(x) \quad u: \Omega \times (0, \infty) \rightarrow \mathbb{R} \quad (3.7)$$

Όπου D ένας τελεστής στο χώρο και $\Omega \subset \mathbb{R}$ είναι το πεδίο του χώρου. Η μέθοδος Galerkin προσεγγίζει τη λύση $x(\mathbf{r}, t)$, όπου \mathbf{r} είναι η θέση στο χώρο και t ο χρόνος, ως εξής:

$$\hat{x}(\mathbf{r}, t) = \sum_{j=1}^d a_j(t) \varphi_j(\mathbf{r}) \quad (3.8)$$

Από τη σχέση (3.8) προκύπτει ότι η εξάρτηση της προσέγγισης \hat{x} από το χώρο καθορίζεται από τις συναρτήσεις $\varphi_j(\mathbf{r})$, οι οποίες είναι σταθερές ως προς το χρόνο και σχηματίζουν τη βάση ενός d -διάστατου συναρτησιακού χώρου από τον οποίο προσεγγίζεται το αρχικό σύστημα, ενώ η εξάρτηση της προσέγγισης \hat{x} από το χρόνο καθορίζεται από τους συντελεστές $a_j(t)$ οι οποίοι επιλέγονται ούτως να ελαχιστοποιείται το r στην παρακάτω σχέση: [9]

$$r(\mathbf{r}, t) = \frac{\partial \hat{x}}{\partial t} - D(\hat{x}) \quad (3.9)$$

Με συνδυασμό των σχέσεων (3.8) και (3.9) προκύπτει:

$$r(\mathbf{r}, t) = \sum_{j=1}^d a_j(t) \varphi_j(\mathbf{r}) - D\left(\sum_{j=1}^d a_j(t) \varphi_j(\mathbf{r})\right) \quad (3.10)$$

Η ελαχιστοποίηση της σχέσης (3.10) ικανοποιείται όταν το υπόλοιπο $r(\mathbf{r}, t)$ είναι ορθογώνιο ως προς όλες τις συναρτήσεις βάσης $\varphi_j(\mathbf{r})$, οπότε τελικά προκύπτει το ακόλουθο σύστημα d συνήθων διαφορικών εξισώσεων ως προς του συντελεστές που περιγράφουν τη χρονική εξάρτηση $a_j(t)$: [9]

$$\frac{da_i}{dt} = \int D\left(\sum_{j=1}^d a_j(t) \varphi_j(\mathbf{r})\right) \varphi_i(\mathbf{r}) d^n r \quad i = 1 \dots d \quad (3.11)$$

Με αρχικές συνθήκες για τις d διαφορικές εξισώσεις:

$$a_i(0) = \int x(\mathbf{r}, 0) \varphi_i(\mathbf{r}) d^n x \quad (3.12)$$

Έχει προκύψει δηλαδή ένα μειωμένης τάξης μοντέλο διάστασης d . Συνεπώς, το διάνυσμα κατάστασης $\mathbf{x}(t)$ προκύπτει εάν τοποθετηθούν όλες οι εξαρτημένες μεταβλητές x σε όλες τις θέσεις του χώρου \mathbf{r} για μια χρονική στιγμή t . [8]

$$\hat{\mathbf{x}}(t) = \sum_{j=1}^d a_j(t) \boldsymbol{\varphi}_j \quad (3.13)$$

Όπου $\boldsymbol{\varphi}_j$ είναι τα αντίστοιχα διανύσματα βάσης, κατασκευασμένα από τις συναρτήσεις βάσεις $\varphi_j(\mathbf{r})$ με την ίδια μεθοδολογία που κατασκευάστηκε το διάνυσμα κατάστασης $\mathbf{x}(t)$.

Σημειώνεται ότι στην παρούσα εργασία θα χρησιμοποιηθεί μόνο η σχέση (3.13), καθώς οι συντελεστές $a_j(t)$ θα προσδιοριστούν με τη χρήση τεχνητών νευρωνικών δικτύων.

3.4 Ανάλυση Κύριων Συνιστωσών

Η ανάλυση κύριων συνιστωσών (principal component analysis, PCA) είναι μια μέθοδος μείωσης τάξης μέσω της οποίας υπολογίζεται η βέλτιστη ορθοκανονική βάση, σύμφωνα με τη μέθοδο των ελαχίστων τετραγώνων, για δοθέντα θεωρητικά, πειραματικά ή υπολογιστικά δεδομένα [8].

Κατά την εφαρμογή της μεθοδολογίας της ανάλυσης κύριων συνιστωσών, το δυναμικό σύστημα προβάλλεται σε έναν υπόχωρο του πλήρους χώρου λύσεων. Σε συνδυασμό με τη μέθοδο Galerkin, παρέχει ένα χρήσιμο εργαλείο για την πρόβλεψη συμπεριφοράς δυναμικών συστημάτων μεγάλης ή ακόμα και άπειρης διάστασης αφού ο υπόχωρος κατασκευάζεται μέσω συναρτήσεων βάσης οι οποίες διατηρούν τα ειδικά χαρακτηριστικά του πλήρους τάξης συστήματος, σε αντίθεση με άλλες μεθόδους (π.χ. πεπερασμένα στοιχεία) όπου οι συναρτήσεις βάσεις είναι ανεξάρτητες της δυναμικής του συστήματος. Επιπλέον, η μέθοδος απαιτεί για την αναπαραγωγή της συμπεριφοράς του μελετώμενου συστήματος απλές πράξεις μεταξύ πινάκων, ακόμη και εάν το σύστημα είναι μη γραμμικό. Στην περίπτωση αυτή, το μειωμένης τάξης σύστημα που προκύπτει είναι επίσης μη γραμμικό, διατηρεί δηλαδή τα χαρακτηριστικά του. Τέλος, η μέθοδος αυτή επεξεργάζεται αποτελέσματα πειραμάτων ή προσομοιώσεων, δεν απαιτεί δηλαδή καμιά πρότερη γνώση του μελετώμενου συστήματος. Για το λόγο αυτό έχει χρησιμοποιηθεί σε πεδία όπως η αναγνώριση προτύπων (pattern recognition), η ρευστοδυναμική, η ανάλυση σημάτων και στο σχεδιασμό ρυθμιστή για χρήση σε πραγματικό χρόνο [8].

Τονίζεται ότι ακριβώς επειδή πρόκειται για μια μέθοδο επεξεργασίας αποτελεσμάτων, η ορθοκανονική βάση που κατασκευάζεται από την PCA προσεγγίζει ικανοποιητικά τη δυναμική συμπεριφορά του υποσυνόλου του συστήματος από το οποίο ελήφθησαν τα πειραματικά – υπολογιστικά δεδομένα και όχι όλου του συστήματος [9].

3.4.1 Κατασκευή Βάσης PCA

Έστω ότι τα αποτελέσματα πειραμάτων ή/και προσομοιώσεων προέρχονται από έναν διανυσματικό χώρο $V = \mathbb{R}^N$. Επιπλέον, έστω ένα σύνολο δεδομένων Y αποτελούμενο από m τροχιές (trajectories) [9], [8].

$$Y = [\mathbf{y}_1(t) \quad \cdots \quad \mathbf{y}_m(t)] \quad (3.14)$$

Όπου $y_i(t) \in \mathbb{R}^N, i = 1, \dots, m$ και $t \in [0, T]$

Στη συνέχεια, χρησιμοποιείται η μέθοδος της ανάλυσης κύριων συνιστωσών για την εύρεση ενός d -διάστατου υπόχωρου $V_d \subset V$ με κύριο στόχο την κατασκευή μιας ορθής προβολής $\Pi_d : V \rightarrow V_d$ σταθερής διάστασης $d \ll m$, ελαχιστοποιώντας την ακόλουθη απόσταση ελαχίστων τετραγώνων: [8]

$$\|Y - \Pi_d Y\|^2 = \sum_{i=1}^m \int_0^T \|\mathbf{y}_i(t) - \Pi_d \mathbf{y}_i(t)\|^2 dt \quad (3.15)$$

Η λύση του προβλήματος προκύπτει με τη χρήση του πίνακα συσχέτισης (correlation matrix) διάστασης $N \times N$:

$$K = \sum_{i=1}^m \int_0^T \mathbf{y}_i(t) \mathbf{y}_i(t)^T dt \quad (3.16)$$

Εξ' ορισμού, ο πίνακας K είναι θετικά ημι-ορισμένος με πραγματικές και μη αρνητικές ιδιοτιμές $\lambda_1 \geq \dots \geq \lambda_n \geq 0$. Έστω ότι \mathbf{u}_j είναι τα αντίστοιχα μη μηδενικά ιδιοδιανύσματα: [8]

$$K \mathbf{u}_j = \lambda_j \mathbf{u}_j \quad (3.17)$$

Τα ιδιοδιανύσματα \mathbf{u}_j διάστασης N επιλέγονται ούτως ώστε να είναι ορθοκανονικά, δηλαδή να ικανοποιούν τη σχέση:

$$\mathbf{u}_i^T \mathbf{u}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (3.18)$$

Κάθε διάνυσμα αναλύεται σε συνιστώσες ως εξής:

$$\mathbf{y}_i(t) = \sum_{j=1}^N y_{ij}(t) \mathbf{u}_j \quad (3.19)$$

Όπου:

$$y_{ij}(t) = \mathbf{y}_i(t)^T \mathbf{u}_j \quad (3.20)$$

3.4.2 Επιλογή Διάστασης Βάσης

Η διάσταση d της ορθοκανονικής βάσης επιλέγεται ούτως ώστε η πρόβλεψη του συστήματος που θα προκύψει από την εφαρμογή της μεθοδολογίας να είναι ικανοποιητική και συγχρόνως να είναι αρκούντως μικρή, δηλαδή να ισχύει $d \ll m$. Σύμφωνα με τη μεθοδολογία της ανάλυσης κύριων συνιστωσών αποδεικνύεται ότι αν επιλεγεί υπόχωρος $V_d = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$, η τιμή της απόστασης (3.15) που είναι μικρότερη δυνατή για δεδομένο d , είναι ίση με: [8], [9]

$$\min \|Y - \Pi_d Y\|^2 = \sum_{i=n-d+1}^n \lambda_i \quad (3.21)$$

Η τιμή της απόστασης σύμφωνα με την παραπάνω σχέση ισούται με το άθροισμα των ιδιοτιμών των ιδιοδιανυσμάτων που δεν χρησιμοποιούνται στην κατασκευή της βάσης.

Ειδικότερα, μπορεί να θεωρηθεί ως το πληροφοριακό περιεχόμενο που δεν λαμβάνεται υπόψη στην παραγόμενη πρόβλεψη του μοντέλου μειωμένης τάξης. Επομένως, η πληροφορία που περιέχεται στην προσέγγιση μπορεί να οριστεί ως το άθροισμα των d πρώτων ιδιοτιμών:

$$\sum_{i=1}^d \lambda_j \quad (3.22)$$

Το σύνολο της πληροφορίας του συστήματος αντιστοιχεί σε $d = N$ όπου η απόσταση (3.15) μηδενίζεται. Το σχετικό πληροφοριακό περιεχόμενο μίας βάσης διάστασης d δίνεται ως εξής: [8], [9]

$$I(d) = \frac{\sum_{i=1}^d \lambda_j}{\sum_{j=1}^n \lambda_j} \quad (3.23)$$

Στα δυναμικά συστήματα, οι μεγάλες ιδιοτιμές αντιστοιχούν σε κύρια και σημαντικά χαρακτηριστικά – ιδιότητες του συστήματος ενώ μικρότερες ιδιοτιμές αντιστοιχούν σε μικρές συνεισφορές της ολικής δυναμικής συμπεριφοράς. Κύριος στόχος θα πρέπει να είναι η επιλογή της διάστασης d της βάσης ούτως ώστε το σχετικό πληροφοριακό περιεχόμενο να τείνει στη μονάδα. Συνεπώς, επιλέγεται από το σχεδιαστή ένα κλάσμα p της πληροφορίας που είναι επιθυμητό να έχει η προσέγγιση [9], [8].

$$I(d) = \min\{d \in N | I(d) \geq p\} \quad (3.24)$$

Είναι ένας εύκολος τρόπος να προεπιλεγεί ένα ποσοστό πληροφορίας δηλαδή της διάστασης d της βάσης ούτως ώστε να περιέχονται σημαντικές πληροφορίες της δυναμικής συμπεριφοράς του συστήματος που επιχειρείται να προσεγγιστεί.

Η βέλτιστη προβολή είναι τελικά:

$$\Pi_d = \sum_{j=1}^d \mathbf{u}_j \mathbf{u}_j^T \quad (3.25)$$

Η δράση της προβολής Π_d σε κάθε διάνυσμα $\mathbf{y}_i(t)$ που αποτελεί και την προσέγγιση με τη μέθοδο ανάλυσης κύριων συνιστωσών, είναι η εξής: [8]

$$\Pi_d \mathbf{y}_i(t) = \hat{\mathbf{y}}_i(t) = \sum_{j=1}^d y_{ij}(t) \mathbf{u}_j \quad (3.26)$$

Διατηρούνται δηλαδή οι πρώτες d συνιστώσες από τις συνολικά m της αρχικής σχέσης του πίνακα \mathbf{Y} οι οποίες περιλαμβάνουν το επιλεγόμενο κάθε φορά πληροφοριακό περιεχόμενο του συστήματος.

3.5 Μέθοδος των Στιγμιοτύπων

Η επίλυση προβλημάτων της σύγχρονης μηχανικής μπορεί να οδηγήσει πολλές φορές σε μεγάλες τιμές της αρχικής διάστασης N , συνεπώς η εφαρμογή της μεθόδου των κύριων συνιστωσών και ειδικότερα η εύρεση του πίνακα συσχέτισης \mathbf{K} καθίσταται από υπολογιστικής άποψης δύσκολη ή αδύνατη. Για την αντιμετώπιση της δυσκολίας αυτής, ο L.

Sirovich πρότεινε τη μέθοδο των στιγμιοτύπων (method of snapshots), παραλλαγή της ανάλυσης κύριων συνιστωσών [8], [10].

Αντίθετα με την κλασική μεθοδολογία όπου χρησιμοποιούνται οι τροχιές $\mathbf{y}_i(\mathbf{t})$ σε συνεχές χρονικό διάστημα, στη μέθοδο των στιγμιοτύπων χρησιμοποιούνται τα στιγμιότυπα των τροχιών σε διακριτές χρονικές στιγμές. Το σύνολο των δεδομένων μετατρέπεται στον πίνακα ως εξής:

$$\mathbf{Y} = [\mathbf{Y}_1(t) \ \cdots \ \mathbf{Y}_E(t)] \in \mathbb{R}^{N \times m} \quad (3.27)$$

Όπου $\mathbf{Y}_i(t) \in \mathbb{R}^{N \times m_i}$, $i = 1, \dots, E$ είναι ο πίνακας της i τροχιάς και E ο αριθμός του συνόλου των επιβολών – τροχιών. Κάθε τροχιά αποτελείται από m_i στιγμιότυπα. Άρα το σύνολο των στιγμιοτύπων που περιλαμβάνονται στον πίνακα \mathbf{Y} ορίζεται ως εξής:

$$m = \sum_{i=1}^E m_i \quad (3.28)$$

Είναι δηλαδή ο συνολικός αριθμός των στιγμιοτύπων. Κάθε στήλη του πίνακα \mathbf{Y} αποτελεί ένα στιγμιότυπο του συστήματος [8].

Ο πίνακας συσχέτισης μπορεί να υπολογισθεί ως εξής:

$$\mathbf{K} = \mathbf{Y}\mathbf{Y}^T \in \mathbb{R}^{N \times N} \quad (3.29)$$

του οποίου οι ιδιοτιμές ταυτίζονται με του πίνακα:

$$\mathbf{K}' = \mathbf{Y}^T\mathbf{Y} \in \mathbb{R}^{m \times m} \quad (3.30)$$

Οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα \mathbf{K}' είναι ευκολότερο από υπολογιστικής πλευράς να βρεθούν διότι συνήθως ισχύει $m \ll N$. Ο υπολογισμός τους πραγματοποιείται ως εξής:

$$\mathbf{K}'\mathbf{u}_j = \lambda_j\mathbf{u}_j \quad (3.31)$$

Όπου τα ιδιοδιανύσματα \mathbf{u}_j επιλέγονται να είναι ορθοκανονικά. Υπενθυμίζεται ότι οι ιδιοτιμές είναι διατεταγμένες σε φθίνουσα σειρά: $\lambda_1 \geq \dots \geq \lambda_n \geq 0$. Στη συνέχεια, ανακτώνται τα ιδιοδιανύσματα \mathbf{v}_j του πίνακα \mathbf{K} , από τα οποία επιλέγονται τα d πρώτα ως η βάση του μειωμένου τάξης μοντέλου. Ο υπολογισμός τους πραγματοποιείται ως εξής: [8]

$$\mathbf{v}_j = \frac{1}{\sqrt{\lambda_j}}\mathbf{Y}\mathbf{u}_j, \quad j = 1, \dots, m \quad (3.32)$$

Κάθε στιγμιότυπο προσεγγίζεται από τη σχέση (3.13) ως εξής:

$$\mathbf{y}_i = \sum_{j=1}^d a_{ij}\mathbf{v}_j \quad (3.33)$$

όπου οι συντελεστές a_{ij} υπολογίζονται ως εξής:

$$a_{ij} = \mathbf{y}_i^T\mathbf{v}_j \quad (3.34)$$

Στη συνέχεια μπορεί να ορισθεί η ορθοκανονική βάση από τα d πρώτα ιδιοδιανύσματα ως εξής:

$$\mathbf{Z} = [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_d] \in \mathbb{R}^{N \times d} \quad (3.35)$$

και \mathbf{a}_i το διάνυσμα των συντελεστών που περιγράφουν τη χρονική εξάρτηση και αντιστοιχούν στο στιγμιότυπο i :

$$\mathbf{a}_i = [a_{i1} \quad \dots \quad a_{id}] \in \mathbb{R}^d \quad (3.36)$$

Επίσης, λόγω της ορθοκανονικότητας των διανυσμάτων βάσης, ισχύει επιπροσθέτως:

$$\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d \quad (3.37)$$

Όπου \mathbf{I}_d είναι ο τετραγωνικός μοναδιαίος πίνακας διάστασης d .

Επιπλέον μπορεί να ορισθεί ο πίνακας των συντελεστών \mathbf{A} που αντιστοιχούν σε όλα τα στιγμιότυπα:

$$\mathbf{A} = [\mathbf{a}_1 \quad \dots \quad \mathbf{a}_m] \in \mathbb{R}^{d \times m} \quad (3.38)$$

Γενικεύοντας, ορίζεται η προσέγγιση που αντιστοιχεί σε ολόκληρο τον πίνακα \mathbf{Y} :

$$\hat{\mathbf{Y}} = \mathbf{Z}\mathbf{A} \quad (3.39)$$

$$\mathbf{A} = \mathbf{Z}^T \mathbf{Y} \quad (3.40)$$

Οι παραπάνω σχέσεις μπορεί να αναφέρονται σε ολόκληρο τον πίνακα \mathbf{Y} είτε στον πίνακα μια τροχιάς, όπου πλέον m θα είναι ο αριθμός των στιγμιότυπων της τροχιάς.

3.5.1 Κανονικοποίηση Δεδομένων

Κατά την εφαρμογή της μεθόδου της ανάλυσης των κύριων συνιστωσών συνιστάται η τροποποίηση των δεδομένων – πειραματικών ή υπολογιστικών – ούτως ώστε τα στοιχεία του πίνακα \mathbf{Y} να είναι ομοιόμορφα καταμεμημένα [6].

Ειδικότερα, στην παρούσα εργασία, η λύση του συστήματος είναι οι κατανομές στο χώρο και στο χρόνο των πεδίων $P(r, 0)$, $\mathbf{u}(r, 0)$ και $T(r, 0)$, ένα διάνυσμα κατάστασης διάστασης $N = 60264$. Οι τιμές των θερμοκρασιών στο χώρο διαφέρουν σημαντικά από τις τιμές της πίεσης ή της ταχύτητας, συνεπώς μη κανονικοποίηση του πίνακα \mathbf{Y} μπορεί να οδηγήσει σε σημαντικές αποκλίσεις της τελικής προσέγγισης, δηλαδή να περιγράψετε σωστά μόνο την κατάσταση της θερμοκρασίας [6].

Αρχικά, υπολογίζεται η μέση τιμή για κάθε γραμμή του πίνακα \mathbf{Y} , δηλαδή για την τιμή κάθε μεγέθους σε κάθε θέση του συστήματος:

$$\bar{x}_i = \frac{1}{m} \sum_{j=1}^m (\mathbf{Y})_{ij}, i = 1, \dots, N \quad (3.41)$$

Επίσης, υπολογίζεται και η τυπική απόκλιση για κάθε γραμμή του πίνακα Y :

$$s_i = \left(\frac{1}{m-1} \sum_{j=1}^m ((Y)_{ij} - \bar{x}_i) \right)^{0.5} \quad i = 1, \dots, N \quad (3.42)$$

Κατά την κανονικοποίηση, αφαιρείται η μέση τιμή \bar{x}_i από κάθε αντίστοιχο στοιχείο του κάθε στιγμιότυπου και στη συνέχεια κάθε στοιχείου το διάνυσματος που έχει προκύψει διαιρείται με την αντίστοιχη τυπική απόκλιση. Η δεύτερη περίπτωση της σχέσης (3.43) είναι πρακτικά επέκταση της 1^{ης} διότι όταν $s_i = 0$ συνεπάγεται από τη σχέση (3.42) ότι $(Y)_{ij} = \bar{x}_i$ δηλαδή κάθε στοιχείο της γραμμής i ισούται με τη μέση τιμή \bar{x}_i . Η περίπτωση αυτή πραγματοποιείται μόνο σε σύνορα του πεδίου όπου επιβάλλονται σταθερές τιμές συγκεκριμένων μεγεθών. Στο συγκεκριμένο αντιδραστήρα αυτό πραγματοποιείται στα ψυχόμενα τοιχώματα, όπου επιβάλλεται σταθερή θερμοκρασία 300 Κ. Μαθηματικά η τροποποίηση των δεδομένων συντελείται ως εξής: [6]

$$u_{ij} = \begin{cases} \frac{(Y)_{ij} - \bar{x}_i}{s_i}, & s_i \neq 0 \\ 0, & s_i = 0 \end{cases} \quad i = 1, \dots, N, j = 1, \dots, m \quad (3.43)$$

Μετά την εύρεση της ορθοκανονικής βάσης καθώς και της προσέγγισης του πίνακα Y , θα πρέπει να ακολουθηθεί η αντίστροφη λογική, δηλαδή η προσέγγιση να μην παρέχεται σε κανονικοποιημένη μορφή αλλά στην πραγματική. Συγκεκριμένα, σε κάθε στοιχείο της προσέγγισης του στιγμιότυπου θα πρέπει να πολλαπλασιαστεί με αντίστοιχη τυπική απόκλιση s_i και να προστεθεί στη συνέχεια η μέση τιμή, \bar{x}_i [6].

$$\widehat{Y}_N = ZA \quad (3.44)$$

$$\widehat{(Y)}_{ij} = \bar{x}_i + s_i \widehat{Y}_N \quad (3.45)$$

όπου \widehat{Y}_N η προσέγγιση του κανονικοποιημένου πλέον πίνακα.

Σημειώνεται ότι η σχέση (3.45) ισχύει τόσο για τροχιές οι οποίες χρησιμοποιήθηκαν για την κατασκευή της ορθοκανονικής βάσης, ανήκουν δηλαδή στον πίνακα Y , αλλά και σε άλλες που δεν ανήκουν. Στη δεύτερη περίπτωση το μέσο στιγμιότυπο \bar{x}_i και η τυπική απόκλιση s_i δεν προκύπτουν από την κανονικοποίηση της συγκεκριμένης τροχιάς αλλά ολόκληρου του πίνακα Y πριν την κατασκευή της βάσης Z .

3.6 Αξιολόγηση της μείωσης τάξης – Σφάλματα

Κατά τη διάρκεια των υπολογισμών, για την αξιολόγηση μιας υποψήφιας βάσης χρησιμοποιείται το σχετικό σφάλμα της εκτίμησης ενός στιγμιότυπου – διάνυσμα κατάστασης $x(t)$ - το οποίο ορίζεται ως εξής:

$$\mathbf{Error}(x(t)) = \frac{\|\widehat{x}(t) - x(t)\|}{\|x(t)\|} \quad (3.46)$$

Επίσης, εκτός από την εξέλιξη του σχετικού σφάλματος συναρτήσει του χρόνου υπολογίστηκε και το μέσο τετραγωνικό σφάλμα το οποίο ορίζεται ως εξής:

$$\mathbf{avererror}(t) = \left(\frac{1}{m} \sum_{j=1}^m (\mathbf{Error}(x(t)))^2 \right)^{0.5} \quad (3.47)$$

Καθώς και το μέγιστο σχετικό σφάλμα το οποίο ορίζεται:

$$\mathbf{maxerror}(t) = \max(\mathbf{Error}(x(t))) \quad (3.48)$$

Σχεδιασμός Νευρωνικού Δικτύου

4.1 Εισαγωγή

Η υπολογιστική νοημοσύνη (computational intelligence) αποτελεί έναν από τους σημαντικότερους και πλέον εξελισσόμενους κλάδους της πληροφορικής αλλά και της τεχνητής νοημοσύνης γενικότερα. Στον κλάδο της υπολογιστικής νοημοσύνης ανήκουν και τα τεχνητά νευρωνικά δίκτυα (artificial neural networks) [11].

Η μελέτη της λειτουργίας του ανθρώπινου εγκεφάλου αποτέλεσε το έναυσμα για την έρευνα και εφαρμογή μαθηματικών μοντέλων και αλγορίθμων, σε μια προσπάθεια προσομοίωσης της λειτουργίας αυτής. Η νευροεπιστήμη (neuroscience) είναι εκείνος ο κλάδος που αρχικά προσπάθησε να εξηγήσει τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου, χωρίς ωστόσο κάτι τέτοιο να είναι απόλυτα ξεκάθαρο ακόμα και σήμερα. Πάνω σε αυτή τη λογική βασίζεται και η φιλοσοφία των τεχνητών νευρωνικών δικτύων, ως μια προσπάθεια προσομοίωσης, με τη βοήθεια υπολογιστών, του ανθρώπινου νευρικού συστήματος [11].

Το νευρικό σύστημα των ζώντων οργανισμών αναπτύσσεται και εκπαιδεύεται συνεχώς μέσω ερεθισμάτων που λαμβάνει από το περιβάλλον, ώστε να επεξεργάζεται τις πληροφορίες με το σωστό τρόπο. Αντίστοιχα, τα τεχνητά νευρωνικά δίκτυα εκπαιδεύονται στο τρόπο με τον οποίο θα επεξεργάζονται τα δεδομένα, μέσω αλγορίθμων εκπαίδευσης και δεδομένα εκπαίδευσης [11].

Σημειώνεται ότι τα νευρωνικά δίκτυα δεν απαιτούν καμία πρότερη γνώση του φυσικού συστήματος το οποίο καλούνται να προσεγγίσουν, δηλαδή των εξισώσεων και των μηχανισμών που το διέπουν. Το νευρωνικό δίκτυο εκπαιδεύεται με χρήση θεωρητικών, πειραματικών ή υπολογιστικών δεδομένων, και μπορεί στη συνέχεια να αναπαραγάγει και να γενικεύσει τη συμπεριφορά του συστήματος, στη συγκεκριμένη περιοχή εκπαίδευσης [6].

Ως συμπέρασμα προκύπτει ότι ένα σύστημα μπορεί να προσομοιωθεί από το συνδυασμό των νευρωνικών δικτύων και την ανάλυση κύριων συνιστωσών χωρίς καμία γνώση του λεπτομερούς συστήματος, με μοναδική απαίτηση αποτελέσματα κατάλληλων πειραμάτων ή προσομοιώσεων από το σύστημα [11].

4.2 Βιολογικά Νευρωνικά Δίκτυα

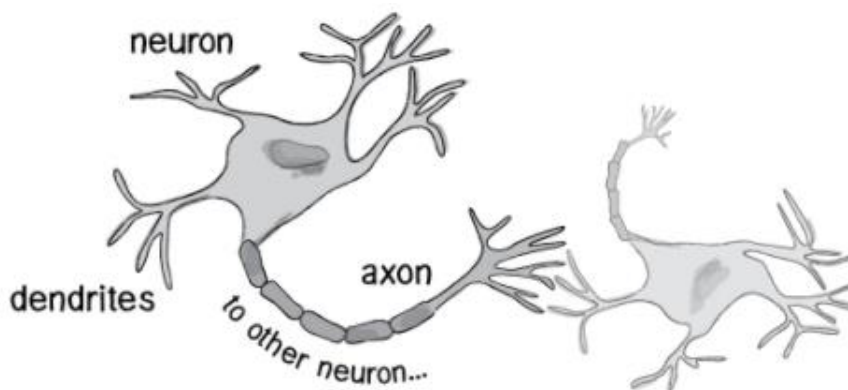
Τα βιολογικά νευρωνικά δίκτυα που υπάρχουν στον ανθρώπινο εγκέφαλο αποτελούνται από νευρώνες. Ως νευρώνας νοείται το μικρότερο τμήμα του εγκεφάλου που είναι ικανό να επεξεργάζεται πληροφορίες, η ύπαρξη του οποίου διαφοροποιεί τα ζώα από τα φυτά, τα οποία δεν έχουν νευρώνες. Ένας νευρώνας λειτουργεί όπως και τα άλλα κύτταρα ενός οργανισμού και ο απόλυτος αριθμός τους σε έναν ανθρώπινο εγκέφαλο εκτιμάται περίπου στους 10^{10} [11].

Κάθε νευρώνας συνδέεται με πολλούς άλλους νευρώνες με τις συνδέσεις αυτές να ονομάζονται συνάψεις. Κάθε νευρώνας συμμετέχει σε περίπου 10^4 συνάψεις. Έτσι, ως νευρωνικό δίκτυο ορίζεται ένας αριθμός νευρώνων μαζί με τις διασυνδέσεις τους. Το συνολικό σύστημα των νευρωνικών δικτύων στον ανθρώπινο οργανισμό ονομάζεται

κεντρικό νευρικό σύστημα (central nervous system). Ένας νευρώνας αποτελείται από: [12], [11]

- Το σώμα του κυττάρου, στο οποίο βρίσκεται ο πυρήνας και άρα το γενετικό υλικό του οργανισμού. Εκεί καταλήγουν όλοι οι δενδρίτες, όπου συγκεντρώνονται όλα τα ερεθίσματα, αθροίζονται, δημιουργώντας έτσι το τελικό σήμα που δέχεται ο νευρώνας. Η τιμή αυτή καθορίζει κατά πόσο θα ενεργοποιηθεί ή όχι ο νευρώνας, το οποίο συντελείται όταν αυτή ξεπεράσει μια συγκεκριμένη τιμή-κατώφλι (threshold) συγκεκριμένη για κάθε κύτταρο. Η ύπαρξη του κατωφλίου σε κάθε κύτταρο είναι απαραίτητη για να μην πραγματοποιείται τυχαία ενεργοποίηση του νευρώνα από διακυμάνσεις του κυτταρικού δυναμικού.
- Τον άξονα, ο οποίος μεταδίδει τα εξερχόμενα σήματα από το σώμα του κυττάρου σε άλλους νευρώνες. Μέσω του άξονα και χάρη στη δομή του, μεταδίδονται τα σήματα μεταξύ των νευρώνων διατηρώντας αναλλοίωτο το πληροφοριακό τους περιεχόμενο.
- Τους δενδρίτες, οι οποίοι αποτελούν λεπτές επεκτάσεις του κυττάρου, και έρχονται σε επαφή με άλλους νευρώνες δεχόμενοι τα εισερχόμενα σήματα από το εξωτερικό ή/και εσωτερικό περιβάλλον. Οι δενδρίτες αναπτύσσονται σε διακλαδωμένες μορφές που ονομάζονται δενδριτικά δένδρα. Με αυτή τη διαδικασία, κάθε νευρώνας έχει τη δυνατότητα να δέχεται πληροφορίες από πολλές πηγές ταυτόχρονα.
- Τις συνάψεις, οι οποίες αποτελούν τις δομές οι οποίες βρίσκονται στις απολήξεις του άξονα και ευθύνονται για τη μετάδοση πληροφορίας μεταξύ των κυττάρων. Κάθε σύναψη συνδέεται με δενδρίτες άλλων κυττάρων ή με κάποιο όργανο απόκρισης, πχ ένα μύ. Σε κάθε σύναψη υπάρχουν αποθηκευμένοι σε κυστίδια νευροδιαβιβαστές, οι οποίοι απελευθερώνονται όταν αντιληφθούν ένα ισχυρό σήμα και προσκολλώνται στον αντίστοιχο δενδρίτη, μεταφέροντας έτσι την πληροφορία στα υπόλοιπα κύτταρα και κατά συνέπεια σε όλον τον ανθρώπινο οργανισμό.

Στην Εικόνα 4-1 παρουσιάζεται σχηματικά ένας τυπικός βιολογικός νευρώνας.



Εικόνα 4-1: Βιολογικός Νευρώνας

Αξίζει να αναφερθεί ότι οι νευρώνες διακρίνονται σε δύο ήδη: [6]

- Στους νευρώνες με χημικές συνάψεις όπου η πληροφορία διαδίδεται μέσω της ροής των νευροδιαβιβαστών (neurotransmitters) από το κύτταρο – πομπό στο κύτταρο – δέκτη.

- Στους νευρώνες με ηλεκτρικές συνάψεις όπου η πληροφορία διαδίδεται μέσω της ροής ηλεκτρονίων με την επαφή δύο κυττάρων.

4.3 Τεχνητά Νευρωνικά Δίκτυα

Με τα τεχνητά νευρωνικά δίκτυα επιχειρείται η προσέγγιση της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή καθώς έχει την ικανότητα να εκτελεί υπολογισμούς με μαζικό παράλληλο τρόπο. Τα δίκτυα αυτά βασίζονται στη μαθηματική μοντελοποίηση των νευρώνων με χημικές συνάψεις.

Ένα τεχνητό νευρωνικό δίκτυο προσομοιώνει τη λειτουργία του ανθρώπινου εγκεφάλου που σκοπό έχει την αποτελεσματικότερη, ταχύτερη και παράλληλη επεξεργασία πληροφοριών. Επί της ουσίας τα δίκτυα αυτά είναι μία αλληλουχία διασυνδεδεμένων στοιχείων επεξεργασίας τα οποία ονομάζονται νευρώνες κατά αντιστοιχία των βιολογικών νευρώνων. Οι νευρώνες συνδέονται μεταξύ τους και είναι οργανωμένοι σε επίπεδα (layers). Κάθε επιμέρους νευρώνας μπορεί να έχει πολλές εισόδους (inputs) για εισερχόμενη πληροφορία αλλά μόνο μία έξοδο (output) απ' όπου και εξέρχεται το σήμα εξόδου.

Επιπλέον χαρακτηριστικά στοιχεία ενός τεχνητού νευρώνα αποτελούν οι συνάψεις, ο αθροιστής και η συνάρτηση ενεργοποίησης. Οι συνάψεις αποτελούν τη σύνδεση των νευρώνων μεταξύ τους και χαρακτηρίζονται από το αντίστοιχο συναπτικό βάρος κάθε σήματος εισόδου. Ο αθροιστής αθροίζει όλα τα σήματα εισόδου αφού προηγουμένως έχουν σταθμιστεί με τα αντίστοιχα βάρη, και η συνάρτηση ενεργοποίησης μετασχηματίζει τα σήματα εισόδου σε ένα σήμα εξόδου. Οι συναρτήσεις ενεργοποίησης μπορεί να είναι είτε γραμμικές είτε μη γραμμικές.

Η βασική ομοιότητα των τεχνητών νευρωνικών δικτύων σε σχέση με τα βιολογικά νευρωνικά δίκτυα εντοπίζεται στον τρόπο με τον οποίο μεταχειρίζονται την πληροφορία την οποία καλούνται να αξιοποιήσουν. Συγκεκριμένα ένα τεχνητό δίκτυο λαμβάνει πληροφορίες από κάποιο συγκεκριμένο χώρο δεδομένων – υπολογιστικών ή πειραματικών – μέσω μιας διαδικασίας εκμάθησης (training phase). Τέλος, η πληροφορία – γνώση, αποθηκεύεται στο τεχνητό δίκτυο μέσω των συνάψεων και των αντίστοιχων βαρών, με στόχο την χρήση της γνώσης αυτής σε μετέπειτα εφαρμογές [11].

4.4 Μαθηματική Μοντελοποίηση Τεχνητού Νευρωνικού Δικτύου

Το μαθηματικό μοντέλο του νευρώνα παρουσιάζεται γραφικά στην Εικόνα 4-2 και ορίζεται μαθηματικά από την εξής σχέση: [11], [12]

$$y(t) = \varphi \left(\sum_{j=1}^n w_j x_j(t) + \theta \right) \quad (4.1)$$

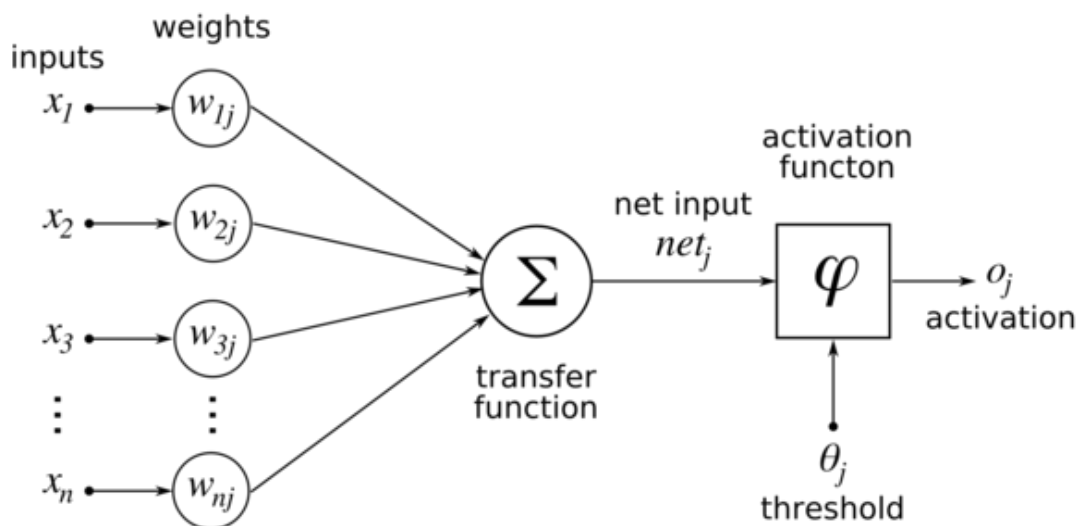
Στα βασικά χαρακτηριστικά ενός τεχνητού νευρώνα περιλαμβάνονται: [11], [12]

- τα σήματα εισόδου $x_j(t)$, τα οποία προέρχονται είτε από το περιβάλλον είτε από την έξοδο άλλων νευρώνων.
- τα συναπτικά βάρη w_j , οι συντελεστές με τους οποίους πολλαπλασιάζονται τα σήματα εισόδου. Όσο μεγαλύτερη είναι η τιμή τους, τόσο σημαντικότερη είναι η συνεισφορά τους στη διαμόρφωση του τελικού αποτελέσματος. Θετική τιμή ενός συναπτικού βάρους αντιστοιχεί σε ενισχυτικό σήμα, ενώ αρνητική σε παρεμποδιστικό.

- Η συνάρτηση ενεργοποίησης $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, η οποία μετατρέπει το σήμα εισόδου σε σήμα εξόδου, $y(t)$. Ο βασικός στόχος της συνάρτησης ενεργοποίησης είναι να μοντελοποιήσει και να προσεγγίσει, στο πλαίσιο της παρούσας εργασίας, τη συμπεριφορά του συστήματος. Μια συνάρτηση ενεργοποίησης μπορεί να είναι μεταξύ άλλων γραμμική, ημι-γραμμική, σιγμοειδής ή εφασπτομενική.
- Συνάρτηση Κατωφλιού θ (Threshold function), η οποία ορίζει το συγκεκριμένο όριο που θα πρέπει να ξεπεράσει ο όρος των συναπτικών βαρών ούτως ώστε να ενεργοποιηθεί η συνάρτηση ενεργοποίησης φ . Η βηματική συνάρτηση ενεργοποίησης μπορεί να γραφεί ως εξής:

$$\varphi(u) = \begin{cases} 0 & \sum_{j=1}^n w_j x_j(t) \leq -\theta \\ 1 & \sum_{j=1}^n w_j x_j(t) > -\theta \end{cases} \quad (4.2)$$

- Το σήμα εξόδου του νευρώνα $y(t)$, το οποίο είναι το αποτέλεσμα της επεξεργασίας των σημάτων εισόδου με βάση τον αλγόριθμο που έχει εκπαιδευτεί για συγκεκριμένα δεδομένα εκπαίδευσης.



Εικόνα 4-2: Γραφική Αναπαράσταση Τεχνητού Νευρωνικού Δικτύου

Με βάση τα παραπάνω τα σήματα εισόδου εκφράζονται σε διανυσματική μορφή, θεωρώντας τον όρο θ ως το συναπτικό βάρος ενός επιπλέον σταθερού σήματος εισόδου $x_0(t) = 1$, ως εξής: [6]

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} 1 \\ x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad (4.3)$$

Το διάνυσμα των συναπτικών βαρών ορίζεται:

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \theta \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \quad (4.4)$$

Οπότε η σχέση (4.1) γράφεται:

$$y(t) = \varphi(\mathbf{w}^T \mathbf{x}(t)) \quad (4.5)$$

4.5 Αρχιτεκτονική Τεχνητού Νευρωνικού Δικτύου

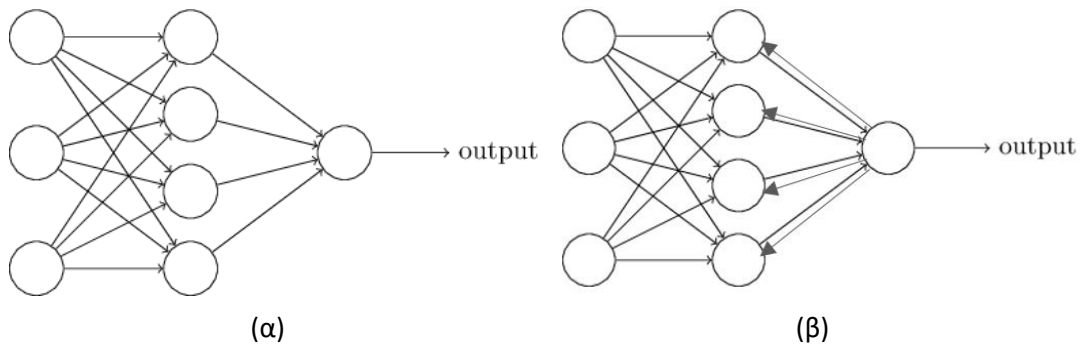
Οι τεχνητοί νευρώνες οργανώνονται σε διάφορα επίπεδα παράλληλης διάταξης. Η αλληλουχία των παράλληλων αυτών επιπέδων είναι αυτή που αποτελεί και το σύνολο ενός τεχνητού νευρωνικού δικτύου. Φυσικά, δεν είναι όλα τα επίπεδα ίδια αλλά διαφοροποιούνται ως εξής: [11]

- Το επίπεδο εισόδου (input layer) αποτελεί το 1^ο επίπεδο ενός τεχνητού δικτύου, όπου σε αυτό εμφανίζονται τα δεδομένα εισόδου. Συγχρόνως, ο αριθμός των νευρώνων του είναι ίσος με τις μεταβλητές των δεδομένων.
- Τα κρυφά επίπεδα (hidden layers) συνδέουν το επίπεδο εισόδου με το επίπεδο εξόδου των νευρώνων. Ο αριθμός των κρυφών επιπέδων που περιέχονται σε ένα δίκτυο, ορίζονται από το σχεδιαστή – χρήστη και επηρεάζει το ‘βάθος’ και τη πολυπλοκότητα του.
- Το επίπεδο εξόδου (output layer) αποτελεί το τελευταίο επίπεδο ενός τεχνητού νευρωνικού δικτύου. Σε αυτό εμφανίζονται τα τελικά αποτελέσματα μετά το τέλος της εκμάθησης και ο αριθμός των νευρώνων του είναι ίσος με τις πιθανές μεταβλητές εξόδου των αποτελεσμάτων.

Τα τεχνητά δίκτυα κατηγοριοποιούνται ανάλογα με την αρχιτεκτονική τους και τον τρόπο με τον οποίο συνδέονται οι νευρώνες μεταξύ τους:

- Σε νευρωνικά δίκτυα με πρόσθια τροφοδότηση (feed-forward networks). Σε αυτή την περίπτωση το σήμα μεταφέρεται, μέσω των νευρώνων, από την είσοδο έως την έξοδο χωρίς όμως να μπορεί να συμβεί το αντίστροφο. Αυτά τα δίκτυα μπορούν με τη σειρά τους να διακριθούν σε μονοεπίπεδα δίκτυα (π.χ. perceptron) ή πολυεπίπεδα δίκτυα (π.χ. multilayer perceptron), ανάλογα με το εάν σε αυτά περιλαμβάνεται και κάποιο ή κάποια κρυφά επίπεδα. Ένα δίκτυο τέτοιου είδους χαρακτηρίζεται από στατικότητα και δεν έχει μνήμη δηλαδή η απόκριση σε ένα σήμα εισόδου εξαρτάται αποκλειστικά από αυτό, και είναι ανεξάρτητη από τα προηγούμενα σήματα. Συνεπώς, κάθε σήμα εισόδου μπορεί να οδηγήσει σε ένα σήμα απόκρισης.
- Σε νευρωνικά δίκτυα με ανατροφοδότηση (feedback network or recurrent network). Σε αυτή την περίπτωση τα νευρωνικά δίκτυα περιέχουν τουλάχιστον ένα βρόγχο

ανατροφοδότησης (feedback loop). Έτσι, ένα επίπεδο νευρώνων έχει το ρόλο τροφοδότησης του σήματος εξόδου πίσω στις εισόδους όλων των άλλων νευρώνων. Το δίκτυο αυτό χαρακτηρίζεται από δυναμικότητα, αφού η απόκριση επηρεάζεται εκτός από τα δεδομένα εισόδου και από ένα ή περισσότερα προηγούμενα αποτελέσματα τα οποία μπορεί να είναι είτε διανύσματα εξόδου είτε ενδιάμεσα διανύσματα. Συνεπώς, ένα διάνυσμα απόκρισης δημιουργείται με χρήση τόσο των σημάτων εισόδου όσο και αυτών των διανυσμάτων απόκρισης των προηγούμενων χρονικών στιγμών.



Εικόνα 4-3: (α) Νευρωνικό Δίκτυο Πρόσθιας Τροφοδότησης
(β) Νευρωνικό Δίκτυο με ανατροφοδότηση

Στην Εικόνα 4-3 παρουσιάζονται γραφικά ένα δίκτυο πρόσθιας τροφοδότησης και ένα δίκτυο με ανατροφοδότηση. Οι κύκλοι αριστερά αναπαριστούν τους κόμβους εισόδου, είναι δηλαδή ένα επίπεδο εισόδου, οι κύκλοι δεξιά αναπαριστούν το επίπεδο εξόδου, ενώ οι κύκλοι που βρίσκονται στη μέση αναπαριστούν τα κρυφά επίπεδα. Για τη σύνδεση μεταξύ 2 επιπέδων απαιτείται η ύπαρξη ενός συναπτικού βάρους, τα οποία αναπαρίστανται από τις συνδέσεις μεταξύ των κύκλων.

4.5.1 Δίκτυο Πρόσθιας Τροφοδότησης

Το απλούστερο τεχνητό νευρωνικό δίκτυο είναι αυτό της πρόσθιας τροφοδότησης με ένα επίπεδο νευρώνων, όπου το σήμα εισόδου μετατρέπεται απευθείας σε σήμα εξόδου. Στην περίπτωση που το δίκτυο αποτελείται από L νευρώνες, το σήμα εξόδου μετατρέπεται με τη βοήθεια της συνάρτησης ενεργοποίησης και των συναπτικών βαρών σε L σήματα εξόδου. Το σήμα εξόδου του k νευρώνα αποτελεί σήμα εισόδου του $k + 1$ νευρώνα. Το σήμα εξόδου κάθε νευρώνα ορίζονται ως εξής:

$$y(t) = \varphi \left(\sum_{j=1}^n w_{k,j} x_j(t) + \theta_k \right), \quad k = 1, \dots, L \quad (4.6)$$

όπου θ_k το συναπτικό βάρος που οδηγεί το j σήμα εισόδου στον k νευρώνα και $\theta_k = w_{k,0}$ είναι η παράμετρος μεροληψίας του k νευρώνα. Στην Εικόνα 4-3 (α) παρουσιάστηκε αναλυτικά η δομή ενός δικτύου πρόσθιας τροφοδότησης.

Η μετατροπή του σήματος εισόδου σε σήμα απόκρισης καθορίζεται πλήρως από τον πίνακα των συναπτικών βαρών, κάθε στήλη του οποίου αποτελεί το συναπτικό βάρος κάθε νευρώνα. Όσο πιο πολύπλοκο είναι το φυσικό σύστημα το οποίο θα πρέπει να προσεγγιστεί, τόσο περισσότερα κρυφά επίπεδα περιέχει το νευρωνικό δίκτυο που

εκπαιδεύεται. Ωστόσο, θα πρέπει ο αριθμός των κρυφών επιπέδων να είναι αρκούντως μικρός ούτως ώστε να αποφευχθεί το φαινόμενο της υπερ-προσαρμογής. Το φαινόμενο αυτό παρατηρείται στην περίπτωση κατά την οποία δημιουργούνται πολυώνυμα μεγάλου βαθμού για να περιγράψουν φαινόμενα πολύ μικρότερης τάξης.

4.5.2 Δίκτυο με Ανατροφοδότηση - Δίκτυο NARX

Από τα δίκτυα με ανατροφοδότηση παρουσιάζεται εκτενώς μόνο το δίκτυο NARX (Nonlinear Auto Regressive with exogenous inputs model), το οποίο θα χρησιμοποιηθεί στο πλαίσιο της παρούσας μελέτης. Το συγκεκριμένο δίκτυο έχει πλεονεκτήματα όσον αφορά τα υπόλοιπα δίκτυα με ανατροφοδότηση καθώς έχει την απλούστερη δομή και συγχρόνως συγκλίνει γρηγορότερα κάνοντας παράλληλα καλύτερη γενίκευση [11], [13], [12].

Μαθηματικά το δίκτυο NARX αναπαρίσταται ως εξής:

$$\mathbf{y}(t_k) = \mathbf{f}(\mathbf{x}(t_{k-n_x}), \dots, \mathbf{x}(t_{k-1}), \mathbf{x}(t_k), \mathbf{y}(t_{k-n_y}), \dots, \mathbf{y}(t_{k-1})) \quad (4.7)$$

όπου $\mathbf{x}(t)$ και $\mathbf{y}(t)$ τα σήματα εισόδου και εξόδου αντίστοιχα και \mathbf{f} μια οποιαδήποτε γραμμική ή μη συνάρτηση που μπορεί να υπολογιστεί από ένα δίκτυο πρόσθιας τροφοδότησης πολλαπλών επιπέδων.

Ειδικότερα, το εκπαιδευόμενο δίκτυο μπορεί να υπολογίσει μια εκτίμηση $\hat{\mathbf{y}}(t_k)$ τη χρονική στιγμή t_k , με χρήση του διανύσματος εισόδου τη χρονική στιγμή t_k , αλλά και τις προηγούμενες χρονικές στιγμές n_x καθώς και τα διανύσματα εξόδου τις προηγούμενες χρονικές στιγμές n_y . Η εφαρμογή του δικτύου NARX έγινε μέσω της εργαλειοθήκης του Matlab όπου χρησιμοποιείται μια παραλλαγμένη σχέση του ορισμού (4.7) η οποία ορίζεται ως εξής: [11] [13]

$$\mathbf{y}(t_k) = \mathbf{f}(\mathbf{x}(t_{k-n_x}), \dots, \mathbf{x}(t_{k-1}), \mathbf{y}(t_{k-n_y}), \dots, \mathbf{y}(t_{k-1})) \quad (4.8)$$

Από τη σχέση απουσιάζει ο όρος $\mathbf{x}(t_k)$. Παρ' όλα αυτά οι σχέσεις (4.7) και (4.8) δεν διαφέρουν ουσιαστικά εάν μετατοπιστεί το διάνυσμα $\mathbf{x}(t_k)$ κατά μία χρονική στιγμή, μετατραπεί δηλαδή σε $\mathbf{x}(t_{k+1})$. Στο πρόβλημα που επιλύεται στην παρούσα εργασία όπως ορίστηκε στο Κεφάλαιο 2, η μορφή των εξισώσεων (2.1)-(2.4) καθώς και η διακριτοποίηση στο χρόνο με την οποία επιλύονται δείχνουν ότι δεν υπάρχει εξάρτηση από πολλές χρονικές στιγμές στο παρελθόν. Για το λόγο αυτό επιλέγεται $n_x = n_y = 1$ οπότε η σχέση (4.8) μετατρέπεται ως εξής:

$$\mathbf{y}(t_k) = \mathbf{f}(\mathbf{x}(t_{k-1}), \mathbf{y}(t_{k-1})) \quad (4.9)$$

Επομένως κάθε διάνυσμα απόκρισης εξαρτάται μόνο από τη συμπεριφορά του συστήματος την ακριβώς προηγούμενη στιγμή, δηλαδή το διάνυσμα εισόδου και απόκρισης. [12]

Τέλος, σημειώνεται ότι το δίκτυο NARX, επιτρέπει σε σχέση με άλλα δίκτυα καλύτερη προσέγγιση σχέσεων με μακροπρόθεσμη εξάρτηση, δηλαδή το διάνυσμα απόκρισης εξαρτάται από τα διανύσματα εισόδου και εξόδου για αρκετές παρελθοντικές στιγμές.

Η αρχιτεκτονική του δικτύου NARX παρουσιάζει δύο εκδοχές: [6], [11], [13]

- Τη σειριακή – παράλληλη αρχιτεκτονική, η οποία χρησιμοποιείται κατά την εκπαίδευση του δικτύου. Ως είσοδος, παρέχονται γνωστές τιμές της ακολουθίας απόκρισης y για κάθε γνωστή τιμή του διανύσματος εισόδου x και στην έξοδο υπολογίζονται οι προσεγγίσεις του διανύσματος \hat{y} . Στην πραγματικότητα πρόκειται για ένα δίκτυο πρόσθιας τροφοδότησης, και συνεπώς μπορεί να προσεγγιστεί μόνο η ακριβώς επόμενη τιμή της απόκρισης εξόδου \hat{y} .
- Στην αρχιτεκτονική με ανατροφοδότηση, η οποία χρησιμοποιείται κατά τη λειτουργία του δικτύου, όταν δεν μπορεί να υπάρχουν γνωστές τιμές του διανύσματος απόκρισης. Ως είσοδος του δικτύου, παρέχονται οι τιμές του διανύσματος εισόδου x , οι τιμές της εκτίμησης της απόκρισης \hat{y} για την προηγούμενη χρονική στιγμή καθώς και το δίκτυο που έχει εκπαιδευτεί προηγουμένως ενώ υπολογίζονται οι εκτιμήσεις της απόκρισης \hat{y} για χρονικές στιγμές μεγαλύτερες του ενός βήματος. Η διαδικασία αυτή εγκυμονεί το κίνδυνο συσσώρευσης σφάλματος της προσέγγισης \hat{y} σε κάθε βήμα, καθώς στην εκτίμηση που υπολογίζεται στην k επανάληψη της μεθόδου υπεισέρχεται k φορές το σφάλμα της συνάρτησης F της πρόσθιας τροφοδότησης, το οποίο περιέχεται στην εκτίμηση της απόκρισης \hat{y} που ανατροφοδοτείται στο δίκτυο. Συνεπώς, η συνάρτηση F θα πρέπει να υπολογιστεί όσο το δυνατόν καλύτερα, με μικρότερο δηλαδή σφάλμα, το οποίο εξασφαλίζεται με τη διαδικασία εκπαίδευσης του δικτύου.

4.6 Εκπαίδευση Νευρωνικού Δικτύου

Ένα τεχνητό νευρωνικό δίκτυο, όταν υποβάλλεται σε μια διαδικασία εκπαίδευσης ξεκινά από μία κατάσταση κατά την οποία δεν υπάρχει καμία γνώση πάνω στο πρόβλημα μελέτης. Στη συνέχεια όμως, κατά τη διαδικασία εκμάθησης παρουσιάζονται τα διάφορα πρότυπα και η ακολουθία δεδομένων εισόδου – εξόδου με αποτέλεσμα το δίκτυο να εκλαμβάνει τη νέα αυτή πληροφορία ως ‘γνώση’, την οποία και αποθηκεύει στα συναπτικά του βάρη. Με αυτόν τον τρόπο θα μπορέσει στη συνέχεια κατά τη διάρκεια λειτουργίας να προσεγγίσει με όσο το δυνατόν μεγαλύτερη ακρίβεια το μελετώμενο σύστημα.

4.6.1 Μέθοδοι εκπαίδευσης

Τα διάφορα τεχνητά δίκτυα μπορούν να διαχωριστούν με βάση τον τρόπο με τον οποίο επιτελείται η διαδικασία εκμάθησής τους. Αυτές οι μέθοδοι μπορούν να συνοψιστούν ως εξής:

- Επιβλεπόμενη εκμάθηση (supervised training): Κατά τη διάρκεια αυτής είναι απαραίτητο να τροφοδοτούνται στο δίκτυο κάποια ‘πρότυπα εκπαίδευσης’ τα οποία αποτελούν το επιθυμητό αποτέλεσμα για τα συγκεκριμένα δεδομένα. Με βάση αυτά τα γνωστά παραδείγματα καθώς και με τη διαφορά πραγματικής και υπολογιζόμενης εξόδου σε κάθε επανάληψη, το δίκτυο προσαρμόζει τα συναπτικά βάρη με στόχο τη σταδιακή μείωση του σφάλματος της προσέγγισης σε σχέση με την πραγματική τιμή η οποία τροφοδοτείται στο δίκτυο ως δεδομένο εισόδου. Η διαδικασία αυτή επαναλαμβάνεται έως ότου το σφάλμα αυτό μηδενιστεί, ή θεωρηθεί αποδεκτό.

- Μη – Επιβλεπόμενη εκμάθηση (unsupervised training): Στις εφαρμογές μη επιβλεπόμενης εκμάθησης ένα τεχνητό δίκτυο καλείται να προσεγγίσει ένα αποτέλεσμα το οποίο δεν μπορεί είναι γνωστό, αφού δεν τροφοδοτούνται σε αυτό γνωστά παραδείγματα - αποτελέσματα. Έτσι, το δίκτυο προσπαθεί να συλλέξει πληροφορίες μέσα από στατιστικά μεγέθη και πρότυπα των δεδομένων, και με αυτά να οδηγηθεί στην αναπροσαρμογή των συναπτικών βαρών του.
- Ενισχυμένη εκμάθηση (reinforced training): Ούτε σε αυτή τη μέθοδο τροφοδοτείται το δίκτυο με επιθυμητά παραδείγματα – αποτελέσματα, αλλά αντιθέτως χρησιμοποιεί ένα σύνολο μέτρων επάρκειας του εξαγόμενου κάθε φορά αποτελέσματος το οποίο ονομάζεται ενισχυτικό σήμα (reinforcement signal). Το σήμα ‘επιβραβεύει’ τις ορθές συμπεριφορές του δικτύου, ‘τιμωρεί’ τις λανθασμένες και σταδιακά μεταβαίνει προς το επιθυμητό αποτέλεσμα.

Στην παρούσα εργασία χρησιμοποιήθηκε η μέθοδος της επιβλεπόμενης εκμάθησης, και πιο συγκεκριμένα ο αλγόριθμος Bayesian regularization back-propagation μέσω του εργαλείου εκπαίδευσης νευρωνικών δικτύων της Matlab. Στόχος του συγκεκριμένου αλγόριθμου είναι η ελαχιστοποίηση της αντικειμενικής συνάρτησης: [14], [15], [16]

$$F = \alpha E_w + \beta E_d \quad (4.10)$$

όπου E_w είναι το άθροισμα των τετραγώνων των συναπτικών βαρών, E_d είναι το άθροισμα του σφάλματος του δικτύου και οι συντελεστές α, β της αντικειμενικής συνάρτησης. Πιο αναλυτικά:

$$E_w = \sum_{i=1}^n w_i^2 \quad (4.11)$$

$$E_d = \sum_{i=1}^n (\bar{t}_i - t_i)^2 \quad (4.12)$$

Προφανώς, όσο $\alpha \gg \beta$ τόσο σημαντικότερη είναι η σωστή πρόβλεψη των συναπτικών βαρών ενώ συγχρόνως ένα μεγαλύτερο σφάλμα δεν θα επηρεάσει το αποτέλεσμα σημαντικά. Αντίθετα, όσο $\beta \gg \alpha$ παρατηρείται υπερ-προσαρμογή των δεδομένων. Συνεπώς, στόχος είναι η εύρεση του βέλτιστου συνδυασμού των συντελεστών αυτών, ο οποίος επιτυγχάνεται μέσω του θεωρήματος Bayes: [16], [14], [15]

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (4.13)$$

Στη συνέχεια, μετά την εύρεση των βέλτιστων συντελεστών α, β , επικαιροποιείται το διάνυσμα των συναπτικών βαρών. Τέλος, εάν δεν έχει επέλθει σύγκλιση, πραγματοποιείται εκ νέου υπολογισμός των α, β και η όλη διαδικασία επαναλαμβάνεται μέχρι τη σύγκλιση.

Σημαντική επίδραση στην εκπαίδευση του δικτύου έχει το εύρος του πεδίου ορισμού της επιθυμητής συνάρτησης. Όσο μεγαλύτερο είναι αυτό, τόσο περισσότερα δεδομένα εισόδου και εξόδου απαιτούνται από διάφορες περιοχές του πεδίου ορισμού. Η προσέγγιση είναι καλύτερη στις περιοχές από όπου έχουν ληφθεί τα περισσότερα

δεδομένα. Συγχρόνως, όσο πιο περίπλοκη είναι η επιθυμητή συνάρτηση – περιοχές ασυνέχειας, πολλαπλότητα του χώρου λύσεων κ.α. – τόσο πιο δύσκολη είναι η σωστή πρόβλεψη της συμπεριφοράς του συστήματος. Τέλος, η ποιότητα των δεδομένων εκπαίδευσης – πειραματικά σφάλματα, θόρυβος αποτελεσμάτων – επηρεάζει σημαντικά τη διαδικασία εκπαίδευση και την τελική συνάρτηση προσέγγισης [16].

Επισημαίνεται η σημασία του αριθμού νευρώνων των κρυφών επιπέδων L στην ποιότητα του εκπαιδευόμενου δικτύου. Σε κάθε επαναληπτική μέθοδο εκπαίδευσης, υπολογίζονται τα συναπτικά βάρη w , τα οποία μετατρέπουν το διάνυσμα εισόδου σε διάνυσμα εξόδου. Όσο μεγαλύτερος είναι ο αριθμός L , τόσο περισσότερες είναι οι συνάψεις και τα συναπτικά βάρη. Σωστή προσέγγιση του διανύσματος απόκρισης, δεν μπορεί να προκύψει σε δύο περιπτώσεις: όταν ο αριθμός των κρυφών νευρώνων L είναι μικρός όπου ο αριθμός των συναπτικών βαρών δεν επαρκεί και όταν ο αριθμός L είναι υπερβολικά μεγάλος όπου το σύστημα είναι ακριβώς ορισμένο ή υποορισμένο. Στη δεύτερη περίπτωση, το δίκτυο εκπαιδεύεται ώστε να περιγράφει απόλυτα το σύστημα. Παρόλα αυτά οι προβλέψεις που λαμβάνονται ως έξοδος παρουσιάζουν μεγάλο σφάλμα, το οποίο οφείλεται στην υπερπροσαρμογή (overfitting) του δικτύου. Η χρυσή τομή επιτυγχάνεται μεταξύ των δύο ακραίων περιπτώσεων που αναφέρθηκαν παραπάνω, με αποτέλεσμα οι προβλέψεις εξόδου \hat{y} να παρουσιάζουν μικρό σφάλμα [14] [15] [11].

Κατά τη διαδικασία εκπαίδευσης, σημαντικό είναι να μην χρησιμοποιούνται όλα τα δεδομένα – πειραματικά ή υπολογιστικά – ούτως ώστε στο τέλος της διαδικασίας να ελεγχθεί κατά πόσο το δίκτυο έχει εκπαιδευτεί σωστά. Στην παρούσα εργασία, τα δεδομένα που χρησιμοποιούνται έχουν συλλεχθεί με προσομοίωση του συστήματος και όχι με πραγματικό πείραμα και συνεπώς υπάρχει η δυνατότητα να ληφθεί μεγάλο αριθμός αποτελεσμάτων χωρίς τυχόν πειραματικά σφάλματα λόγω αστοχίας εξοπλισμού, ανθρωπίνου λάθους ή άλλων τυχαίων παραγόντων που επηρεάζουν τα αποτελέσματα [11].

Σημειώνεται ότι, επειδή το νευρωνικό δίκτυο εκπαιδεύεται χωρίς καμία γνώση των εξισώσεων που διέπουν το σύστημα, τα αποτελέσματα που παράγει έχουν ισχύ μόνο εντός του πεδίου των δεδομένων εκπαίδευσης. Αστοχούν σε πολύ μεγάλο βαθμό να προβλέψουν τη συμπεριφορά της ακολουθίας απόκρισης εκτός της περιοχής αυτής [11].

Στην παρούσα εργασία θα πραγματοποιηθεί εκπαίδευση του δικτύου NARX σε σειριακή – παράλληλη αρχιτεκτονική ενώ στη συνέχεια, θα λειτουργεί με ανατροφοδότηση. Επομένως, η διαδικασία του ελέγχου θα πραγματοποιηθεί στη λειτουργία με ανατροφοδότηση, η οποία θα τροφοδοτείται με τα δεδομένα εισόδου και θα παράγει τις εκτιμήσεις των αποκρίσεων εξόδου \hat{y} , οι οποίες στη συνέχεια θα συγκρίνονται με τις τιμές από γνωστές αποκρίσεις εξόδου y .

4.7 Εφαρμογή

Το νευρωνικό δίκτυο θα χρησιμοποιηθεί για την εύρεση των συντελεστών που περιγράφουν τη χρονική εξέλιξη της συμπεριφοράς $\alpha(t)$ του μοντέλου μειωμένης τάξης, όπως περιεγράφηκε αναλυτικά στο Κεφάλαιο 3. Η χρονική εξέλιξη της παροχής εισόδου, δρα ως ανεξάρτητη μεταβλητή, ενώ το διάνυσμα εξόδου $y(t)$ ως εξαρτημένη μεταβλητή. Το διάνυσμα εξόδου $y(t)$ προσεγγίζεται μέσω των συντελεστών $\alpha(t)$ σύμφωνα με τη

σχέση (3.39) και συνεπώς το δίκτυο θα εκπαιδευτεί ούτως ώστε να μπορεί να προβλέπει του συντελεστές αυτούς [6].

Έστω μια άγνωστη τροχιά $y_i(t)$, η οποία αποτελείται από m χρονικές στιγμές, για την οποία είναι γνωστά μόνο η αρχική κατάσταση $x(t_0)$, και συνεπώς οι αρχικοί χρονικοί συντελεστές $\alpha(t_0)$, καθώς και το διάνυσμα της παροχής εισόδου $\dot{m}(t)$ σε κάθε χρονική στιγμή. Άρα, όταν επιβάλλεται μια βηματική μεταβολή από μία αρχική κατάσταση $x(t_0)$ για γνωστή παροχή $\dot{m}(t_0)$ σε μία άλλη $\dot{m}(t)$ τότε στόχος είναι η εύρεση της χρονικής εξέλιξης του διανύσματος κατάστασης σε όλο το χώρο του αντιδραστήρα από μία μόνιμη κατάσταση σε μία άλλη.

Από την εφαρμογή του εκπαιδευμένου δικτύου προκύπτουν, λοιπόν, οι συντελεστές $\alpha(t)$, δηλαδή προκύπτει το διάνυσμα:

$$\hat{\mathbf{A}} = [\hat{\alpha}(t_0) \quad \dots \quad \hat{\alpha}(t_{m-1})] \quad (4.14)$$

οι οποίοι σε συνδυασμό με την ορθοκανονική βάση \mathbf{Z} μπορούν να δώσουν την προσέγγιση του πίνακα τροχιών $\hat{\mathbf{Y}}$ σύμφωνα με τη σχέση (3.39).

4.8 Αξιολόγηση Νευρωνικού Δικτύου – Σφάλματα

Στο τελικό στάδιο, το δίκτυο που έχει εκπαιδευτεί θα πρέπει να αξιολογηθεί ως προς την προσέγγιση του πίνακα \mathbf{Y} , δηλαδή του μοντέλου πλήρους τάξης αλλά και ως προς την προσέγγιση του πίνακα $\hat{\mathbf{Y}}$ όπως προκύπτει από το μοντέλο μειωμένης τάξης. Στόχος της όλης ανάλυσης είναι η σωστή εκτίμηση του μοντέλου πλήρους τάξης, με μόνη γνώση τη μεταβλητή εκ χειρισμού, δηλαδή την παροχή εισόδου $\dot{m}(t)$.

Όπως ορίστηκε στο προηγούμενο κεφάλαιο, έτσι και τώρα, ορίζεται το σχετικό σφάλμα της πρόβλεψης του νευρωνικού δικτύου ως προς την εκτίμηση του μοντέλου πλήρους τάξης, όσο και ως προς την εκτίμηση του μοντέλου μειωμένης τάξης, ως εξής:

$$Error_{Full} = \frac{\|\bar{\mathbf{x}}(t) - \mathbf{x}(t)\|}{\|\bar{\mathbf{x}}(t)\|} \quad (4.15)$$

$$Error_{POD} = \frac{\|\hat{\mathbf{x}}(t) - \bar{\mathbf{x}}(t)\|}{\|\hat{\mathbf{x}}(t)\|} \quad (4.16)$$

Όπου $\bar{\mathbf{x}}(t)$ η προσέγγιση που προκύπτει μετά την εφαρμογή του νευρωνικού δικτύου ενώ $\hat{\mathbf{x}}(t)$, η προσέγγιση που προκύπτει με εφαρμογή της ανάλυσης κύριων συνιστωσών.

Αντίστοιχα, θα υπολογιστούν και το μέγιστο σχετικό σφάλμα και το μέσο σχετικό σφάλμα ως προς την εκτίμηση του μοντέλου πλήρους τάξης, όσο και ως προς την εκτίμηση του μοντέλου μειωμένης τάξης, ως εξής:

$$avererror(t) = \left(\frac{1}{m} \sum_{j=1}^m (Error_{POD}(t))^2 \right)^{0.5} \quad (4.17)$$

$$avererror(t) = \left(\frac{1}{m} \sum_{j=1}^m (Error_{Full}(t))^2 \right)^{0.5} \quad (4.18)$$

$$\mathbf{maxerror}(t) = \max(\mathbf{Error}_{POD}(t)) \quad (4.19)$$

$$\mathbf{maxerror}(t) = \max(\mathbf{Error}_{Full}(t)) \quad (4.20)$$

4.9 Συμπεράσματα

Μετά και την ολοκλήρωση του θεωρητικού υποβάθρου της μεθόδου, συμπεραίνεται ότι είναι εφικτός ο σχεδιασμός ενός μοντέλου μειωμένης τάξης το οποίο θα περιγράφει αρκούτως ικανοποιητικά το μελετώμενο σύστημα. Το σύστημα πλήρους τάξης παρουσιάζει τόσο χωρική όσο και χρονική εξάρτηση. Υπενθυμίζεται ότι με τη μεθοδολογία η οποία παρουσιάστηκε στο κεφάλαιο 3 υπολογίζεται η εξάρτηση από το χώρο, και με την εκπαίδευση και την εφαρμογή ενός νευρωνικού δικτύου δημιουργούνται οι συντελεστές οι οποίοι περιγράφουν την εξάρτηση από το χρόνο. Συνεπώς, θα σχεδιαστεί ένα μοντέλο το οποίο θα μπορεί να περιγράψει σε πολύ μεγάλο βαθμό το σύστημα της χημικής απόθεσης από ατμό χρησιμοποιώντας ελάχιστο υπολογιστικό χρόνο και μνήμη, χωρίς καμία γνώση του συστήματος πέρα των υπολογιστικών πειραματικών αποτελεσμάτων.

Χημικές Αντιδράσεις

5.1 Χημικό Σύστημα

Το μοντέλο μειωμένης τάξης το οποίο σχεδιάζεται σύμφωνα με τα δύο προηγούμενα κεφάλαια δεν λαμβάνει υπόψη τις χημικές αντιδράσεις που συντελούνται στον αντιδραστήρα προς το σχηματισμό ενός υμενίου. Σύμφωνα με προηγούμενη έρευνα, [1], [2], όταν το αέριο μίγμα στην είσοδο του αντιδραστήρα είναι πολύ αραιό, ο χώρος λύσεων που προκύπτει παραμένει πρακτικά ίδιος, δηλαδή δεν επηρεάζεται με την προσθήκη του χημικού προτύπου και την ύπαρξη χημικών αντιδράσεων στον αντιδραστήρα. Οπότε, η μελέτη που θα πραγματοποιηθεί με τη χρήση της ανάλυσης κύριων συστατικών όσο και με τα νευρωνικά δίκτυα μπορεί εύκολα να επεκταθεί στη δυναμική μελέτη του συστήματος όταν σε αυτό συμπεριλαμβάνεται ένα χημικό μοντέλο.

Στην παρούσα εργασία θα μελετηθεί ο σχηματισμός υμενίου πυριτίου (Si) από την πρόδρομη ένωση σιλάνιο (SiH_4), όπου φέρον αέριο είναι το N_2 . Θεωρούμε ότι το χημικό σύστημα υπακούει σε ένα απλοποιημένο πρότυπο το οποίο είναι ενσωματωμένο στο Fluent, και αποτελεί απλοποιημένο πρότυπο των 32 αντιδράσεων των Coltrin et al. [17]. Οι χημικές αντιδράσεις οι οποίες θεωρείται ότι λαμβάνουν χώρα στον αντιδραστήρα παρουσιάζονται στον πίνακα 5-1: [2], [5], [17]

Πίνακας 5-1: Απλοποιημένο χημικό πρότυπο Απόθεσης Si από SiH_4 το οποίο θα χρησιμοποιηθεί για τη μελέτη του χημικού συστήματος

G1	$\text{SiH}_{4(g)} \rightarrow \text{SiH}_{2(g)} + \text{H}_{2(g)}$
S1	$\text{SiH}_{4(g)} \rightarrow \text{Si}_{(s)} + 2\text{H}_{2(g)}$
S2	$\text{SiH}_{2(g)} \rightarrow \text{Si}_{(s)} + \text{H}_{2(g)}$

Η αντίδραση G1 πραγματοποιείται στην αέρια φάση, ενώ οι αντιδράσεις S1 και S2 λαμβάνουν χώρα στην επιφάνεια του στερεού δισκίου, πρόκειται δηλαδή για ετερογενείς αντιδράσεις. Οι συγκεκριμένες αντιδράσεις, όσον αφορά την κινητική τους, υπακούουν στον τροποποιημένο νόμο Arrhenius, ο οποίος ορίζεται σύμφωνα με τη σχέση:

$$r = kT^b \exp\left(-\frac{E_a}{RT}\right) \quad (5.1)$$

Όπου r ο ρυθμός της χημικής αντίδρασης σε $\text{kmol}/(\text{m}^2\text{s})$. Οι τιμές για τον προ-εκθετικό παράγοντα k , για το συντελεστή b και για την ενέργεια ενεργοποίησης E_a δίδονται στον παρακάτω πίνακα: [2]

Πίνακας 5-2: Κινητικές παράμετροι των αντιδράσεων του χημικού συστήματος το οποίο παρουσιάστηκε στον Πίνακα 5-1

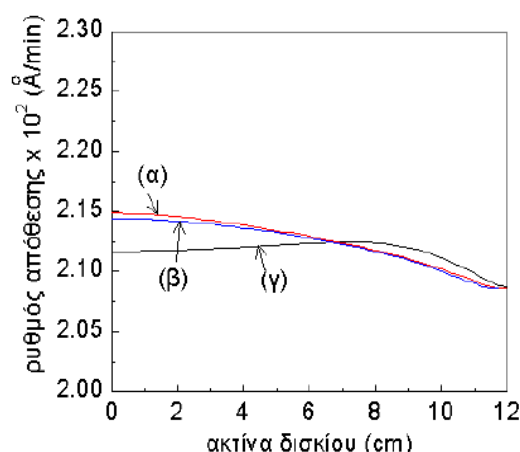
α/α	k	b	$E_a(\text{J/kmol})$
G1	$2.110 \cdot 10^{15}$	0	$2.590 \cdot 10^8$
S1	0.334	0.500	$7.815 \cdot 10^7$
S2	$1.000 \cdot 10^{15}$	0	100.0

Σημειώνεται ότι στον πίνακα 5-1, οι μονάδες του προ-εκθετικού παράγοντα k δίνονται σε $kmol/(m^2s)$ για τις αντιδράσεις G1 και S2, ενώ για την αντίδραση S1 σε $kmol/(m^2s\sqrt{K})$, ούτως ώστε να ορίζεται ο ρυθμός αντίδρασης στην επιφάνεια του υποστρώματος σε $kmol/(m^2s)$.

Οι συνοριακές συνθήκες για τις εξισώσεις που περιγράφουν τα φαινόμενα μεταφοράς παραμένουν οι ίδιες όπως στην ανάλυση που πραγματοποιήθηκε στο κεφάλαιο 2. Για την εξίσωση διατήρησης των χημικών συστατικών, η τιμή του μοριακού κλάσματος μάζας του σιλανίου (SiH_4) στην είσοδο του αντιδραστήρα είναι 10^{-3} ενώ η υπόλοιπη ποσότητα αντιστοιχεί στο αδρανές αέριο άζωτο (N_2). Εξαιτίας της μικρής ποσότητας του SiH_4 , όλα τα υπόλοιπα χημικά είδη τα οποία είτε αντιδρούν είτε παράγονται εμφανίζονται σε ίχνη, συνεπώς το μίγμα μπορεί να θεωρηθεί αραιό. Η ροή των χημικών συστατικών είναι μηδενική στα τοιχώματα του αντιδραστήρα με εξαίρεση την επιφάνεια του δισκίου, όπου πραγματοποιούνται οι χημικές αντιδράσεις. Υπενθυμίζεται ότι στα τοιχώματα του αντιδραστήρα επιβάλλεται σταθερή θερμοκρασία 300 K, ούτως ώστε να μην ενεργοποιούνται οι επιφανειακές αντιδράσεις προς σχηματισμό υμενίου στα τοιχώματα, ενώ η επιφάνεια του δισκίου θερμαίνεται στους 700 K [2], [1].

Για τις ιδιότητες των χημικών συστατικών χρησιμοποιήθηκε η βιβλιοθήκη του Fluent, ενώ για τον υπολογισμό των ιδιοτήτων του μίγματος χρησιμοποιούνται οι μέσοι όροι με βάση τη σύσταση του μίγματος.

Όπως έχει αναλυθεί προηγουμένως, η ομοιομορφία του παραγόμενου υμενίου εξαρτάται από τις συνθήκες κυκλοφορίας που επικρατούν στον αντιδραστήρα. Στην Εικόνα 5-1 παρουσιάζονται σχηματικά οι συγκεκριμένες διαφορές στην ανομοιομορφία του υμενίου στις τρεις μόνιμες καταστάσεις του αντιδραστήρα. Παρατηρείται ότι τα υμένια που σχηματίζονται στον πάνω ευσταθή κλάδο παρουσιάζουν μεγαλύτερη ομοιομορφία σε σχέση με αυτά του κάτω ή του ασταθούς κλάδου [1], [2].



Εικόνα 5-1: Ρυθμοί απόθεσης του υμενίου κατά μήκος του δισκίου απόθεσης για τον (α) κάτω ευσταθή κλάδο (β) ασταθή κλάδο (γ) πάνω ευσταθή κλάδο

Η ανομοιομορφία του υμενίου κατά μήκος του δισκίου απόθεσης ορίζεται ως εξής:
[2]

$$A = \frac{DR_{max} - DR_{min}}{DR_{max}} \% \quad (5.2)$$

Όπου DR_{max} και DR_{min} είναι οι τιμές του μέγιστου και του ελάχιστου του ρυθμού απόθεσης. Ειδικότερα, στον κάτω κλάδο που αντιστοιχεί στη φυσική κυκλοφορία η ανομοιομορφία του υμενίου έχει υπολογιστεί ίση με 2.6% ενώ στην εξαναγκασμένη συναγωγή 1.7% [2], [1].

5.2 Εφαρμογή

Στόχος της συγκεκριμένης ανάλυσης είναι η μελέτη της διαφοράς που παρουσιάζει στο ρυθμό απόθεσης του πυριτίου (Si) κατά μήκος του δισκίου, η ακριβής λύση σε σχέση με την προσέγγιση που προκύπτει με χρήση της ανάλυσης κύριων συνιστωσών και του νευρωνικού δικτύου, όπως αναλύθηκαν στα προηγούμενα κεφάλαια. Ειδικότερα, συγκρίνεται κάθε φορά η πρόβλεψη του ρυθμού απόθεσης στο δισκίο με την ακριβή λύση που προκύπτει από την επίλυση του προβλήματος με τον CFD κώδικα Fluent.

Έχει αποδειχθεί [2] ότι το διάλυμα κατάστασης του αντιδραστήρα, εάν βρίσκεται σε μόνιμη κατάσταση, διατηρείται πρακτικά σταθερό είτε λαμβάνουν χώρα χημικές αντιδράσεις είτε όχι, όταν το μίγμα εισόδου στον αντιδραστήρα είναι αραιό. Συνεπώς, για την παρατήρηση των φαινομένων που λαμβάνουν χώρα στον αντιδραστήρα παρουσία χημικών συστατικών αρκεί κάθε φορά η εισαγωγή ως αρχικής συνθήκης της κατάστασης για την οποία αναζητείται ο ρυθμός απόθεσης, και στη συνέχεια η επίλυση των εξισώσεων διατήρησης μάζας των χημικών συστατικών από τον Fluent, διαδικασία η οποία απαιτεί πολύ χρόνο.

Προκειμένου να βρεθεί η κατανομή του ρυθμού απόθεσης κατά μήκος της ακτίνας του υποστρώματος που αντιστοιχεί στην προσεγγιστική λύση που δίνει το μειωμένης τάξης μοντέλο, αρκεί να εκτελεστούν κάποιες λίγες μόνο επαναλήψεις με αρχική εκτίμηση, την προσεγγιστική λύση. Εκτελώντας περισσότερα χρονικά βήματα, τελικά θα προκύψει η ακριβής λύση που θα έδινε το λεπτομερές μοντέλο του Fluent.

Για την εύρεση του ρυθμού απόθεσης κατά μήκος του δισκίου για την κάθε προσεγγιστική κατάσταση η οποία έχει υπολογιστεί, ο Fluent δεν αφήνεται να φτάσει σε σύγκλιση. Αντίθετα, επιβάλλονται 5 χρονικά βήματα, το καθένα από τα οποία αποτελείται από 200 επαναλήψεις. Με αυτό τον τρόπο, η κατάσταση απομακρύνεται πρακτικά από την προσέγγιση ελάχιστα, και έτσι προκύπτει και η προσέγγιση του ρυθμού απόθεσης κατά μήκος του δισκίου. Επισημαίνεται ότι, σε κάθε χρονικό βήμα η παροχή εισόδου διατηρείται σταθερή, σε εκείνη την τιμή όπου είχε πραγματοποιηθεί η βηματική επιβολή στο σύστημα χωρίς αντιδράσεις.

Αποτελέσματα

6.1 Συλλογή Αποτελεσμάτων

Αρχικά, υπολογίστηκαν οι μόνιμες καταστάσεις για τιμές της παροχής στο εύρος $1.5 \cdot 10^{-5} - 6.5 \cdot 10^{-5}$ kg/s για τους δύο ευσταθείς κλάδους που παρουσιάζει το σύστημα, όπως αναλύθηκε διεξοδικά στο Κεφάλαιο 2 (βλ. Εικόνα 2-3).

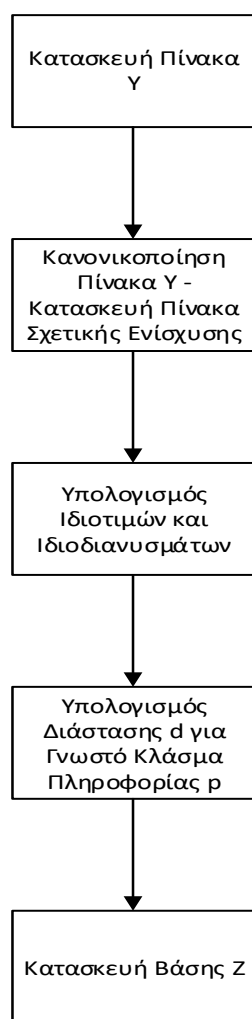
Υπενθυμίζεται ότι ο κάτω κλάδος αντιστοιχεί σε ροή όπου κυριαρχεί η άνωση και ο πάνω κλάδος σε εξαναγκασμένη κυκλοφορία. Στη συνέχεια, πραγματοποιήθηκαν προσομοιώσεις σε μη μόνιμες συνθήκες, και πιο συγκεκριμένα, βηματικές επιβολές από μία μόνιμη κατάσταση σε μια συγκεκριμένη παροχή $\dot{m}(t)$ και έναν κλάδο πολλαπλότητας σε μία άλλη μόνιμη κατάσταση, η οποία μπορεί να βρίσκεται είτε στον ίδιο είτε στον άλλο κλάδο.

Το εύρος των επιβολών αυτών επιλέχθηκε ούτως ώστε να περιγράφεται ικανοποιητικά όλο το εύρος της περιοχής πολλαπλότητας. Σημειώνεται ότι ο χρόνος σύγκλισης αυξάνεται σημαντικά όσο η κατάσταση του συστήματος τείνει στα σημεία αλλαγής ευστάθειας, και άρα οι επιβολές εναλλαγής κλάδων συνοδεύονται από μεγάλους χρόνους σύγκλισης.

Υπενθυμίζεται ότι οι βηματικές επιβολές συλλέχθηκαν με σύζευξη του Fluent με Matlab όπως αναλύθηκε στο Κεφάλαιο 2. Μερικές από αυτές χρησιμοποιήθηκαν για το σχεδιασμό της ορθοκανονικής βάσης σύμφωνα με τη μεθοδολογία που αναπτύχθηκε στο Κεφάλαιο 3, άλλες για τη εκπαίδευση του νευρωνικού δικτύου όπως παρουσιάστηκε στο κεφάλαιο 4 ενώ άλλες για αξιολόγηση των παραπάνω.

Κύριος στόχος της παρούσας μελέτης είναι ο σχεδιασμός ενός μοντέλου που να περιγράφει ικανοποιητικά τους δύο ευσταθείς κλάδους πολλαπλότητας, καθώς και τις εναλλαγές μεταξύ αυτών. Όσον αφορά τις επιβολές εναλλαγής, ξεκινώντας από μια γνωστή κατάσταση σε έναν από τους δύο κλάδους πολλαπλότητας θα πρέπει να δοθεί επιβολή εκτός αυτού, ούτως ώστε να συγκλίνει ο Fluent στον άλλο κλάδο πολλαπλότητας. Αυτό συμβαίνει διότι ο υπολογιστικός κώδικας συγκλίνει στην πιο κοντινή στην αρχική λύση. Αν για παράδειγμα η αρχική κατάσταση αντιστοιχεί σε τιμή της μαζικής παροχής εισόδου $\dot{m}(t) = 4 \cdot 10^{-5}$ kg/s στον πάνω κλάδο και δοθεί βηματική επιβολή $-1 \cdot 10^{-5}$ kg/s και άρα $\dot{m}(t) = 3 \cdot 10^{-5}$ kg/s, ο κώδικας θα συγκλίνει στην πιο κοντινή κατάσταση, δηλαδή στον πάνω κλάδο πολλαπλότητας λύσεων. Συνεπώς, είναι δύσκολο να συλλεχθούν πολλά υπολογιστικά δεδομένα που να περιγράφουν την εναλλαγή μεταξύ των κλάδων πολλαπλότητας και συγχρόνως να βρίσκονται στη μελετώμενη περιοχή.

6.2 Μείωση Τάξης



Εικόνα 6-1:
Διάγραμμα Ροής
Διαδικασίας Μείωσης
Τάξης

Στην Εικόνα 6-1 παρουσιάζεται ο αλγόριθμος της διαδικασίας που ακολουθήθηκε για την κατασκευή της ορθοκανονικής βάσης Z . Πιο αναλυτικά, ο πίνακας Y κατασκευάζεται από την ένωση όλων των τροχιών (trajectories) που θα χρησιμοποιηθούν για την κατασκευή της βάσης σύμφωνα με τη σχέση (3.27). Οι τροχιές αυτές είναι βηματικές επιβολές που συλλέχθηκαν όπως αναφέρθηκε στο Κεφάλαιο 2.

Οι επιβολές αυτές θα επιλεγούν ούτως ώστε να περιέχουν πληροφορίες για το μελετώμενο σύστημα, να περιγράφουν δηλαδή όλη την περιοχή λειτουργίας του μελετώμενου συστήματος. Η περιοχή αυτή κυμαίνεται για τις τιμές της παροχής εισόδου $\dot{m}(t)$ στο εύρος $1.5 \cdot 10^{-5} - 6.0 \cdot 10^{-5}$ kg/s και θα πρέπει να περιλαμβάνει επιβολές που βρίσκονται σε περιοχή όπου κυριαρχεί η εξαναγκασμένη ροή, η φυσική ροή καθώς και οι εναλλαγές μεταξύ αυτών.

Στη συνέχεια, υπολογίζεται η μέση τιμή κάθε γραμμής του πίνακα τροχιών Y , ή οποία αφαιρείται στη συνέχεια από κάθε στοιχείο της, καθώς και η τυπική απόκλιση κάθε γραμμής του πίνακα Y , ή οποία διαιρείται από κάθε στοιχείο σύμφωνα με τη σχέση (3.43), ούτως ώστε να κανονικοποιηθεί ο πίνακας Y για να μην παρουσιαστούν προβλήματα στις προσεγγίσεις των αποκρίσεων που να οφείλονται στις διαφορές τάξης μεγέθους των στοιχείων του πίνακα τροχιών Y .

Έπειτα, κατασκευάζεται ο πίνακας σχετικής ενίσχυσης K' διάστασης $m \ll N$, όπου m ο αριθμός των τροχιών που χρησιμοποιούνται συνολικά, υπολογίζονται οι ιδιοτιμές και τα ιδιοδιανύσματα του, και με βάση το κλάσμα p της πληροφορίας που επιλέγεται κάθε φορά, υπολογίζεται η διάσταση d της ορθοκανονικής βάσης σύμφωνα με τη σχέση (3.23). Τέλος, σύμφωνα με τη σχέση (3.32) υπολογίζονται τα d πρώτα

ιδιοδιανύσματα του πίνακα K , τα οποία αποτελούν την ορθοκανονική βάση διάστασης d, Z_d .

6.3 Σχεδιασμός Ορθοκανονικής Βάσης

Πρώτο βήμα της συγκεκριμένης μελέτης αποτελεί η διερεύνηση του αριθμού των στιγμιότυπων που απαιτούνται για το σχεδιασμό μιας ορθοκανονικής βάσης, δηλαδή της διάστασης m του πίνακα τροχιών Y , ο οποίος να περιγράφει ικανοποιητικά το μελετώμενο σύστημα. Στην πρώτη φάση επιλέχθηκαν τυχαία αρκετές βηματικές επιβολές που περιγράφουν ικανοποιητικά όλα το εύρος πολλαπλότητας (Διερεύνηση 1), ενώ στη συνέχεια επιχειρήθηκε η μείωση της διάστασης m με κύριο στόχο τη διερεύνηση του

απαιτούμενου αριθμού δεδομένων για μια ικανοποιητική απεικόνιση του δυναμικού συστήματος.

6.3.1 Διερεύνηση 1: Τυχαία επιλογή βηματικών επιβολών από όλο το εύρος λύσεων για την κατασκευή της βάσης

Ως πρώτο βήμα της διερεύνησης επελέγησαν συνολικά 40 βηματικές επιβολές, 23 επιβολές οι οποίες πραγματοποιούνται στον πάνω κλάδο, 7 επιβολές στον κάτω κλάδο και 10 επιβολές με εναλλαγή κλάδου για την κατασκευή της ορθοκανονικής βάσης με τη μέθοδο των στιγμιοτύπων που θα αποτυπώνει αρκούντως ικανοποιητικά τη συμπεριφορά στις δύο περιοχές πολλαπλότητας λύσεων. Η διαφορά στο πλήθος των βηματικών επιβολών οφείλεται στη διερεύνηση η οποία πραγματοποιήθηκε από τον Θ.Ν. Παπαπέτρου [6]. Όπως αναφέρθηκε προηγουμένως, οι βηματικές επιβολές πραγματοποιούνται από μια αρχική μόνιμη κατάσταση έως επίτευξης νέας μόνιμης κατάστασης, συνεπώς ως στιγμιότυπο τροχιάς ορίζεται η απόκριση του δυναμικού συστήματος σε κάθε επιβολή.

Στον πίνακα 6-1 καταγράφεται το σχετικό πληροφοριακό περιεχόμενο κάθε βάσης PCA διαφορετικής διάστασης d , όπως προέκυψε μετά την εφαρμογή της μεθοδολογίας η οποία παρουσιάζεται στην Εικόνα 6-1. Σημειώνεται ότι ο δείκτης της βάσης αφορά το πλήθος ιδιοδιανυσμάτων της σχεδιαζόμενης βάσης, δηλαδή τη διάσταση d .

Πίνακας 6-1: Σχετικό Πληροφοριακό Περιεχόμενο για διαφορετικές διαστάσεις της βάσης PCA για τις 40 βηματικές επιβολές

Διάσταση Βάσης PCA, d	Z_1	Z_2	Z_3
Σχετικό Πληροφοριακό Περιεχόμενο	0.8948	0.9629	0.9949

Η βάση Z_1 που περιέχει ένα ιδιοδιάνυσμα του πίνακα K περιέχει το 89,48% της πληροφορίας των 40 επιβολών. Αντίστοιχα, η βάση που περιέχει δύο ιδιοδιανύσματα περιέχει το 96,29% ενώ η βάση που αποτελείται από τρία ιδιοδιανύσματα περιέχει το 99,49% της πληροφορίας των 40 επιβολών.

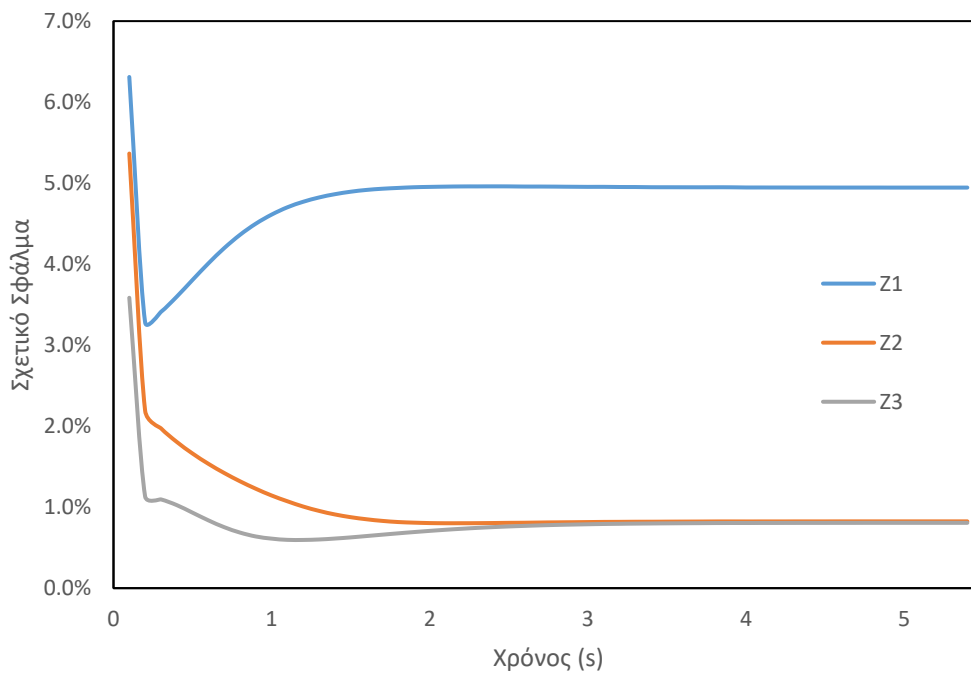
Στη συνέχεια αξιολογήθηκαν οι τρεις παραπάνω βάσεις ως προς το σφάλμα που παρουσιάζουν στην προσέγγιση γνωστών επιβολών σύμφωνα με τη σχέση (3.46). Τα σφάλματα αυτά υπολογίστηκαν για τις επιβολές οι οποίες παρουσιάζονται στον πίνακα 6-2. Η αξιολόγηση αυτή έγινε με τον υπολογισμό των συντελεστών $\alpha(t)$ για κάθε βάση σύμφωνα με τη σχέση (3.40), και στη συνέχεια με υπολογισμό της προσέγγισης $\hat{y}(t)$.

Σημειώνεται ότι στους συντελεστές $\alpha(t)$ της βάσης Z_3 περιέχονται οι συντελεστές των άλλων δύο μικρότερων σε διάσταση βάσεων. Με βάση την παραπάνω ανάλυση κατασκευάστηκαν τα παρακάτω διαγράμματα για τις επιβολές που παρουσιάζονται στον πίνακα 6-2. Οι επιβολές αυτές, θα χρησιμοποιούνται καθόλη την αξιολόγηση των αποτελεσμάτων για να μπορούν να συγκριθούν ευκολότερα τα αποτελέσματα. Παρόλα αυτά δεν πρόκειται για ολοκληρωμένη προσέγγιση διότι για τον υπολογισμό των συντελεστών $\alpha(t)$ απαιτείται η γνώση του κανονικοποιημένου πίνακα τροχιών Y_N .

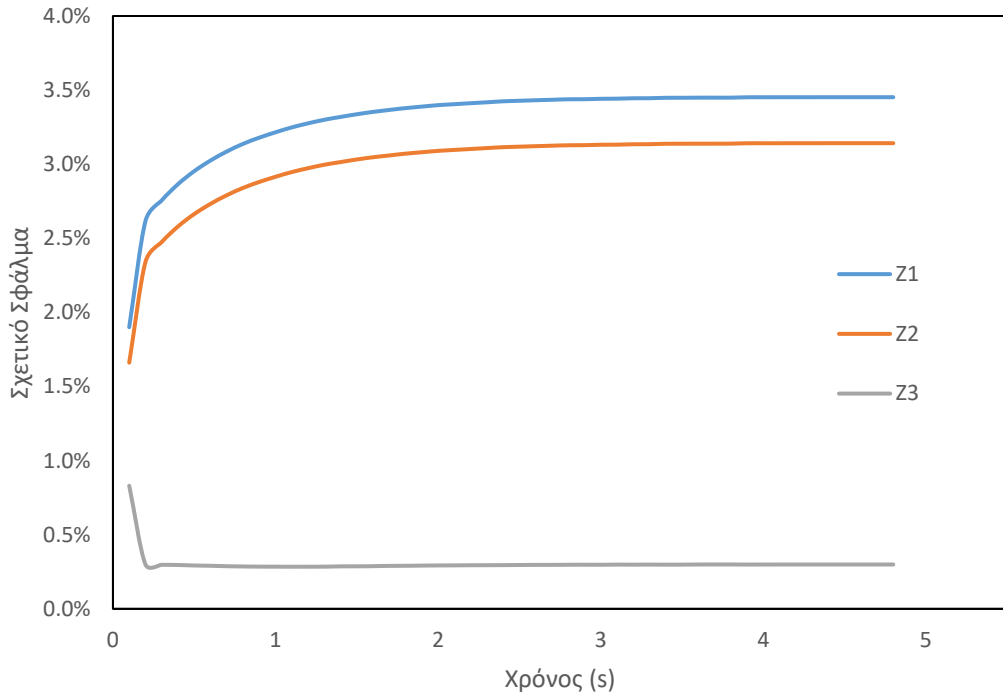
Πίνακας 6-2: Επιβολές που χρησιμοποιούνται για την αξιολόγηση των αποτελεσμάτων

α/α	Αρχική Κατάσταση \dot{m}_0 (10^{-5} kg/s)	Κλάδος	Τελική Κατάσταση \dot{m}_0 (10^{-5} kg/s)	Κλάδος
1	4.0	Πάνω	6.0	Πάνω
3	2.0	Κάτω	1.5	Κάτω
4	4.5	Πάνω	2.0	Κάτω

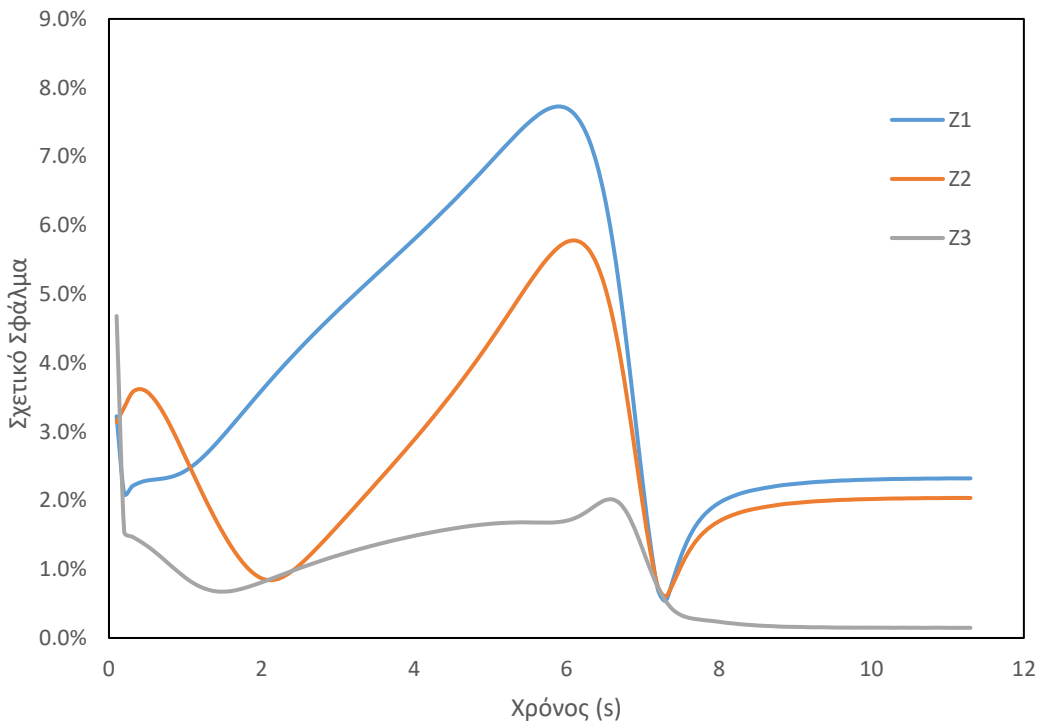
Τα διαγράμματα 6-2 έως 6-4 παρουσιάζουν το σχετικό σφάλμα της κάθε επιβολής για τις 3 διαφορετικές βάσεις, συναρτήσει του χρόνου.



Εικόνα 6-2: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=6 \cdot 10^{-5}$ kg/s



Εικόνα 6-3:Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=2\cdot 10^{-5}$ kg/s σε $\dot{m}(t)=1.5\cdot 10^{-5}$ kg/s



Εικόνα 6-4:Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4.5\cdot 10^{-5}$ kg/s σε $\dot{m}(t)=2.0\cdot 10^{-5}$ kg/s

Στα παραπάνω διαγράμματα παρατηρείται ότι το σχετικό σφάλμα είναι πολύ μικρό για τις μελετώμενες επιβολές ιδιαίτερα για μεγαλύτερους χρόνους όπου η επιβολή πλησιάζει στη νέα κατάσταση ισορροπίας. Το σχετικό σφάλμα είναι σχεδόν σταθερό στις

δύο εξεταζόμενες επιβολές όπου η βηματική επιβολή δεν πραγματοποιεί εναλλαγή κλάδων πολλαπλότητας. Αντίθετα, το σχετικό σφάλμα της βηματικής επιβολής εναλλαγής παρουσιάζει μέγιστο περίπου στη μέση της χρονικής διάρκειας της. Αυτό εικάζεται ότι οφείλεται στο γεγονός ότι σε αυτό το σημείο πραγματοποιείται η εναλλαγή του κλάδου πολλαπλότητας. Παρόλα αυτά επειδή το σχετικό σφάλμα μειώνεται όσο η επιβολή πλησιάζει στη σύγκλιση υποδεικνύει ότι η εναλλαγή κλάδου έχει πραγματοποιηθεί. Το μικρότερο σφάλμα παρουσιάζει η βάση Z_3 , για τις τρεις επιβολές, που περιέχει και το μεγαλύτερο σχετικό πληροφοριακό περιεχόμενο, όπως καταγράφηκε στον πίνακα 6-1.

Στους πίνακες 6-3 και 6-4 καταγράφονται το μέσο και το μέγιστο σχετικό σφάλμα, αντίστοιχα για τις τρεις εξεταζόμενες βηματικές επιβολές.

Πίνακας 6-3: Μέσο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

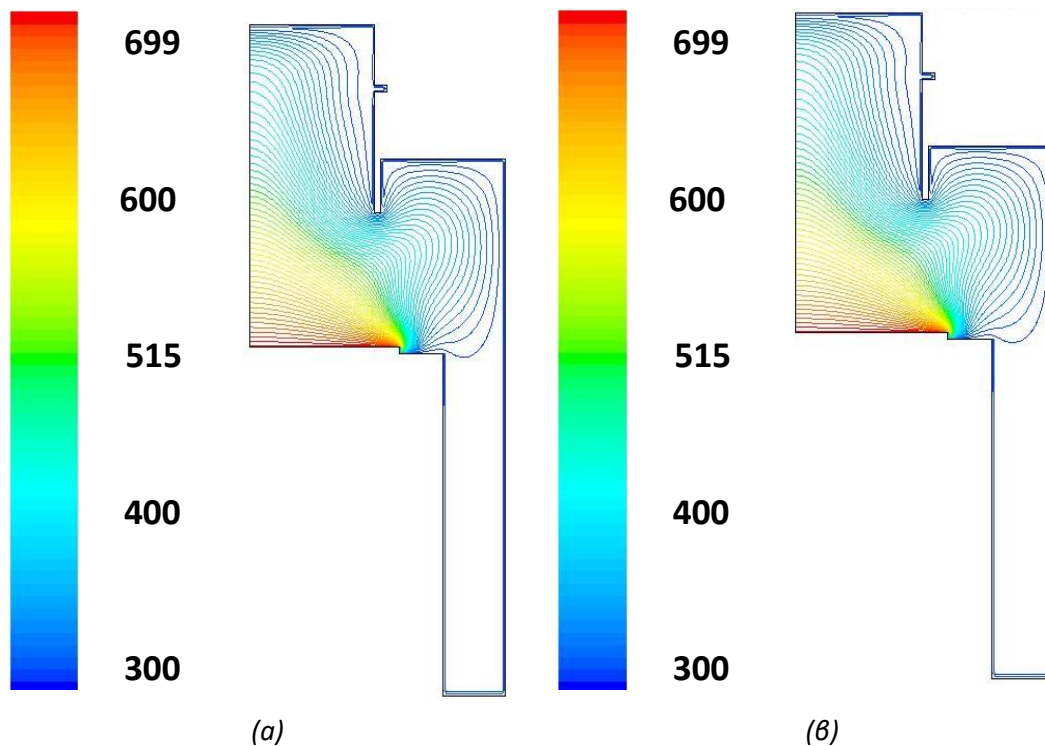
α/α	Z_1	Z_2	Z_3
1	4.89%	0.87%	0.75%
2	3.30%	3.00%	0.31%
3	2.19%	1.92%	0.17%

Πίνακας 6-4: Μέγιστο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

α/α	Z_1	Z_2	Z_3
1	4.96%	1.42%	0.81%
2	3.45%	3.14%	0.83%
3	2.32%	2.04%	0.33%

Για να συγκριθεί η προσεγγιστική λύση με αυτή που προκύπτει από το λεπτομερές μοντέλο του Fluent, ακολουθείται μια διαδικασία «ανόρθωσης», κατά την οποία επιλέγεται ένα στιγμιότυπο (διάγραμμα) του πίνακα \hat{Y} , και αυτό εισάγεται στον Fluent. Με αυτόν τρόπο παρουσιάζονται οι κατανομές των πεδίων ροής στο χωρίο του αντιδραστήρα που έχουν υπολογιστεί από το προσεγγιστικό μοντέλο.

Ενδεικτικά, η προσέγγιση που προέκυψε με τη βάση Z_3 για την επιβολή εναλλαγής, η οποία είναι και πιο δύσκολα προσεγγίσιμη, εισήχθη στον Fluent με στόχο τη σύγκρισή της με την πραγματική λύση. Το προφίλ της θερμοκρασίας που προκύπτει για την πραγματική και την προσεγγιστική λύση είναι σχεδόν ίδια, γεγονός που επιβεβαιώνει το μικρό σχετικό σφάλμα μεταξύ των δύο. Στην παρακάτω εικόνα παρουσιάζονται τα συγκεκριμένα αποτελέσματα.



Εικόνα 6-5: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει από την προσέγγιση της βάσης Z_3 της επιβολής εναλλαγής

Παρόλο που τα αποτελέσματα της συγκεκριμένης βάσης κρίνονται ικανοποιητικά, θα πραγματοποιηθεί διερεύνηση σχετικά με το πλήθος των βηματικών επιβολών που απαιτούνται για την περιγραφή του συστήματος. Άλλωστε, βασικός σκοπός αυτής της διπλωματικής εργασίας είναι η δημιουργία ενός μοντέλου μειωμένης τάξης το οποίο θα προσομοιώνει ικανοποιητικά τους δύο κλάδους πολλαπλότητας λύσεων καθώς και τις εναλλαγές μεταξύ αυτών, δηλαδή θα πρέπει η βάση να περιέχει επιβολές οι οποίες πραγματοποιήθηκαν στο πάνω κλάδο πολλαπλότητας λύσεων, στον κάτω όσο και επιβολές από τον πάνω κλάδο στον κάτω κλάδο και αντίστροφα. Η βάση που σχεδιάστηκε παραπάνω διαθέτει τα παραπάνω χαρακτηριστικά.

Εύλογο όμως είναι το ερώτημα κατά πόσο απαιτείται τόσο μεγάλος αριθμός στιγμιοτύπων για το σχεδιασμό ενός μειωμένης τάξεως μοντέλου. Για την επίτευξη αυτού του στόχου, θεωρήθηκε ως δεδομένο η διερεύνηση του Θ.Ν. Παπαπέτρου για τις επιβολές που χρησιμοποιούνται για το σχεδιασμό της βάσης στον κάτω κλάδο (4 επιβολές). Η συγκεκριμένη διερεύνηση απαρτίζεται από δύο σκέλη. Καταρχάς, θα δημιουργηθεί μια ορθοκανονική βάση που να περιέχει 4 επιβολές από τον κάτω κλάδο και 5 επιβολές από τον πάνω κλάδο. Στη συνέχεια η διερεύνηση αυτή θα επεκταθεί, και θα προστεθούν 4 ακόμη επιβολές από τον πάνω κλάδο στον κάτω ή αντίστροφα.

6.3.2 Διερεύνηση 2: Επιλογή βηματικών επιβολών από τους 2 διαφορετικούς κλάδους πολλαπλότητας για την κατασκευή της Βάσης

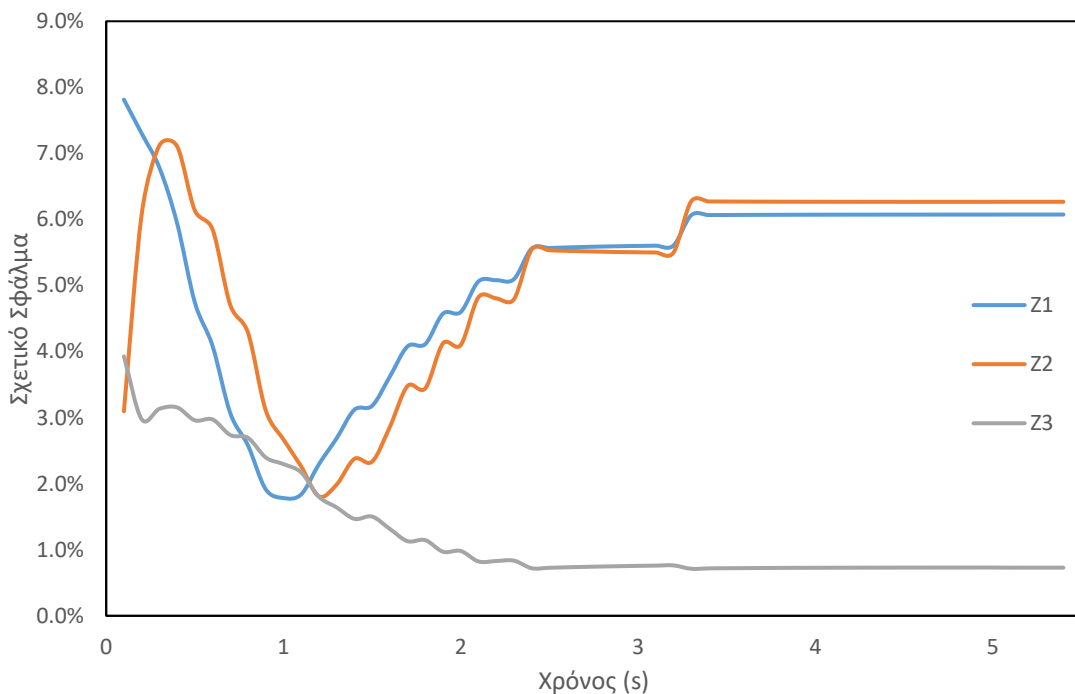
Όπως αναλύθηκε παραπάνω, θα σχεδιαστεί ορθοκανονική βάση που θα περιέχει στιγμιότυπα από τον κάτω και τον πάνω κλάδο. Στη συνέχεια, θα αξιολογηθεί η συγκεκριμένη βάση και θα συγκριθεί με την προηγούμενη, καθώς και κατά πόσο μπορεί να προσομοιώσει εναλλαγές μεταξύ αυτών. Οι επιβολές που χρησιμοποιήθηκαν αποτυπώνουν ικανοποιητικά όλο το εύρος του πάνω κλάδου και του κάτω κλάδου, όπως αυτό διερευνήθηκε από τον Θ.Ν. Παπαπέτρου. Το σχετικό πληροφοριακό περιεχόμενο της συγκεκριμένης βάσης παρουσιάζεται στον πίνακα 6-5.

Πίνακας 6-5: Σχετικό Πληροφοριακό Περιεχόμενο για διαφορετικές διαστάσεις της βάσης PCA για τις 9 βηματικές επιβολές

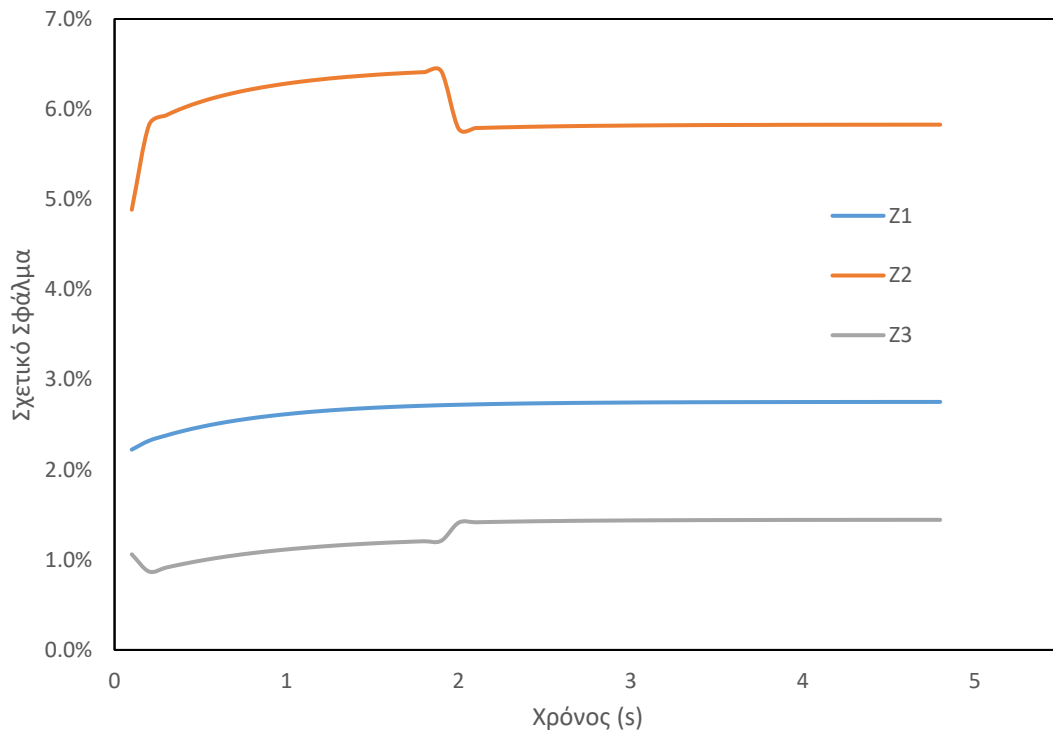
Διάσταση Βάσης PCA, d	Z_1	Z_2	Z_3
Σχετικό Πληροφοριακό Περιεχόμενο	0.7309	0.9367	0.9819

Η αξιολόγηση της βάσης θα πραγματοποιηθεί, όπως στο προηγούμενο βήμα, για τις επιβολές οι οποίες παρουσιάστηκαν στον πίνακα 6-2, και στη συνέχεια θα ακολουθηθεί η ίδια διαδικασία.

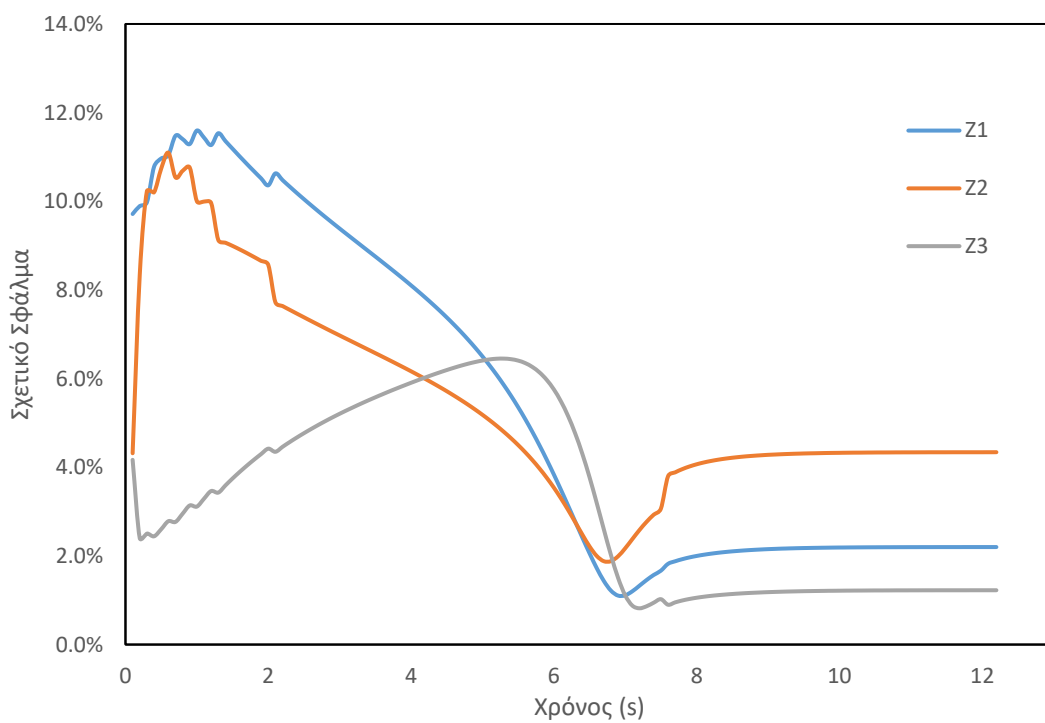
Τα διαγράμματα 6-6 έως 6-8 παρουσιάζουν το σχετικό σφάλμα της κάθε επιβολής για τις 3 διαφορετικές βάσεις, συναρτήσει του χρόνου.



Εικόνα 6-6: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4.0 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=6.0 \cdot 10^{-5}$ kg/s



Εικόνα 6-7: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=2.0 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=1.5 \cdot 10^{-5}$ kg/s



Εικόνα 6-8: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4.5 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=2.0 \cdot 10^{-5}$ kg/s

Στους πίνακες 6-6 και 6-7 καταγράφονται το μέσο και το μέγιστο σχετικό σφάλμα, αντίστοιχα για τις τρεις εξεταζόμενες βηματικές επιβολές.

Πίνακας 6-6: Μέσο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

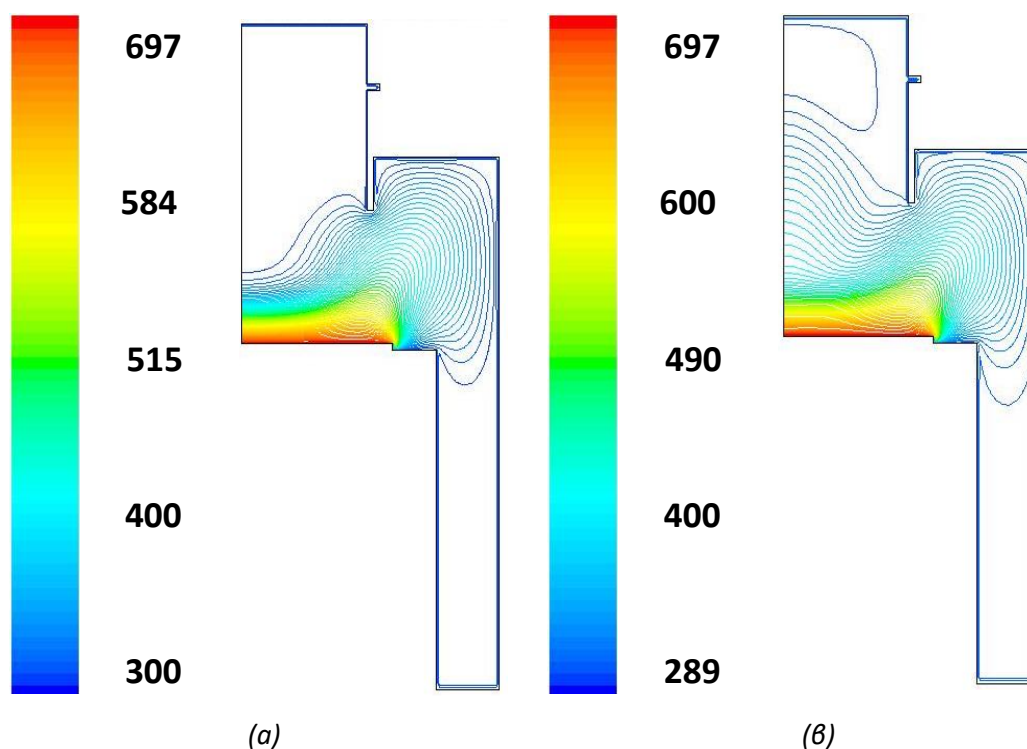
α/α	Z_1	Z_2	Z_3
1	5.05%	5.11%	1.03%
2	2.67%	5.96%	1.30%
3	2.13%	4.24%	1.17%

Πίνακας 6-7: Μέγιστο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

α/α	Z_1	Z_2	Z_3
1	6.07%	6.27%	2.74%
2	2.75%	6.42%	1.44%
3	2.20%	4.34%	1.23%

Το σχετικό σφάλμα παρουσιάζεται αυξημένο σε σχέση με τον προηγούμενο σχεδιασμό. Αυτό εικάζεται ότι οφείλεται στο μικρό αριθμό βηματικών επιβολών που χρησιμοποιούνται στην κατασκευή της βάσης. Το μικρότερο σφάλμα παρουσιάζει η βάση Z_3 , για τις τρεις επιβολές, που περιέχει το μεγαλύτερο σχετικό πληροφοριακό περιεχόμενο, ενώ το μεγαλύτερο η βάση Z_2 . Το συγκεκριμένο φαινόμενο πιθανόν οφείλεται σε θόρυβο που μπορεί να περιέχει η βάση αυτή. Συγχρόνως, παρουσιάζεται δυσκολία στην προσέγγιση του πάνω κλάδου πολλαπλότητας.

Ενδεικτικά, η προσέγγιση που προέκυψε με τη βάση Z_2 για την επιβολή στον πάνω κλάδο πολλαπλότητας, η οποία παρουσίασε το μεγαλύτερο σχετικό σφάλμα, εισήχθη στον Fluent με στόχο τη σύγκρισή της με την πραγματική λύση. Το προφίλ της θερμοκρασίας στην πραγματική λύση και την προσεγγιστική είναι σχεδόν ίδια, γεγονός που επιβεβαιώνει τα παραπάνω αποτελέσματα. Στην Εικόνα 6-9 παρουσιάζονται συγκριτικά οι δύο κατανομές.



Εικόνα 6-9: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει από την προσέγγιση της βάσης Z_2

Οι κατανομές της θερμοκρασίας παρουσιάζουν αρκετές διαφορές τόσο ως προς την κατανομή της στον αντιδραστήρα, αλλά και ως προς το θερμοκρασιακό εύρος. Η βάση η οποία αποτελείται από 3 ιδιοδιανύσματα παρουσιάζει βελτιωμένα αποτελέσματα, παρ' όλα αυτά η διερεύνηση θα συνεχιστεί με στόχο την εύρεση καλύτερης προσέγγισης, δηλαδή να απεικονίζεται ακόμα καλύτερα η μελετώμενη περιοχή.

6.3.3 Διερεύνηση 3: Εμπλουτισμός των Δεδομένων με Επιβολές Εναλλαγής μεταξύ των Κλάδων Πολλαπλότητας

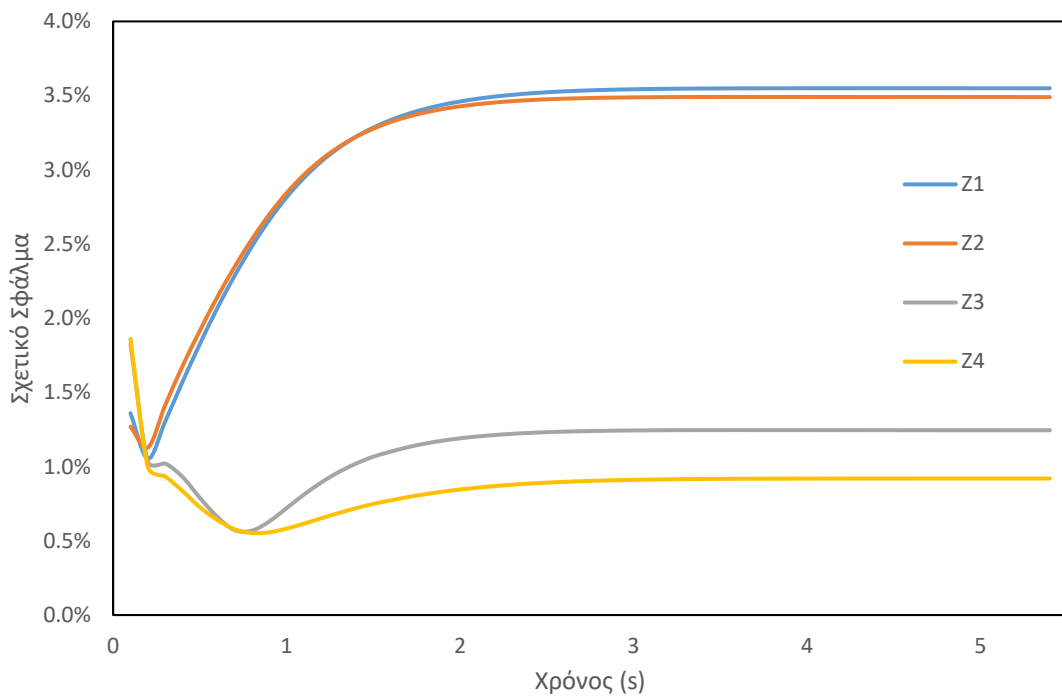
Όπως αναλύθηκε παραπάνω, θα σχεδιαστεί ορθοκανονική βάση που θα περιέχει στιγμιότυπα από τον κάτω, τον πάνω κλάδο πολλαπλότητας καθώς και ισάριθμες επιβολές εναλλαγής. Στη συνέχεια, θα αξιολογηθεί η συγκεκριμένη βάση και θα συγκριθεί με τις δύο προηγούμενες, καθώς και κατά πόσο μπορεί να προσομοιώσει εναλλαγές μεταξύ αυτών. Οι επιβολές που χρησιμοποιούνται είναι αυτές της προηγούμενης βάσης, με προσθήκη τεσσάρων επιβολών εναλλαγής (από τον πάνω στον κάτω κλάδο πολλαπλότητας και αντίστροφα). Επίσης, έγινε προσθήκη ενός επιπλέον ιδιο-διανύσματος και έτσι δημιουργήθηκε επιπλέον η βάση Z_4 . Το σχετικό πληροφοριακό περιεχόμενο των βάσεων παρουσιάζεται στον πίνακα 6-8.

Πίνακας 6-8: Σχετικό Πληροφοριακό Περιεχόμενο για διαφορετικές διαστάσεις της βάσης PCA για τις 13 βηματικές επιβολές

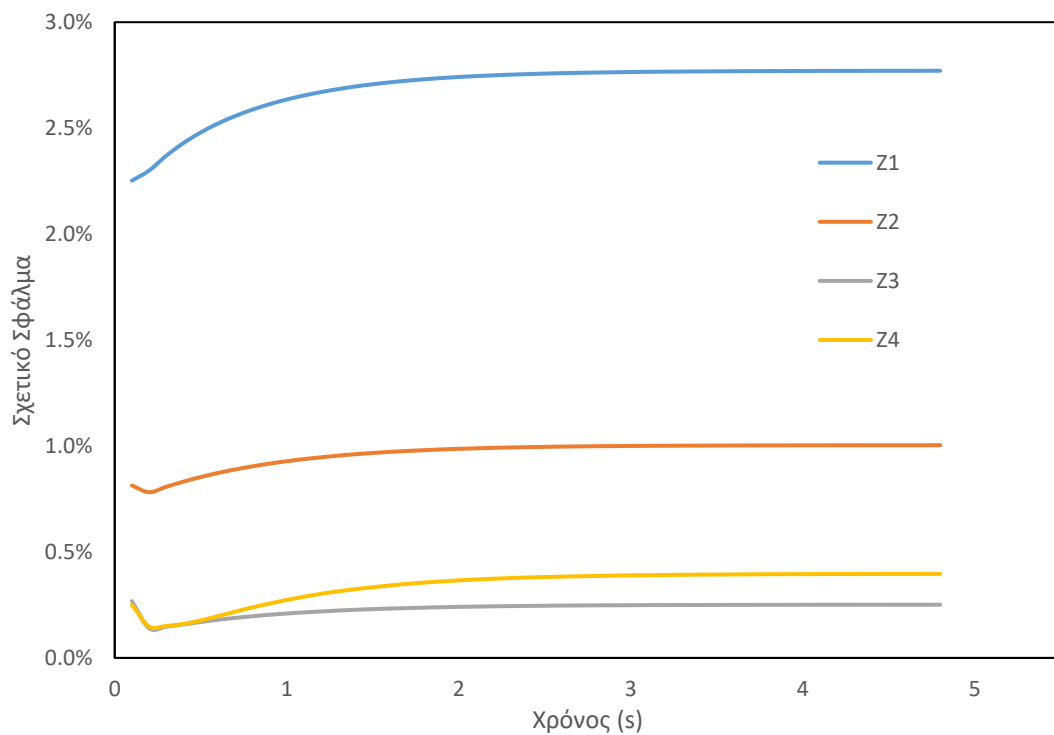
Διάσταση Βάσης PCA, d	Z_1	Z_2	Z_3	Z_4
Σχετικό Πληροφοριακό Περιεχόμενο	0.7695	0.9301	0.9842	0.9977

Η αξιολόγηση της βάσης θα πραγματοποιηθεί, όπως στις δύο προηγούμενες διερευνήσεις, για τις επιβολές οι οποίες παρουσιάζονται στον πίνακα 6-2, και θα ακολουθηθεί η ίδια διαδικασία.

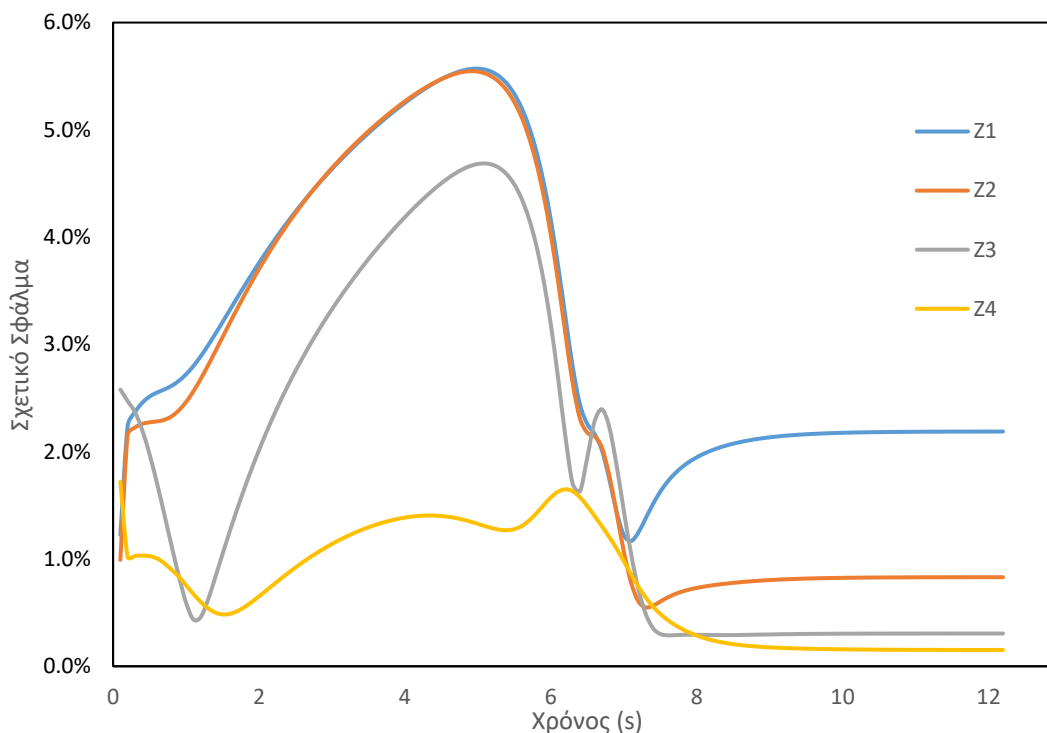
Τα διαγράμματα 6-10 έως 6-12 παρουσιάζουν το σχετικό σφάλμα της κάθε επιβολής για τις 4 διαφορετικές βάσεις, συναρτήσει του χρόνου.



Εικόνα 6-10: Σχετικό σφάλμα κάθε βάσης για την επιβολή από από $\dot{m}(t)=4\cdot 10^{-5}$ kg/s σε $\dot{m}(t)=6\cdot 10^{-5}$ kg/s



Εικόνα 6-11: Σχετικό σφάλμα κάθε βάσης για την επιβολή από από $\dot{m}(t)=2\cdot 10^{-5}$ kg/s σε $\dot{m}(t)=1.5\cdot 10^{-5}$ kg/s



Εικόνα 6-12: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4.5 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=2.0 \cdot 10^{-5}$ kg/s

Στους πίνακες 6-9 και 6-10 καταγράφονται το μέσο και το μέγιστο σχετικό σφάλμα, αντίστοιχα για τις τρεις εξεταζόμενες βηματικές επιβολές.

Πίνακας 6-9: Μέσο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	3.40%	3.36%	1.15%	0.85%
2	2.69%	0.96%	0.23%	0.34%
3	2.11%	0.80%	0.30%	0.20%

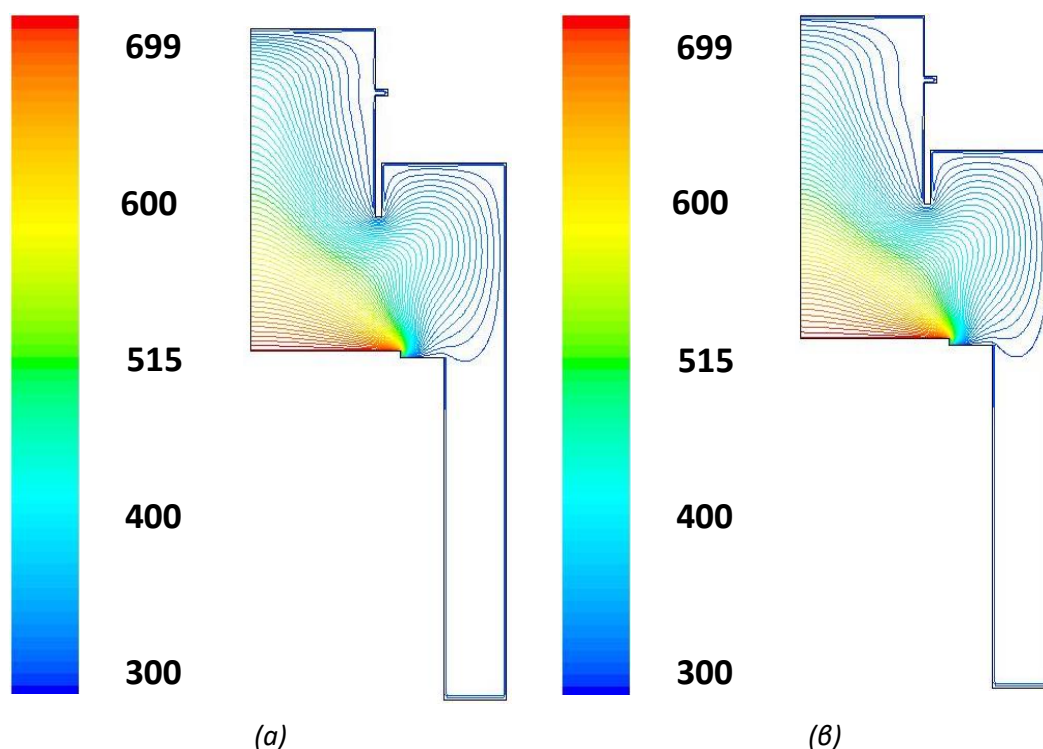
Πίνακας 6-10: Μέγιστο Σχετικό Σφάλμα για τη σχεδιαζόμενη βάση

α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	3.55%	3.49%	1.25%	0.92%
2	2.77%	1.00%	0.27%	0.40%
3	2.19%	0.83%	0.31%	0.49%

Για τις βάσεις Z₃ και Z₄ το μέγιστο σφάλμα δεν ξεπερνάει το 0.31% για την πρώτη περίπτωση και το 0.5% για τη δεύτερη, ενώ το μέσο σφάλμα κυμαίνεται στο 0.3% και 0.2%, αντίστοιχα. Από τα παραπάνω διαγράμματα παρατηρείται ότι το σχετικό σφάλμα είναι πολύ μικρό για τις μελετώμενες επιβολές ιδιαίτερα για μεγαλύτερους χρόνους όπου η επιβολή πλησιάζει σε κατάσταση ισορροπίας. Το σχετικό σφάλμα των βάσεων Z₃ και Z₄, είναι παρόμοιο για τις τρεις επιβολές παρόλα αυτά, θα χρησιμοποιείται η βάση Z₃, η οποία περιέχει μικρότερη πληροφορία και συνεπώς απαιτεί μικρότερο υπολογιστικό χρόνο

παράγοντας ίδιας ή και καλύτερης ποιότητας αποτελέσματα. Το γεγονός αυτό, αποδεικνύει την καταλληλότητα της συγκεκριμένης βάσης για τη χρήση της στην προσέγγιση του σχεδιαζόμενου μοντέλου.

Ενδεικτικά, η προσέγγιση που προέκυψε με τη βάση Z_3 για την επιβολή εναλλαγής, η οποία είναι και πιο δύσκολα προσεγγίσιμη, εισήχθη στον Fluent με στόχο τη σύγκρισή της με την πραγματική λύση. Το προφίλ της θερμοκρασίας στην πραγματική λύση και την προσεγγιστική είναι σχεδόν ίδια, γεγονός που επιβεβαιώνει τα παραπάνω αποτελέσματα. Στην παρακάτω εικόνα παρουσιάζονται τα συγκεκριμένα αποτελέσματα.



Εικόνα 6-13: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει για την επιβολή εναλλαγής από την προσέγγιση της βάσης Z_3

Τα αποτελέσματα της συγκεκριμένης διερεύνησης θεωρούνται ικανοποιητικά. Ειδικότερα, το σχετικό σφάλμα παρουσιάζεται μειωμένο σε σχέση με την προηγούμενη διερεύνηση για κάθε βάση ίδιας διάστασης.

Συμπέρασμα της παραπάνω διαδικασίας είναι ότι το σχεδιαζόμενο μοντέλο μειωμένης τάξης μπορεί να περιγράψει ικανοποιητικά την κατάσταση του αντιδραστήρα. Πιο συγκεκριμένα, η βάση η οποία κατασκευάστηκε με χρήση 13 βηματικών επιβολών κρίθηκε κατάλληλη και θα χρησιμοποιηθεί στην ανάλυση που θα ακολουθήσει για τον ολοκληρωμένο σχεδιασμό ενός μειωμένης τάξης μοντέλου, το οποίο θα προσομοιώνει τη διεργασία της χημικής απόθεσης από ατμό.

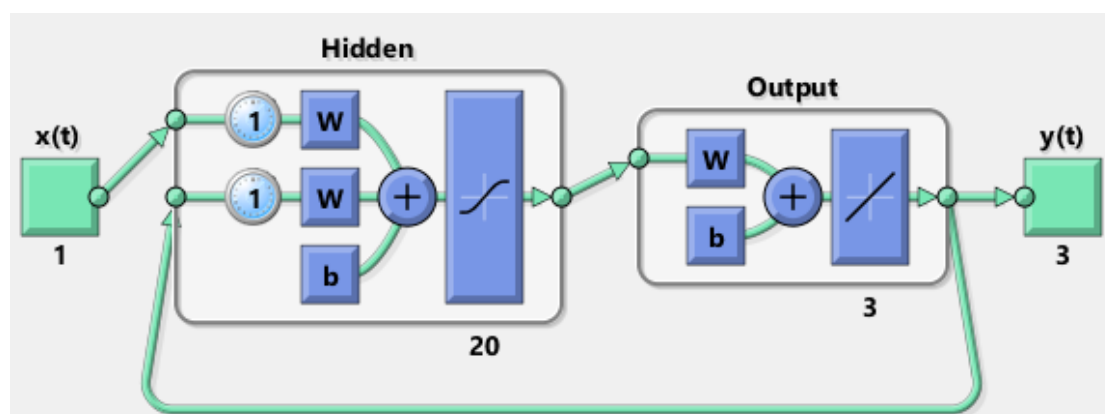
6.4 Σχεδιασμός Νευρωνικού Δικτύου

6.4.1 Εκπαίδευση δικτύου

Στην προηγούμενη ενότητα, επιλέχθηκε βάση η οποία σχεδιάστηκε με χρήση 13 βηματικών επιβολών οι οποίες περιγράφουν με ακρίβεια όλο το εύρος πολλαπλότητας

λύσεων. Για την εκπαίδευση του δικτύου χρησιμοποιήθηκε το toolbox της Matlab, όπως αναλύθηκε διεξοδικά στο κεφάλαιο 4.

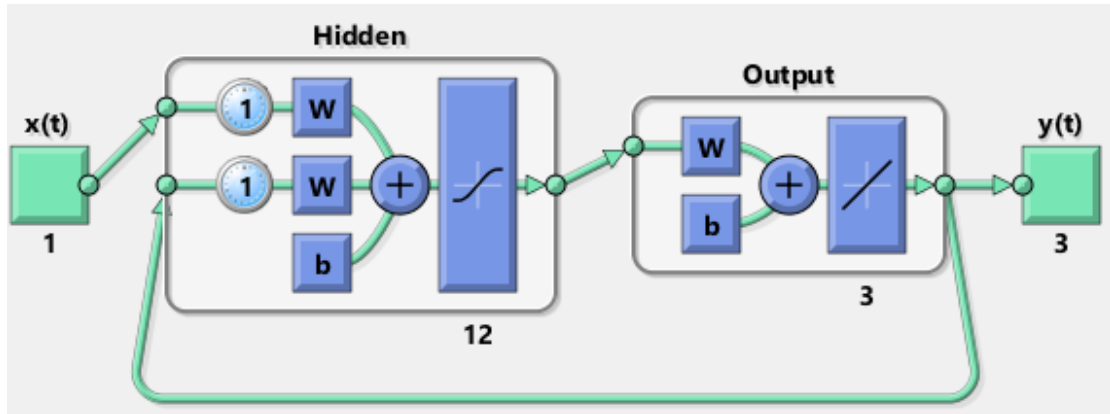
Στο αρχικό στάδιο της μελέτης πραγματοποιήθηκε διερεύνηση για τον απαραίτητο αριθμό επιβολών εκπαίδευσης. Σε πρώτη φάση, διερευνήθηκε η ισάριθμη κατανομή των τριών ειδών επιβολών εκπαίδευσης, δηλαδή 6 από κάθε ξεχωριστό κλάδο και 6 επιβολές εναλλαγής. Εκπαιδεύτηκε δηλαδή ένα δίκτυο με συνολικά 18 βηματικές επιβολές με αποκλειστική χρήση της βάσης Z_3 η οποία δημιουργήθηκε προηγουμένως. Το συγκεκριμένο δίκτυο αποτελείται από $L_3=20$ κρυφούς νευρώνες. Στην Εικόνα 6-14 παρουσιάζεται ενδεικτικά το κλειστό νευρωνικό δίκτυο που εκπαιδεύτηκε για τη βάση Z_3 . Υπενθυμίζεται ότι το νευρωνικό δίκτυο εκπαιδεύεται σε παράλληλη διάταξη ενώ στη συνέχεια λειτουργεί με ανατροφοδότηση.



Εικόνα 6-14: Κλειστό Νευρωνικό Δίκτυο το οποίο εκπαιδεύτηκε με χρήση 18 βηματικών επιβολών σε συνδυασμό με τη βάση Z_3

Με τη χρήση του συγκεκριμένου δικτύου παρουσιάστηκαν κάποιες προβληματικές προσεγγίσεις κυρίως στις επιβολές εναλλαγής. Εξαιτίας αυτού αποφασίστηκε να εμπλουτιστούν οι επιβολές εκπαίδευσης με επιπλέον υπολογιστικά δεδομένα. Το δίκτυο, εκπαιδεύτηκε ξανά με χρήση 24 βηματικών επιβολών, οι οποίες περιγράφουν όλο το εύρος πολλαπλότητας λύσεων. Από αυτές, οι 9 πραγματοποιούνται στον κάτω κλάδο πολλαπλότητας, οι 6 στον πάνω κλάδο ενώ οι υπόλοιπες 8 πραγματοποιούν επιβολές εναλλαγής μεταξύ των κλάδων πολλαπλότητας. Εξ' αιτίας του διερευνητικού χαρακτήρα της παρούσας εργασίας, εκπαιδεύτηκαν συνολικά, τέσσερα νευρωνικά δίκτυα για τις βάσεις Z_1 , Z_2 , Z_3 και Z_4 διαδοχικά, οι οποίες έχουν σχεδιαστεί προηγουμένως.

Τέλος, τα νευρωνικά δίκτυα εκπαιδεύτηκαν με $L_1=10$, $L_2=3$, $L_3=12$ και $L_4=14$ κρυφούς νευρώνες αντίστοιχα για τις βάσεις Z_1 , Z_2 , Z_3 και Z_4 . Η επιλογή του πλήθους των κρυφών νευρώνων έγινε μετά από διάφορες δοκιμές και με κύριο γνώμονα την ποιότητα των παραγόμενων προσεγγίσεων. Στην Εικόνα 6-15 παρουσιάζεται ενδεικτικά το κλειστό νευρωνικό δίκτυο που εκπαιδεύτηκε για τη βάση Z_3 .



Εικόνα 6-15: Κλειστό Νευρωνικό Δίκτυο το οποίο εκπαιδεύτηκε με χρήση 24 βηματικών επιβολών σε συνδυασμό με τη βάση Z_3

6.4.2 Αξιολόγηση Νευρωνικών Δικτύων

Τα εκπαιδευόμενα Νευρωνικά Δίκτυα θα αξιολογηθούν ως προς το σχετικό σφάλμα που παρουσιάζει η προσέγγιση της κάθε επιβολής σε σχέση με την πραγματική κατάσταση, τόσο όσο προς τις επιβολές εκπαίδευσης, όσο και ως προς τις επιβολές αξιολόγησης οι οποίες παρουσιάστηκαν στον πίνακα 6-2.

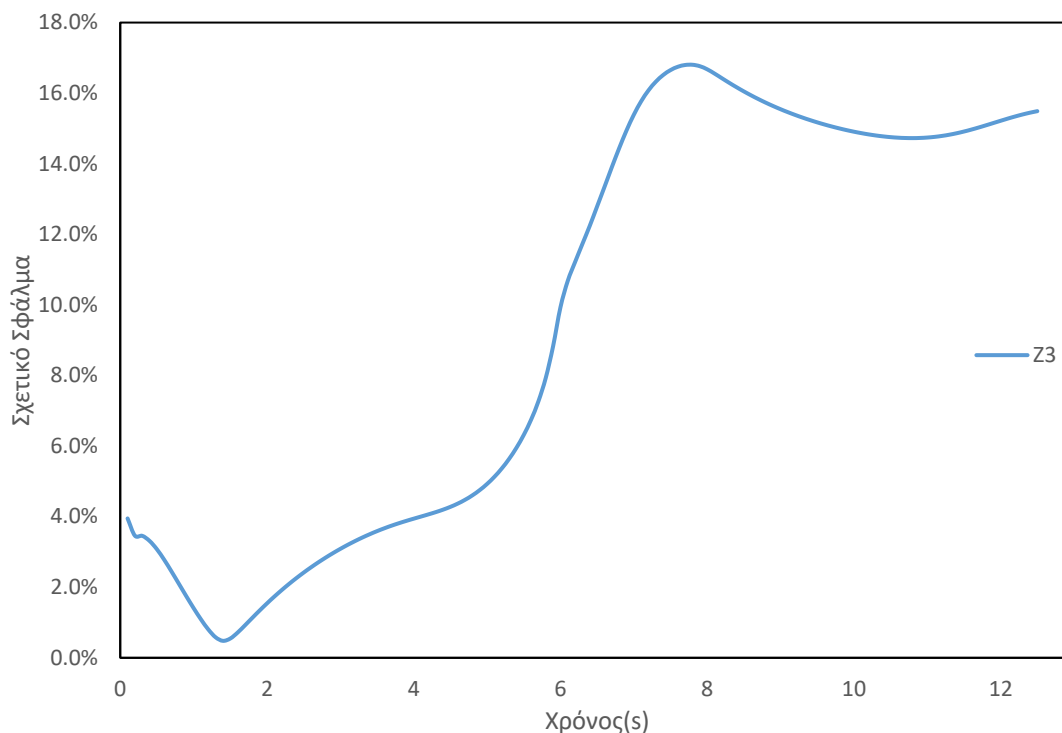
6.4.2.1 Διερεύνηση Νευρωνικού Δικτύου 1: Επιλογή ισάριθμων επιβολών εκπαίδευσης

Όπως έχει ήδη προαναφερθεί, στην πρώτη φάση της διερεύνησης εκπαιδεύτηκε νευρωνικό δίκτυο με χρήση 18 βηματικών επιβολών το οποίο παρουσίασε συγκεκριμένα προβλήματα, κυρίως όσον αφορά τις επιβολές εναλλαγής, όσο και τις επιβολές οι οποίες βρίσκονται στα σημεία στροφής, εκεί δηλαδή όπου παρατηρείται αλλαγή στην πολλαπλότητα λύσεων. Ειδικότερα, υπολογίστηκαν τα σχετικά σφάλματα όλων των προσεγγίσεων των επιβολών εκπαίδευσης σε σχέση με τις πραγματικές λύσεις που τους αντιστοιχούν και προέκυψε ότι το συγκεκριμένο δίκτυο μπορεί να προσομοιώσει ικανοποιητικά τόσο τον πάνω κλάδο λύσεων όσο και τον κάτω, με το σχετικό σφάλμα να κυμαίνεται στο 1-2%. Αστοχεί όμως σημαντικά να περιγράψει τις επιβολές εναλλαγής μεταξύ των κλάδων πολλαπλότητας. Ενδεικτικά, παρουσιάζεται το σχετικό σφάλμα της επιβολής εναλλαγής η οποία παρουσιάζεται στον πίνακα 6-11.

Πίνακας 6-11: Εξεταζόμενη επιβολή εκπαίδευσης

Αρχική Κατάσταση m_{10} (10^{-5} kg/s)	Κλάδος	Τελική Κατάσταση m_{10} (10^{-5} kg/s)	Κλάδος
6.0	Πάνω	2.0	Κάτω

Στο διάγραμμα 6-16 παρουσιάζεται το σχετικό σφάλμα της επιβολής εκπαίδευσης για τη βάση Z_3 .



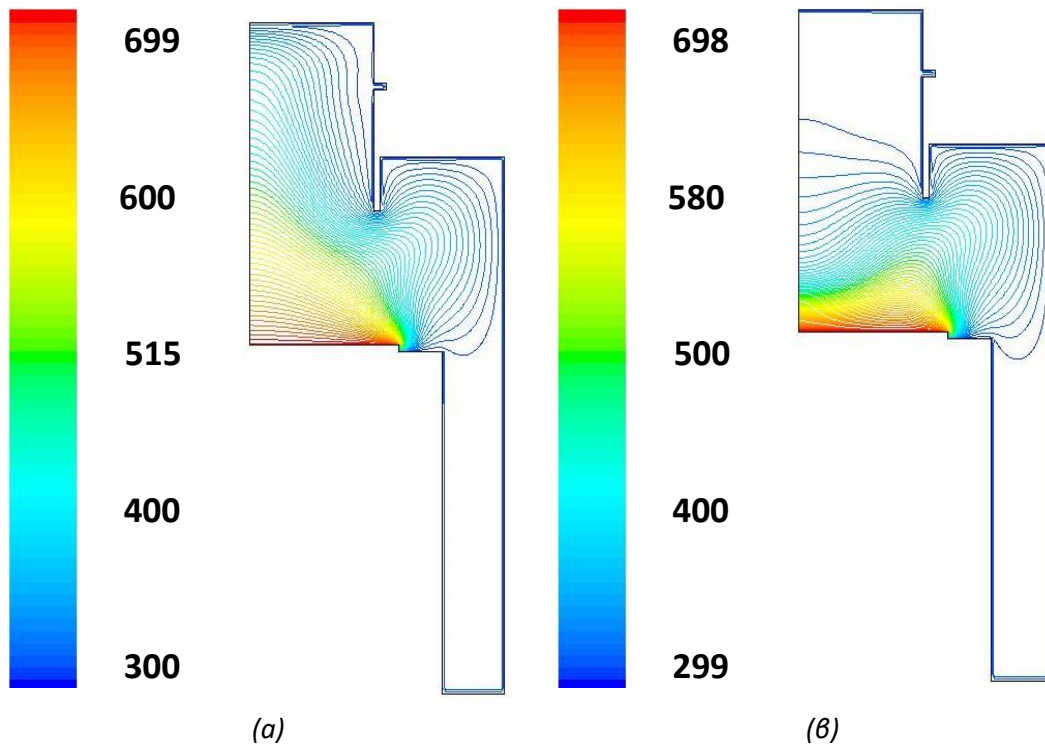
Εικόνα 6-16: Σχετικό σφάλμα του νευρωνικού δικτύου που εκπαιδεύτηκε στην πρώτη φάση της διερεύνησης την επιβολή από $\dot{m}(t)= 6.0 \cdot 10^{-5} \text{ kg/s}$ σε $\dot{m}(t)=2 \cdot 10^{-5} \text{ kg/s}$

Στον πίνακα 6-12 παρουσιάζεται το μέσο και το μέγιστο σχετικό σφάλμα της συγκεκριμένης επιβολής εκπαίδευσης.

Πίνακας 6-12: Μέσο και Μέγιστο Σχετικό Σφάλμα για την Επιβολή Εκπαίδευσης

α/α	Μέσο Σχετικό Σφάλμα	Μέγιστο Σχετικό Σφάλμα
1	11.28%	16.80%

Από τη μελέτη του διαγράμματος 6-16 σε συνδυασμό με τον πίνακα 6-12 εικάζεται ότι η επιβολή αυτή δεν μπορεί να προσεγγιστεί με χρήση του δικτύου που έχει εκπαιδευτεί. Για το λόγο αυτό, εισάγεται στον Fluent η προσέγγιση της λύσης τη χρονική στιγμή $t=12.5 \text{ s}$. Στην εικόνα 6-17 παρουσιάζεται συγκριτικά η προσέγγιση που προέκυψε σε σχέση με την πραγματική λύση.



Εικόνα 6-17: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει από την προσέγγιση του δικτύου που σχεδιάστηκε με χρήση 18 βηματικών επιβολών

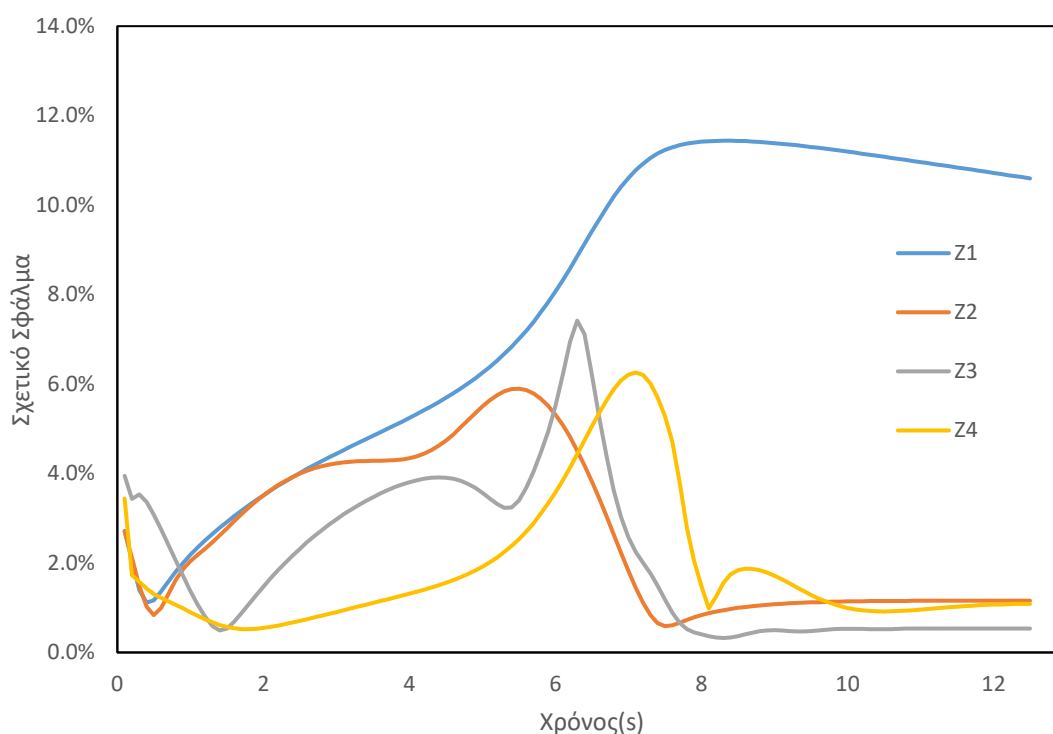
Από την Εικόνα 6-17 συμπεραίνεται ότι η προσεγγιστική λύση δεν έχει κάποια φυσική σημασία, διότι φαίνεται να βρίσκεται στην περιοχή όπου επικρατούν φαινόμενα εξαναγκασμένης συναγωγής, σε αναντιστοιχία με τη συγκεκριμένη παροχή εισόδου, η οποία έχει αποδειχτεί ότι βρίσκεται εκτός της περιοχής πολλαπλότητας στον κλάδο όπου επικρατεί φυσική ροή.

Το συγκεκριμένο δίκτυο δεν καταφέρνει να προσεγγίσει τις επιβολές εναλλαγής μεταξύ των δύο κλάδων πολλαπλότητας, άρα κρίνεται ακατάλληλο για τη συγκεκριμένη μελέτη, αφού επιζητείται ο σχεδιασμός ενός μοντέλου το οποίο θα περιγράφει εκτός των δύο ευσταθών κλάδων ισορροπίας, και τις αντίστοιχες εναλλαγές μεταξύ αυτών. Για αυτόν ακριβώς το λόγο αποφασίστηκε ο εμπλουτισμός των επιβολών εκπαίδευσης με επιπλέον υπολογιστικά δεδομένα, τα οποία περιλαμβάνουν κυρίως επιβολές εναλλαγής.

6.4.2.2 Διερεύνηση Νευρωνικού Δικτύου 2: Εμπλουτισμός δεδομένων εκπαίδευσης

Μετά την προσθήκη των βηματικών επιβολών όπως αναλύθηκε στην προηγούμενη ενότητα, εκπαιδεύτηκαν τα τέσσερα νευρωνικά δίκτυα και σε συνδυασμό με τις βάσεις Z_1 , Z_2 , Z_3 και Z_4 υπολογίστηκαν τα σχετικά σφάλματα των προσεγγίσεων των τροχιών εκπαίδευσης, σε σχέση με την πραγματική τροχιά. Ειδικότερα, τα εκπαιδευόμενα δίκτυα εμφανίζουν μικρά σφάλματα στις επιβολές οι οποίες πραγματοποιούνται στον πάνω και στον κάτω κλάδο πολλαπλότητας λύσεων, και με εξαίρεση το δίκτυο το οποίο σχεδιάστηκε για τη βάση Z_1 , μπορούν να προσεγγίσουν με ικανοποιητικό τρόπο τις επιβολές εναλλαγής.

Ενδεικτικά, παρουσιάζεται το σχετικό σφάλμα των τεσσάρων βάσεων για την επιβολή εκπαίδευσης η οποία παρουσιάστηκε στον πίνακα 6-11. Στο διάγραμμα 6-18 παρουσιάζονται συγκριτικά τα σχετικά σφάλματα της επιβολής εκπαίδευσης.



Εικόνα 6-18: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t) = 6.0 \cdot 10^{-5}$ kg/s σε $\dot{m}(t) = 2.0 \cdot 10^{-5}$ kg/s

Ενδεικτικά στους πίνακες 6-13 και 6-14 καταγράφονται το μέσο και το μέγιστο σχετικό σφάλμα, αντίστοιχα για την εξεταζόμενη βηματική επιβολή εναλλαγής για τα δίκτυα που έχουν εκπαιδευτεί για βάσεις διαφορετικής διάστασης d .

Πίνακας 6-13: Μέσο Σχετικό Σφάλμα για τα εκπαιδευόμενα δίκτυα

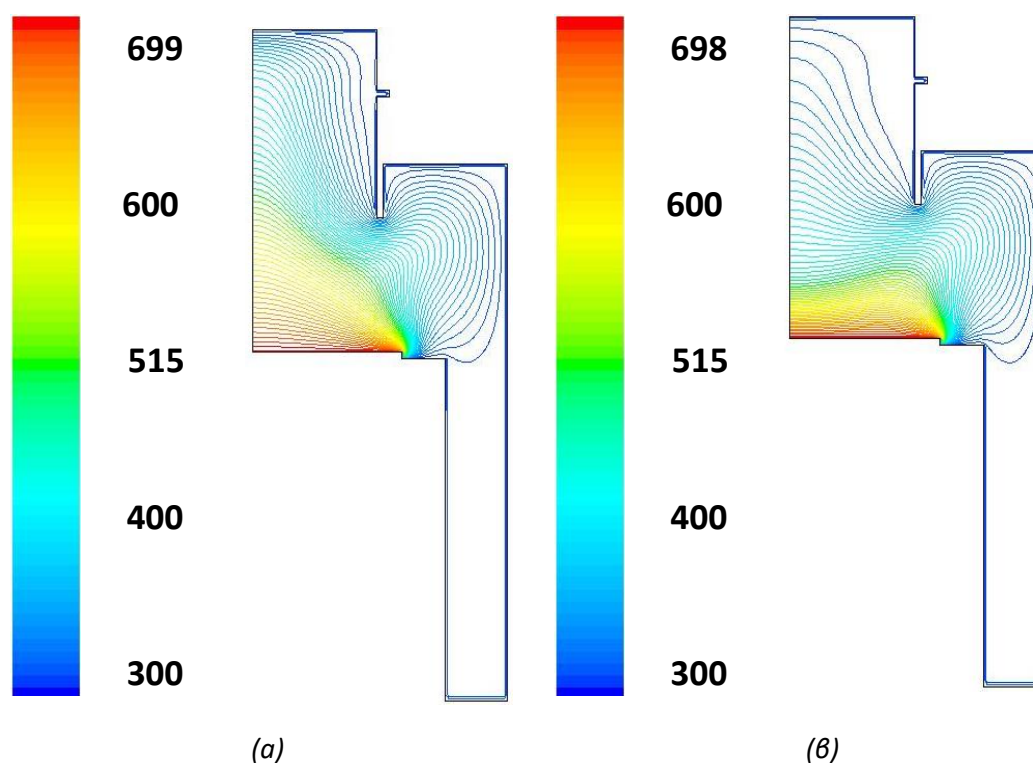
α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	8.52%	3.13%	2.70%	2.39%

Πίνακας 6-14: Μέγιστο Σχετικό Σφάλμα για τα εκπαιδευόμενα δίκτυα

α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	11.44%	5.90%	7.42%	6.26%

Γενικά, στο πλαίσιο αυτής της εργασίας επιδιώκεται το σφάλμα να είναι μικρότερο όσο η επιβολή πλησιάζει στη νέα μόνιμη κατάσταση. Συνεπώς, δεν είναι ανησυχητικά τα μέγιστα που παρουσιάζονται στο διάγραμμα 6-18 στη μέση του χρονικού διαστήματος. Αντίθετα, η σταθεροποίηση του σφάλματος του δικτύου που σχεδιάστηκε για τη βάση Z₁ σε σχετικό σφάλμα που ξεπερνά το 10% υποδεικνύει τη μη πραγματοποίηση της επιβολής αυτής. Ειδικότερα, υπολογίζεται μια υβριδική κατάσταση μεταξύ των δύο κλάδων, η οποία δεν έχει φυσική σημασία. Ενδεικτικά, παρουσιάζεται στην Εικόνα 6-19 η κατανομή της

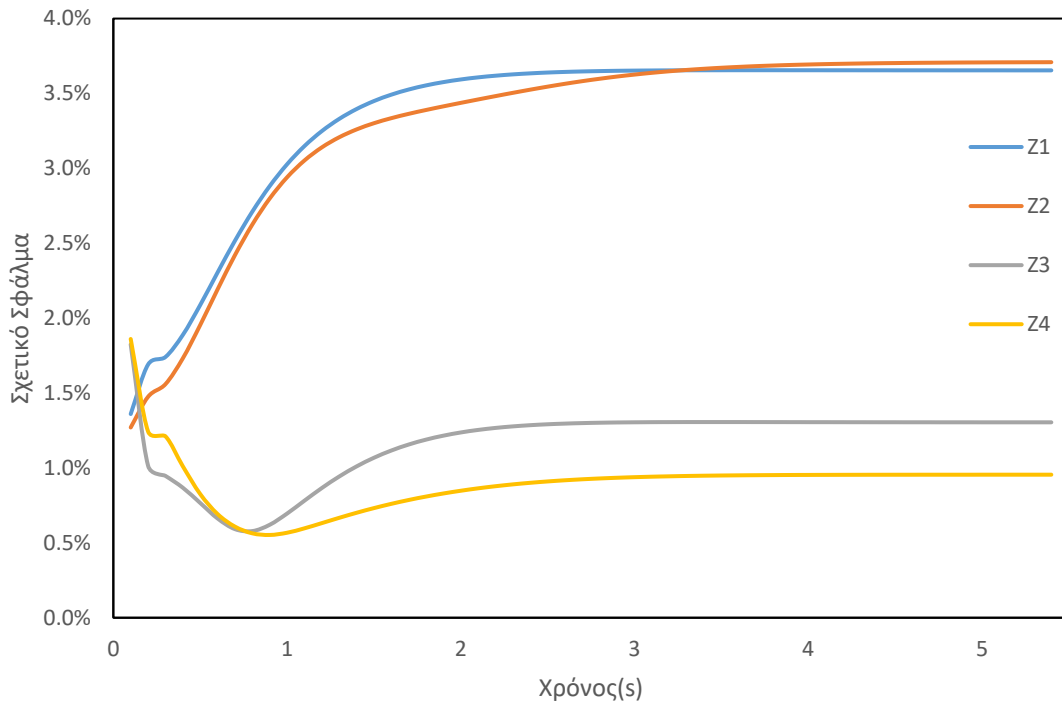
θερμοκρασίας της προσέγγισης που προκύπτει από το συγκεκριμένο δίκτυο σε σύγκριση με την πραγματική λύση.



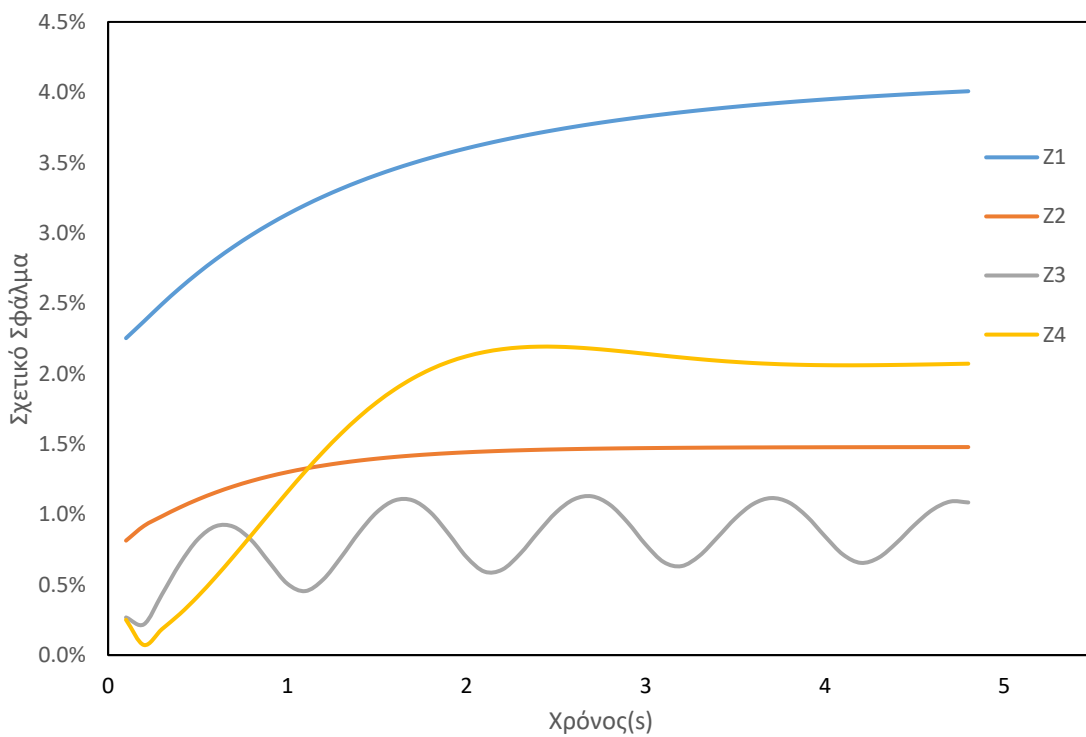
Εικόνα 6-19: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει από την προσέγγιση του δικτύου που σχεδιάστηκε σε συνδυασμό με τη βάση Z_1

Στην Εικόνα 6-19 παρουσιάζεται η συγκεκριμένη κατάσταση η οποία αναλύθηκε προηγουμένως. Είναι εμφανής η διαφορά της κατανομής της θερμοκρασίας, όπως επίσης η ομοιότητα της συγκεκριμένης κατανομής με αυτή η οποία παρουσιάστηκε στο κεφάλαιο 2, στην Εικόνα 2-4 η οποία παρουσιάζει την κατάσταση της θερμοκρασίας στον πάνω κλάδο. Το εύρος θερμοκρασιών στον αντιδραστήρα έχει υπολογιστεί σωστά, αλλά οι διαφορές που εμφανίζονται είναι σημαντικές. Συνεπώς το συγκεκριμένο δίκτυο παρουσιάζει προβλήματα ως προς την προσέγγιση των επιβολών εναλλαγής μεταξύ των κλάδων πολλαπλότητας.

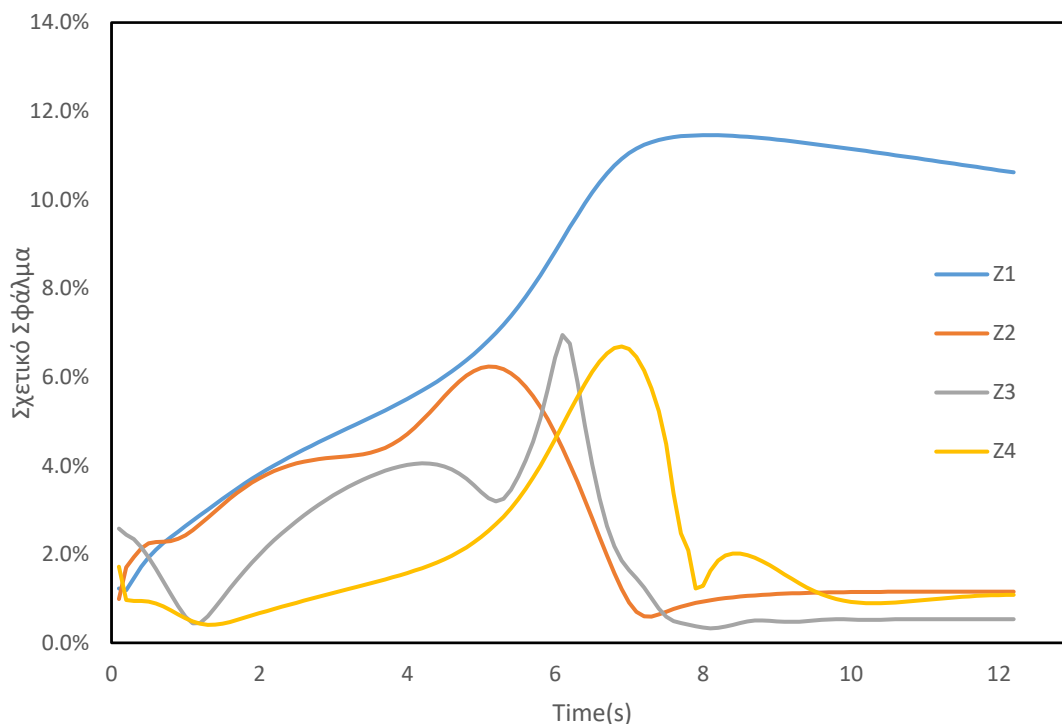
Εν συνεχεία, θα υπολογιστούν τα σχετικά σφάλματα των επιβολών ελέγχου, οι οποίες παρουσιάστηκαν προηγουμένως στον πίνακα 6-2. Τα συγκεκριμένα αποτελέσματα, καταγράφονται στα διαγράμματα 6-20 έως 6-22 συγκριτικά για τα τέσσερα δίκτυα τα οποία έχουν εκπαιδευτεί.



Εικόνα 6-20: Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t) = 4.0 \cdot 10^{-5} \text{ kg/s}$ σε $\dot{m}(t) = 6.0 \cdot 10^{-5} \text{ kg/s}$



Εικόνα 6-21: : Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t) = 2.0 \cdot 10^{-5} \text{ kg/s}$ σε $\dot{m}(t) = 1.5 \cdot 10^{-5} \text{ kg/s}$



Εικόνα 6-22: : Σχετικό σφάλμα κάθε βάσης για την επιβολή από $\dot{m}(t)=4.5 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=2.0 \cdot 10^{-5}$ kg/s

Στους πίνακες 6-15 και 6-16 καταγράφονται το μέσο και το μέγιστο σχετικό σφάλμα κάθε δικτύου που έχει εκπαιδευτεί, αντίστοιχα για τις τρεις εξεταζόμενες βηματικές επιβολές.

Πίνακας 6-15: Μέσο Σχετικό Σφάλμα για τα εκπαιδευόμενα δίκτυα

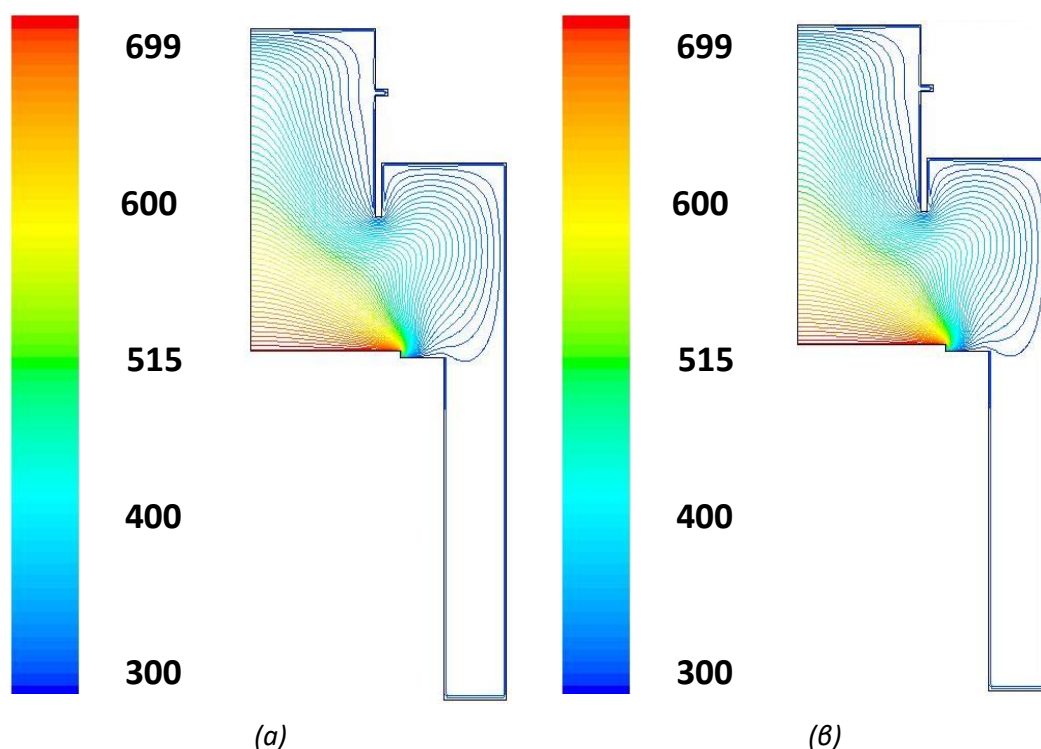
α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	3.40%	3.36%	1.20%	0.92%
2	3.57%	1.39%	0.85%	1.84%
3	8.63%	3.25%	2.61%	2.60%

Πίνακας 6-16: Μέγιστο Σχετικό Σφάλμα για τα εκπαιδευόμενα δίκτυα

α/α	Z ₁	Z ₂	Z ₃	Z ₄
1	3.66%	3.71%	1.82%	1.86%
2	4.01%	1.48%	1.13%	2.19%
3	11.46%	6.24%	6.95%	6.69%

Τα σχετικά σφάλματα των εκπαιδευόμενων δικτύων για τις επιβολές 1,2 οι οποίες πραγματοποιούνται στον πάνω και στον κάτω κλάδο δεν ξεπερνούν το 4.00% το οποίο κρίνεται ικανοποιητικό, δηλαδή μπορεί να θεωρηθεί ότι το σύστημα περιγράφεται με αρκετά μεγάλη ακρίβεια. Παρόλα αυτά το σφάλμα της επιβολής εναλλαγής για το δίκτυο της βάσης Z₁ είναι αρκετά μεγάλο και σταθεροποιείται στο τέλος της επιβολής όπου πλησιάζει στη μόνιμη κατάσταση περίπου στο 11%. Η μεγάλη αυτή απόκλιση στο τέλος της

επιβολής υποδηλώνει, όπως αναλύθηκε και για την επιβολή εκπαίδευσης, ότι δεν πραγματοποιείται η εναλλαγή των κλάδων και υπολογίζεται μια υβριδική κατάσταση μεταξύ των δύο κλάδων, η οποία δεν έχει φυσική σημασία. Το σφάλμα των υπόλοιπων τριών βάσεων κρίνεται αρκούτως μικρό και προσομοιώνει ικανοποιητικά την επιβολή εναλλαγής 3. Ενδεικτικά, παρουσιάζεται στην Εικόνα 6-23 η κατανομή της θερμοκρασίας της προσέγγισης που προκύπτει από το δίκτυο που σχεδιάστηκε για τη βάση Z_3 σε σύγκριση με την πραγματική λύση, τη χρονική στιγμή $t=12.2s$.



Εικόνα 6-23: Κατανομή Θερμοκρασίας (α) για τη λύση που προκύπτει μέσω Fluent και (β) για τη λύση που προκύπτει από την προσέγγιση της βάσης Z_3 για την επιβολή εναλλαγής

Τόσο το εύρος της θερμοκρασίας, όσο και η κατανομή αυτής στον αντιδραστήρα παρουσιάζουν ελάχιστες διαφορές, οι οποίες όταν ο υπολογιστικός κώδικας έχει συγκλίνει είναι πάρα πολύ μικρές. Υπενθυμίζεται ότι τα στιγμιότυπα αυτά λήφθηκαν τη χρονική στιγμή όπου το σύστημα έχει συγκλίνει στη νέα μόνιμη κατάσταση.

Συμπερασματικά, τα δίκτυα που σχεδιάστηκαν για τις βάσεις Z_3 και Z_4 παρουσιάζουν τα μικρότερα σφάλματα και την καλύτερη προσέγγιση της κατάστασης στον αντιδραστήρα ΧΑΑ. Το δίκτυο που κατασκευάστηκε για τη βάση Z_3 αντιστοιχεί σε μικρότερο σχετικό πληροφοριακό περιεχόμενο συγκριτικά με τη βάση Z_4 παράγοντας περίπου ίδιας ποιότητας αποτελέσματα, αν όχι καλύτερα. Συνεπώς, προτείνεται και επιλέγεται η χρήση του πρώτου δικτύου αφού τόσο οι απαιτήσεις υπολογιστικού χρόνου όσο και αυτές της μνήμης είναι μικρότερες.

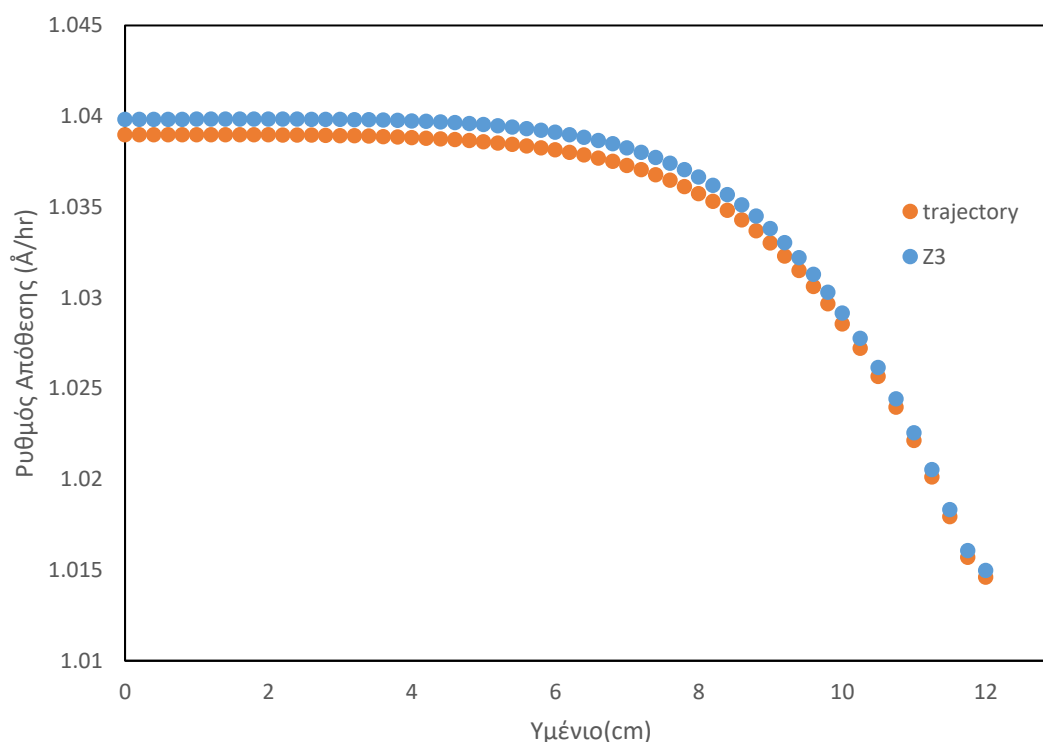
6.5 Μελέτη Χημικού Συστήματος

Το μοντέλο της μειωμένης τάξης το οποίο έχει σχεδιαστεί μπορεί να προσεγγίσει με ελάχιστο σφάλμα τη δυναμική κατάσταση του μελετώμενου αντιδραστήρα στην περιοχή πολλαπλότητας χώρου λύσεων. Παρόλα αυτά το μοντέλο αυτό, σχεδιάστηκε απουσία χημικών αντιδράσεων, συνεπώς δεν μπορούν να προκύψουν συμπεράσματα σχετικά με την παραγωγή του υμενίου, το ρυθμό απόθεσης του, την ομοιομορφία ή το πάχος του.

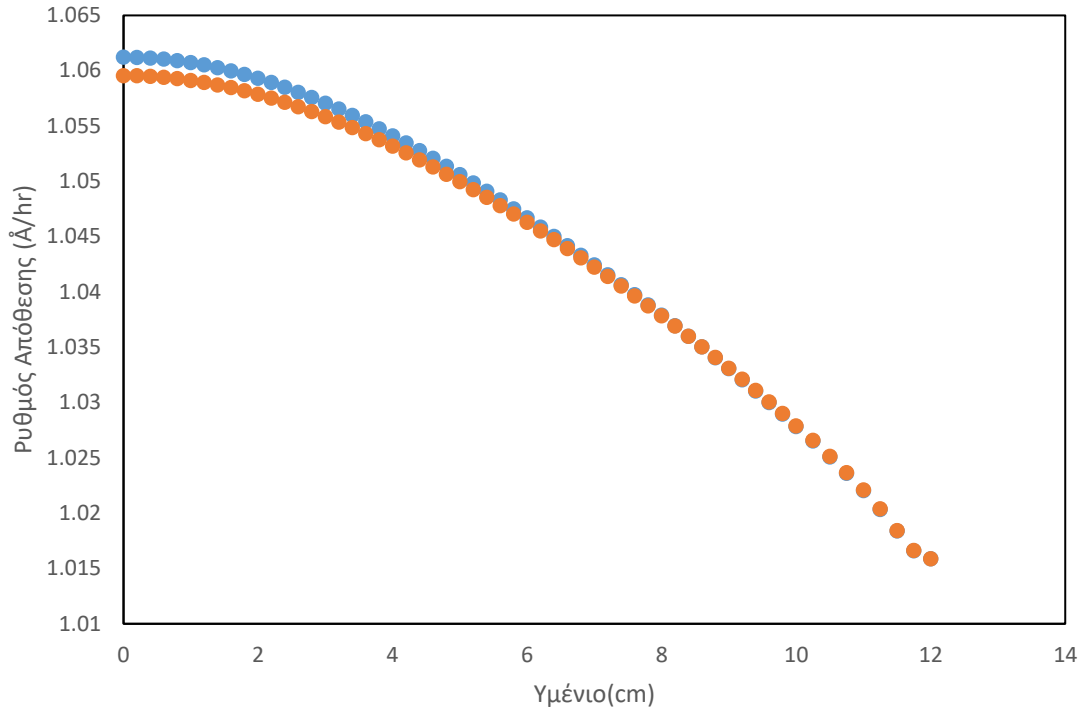
Όπως έχει αναφερθεί στο Κεφάλαιο 5 και σύμφωνα με τη βιβλιογραφία [2], το πεδίο ροής ελάχιστα διαφέρει είτε στον αντιδραστήρα λαμβάνουν χώρα χημικές αντιδράσεις είτε όχι. Συνεπώς, για τον υπολογισμό των ποσοτήτων που συνδέονται με τη χημεία απαιτούνται ελάχιστες επαναλήψεις από το λογισμικό Fluent για να ικανοποιηθούν («κλείσουν») τα ισοζύγια μάζας των χημικών συστατικών.

Σύμφωνα με τη μεθοδολογία που αναλύθηκε προηγουμένως, εισάγεται στο λογισμικό Fluent η προσέγγιση και στη συνέχεια αφήνεται για λίγες επαναλήψεις χωρίς αυτός να συγκλίνει, αλλά απλώς να προκύψει ο ρυθμός απόθεσης για την προσεγγιστική λύση.

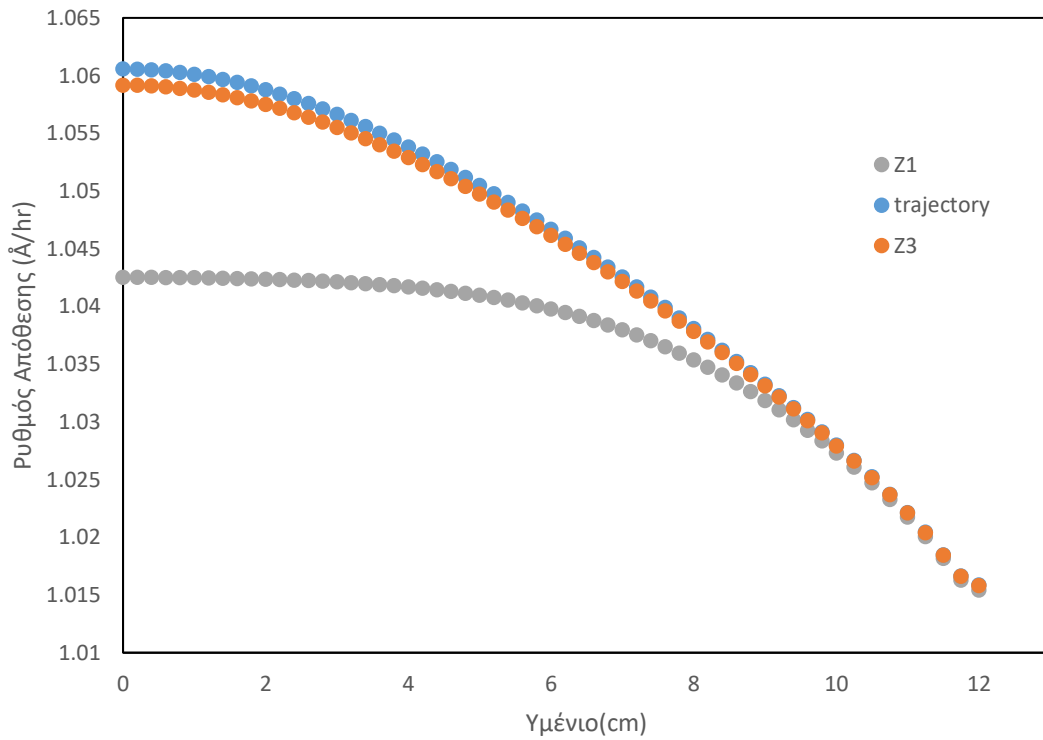
Συγκεκριμένα, συγκρίθηκαν οι ρυθμοί απόθεσης των προσεγγίσεων και των πραγματικών καταστάσεων για τις επιβολές του πίνακα 6-2. Παρακάτω, παρουσιάζονται συγκριτικά τα αποτελέσματα:



Εικόνα 6-24: Συγκριτική απεικόνιση του ρυθμού απόθεσης της πραγματικής και της προσεγγιστικής λύσης για την επιβολή από $\dot{m}(t)=4.0 \cdot 10^{-5}$ kg/s σε $\dot{m}(t)=6.0 \cdot 10^{-5}$ kg/s



Εικόνα 6-25: Συγκριτική απεικόνιση του ρυθμού απόθεσης της πραγματικής και της προσεγγιστικής λύσης για την επιβολή από $\dot{m}(t) = 2.0 \cdot 10^{-5} \text{ kg/s}$ σε $\dot{m}(t) = 1.5 \cdot 10^{-5} \text{ kg/s}$



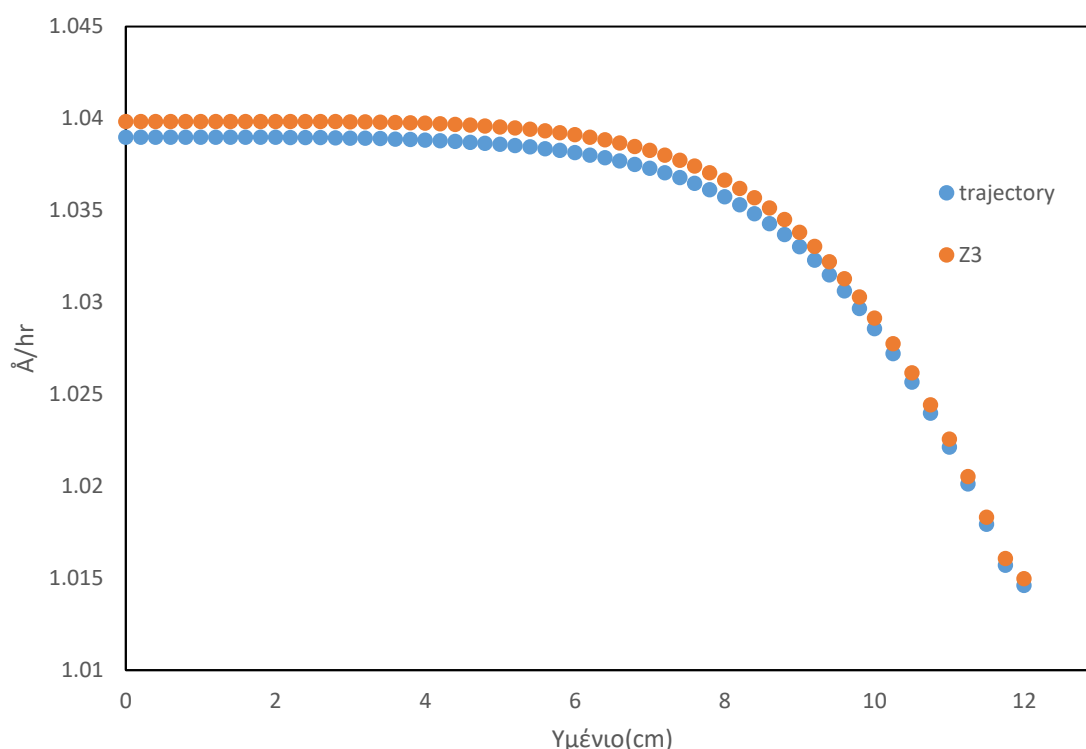
Εικόνα 6-26: Συγκριτική απεικόνιση του ρυθμού απόθεσης της πραγματικής και της προσεγγιστικής λύσης για δύο διαφορετικά δίκτυα για την επιβολή από $\dot{m}(t) = 4.5 \cdot 10^{-5} \text{ kg/s}$ σε $\dot{m}(t) = 2.0 \cdot 10^{-5} \text{ kg/s}$

Λόγω της μεγαλύτερης ομοιομορφίας που επιτυγχάνεται στον πάνω κλάδο πολλαπλότητας επιδιώκεται η λειτουργία σε αυτές τις καταστάσεις. Γι' αυτό το λόγο, παρουσιάζεται μία επιπλέον επιβολή εναλλαγής η οποία πραγματοποιείται από τον κάτω κλάδο ισορροπίας στον πάνω. Η συγκεκριμένη επιβολή παρουσιάζεται στον πίνακα 6-17:

Πίνακας 6-17: Εξεταζόμενη επιβολή εναλλαγής από τον κάτω κλάδο πολλαπλότητας στον πάνω

α/α	Αρχική Κατάσταση \dot{m}_0 (10^{-5} kg/s)	Κλάδος	Τελική Κατάσταση \dot{m}_0 (10^{-5} kg/s)	Κλάδος
1	2.0	Κάτω	6.0	Πάνω

Στη συνέχεια, παρουσιάζεται το συγκριτικό διάγραμμα του ρυθμού απόθεσης της πραγματικής και της προσεγγιστικής κατάστασης.



Εικόνα 6-27: Συγκριτική απεικόνιση του ρυθμού απόθεσης της πραγματικής και της προσεγγιστικής λύσης για δύο διαφορετικά δίκτυα για την επιβολή από $\dot{m}(t) = 2.0 \cdot 10^{-5}$ kg/s σε $\dot{m}(t) = 6.0 \cdot 10^{-5}$ kg/s

Σε γενικές γραμμές οι διαφορές οι οποίες καταγράφηκαν στα διαγράμματα 6-24 έως 6-27 είναι πολύ μικρές. Η προσέγγιση η οποία έχει κατασκευαστεί με χρήση της ανάλυσης κύριων συνιστωσών και των νευρωνικών δικτύων έχει τη δυνατότητα εκτός από την περιγραφή της κατάστασης του αντιδραστήρα, να προσδιορίσει σε μεγάλο βαθμό τη απόθεση του υμενίου.

Ενδιαφέρον αποτέλεσμα είναι ο ρυθμός απόθεσης ο οποίος υπολογίστηκε για την κακή πρόβλεψη του νευρωνικού δικτύου της βάση Z_1 . Ειδικότερα, η μορφή του ρυθμού απόθεσης παραπέμπει σε υμένιο το οποίο δημιουργείται στον κλάδο της εξαναγκασμένης κυκλοφορίας παρόλο που στη συγκεκριμένη παροχή εισόδου $\dot{m}(t) = 2 \cdot 10^{-5}$ kg/s δεν υφίσταται η συγκεκριμένη κατάσταση.

Τέλος, υπολογίστηκε η ανομοιομορφία των παραγόμενων υμενίων σύμφωνα με τη σχέση (5.2) για τις εξεταζόμενες επιβολές. Τα αποτελέσματα παρουσιάζονται στον πίνακα 6-18:

Πίνακας 6-18: Ανομοιομορφία Παραγόμενου Υμενίου για τις 4 Εξεταζόμενες Επιβολές

Αρχική Κατάσταση $\dot{m}(t)$ (10^{-5} kg/s)	Τελική Κατάσταση $\dot{m}(t)$ (10^{-5} kg/s)	Ανομοιομορφία Πραγματικής Λύσης (%)	Ανομοιομορφία Προσεγγιστικής Λύσης (%)
4.0 (Πάνω Κλάδος)	6.0 (Πάνω Κλάδος)	2.35%	2.39%
2.0 (Κάτω Κλάδος)	1.5 (Κάτω Κλάδος)	4.27%	4.12%
4.5 (Πάνω Κλάδος)	2.0 (Κάτω Κλάδος)	4.22%	4.09%
2.0 (Κάτω Κλάδος)	6.0 (Πάνω Κλάδος)	2.35%	2.39%

Από τη μελέτη του πίνακα 6-18 σε συνδυασμό με τα διαγράμματα 6-24 έως 6-27 επιβεβαιώνεται ότι το μοντέλο μειωμένης τάξης μπορεί να προσομοιώσει σε ικανοποιητικό βαθμό τη μορφή του υμενίου πυριτίου, δηλαδή εκτός του ότι έχει τη δυνατότητα να υπολογίσει την κατάσταση του αντιδραστήρα, μπορεί συγχρόνως να παράγει αποτελέσματα σχετικά με την ποιότητα του προϊόντος. Πρόκειται συνεπώς για ένα πολύ χρήσιμο εργαλείο για τη βιομηχανία παραγωγής υμενίων, αφού ανά πάσα στιγμή μπορεί να εξαχθούν συμπεράσματα σχετικά με τη λειτουργία του αντιδραστήρα, καθώς και για το προϊόν, με ασήμαντη κατανάλωση υπολογιστικού χρόνου και μνήμης.

Από τα στοιχεία του 6-18 γίνεται εμφανής η διαφορά στην ανομοιομορφία η οποία δημιουργείται στους δύο διαφορετικούς κλάδους λύσεων. Συγκεκριμένα, στον κάτω κλάδο πολλαπλότητας λύσεων η ανομοιομορφία του υμενίου προσδιορίζεται περίπου στο 4.2% ενώ αντίθετα, στον πάνω κλάδο περίπου στο 2.3%. Σημειώνεται ότι, η ανομοιομορφία της επιβολής 2 και 4 είναι ίδια διότι το σύστημα έχει συγκλίνει στην ίδια μόνιμη κατάσταση $\dot{m}(t)=6 \cdot 10^{-5}$ kg/s.

Συμπεράσματα

Η διεργασία της χημικής απόθεσης από ατμό είναι μια ευρέως χρησιμοποιούμενη μέθοδος στη βιομηχανία με πλήθος εφαρμογών όπως η παραγωγή ημιαγωγικών υλικών, υλικών μικροηλεκτρονικής κ.α. Το γεγονός αυτό καθιστά απαραίτητη τη διεξοδική μελέτη των αντιδραστήρων, των συνθηκών λειτουργίας καθώς και της δυναμικής τους συμπεριφοράς, καθώς οι προδιαγραφές ως προς την ομοιομορφία της ανάπτυξης του μενίου είναι σε αρκετές περιπτώσεις αυστηρές.

Κατά την εκπόνηση της συγκεκριμένης εργασίας πραγματοποιήθηκε εκτενής μελέτη της δυναμικής συμπεριφοράς ενός αντιδραστήρα χημικής απόθεσης από ατμό, απουσία χημικών αντιδράσεων. Το χαρακτηριστικό του μελετώμενου συστήματος το οποίο το καθιστά ιδιαίτερα ενδιαφέρον τόσο από θεωρητική όσο και από πρακτική σκοπιά, είναι η πολλαπλότητα χώρου λύσεων που παρουσιάζει. Ειδικότερα, μπορούν να παρατηρηθούν τρεις διαφορετικές καταστάσεις, δύο ευσταθείς και μία ασταθής, στον αντιδραστήρα για μια συγκεκριμένη τιμή παροχής εισόδου $\dot{m}(t)$ η οποία κυμαίνεται μεταξύ $2.3 \cdot 10^{-5} - 5 \cdot 10^{-5}$ kg/s. Η συγκεκριμένη μελέτη πραγματοποιήθηκε με χρήση του εμπορικού λογισμικού υπολογιστικής ρευστοδυναμικής ANSYS Fluent 13.0.

Κύριος στόχος της παρούσας διπλωματικής εργασίας ήταν η δημιουργία ενός μειωμένης τάξεως μοντέλου, ο οποίος επιτεύχθηκε με χρήση της ανάλυσης κύριων συνιστωσών σε συνδυασμό με κατάλληλα εκπαιδευόμενα τεχνητά νευρωνικά δίκτυα. Το φυσικό σύστημα παρουσιάζει εξάρτηση τόσο από το χώρο όσο και από το χρόνο. Ειδικότερα, ο σχεδιασμός αυτός πραγματοποιήθηκε με χρήση αποτελεσμάτων προσομοιώσεων και με εφαρμογή της μεθόδου στιγμιοτύπων, και την κατάλληλη εκπαίδευση νευρωνικού δικτύου με ανατροφοδότηση. Η κατάσταση του αντιδραστήρα προσεγγίζεται ως γραμμικός συνδυασμός της ορθοκανονικής βάσης, από όπου προκύπτει η χωρική εξάρτηση, και των συντελεστών $\alpha(t)$, από όπου προκύπτει η χρονική εξάρτηση. Μετά από διερεύνηση ως προς την ποιότητα των παραγόμενων προσεγγίσεων επιλέχθηκε και χρησιμοποιήθηκε η ορθοκανονική βάση η οποία αποτελείται από τρεις ιδιοκαταστάσεις σε συνδυασμό με νευρωνικό δίκτυο το οποίο περιέχει 12 κρυφά επίπεδα.

Το μοντέλο αυτό έχει τη δυνατότητα να προσομοιώσει και να προσεγγίσει σε πολύ μεγάλο βαθμό την κατάσταση του αντιδραστήρα μετά την επιβολή κάποιας διαταραχής, τόσο κατά τη διάρκεια της επιβολής όσο και στη νέα μόνιμη κατάσταση. Οι επιβολές αυτές μπορεί να πραγματοποιούνται είτε στους δύο ευσταθείς κλάδους ισορροπίας, είτε σε εναλλαγές μεταξύ αυτών. Σημαντικό πλεονέκτημα της συγκεκριμένης μεθόδου αποτελεί το γεγονός ότι δεν απαιτείται προβολή τύπου Galerkin των εξισώσεων πάνω στη βάση για να προκύψουν οι συντελεστές. Συνεπώς, από την στιγμή που το μοντέλο μειωμένης τάξης κατασκευαστεί δεν απαιτεί καμία γνώση των εξισώσεων και των μηχανισμών που επικρατούν στο σύστημα, για την προσέγγιση της κατάστασης στον αντιδραστήρα για μια συγκεκριμένη μαζική παροχή εισόδου.

Στην τελική φάση της εργασίας, λήφθηκε υπόψη ένα απλοποιημένο χημικό σύστημα για την απόθεση πυριτίου(Si) από πρόδρομη ένωση σιλάνιο (SiH_4), του οποίου το κλάσμα μάζας στην είσοδο του αντιδραστήρα είναι 10^{-3} . Κύριος στόχος της συγκεκριμένης

ανάλυσης είναι ο προσδιορισμός της επίδρασης της προσεγγιστικής λύσης, όπως παράγεται από το μειωμένης τάξης μοντέλο, στο ρυθμό απόθεσης και κατ' επέκταση στο παραγόμενο προϊόν. Μετά από συγκριτική μελέτη της μορφής και της ανομοιομορφίας του υμενίου πυριτίου όπως υπολογίζεται από την πραγματική και την προσεγγιστική κατάσταση προέκυψε ότι οι διαφορές είναι μηδαμινές.

Συμπερασματικά, το μοντέλο μειωμένης τάξης το οποίο σχεδιάστηκε μπορεί να προβλέψει με ικανοποιητική ακρίβεια όλες τις ενδιαφέρουσες πτυχές της δυναμικής συμπεριφοράς του συστήματος, με μόνη "γνώση" την αρχική κατάσταση του συστήματος καθώς και τη νέα τιμή της παροχής εισόδου, χωρίς καμία περαιτέρω πληροφορία για τις εξισώσεις οι οποίες διέπουν το σύστημα. Μπορεί να αποτελέσει ένα σημαντικό εργαλείο για την πρόβλεψη της συμπεριφοράς σε πραγματικό χρόνο, αφού απαιτεί ελάχιστο χρόνο για τον υπολογισμό της νέας κατάστασης. Τέλος, η πρόβλεψη του μειωμένης τάξης μοντέλου στη μόνιμη κατάσταση μπορεί να χρησιμοποιηθεί ως αρχική τιμή για την πιο γρήγορη επίλυση του συστήματος από το εμπορικό λογισμικό υπολογιστικής ρευστοδυναμικής ANSYS Fluent, αφού η προσέγγιση της λύσης διαφέρει ελάχιστα από την πραγματική.

Προτάσεις για Περαιτέρω Έρευνα

Τα αποτελέσματα της δυναμικής ανάλυσης του συστήματος χημικής απόθεσης από ατμό βοηθούν στην περαιτέρω κατανόησή του, και δίνουν έναυσμα στην προσπάθεια βελτίωσης του.

Όπως έχει τονιστεί, το συγκεκριμένο σύστημα δημιουργήθηκε με μόνη μεταβαλλόμενη τιμή την παροχή εισόδου στον αντιδραστήρα. Ένα βήμα για τη δημιουργία ενός πιο ρεαλιστικού μοντέλου θα ήταν η δυνατότητα να μεταβάλλονται με το χρόνο και άλλες τιμές καθοριστικές για τη λειτουργία του αντιδραστήρα όπως η θερμοκρασία των τοιχωμάτων ή/και του υποστρώματος, η θερμοκρασία εισόδου του μίγματος και η πίεση λειτουργίας. Με αυτόν τον τρόπο θα μπορούσε να μελετηθεί διεξοδικά η επίδραση των παραπάνω μεγεθών στο πεδίο ροής και στην πολλαπλότητα των λύσεων. Το μοντέλο μειωμένης τάξης το οποίο θα παραχθεί θα μπορεί να προσομοιώσει πιο ρεαλιστικά την κατάσταση του αντιδραστήρα. Ειδικότερα, θα δέχεται ως μεταβλητές εισόδου τις τιμές των μελετώμενων διαταραχών, όπως στη συγκεκριμένη περίπτωση την παροχή εισόδου, και θα παράγει τη νέα κατάσταση του αντιδραστήρα.

Το μοντέλο το οποίο έχει σχεδιαστεί μπορεί να προσομοιώσει ικανοποιητικά τη δυναμική συμπεριφορά του συστήματος χημικής απόθεσης από ατμό. Προτείνεται η χρήση του μοντέλου αυτού σε ένα σύστημα ρύθμισης προβλεπτικού μοντέλου. Τα συστήματα αυτόματης ρύθμισης αποτελούν αναπόσπαστο στοιχείο πολλών χημικών διεργασιών, αλλά η ρύθμιση στους αντιδραστήρες χημικής απόθεσης από ατμό πραγματοποιείται με μη αυτόματους τρόπους, και βασίζεται κυρίως στην εμπειρία του χειριστή. Συνεπώς, απαραίτητη είναι η διεξοδική μελέτη ενός ρυθμιστή προβλεπτικού μοντέλου και κατά πόσο αυτός μπορεί να δώσει πιο αξιόπιστα αποτελέσματα σε σχέση με τα τρέχοντα στη βιομηχανία.

Τέλος, προτείνεται η επαύξηση του μοντέλου με πειραματικά δεδομένα, δηλαδή να διερευνηθεί κατά πόσο πειραματικές τιμές από κάποιο σημείο –χώρο– του αντιδραστήρα θα μπορούσαν να προστεθούν στο μοντέλο και μαζί να προβλέψουν την κατάσταση που επικρατεί στον αντιδραστήρα τη συγκεκριμένη χρονική στιγμή και κατ' επέκταση στην ποιότητα του παραγόμενου προϊόντος. Βέβαια, σε μια πειραματική διάταξη δεν μπορούν να συλλεχθούν πληροφορίες σχετικά με την κατάσταση του αντιδραστήρα σε μεγάλη έκταση, όπως στην προσομοίωση, αλλά μεμονωμένες τιμές οι οποίες μπορεί να αφορούν το πάχος του υμενίου, τη θερμοκρασία κ.α. σε μερικά σημεία του χώρου. Το συγκεκριμένο μοντέλο θα αποτελούσε σημαντικό και χρήσιμο εργαλείο για τη βιομηχανία, αφού θα μπορούσε ανά πάσα στιγμή να γνωρίζει την κατάσταση στον αντιδραστήρα.

Βιβλιογραφία

- [1] N. Cheimarios, E. D. Koronaki and A. G. Boudouvis, "Enabling a commercial computational fluid dynamics code to perform certain nonlinear analysis tasks," *Computers and Chemical Enigneering* **35**, pp. 2632-2645, 2011.
- [2] N. Χειμαριός, "Προσομοίωση πολλαπλών χωρικών κλιμάκων και συστημική ανάλυση διεργασιών χημικής απόθεσης από ατμό," Διδακτορική Διατριβή, Σχολή Χημικών Μηχανικών ΕΜΠ, Αθήνα, 2012.
- [3] A. Fluent, Theory Guide, Ansys Inc 5, 2009.
- [4] H. Pedersen and S. D. Elliot, "Studying chemical vapor deposition processes with theoretical chemistry," *Theoretical Chemistry Accounts* **133**, p. 1476, 2014.
- [5] C. R. Kleijn, "A mathematical model of the hydrodynamics and gas-phase reaction in silicon LPCVD in a single-wafer reactor," *Journal of the Electrochemical Society* **138**, pp. 2190-2200, 1991.
- [6] Θ. Ν. Παπαπέτρου, "Ανάπτυξη μειωμένης τάξης μοντέλου αντιδραστήρα χημικής απόθεσης από ατμό με τη μέθοδο της ανάλυσης κύριων συνιστωσών και τεχνητά νευρωνικά δίκτυα," Διπλωματική Εργασία, Σχολή Χημικών Μηχανικών ΕΜΠ, Αθήνα, 2016.
- [7] M. R2016a. [Online]. Available: <http://www.mathworks.com/products/matlab/>.
- [8] W. Schilders, H. vdn der Vorst and J. Rommes, Model order reduction theory, Springer-Verlag Berlin Heidelberg, 2008.
- [9] A. J. Newman, "Model reduction via the Karhunen-Loeve expansion part II: Some elementary examples," Technical Report T.R. 96-33, Inst. Systems Reasearch, April, 1996.
- [10] L. Sirovich, "Turbulence and the dynamics of coherent structures. Part I-III," *Quarterly of Applied Mathematics* **45**, pp. 561-590, 1987.
- [11] F. L. Lewis, S. Jagannathan and A. Yesildirek, Neural network control of robot manipulators and nonlinear systems, London: Taylor & Francis, 1999.
- [12] M. T. Hagan, H. B. Demuth, O. De Jesus and M. H. Beale, Neural Network Design.
- [13] L. Tsunganan, C. Lee Giles, B. G. Horne and S. Y. Kung, "A delay damage model selection algorithm for NARX neural network," U.of Maryland Technical Report CS-TR-3707.
- [14] O. Baghirli, "Comparison of Levenberg-Marquardt, Scaled conjugate gradient and bayesian regularization backpropagation algorithms for multistep ahead wind speed forecasting using multilayer percptron feedforward neural network," Uppsala Universitet, June 2015.
- [15] M. Kayri, "Predictive abilities of Bayesian Regularization and Levenberg-Marquardt algorithms in artificial neural networks: A comparative empirical study on social data,"

Mathematical and Computational Applications **21**, pp. 1-11, 2016.

- [16] P. Kumar, S. N. Merchant and U. B. Desai, "Improving performance in pulse radar detection using Bayesian regularization for neural network training," *Digital Signal Processing* **14**, pp. 438-448, 2004.
- [17] M. E. Coltrin, R. J. Kee and J. A. Miller, "A mathematical model of the coupled fluid mechanics and chemical kinetics in a chemical vapor deposition reactor," *Journal of the Electrochemical Society* **131**, pp. 425-434, 1984.
- [18] H. Demuth, M. Beale and M. Hagan, Neural Network Toolbox 6, User Guide.
- [19] P. Nigro, M. Anndif, Y. Teixeira, P. M. Pimeta and P. Wriggers, "An adaptive model order reduction by proper snapshot selection for nonlinear dynamical problems," *Computational Mechanics* **57**, pp. 537-554, 2016.
- [20] N. Cheimarios, G. Kokkoris and A. G. Boudouvis, "Multiscale modeling in chemical vapor deposition processes: Coupling reactor scale with feature scale computations," *Chemical Engineering Science* **65**, pp. 5018-5028, 2010.
- [21] P. A. Gkinis, I. G. Aviziotis, E. D. Koronaki and A. G. Boudouvis, "The effects of flow multiplicity on GaN deposition in a rotating disk CVD reactor," *Journal of Crystal Growth* **458**, pp. 140-148, 2017.
- [22] N. Cheimarios, E. D. Koronaki and A. G. Boudouvis, "Illuminating nonlinear dependence of film deposition rate in a CVD reactor on operating condition," *Chemical Engineering Journal* **181**, pp. 516-523, 2012.
- [23] W. Xie, I. Bonis and C. Theodoropoulos, "Linear MPC based on data-driven artificial neural networks for large-scale nonlinear distributed parameter systems," in *22nd European Symposium on Computer Aided Process Engineering*, London, 17 - 20 June 2012.