



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Χημικών Μηχανικών

Τομέας II – Ανάλυσης, Σχεδιασμού και Ανάπτυξης Διεργασιών
και Συστημάτων

**Εναλλακτική Υλοποίηση της Ανάλυσης Κυρίων Συνιστωσών
για τη διερεύνηση μη γραμμικής συμπεριφοράς
της Χημικής Απόθεσης από Ατμό**

Διπλωματική Εργασία

Μπαΐλη Κυριακή Βαλεντίνα

Επιβλέπων Καθηγητής

Ανδρέας Μπουντουβής

Αθήνα, 2019

Ευχαριστίες

Ολοκληρώνοντας την εργασία αυτή και κατά συνέπεια και τις σπουδές μου, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή Ανδρέα Μπουντουβή για την ευκαιρία που μου έδωσε να εκπονήσω τη διπλωματική μου εργασία υπό την επίβλεψη του ιδίου και της ερευνητικής του ομάδας. Θα ήθελα να τον ευχαριστήσω για την εμπιστοσύνη που μου έδειξε και την στήριξη του όλο αυτό το διάστημα, ώστε να μπορέσω να φέρω εις πέρας την εργασία. Τον ευχαριστώ ακόμα για την προσοχή του και τις εκτενείς διορθώσεις του στην παρούσα εργασία.

Ευχαριστώ επίσης την Δρα Ελένη Κορωνάκη για την συνεχή και πολύτιμη καθοδήγηση της κατά την διάρκεια της εκπόνησης της εργασίας. Θα ήθελα ακόμα να την ευχαριστήσω για τη βοήθεια της πάνω στο κομμάτι της υπολογιστικής ρευστοδυναμικής και στην ανάπτυξη μοντέλων μειωμένης τάξης. Η μεγάλη γνώση της πάνω στο μελετώμενο σύστημα καθώς και οι συμβουλές της με βοήθησαν πολύ να συνεχίσω.

Επιπλέον, θα ήθελα να ευχαριστήσω τον υποψήφιο διδάκτορα Παύλο Γκίνη για την βοήθεια και τις συμβουλές του, από την αρχή της εργασίας μέχρι και την ολοκλήρωση της. Οι προτάσεις και οι υποδείξεις του στην ανάπτυξη μοντέλων μειωμένης τάξης καθώς και στην εκπαίδευση των τεχνητών νευρωνικών δικτύων με βοήθησαν να ολοκληρώσω την εργασία.

Επιπροσθέτως, ευχαριστώ τον Θοδωρή Παπαπέτρου για τις γνώσεις που μου έδωσε μέσω των γραπτών του σχετικά με το μελετώμενο σύστημα, καθώς και για όλο το υλικό που μπόρεσα να χρησιμοποιήσω σε ότι αφορά την προσομοίωση του αντιδραστήρα.

Με την ολοκλήρωση των σπουδών μου θα ήθελα να ευχαριστήσω τους φίλους μου που ήταν δίπλα μου όλα αυτά τα χρόνια και που με έκαναν να γεμίσω εμπειρίες και όμορφες αναμνήσεις.

Τέλος, ένα μεγάλο ευχαριστώ οφείλω στην οικογένεια μου, τους γονείς μου Μαρία και Γιώργο, και την αδερφή μου Βασιλική για όλη την αγάπη και την στήριξη που μου έχουν δώσει. Γνωρίζω ότι είναι πάντα δίπλα μου, με κάθε τρόπο, και αυτό με κάνει να προχωρώ.

Περιεχόμενα

Περίληψη.....	6
Abstract	8
1. Εισαγωγή	10
2. Το φυσικό σύστημα.....	12
2.1 Διεργασία χημικής απόθεσης από ατμό.....	12
2.2 Περιγραφή του αντιδραστήρα	13
2.3 Μαθηματική διατύπωση του προβλήματος.....	14
2.4 Επίλυση του προβλήματος.....	17
2.5 Πολλαπλότητα χώρου λύσεων.....	18
3. Μείωση Τάξης Μοντέλου.....	21
3.1 Εισαγωγή	21
3.2 Βασική αρχή της Μείωσης Τάξης Δυναμικών Συστημάτων.....	22
3.3 Ανάλυση Κύριων Συνιστωσών.....	23
3.3.1 Εισαγωγή	23
3.3.2 Βασικές εξισώσεις μεθόδου	24
3.3.3 Κατασκευή Βάσης και Προβολής	25
3.3.4 Επιλογή διάστασης βάσης.....	26
3.4 Μέθοδος των στιγμιότυπων	27
3.4.1 Περιγραφή μεθόδου	27
3.4.2 Κανονικοποίηση δεδομένων.....	30
3.5 Αξιολόγηση μείωσης τάξης – Σφάλματα.....	31
3.6 Αλγόριθμος για την κατασκευή της βάσης	32
3.7 Εναλλακτικές προσεγγίσεις ανάλυσης κυρίων συνιστωσών	33
3.7.1. Εισαγωγή	33
3.7.2. Επιλογή παραμέτρων	33
3.7.3. Εφαρμογή της προσέγγισης στο μελετώμενο σύστημα	34
4. Τεχνητά Νευρωνικά Δίκτυα.....	36
4.1. Εισαγωγή	36
4.2. Ο νευρώνας	37
4.2.1. Βιολογικός νευρώνας	37
4.2.2. Μαθηματική Μοντελοποίηση Νευρώνα	39
4.3. Αρχιτεκτονική νευρωνικών δικτύων	42
4.4. Δίκτυα πρόσθιας τροφοδότησης	44
4.5. Δίκτυα με Ανατροφοδότηση - Δίκτυο NARX	45
4.6. Εκπαίδευση νευρωνικού δικτύου	47

4.7.	Εφαρμογή.....	50
4.8.	Αξιολόγηση νευρωνικού δικτύου.....	51
5.	Αποτελέσματα	53
5.1	Εισαγωγή.....	53
5.2	Ρύθμιση παραμέτρων στον υπολογιστικό κώδικα	54
5.3	Διερευνήσεις	56
A)	1 ^η Διερεύνηση.....	56
B)	2η Διερεύνηση.....	65
Γ)	3η Διερεύνηση.....	68
Δ)	4η Διερεύνηση.....	70
6.	Συμπεράσματα & Προτάσεις	82
6.1	Συμπεράσματα	82
6.2	Προτάσεις για περαιτέρω έρευνα.....	83
7.	Βιβλιογραφία.....	86

Περίληψη

Στην παρούσα εργασία μελετάται η δυναμική συμπεριφορά ενός αντιδραστήρα χημικής απόθεσης από ατμό (Chemical Vapor Deposition, CVD). Για τις δεδομένες συνθήκες λειτουργίας του αντιδραστήρα, εμφανίζονται στο υπάρχον σύστημα δύο αντιμαχόμενοι μηχανισμοί ροής, ο μηχανισμός της φυσικής κυκλοφορίας λόγω άνωσης και αυτός της εξαναγκασμένης ροής. Το σύστημα χαρακτηρίζεται συνεπώς από την ύπαρξη πολλαπλότητας στις καταστάσεις που ισορροπεί, για παροχές εισόδου που κυμαίνονται από $2,3 \cdot 10^{-5}$ kg/s έως και $5,0 \cdot 10^{-5}$ kg/s. Οι μη γραμμικές εξισώσεις που εκφράζουν τα φαινόμενα μεταφοράς και τις κινητικές των αντιδράσεων που λαμβάνουν χώρα στον αντιδραστήρα επιλύονται με χρήση του κώδικα υπολογιστικής ρευστοδυναμικής ANSYS Fluent. Το μοντέλο που προκύπτει χαρακτηρίζεται από μεγάλη πολυπλοκότητα και η επίλυση του αποτελεί χρονοβόρα διαδικασία λόγω του μεγάλου αριθμού των εξισώσεων που απαιτούνται για την περιγραφή του. Κύρια ιδέα της εργασίας είναι η ανάπτυξη ενός μοντέλου μειωμένης τάξης (Reduced Order Model, ROM), μέσω μιας εναλλακτικής υλοποίησης της μεθόδου ανάλυσης κυρίων συνιστωσών (Principal Component Analysis, PCA) και εν συνεχεία μέσω της εκπαίδευσης τεχνητών νευρωνικών δικτύων (Artificial Neural Networks, ANN), με στόχο την περιγραφή της δυναμικής συμπεριφοράς του πλήρους συστήματος χρησιμοποιώντας δεδομένα από μη-μόνιμες καταστάσεις. Για τον σκοπό αυτό, λήφθηκαν δεδομένα της δυναμικής απόκρισης των πεδίων ροής και θερμοκρασίας του λεπτομερούς μοντέλου, κατά την χρονική του εξέλιξη (τροχιά) από διάφορες αρχικές καταστάσεις με τιμή μαζικής παροχής εντός του εύρους $1,5 \cdot 10^{-5}$ kg/s – $6,5 \cdot 10^{-5}$ kg/s, στο οποίο εμφανίζεται η πολλαπλότητα των λύσεων, προς την τελική μόνιμη κατάσταση με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s. Έπειτα, πραγματοποιήθηκε ανάλυση ως προς τον αριθμό των αρχικών καταστάσεων που χρησιμοποιούνται και του χρονικού διαστήματος στο οποίο αφήνονται να εξελιχθούν, χωρίς να φτάνουν σε μόνιμη κατάσταση, έτσι ώστε να προσεγγίζεται με ικανοποιητική ακρίβεια η προσομοίωση του αντιδραστήρα. Το καλύτερο μοντέλο μειωμένης τάξης επιλέχθηκε με γνώμονα την υπολογιστική ακρίβεια σε συνδυασμό με τον απαιτούμενο όγκο δεδομένων για τον σχεδιασμό του. Το μοντέλο που επιλέχθηκε είναι αυτό που χρησιμοποιεί την ορθοκανονική βάση η οποία αποτελείται από 4 ιδιοσυναρτήσεις σε συνδυασμό με ένα νευρωνικό δίκτυο που αποτελείται από 10 κρυφούς νευρώνες. Η βάση σχεδιάστηκε με χρήση 23 βηματικών επιβολών, ενώ το νευρωνικό δίκτυο που χρησιμοποιείται εκπαιδεύτηκε με 22 βηματικές επιβολές. Οι αρχικές καταστάσεις επιλέχθηκαν έτσι ώστε να καλύπτουν όλο το φάσμα της πολλαπλότητας των λύσεων, ενώ το χρονικό διάστημα στο οποίο αφήνεται να εξελιχθεί κάθε αρχική κατάσταση, είναι τέτοιο ώστε να επιτευχθεί μια συγκεκριμένη τιμή σφάλματος, 10^{-3} , μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης κατά την τροχιά του συστήματος προς την τελική κατάσταση. Το χρονικό αυτό διάστημα αντιπροσωπεύεται με τον αριθμό snapshots από τον οποίο αποτελείται κάθε επιβολή και έχει ελάχιστη επιτρεπτή τιμή τα 10 snapshots. Με την κατασκευή του μοντέλου μειωμένης τάξης, επιτυγχάνεται η προσέγγιση της δυναμικής του φυσικού συστήματος χρησιμοποιώντας πολύ μικρότερο αριθμό

δεδομένων. Το μοντέλο που σχεδιάστηκε σε αυτήν την εργασία, έχει την δυνατότητα να προσεγγίζει με ικανοποιητική ακρίβεια, μόνιμες καταστάσεις του πλήρους συστήματος, ενώ το ίδιο έχει κατασκευαστεί από δεδομένα μη μόνιμων καταστάσεων. Το υπολογιστικό όφελος από την κατασκευή ενός τέτοιου είδους μοντέλου είναι σημαντικό, καθώς ο όγκος των δεδομένων που απαιτούνται για την ανάπτυξη του μοντέλου και η ταχύτητα επίλυσης του μειώνονται κατά πολύ.

Alternative implementation of Principal Component Analysis in the investigation of the nonlinear behavior of a Chemical Vapor Deposition process

Diploma Thesis

by

Kiriaki Valentina Baili

Abstract

This thesis is about the study of the dynamic behavior of a chemical vapor deposition reactor (CVD). For given operating conditions of the reactor, two competing mechanisms appear, the free convection mechanism due to buoyancy and the forced convection. The system is nonlinear and is characterized by a multiplicity of steady states for inlet mass flow rates between $2.3 \cdot 10^{-5}$ kg/s and $5.0 \cdot 10^{-5}$ kg/s, due to the coexistence of the competing mechanisms. The nonlinear equations describing the phenomena that occur in the reactor are solved using the ANSYS Fluent Computational Fluid Dynamics (CFD) code. Due to the large number of equations, these systems are characterized by great complexity and their solution is a time consuming process. The aim of the thesis is the construction of a Reduced Order Model (ROM), by using an alternative implementation of the Principal Component Analysis (PCA) method combined with properly trained Artificial Neural Networks (ANN), to describe the dynamic behavior of the complete system by using solely data from non-steady states. For this purpose, data of the dynamic response of the flow, temperature and pressure fields were obtained during its time evolution (trajectory) from various initial states with a mass flow rate within the range of $1.5 \cdot 10^{-5}$ kg/s - $6.5 \cdot 10^{-5}$ kg/s, in which the multiplicity of solutions is observed, to the final steady state with a mass flow rate of $3 \cdot 10^{-5}$ kg/s. Then, an analysis was carried out regarding the number of initial states used and the time allowed to evolve without reaching the final steady state, so that the accuracy of simulation is satisfactory. The choice of the best model was based on the quality of the approximation and the size of the required data for its design. The reduced order model that was chosen includes a PCA basis with four eigenvectors combined with a neural network containing ten hidden layers. The design of the PCA basis uses 23 step changes, while the training of the neural network uses 26. The initial conditions were selected in order to describe the entire range of multiplicity, while the transients are allowed to evolve until an error tolerance (10^{-3}) is reached. The error is calculated between two successive transient states. This time interval is represented by the number of snapshots of each step change; the number of snapshots was kept higher than 10. By constructing the reduced order model, the physical system dynamic analysis required a highly reduced number of data. The

model designed in this thesis, is capable of approximating with sufficient accuracy, steady states of the full model, while it was built from data of non-steady states. The computational advantage of constructing such a reduced order model is significant, since the number of the required data and the computational cost are greatly reduced.

1. Εισαγωγή

Η χημική απόθεση από ατμό είναι μια διεργασία ευρύτατα διαδεδομένη στην βιομηχανία για την παραγωγή λεπτών υμενίων. Τα υμένια αυτά, χρησιμοποιούνται σε ποικιλία προϊόντων όπως σε εργαλεία, γυάλινες φιάλες για φάρμακα, ημιαγωγούς, διακοσμητικά αντικείμενα κτλ. Η χημική απόθεση από ατμό παρουσιάζει μεγάλο ενδιαφέρον τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο, καθώς, παρά το γεγονός ότι αποτελεί μια εξαιρετικά πολύπλοκη διεργασία, το τελικό προϊόν είναι μεγάλης αξίας και οι πρώτες ύλες ακριβές, η ρύθμιση της στην βιομηχανία πραγματοποιείται συνήθως εμπειρικά από τον χειριστή. Η πολυπλοκότητα της διεργασίας οφείλεται στα μη γραμμικά φαινόμενα και το σύνθετο δίκτυο χημικών αντιδράσεων που την χαρακτηρίζουν, ενώ έχει διαπιστωθεί ότι για τις ίδιες συνθήκες λειτουργίας του αντιδραστήρα, εμφανίζονται τρεις καταστάσεις ισορροπίας στο σύστημα, δύο ευσταθείς και μία ασταθής.

Συνεπώς η λεπτομερής μελέτη του συστήματος καθίσταται αναγκαία για τη σωστή κατανόησή του, αλλά και την παραγωγή υμενίων επιθυμητών προδιαγραφών. Η μελέτη αυτή, είναι δυνατό να επιτευχθεί με χρήση κάποιων ειδικών, εμπορικών λογισμικών που έχουν αναπτυχθεί τα τελευταία χρόνια, για την ακριβή προσομοίωση τέτοιων φυσικών συστημάτων. Ο τρόπος με τον οποίο λειτουργούν είναι η δημιουργία ενός πλέγματος που διακρίτοποιεί το χώρο, ή και τον χρόνο, και η επίλυση των κατάλληλων εξισώσεων σε κάθε κόμβο του πλέγματος. Ο αριθμός των κόμβων του πλέγματος εξαρτάται από την επιθυμητή ακρίβεια και καθορίζει το κόστος και τον χρόνο επίλυσης. Όσο μεγαλύτερη ακρίβεια επιδιώκεται, τόσο πιο πυκνό πλέγμα απαιτείται, που συνεπάγεται με αύξηση του πλήθους των εξισώσεων που πρέπει να επιλυθούν και συνεπώς και του συνολικού χρόνου επίλυσης.

Στόχος της εργασίας είναι η ανάπτυξη ενός μοντέλου μειωμένης τάξης (Reduced Order Model, ROM), το οποίο θα περιγράφει τη δυναμική συμπεριφορά του πλήρους συστήματος, δηλαδή των δύο ευσταθών κλάδων καθώς και εναλλαγές μεταξύ αυτών, χρησιμοποιώντας για την κατασκευή του δεδομένα του συστήματος από μη μόνιμες καταστάσεις. Η ιδιαιτερότητα της παρούσας εργασίας προκύπτει από το γεγονός ότι επιδιώκεται η κατασκευή ενός μοντέλου μειωμένης τάξης, που να προσεγγίζει ικανοποιητικά μόνιμες καταστάσεις του συστήματος, ενώ το ίδιο χρησιμοποιεί για την κατασκευή του δεδομένα από μη μόνιμες καταστάσεις. Η ιδέα αυτή βασίστηκε στην μελέτη που πραγματοποίησαν οι Michael D. Graham και Ioannis G. Kevrekidis (1995) [3], ως προς κάποιες εναλλακτικές προσεγγίσεις την ανάλυσης κυρίων συνιστωσών η οποία θα αναπτυχθεί συνοπτικά στην Παράγραφο 3.7. Με τον τρόπο αυτό επιδιώκεται να επιτευχθεί η κατασκευή ενός αξιόπιστου μοντέλου, που να μπορεί να προβλέπει την συμπεριφορά του συστήματος με μεγάλη ακρίβεια, αλλά και με μεγάλη ταχύτητα υπολογισμών που είναι απαραίτητη για την επίλυση τέτοιου είδους φυσικών συστημάτων σε πραγματικό χρόνο.

Τα αποτελέσματα που χρησιμοποιούνται λαμβάνονται από ένα μεγάλης τάξης, λεπτομερές μοντέλο, το οποίο επιλύεται με χρήση του λογισμικού υπολογιστικής ρευστοδυναμικής ANSYS Fluent 13 [4]. Η αναζητούμενη λύση παρουσιάζει εξάρτηση

από το χώρο και από το χρόνο. Η ανάπτυξη του μοντέλου μειωμένης τάξης πραγματοποιείται σε δύο στάδια, στα οποία επιδιώκεται να αντιμετωπισθεί η χωρική και η χρονική εξάρτηση αντίστοιχα. Στο πρώτο στάδιο χρησιμοποιείται η ανάλυση κυρίων συνιστωσών, και συγκεκριμένα η μέθοδος των στιγμιοτύπων (Snapshots), με την βοήθεια της οποίας καλύπτεται η χωρική εξάρτηση των μεταβλητών και συνεπώς μειώνεται η τάξη του μοντέλου. Συγκεκριμένα υπολογίζεται ένας αριθμός ιδιοδιανυσμάτων τα οποία περιέχουν μεγάλο ποσοστό πληροφορίας του συστήματος και που με κατάλληλους συντελεστές μπορούν να αποδώσουν ως γραμμικό συνδυασμό μια ικανοποιητική προσέγγιση της λύσης. Στο δεύτερο στάδιο, με χρήση των τεχνητών νευρωνικών δικτύων αντιμετωπίζεται η χρονική εξάρτηση του συστήματος, καθώς υπολογίζεται η χρονική εξέλιξη των συντελεστών αυτών κι εν συνεχεία προσεγγίζεται η δυναμική του.

Σημειώνεται ότι σε όλα τα υπολογιστικά πειράματα που διεξάγονται για την επιλογή του καταλληλότερου μοντέλου μειωμένης τάξης, η κατασκευή της βάσης και η εκπαίδευση του νευρωνικού δικτύου επιτυγχάνονται χρησιμοποιώντας ως μόνα δεδομένα τα αποτελέσματα υπολογιστικών προσομοιώσεων. Δεν απαιτείται καμία γνώση των εξισώσεων που διέπουν το φυσικό σύστημα, παρά μόνο τα αποτελέσματα από τον υπολογιστικό κώδικα. Η επιτυχία της μεθόδου βασίζεται στην επιλογή κατάλληλων προσομοιώσεων, κατάλληλου αριθμού προσομοιώσεων και κατάλληλου χρονικού διαστήματος στο οποίο αφήνονται να εξελιχθούν οι αρχικές καταστάσεις που επιλέγονται, ώστε να ενεργοποιηθούν όλες οι σημαντικές «πτυχές» του συστήματος. Επίσης σημαντικό ρόλο παίζει και η σωστή επιλογή παραμέτρων του νευρωνικού δικτύου.

Η εργασία δομείται ως εξής: Στο Κεφάλαιο 2 περιγράφεται με λεπτομέρεια το φυσικό σύστημα και ο τρόπος λήψης δεδομένων από το λεπτομερές μοντέλο. Στη συνέχεια, αναπτύσσεται το θεωρητικό υπόβαθρο της μεθοδολογίας της κατασκευής του μοντέλου μειωμένης τάξης: Αρχικά αναπτύσσεται η ανάλυση κυρίων συνιστωσών (Κεφάλαιο 3) κι έπειτα, τα τεχνητά νευρωνικά δίκτυα (Κεφάλαιο 4). Κλείνοντας παρουσιάζονται τα αποτελέσματα των υπολογισμών που πραγματοποιήθηκαν (Κεφάλαιο 5), τα συμπεράσματα που προέκυψαν καθώς και προτάσεις για περαιτέρω έρευνα (Κεφάλαιο 6).

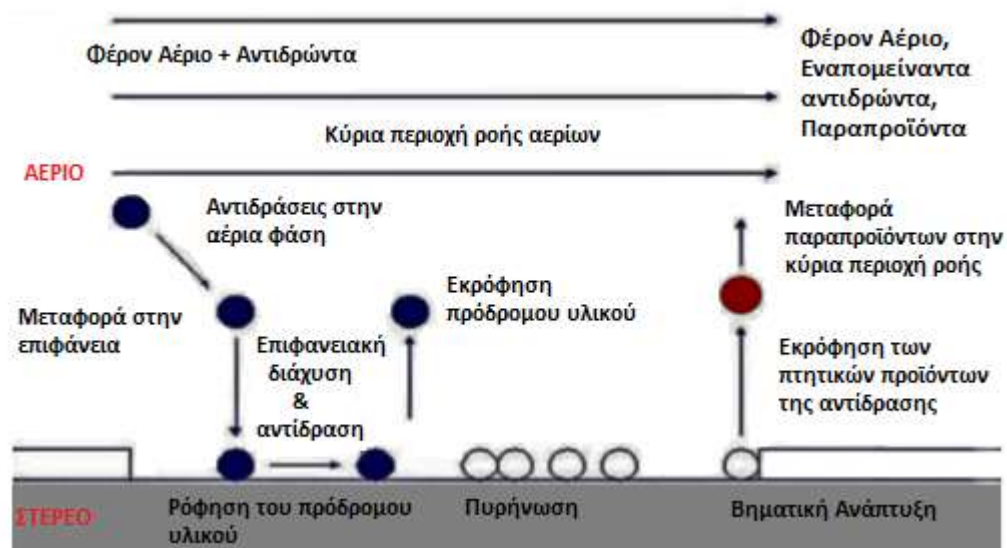
2. Το φυσικό σύστημα

2.1 Διεργασία χημικής απόθεσης από ατμό

Χημική απόθεση από ατμό (ΧΑΑ) (Chemical Vapor Deposition, CVD) ονομάζεται η διεργασία σχηματισμού λεπτών στερεών υμενίων (films) με χρήση αντιδρώντων αερίων. Κατά την διεργασία αυτή λεπτά στερεά υμένια αποτίθενται πάνω σε ειδικές επιφάνειες μέσω χημικών αντιδράσεων. Χάρη στην εύκολη μεταβολή και προσαρμογή των συνθηκών παραγωγής των υμενίων στις εκάστοτε απαιτήσεις της βιομηχανίας, η διεργασία της χημικής απόθεσης από ατμό βρίσκει ευρεία εφαρμογή για την παραγωγή μικρο ή νάνο-ηλεκτρονικών συσκευών. [2]

Η χημική απόθεση από ατμό αποτελεί μια ετερογενή διεργασία που λαμβάνει χώρα στην αέρια και στη στερεά φάση μέσα στον αντιδραστήρα (αντιδραστήρας ΧΑΑ). Τα αντιδρώντα, τα οποία περιέχονται στο πρόδρομο υλικό (precursor), μεταφέρονται μέσω ενός φέροντος αερίου από την είσοδο του αντιδραστήρα σε στερεά δισκία (wafers) που βρίσκονται στο εσωτερικό του. Στα δισκία αυτά, τα οποία στηρίζονται σε ειδικές επιφάνειες (υποστρώματα, substrates), πραγματοποιείται απόθεση των αντιδρώντων και μέσω χημικών αντιδράσεων σχηματίζονται υμένια μετάλλων, ημιαγωγών, μονωτών και επιστρώσεων. Ένα ακόμη φαινόμενο που παρατηρείται κατά την διεργασία αυτή, εκτός από τις αντιδράσεις που λαμβάνουν χώρα στα στερεά δισκία, είναι αντιδράσεις που πραγματοποιούνται στην αέρια φάση στον κύριο όγκο του αντιδραστήρα. Η μελέτη της διεργασίας της χημικής απόθεσης με ατμό είναι ιδιαίτερα σύνθετη καθώς περιλαμβάνει περίπλοκους μηχανισμούς χημικής κινητικής και πραγματοποιούνται ταυτόχρονα φαινόμενα μεταφοράς ορμής, θερμότητας και μάζας.

Στο Σχήμα 2.1 παρουσιάζεται σχηματικά η διεργασία της χημικής απόθεσης από ατμό.



Σχήμα 2. 1: Διεργασία χημικής απόθεσης από ατμό

Ο τύπος του παραγόμενου υμενίου και οι συνθήκες λειτουργίες του αντιδραστήρα παίζουν καθοριστικό ρόλο στην κατηγοριοποίηση των διεργασιών χημικής απόθεσης.

Τα υμένια παράγονται κάτω από συγκεκριμένες συνθήκες λειτουργίας και θα πρέπει να έχουν ελεγχόμενες ιδιότητες όπως και να μπορούν να αναπαράγονται. Το πάχος, η ομοιομορφία, η μορφολογία της επιφάνειας, η σύσταση και η καθαρότητα των υμενίων αποτελούν βασικές ιδιότητες που καθορίζουν την ποιότητα των παραγόμενων υμενίων. Οι παραπάνω ιδιότητες έχουν ένα εύρος αποδεκτών ορίων το οποίο εξαρτάται από το είδος της διεργασίας, το υλικό καθώς και τη χρήση για την οποία προορίζεται το υμένιο. [1]

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 φωτονικά 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 Si_4-xCl_x ή από SiH_4 , GaAs από $Ga(CH_3)_3$, InP από $In(CH_3)_3$ και PH_3 κ.α. [2]

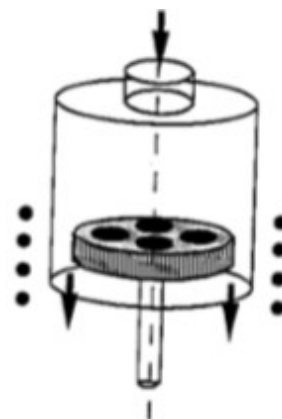
Οι αντιδραστήρες Χημικής Απόθεσης από Ατμό απαρτίζονται από τέσσερα κύρια μέρη: το σύστημα εισόδου, τον κύριο αντιδραστήρα, την πηγή θέρμανσης και το σύστημα εξαγωγής αερίων. Ανάλογα με τις εκάστοτε ιδιότητες των παραγόμενων υμενίων, έχουν αναπτυχθεί πολλά είδη αντιδραστήρων, οι οποίοι διαφοροποιούνται ως προς το μέγεθος, το σχήμα, τον 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]

2.2 Περιγραφή του αντιδραστήρα

Ο αντιδραστήρας που χρησιμοποιείται ανήκει στην κατηγορία αντιδραστήρων που χρησιμοποιούνται ευρέως στην έρευνα των διεργασιών χημικής απόθεσης από ατμό και στην παραγωγή σύνθετων ημιαγωγών και μικροηλεκτρονικών με βάση το πυρίτιο.

Πρόκειται για έναν κυλινδρικό, κατακόρυφο αντιδραστήρα, που παρουσιάζει αξονική συμμετρία και έχει ένα θερμαινόμενο υπόστρωμα (στους 700 K) και ψυχόμενα τοιχώματα (στους 300 K). Στο σχήμα 2.2 παρουσιάζεται μια τρι-διάστατη απεικόνιση ενός τέτοιου αντιδραστήρα ΧΑΑ. [2]

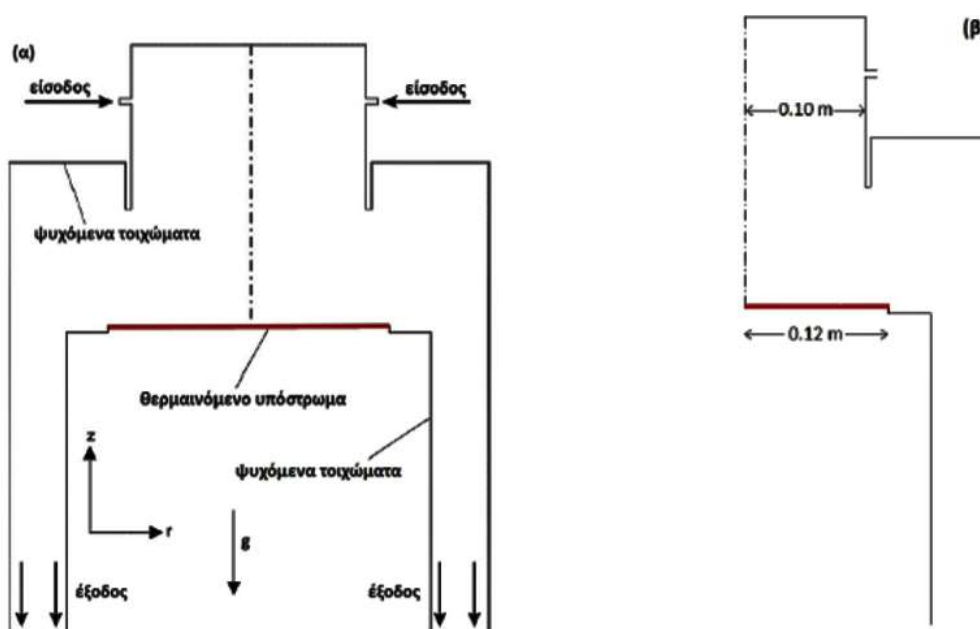
Κατά την χημική απόθεση από ατμό, το αέριο μίγμα εισέρχεται στον αντιδραστήρα από την πάνω πλευρά του σωλήνα έγχυσης, διαμέτρου 0,20m, και στη συνέχεια οδηγείται στο υπόστρωμα, διαμέτρου 0,24 m. [5] Εκεί, αποτίθενται τα στερεά προϊόντα της αντίδρασης και



Σχήμα 2. 2: Τρι-διάστατη απεικόνιση αντιδραστήρα ΧΑΑ

σχηματίζεται το υμένιο, όπως παρουσιάστηκε παραπάνω, ενώ τα παραπροϊόντα και τα αδρανή οδηγούνται στην δακτυλιοειδή έξοδο του αντιδραστήρα. Για να αποφευχθούν κάποιες από τις μη επιθυμητές αντιδράσεις της διεργασίας, τα τοιχώματα του αντιδραστήρα ψύχονται έτσι ώστε να μην ενεργοποιούνται αντιδράσεις απόθεσης και συνεπώς να μην αποτίθεται προϊόν στα τοιχώματα του αντιδραστήρα. [2][5]

Όπως έχει αναφερθεί στην παράγραφο 2.1 η χημική απόθεση από ατμό είναι μια ιδιαίτερα περίπλοκη διεργασία και η μελέτη της μέσω προσομοιώσεων υπολογιστικής μηχανικής απαιτεί πολύ χρόνο. Στα πλαίσια της παρούσας διπλωματικής εργασίας χρησιμοποιείται μια δι-διάστατη αξονοσυμμετρική προσομοίωση του αντιδραστήρα. Στο Σχήμα 2.3 απεικονίζεται η τομή του αντιδραστήρα ΧΑΑ, η γεωμετρία που χρησιμοποιήθηκε για την μελέτη του συστήματος και οι παραπάνω αναφερόμενες διαστάσεις του.



Σχήμα 2. 3 (α) Τομή του αντιδραστήρα ΧΑΑ παράλληλη στην ροή, (β) Γεωμετρία του αντιδραστήρα που χρησιμοποιήθηκε με σημειωμένες τις διαστάσεις του. Με την περιστροφή του σχήματος γύρω από τον άξονα συμμετρίας προκύπτει ο αντιδραστήρας.

2.3 Μαθηματική διατύπωση του προβλήματος

Στο δεδομένο πρόβλημα, μελετάται η ταυτόχρονη ροή του αδρανούς αερίου N_2 και η μεταφορά ενέργειας στον παραπάνω αντιδραστήρα. Η ποσότητα των αντιδρώντων που εισέρχονται στον αντιδραστήρα μέσω του N_2 , το οποίο δρα ως φέρον αέριο, είναι πολύ μικρή. Παράλληλα, τα παρόμοια μοριακά βάρη όλων των συστατικών του αερίου μίγματος, κάνουν εφικτή την απλοποίηση του μελετώμενου συστήματος. Ειδικότερα για την μαθηματική διατύπωση του προβλήματος δεν θα ληφθούν υπόψη οι χημικές αντιδράσεις που πραγματοποιούνται στην αέρια φάση και στην στερεή επιφάνεια. [1]

Η πίεση λειτουργίας του αντιδραστήρα διατηρείται σταθερή στα 1300 Pa και έτσι μπορεί να θεωρηθεί ότι τα αέρια στο εσωτερικό του συμπεριφέρονται ως ιδανικά καθώς και ότι η ροή είναι στρωτή. [1][5]

Με δεδομένα τα παραπάνω, παρουσιάζονται οι εξισώσεις που χρησιμοποιήθηκαν για την μαθηματική διατύπωση του προβλήματος σε μεταβατική κατάσταση. Οι εξισώσεις αυτές περιλαμβάνουν τους μηχανισμούς των φαινομένων μεταφοράς μάζας, ορμής και ενέργειας, έχουν διανυσματική μορφή και ισχύουν σε όλο τον όγκο του αντιδραστήρα. [1][2].

i. Εξίσωση συνέχειας:

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

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

ii. Εξίσωση ορμής:

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

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

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

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

iii. Εξίσωση ενέργειας:

$$\frac{\partial (\rho e)}{\partial t} + \nabla \cdot (\mathbf{v}(\rho e + p)) = \nabla \cdot (k \nabla T + \boldsymbol{\tau} \cdot \mathbf{v}) \quad (2.4)$$

Όπου k η θερμική αγωγιμότητα και T η θερμοκρασία. Η ειδική εσωτερική ενέργεια, e , ορίζεται ως εξής:

$$e = h - \frac{p}{\rho} + \frac{\mathbf{v} \cdot \mathbf{v}}{2} \quad (2.5)$$

Όπου h η ειδική ενθαλπία που ορίζεται ως εξής:

$$h = \int_{T_{ref}}^T c_p dT \quad (2.6)$$

Όπου, c_p η ειδική θερμοχωρητικότητα υπό σταθερή πίεση και θερμοκρασία αναφοράς, $T_{ref} = 298.15$ K.

iv. Εξίσωση ιδανικών αερίων:

$$\rho = \frac{pM}{RT} \quad (2.7)$$

Όπου M το μοριακό βάρος του αζώτου και R η παγκόσμια σταθερά των ιδανικών αερίων

Με την ταυτόχρονη επίλυση των παραπάνω εξισώσεων προκύπτει το πεδίο ροής και η κατανομή της θερμοκρασίας και της πίεσης στο εσωτερικό του αντιδραστήρα ΧΑΑ. Για την επίλυση τους απαιτούνται οι εξής συνοριακές συνθήκες: [2][5]

- i. Η θερμοκρασία εισόδου του αερίου στον αντιδραστήρα είναι σταθερή και ίση με 298K. Η ταχύτητα εισόδου του αερίου στον αντιδραστήρα είναι ομοιόμορφη, κάθετη στην είσοδο του αντιδραστήρα και υπολογίζεται από τον τύπο:

$$\mathbf{n} \cdot \mathbf{v} = \frac{\dot{m}(t)}{2\pi L_{in}} \quad \text{και} \quad \mathbf{n} \times \mathbf{v} = 0 \quad (2.8)$$

Όπου \mathbf{n} το μοναδιαίο κάθετο διάνυσμα, \dot{m} η μαζική παροχή και L_{in} το πάχος του δακτυλιοειδούς στομίου εισόδου.

Η εξίσωση (2.8) συνδέει την ταχύτητα με την μαζική παροχή η οποία μεταβάλλεται με τον χρόνο. Η μαζική παροχή κυμαίνεται μεταξύ των τιμών $1,5 \cdot 10^{-5}$ kg/s και $6,5 \cdot 10^{-5}$ 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:

$$\mathbf{n} \cdot (\nabla \rho \mathbf{v}) = 0 \quad \text{και} \quad \mathbf{n} \times \mathbf{v} = \mathbf{0} \quad (2.9)$$

$$\mathbf{n} \cdot \nabla T = 0 \quad (2.10)$$

- iii. Στα τοιχώματα του αντιδραστήρα ισχύει η συνθήκη της μη-ολίσθησης για την ταχύτητα:

$$\mathbf{v} = \mathbf{0} \quad (2.11)$$

και η θερμοκρασία διατηρείται σταθερή και ίση με 700 K για το θερμαινόμενο υπόστρωμα και με 300 K στα ψυχόμενα τοιχώματα του αντιδραστήρα.

- iv. Στον άξονα συμμετρίας, η ταχύτητα έχει μηδενική συνιστώσα κάθετη στον άξονα και η βαθμίδα κάθε μεταβλητής είναι μηδενική στη διεύθυνση κάθετη στον άξονα.

$$\mathbf{n} \cdot (\nabla \rho \mathbf{v}) = 0 \quad \text{και} \quad \mathbf{n} \cdot \mathbf{v} = 0 \quad (2.12)$$

$$\mathbf{n} \cdot \nabla T = 0 \quad (2.13)$$

Υπενθυμίζεται ότι η παροχή $\dot{m}(t)$ είναι η μοναδική από τις συνοριακές συνθήκες που μεταβάλλεται με το χρόνο. Λόγω της φύσης του προβλήματος και των περιορισμών που τίθενται λόγω των χημικών αντιδράσεων οι υπόλοιπες μεταβλητές εισόδου, δηλαδή η θερμοκρασία του υποστρώματος και η θερμοκρασία των τοιχωμάτων του αντιδραστήρα παραμένουν σταθερές.

Η ύπαρξη αρχικών συνθηκών είναι απαραίτητη για την επίλυση του προβλήματος. Οι γνωστές κατανομές στο χώρο των πεδίων $P(\mathbf{r},0)$, $\mathbf{v}(\mathbf{r},0)$ και $T(\mathbf{r},0)$, στο χρόνο $t=0$, αποτελούν τις αρχικές συνθήκες και έχουν προκύψει από επίλυση του προβλήματος

στη μόνιμη κατάσταση για κάποια τιμή της παροχής εισόδου. Οι κατανομές στο χώρο και στο χρόνο των πεδίων $P(\mathbf{r},t)$, $\mathbf{v}(\mathbf{r},t)$ και $T(\mathbf{r},t)$, αποτελούν την λύση του προβλήματος.

2.4 Επίλυση του προβλήματος

Το πρόβλημα που αναπτύχθηκε στην προηγούμενη ενότητα, επιλύεται σε κυλινδρικές συντεταγμένες με χρήση του εμπορικού κώδικα υπολογιστικής ρευστοδυναμικής ANSYS FLUENT (στο εξής Fluent), με χρήση της μεθόδου των πεπερασμένων όγκων [4]. Επισημαίνεται ότι, ο Fluent διαθέτει βιβλιοθήκες και βάσεις δεδομένων οι οποίες χρησιμοποιούνται για τον υπολογισμό του ιξώδους και της ειδικής θερμοχωρητικότητας υπό σταθερή πίεση του αζώτου. Η προσομοίωση του αντιδραστήρα που χρησιμοποιείται για την μελέτη του συστήματος Σχήμα 2.3(β) διαμερίζεται σε 15066 κελιά. Σε καθένα από αυτά υπολογίζεται κάθε χρονική στιγμή η πίεση, η θερμοκρασία και το διάνυσμα της ταχύτητας (αξονική και ακτινική συνιστώσα), δημιουργώντας έτσι ένα διάνυσμα κατάστασης $\mathbf{x}(t) \in \mathbb{R}^N$, με $N=60264$. [2] [6]

Κατά τη διάρκεια της δυναμικής μελέτης του αντιδραστήρα ΧΑΑ, ο Fluent δεν αφήνεται να συγκλίνει για κάθε χρονικό βήμα Δt , στη λύση της μεταβατικής κατάστασης, αλλά αντίθετα καλείται επανειλημμένα εξωτερικά από το περιβάλλον Matlab. [7] Για κάθε επανάληψη k , εισάγονται ως αρχικές συνθήκες τα στοιχεία του διανύσματος κατάστασης της $\mathbf{x}(t_{k-1})$ με παροχή εισόδου $\dot{m}(t_k)$ και πραγματοποιούνται 50 βήματα της επαναληπτικής μεθόδου. Η χρονική διάρκεια κάθε επαναληπτικού βήματος θα πρέπει να είναι αρκετή 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} . Ο χρόνος διακριτοποιείται σε χρονικά βήματα διάρκειας $\Delta t=0,1$ s και ισχύουν οι σχέσεις:

$$t_0 = 0 \text{ s} \quad (2.14)$$

$$t_k = t_{k-1} + \Delta t, \quad k = 1, 2, \dots \quad (2.15)$$

Η παροχή $\dot{m}(t)$ αποτελεί συνεπώς την μεταβλητή εισόδου του συστήματος, ενώ το διάνυσμα κατάστασης $\mathbf{x}(t)$ που υπολογίζεται, αποτελεί την μεταβλητή εξόδου.

Στην παρούσα εργασία μελετάται η δυναμική απόκριση των πεδίων ροής και θερμοκρασίας του συστήματος για διάφορες βηματικές επιβολές στην παροχή. Μελετάται συνεπώς η συμπεριφορά του συστήματος κατά την μετάβαση του από μια αρχική μόνιμη κατάσταση, με δεδομένη τιμή μαζικής παροχής, σε μια νέα μόνιμη κατάσταση με διαφορετική τιμή μαζικής παροχής. Η μετάβαση αυτή οφείλεται στην επιβολή μιας νέας τιμής μαζικής παροχής στο σύστημα και αλλάζει για κάθε περίπτωση. Αρχικά, υπολογίζεται η δυναμική απόκριση του, όταν αφήνεται να εξελιχθεί μέχρι την νέα μόνιμη κατάσταση, ενώ, έπειτα μελετάται η απόκριση του συστήματος κατά την εξέλιξη του για ένα μικρό αριθμό χρονικών στιγμών κοντά στην αρχική του κατάσταση.

Σε μια επιβολή με m_i χρονικά βήματα υπολογίζονται διαδοχικά m_i διανύσματα κατάστασης $\mathbf{x}(t)$, καθένα από τα οποία αποτελεί ένα στιγμιότυπο (snapshot) του συστήματος, και όλα μαζί απαρτίζουν την τροχιά (trajectory) του συστήματος στο χρόνο. Με αυτόν τον τρόπο δημιουργείται ένας πίνακας στιγμιότυπων, ο οποίος καταγράφει τη δυναμική συμπεριφορά του συστήματος κατά την διάρκεια των χρονικών βημάτων m_i . Ο πίνακας αυτός αποτελεί τον πίνακα τροχιάς της συγκεκριμένης επιβολής και η μορφή του είναι η εξής:

$$Y_i = [\mathbf{x}(t_0)\mathbf{x}(t_1) \dots \mathbf{x}(t_{m_i-1})] \in \mathbb{N}^{N \times m_i} \quad (2.16)$$

Στην περίπτωση που πραγματοποιείται βηματική επιβολή στην παροχή:

$$\dot{m}(t_k) = \begin{cases} \dot{m}_0 & k = 0 \\ \dot{m}_f & k > 0 \end{cases} \quad (2.17)$$

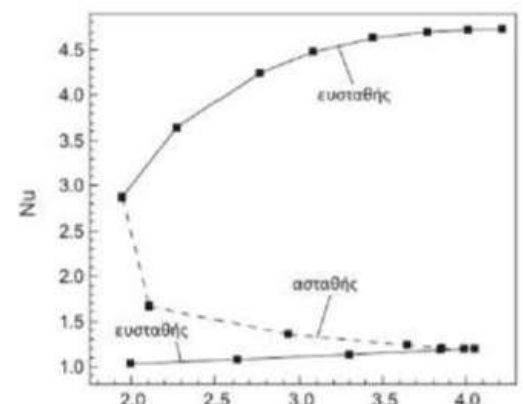
και η επίλυση του συστήματος γίνεται μέχρι τελικής σύγκλισης σε μόνιμη κατάσταση, η σύγκλιση αυτή ελέγχεται εξωτερικά του Fluent, μέχρις ότου η νόρμα της διαφοράς δύο διαδοχικών διανυσμάτων κατάστασης να είναι μικρότερη από ένα συγκεκριμένο όριο ανοχής (tolerance): [6]

$$\|\mathbf{x}(t_k) - \mathbf{x}(t_{k-1})\| < tol \quad (2.18)$$

Μετά από δοκιμές που πραγματοποιήθηκαν, η ανοχή που επιλέχθηκε και στην οποία θεωρείται ότι το σύστημα συγκλίνει στη μόνιμη κατάσταση, είναι $tol = 10^{-5}$.

2.5 Πολλαπλότητα χώρου λύσεων

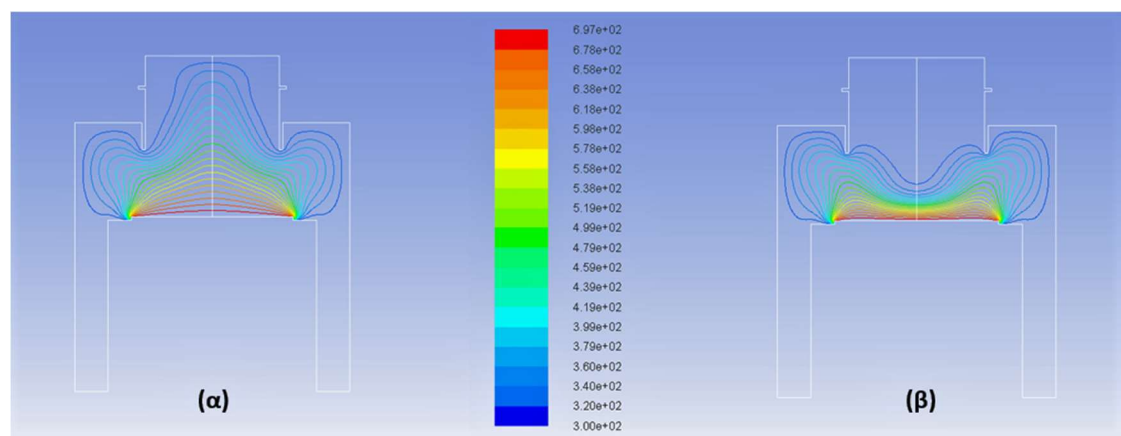
Χαρακτηριστικό του συγκεκριμένου συστήματος είναι η πολλαπλότητα λύσεων που παρουσιάζει στις καταστάσεις ισορροπίας, γεγονός που οφείλεται στους μη γραμμικούς όρους των εξισώσεων (2.2) και (2.4). Πιο συγκεκριμένα, για κάθε τιμή της μαζικής παροχής $\dot{m}(t)$ που κυμαίνεται από $2,3 \cdot 10^{-5}$ kg/s έως $5,0 \cdot 10^{-5}$ kg/s, υπάρχουν τρεις μόνιμες καταστάσεις, δύο ευσταθείς και μία ασταθής, στις οποίες μπορεί να συγκλίνει το σύστημα σε μια βηματική επιβολή. Η κατάσταση αυτή εξαρτάται από την αρχική κατάσταση ισορροπίας στην οποία βρίσκεται το σύστημα την εκάστοτε χρονική στιγμή. Η γραφική αναπαράσταση του συνόλου των λύσεων της μόνιμης κατάστασης αποτελείται από τρεις κλάδους όπως φαίνεται στο Σχήμα 2.4, όπου η παροχή έχει αντικατασταθεί με έναν αντίστοιχο αριθμό Reynolds και η μόνιμη κατάσταση του συστήματος χαρακτηρίζεται από τον αριθμό Nusselt. [1] [2]



Σχήμα 2. 4: Γραφική αναπαράσταση της πολλαπλότητας λύσεων του μελετώμενου συστήματος

Η παρουσία τριών μόνιμων καταστάσεων ισορροπίας κατά την μαθηματική επίλυση του συστήματος, μπορεί να εξηγηθεί φυσικά λόγω των δύο αντιμαχόμενων μηχανισμών που υπάρχουν στο σύστημα. Συγκεκριμένα, η φυσική κυκλοφορία κατά την οποία κυριαρχεί η άνωση λόγω της βαθμίδας θερμοκρασίας μεταξύ του θερμαινόμενου υποστρώματος και του κύριου όγκου του αντιδραστήρα, ανταγωνίζεται την εξαναγκασμένη κυκλοφορία όπου κυριαρχεί η ορμή του

εισερχόμενου ρευστού. Η φυσική κυκλοφορία παρατηρείται στον κάτω ευσταθή κλάδο, για χαμηλές τιμές μαζικής παροχής (Σχήμα 2.5(α)), ενώ η εξαναγκασμένη κυκλοφορία παρατηρείται στον άνω ευσταθή κλάδο, για μεγάλες τιμές μαζικής παροχής. (Σχήμα 2.5(β)). Για κάθε παροχή $\dot{m}(t) \leq 2,3 \cdot 10^{-5}$ kg/s υπάρχει μια μοναδική μόνιμη κατάσταση στην οποία υπερισχύει ο μηχανισμός της φυσικής κυκλοφορίας, ενώ για κάθε παροχή $\dot{m}(t) \geq 5,0 \cdot 10^{-5}$ kg/s υπάρχει και πάλι μια μοναδική μόνιμη κατάσταση στην οποία υπερισχύει ο μηχανισμός της εξαναγκασμένης κυκλοφορίας. Συνεπώς, ο κάτω ευσταθής κλάδος εκτείνεται για κάθε $\dot{m}(t) \leq 5,0 \cdot 10^{-5}$ kg/s, ενώ ο άνω κλάδος εκτείνεται για κάθε $\dot{m}(t) \geq 2,3 \cdot 10^{-5}$ kg/s. [1] [2]



Σχήμα 2. 5: Κατανομή θερμοκρασίας στον αντιδραστήρα για $\dot{m} = 3.75 \cdot 10^{-5}$ kg/s (α) στον κάτω ευσταθή κλάδο και (β) στον άνω ευσταθή κλάδο

Στο σχήμα 2.5 παρουσιάζεται η κατανομή της θερμοκρασίας στον αντιδραστήρα για την ίδια τιμή μαζικής παροχής εντός της περιοχής όπου εμφανίζεται η πολλαπλότητα των λύσεων. Αν στο σύστημα επιβληθεί μια βηματική επιβολή, που ξεκινά από μια μόνιμη κατάσταση για μια παροχή που ανήκει στον κάτω κλάδο και είναι εκτός της περιοχής που υπάρχει πολλαπλότητα λύσεων, και καταλήγει σε μια παροχή εντός της περιοχής πολλαπλότητας, ο Fluent θα συγκλίνει στη λύση η οποία βρίσκεται πιο κοντά, δηλαδή στη μόνιμη κατάσταση στον κάτω κλάδο. Το αντίστοιχο θα συμβεί αν στην βηματική επιβολή που θα επιβληθεί η αρχική παροχή βρίσκεται στον άνω ευσταθή κλάδο, εκτός του χώρου πολλαπλότητας λύσεων και η τελική τιμή παροχής βρίσκεται εντός της περιοχής πολλαπλότητας, το σύστημα θα συγκλίνει στη μόνιμη κατάσταση στον άνω κλάδο. Επίσης, αν η βηματική επιβολή ξεκινήσει από μια τυχαία τιμή εντός του εύρους πολλαπλότητας λύσεων, ο Fluent θα συγκλίνει είτε πάνω είτε κάτω.

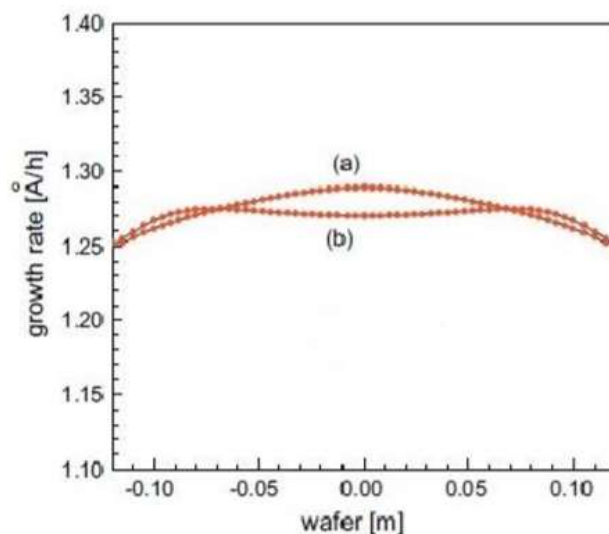
Τέλος, όταν το σύστημα προσπαθεί να συγκλίνει σε τιμή παροχής στην οποία πραγματοποιείται στροφή από τον έναν ευσταθή κλάδο στον άλλον, δηλαδή όσο πλησιάζει την περιοχή αστάθειας, ο χρόνος σύγκλισης αυξάνει σημαντικά. [1] [2] Ειδικότερα, αν το σύστημα ξεκινήσει από μια αρχική μόνιμη κατάσταση στον άνω κλάδο και καταλήγει σε μια τιμή παροχής $\dot{m} \geq 2,3 \cdot 10^{-5}$ kg/s, ο αριθμός των επαναλήψεων που εκτελεί το σύστημα για να συγκλίνει και άρα και ο χρόνος σύγκλισης αυξάνεται δραματικά, καθώς το σύστημα προσπαθεί να συγκλίνει σε

σημείο που παρατηρείται στροφή από τον άνω ευσταθή κλάδο στον ασταθή. Ομοίως, αν η αρχική παροχή βρίσκεται στον κάτω ευσταθή κλάδο και $\dot{m} \leq 5,0 \cdot 10^{-5}$ kg/s.

Σε αντίθεση με τα πεδία ροής και θερμοκρασίας που δεν επηρεάζονται από την ύπαρξη άλλων συστατικών πέραν του Αζώτου, την διάχυση αυτών και τις μεταξύ τους αντιδράσεις, όταν αυτά βρίσκονται σε ίχνη μέσα στο μίγμα, ο ρυθμός απόθεσης καθώς και η ομοιομορφία του σχηματιζόμενου υμενίου εξαρτώνται σημαντικά από το μηχανισμό ροής, δηλαδή τον κλάδο στον οποίο βρίσκεται η λύση. [2]

Στο Σχήμα 2.5 παρουσιάζονται οι διαφορές στην ομοιομορφία του παραγόμενου υμενίου για τους δύο αντιμαχόμενους μηχανισμούς του συστήματος. Λόγω της πολλαπλότητας των λύσεων, η ομοιομορφία του παραγόμενου υμενίου και ο ρυθμός απόθεσης δεν μπορούν να καθοριστούν πλήρως μέσω της παροχής. Όταν ο αντιδραστήρας λειτουργεί κοντά στο σημείο στροφής του κλάδου, αυτές οι ιδιότητες είναι ακόμα πιο δύσκολο να ελεγχθούν σε μικρές αλλαγές της παροχής. Τέλος, τα υμένια τα οποία παράγονται σε συνθήκες εξαναγκασμένης κυκλοφορίας χαρακτηρίζονται από μεγαλύτερη ομοιομορφία. [2] [1]

Από την παραπάνω περιγραφή του φαινομένου, συμπεραίνεται ότι για την παραγωγή υμενίων που να ικανοποιούν τις προδιαγραφές της εκάστοτε βιομηχανίας, είναι σημαντικό το σύστημα ρύθμισης του αντιδραστήρα να λαμβάνει υπόψη την πλήρη δυναμική συμπεριφορά του συστήματος, την ύπαρξη των αντιμαχόμενων μηχανισμών, τις συνέπειες αυτών στο προϊόν, καθώς και την συμπεριφορά του συστήματος στα σημεία στροφής όπου μια μικρή αλλαγή στη παροχή μπορεί να αλλάξει ριζικά όλα τα χαρακτηριστικά της ροής του προϊόντος.



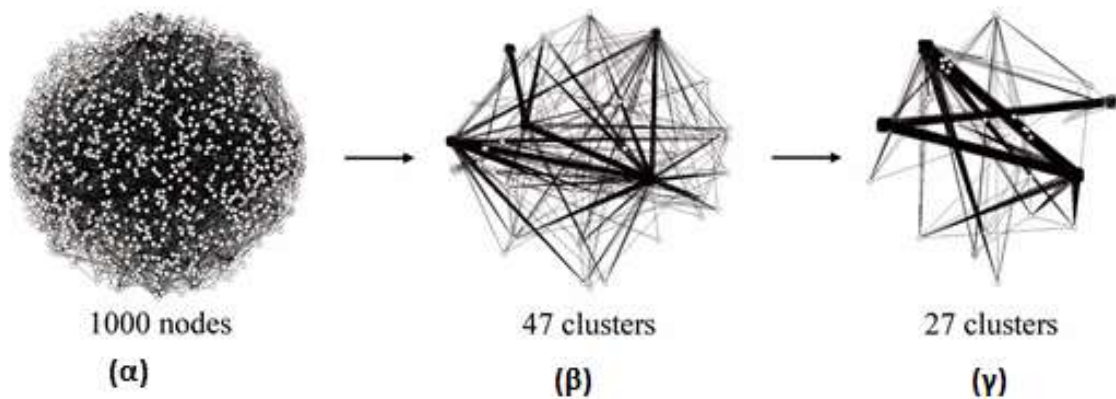
Σχήμα 2. 6: Ο ρυθμός απόθεσης του σχηματιζόμενου υμενίου συναρτήσει της θέσης στο υπόστρωμα στον (α) κάτω ευσταθή κλάδο και (β) στον άνω ευσταθή κλάδο

3. Μείωση Τάξης Μοντέλου

3.1 Εισαγωγή

Τα πολύπλοκα συστήματα υψηλής τάξης προέρχονται συνήθως από υπολογιστική προσομοίωση πραγματικών συστημάτων. Οι υπολογιστικές προσομοιώσεις αποτελούν πλέον τον τρίτο τρόπο μελέτης συστημάτων και διεργασιών, μαζί με την θεωρία και τα πειράματα. Η πολυπλοκότητα των δυναμικών μοντέλων υψηλής τάξης έγκειται στο γεγονός ότι για την περιγραφή τους απαιτείται μεγάλος αριθμός εξισώσεων και μεταβλητών (10^4 - 10^9). Η μείωση τάξης μοντέλου (Model Order Reduction, MOR) αποτελεί μια μέθοδο που αναπτύχθηκε στο πλαίσιο της θεωρίας συστημάτων και ρύθμισης με στόχο την απλοποίηση αυτών των δυναμικών μοντέλων, χωρίς όμως να αλλοιώνονται οι βασικές ιδιότητες του πλήρους τάξης μοντέλου (Full Order Model), ως προς τις μεταβλητές εισόδου και εξόδου.

Το μοντέλο μειωμένης τάξης (Reduced Order Model, ROM) που προκύπτει, περιγράφεται από έναν σημαντικά μικρότερο αριθμό εξισώσεων και μεταβλητών και συνεπώς μπορεί να πραγματοποιεί προσομοιώσεις του πλήρους συστήματος σε πολύ μικρότερο χρονικό διάστημα. Συνεπώς μειώνεται και η υπολογιστική μνήμη που απαιτείται για τις προσομοιώσεις, ενώ το σύστημα δεν παύει να περιγράφεται με ικανοποιητική ακρίβεια. Στο Σχήμα 3.1 παρουσιάζεται η βασική ιδέα της μείωσης τάξης μοντέλου, δηλαδή την μικρή απαίτηση σε πληροφορία, για την περιγραφή των διασυνδέσεων ενός συστήματος. [8]



Σχήμα 3. 1: Γραφική αναπαράσταση της Μείωσης Τάξης Μοντέλου με στόχο τη διατήρηση των διασυνδέσεων των συμπλεγμάτων. (α) Αρχικό μοντέλο αποτελούμενο από 1000 κόμβους (β) Μειωμένης τάξης μοντέλο αποτελούμενο από σύμπλεγμα 47 κόμβων (γ) Μειωμένης τάξης μοντέλο αποτελούμενο από σύμπλεγμα 27 κόμβων

Χωρίς την εφαρμογή της μεθόδου της μείωσης τάξης μοντέλων η μελέτη, η ανάλυση και η προσομοίωση πολύπλοκων συστημάτων είναι πολύ δύσκολη έως και αδύνατη, λόγω του μεγάλου πλήθους εξισώσεων που τα χαρακτηρίζουν. Τα μοντέλα μειωμένης τάξης που προκύπτουν από την εφαρμογή της μεθόδου, διατηρούν τα κύρια χαρακτηριστικά του συστήματος και βοηθούν σημαντικά στην κατανόηση και στη ρύθμιση των διεργασιών.

Τα πολύπλοκα δυναμικά μοντέλα που περιγράφουν σύνθετες διεργασίες μηχανικής μπορούν πλέον να λυθούν γρήγορα, χωρίς απλοποιήσεις και με μεγάλη ακρίβεια. Για τον λόγο αυτό είναι εύλογος ο ισχυρισμός ότι λόγω της σύγχρονης τεχνολογίας δεν υπάρχει ανάγκη για μείωση της τάξης των μοντέλων. Παρ' όλα αυτά η επίλυση δυναμικών μοντέλων πλήρους τάξης είναι μια χρονοβόρα διαδικασία και απαιτεί μεγάλη υπολογιστική μνήμη. Όσο αυξάνεται η ζητούμενη ακρίβεια επίλυσης αυξάνονται και ο χρόνος και ο υπολογιστικός χώρος που απαιτούνται, συνεπώς η επίλυση πλήρων μοντέλων σύνθετων διεργασιών σε πραγματικό χρόνο δεν ενδείκνυται για την αυτόματη ρύθμιση ενός συστήματος. Ο μεγάλος όγκος αποτελεσμάτων που προκύπτουν από την επίλυση των παραπάνω μοντέλων συντελούν στην πιο σφαιρική κατανόηση του προβλήματος, όμως μοντέλα που προσεγγίζουν την κατάσταση του συστήματος σε μικρό χρονικό διάστημα και με μικρή απαίτηση σε υπολογιστική μνήμη, κρίνονται απαραίτητα για την ανάλυση και τον έλεγχο τέτοιων συστημάτων. [8][3]

3.2 Βασική αρχή της Μείωσης Τάξης Δυναμικών Συστημάτων

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]

$$\frac{dx}{dt} = \mathbf{f}(x, \mathbf{u}) \quad \mathbf{f}: \mathbb{R}^{n+n_u} \rightarrow \mathbb{R}^n \quad (3.1)$$

$$\mathbf{y} = \mathbf{g}(x, \mathbf{u}) \quad \mathbf{g}: \mathbb{R}^{n+n_u} \rightarrow \mathbb{R}^{n_y} \quad (3.2)$$

Όπου x η μεταβλητή κατάστασης διάστασης n , \mathbf{u} η είσοδος του συστήματος διάστασης n_u και \mathbf{y} η έξοδος διάστασης n_y . Ο τελεστής \mathbf{f} μπορεί να περιλαμβάνει γραμμικούς και μη γραμμικούς όρους που στην περίπτωση των διεργασιών ΧΑΑ αφορούν τη διάχυση, τη συναγωγή και την αντίδραση και ο τελεστής \mathbf{g} περιγράφει την σχέση μεταξύ της εισόδου με την έξοδο του συστήματος. Η διάσταση n του διανύσματος x , χαρακτηρίζει το πλήθος των μεταβλητών κατάστασης και συνεπώς την πολυπλοκότητα του συστήματος [8]. Η μείωση της τάξης ενός μοντέλου αφορά τη μείωση της διάστασης n , δηλαδή την εύρεση ενός νέου δυναμικού συστήματος που ορίζεται ως εξής:

$$\frac{d\hat{x}}{dt} = \hat{\mathbf{f}}(\hat{x}, \mathbf{u}) \quad \hat{\mathbf{f}}: \mathbb{R}^{d+n_u} \rightarrow \mathbb{R}^d \quad (3.3)$$

$$\mathbf{y} = \hat{\mathbf{g}}(\hat{x}, \mathbf{u}) \quad \hat{\mathbf{g}}: \mathbb{R}^{d+n_u} \rightarrow \mathbb{R}^{n_y} \quad (3.4)$$

Όπου η διάσταση d του νέου διανύσματος κατάστασης \hat{x} είναι πολύ μικρότερη του n και $\hat{\mathbf{f}}, \hat{\mathbf{g}}$ οι προσεγγίσεις των συναρτήσεων \mathbf{f}, \mathbf{g} αντίστοιχα. Για την εύρεση μιας καλής προσέγγισης του αρχικού δυναμικού συστήματος, θα πρέπει το σφάλμα της προσέγγισης να είναι μικρό, να διατηρούνται οι ιδιότητες του αρχικού συστήματος, όπως η ευστάθεια, και η συνολική διαδικασία μείωσης να είναι αποτελεσματική.

Οι δύο εξισώσεις που ακολουθούν αποτελούν το απλοποιημένο σύστημα εξισώσεων που θα μελετηθεί στην παρούσα εργασία:

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

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

όπου \dot{m} είναι η μαζική παροχή στην είσοδο του αντιδραστήρα.

3.3 Ανάλυση Κύριων Συνιστωσών

3.3.1 Εισαγωγή

Η Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis, PCA), αποτελεί την απλούστερη και πλέον διαδεδομένη πολυμεταβλητή ανάλυση για τον υπολογισμό, μέσω της μεθόδου των ελαχίστων τετραγώνων, της βέλτιστης ορθοκανονικής βάσης δεδομένων ενός πολυμεταβλητού συστήματος. Ανάλογα με το πεδίο εφαρμογής της είναι γνωστή και ως ορθή ορθογώνια αποσύζευξη (Proper Orthogonal Decomposition, POD) ή και Karhunen-Loève Αποσύζευξη (Karhunen-Loève Decomposition). [3] Η βασική ιδέα στην οποία βασίζεται η Ανάλυση Κυρίων Συνιστωσών είναι ότι η χρονική απόκριση ενός συστήματος σε μια συγκεκριμένη επιβολή στην είσοδο του, εμπεριέχει τη συμπεριφορά και τις βασικές ιδιότητες του δεδομένου συστήματος. Συνεπώς, με κατάλληλη μαθηματική επεξεργασία της εξόδου του συστήματος είναι εφικτή η δημιουργία ενός μοντέλου μειωμένης τάξης που θα εμπεριέχει τα παραπάνω, γεγονός μεγάλης σημασίας για την ακρίβεια του μοντέλου. [8]

Με την εφαρμογή της μεθόδου οικοδομείται υπόχωρος από συναρτήσεις βάσης οι οποίες διατηρούν ειδικά χαρακτηριστικά της δυναμικής του μελετώμενου συστήματος σε αντίθεση με άλλες μεθόδους (πχ πεπερασμένα στοιχεία), όπου οι συναρτήσεις βάσης είναι ανεξάρτητες του πλήρους συστήματος. Ακόμη, με την ανάλυση κυρίων συνιστωσών γίνεται εφικτή η εύρεση μοντέλων μειωμένης τάξης για δυναμικά συστήματα, μεγάλης ή και άπειρης διάστασης, χρησιμοποιώντας μόνο πράξεις μεταξύ πινάκων. Στην περίπτωση ενός μη γραμμικού συστήματος η μέθοδος είναι ικανή να παράγει μοντέλο μειωμένης τάξης, το οποίο θα είναι και αυτό μη γραμμικό, διατηρεί δηλαδή τα βασικά χαρακτηριστικά του. Ένα σημαντικό πλεονέκτημα της μεθόδου είναι ότι για την εφαρμογή της δεν είναι αναγκαία η πρότερη γνώση του πλήρους συστήματος, αλλά μόνο των αποτελεσμάτων από πειράματα ή προσομοιώσεις. Για τον λόγο αυτό η ανάλυση κυρίων συνιστωσών έχει ευρεία εφαρμογή σε πεδία όπως η αναγνώριση προτύπων (pattern recognition), η ρευστοδυναμική, η ανάλυση σημάτων και ο σχεδιασμός ρυθμιστή για χρήση σε πραγματικό χρόνο. [8]

Πέρα από τα πλεονεκτήματα που αναφέρθηκαν παραπάνω, η μέθοδος παρουσιάζει κάποιους περιορισμούς σχετικά με την προσέγγιση του πλήρους συστήματος. Οι περιορισμοί προέρχονται από την επιλογή του υπόχωρου από τον οποίο λαμβάνονται τα πειραματικά ή υπολογιστικά δεδομένα. Πιο συγκεκριμένα τα ιδιοδιανύσματα που παράγονται από την μέθοδο PCA και η διάταξη τους στο χώρο,

προσεγγίζουν μόνο τη δυναμική του υποσυνόλου του συστήματος από το οποίο ελήφθησαν τα δεδομένα και όχι του πλήρους συστήματος. Επιπλέον, κάποια σημαντικά φαινόμενα για την δυναμική του συστήματος που όμως δεν συμβαίνουν με μεγάλη συχνότητα, αγνοούνται από την ανάλυση αυτή. [3]

3.3.2 Βασικές εξισώσεις μεθόδου

Η PCA μέθοδος βασίζεται στο ότι οι μεταβλητές κατάστασης x (Παράγραφος 3.2) μπορούν να προσεγγιστούν από τις \hat{x} ως το γινόμενο συναρτήσεων βάσης οι οποίες διατηρούν την χωρική εξάρτηση του συστήματος αλλά είναι χρονικά ανεξάρτητες, και των συντελεστών που αντίστοιχα διατηρούν την χρονική εξάρτηση αλλά όχι την χωρική.

Αν $\hat{x}(\mathbf{r}, t)$ η προσέγγιση της μεταβλητής κατάστασης του συστήματος στην θέση \mathbf{r} την χρονική στιγμή t , τότε: [8][19][20]:

$$\hat{x}(\mathbf{r}, t) = \sum_{j=1}^d \alpha_j(t) \varphi_j(\mathbf{r}) \quad (3.7)$$

όπου η εξάρτηση της προσέγγισης \hat{x} από το χώρο καθορίζεται από τις συναρτήσεις $\varphi_j(\mathbf{r})$ οι οποίες όπως αναφέρθηκε δεν μεταβάλλονται στο χρόνο και σχηματίζουν τη βάση ενός d -διάστατου συναρτησιακού χώρου από τον οποίο προσεγγίζεται το πλήρες σύστημα.

Το διάνυσμα κατάστασης $\hat{\mathbf{x}}(t)$ προκύπτει εάν, για μια χρονική στιγμή t , τοποθετηθούν οι τιμές όλων των εξαρτημένων μεταβλητών x , σε όλες τις θέσεις του χώρου \mathbf{r} και η εξίσωση του διανύσματος γράφεται ως εξής:[8]

$$\hat{\mathbf{x}}(t) = \sum_{j=1}^d \alpha_j(t) \boldsymbol{\varphi}_j \quad (3.8)$$

Όπου $\boldsymbol{\varphi}_j$ είναι τα αντίστοιχα διανύσματα βάσης, που κατασκευάζονται από τις συναρτήσεις βάσεις $\varphi_j(\mathbf{r})$, με τον ίδιο τρόπο που κατασκευάστηκε το διάνυσμα κατάστασης $\hat{\mathbf{x}}(t)$.

Η ανάλυση κυρίων συνιστωσών, ενώ βασίζεται στις δύο παραπάνω εξισώσεις, μπορεί να υπολογίζει μόνο την βέλτιστη βάση του μοντέλου. Ο υπολογισμός της βέλτιστης βάσης επιτυγχάνεται με προβολή των διανυσμάτων που περιέχουν τα δεδομένα για την κατασκευή του Μοντέλου Μειωμένης Τάξης από το συνολικό χώρο V διάστασης n στον υπόχωρο V_d διάστασης $d \ll n$, ενώ για την επίτευξη της ελαχιστοποίησης του σφάλματος προσέγγισης πρέπει η προβολή να είναι ορθή και επομένως και η βάση να είναι ορθοκανονική, υπό την έννοια των ελαχίστων τετραγώνων. Σημαντικό ρόλο παίζει επίσης η επιλογή των δεδομένων που θα προβληθούν η οποία πολύ συχνά προσαρμόζεται σε στρατηγικές διερεύνησης ανάλογα με την πτυχή του συστήματος που διερευνάται. Μια σημαντική παραλλαγή της ανάλυσης κυρίων συνιστωσών αποτελεί η μέθοδος των στιγμιότυπων κατά την οποία το σύνολο των δεδομένων είναι διακριτοποιημένο στο χρόνο και αποτελείται από ξεχωριστές καταστάσεις του

συστήματος για κάθε χρονική στιγμή. Οι καταστάσεις αυτές ονομάζονται στιγμιότυπα, και ουσιαστικά αναιρούν την ύπαρξη μιας συνεχόμενης χρονικής τροχιάς, όπως συμβαίνει στην γενική μέθοδο PCA. Η μέθοδος αυτή θα αναπτυχθεί αναλυτικότερα παρακάτω.

Ο υπολογισμός των συντελεστών που καθορίζουν την χρονική εξάρτηση της προσέγγισης του διανύσματος κατάστασης, συνήθως επιτυγχάνεται με την Μέθοδο Galerkin. Όπως και η μέθοδος PCA, βασίζεται στις εξισώσεις 3.7 και 3.8 και αποτελεί μέθοδο προβολής, η συνήθης εφαρμογή της οποίας είναι συνδυαστικά με την μείωση τάξης ενός μοντέλου. Σύμφωνα με την μέθοδο αυτή, οι μερικές διαφορικές εξισώσεις του μοντέλου πλήρους τάξης προβάλλονται πάνω στην ορθοκανονική βάση που υπολογίζεται μέσω της PCA, επιδιώκοντας και εδώ την ελαχιστοποίηση της διαφοράς των τετραγώνων. Σημαντική διαφορά των δύο μεθόδων είναι ότι η μέθοδος Galerkin βασίζεται σε εξισώσεις σε αντίθεση με την PCA που βασίζεται σε δεδομένα. Συνεπώς, στην εξέλιξη της παρούσας εργασίας ο προσδιορισμός των συντελεστών α_j (3.7 & 3.8), επιτυγχάνεται με χρήση μιας στατικής μεθόδου η οποία απαιτεί την γνώση δεδομένων και όχι εξισώσεων, τα Τεχνητά Νευρωνικά Δίκτυα.

Έτσι καθίσταται εφικτή η κατασκευή ενός μοντέλου μειωμένης τάξης, με χρήση της μεθόδου ανάλυσης κυρίων συνιστωσών και των τεχνητών νευρωνικών δικτύων, για την κατασκευή του οποίου χρησιμοποιούνται μόνο αποτελέσματα από προσομοιώσεις του αναλυτικού μοντέλου. Τα τεχνητά νευρωνικά δίκτυα θα αναπτυχθούν εκτενέστερα στο επόμενο κεφάλαιο.

3.3.3 Κατασκευή Βάσης και Προβολής

Έστω ένα σύνολο διανυσμάτων προερχόμενα από έναν διανυσματικό χώρο $V = \mathbb{R}^N$. Επιπλέον, έστω ένα σύνολο δεδομένων Y που αποτελείται από m τροχιές (trajectories). [9], [8]

$$Y = \{\mathbf{y}_1(t), \dots, \mathbf{y}_m(t)\} \quad (3.9)$$

όπου $y_i(t) \in V, i = 1, \dots, m, t \in [0, T]$.

Στόχος είναι η κατασκευή μιας ορθής προβολής $\Pi_d : V \rightarrow V_d$, η οποία να είναι συγκεκριμένης διάστασης d και μέσω της οποίας η ακόλουθη απόσταση ελαχίστων τετραγώνων να ελαχιστοποιείται:

$$\|Y - \Pi_d Y\|^2 = \sum_{i=1}^m \int_0^T \|\mathbf{y}_i(t) - \Pi_d \mathbf{y}_i(t)\|^2 dt \quad (3.10)$$

Με χρήση του πίνακα συσχέτισης (correlation matrix) προκύπτει η λύση του προβλήματος:

$$K = \sum_{i=1}^m \int_0^T \mathbf{y}_i(t) \mathbf{y}_i(t)^T dt \in \mathbb{R}^{N \times N} \quad (3.11)$$

ο πίνακας \mathbf{K} είναι διάστασης $N \times N$ κι εξ' ορισμού συμμετρικός και θετικά ημιορισμένος. Επομένως έχει μη αρνητικές ιδιοτιμές, $\lambda_1 \geq \dots \geq \lambda_N \geq 0$, που ικανοποιούν την παρακάτω σχέση:

$$\mathbf{K}\mathbf{u}_j = \lambda_j\mathbf{u}_j, \quad j = 1, \dots, N \quad (3.12)$$

όπου, τα ιδιοδιανύσματα \mathbf{u}_j είναι μη μηδενικά, διάστασης N και επιλέγονται ούτως ώστε να είναι ορθοκανονικά. Ικανοποιούν δηλαδή την εξής σχέση:

$$\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (3.13)$$

ενώ κάθε διάνυσμα αναλύεται στις παρακάτω συνιστώσες:

$$\mathbf{y}_i(t) = \sum_{j=1}^N y_{ij}(t)\mathbf{u}_j \quad (3.14)$$

όπου,

$$y_{ij}(t) = \mathbf{y}_i(t)^T \mathbf{u}_j \quad (3.15)$$

Ο βέλτιστος υπόχωρος διάστασης d , μέσω του οποίου ελαχιστοποιείται η απόσταση (3.10), σύμφωνα με την μέθοδο ανάλυσης κύριων συνιστωσών, είναι το γραμμικό περίβλημα των πρώτων d ιδιοδιανυσμάτων, $\text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$, ενώ η βέλτιστη προβολή είναι η:

$$\Pi_d = \sum_{j=1}^d \mathbf{u}_j \mathbf{u}_j^T \in \mathbb{R}^{N \times N} \quad (3.16)$$

Τα ιδιοδιανύσματα αυτά ονομάζονται κυρίαρχα διανύσματα (PCA modes) κι αυτό γιατί η προσέγγιση μειωμένης τάξης του $\mathbf{y}_i(t)$ είναι σύμφωνα με την μέθοδο PCA, η προβολή Π_d σε κάθε διάνυσμα $\mathbf{y}_i(t)$. Η προβολή αυτή διατηρεί στην ουσία τις πρώτες d από τις συνολικά N συνιστώσες τις ακριβούς σχέσης (3.14), οι οποίες περιέχουν την κύρια πληροφορία για την δομή των διανυσμάτων και είναι η εξής:

$$\Pi_d \mathbf{y}_i(t) = \hat{\mathbf{y}}_i(t) = \sum_{j=1}^d y_{ij}(t)\mathbf{u}_j \quad (3.17)$$

Συνεπώς στην περίπτωση που χρησιμοποιείται η μέθοδος Galerkin, μέσω της μεθόδου PCA, προτείνονται ως διανύσματα βάσης τα πρώτα d ιδιοδιανύσματα του πίνακα συσχέτισης και ως συντελεστές της οι προβολές της σχέσης (3.15).

3.3.4 Επιλογή διάστασης βάσης

Η διάσταση d της ορθοκανονικής βάσης επιλέγεται έτσι ώστε η προβολή των διανυσμάτων $\mathbf{y}_i(t)$ (3.17) να είναι επαρκής. Βάσει της μεθόδου ανάλυσης κύριων συνιστωσών, σε δεδομένο υπόχωρο $V_d = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$, η τιμή της απόστασης (3.10), που είναι η μικρότερη δυνατή για δεδομένο d , είναι ίση με:

$$\min_{V_d} \|Y - \Pi_d Y\|^2 = \sum_{j=n-d+1}^n \lambda_j \quad (3.18)$$

Η απόσταση αυτή είναι ισούται με το άθροισμα των ιδιοτιμών των ιδιοδιανυσμάτων που δεν χρησιμοποιούνται στην κατασκευή της βάσης. Μπορεί να θεωρηθεί ως το μέρος της «ενέργειας», του πληροφοριακού περιεχομένου, δηλαδή, των διανυσμάτων που δεν λαμβάνεται υπόψη στη προσέγγιση (3.17). [8][9]

Το άθροισμα των πρώτων d ιδιοτιμών μπορεί να αποτελέσει την ποσοτικοποίηση της πληροφορίας της παραπάνω προσέγγισης:

$$\sum_{j=1}^d \lambda_j \quad (3.19)$$

Το σύνολο της πληροφορίας του συστήματος αντιστοιχεί σε $d=N$, όπου μηδενίζεται και η απόσταση (3.18). Έτσι, για κάθε προσέγγιση που υπολογίζεται με την μέθοδο ανάλυσης κύριων συνιστωσών, διάστασης d , το σχετικό πληροφοριακό περιεχόμενο ορίζεται ως:

$$I(d) = \frac{\sum_{j=1}^d \lambda_j}{\sum_{j=1}^n \lambda_j} \quad (3.20)$$

Κατά τον υπολογισμό του παραπάνω πληροφοριακού περιεχομένου, παρατηρείται αρχικά μια απότομη αύξηση καθώς προστίθενται πρώτα οι μεγαλύτερες ιδιοτιμές που εκφράζουν τα κύρια χαρακτηριστικά μιας δυναμικής ενώ στη συνέχεια ο ρυθμός αύξησης του περιεχομένου συνεχώς μικραίνει καθώς οι ιδιοτιμές που προστίθενται είναι μικρότερες, αντιστοιχούν δηλαδή σε διαταραχές της δυναμικής του συστήματος. [8]

Η διάσταση d της βάσης μπορεί να υπολογισθεί εάν προεπιλεγεί η τιμή του κλάσματος του πληροφοριακού περιεχομένου που είναι επιθυμητό να περιέχει η προσέγγιση, p ίση με 0.99 και άνω. [10]

$$d = \min\{d \in \mathbb{N}^* | I(d) \geq p\} \quad (3.21)$$

Ο παραπάνω τρόπος υπολογισμού της διάστασης d της βάσης εξασφαλίζει ότι η προσέγγιση θα παρέχει τις σημαντικές παραμέτρους της δυναμικής του συστήματος, είναι όμως πιθανό λόγω της προεπιλογής του p να συμπεριληφθούν περισσότερα ιδιοδιανύσματα από τα απαραίτητα λόγω υπολογιστικού θορύβου, που να μειώνουν την ακρίβεια των αποτελεσμάτων.

3.4 Μέθοδος των στιγμιοτύπων

3.4.1 Περιγραφή μεθόδου

Συχνά τα περίπλοκα δυναμικά συστήματα έχουν μεγάλες τιμές αρχικής διάστασης N των διανυσμάτων τροχιάς, με αποτέλεσμα ο πίνακας συσχέτισης (3.11) να αποκτά μεγάλο μέγεθος και συνεπώς η εύρεση των ιδιοδιανυσμάτων και ιδιοτιμών του (3.12) να είναι από υπολογιστικής πλευράς δύσκολη ή και ακατόρθωτη. Για την απλοποίηση

της ανάλυσης μεγάλου όγκου δεδομένων, ο L. Sirovich πρότεινε την μέθοδο των στιγμιοτύπων (method of snapshots), που αποτελεί μια παραλλαγή της μεθόδου ανάλυσης κυρίων συνιστωσών. [11]

Η απλοποίηση της ανάλυσης επιτυγχάνεται με τη χρήση στιγμιοτύπων των τροχιών σε διακριτές χρονικές στιγμές έναντι των τροχιών $y_i(t)$ που ορίζονται σε συνεχή χρονικά διαστήματα. Έτσι το σύνολο δεδομένων (3.9) μετατρέπεται στον πίνακα: [8]

$$\mathbf{Y} = [\mathbf{Y}_1 \dots \mathbf{Y}_{N_R}] \in \mathbb{R}^{N \times m} \quad (3.22)$$

Όπου $\mathbf{Y}_i(t) \in \mathbb{R}^{N \times m_i}$, $i = 1, \dots, N_R$ είναι ο πίνακας της i τροχιάς και N_R ο συνολικός αριθμός των τροχιών που εκτελεί το σύστημα. Κάθε τροχιά αποτελείται από m_i στιγμιότυπα. Άρα ο συνολικός αριθμός των στιγμιοτύπων ορίζεται ως το εξής άθροισμα:

$$m = \sum_{i=1}^{N_R} m_i \quad (3.23)$$

Κάθε στήλη του πίνακα δεδομένων που προέκυψε παραπάνω (3.22) αποτελεί ένα στιγμιότυπο του συστήματος. Ο πίνακας συσχέτισης υπολογίζεται πλέον ως εξής:

$$\mathbf{K} = \mathbf{Y}\mathbf{Y}^T \in \mathbb{R}^{N \times N} \quad (3.24)$$

του οποίου οι ιδιοτιμές είναι ίσες με του πίνακα:

$$\mathbf{K}_1 = \mathbf{Y}^T\mathbf{Y} \in \mathbb{R}^{m \times m} \quad (3.25)$$

όπως έχει αναφερθεί η μέθοδος των στιγμιοτύπων εφαρμόζεται με στόχο την μείωση της τιμής του N , και συνεπώς είναι εύλογος ο ισχυρισμός ότι $m \ll N$. Έτσι προκύπτει ότι τα ιδιοδιανύσματα του πίνακα \mathbf{K}_1 μπορούν να υπολογισθούν με μεγαλύτερη ευκολία:

$$\mathbf{K}_1 \mathbf{v}_j = \lambda_j \mathbf{v}_j \quad (3.26)$$

όπου $j=1, \dots, m$. Όπως και για τον πίνακα συσχέτισης (3.11), οι ιδιοτιμές είναι μη αρνητικές και πιο συγκεκριμένα $\lambda_1 \geq \dots \geq \lambda_m \geq 0$ και τα ιδιοδιανύσματα $\mathbf{v}_j \in \mathbb{R}^m$ του \mathbf{K}_1 επιλέγονται να είναι ορθοκανονικά, μη μηδενικά. Στη συνέχεια πραγματοποιείται ανάκτηση των ιδιοδιανυσμάτων \mathbf{u}_j του πίνακα \mathbf{K} , από τα οποία επιλέγονται τα d πρώτα, ως βάση του μειωμένου τάξης μοντέλου, με χρήση της σχέσης: [8]

$$\mathbf{u}_j = \frac{1}{\sqrt{\lambda_j}} \mathbf{Y} \mathbf{v}_j \quad (3.27)$$

όπου $j=1, \dots, m$.

Η προσέγγιση κάθε στιγμιοτύπου y_i υλοποιείται με χρήση της σχέσης (3.17), ως εξής:

$$\hat{y}_i = \sum_{j=1}^d \alpha_{ij} \mathbf{u}_j \quad (3.28)$$

όπου οι συντελεστές α_{ij} υπολογίζονται από τη σχέση (3.15) ως εξής:

$$a_{ij} = \mathbf{y}_i^T \mathbf{u}_j \quad (3.29)$$

Αν ορισθεί ως \mathbf{Z} η ορθοκανονική βάση αποτελούμενη από τα d πρώτα ιδιοδιανύσματα, δηλαδή τα κύρια PCA διανύσματα, τότε:

$$\mathbf{Z} := [\mathbf{u}_1 \dots \mathbf{u}_d] \in \mathbb{R}^{N \times d} \quad (3.30)$$

και α_i το διάνυσμα των συντελεστών που αντιστοιχούν στο στιγμιοτύπο i :

$$\alpha_i = [a_{i1} \dots a_{id}] \in \mathbb{R}^d \quad (3.31)$$

Κι επειδή τα διανύσματα της βάσης έχουν επιλεγθεί να είναι ορθοκανονικά, ισχύει η σχέση:

$$\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_d \quad (3.32)$$

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

Οι σχέσεις (3.28) και (3.29) μπορούν τώρα να γραφούν με μια πιο απλή μορφή:

$$\hat{y}_i = \mathbf{Z} \alpha_i \quad (3.33)$$

$$\alpha_i = \mathbf{Z}^T y_i \quad (3.34)$$

Από τις παραπάνω σχέσεις προκύπτει το συμπέρασμα ότι το $\alpha_i \in \mathbb{R}^d$ αποτελεί μια μειωμένης τάξης προσέγγιση του διανύσματος $y_i \in \mathbb{R}^N$

Επιπλέον, μπορεί να ορισθεί ο πίνακας \mathbf{A} , των συντελεστών όλων των ιδιοδιανυσμάτων ως εξής:

$$\mathbf{A} = [\alpha_1 \dots \alpha_m] \in \mathbb{R}^{d \times m} \quad (3.35)$$

Και άρα μια γενικευμένη μορφή των εξισώσεων (3.33) και (3.34) είναι η εξής:

$$\hat{\mathbf{Y}} = \mathbf{Z} \mathbf{A} \quad (3.36)$$

$$\mathbf{A} = \mathbf{Z}^T \mathbf{Y} \quad (3.37)$$

Οι σχέσεις (3.35)-(3.37) γίνεται να αναφέρονται είτε σε ολόκληρο τον πίνακα \mathbf{Y} , είτε στον πίνακα μιας τροχιάς, όπου m θα είναι πλέον ο αριθμός στιγμιοτύπων της τροχιάς.

3.4.2 Κανονικοποίηση δεδομένων

Ένα πολύ συχνό φαινόμενο κατά την εφαρμογή της μεθόδου ανάλυσης κυρίων συνιστωσών, είναι η τροποποίηση των δεδομένων κατά τον υπολογισμό της μέσης τιμής του πίνακα \mathbf{Y} , δηλαδή της τιμής κάθε μεγέθους σε κάθε θέση:[8]

$$\bar{x}_i = \frac{1}{m} \sum_{j=1}^m (\mathbf{Y})_{ij} \quad , \quad i = 1, \dots, N \quad (3.38)$$

με την αφαίρεση της μέσης τιμής που υπολογίζεται από κάθε αντίστοιχο στοιχείο όλων των στιγμιότυπων, υπολογίζεται διανυσματικά το μέσο στιγμιότυπο:

$$\bar{\mathbf{x}}_i := [\bar{x}_1 \dots \bar{x}_N]^T \in \mathbb{R}^N \quad (3.39)$$

το οποίο στη συνέχεια αφαιρείται από κάθε στιγμιότυπο του πίνακα \mathbf{Y} .

Επιπλέον, πραγματοποιείται ο υπολογισμός της τυπικής απόκλισης για κάθε γραμμή του πίνακα \mathbf{Y} , ως εξής:

$$s_i = \left(\frac{1}{m-1} \sum_{j=1}^m ((\mathbf{Y})_{ij} - \bar{x}_i)^2 \right)^{\frac{1}{2}} \quad , \quad i = 1, \dots, N \quad (3.40)$$

Διαιρώντας κάθε στοιχείο του προηγούμενου διανύσματος με την αντίστοιχη τυπική απόκλιση, οι τροχιές κανονικοποιούνται ώστε κάθε στοιχείο να έχει την μορφή:

$$v_{ij} = \begin{cases} \frac{(\mathbf{Y})_{ij} - \bar{x}_i}{s_i} & , s_i \neq 0, i = 1, \dots, N, j = 1, \dots, m \\ 0 & , s_i = 0 \end{cases} \quad (3.41)$$

Η κανονικοποίηση των δεδομένων πραγματοποιείται για να μειωθεί η πιθανότητα οι ιδιοτιμές να διαφέρουν σε τάξη μεγέθους, λόγω διαφοράς στα αρχικά δεδομένα του συστήματος. Αυτό συμβαίνει στην παρούσα διπλωματική εργασία όπου η λύση του συστήματος που μελετάται είναι οι κατανομές στο χώρο και στον χρόνο των πεδίων $P(\mathbf{r},0)$, $\mathbf{u}(\mathbf{r},0)$ και $T(\mathbf{r},0)$ και οι οποίες διαμορφώνουν ένα διάνυσμα κατάστασης διάστασης $N=60264$. Οι τιμές της θερμοκρασίας στον χώρο διαφέρουν κατά πολύ από τις τιμές της πίεσης ή της ταχύτητας. Για τον λόγο αυτό μπορεί να προκληθούν σημαντικές αποκλίσεις της τελικής προσέγγισης αν δεν πραγματοποιηθεί κανονικοποίηση του πίνακα \mathbf{Y} . Τέτοιες αποκλίσεις μπορεί να είναι η προσέγγιση να περιγράψει σωστά μόνο την κατανομή της θερμοκρασίας. [6]

Η δεύτερη περίπτωση της σχέσης (3.41) είναι επέκταση της πρώτης διότι όταν $s_i = 0$ συνεπάγεται από τη σχέση (3.40) ότι $(\mathbf{Y})_{ij} = \bar{x}_i$ δηλαδή κάθε στοιχείο της γραμμής i ισούται με τη μέση τιμή \bar{x}_i . Η περίπτωση αυτή πραγματοποιείται μόνο για σταθερές τιμές μεταβλητών όπως συμβαίνει στα σύνορα του πεδίου όπου επιβάλλεται κάποια μεγέθη να είναι σταθερά. Στον αντιδραστήρα που μελετάται στην παρούσα εργασία αυτό συμβαίνει στα τοιχώματα του όπου επιβάλλεται σταθερή θερμοκρασία 300K.

Η ορθοκανονική βάση που προκύπτει καθώς και η προσέγγιση του πίνακα Y , όπως έχει αναφερθεί και παραπάνω, είναι σε κανονικοποιημένη μορφή. Το επόμενο βήμα που ακολουθεί είναι η αντίστροφη διαδικασία δηλαδή ο υπολογισμός της πραγματικής προσέγγισης του συστήματος. Αυτό επιτυγχάνεται αν κάθε στοιχείο της προσέγγισης του στιγμιότυπου, πολλαπλασιαστεί με την αντίστοιχη τυπική απόκλιση s_i κι έπειτα προστεθεί η μέση τιμή, \bar{x}_i . [6]

$$\hat{Y}_N = ZA \quad (3.42)$$

$$(\hat{Y})_{ij} = \bar{x}_i + s_i \hat{Y}_N \quad (3.43)$$

Όπου \hat{Y}_N αποτελεί την προσέγγιση του κανονικοποιημένου πίνακα.

Η τελευταία σχέση δεν περιορίζει την εφαρμογή της μόνο σε τροχιές που έχουν χρησιμοποιηθεί για την κατασκευή της βάσης, δηλαδή για την κατασκευή του πίνακα Y . Για τροχιές που δεν περιλαμβάνονται στον πίνακα Y , το μέσο στιγμιότυπο (3.39) και η τυπική απόκλιση (3.44) δεν υπολογίζονται από την κανονικοποίηση της δεδομένης τροχιάς αλλά ολόκληρου του πίνακα Y , πριν την κατασκευή της βάσης Z .

3.5 Αξιολόγηση μείωσης τάξης – Σφάλματα

Το σχετικό σφάλμα της εκτίμησης ενός στιγμιότυπου $\mathbf{x}(t)$, που ορίζεται από τη σχέση (3.44), χρησιμοποιείται κατά την διάρκεια των υπολογισμών για την αξιολόγηση της βάσης.

$$e_r(\mathbf{x}(t)) = \frac{\|\hat{\mathbf{x}}(t) - \mathbf{x}(t)\|}{\|\mathbf{x}(t)\|} \quad (3.44)$$

Δύο επιπλέον τρόποι αξιολόγησης της περιγραφής μιας γνωστής τροχιάς από την βάση που κατασκευάστηκε, αποτελούν ο υπολογισμός του μέσου τετραγωνικού σχετικού σφάλματος mse_r , που υπολογίζεται από τη σχέση (3.45) καθώς και του μέγιστου σχετικού σφάλματος max_{er} , το οποίο ορίζεται από τη σχέση (3.46).

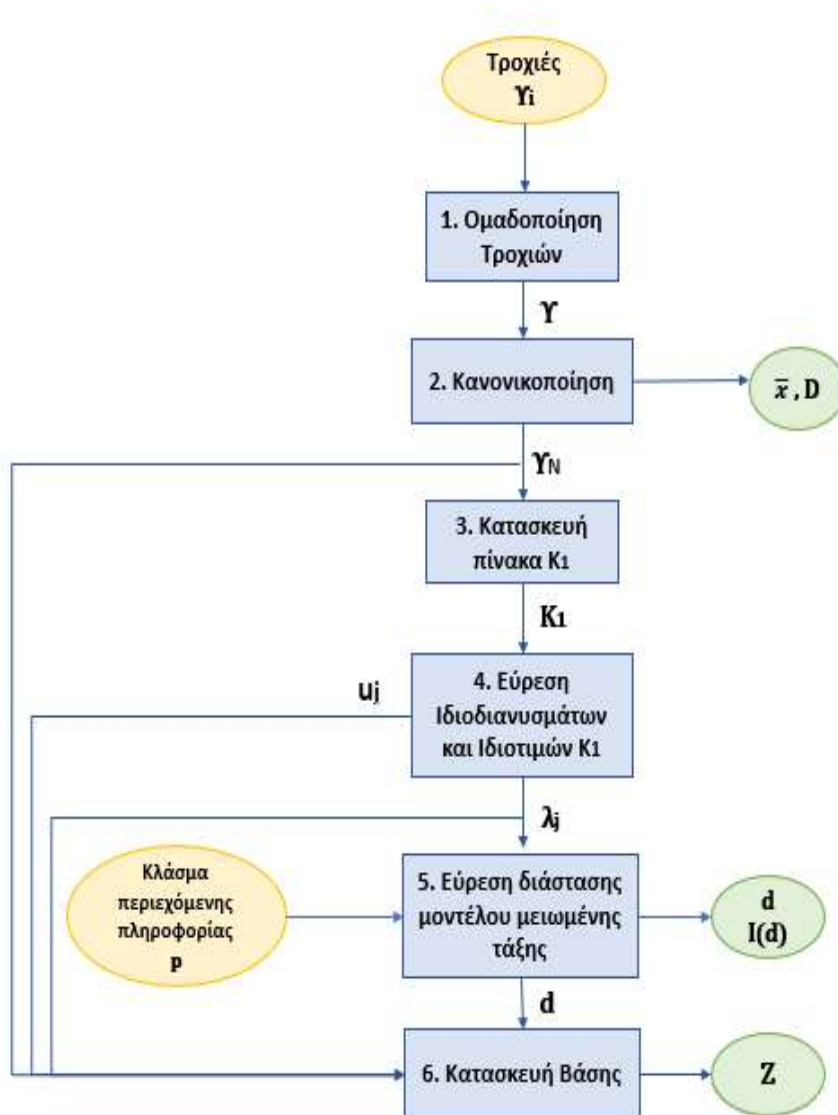
$$mse_{rr} = \left(\frac{1}{m} \sum_{j=1}^m e_r^2(\mathbf{x}(t_j)) \right)^{\frac{1}{2}} \quad (3.45)$$

$$max_{er} = \max \left(e_r(\mathbf{x}(t_j)) \right), \quad j = 1, \dots, m \quad (3.46)$$

Το σχετικό σφάλμα (3.44) ισχύει και για καταστάσεις ισορροπίας που δεν αποτελούν στιγμιότυπα του συστήματος. Για τα στιγμιότυπα αυτά, οι σχέσεις (3.45) και (3.46), μπορούν να χρησιμοποιηθούν εάν η μεταβλητή m αποτελεί πλέον το πλήθος των καταστάσεων ισορροπίας και αν ο δείκτης j αριθμεί τις καταστάσεις ισορροπίας, αντί για διαδοχικές στιγμές. Τα σφάλματα αυτά αξιολογούν σε αυτή την περίπτωση την περιγραφή ενός συνόλου καταστάσεων ισορροπίας από μια υποψήφια βάση.

3.6 Αλγόριθμος για την κατασκευή της βάσης

Στο Σχήμα 3.3 παρουσιάζεται ο αλγόριθμος, που περιέχει συνοπτικά όλα τα βήματα που ακολουθούνται για την κατασκευή της βάσης Z , του μοντέλου μειωμένης τάξης.



Σχήμα 3. 2: Αλγόριθμος κατασκευής βάσης μοντέλου μειωμένης τάξης. Με κίτρινο χρώμα σημειώνονται τα αρχικά δεδομένα, ενώ με πράσινο τα τελικά αποτελέσματα. Τα βήματα του αλγορίθμου σημειώνονται με μπλε χρώμα.

Αρχικά κατασκευάζεται ο πίνακας Y , από την ένωση όλων των τροχιών (Trajectories) που χρησιμοποιούνται για την κατασκευή της βάσης (Σχέση 3.22). Οι τροχιές αυτές αποτελούν βηματικές επιβολές που εκτελούν συγκεκριμένο αριθμό επαναλήψεων (Παράγραφος 2) και έχουν επιλεχθεί με τρόπο ώστε να περιγράφεται όλο το εύρος της πολλαπλότητας λύσεων που εμφανίζει το σύστημα.

Έπειτα υπολογίζεται η μέση τιμή κάθε γραμμής του πίνακα τροχιών Y και στη συνέχεια αφαιρείται από κάθε στοιχείο της. Υπολογίζεται επίσης η τυπική απόκλιση κάθε γραμμής του πίνακα Y και διαιρείται με κάθε στοιχείο σύμφωνα με την (3.41), ώστε να κανονικοποιηθεί ο πίνακας και να μην παρουσιαστούν στη συνέχεια προβλήματα στις προσεγγίσεις λόγω μεγάλων διαφορών στις τάξεις μεγέθους των στοιχείων του πίνακα Y .

Ακολουθεί ο υπολογισμός του πίνακα σχετικής ενίσχυσης K_1 , που υπενθυμίζεται είναι μικρότερης διάστασης από τον πίνακα συσχέτισης K (3.11), των ιδιοδιανυσμάτων και ιδιοτιμών του, και με βάση το κλάσμα περιεχόμενης πληροφορίας p που επιλέγεται κάθε φορά, υπολογίζεται η διάσταση d της ορθοκανονικής βάσης (3.21). Τέλος, υπολογίζονται τα d πρώτα ιδιοδιανύσματα του πίνακα K , τα οποία αποτελούν την ορθοκανονική βάση διάστασης d , Z_d .

3.7 Εναλλακτικές προσεγγίσεις ανάλυσης κυρίων συνιστωσών

3.7.1. Εισαγωγή

Όπως αναφέρθηκε στην Παράγραφο 3.4.1 η ανάλυση κυρίων συνιστωσών αδυνατεί να παρέχει πληροφορίες για το πλήρες μελετώμενο σύστημα και συνεπώς ενέργειες όπως η αναπαράσταση όλων των πιθανών δυναμικών του συστήματος ή ο αποτελεσματικός σχεδιασμός ενός μη γραμμικού συστήματος ρύθμισης, δεν μπορούν να πραγματοποιηθούν. [3] Στην προσπάθεια να αντιμετωπιστεί ο περιορισμός αυτός της μεθόδου, οι Michael D. Graham και Ioannis G. Kevrekidis (1995) [3], πραγματοποίησαν κάποιες εναλλακτικές προσεγγίσεις ως προς τον υπόχωρο από τον οποίο αντλούνται τα πειραματικά ή υπολογιστικά δεδομένα για ένα σύστημα διάχυσης (Παράγραφος 2). Στην παραπάνω μελέτη βασίστηκε η ιδέα της παρούσας εργασίας.

Για την διευκόλυνση στην κατανόηση των αλλαγών που πραγματοποίησαν κατά την έρευνα τους, στη συνέχεια αναπτύσσεται συνοπτικά η πιο συνήθης εφαρμογή της ανάλυσης κυρίων συνιστωσών για ένα δυναμικό σύστημα.

Ένα δυναμικό σύστημα, μεταβαλλόμενο ως προς τον χρόνο και τον χώρο, όπως είναι και το παρόν σύστημα, ξεκινάει από μια αρχική μεταβαλλόμενη θέση και καταλήγει σε μια σταθερή κατάσταση που δεν αποτελεί απαραίτητα και κατάσταση ισορροπίας. Αυτή η θέση θα καλείται στη συνέχεια “τελική κατάσταση”. Τα δεδομένα αυτής της διαδικασίας λαμβάνονται με την μορφή διακριτοποιημένων εικόνων (Snapshots) ακολουθώντας την χρονική εξέλιξη (Τροχιά) μιας ή περισσοτέρων αρχικών καταστάσεων (Initial Solution), με σκοπό να διατηρηθούν μερικά δεδομένα για την προσέγγιση της τελικής κατάστασης. Έπειτα με χρήση αυτών των δεδομένων, υπολογίζεται η βέλτιστη ορθοκανονική βάση η οποία αποτελείται από έναν δεδομένο αριθμό ιδιοδιανυσμάτων.

Η ιδέα της παρούσας εργασίας, βασίστηκε στην αναπαράσταση της δυναμικής του συστήματος με την κατασκευή μιας ορθοκανονικής βάσης, από δεδομένα τα οποία λαμβάνονται από ένα σύνολο σημείων που βρίσκονται κοντά σε μια αρχική κατάσταση. Το σύνολο των δεδομένων προς επεξεργασία προκύπτει με την επιλογή τυχαίων αρχικών καταστάσεων που αφήνονται να εξελιχθούν για ένα σύντομο χρονικό διάστημα προς την τελική κατάσταση που αναφέρθηκε παραπάνω. Με αυτόν τον τρόπο επιδιώκεται η πιο σφαιρική προσέγγιση του υπό μελέτη συστήματος.

Η επιτυχία της μεθόδου εξαρτάται από την επιλογή του συνόλου από το οποίο λαμβάνονται οι αρχικές καταστάσεις και το χρονικό διάστημα στο οποίο αφήνεται το σύστημα να εξελιχθεί.

3.7.2. Επιλογή παραμέτρων

Η επιλογή του δείγματος δεδομένων για τον υπολογισμό της κατάστασης αυτής, πραγματοποιείται με στόχο την επίτευξη καθολικής πληροφόρησης. Συγκεκριμένα, επιλέγονται κάποιες αρχικές καταστάσεις από έναν υπόχωρο επιλεγμένο με τέτοιο τρόπο ώστε να εμπεριέχει όλο το εύρος στο οποίο αναπαρίστανται οι δυναμικές του

συστήματος (παράγραφος 2.5) και αφήνονται να εξελιχθούν για ένα σύντομο χρονικό διάστημα τ .

Το χρονικό διάστημα, τ , θα πρέπει να είναι αρκετά μεγάλο ώστε οι αρχικές καταστάσεις να εξελιχθούν προς την “τελική κατάσταση”, αλλά και σύντομο αρκετά ώστε να μην προλάβουν οι τροχιές να συγκλίνουν σε μια συγκεκριμένη θέση. Αν το χρονικό διάστημα είναι πολύ μικρό η προσέγγιση της “τελικής κατάστασης” θα είναι ουσιαστικά όλος ο υπόχωρος που επιλέχθηκε αρχικά, ενώ αν είναι πολύ μεγάλο δεν θα είναι εφικτή η παροχή πληροφοριών για το πλήρες σύστημα που είναι ο στόχος της διερεύνησης αυτής.

Συνεπώς ο υπόχωρος από τον οποίο αντλούνται οι αρχικές καταστάσεις, το πλήθος τους καθώς και το χρονικό διάστημα στο οποίο αφήνονται να εξελιχθούν παίζουν πολύ σημαντικό ρόλο για την επιτυχία της μεθόδου. Έπειτα με χρήση των δεδομένων που αντλούνται με τον παραπάνω τρόπο υπολογίζεται η βέλτιστη ορθοκανονική βάση, η οποία μπορεί να προσεγγίζει ικανοποιητικά την δυναμική του πλήρους συστήματος.

3.7.3. Εφαρμογή της προσέγγισης στο μελετώμενο σύστημα

Στην παρούσα εργασία, πραγματοποιήθηκε η ανάπτυξη ενός μοντέλου μειωμένης τάξης για την περιγραφή της δυναμικής συμπεριφοράς του πλήρους συστήματος, μέσω της εναλλακτικής υλοποίησης της μεθόδου ανάλυσης κυρίων συνιστωσών και την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου (Παράγραφος 4), χρησιμοποιώντας δεδομένα από μη μόνιμες καταστάσεις.

Συγκεκριμένα για την κατασκευή του μοντέλου χρησιμοποιήθηκαν δεδομένα από τις χρονικές εξελίξεις του συστήματος (τροχιές), καθώς μεταβαίνει από ένα πλήθος αρχικών καταστάσεων προς την ίδια τελική μόνιμη κατάσταση, χωρίς όμως να φτάνει στην κατάσταση αυτή. Με χρήση της προσομοίωσης του αντιδραστήρα στο λογισμικό Fluent, πραγματοποιήθηκαν κάποια υπολογιστικά πειράματα κατά την διάρκεια των οποίων επιβάλλονταν μια βηματική επιβολή στην μαζική παροχή του αντιδραστήρα, που είχε σε όλα τα την τιμή $3 \cdot 10^{-5}$ kg/s, και μελετήθηκε έπειτα η εξέλιξη του προς την νέα αυτή μόνιμη κατάσταση. Οι αρχικές καταστάσεις επιλέγονται με τέτοιο τρόπο ώστε να καλύπτεται όλο το φάσμα της πολλαπλότητας των λύσεων του συστήματος (Παράγραφος 2), ενώ τα δεδομένα για την κατασκευή του μοντέλου αντλούνται από ένα σύνολο σημείων που βρίσκονται κοντά στις αρχικές αυτές καταστάσεις, κατά την εξέλιξη του συστήματος προς την τελική μόνιμη κατάσταση.

Αυτό πραγματοποιείται με την βοήθεια του λογισμικού Matlab, το οποίο καλεί επανειλημμένα, εξωτερικά, τον κώδικα υπολογιστικής ρευστοδυναμικής Ansys Fluent. Το Matlab καθιστά εφικτή τη διακοπή της εξέλιξης του συστήματος μετά από ένα συγκεκριμένο χρονικό διάστημα. Το χρονικό αυτό διάστημα, όπως και το χρονικό διάστημα τ , δεν θα πρέπει να είναι ούτε πολύ μεγάλο ώστε το σύστημα να πλησιάζει την τελική μόνιμη κατάσταση, ούτε πολύ μικρό ώστε τα δεδομένα για την κατασκευή του μοντέλου να αποτελούν τις αρχικές καταστάσεις από τις οποίες ξεκίνησε το σύστημα.

Κι εδώ η επιτυχία της μεθόδου εξαρτάται από το πλήθος των αρχικών καταστάσεων που θα επιλεγθούν, καθώς και το χρονικό διάστημα στο οποίο αφήνονται να εξελιχθούν οι καταστάσεις αυτές, κατά την τροχιά του συστήματος προς την νέα μόνιμη κατάσταση.

4. Τεχνητά Νευρωνικά Δίκτυα

4.1. Εισαγωγή

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks, ANN) ανήκουν στον κλάδο της τεχνητής νοημοσύνης. Αποτελούνται από ένα σύνολο απλών, διασυνδεδεμένων και προσαρμοστικών μονάδων, οι οποίες συνιστούν ένα παράλληλο πολύπλοκο υπολογιστικό μοντέλο. Με χρήση αυτού του μαθηματικού μοντέλου, γίνεται προσπάθεια μίμησης της λειτουργίας των βιολογικών νευρώνων, που αποτελούν θεμελιώδη μονάδα του εγκεφάλου, και της δομής των βιολογικών Νευρωνικών Δικτύων. Μέσω των τεχνητών νευρωνικών δικτύων έχει επιτευχθεί η επιτυχής αντιμετώπιση προβλημάτων όπως η ταξινόμηση, η αναγνώριση προτύπων, η προσέγγιση συναρτήσεων και η προσομοίωση δυναμικών συστημάτων. [11] [12]

Το νευρικό σύστημα των έμβιων 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 έτσι ώστε να χρησιμοποιούν τον κατάλληλο τρόπο επεξεργασίας δεδομένων, μέσω αλγορίθμων εκπαίδευσης και δεδομένων εκπαίδευσης που βοηθούν στον υπολογισμό κατάλληλων παραμέτρων που καθορίζουν τα τεχνητά νευρωνικά δίκτυα [12].

Τα Τεχνητά Νευρωνικά Δίκτυα «μοιάζουν» με τον ανθρώπινο εγκέφαλο στα εξής σημεία: [11]

- Η γνώση αποκτάται από το δίκτυο μέσω μιας διαδικασίας μάθησης – εκπαίδευσης
- Η γνώση αποθηκεύεται στις δυνάμεις σύνδεσης των νευρώνων, οι οποίες είναι τα συναπτικά βάρη

Ένα πολύ σημαντικό πλεονέκτημα των νευρωνικών δικτύων είναι πως δεν απαιτείται καμία πρότερη γνώση των εξισώσεων και των μηχανισμών που περιγράφουν το φυσικό σύστημα το οποίο καλούνται να προσεγγίσουν. Ένα τεχνητό νευρωνικό δίκτυο που εκπαιδεύεται από ένα σύνολο δεδομένων μπορεί να αναπαράγει και να γενικεύσει τη συμπεριφορά του υπό μελέτη συστήματος εντός του πεδίου που ανήκουν τα δεδομένα εκπαίδευσης. [6] Αυτό επιτυγχάνεται έχοντας ως μόνη γνώση τα θεωρητικά, πειραματικά ή υπολογιστικά δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση του.

Με βάση όσα αναφέρθηκαν στην προηγούμενη παράγραφο, το διάνυσμα κατάστασης ενός δυναμικού συστήματος μπορεί να προσεγγιστεί από το γραμμικό συνδυασμό διανυσμάτων βάσης που περιέχουν την χωρική εξάρτηση, με κατάλληλους συντελεστές που περιέχουν την χρονική εξάρτηση της κατάστασης. Τα διανύσματα της βάσης, τα οποία ονομάζονται και κύρια PCA διανύσματα, προσδιορίζονται σύμφωνα με την Ορθή Ορθογώνια Αποσύνθεση ενός πίνακα δεδομένων που μπορεί να προκύψει είτε από πειράματα, είτε από υπολογιστικές προσομοιώσεις του πλήρους συστήματος. Ο υπολογισμός των χρονικών συντελεστών πραγματοποιείται με χρήση των Τεχνητών Νευρωνικών Δικτύων, κατά

την εφαρμογή των οποίων απαιτείται και πάλι η γνώση δεδομένων και όχι εξισώσεων.

Επομένως προκύπτει το συμπέρασμα, ότι η κατασκευή ενός μοντέλου μειωμένης τάξης, που επιτυγχάνεται μέσω της ανάλυσης κυρίων συνιστωσών και την εκπαίδευση τεχνητών νευρωνικών δικτύων, για την προσέγγιση ενός πολύπλοκου δυναμικού συστήματος, μπορεί να πραγματοποιηθεί με χρήση αποτελεσμάτων από κατάλληλα πειράματα ή προσομοιώσεις του συστήματος και χωρίς καμία άλλη γνώση του λεπτομερούς συστήματος. [12]

4.2. Ο νευρώνας

4.2.1. Βιολογικός νευρώνας

Το έργο στο επιστημονικό πεδίο των Τεχνητών Νευρωνικών Δικτύων βασίζεται στο γεγονός ότι ο ανθρώπινος εγκέφαλος εκτελεί τους υπολογισμούς με εντελώς διαφορετικό τρόπο από το συμβατικό υπολογιστή. Η αναλογία μεταξύ της νευροφυσιολογίας του ανθρώπινου εγκεφάλου και των τεχνητών νευρωνικών δικτύων είναι πολύ σημαντική για την κατανόηση των δικτύων αυτών. [11]

Ο βιολογικός νευρώνας είναι ένας εξειδικευμένος τύπος κυττάρου που απαρτίζει τη βασική μονάδα του συστήματος επεξεργασίας πληροφοριών που δομούν το νευρικό σύστημα όλων των ζωντανών οργανισμών. 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. Ανάλογα με το μέσο με το οποίο μεταφέρεται η πληροφορία από κύτταρο σε κύτταρο, οι νευρώνες διακρίνονται σε δύο είδη: [11][13]

- Νευρώνες με χημικές συνάψεις

Σε αυτό το είδος νευρώνα η πληροφορία διαδίδεται μέσω της ροής συγκεκριμένων χημικών ενώσεων, των νευροδιαβιβαστών (neurotransmitters), από το κύτταρο πομπό στο κύτταρο δέκτη και η φορά διάδοσης της πληροφορίας είναι μία, συγκεκριμένη και προκαθορισμένη.

- Νευρώνες με ηλεκτρικές συνάψεις

Η πληροφορία διαδίδεται μέσω της ροής ηλεκτρονίων με σχεδόν απευθείας επαφή των δύο κυττάρων. Στην περίπτωση αυτή οι ρόλοι μεταξύ κυττάρου-πομπού και κυττάρου-δέκτη εναλλάσσονται κατά περίπτωση καθώς η πληροφορία μπορεί να διαδοθεί και προς τις δύο κατευθύνσεις.

Τα τεχνητά νευρωνικά δίκτυα επιδιώκουν την μοντελοποίηση των βιολογικών νευρώνων με χημικές συνάψεις. Με τη βοήθεια του Σχήματος 4.1 και ακολουθώντας

την κατεύθυνση μετάδοσης της πληροφορίας, αναπτύσσονται συνοπτικά τα κύρια μέρη από τα οποία αποτελούνται οι νευρώνες: [14]



Σχήμα 4. 1: Βιολογικός νευρώνας. Αποτελείται από το σώμα, που αποτελεί τον πυρήνα του, τους δενδρίτες μέσω των οποίων λαμβάνει σήματα από γειτονικούς νευρώνες (σημεία εξόδου) και τον άξονα που είναι η έξοδος του νευρώνα και το μέσο σύνδεσης του με άλλους νευρώνες. Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζεται σύναψη.

- Δενδρίτες (dendrites), αποτελούν τις πύλες εισόδου του νευρώνα καθώς είναι οι θέσεις από τις οποίες το κύτταρο λαμβάνει ερεθίσματα από το εξωτερικό του περιβάλλον, μέσω των νευροδιαβιβαστών. Πρόκειται για λεπτές επεκτάσεις του κυττάρου που αναπτύσσονται σε περίπλοκες διακλαδωμένες δομές τα λεγόμενα δενδριτικά δέντρα, και που κάθε μία δέχεται ένα ερέθισμα από διαφορετική πηγή. Χάρη στους δενδρίτες κάθε νευρώνας έχει τη δυνατότητα να δέχεται πληροφορίες από πολλές πηγές ταυτόχρονα. Το ερέθισμα που δέχεται ένας δενδρίτης, αποτελεί μια διαφορά δυναμικού που προκάλεσε τον νευρώνα-πομπό να ενεργοποιηθεί και να αποστείλει τους νευροδιαβιβαστές προς το νευρώνα-δέκτη. Το ερέθισμα αυτό δεν ισούται με την μεταβολή του δυναμικού που συμβαίνει στο εσωτερικό του νευρώνα, συνεπώς μπορεί να θεωρηθεί ότι στους δενδρίτες το ερέθισμα που λαμβάνεται πολλαπλασιάζεται με έναν συντελεστή. Το σήμα που στέλνει ένας δενδρίτης στο σώμα του νευρώνα, μπορεί να είναι είτε ενισχυτικό δηλαδή να τείνει να αυξήσει το ηλεκτρικό δυναμικό στο εσωτερικό του κυττάρου σε σχέση με το περιβάλλον του και συνεπώς να οδηγήσει στην ενεργοποίηση του νευρώνα, είτε παρεμποδιστικό δηλαδή να μειώνει το ηλεκτρικό δυναμικό του κυττάρου και να αποτρέπει δηλαδή την ενεργοποίηση του νευρώνα.
- Το Σώμα του κυττάρου. Εκεί βρίσκεται ο πυρήνας και άρα περιέχει το γενετικό υλικό του οργανισμού, όπως και όλα τα οργανίδια που έχουν και τα άλλα κύτταρα. Στο σώμα καταλήγουν όλα τα δενδριτικά δέντρα ενός νευρώνα, και σε εκείνο το σημείο συγκεντρώνονται και αθροίζονται όλα τα ερεθίσματα, δημιουργώντας το τελικό σήμα που δέχεται το κύτταρο. Η τιμή αυτή είναι

καθοριστική για το αν ενεργοποιηθεί ο νευρώνας ή όχι, γεγονός που συμβαίνει όταν η τιμή αυτή ξεπεράσει μια συγκεκριμένη τιμή-κατώφλι (threshold) διαφορετική για κάθε νευρώνα. Σε καταστάσεις αδράνειας (αρνητική τάση) ή όταν η τάση αυξάνεται ελάχιστα εκπέμπονται από ασθενείς ηλεκτρικοί παλμοί ανάλογοι της τάσης, οι οποίοι δεν περιέχουν κάποια πληροφορία. Αντίθετα, αν η τάση αυτή ξεπεράσει την τιμή που έχει το κατώφλι τότε ο νευρώνας «πυροδοτείται», εκπέμπεται δηλαδή ένας ισχυρός παλμός ανεξάρτητος της τάσης που ενεργοποιεί την αποστολή πληροφοριών προς άλλους νευρώνες. Η ύπαρξη του κατωφλίου είναι απαραίτητη για να μην πραγματοποιείται τυχαία ενεργοποίηση του νευρώνα από διακυμάνσεις του κυτταρικού δυναμικού, αν δεν υπάρχει πραγματικά πληροφορία.

- Άξονας, μέσω του οποίου μεταδίδονται τα εξερχόμενα σήματα από το σώμα του κυττάρου σε άλλους νευρώνες. Αποτελεί μια λεπτή, κυλινδρική απόφυση, η οποία ξεκινάει από το σώμα του νευρώνα και καταλήγει στις συνάψεις, οι οποίες όπως θα αναφερθεί παρακάτω αποτελούν ένα πλέγμα απολήξεων. Μέσω του άξονα, μεταδίδονται τα σήματα μεταξύ των νευρώνων και χάρη στη δομή του διατηρείται αναλλοίωτο το πληροφοριακό περιεχόμενο των ισχυρών παλμών. Οι ασθενείς παλμοί που μεταδίδονται με ελαττούμενο πλάτος κατά μήκος του άξονα δεν καταφέρνουν να φτάσουν μέχρι τις συνάψεις.
- Συνάψεις, αποτελούν τις δομές που βρίσκονται στις απολήξεις του άξονα και ευθύνονται για τη μετάδοση πληροφορίας μεταξύ των κυττάρων. Κάθε σύναψη συνδέεται, μέσω ενός μικρού διακυτταρικού χώρου, με έναν δενδρίτη ενός άλλου κυττάρου ή με κάποιο όργανο απόκρισης, πχ ένα μύ. Στις συνάψεις υπάρχουν κυστίδια μέσα στο οποία είναι αποθηκευμένοι νευροδιαβιβαστές. Όταν αντιληφθούν κάποιο ισχυρό σήμα, απελευθερώνονται οι νευροδιαβιβαστές και προσκολλώνται στον επόμενο δενδρίτη ή στο όργανο απόκρισης, μεταφέροντας έτσι την πληροφορία σε ένα πλήθος κυττάρων.

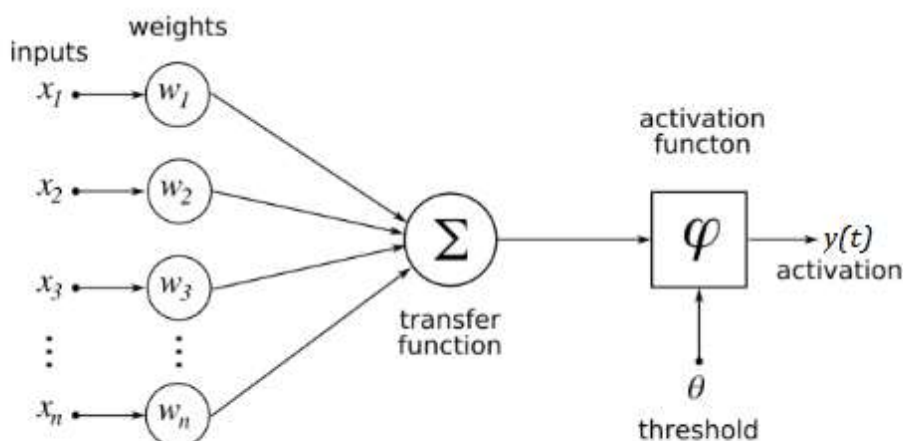
4.2.2. Μαθηματική Μοντελοποίηση Νευρώνα

Όπως αναφέρθηκε στην παράγραφο 4.1 με τα τεχνητά νευρωνικά δίκτυα επιχειρείται η προσέγγιση της λειτουργίας του ανθρώπινου εγκεφάλου και συνεπώς επιχειρείται η μοντελοποίηση των βιολογικών νευρώνων με χημικές συνάψεις που αναπτύχθηκαν στην παραπάνω ενότητα.[11]

Ένας τεχνητός νευρώνας αποτελεί ένα μαθηματικό μοντέλο που περιγράφεται από τη σχέση: [12]

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

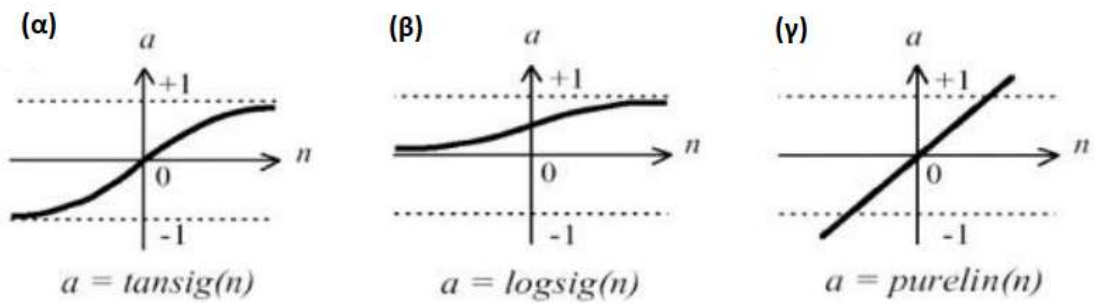
Για την διευκόλυνση στην κατανόηση των όρων της συνάρτησης, παρακάτω παρουσιάζεται και η σχηματική απεικόνιση ενός τεχνητού νευρώνα:[12]



Σχήμα 4. 2: Γραφική αναπαράσταση Τεχνητού Νευρώνα

όπου,

- $x_j(t)$, περιγράφει τα σήματα εισόδου (inputs), που προέρχονται είτε από το εξωτερικό περιβάλλον είτε ως σήματα εξόδου άλλων νευρώνων και το πλήθος τους δηλώνει η τιμή n.
- w_j , είναι τα συναπτικά βάρη (weights) και αποτελούν τους συντελεστές με τους οποίους πολλαπλασιάζονται τα σήματα εισόδου στους δενδρίτες. Όσο μεγαλύτερη κατά απόλυτη τιμή έχουν, τόσο σημαντικότερη είναι η συνεισφορά τους στην διαμόρφωση του τελικού αποτελέσματος του νευρώνα. Θετική τιμή ενός συναπτικού βάρους αντιστοιχεί σε ενισχυτικό σήμα, ενώ αρνητική σε παρεμποδιστικό.
- $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, είναι η συνάρτηση ενεργοποίησης (activation function) που καθορίζει αν θα ενεργοποιηθεί ή όχι ο νευρώνας. Μετατρέπει το σήμα εισόδου σε σήμα εξόδου, $y(t)$, και βασικός της στόχος στα πλαίσια της παρούσας εργασίας είναι να μοντελοποιεί και να προσεγγίζει την συμπεριφορά του υπό μελέτη δυναμικού συστήματος. Μια συνάρτηση ενεργοποίησης μπορεί να έχει διάφορες μορφές, κάποιες από αυτές είναι η γραμμική, ημι-γραμμική, λογαριθμική, σιγμοειδής ή εφαπτομενική. Στο Σχήμα 4.3 παρουσιάζονται τα διαγράμματα τριών από τις προαναφερθείσες συναρτήσεις ενεργοποίησης.



Σχήμα 4. 3: Γραφικές παραστάσεις των συναρτήσεων: α) Εφαπτομενική, β) Λογαριθμική, γ)Γραμμική

Στην περίπτωση των βιολογικών νευρώνων η συνάρτηση ενεργοποίησης είναι βηματική, εφόσον ο νευρώνας είτε θα ενεργοποιηθεί είτε όχι, δηλαδή το σήμα εξόδου είναι δυαδικό:

$$\varphi(u) = \begin{cases} 0, & u \leq 0 \\ 1, & u > 0 \end{cases} \quad (4.2)$$

- θ , συνάρτηση κατωφλιού (Threshold function), αποτελεί τον σταθερό όρο της συνάρτησης 4.1 και ορίζει το όριο που θα πρέπει να ξεπεράσει η τιμή των συναπτικών βαρών έτσι ώστε να ενεργοποιηθεί η συνάρτηση ενεργοποίησης φ . Η βηματική συνάρτηση 4.2 μπορεί να γραφεί και ως εξής:

$$\varphi(u) = \begin{cases} 0, & \sum_{j=1}^n w_j x_j(t) \leq -\theta \\ 1, & \sum_{j=1}^n w_j x_j(t) > -\theta \end{cases} \quad (4.3)$$

Παρατηρείται από την εξίσωση 4.3 ότι η συνάρτηση κατωφλιού θ , αποτελεί το αντίθετο από το κατώφλι των βιολογικών νευρώνων.

- $\mathbf{y}(t)$, είναι το σήμα εξόδου, και αποτελεί αποτέλεσμα που προκύπτει από την επεξεργασία όλων των σημάτων εισόδου του νευρώνα. Με χρήση του αλγορίθμου με τον οποίο έχει εκπαιδευτεί ο νευρώνας, λαμβάνει μια απόφαση και την μεταδίδει στους επόμενους νευρώνες.

Αν θεωρηθεί ότι ο σταθερός όρος, θ , της συνάρτησης φ , αποτελεί το συναπτικό βάρος ενός επιπλέον σταθερού σήματος εισόδου του $x_0(t) = 1$, τότε όλα τα σήματα εισόδου μπορούν να εκφραστούν ως ένα διάνυσμα του \mathbb{R}^{n+1} . [6]

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} 1 \\ x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad (4.4)$$

Και αντίστοιχα το διάνυσμα, \mathbf{w} , των συναπτικών βαρών ορίζεται στο πεδίο \mathbb{R}^{n+1} ως εξής:

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \theta \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \quad (4.5)$$

και άρα η συνάρτηση 4.1, μπορεί να γραφεί σε διανυσματική μορφή ως εξής:

$$\mathbf{y}(t) = \varphi(\mathbf{w}^T \mathbf{x}(t)) \quad (4.6)$$

4.3. Αρχιτεκτονική νευρωνικών δικτύων

Τα νευρωνικά δίκτυα χαρακτηρίζονται από την αρχιτεκτονική τους, τη λειτουργία που επιτελούν και την μέθοδο εκπαίδευσής τους. Η αρχιτεκτονική ενός δικτύου, καθορίζει τον αριθμό των νευρώνων από τους οποίους αποτελείται, τον τύπο των νευρώνων, καθώς και τη διάταξη των συνδέσεων τους. [11]

Σε αντίθεση με τα βιολογικά νευρωνικά δίκτυα που είναι εξαιρετικά πολύπλοκα, αρκεί να αναφερθεί ότι ο εγκέφαλος ενός νεογέννητου ανθρώπου αποτελείται από περίπου 100 δισεκατομμύρια νευρώνες κάθε ένας από τους οποίους συνδέεται με περίπου 1000 άλλους νευρώνες, τα τεχνητά νευρωνικά δίκτυα χρησιμοποιούν πολύ απλοποιημένα μοντέλα με πολύ μικρότερο αριθμό νευρώνων. Διατηρούν μόνο τα αδρά χαρακτηριστικά των λεπτομερών μοντέλων που χρησιμοποιούνται στη νευρολογία και συνεπώς έχουν ελάχιστη σχέση με τα βιολογικά νευρωνικά συστήματα. Αυτό συμβαίνει διότι οι υπολογιστές δεν μπορούν να υποστηρίξουν τόσο μεγάλο αριθμό νευρώνων καθώς και τέτοια απαίτηση σε υπολογιστική ισχύ. [11]

Οι νευρώνες ενός τεχνητού νευρωνικού δικτύου οργανώνονται σε μορφή παράλληλων επιπέδων (layers) και ο τρόπος με τον οποίο είναι δομημένοι σχετίζεται με τον αλγόριθμο εκμάθησης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Η παράλληλη διαμόρφωση των νευρώνων επιτρέπει την ταυτόχρονη επεξεργασία των ίδιων δεδομένων από πολλούς διαφορετικούς νευρώνες, οι οποίοι εξάγουν στην στην πλειοψηφία των περιπτώσεων διαφορετικό αποτέλεσμα, επιτρέποντας έτσι την ταχύτερη επεξεργασία πληροφορίας με τρόπο παρόμοιο με τα βιολογικά νευρωνικά δίκτυα.

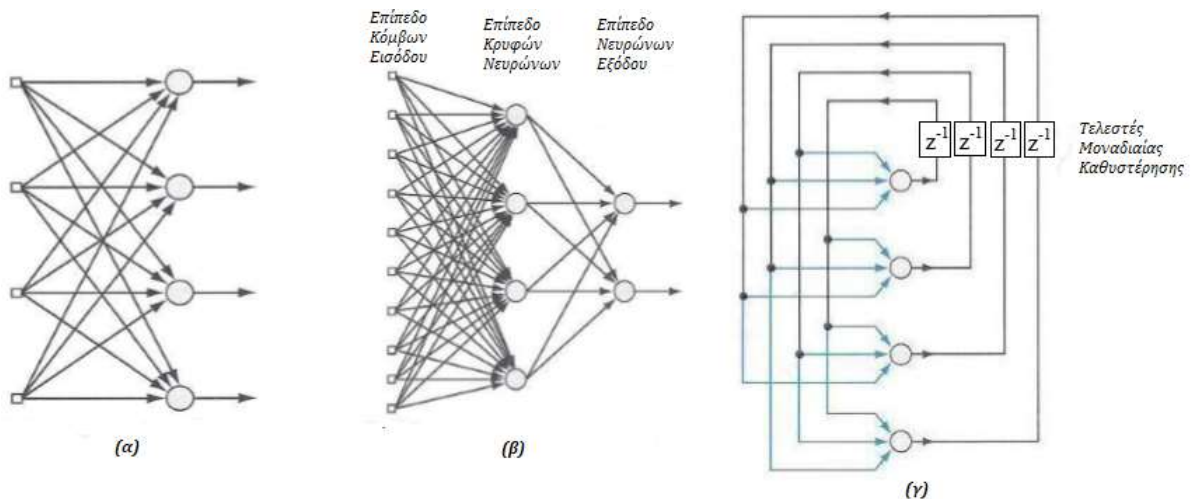
Τα επίπεδα των νευρωνικών δικτύων χωρίζονται σε τρεις βασικές κατηγορίες: [12]

- Επίπεδο εισόδου (input layer), αποτελεί το πρώτο επίπεδο ενός τεχνητού δικτύου και σε αυτό βρίσκονται τα δεδομένα εισόδου, τα οποία τροφοδοτούνται διαμέσου αυτού στο δίκτυο. Είναι το μόνο επίπεδο που δεν εμπεριέχει νευρώνες και συνεπώς δεν πραγματοποιείται κανένας υπολογισμός εντός αυτού. Κάθε μονάδα του επιπέδου αυτού ονομάζεται κόμβος εισόδου και αντιστοιχεί σε μια μεταβλητή εισόδου. Η διάσταση αυτού του επιπέδου, δηλαδή το πλήθος των κόμβων που περιέχει, ισούται πάντα με την διάσταση του διανύσματος εισόδου, δηλαδή το πλήθος των μεταβλητών εισόδου.

- Κρυφό επίπεδο (hidden layers), συνδέει το επίπεδο εισόδου με το επίπεδο εξόδου. Το πρώτο από τα κρυφά επίπεδα δέχεται όλες τις μεταβλητές εισόδου και ο κάθε νευρώνας του οποίου εξάγει μια ενδιάμεση μεταβλητή εξόδου, το σύνολο των οποίων εισέρχεται στο επόμενο επίπεδο. Αυτή η διαδικασία επαναλαμβάνεται ως το τελευταίο επίπεδο του δικτύου. Μέσω των συναπτικών βαρών τα επίπεδα συνδέονται πλήρως μεταξύ τους. Οι νευρώνες ενός κρυφού επιπέδου, όπως συνήθως και όλα τα κρυφά επίπεδα του συνολικού δικτύου, αποτελούνται από την ίδια συνάρτηση ενεργοποίησης. Ο αριθμός των κρυφών επιπέδων και η διάστασή τους, ορίζονται από τον σχεδιαστή και αποτελούν δύο από τις κρίσιμότερες σχεδιαστικές παραμέτρους ενός δικτύου. Αυτό συμβαίνει καθώς οι παράμετροι αυτοί καθορίζουν την πολυπλοκότητα και την ευελιξία του δικτύου, οπότε έμμεσα την αποτελεσματικότητα και την αποδοτικότητα του.
- Επίπεδο εξόδου (output layer) αποτελεί το τελευταίο επίπεδο ενός τεχνητού νευρωνικού δικτύου και δέχεται ως μεταβλητές εισόδου τις μεταβλητές εξόδου του τελευταίου κρυφού επιπέδου. Στο επίπεδο αυτό εμφανίζονται τα αποτελέσματα του δικτύου μετά το τέλος της συνολικής υπολογιστικής διαδικασίας. Όλοι οι νευρώνες του επιπέδου έχουν την ίδια συνάρτηση ενεργοποίησης, η οποία όμως στην πλειονότητα των περιπτώσεων διαφέρει από αυτήν των κρυφών επιπέδων. Στο επίπεδο εξόδου χρησιμοποιείται συνήθως η γραμμική συνάρτηση λόγω της απλότητας που την διακατέχει. Ο αριθμός των νευρώνων του είναι ίσος με τις πιθανές μεταβλητές εξόδου, δηλαδή η διάσταση του επιπέδου ταυτίζεται με αυτήν του διανύσματος εξόδου.

Ο αριθμός των επιπέδων ενός δικτύου καθώς και ο αριθμός των νευρώνων που περιέχει κάθε επίπεδο, παίζει σημαντικό ρόλο για ένα νευρωνικό δίκτυο.[12][15]

Ανάλογα με την δομή τους και τον τρόπο με τον οποίο συνδέονται οι νευρώνες μεταξύ τους, τα νευρωνικά δίκτυα διακρίνονται σε δύο ευρείες κατηγορίες, τα *δίκτυα πρόσθιας τροφοδότησης* (feed-forward networks) και τα *δίκτυα με ανατροφοδότηση* (feed-back networks) ή *αναδρομικά* (recurrent networks). [11]



Σχήμα 4. 4: Τεχνητά νευρωνικά δίκτυα: α) Πρόσθιας τροφοδότησης ενός επιπέδου. Οι κύκλοι δεξιά αναπαριστούν τεχνητούς νευρώνες (επίπεδο εξόδου), ενώ τα τετράγωνα αριστερά αναπαριστούν τους κόμβους εισόδου (επίπεδο εισόδου). Για τη σύνδεση μεταξύ 2 επιπέδων απαιτείται η ύπαρξη ενός συναπτικού βάρους, τα οποία αναπαρίστανται με τα βέλη που συνδέουν τους κόμβους εισόδου με τους νευρώνες εξόδου. β) Δίκτυο Πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο γ) Δίκτυο με ανατροφοδότηση , χωρίς αυτό-ανάδραση και χωρίς κρυφούς νευρώνες. Τα στοιχεία μοναδιαίας καθυστέρησης συμβολίζονται με z^{-1} .

4.4. Δίκτυα πρόσθιας τροφοδότησης

Σε αυτή την κατηγορία η πληροφορία μεταφέρεται από την είσοδο προς την έξοδο, μέσω των νευρώνων, χωρίς να μπορεί να συμβεί το αντίστροφο. Είναι στατικά δίκτυα και χωρίς μνήμη, αυτό σημαίνει ότι η απόκριση σε ένα σήμα εισόδου εξαρτάται αποκλειστικά από αυτό και δεν επηρεάζεται από προηγούμενα σήματα. Συνεπώς κάθε διάνυσμα εισόδου μπορεί αναπαράγει ένα διάνυσμα απόκρισης. Τα δίκτυα πρόσθιας τροφοδότησης μπορούν να διακριθούν σε δύο επιμέρους κατηγορίες, τα δίκτυα ενός επιπέδου και τα πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης, ανάλογα με τον αριθμό των κρυφών επιπέδων που έχει το κάθε δίκτυο. [11]

Τα δίκτυα πρόσθιας τροφοδότησης ενός επιπέδου, στα οποία το σήμα εισόδου μετατρέπεται απευθείας σε σήμα εξόδου, αποτελούν την πιο απλή μορφή ενός δικτύου και θα αναπτυχθούν περαιτέρω στη συνέχεια. [12] Στο Σχήμα 4.4(α) απεικονίζεται αναλυτικά η δομή ενός τέτοιου δικτύου.

Αν στο δίκτυο αυτό, το επίπεδο εισόδου αποτελείται από L_0 κόμβους εισόδου, και το επίπεδο εξόδου αποτελείται από L_1 νευρώνες, το σήμα εισόδου μετατρέπεται με τη βοήθεια της συνάρτησης ενεργοποίησης (4.1) και των συναπτικών βαρών σε L_1 διαφορετικά σήματα εξόδου $y_1(t), \dots, y_{L_1}(t)$.

Το σήμα εξόδου κάθε νευρώνα ορίζεται τότε ως εξής:

$$y_k(t) = \varphi \left(\sum_{j=1}^n w_{k,j} x_j(t) + \theta_k \right), \quad k = 1, \dots, L_1 \quad (4.7)$$

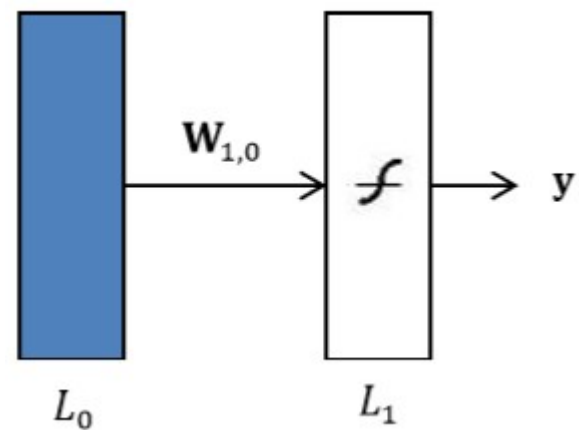
Όπου, $w_{k,j}$ είναι το συναπτικό βάρος που οδηγεί το j σήμα εισόδου στον k νευρώνα. Ενώ, η παράμετρος μεροληψίας του k νευρώνα είναι η $\theta_k = w_{k,0}$.

Η μετατροπή του σήματος εισόδου σε σήμα απόκρισης καθορίζεται πλήρως από τον πίνακα των συναπτικών βαρών, κάθε στήλη του οποίου αποτελεί το διάνυσμα συναπτικών βαρών του αντίστοιχου νευρώνα σύμφωνα με την σχέση (4.5). [12]

Ο πίνακας συναπτικών βαρών του νευρωνικού δικτύου που αναπτύχθηκε παραπάνω είναι ο εξής, όπου $\mathbf{w}_{1,0} \in \mathbb{R}^{(L_0+1) \times L_1}$

$$\mathbf{w}_{1,0} = \begin{bmatrix} w_{1,0} & w_{2,0} & \dots & w_{L_1,1} \\ w_{1,1} & w_{2,1} & \ddots & w_{L_1,2} \\ \vdots & \vdots & & \vdots \\ w_{1,L_0} & w_{2,L_0} & \dots & w_{L_1,L_0} \end{bmatrix} = [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_{L_1}] \quad (4.8)$$

Με τις αλλαγές αυτές το σχήμα 4.3(α) μπορεί να αναπαρασταθεί με μια πιο απλή μορφή, όπως φαίνεται στο Σχήμα 4.4. Το πρώτο παραλληλόγραμμο αποτελεί το επίπεδο εισόδου, το δεύτερο το επίπεδο εξόδου ενώ μεταξύ των δύο επιπέδων αναγράφεται ο πίνακας συναπτικών βαρών του δεύτερου επιπέδου. Κάτω από κάθε επίπεδο αναγράφεται η διάσταση του δηλαδή από πόσους κόμβους εισόδου, αν πρόκειται για το πρώτο ή πόσους νευρώνες αν πρόκειται για το δεύτερο, αποτελείται.



Για την δημιουργία ενός νευρωνικού δικτύου πρόσθιας τροφοδότησης πολλαπλών επιπέδων αρκεί να συνδεθούν περισσότερα του ενός δίκτυα πρόσθιας τροφοδότησης σε σειρά. Οπότε η έξοδος του ενός δικτύου αποτελεί είσοδο του επόμενου. Εκτός από το τελευταίο επίπεδο του οποίου η έξοδος αποτελεί την απόκριση του δικτύου, τα ενδιάμεσα επίπεδα αποτελούν τα κρυφά επίπεδα του δικτύου.

Σχήμα 4. 5: Τεχνητό νευρωνικό δίκτυο πρόσθιας τροφοδότησης ενός επιπέδου. Το σύμβολο εντός του δεύτερου παραλληλόγραμμου φανερώνει την συνάρτηση που έχουν οι νευρώνες του επιπέδου. Σε αυτήν την περίπτωση είναι η σιγμοειδής.

4.5. Δίκτυα με Ανατροφοδότηση - Δίκτυο NARX

Σε αυτή την κατηγορία τα νευρωνικά δίκτυα περιέχουν τουλάχιστον έναν βρόγχο ανατροφοδότησης ή αλλιώς βρόγχο ανάδρασης (feedback loop). Η έξοδος κάθε νευρώνα του δικτύου ανατροφοδοτεί την είσοδο των άλλων του ίδιου επιπέδου και σε κάποιες περιπτώσεις ακόμα και τη δική του είσοδο (αυτό-ανάδραση). Η παρουσία βρόγχων ανατροφοδότησης επηρεάζει την δυνατότητα μάθησης του δικτύου και την απόδοσή του. Οι κλάδοι των παραπάνω βρόγχων περιέχουν στοιχεία μοναδιαίας χρονικής καθυστέρησης που συμβολίζεται με z^{-1} . [11]

Το δίκτυο αυτό αποτελεί ένα δυναμικό σύστημα, 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 για την δημιουργία ενός διανύσματος απόκρισης

χρησιμοποιούνται τόσο σήματα εισόδου όσο και διανύσματα απόκρισης από προηγούμενες στιγμές.

Στην παρούσα εργασία χρησιμοποιείται το δίκτυο NARX, (Nonlinear Auto Regressive with exogenous inputs model), που αποτελεί ένα πολύ ισχυρό δίκτυο της κατηγορίας νευρωνικών δικτύων με ανατροφοδότηση. Η ισχύς του δικτύου οφείλεται στην απλή δομή που έχει συγκριτικά με τα υπόλοιπα δίκτυα της κατηγορίας του, καθώς και στην ταχύτερη σύγκλιση και την παράλληλα καλύτερη γενίκευση που κάνει. [12],[13]

Το δίκτυο NARX προσεγγίζει σχέσεις τις μορφής:

$$\mathbf{y}(t_k) = \boldsymbol{\varphi}(\mathbf{x}(t_{k-n_x}), \dots, \mathbf{x}(t_{k-1}), \mathbf{x}(t_k), \mathbf{y}(t_{k-n_y}), \dots, \mathbf{y}(t_{k-1})) \quad (4.9)$$

Όπου $\mathbf{y}(t)$ είναι τα σήματα εξόδου, $\mathbf{x}(t)$ τα σήματα εισόδου, ενώ η συνάρτηση $\boldsymbol{\varphi}$ αποτελεί μια οποιαδήποτε μη γραμμική συνάρτηση που μπορεί να προσεγγιστεί από ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης πολλαπλών επιπέδων. [16][13]

Το δίκτυο αυτό, έχει δηλαδή την δυνατότητα να υπολογίζει μια εκτιμώμενη τιμή της ακολουθίας απόκρισης \mathbf{y} , για μια δεδομένη χρονική στιγμή t_k , που συμβολίζεται με $\hat{\mathbf{y}}(t_k)$, χρησιμοποιώντας το διάνυσμα εισόδου της χρονικής αυτής στιγμής, προηγούμενων χρονικών στιγμών n_x , καθώς και τα διανύσματα εξόδου από προηγούμενες χρονικές στιγμές n_y .

Η εφαρμογή του δικτύου NARX πραγματοποιήθηκε μέσω της εργαλειοθήκης του Matlab, η οποία χρησιμοποιεί μια παραλλαγμένη σχέση του ορισμού (4.9), απ' όπου απουσιάζει ο όρος $\mathbf{x}(t_k)$.

$$\mathbf{y}(t_k) = \boldsymbol{\varphi}(\mathbf{x}(t_{k-n_x}), \dots, \mathbf{x}(t_{k-1}), \mathbf{y}(t_{k-n_y}), \dots, \mathbf{y}(t_{k-1})) \quad (4.10)$$

Οι δύο αυτές σχέσεις δεν διαφέρουν ουσιαστικά, καθώς η σχέση (4.10) προκύπτει από την (4.9) εάν μετατοπιστεί η ακολουθία κατά μια χρονική στιγμή στο παρελθόν, δηλαδή μετατραπεί σε $\mathbf{x}(t_{k+1})$. Το δίκτυο NARX επιτρέπει, σε σχέση με άλλα δίκτυα, καλύτερη προσέγγιση σχέσεων με μακροπρόθεσμη εξάρτηση, όπου δηλαδή η απόκριση εξαρτάται από διανύσματα εισόδου και αποκρίσεις πολλές στιγμές στο παρελθόν. [23] Το πρόβλημα που μελετάται στην παρούσα εργασία και όπως αυτό ορίστηκε στο Κεφάλαιο 2, δεν παρουσιάζει τέτοιου είδους εξάρτηση. Το συμπέρασμα αυτό προκύπτει από την μορφή που έχουν οι εξισώσεις (2.1)-(2.4) καθώς και τη διακριτοποίηση που γίνεται στον χρόνο ώστε να επιλυθούν. Για τον λόγο αυτό επιλέγεται για την επίλυση της εξίσωσης (4.10) $n_x = n_y = 1$, και συνεπώς η σχέση παίρνει την εξής μορφή:

$$\mathbf{y}(t_k) = \boldsymbol{\varphi}(\mathbf{x}(t_{k-1}), \mathbf{y}(t_{k-1})) \quad (4.11)$$

Από την μορφή της εξίσωσης προκύπτει ότι κάθε διάνυσμα απόκρισης εξαρτάται μόνο από τα διανύσματα εισόδου και εξόδου την ακριβώς προηγούμενη στιγμή.

Η αρχιτεκτονική ενός δικτύου NARX, παρουσιάζει δύο εκδοχές: [6][12][15]

- Τη σειριακή- παράλληλη αρχιτεκτονική, η οποία χρησιμοποιείται κατά την εκπαίδευση του δικτύου. Στην αρχιτεκτονική αυτή τοποθετούνται στο πρώτο επίπεδο νευρώνων γνωστές τιμές της ακολουθίας y , ενώ στην έξοδο υπολογίζονται εκτιμήσεις του διανύσματος \hat{y} . Η αρχιτεκτονική αυτή καθιστά στην ουσία εφικτό τον υπολογισμό μόνο της ακριβώς επόμενης ακολουθίας y και αποτελεί στην πραγματικότητα αρχιτεκτονική δικτύου πρόσθιας τροφοδότησης.
- Την παράλληλη αρχιτεκτονική ή αρχιτεκτονική με ανατροφοδότηση, η οποία χρησιμοποιείται κατά την λειτουργία του δικτύου, όταν δεν υπάρχουν γνωστές τιμές του διανύσματος απόκρισης. Στο πρώτο επίπεδο νευρώνων τροφοδοτούνται οι τιμές του γνωστού διανύσματος x με τα αντίστοιχα συναπτικά βάρη, καθώς και οι τιμές της εκτίμησης της απόκρισης \hat{y} του ακριβώς προηγούμενου χρονικού βήματος. Με αυτόν τον τρόπο είναι εφικτή η πρόβλεψη των τιμών της απόκρισης y για χρονικές στιγμές μεγαλύτερες του ενός βήματος. Κατά την διαδικασία αυτή, η εκτίμηση που υπολογίζεται στην k επανάληψη της μεθόδου εμπεριέχει k φορές το σφάλμα της συνάρτησης φ της πρόσθιας τροφοδότησης, το οποίο περιέχεται στην εκτίμηση της απόκρισης \hat{y} που ανατροφοδοτείται. Γίνεται λοιπόν κατανοητό η διαδικασία αυτή εγκυμονεί τον κίνδυνο συσσώρευσης του σφάλματος της προσέγγισης \hat{y} σε κάθε βήμα. Για να αποφευχθεί αυτό θα πρέπει ο υπολογισμός της συνάρτησης φ να γίνει με το μικρότερο δυνατό σφάλμα, πράγμα που εξασφαλίζεται με την διαδικασία εκπαίδευσης του δικτύου.

4.6. Εκπαίδευση νευρωνικού δικτύου

Η διαδικασία εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου ξεκινά χωρίς αυτό να έχει καμία γνώση – πληροφορία του συστήματος που πρέπει να προσεγγίσει. Κατά την διαδικασία εκπαίδευσης, το δίκτυο λαμβάνει τα δεδομένα που του δίνονται, όπως είναι τα δεδομένα εισόδου – εξόδου, ή διάφορα πρότυπα, και αποθηκεύει την πληροφορία αυτή στον πίνακα συναπτικών βαρών, ώστε να μπορεί στη συνέχεια να προσεγγίσει την επιθυμητή σχέση. Συνεπώς ως εκπαίδευση ορίζεται η διαδικασία κατά την οποία πραγματοποιείται προσδιορισμός της τιμής κάθε συναπτικού βάρους, το σύνολο των οποίων συνδέει τα επίπεδα του δικτύου μεταξύ τους.

Ο υπολογισμός των πινάκων των συνοπτικών βαρών $W_{i,j}$, μεταξύ όλων των επιπέδων ενός νευρωνικού δικτύου πραγματοποιείται συνήθως με χρήση κάποιου επαναληπτικού αλγόριθμου εκπαίδευσης. Οι αρχικές τιμές των συναπτικών βαρών είναι τυχαίες και δεν έχουν καμία σημασία, ενώ οι τιμές μετά την εκπαίδευση υποδεικνύουν τη σχέση που συνδέει την είσοδο με την έξοδο του δικτύου. Υπάρχουν όμως και ελάχιστες περιπτώσεις στις οποίες ο πίνακας συναπτικών βαρών υπολογίζεται απευθείας, ο τρόπος αυτός αποφεύγεται διότι είναι μεγαλύτερες οι πιθανότητες το νευρωνικό δίκτυο να μην προσεγγίζει με μεγάλη ακρίβεια το μελετώμενο σύστημα. [11][12][15]

Οι αλγόριθμοι εκπαίδευσης χωρίζονται σε τρεις ευρείες κατηγορίες που καθορίζουν και τον τρόπο με τον οποίο επιτελείται ο τρόπος εκμάθησης των νευρωνικών δικτύων.

- Η επιβλεπόμενη εκμάθηση (supervised training), αποτελεί την πρώτη κατηγορία, και κατά την διάρκεια της θα πρέπει να τροφοδοτούνται στο δίκτυο κάποια "πρότυπα εκπαίδευσης". Τα πρότυπα αυτά, αποτελούν το επιθυμητό αποτέλεσμα του προβλήματος που καλείται να προσεγγίσει το νευρωνικό δίκτυο. Με χρήση αυτών των παραδειγμάτων, καθώς και με την διαφορά της πραγματικής από την υπολογιζόμενη τιμή εξόδου σε κάθε επανάληψη, το δίκτυο προσαρμόζει τα συναπτικά βάρη έτσι ώστε να επιτευχθεί σταδιακή η μείωση του σφάλματος της προσέγγισης του δικτύου σε σχέση με την πραγματική τιμή που αποτελεί την είσοδο του δικτύου. Η διαδικασία αυτή επαναλαμβάνεται μέχρις ότου το σφάλμα φτάσει μια αποδεκτή τιμή ή και μέχρι να εξαλειφθεί εντελώς.
- Η μη -επιβλεπόμενη εκμάθηση (unsupervised training), κατά την κατηγορία αυτή δεν τροφοδοτούνται στο δίκτυο γνωστά πρότυπα ή αποτελέσματα, αλλά πρέπει το δίκτυο να συλλέξει πληροφορίες μέσα από στατιστικά μεγέθη και πρότυπα των δεδομένων και να υλοποιήσει έτσι την αναπροσαρμογή των συναπτικών βαρών. Στην περίπτωση αυτή το δίκτυο καλείται να προσεγγίσει ένα τελείως άγνωστο για αυτό αποτέλεσμα.
- Ενισχυμένη εκμάθηση (reinforced training), όπως και στην προηγούμενη μέθοδο έτσι και σε αυτή, δεν τροφοδοτούνται στο δίκτυο τα επιθυμητά αποτελέσματα που καλείται να προσεγγίσει. Η μέθοδος αυτή χρησιμοποιεί για κάθε εξαγόμενο αποτέλεσμα ένα ενισχυτικό σήμα, που αποτελείται από το σύνολο των μέτρων επάρκειας του κάθε αποτελέσματος, το οποίο ωθεί το δίκτυο προς το επιθυμητό αποτέλεσμα, "επιβραβεύοντας" το για τις ορθές συμπεριφορές του και "τιμωρώντας" το για τις λανθασμένες.

Ο αλγόριθμος εκπαίδευσης που χρησιμοποιήθηκε στην παρούσα εργασία μέσω του εργαλείου εκπαίδευσης νευρωνικών δικτύων της Matlab, είναι ο αλγόριθμος Levenberg-Marquardt και ανήκει στην κατηγορία των αλγορίθμων επιβλεπόμενης μάθησης. Χρησιμοποιεί ως συνάρτηση για τον υπολογισμό του σφάλματος μεταξύ της προσέγγισης του δικτύου και της πραγματικής τιμής, το άθροισμα τετραγώνων των σφαλμάτων των αποκρίσεων:

$$V(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 \quad (4.12)$$

Και στην περίπτωση αυτή η εκτίμηση της απόκρισης $\hat{\mathbf{y}}$ προκύπτει από τη σχέση (4.11):

$$\hat{\mathbf{y}}(t_k) = \boldsymbol{\varphi}_N(\mathbf{x}(t_{k-1}), \mathbf{y}(t_{k-1})) \quad (4.13)$$

Το εύρος του πεδίου ορισμού της επιθυμητής συνάρτησης, αποτελεί έναν από τους παράγοντες που επηρεάζουν την διαδικασία και την ποιότητα εκπαίδευσης του δικτύου. Όσο μεγαλύτερο είναι τόσο περισσότερα δεδομένα εκπαίδευσης απαιτούνται, από διάφορες περιοχές του πεδίου ορισμού, καθώς η προσέγγιση είναι καλύτερη στις περιοχές απ' όπου έχουν ληφθεί τα περισσότερα δεδομένα. Είναι

σημαντικό να τονισθεί ότι το νευρωνικό δίκτυο δεν έχει καμία γνώση της συνάρτησης κι έτσι η γενίκευση του ισχύει μόνο εντός του πεδίου των δεδομένων εκπαίδευσης του. Ένας ακόμη παράγοντας είναι η *μορφή της επιθυμητής συνάρτησης*. Όσο πιο περίπλοκη είναι, δηλαδή εάν παρουσιάζει περιοχές ασυνέχειας, πολλαπλότητα στον χώρο λύσεων κ.α., τόσο πιο δύσκολη είναι η σωστή προσέγγιση της συμπεριφοράς του συστήματος από το εκπαιδευόμενο δίκτυο. Τέλος, η *ποιότητα των δεδομένων* που χρησιμοποιούνται στην εκπαίδευση του δικτύου παίζουν σημαντικό ρόλο στην διαδικασία εκπαίδευσης καθώς και στην τελική συνάρτηση. Συχνά συναντώνται περιπτώσεις στις οποίες τα δεδομένα περιέχουν πειραματικά σφάλματα, θορύβους αποτελεσμάτων κ.α.

Η αποτελεσματικότητα των δεδομένων εκπαίδευσης κρίνεται με βάση τα εξής κριτήρια:

- Ποσότητα: Απαιτείται αρκετή ποσότητα δεδομένων έτσι ώστε να καλύπτεται όλο το εύρος των μελετώμενων παραμέτρων.

- Ποιότητα: Σχετίζεται με το πόσα διαφορετικά πρότυπα εμπεριέχονται στα δεδομένα και τον αριθμό των δεδομένων που απαιτείται για την ικανοποιητική εκπαίδευση του δικτύου. Πολύ μεγάλος αριθμός δεδομένων μπορεί να προκαλέσει το φαινόμενο της υπερπροσαρμογής, το οποίο θα αναπτυχθεί στη συνέχεια.

- Υπολογιστική πολυπλοκότητα: Επηρεάζεται σημαντικά από την δομή και τον αλγόριθμο εκπαίδευσης του δικτύου, και δηλώνει τον υπολογιστικό χρόνο που απαιτείται για να ολοκληρωθεί η εκπαίδευση του δικτύου.

Στην παρούσα εργασία χρησιμοποιούνται δεδομένα από προσομοίωση πειράματος και όχι από πραγματικό πείραμα, συνεπώς μπορεί να ληφθεί μεγάλος αριθμός αποτελεσμάτων, χωρίς την ύπαρξη πειραματικών σφαλμάτων. Αυτό δεν παύει όμως να σημαίνει ότι δεν θα εμφανιστεί κανένα είδος σφαλμάτων στα δεδομένα, καθώς μπορεί να εμφανιστούν κάποια που οφείλονται στην ίδια την προσομοίωση. Τέτοια σφάλματα μπορεί να είναι η επάρκεια του πλέγματος, η επίτευξη ή μη της σύγκλισης, που είναι σφάλματα πολύ μικρότερης τάξης από τα πειραματικά. [12]

Η ορθή επιλογή κάποιων εκ των βασικών παραγόντων, κατά την εκπαίδευση ενός νευρωνικού δικτύου παίζει σημαντικό ρόλο για την αποφυγή της εμφάνισης του φαινομένου υπερπροσαρμογής. Το φαινόμενο αυτό παρατηρείται στην περίπτωση κατά την οποία δημιουργούνται πολυώνυμα μεγάλου βαθμού για να περιγράψουν φαινόμενα πολύ μικρότερης τάξης. Κατά το φαινόμενο αυτό το νευρωνικό δίκτυο αδυνατεί να γενικεύσει, και παρουσιάζει μεγάλη ακρίβεια μόνο εντός των δεδομένων εκπαίδευσης. Οι παράγοντες που καθορίζουν την εμφάνιση αλλά και την έκταση του φαινομένου της υπερπροσαρμογής είναι πολλοί, οι κρίσιμότεροι είναι όμως α) η ποσότητα και η ποιότητα των δεδομένων εκπαίδευσης, β) ο αριθμός των κρυφών επιπέδων και η διάστασή τους, γ) η μέθοδος εκπαίδευσης που εφαρμόζεται. [6][15]

Ο αριθμός των κρυφών επιπέδων που περιέχει ένα νευρωνικό δίκτυο εξαρτάται από την πολυπλοκότητα του συστήματος που θα πρέπει να προσεγγίσει. Η επιλογή του

αριθμού κρυφών επιπέδων θέλει ιδιαίτερη προσοχή, διότι αν επιλεγθεί μεγάλος αριθμός κρυφών επιπέδων είναι πολύ πιθανό να εμφανιστεί το φαινόμενο υπερπροσαρμογής. Σημαντικό ρόλο για την αποφυγή του φαινομένου, παίζει επίσης και το πλήθος των νευρώνων του κρυφού επιπέδου, L , το οποίο έχει άμεση σχέση με το πλήθος των δεδομένων εκπαίδευσης, καθώς και το πόσο διαφορετικά είναι μεταξύ τους. Όπως έχει αναφερθεί παραπάνω, οι μόνοι άγνωστοι που πρέπει να προσδιοριστούν κατά την εκπαίδευση ενός δικτύου είναι τα συναπτικά βάρη. Όσο περισσότεροι είναι οι κρυφοί νευρώνες, τόσο περισσότερες είναι και οι συνάψεις και τα συναπτικά βάρη, συνεπώς:

- Αν οι κρυφοί νευρώνες, L , είναι πολύ λίγοι τότε το πλήθος των συναπτικών βαρών δεν επαρκεί για την περιγραφή της επιθυμητής συνάρτησης.
- Αν οι κρυφοί νευρώνες είναι υπερβολικά πολλοί, τότε το σύστημα εξισώσεων είναι ακριβώς ορισμένο ή υποορισμένο, υπάρχουν δηλαδή ίσες ή περισσότερες εξισώσεις από συναπτικά βάρη. Στην περίπτωση αυτή το δίκτυο έχει εκπαιδευτεί ώστε να περιγράφει με μεγάλη ακρίβεια την συμπεριφορά των δεδομένων εκπαίδευσης, αλλά αδυνατεί να γενικεύσει.
- Η επιθυμητή κατάσταση για την ικανοποιητική λειτουργία ενός δικτύου, επιτυγχάνεται μεταξύ των δύο παραπάνω άκρων. Στην κατάσταση αυτή τα δεδομένα εκπαίδευσης περιγράφονται ικανοποιητικά, ενώ το δίκτυο έχει επιπλέον την δυνατότητα να γενικεύσει επιτυχώς τα δεδομένα αυτά, δηλαδή οι προβλέψεις του δικτύου να ακολουθούν την επιθυμητή συνάρτηση.

Το νευρωνικό δίκτυο που εκπαιδεύεται στην παρούσα εργασία είναι το NARX σε σειριακή -παράλληλη αρχιτεκτονική ενώ η λειτουργία του πραγματοποιείται με ανατροφοδότηση. Η διαδικασία ελέγχου του δικτύου εφαρμόζεται στην λειτουργία με ανατροφοδότηση, στην οποία εισέρχονται τα δεδομένα εισόδου και στη συνέχεια παράγονται οι εκτιμήσεις των αποκρίσεων \hat{y} , που συγκρίνονται με τη σειρά τους με τις τιμές των γνωστών αποκρίσεων εξόδου y . Είναι σημαντικό να αναφερθεί ότι κατά την εκπαίδευση του νευρωνικού δικτύου δεν απαιτούνται τα ίδια δεδομένα, ούτε το ίδιο πλήθος δεδομένων εκπαίδευσης που χρησιμοποιούνται για την κατασκευή της PCA βάσης, αρκεί όμως να ανήκουν όλα στο ίδιο εύρος τιμών από το οποίο επιλέγονται.

4.7. Εφαρμογή

Η προσέγγιση της χρονικής εξέλιξης των συντελεστών $\alpha(t)$ του μοντέλου μειωμένης τάξης, όπως έχει αναφερθεί στην Παράγραφο 3.3, επιτυγχάνεται μέσω των νευρωνικών δικτύων.

Στο νευρωνικό δίκτυο που εκπαιδεύεται, η παροχή $\hat{m}(t)$, (Παράγραφος 2.4) αποτελεί την ανεξάρτητη μεταβλητή εισόδου, ενώ η κατάσταση $\mathbf{x}(t)$ την εξαρτημένη μεταβλητή εξόδου. Με την εφαρμογή της μείωσης τάξης, η χρονική εξέλιξη της κατάστασης $\mathbf{x}(t)$ υπολογίζεται με την προσέγγιση της χρονικής εξέλιξης των συντελεστών $\alpha(t)$, συνεπώς η εξαρτημένη μεταβλητή την οποία καλείται να προβλέψει το δίκτυο NARX σε βάθος χρόνου είναι αυτή των συντελεστών $\alpha(t)$.

Το μοντέλο μειωμένης τάξης που κατασκευάστηκε αποτελείται από την ανάλυση κυρίων συνιστωσών και την εκπαίδευση του τεχνητού νευρωνικού δικτύου. Έστω μια άγνωστη τροχιά \mathbf{Y} του συστήματος, η οποία διαρκεί m χρονικά βήματα και για την οποία είναι γνωστά μόνο η αρχική κατάσταση $\mathbf{x}(t_0)$ άρα και οι αρχικοί συντελεστές $\alpha(t_0)$ καθώς και το διάνυσμα της παροχής εισόδου $\dot{\mathbf{m}}(t)$ σε κάθε χρονική στιγμή. Αν στο σύστημα επιβληθεί μια βηματική επιβολή στην μαζική παροχή από $\dot{\mathbf{m}}(t_0)$ σε $\dot{\mathbf{m}}(t)$, τότε στόχος είναι η εύρεση της εξέλιξης της μεταβλητής κατάστασης από $\mathbf{x}(t_0)$ σε $\mathbf{x}(t)$, δηλαδή η εύρεση της χρονικής εξέλιξης της μεταβλητής κατάστασης σε όλο τον όγκο του αντιδραστήρα καθώς μεταβαίνει από μια αρχική κατάσταση ισορροπίας σε μια νέα.

Το νευρωνικό δίκτυο υπολογίζει μια προσέγγιση των συντελεστών $\alpha(t)$, με τη μορφή του εξής διανύσματος:

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

Σύμφωνα με τη σχέση (3.40), το παραπάνω διάνυσμα σε συνδυασμό με την ορθοκανονική βάση \mathbf{Z} που κατασκευάστηκε στην προηγούμενη ενότητα, μπορούν να δώσουν μια προσέγγιση του πίνακα τροχιών \mathbf{Y} .

4.8. Αξιολόγηση νευρωνικού δικτύου

Μετά την ολοκλήρωση της εκπαίδευσης του δικτύου, καλείται να αξιολογηθεί ως προς την ακρίβεια στην προσέγγιση του πλήρους συστήματος, με μόνη γνωστή μεταβλητή την μεταβλητή εκ χειρισμού, που είναι η μαζική παροχή $\dot{\mathbf{m}}(t)$. Εφόσον τα δεδομένα εκπαίδευσης του δικτύου είναι γνωστά από την προσομοίωση του αντιδραστήρα, οι εκτιμήσεις που παράγει το νευρωνικό δίκτυο μπορούν να συγκριθούν με τα δεδομένα του μοντέλου πλήρους τάξης, όπως και με αυτά του μοντέλου μειωμένης τάξης.

Το σχετικό σφάλμα της προσέγγισης του νευρωνικού δικτύου ως προς το μοντέλο πλήρους τάξης, για μια δεδομένη χρονική στιγμή (snapshot), υπολογίζεται από τη σχέση:

$$e_{err,full}(t) = \frac{\|\hat{\mathbf{x}}(t) - \mathbf{x}(t)\|}{\|\mathbf{x}(t)\|} \quad (4.15)$$

όπου, $\hat{\mathbf{x}}(t)$ αποτελεί την πρόβλεψη του δικτύου την στιγμή t , ενώ $\mathbf{x}(t)$ είναι το διάνυσμα κατάστασης του πλήρους τάξης μοντέλου τη στιγμή αυτή.

Το σχετικό σφάλμα της προσέγγισης του νευρωνικού δικτύου ως προς το μοντέλο μειωμένης τάξης, για ένα στιγμιότυπο την χρονική στιγμή t , είναι:

$$e_{err,pod}(t) = \frac{\|\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t)\|}{\|\hat{\mathbf{x}}(t)\|} \quad (4.16)$$

όπου $\hat{\mathbf{x}}(t)$ η προσέγγιση του μοντέλου μειωμένης τάξης, για την χρονική στιγμή t .

Χρήσιμοι τύποι σχετικών σφαλμάτων για την αξιολόγηση του δικτύου είναι επίσης το μέσο τετραγωνικό σχετικό σφάλμα (4.17) και το μέγιστο σχετικό σφάλμα (4.18) ως προς την προσέγγιση του μοντέλου πλήρους τάξης:

$$mse_{err,full} = \left(\frac{1}{m} \sum_{j=1}^m e_{rr,full}^2(t) \right)^{\frac{1}{2}} \quad (4.17)$$

$$maxe_{err,full}(t) = \max(e_{rr,full}(t)) \quad , j = 1, \dots, m \quad (4.18)$$

Όπως και τα αντίστοιχα σχετικά σφάλματα της πρόβλεψης του δικτύου ως προς το μοντέλο μειωμένης τάξης:

$$mse_{err,pod} = \left(\frac{1}{m} \sum_{j=1}^m e_{rr,pod}^2(t) \right)^{\frac{1}{2}} \quad (4.19)$$

$$maxe_{err,pod}(t) = \max(e_{rr,pod}(t)) \quad , j = 1, \dots, m \quad (4.20)$$

5. Αποτελέσματα

5.1 Εισαγωγή

Όπως αναφέρθηκε στην Παράγραφο 3.7, στην παρούσα εργασία πραγματοποιείται η κατασκευή ενός Μοντέλου Μειωμένης Τάξης, με το οποίο καθίσταται εφικτή η μελέτη της δυναμικής συμπεριφοράς του μελετώμενου συστήματος, χρησιμοποιώντας δεδομένα μη μόνιμων καταστάσεων. Στόχος είναι η κατασκευή μιας Ορθοκανονικής Βάσης και η εκπαίδευση ενός Τεχνητού Νευρωνικού Δικτύου, που να πληρούν τα εξής χαρακτηριστικά:

- Προσέγγιση του πλήρους συστήματος με μεγάλη ακρίβεια
- Υψηλή ταχύτητα υπολογισμών
- Μικρή απαίτηση σε υπολογιστική μνήμη
- Χαμηλό υπολογιστικό κόστος

Αρχικά, από την επίλυση του συστήματος σε μόνιμη κατάσταση λαμβάνονται οι μόνιμες καταστάσεις (Steady States) για κάθε τιμή παροχής στο εύρος $1.5 - 4.5 \cdot 10^{-5}$ kg/s για τον κάτω ευσταθή κλάδο και $2.5 - 6.5 \cdot 10^{-5}$ kg/s για τον άνω ευσταθή κλάδο, με βήμα $1 \cdot 10^{-5}$ kg/s.

Έπειτα, για την συλλογή των κατάλληλων δεδομένων, πραγματοποιούνται κάποιες προσομοιώσεις επιβολών με χρήση του υπολογιστικού κώδικα ρευστοδυναμικής Ansys Fluent (Παράγραφος 2.4).

Η μόνη μεταβαλλόμενη μεταβλητή στο μελετώμενο σύστημα είναι η μαζική παροχή. Αν επιβληθεί μια βηματική επιβολή στην τιμή της μαζικής παροχής, τότε το σύστημα μεταβαίνει από την αρχική μόνιμη κατάσταση στην οποία βρισκόταν, σε μια νέα, που έχει ως μαζική παροχή την τιμή της επιβολής. Στην συνέχεια της εργασίας με τον όρο «Επιβολή» θα νοείται η παραπάνω μετάβαση του συστήματος από μια αρχική μόνιμη κατάσταση σε μια τελική μόνιμη κατάσταση που προκύπτει από την διαταραχή της μαζικής παροχής.

Έχοντας ως αρχική κατάσταση μια τιμή παροχής στο εύρος από το οποίο λήφθηκαν οι μόνιμες καταστάσεις, επιβάλλεται στο σύστημα μια νέα τιμή μαζικής παροχής και μέσω του Fluent επιλύεται το σύστημα μέχρι να επέλθει ξανά σε μόνιμη κατάσταση. Όλες οι επιβολές έχουν ως τελική κατάσταση την ίδια τιμή παροχής $3 \cdot 10^{-5}$ kg/s και για να επιτευχθεί η συλλογή των δεδομένων από μη μόνιμες καταστάσεις ο κώδικας δεν αφήνεται να συγκλίνει στην λύση, αλλά σταματάει μετά από ένα συγκεκριμένο αριθμό χρονικών βημάτων. Οι αρχικές καταστάσεις που χρησιμοποιήθηκαν εκτείνονται στους δύο ευσταθείς κλάδους που εμφανίζει το σύστημα (Παράγραφος 2.5) με σκοπό να καλύπτεται όλο το φάσμα της πολλαπλότητας των λύσεων, ενώ η τελική παροχή βρίσκεται εντός της περιοχής πολλαπλότητας.

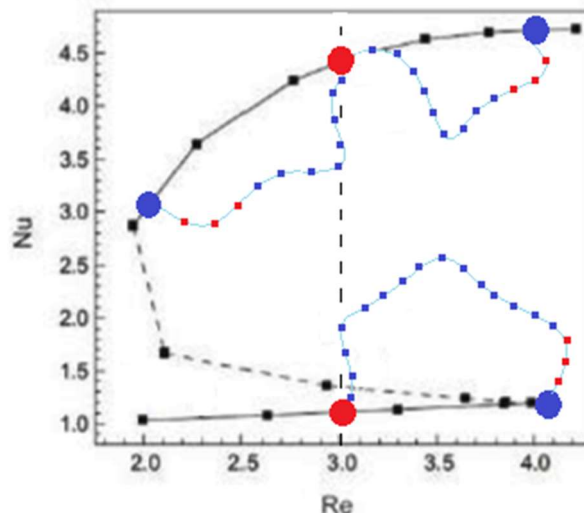
Η μέθοδος που χρησιμοποιήθηκε παρουσιάζεται στο Διάγραμμα 5.1, όπου οι κόκκινοι κύκλοι αποτυπώνουν την τελική κατάσταση $3 \cdot 10^{-5}$ kg/s, στον άνω και στον κάτω κλάδο αντίστοιχα, ενώ οι μπλε κύκλοι τρεις τυχαία επιλεγμένες αρχικές καταστάσεις από τις οποίες ξεκινάει το σύστημα. Έχουν σχεδιαστεί τυχαία οι τροχιές που θα μπορούσε να ακολουθεί η κάθε επιβολή, και τα τετράγωνα επάνω σε αυτές, αποτυπώνουν τα χρονικά βήματα (timesteps) που εκτελούνται μέχρι το σύστημα να ισορροπήσει στη νέα τιμή παροχής. Κάθε τετράγωνο αποτελεί ένα snapshot, μια εικόνα δηλαδή του συστήματος την δεδομένη χρονική στιγμή, στην οποία είναι γνωστές οι τιμές των ταχυτήτων (αξονική και ακτινική), της πίεσης και τις θερμοκρασίας σε όλα τα κελιά της προσομοίωσης του αντιδραστήρα. Για την κατασκευή του Μοντέλου Μειωμένης Τάξης, χρησιμοποιούνται μόνο δεδομένα από τις πρώτες χρονικές στιγμές, τα οποία αποτυπώνονται στο Σχήμα 5.1 με κόκκινο χρώμα.

Η μελέτη συνεπώς επικεντρώνεται στον αριθμό των επιβολών που απαιτούνται για να ικανοποιούνται τα κριτήρια του μοντέλου, και τον αριθμό των snapshots που επιλέγεται από κάθε επιβολή.

5.2 Ρύθμιση παραμέτρων στον υπολογιστικό κώδικα

Ο αντιδραστήρας Χημικής Απόθεσης από Ατμό που μελετάται στην παρούσα εργασία, έχει μελετηθεί σε μια σειρά από προηγούμενες εργασίες [1][6][21][22], συνεπώς είναι γνωστές αρκετές παράμετροι που χρησιμοποιούνται στον υπολογιστικό κώδικα.

Πραγματοποιήθηκαν τέσσερις δοκιμές, στις οποίες προκαλούνταν μια διαταραχή στην μαζική παροχή του συστήματος και το σύστημα αφήνονταν να ισορροπήσει στην νέα τιμή μαζικής παροχής. Και οι τέσσερις δοκιμές υλοποιήθηκαν έχοντας την ίδια, αρχική τιμή μαζικής παροχής $4,4869 \cdot 10^{-5}$ kg/s, η οποία στο σύστημα που μελετάται (στους 700K) είναι παροχή που βρίσκεται στον ασταθή κλάδο. Μεταβάλλοντας τον αριθμό των επαναληπτικών βημάτων που εκτελούνται κατά την διάρκεια ενός timestep στον Fluent, παρατηρούνταν αλλαγή στον συνολικό αριθμό των timesteps που χρειάζονταν το σύστημα για να φτάσει στην μόνιμη κατάσταση.



Διάγραμμα 5. 1: Τροχιές τυχαία επιλεγμένων επιβολών στον άνω και στον κάτω κλάδο του συστήματος. Με μπλε κύκλο συμβολίζεται η αρχική μόνιμη κατάσταση, ενώ με κόκκινο κύκλο η τελική μόνιμη κατάσταση που είναι για όλες τις επιβολές η ίδια, με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s. Οι γαλάζιες καμπύλες αποτελούν τις τροχιές των επιβολών ενώ τα τετράγωνα συμβολίζουν τα snapshots κάθε τροχιάς. Με κόκκινο χρώμα συμβολίζονται αυτά που χρησιμοποιούνται για την κατασκευή του Μοντέλου Μειωμένης Τάξης, ενώ με μπλε όλα τα υπόλοιπα.

Στον Πίνακα 5.1 παρουσιάζονται οι δοκιμές αυτές με τα αντίστοιχα αποτελέσματα.

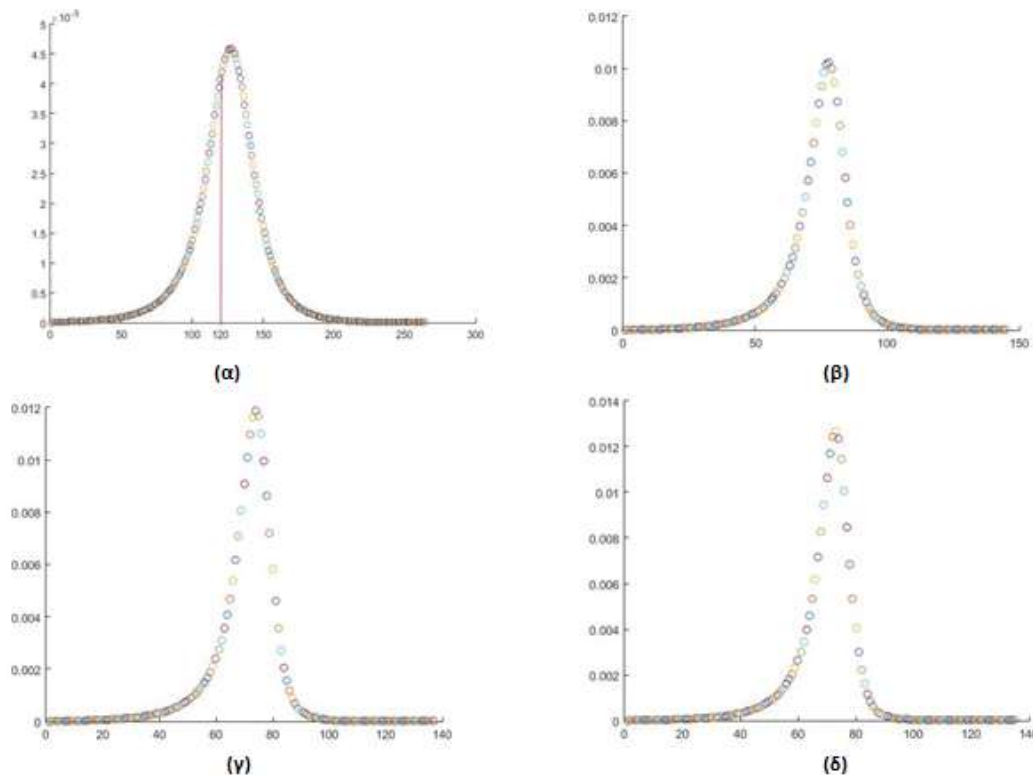
A/A	Αριθμός επαναλήψεων/ timestep	Αριθμός timesteps
1	50	265
2	200	140
3	350	138
4	600	138

Πίνακας 5. 1 Αποτελέσματα από τις 4 δοκιμές που πραγματοποιήθηκαν

Με την επίτευξη της νέας μόνιμης κατάστασης του συστήματος, υπολογίζονται τα διαγράμματα σφαλμάτων (Διάγραμμα 5.2). Η απόκλιση μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης κατά την εξέλιξη του συστήματος προς την νέα μόνιμη κατάσταση υπολογίζεται μέσω του πηλίκου 5.1. Στο Διάγραμμα 5.2 που ακολουθεί παρουσιάζεται η απόκλιση που αναφέρθηκε υπολογισμένη για κάθε ένα χρονικό βήμα (timestep) μέχρι την σύγκλιση της επαναληπτικής διαδικασίας.

$$aa = \frac{\text{norm}(x_1 - x)}{\text{norm}(x)} \quad (5.1)$$

όπου, x είναι το διάνυσμα κατάστασης του συστήματος για χρόνο t , ενώ x_1 το διάνυσμα κατάστασης της προηγούμενης χρονικής στιγμής, δηλαδή για $t-1$.



Διάγραμμα 5. 2: Καμπύλες του πηλίκου aa , για αριθμό επαναλήψεων στο Fluent: α) 50 i/timestep, β) 200 i/timestep, γ) 350 i/timestep και δ) 600 i/timestep. Ο οριζόντιος άξονας συμβολίζει τον αριθμό timesteps του συστήματος, δηλαδή τον χρόνο, ενώ ο κάθετος άξονας αναφέρεται στην τιμή σφάλματος.

Είναι εμφανές ότι στο Διάγραμμα 5.2.(α), που αντιστοιχεί σε αριθμό επαναλήψεων 50 i/timester, το μέγιστο σφάλμα της καμπύλης μειώνεται σημαντικά συγκριτικά με τα υπόλοιπα διαγράμματα. Παρά την μεγάλη μείωση στο ύψος της κορυφής της καμπύλης, ο αριθμός των timesteps που απαιτούνται μέχρι το σύστημα να φτάσει σε μόνιμη κατάσταση αυξάνεται. Η αύξηση αυτή δεν αποτελεί πρόβλημα στην εξέλιξη της εργασίας, αφού όπως έχει αναφερθεί για την κατασκευή του μοντέλου μειωμένης τάξης χρησιμοποιούνται δεδομένα μη μόνιμων καταστάσεων. Συνεπώς επιλέγεται να εκτελούνται 50 επαναλήψεις σε κάθε χρονικό βήμα $\Delta t=0.1s$ του λογισμικού Fluent.

5.3 Διερευνήσεις

A) 1^η Διερεύνηση

i) Επιβολές

Όπως αναφέρθηκε στην Παράγραφο 5.1, πραγματοποιήθηκαν κάποιες προσομοιώσεις επιβολών με στόχο να υπολογισθεί ο χώρος από τον οποίο αντλούνται τα υπολογιστικά δεδομένα που θα χρησιμοποιηθούν για την κατασκευή της βάσης και στη συνέχεια του νευρωνικού δικτύου. Πιο συγκεκριμένα πραγματοποιήθηκαν εννέα επιβολές στο σύστημα ανοιχτού βρόχου, έχοντας ως αρχική κατάσταση μια τιμή παροχής στο εύρος από το οποίο λήφθηκαν οι μόνιμες καταστάσεις, και ως τελική την ίδια τιμή παροχής $3 \cdot 10^{-5} \text{ kg/s}$. Επαναλαμβάνεται, ότι οι αρχικές καταστάσεις που χρησιμοποιήθηκαν εκτείνονταν στους δύο ευσταθείς κλάδους που εμφανίζει το σύστημα (Παράγραφος 2.5) με σκοπό να καλύπτεται όλο το φάσμα της πολλαπλότητας των λύσεων, ενώ η τελική παροχή βρίσκεται εντός της περιοχής πολλαπλότητας.

Στην μελέτη των Graham et al. [3] γίνεται αναφορά στο χρονικό διάστημα, κατά το οποίο αφήνονται οι αρχικές καταστάσεις να εξελιχθούν (Παράγραφος 3.8). Το χρονικό διάστημα δεν θα πρέπει να είναι ούτε πολύ μεγάλο με αποτέλεσμα να πραγματοποιείται σύγκλιση στη λύση του συστήματος, ούτε πολύ μικρό ώστε ο τελικός όγκος από τον οποίο αντλούνται τα δεδομένα να ισούται με τον αυτόν από τον οποίο ελήφθησαν οι αρχικές καταστάσεις. Στην παρούσα εργασία, ο χρόνος που αφήνεται μια αρχική κατάσταση να εξελιχθεί υπολογίζεται με τον αριθμό των timestep που εκτελεί κάθε επιβολή. Αν ο αριθμός των timestep είναι πολύ μεγάλος τότε το σύστημα θα έχει προλάβει να συγκλίνει στην τελική κατάσταση, η οποία είναι ίδια για όλες τις επιβολές και ίση με $3 \cdot 10^{-5} \text{ kg/s}$. Σε αυτή την περίπτωση χάνεται η άντληση πληροφορίας που επιδιώκεται (Παράγραφος 3.7). Αντίστοιχα, αν ο αριθμός των επαναληπτικών βημάτων που επιλεχθεί είναι πολύ μικρός, όλες οι επιβολές θα είναι κοντά στις αρχικές καταστάσεις από τις οποίες ξεκίνησαν, και συνεπώς το μοντέλο θα αποτυγχάνει να προβλέπει με ακρίβεια την συμπεριφορά του μελετώμενου συστήματος.

Για την επιλογή του αριθμού των timesteps που θα εκτελούν οι επιβολές στην πρώτη προσέγγιση, χρησιμοποιήθηκε το Διάγραμμα 5.2.(α). Συγκεκριμένα επιλέχθηκε ένας αριθμός για τον οποίο το σύστημα βρίσκεται εντός της μεταβατικής κατάστασης και

δεν έχει προλάβει να συγκλίνει στη λύση. Ο αριθμός timesteps που επιλέχθηκε να εκτελούν οι προσομοιώσεις των επιβολών είναι 120.

Στον Πίνακα 5.2 παρουσιάζονται οι επιβολές που χρησιμοποιήθηκαν για την μελέτη του συστήματος.

Α/Α	Αρχική κατάσταση		Τελική Κατάσταση	
	Παροχή m_0 ($\cdot 10^{-5}$) kg/s	Κλάδος	Παροχή m_f ($\cdot 10^{-5}$) kg/s	Κλάδος
1	1,5	(Κάτω)	3	Κάτω
2	2,5	Κάτω	3	Κάτω
3	2,5	Άνω	3	Άνω
4	3,5	Κάτω	3	Κάτω
5	3,5	Άνω	3	Άνω
6	4,5	Κάτω	3	Κάτω
7	4,5	Άνω	3	Άνω
8	5,5	(Άνω)	3	Άνω
9	6,5	(Άνω)	3	Άνω

Πίνακας 5. 2: Βηματικές επιβολές στο σύστημα ανοιχτού βρόχου. Ο κλάδος βρίσκεται σε παρένθεση αν η μόνιμη κατάσταση βρίσκεται εκτός της περιοχής πολλαπλότητας.

ii) Κατασκευή PCA Βάσης

Για την κατασκευή της βάσης με τη μέθοδο των στιγμιοτύπων, χρησιμοποιήθηκαν οι οκτώ από τις εννέα επιβολές. Η επιβολή 5 δεν χρησιμοποιήθηκε για την κατασκευή της βάσης και για αυτό από εδώ και στο εξής θα καλείται εξωτερική επιβολή. Οι επιβολές που χρησιμοποιήθηκαν για την κατασκευή της βάσης θα καλούνται εσωτερικές επιβολές.

Χρησιμοποιώντας επομένως ως τροχιές τις δυναμικές αποκρίσεις στις επιβολές αυτές, εφαρμόζοντας τον αλγόριθμο του Σχήματος 3.3 και επιλέγοντας κλάσμα περιεχόμενης πληροφορίας $\rho=0,98$ κατασκευάστηκε μια βάση από ιδιοδιανύσματα.

Από τον παραπάνω αλγόριθμο προκύπτει ότι η βάση, η οποία καλείται Z4, αποτελείται από $d=4$ ιδιοσυναρτήσεις και περιέχει πάνω από το 98.6% της πληροφορίας των εννέα επιβολών. Το ποσοστό αυτό με βάση το θεωρητικό υπόβαθρο της παρούσας διπλωματικής εργασίας, τα δεδομένα και τους στόχους της, είναι ικανοποιητικό.

Στον παρακάτω πίνακα παρουσιάζεται το σχετικό πληροφοριακό περιεχόμενο για κάθε προσέγγιση PCA των τεσσάρων αυτών ιδιοσυναρτήσεων μέχρι τη βάση Z4.

Διάσταση προσέγγισης PCA, d	1	2	3	4
Σχετικό πληροφοριακό περιεχόμενο, I(d)	0,7343	0,9219	0,9779	0,9860

Πίνακας 5. 3: Σχετικό πληροφοριακό περιεχόμενο προσεγγίσεων PCA για διαφορετικές διαστάσεις βάσης. Επιβολές που αποτελούνται από 120 snapshots.

Παρατηρείται ότι το πρώτο ιδιοδιάνυσμα περιέχει μόνο το 73.43% του πληροφοριακού περιεχομένου των εννέα επιβολών, ενώ από το δεύτερο

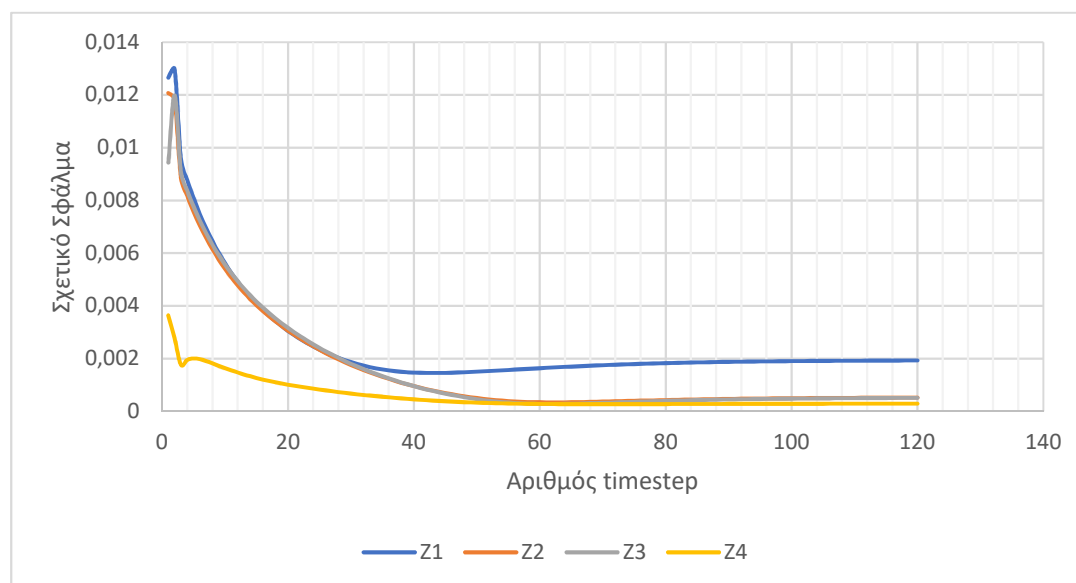
ιδιοδιάνυσμα κι έπειτα το ποσοστό αυξάνεται πάνω από το 90% της πληροφορίας. Η βάση που αποτελείται από το πρώτο ιδιοδιάνυσμα θα καλείται Z1, ενώ Z2 η βάση που αποτελείται από τα δύο πρώτα ιδιοδιανύσματα και Z3, η βάση που αποτελείται από τα 3 πρώτα ιδιοδιανύσματα αντίστοιχα.

Στη συνέχεια αξιολογήθηκαν οι τέσσερις παραπάνω βάσεις ως προς το σφάλμα που παρουσιάζουν στην προσέγγιση γνωστών επιβολών. Τα σφάλματα αυτά υπολογίστηκαν για την εξωτερική επιβολή 5 και για την εσωτερική επιβολή 2 (Πίνακας 5.2) και σημειώνονται στον Πίνακα 5.3. Συγκεκριμένα τα σφάλματα που παρουσιάζονται είναι το ελάχιστο σφάλμα για κάθε βάση, το μέγιστο, καθώς και το τελικό σφάλμα δηλαδή μετά το πέρας των 120 timesteps.

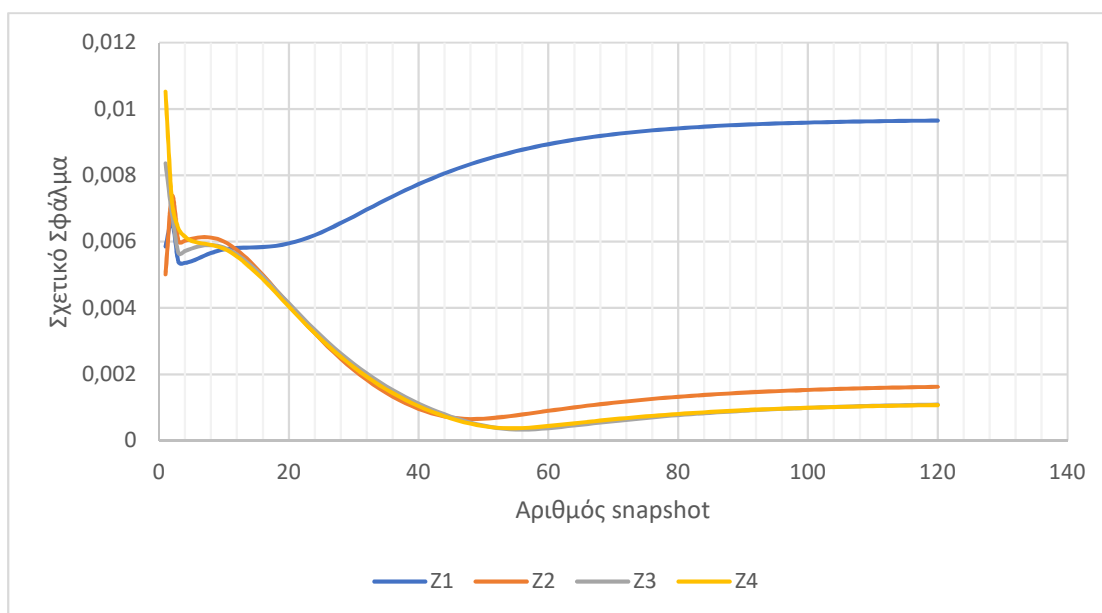
Βάσεις	Εσωτερική Επιβολή 2 $\dot{m}_0=2,5 \cdot 10^{-5}$ kg/s σε $\dot{m}_f=3 \cdot 10^{-5}$ kg/s (Κάτω κλάδος)			Εξωτερική Επιβολή 5 $\dot{m}_0=3,5 \cdot 10^{-5}$ kg/s σε $\dot{m}_f=3 \cdot 10^{-5}$ kg/s (Άνω κλάδος)		
	Min	Max	Τελικό Σφάλμα	Min	Max	Τελικό Σφάλμα
Z1	$1,46 \cdot 10^{-3}$	$1,30 \cdot 10^{-2}$	$1,93 \cdot 10^{-3}$	$8,67 \cdot 10^{-3}$	$5,36 \cdot 10^{-3}$	$9,65 \cdot 10^{-3}$
Z2	$3,42 \cdot 10^{-4}$	$1,211 \cdot 10^{-2}$	$5,23 \cdot 10^{-4}$	$1,44 \cdot 10^{-3}$	$6,53 \cdot 10^{-4}$	$1,62 \cdot 10^{-3}$
Z3	$2,79 \cdot 10^{-4}$	$1,20 \cdot 10^{-2}$	$5,12 \cdot 10^{-4}$	$9,48 \cdot 10^{-4}$	$3,35 \cdot 10^{-4}$	$1,09 \cdot 10^{-3}$
Z4	$2,73 \cdot 10^{-4}$	$3,64 \cdot 10^{-3}$	$2,93 \cdot 10^{-4}$	$3,73 \cdot 10^{-4}$	$1,05 \cdot 10^{-2}$	$1,07 \cdot 10^{-3}$

Πίνακας 5. 4: Σχετικά σφάλματα προσεγγίσεων PCA για διαφορετικών διαστάσεων βάσεις, κατά την προσέγγιση μιας εσωτερικής επιβολής και μιας εξωτερικής. Διερεύνηση με χρήση προσομοιώσεων εννέα επιβολών που αποτελούνται από 120 snapshots.

Στα Διαγράμματα 5.3 και 5.4 παρουσιάζεται το σχετικό σφάλμα κατά την προσέγγιση των δύο επιβολών, για τις τέσσερις διαφορετικές βάσεις συναρτήσει του αριθμού των timesteps που εκτελούνται σε κάθε επιβολή.



Διάγραμμα 5. 3: Σχετικό σφάλμα κάθε βάσης κατά την προσέγγιση της εξωτερικής επιβολής, από $\dot{m}(t)=3,5 \cdot 10^{-5}$ kg/s (Άνω κλάδος) σε $\dot{m}(t)=3 \cdot 10^{-5}$ kg/s. Διερεύνηση με χρήση προσομοιώσεων εννέα επιβολών που αποτελούνται από 120 snapshots..



Διάγραμμα 5. 4: Σχετικό σφάλμα κάθε βάσης κατά την προσέγγιση της εσωτερικής επιβολής, από $\dot{m}(t)=2,5 \cdot 10^{-5}$ kg/s (κάτω κλάδος) σε $\dot{m}(t)=3 \cdot 10^{-5}$ kg/s. Διερεύνηση με χρήση προσομοιώσεων εννέα επιβολών που αποτελούνται από 120 snapshots.

Στα παραπάνω διαγράμματα είναι εμφανές ότι το σχετικό σφάλμα μειώνεται καθώς οι επιβολές εξελίσσονται προς τη νέα μόνιμη κατάσταση. Παρατηρείται ότι και στις δύο περιπτώσεις, η βάση Z4 ή οποία περιέχει το μεγαλύτερο σχετικό πληροφοριακό περιεχόμενο (Πίνακας 5.3), παρουσιάζει το μικρότερο σφάλμα. Συνεπώς για την κατασκευή της βάσης επιλέγεται η χρήση τεσσάρων ιδιοδιανυσμάτων.

iii) Εκπαίδευση Νευρωνικού Δικτύου

Για την εκπαίδευση του νευρωνικού δικτύου, η επιλογή του πλήθους δεδομένων και του αριθμού L1 των νευρώνων του κρυφού επιπέδου (Σχήμα 4.4.β) βασίστηκε σε προηγούμενες εργασίες [6][21], αφού πραγματοποιήθηκε έλεγχος και τα αποτελέσματα ήταν ικανοποιητικά. Υπενθυμίζεται ότι η παρούσα εργασία επικεντρώνεται στην ανάλυση του αριθμού των επιβολών και του αριθμού των snapshots κάθε επιβολής έτσι ώστε να ικανοποιούνται τα κριτήρια που θα πρέπει να πληροί το μοντέλο.

Για την εκπαίδευση του δικτύου, χρησιμοποιήθηκαν οι επτά από τις εννέα επιβολές, με σκοπό να καλυφθούν τμήματα από όλο το εύρος της πολλαπλότητας των λύσεων και να αποκαλυφθεί όσο το δυνατόν μεγαλύτερο μέρος της δυναμικής του συστήματος. Σε προγενέστερη εργασία, οι επτά επιβολές αποτέλεσαν τον μικρότερο αριθμό επιβολών με τον οποίο πραγματοποιήθηκε η εκπαίδευση ενός νευρωνικού δικτύου κι εν συνεχεία ο έλεγχος κατά την πρόβλεψη της συμπεριφοράς του συστήματος από το δίκτυο και τα αποτελέσματα ήταν ικανοποιητικά.

Συγκεκριμένα οι επιβολές που χρησιμοποιήθηκαν ήταν οι 1, 3, 4, 6, 7, 8 και 9 (Πίνακας 5.2) και αποτελούν τις εσωτερικές επιβολές του δικτύου, ενώ οι 2 και 5 δεν χρησιμοποιήθηκαν στην εκπαίδευση και συνεπώς αποτελούν τις εξωτερικές

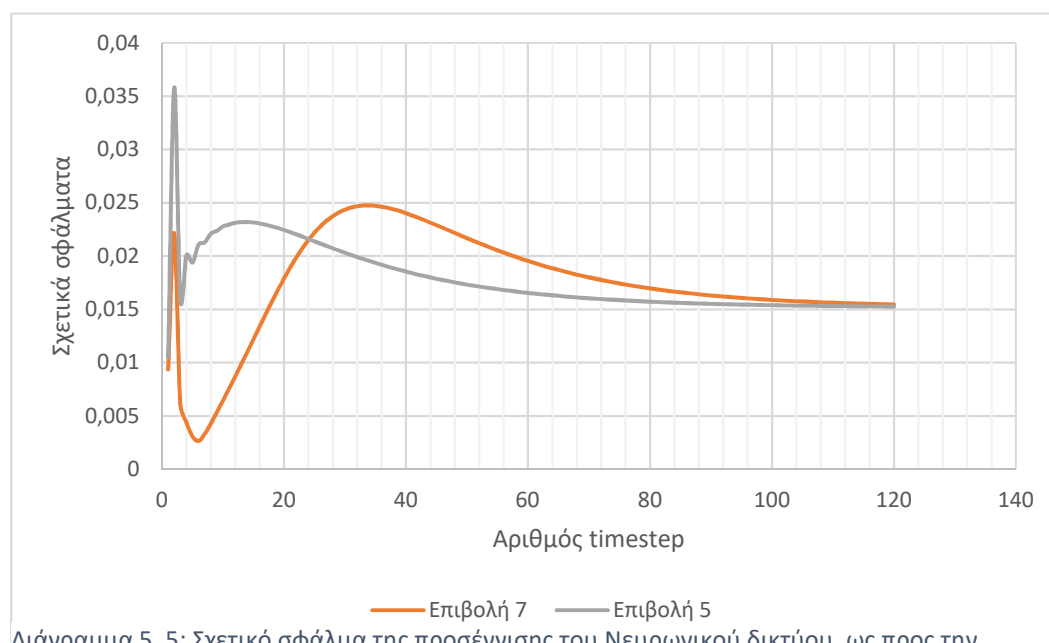
επιβολές. Το δίκτυο που εκπαιδεύτηκε έχει καθυστέρηση $d=1$ και το κρυφό επίπεδο του αποτελείται από $L=10$ νευρώνες.

Στον Πίνακα 5.5 παρουσιάζονται τα σχετικά σφάλματα που προέκυψαν από την πρόβλεψη γνωστών μόνιμων καταστάσεων του πλήρους συστήματος, μέσω του νευρωνικού δικτύου. Συγκεκριμένα παρουσιάζονται το ελάχιστο σφάλμα, το μέγιστο καθώς και το σφάλμα του τελευταίου timestep.

A/A	Επιβολές $\dot{m}_0 (\cdot 10^{-5})$	Κλάδος	Min	Max	Τελικό Σφάλμα
1	1,5	(Κάτω)	$5,13 \cdot 10^{-4}$	$1,20 \cdot 10^{-2}$	$5,13 \cdot 10^{-4}$
2	2,5 (Εξωτερική)	Κάτω	$5,07 \cdot 10^{-4}$	$8,48 \cdot 10^{-3}$	$5,07 \cdot 10^{-4}$
3	2,5	Άνω	$8,57 \cdot 10^{-3}$	$1,89 \cdot 10^{-2}$	$1,51 \cdot 10^{-2}$
4	3,5	Κάτω	$3,95 \cdot 10^{-4}$	$3,09 \cdot 10^{-3}$	$5,02 \cdot 10^{-4}$
5	3,5 (Εξωτερική)	Άνω	$1,05 \cdot 10^{-2}$	$3,58 \cdot 10^{-2}$	$1,53 \cdot 10^{-2}$
6	4,5	Κάτω	$4,36 \cdot 10^{-4}$	$9,75 \cdot 10^{-3}$	$4,97 \cdot 10^{-4}$
7	4,5	Άνω	$2,66 \cdot 10^{-3}$	$2,48 \cdot 10^{-2}$	$1,55 \cdot 10^{-2}$
8	5,5	(Άνω)	$2,66 \cdot 10^{-3}$	$2,48 \cdot 10^{-2}$	$1,55 \cdot 10^{-2}$
9	6,5	(Άνω)	$3,91 \cdot 10^{-3}$	$4,21 \cdot 10^{-2}$	$1,55 \cdot 10^{-2}$

Πίνακας 5. 5: Πίνακας σφαλμάτων του νευρωνικού δικτύου. Ο κλάδος βρίσκεται σε παρένθεση αν η μόνιμη κατάσταση βρίσκεται εκτός της περιοχής πολλαπλότητας. Διερεύνηση με χρήση προσομοιώσεων επιβολών που αποτελούνται από 120 snapshots.

Στο Διάγραμμα 5.5 παρουσιάζονται τα σχετικά σφάλματα που προκύπτουν κατά την προσέγγιση δύο γνωστών μόνιμων καταστάσεων του συστήματος από το νευρωνικό δίκτυο που κατασκευάστηκε. Πιο συγκεκριμένα προβάλλονται τα σφάλματα που προκύπτουν από την πρόβλεψη της Επιβολής 7 και της Επιβολής 5 (Πίνακας 5.2), του μοντέλου πλήρους τάξης.



Διάγραμμα 5. 5: Σχετικό σφάλμα της προσέγγισης του Νευρωνικού δικτύου, ως προς την πρόβλεψη των επιβολών 5 και 7. (Πίνακας 6.2)

Από το παραπάνω διάγραμμα παρατηρείται ότι τα σφάλματα του δικτύου, για μεγάλο αριθμό timesteps έχουν κοντινές τιμές, ενώ η Επιβολή 7, αν και αποτελεί επιβολή εκπαίδευσης του δικτύου, παρουσιάζει αρχικά μεγαλύτερη διαταραχή. Από τον Πίνακα 5.5, φαίνεται επίσης ότι η τάξη μεγέθους των σφαλμάτων είναι μεταξύ 10^{-2} - 10^{-4} που είναι ικανοποιητική για τα δεδομένα της εργασίας.

iv) Υπολογισμός μονίμων καταστάσεων

Με τη χρήση της ορθοκανονικής βάσης Z4 που κατασκευάστηκε στην Παράγραφο 5.3.Aii και του Νευρωνικού δικτύου που εκπαιδεύτηκε στην προηγούμενη παράγραφο, υπολογίστηκαν τέσσερις μόνιμες καταστάσεις του συστήματος. Υπενθυμίζεται ότι οι προσομοιώσεις των επιβολών που χρησιμοποιήθηκαν για την κατασκευή των παραπάνω, αποτελούνται από συγκεκριμένο αριθμό στιγμιοτύπων που λαμβάνονται κατά την εξέλιξη του συστήματος προς την νέα μόνιμη κατάσταση, χωρίς όμως να φτάνει σε αυτή. Πραγματοποιήθηκε λοιπόν, ο υπολογισμός τεσσάρων μόνιμων καταστάσεων του πλήρους συστήματος, από το μοντέλο μειωμένης τάξης που κατασκευάστηκε με χρήση δεδομένων μη μόνιμων καταστάσεων.

Οι μόνιμες καταστάσεις που επιλέχθηκαν να υπολογισθούν είχαν την ίδια τιμή μαζικής παροχής με κάποιες από τις μόνιμες καταστάσεις που είχαν ήδη υπολογισθεί από την προσομοίωση του αντιδραστήρα μέσω του Fluent, όπως αναφέρεται στην Παράγραφο 5.1, ώστε να είναι εφικτή η σύγκριση των αποτελεσμάτων. Συγκεκριμένα, πραγματοποιήθηκε σύγκριση της γνωστής μόνιμης κατάστασης, δηλαδή εκείνης που υπολογίζεται μέσω του Fluent, με εκείνη που υπολογίζεται από το Μοντέλο Μειωμένης Τάξης. Οι αποκλίσεις μεταξύ των δύο μόνιμων καταστάσεων υπολογίστηκαν με χρήση του πηλίκου:

$$SF = \frac{norm(SS_ANN - S_CFD)}{norm(SS_CFD)} \quad (5.2)$$

Όπου,

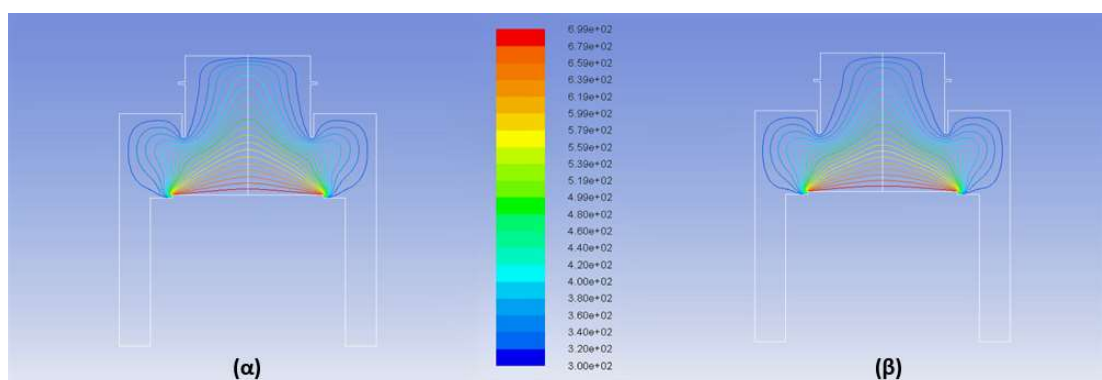
- SS_ANN: Η μόνιμη κατάσταση που υπολογίζεται από το νευρωνικό δίκτυο. Αποτελεί το διάνυσμα κατάστασης του συστήματος, που υπολογίζεται μέσω του νευρωνικού δικτύου, όταν επιτυγχάνεται η σύγκλιση στη λύση.
- SS_CFD: Η μόνιμη κατάσταση που έχει υπολογισθεί στην Παράγραφο 5.1, με χρήση του Fluent. Αποτελεί το διάνυσμα κατάστασης του συστήματος όταν επιτυγχάνεται η σύγκλιση στη λύση.

Στον Πίνακα 5.6 παρουσιάζονται οι τέσσερις μόνιμες καταστάσεις που προσεγγίστηκαν και το σφάλμα που προέκυψε για καθεμία.

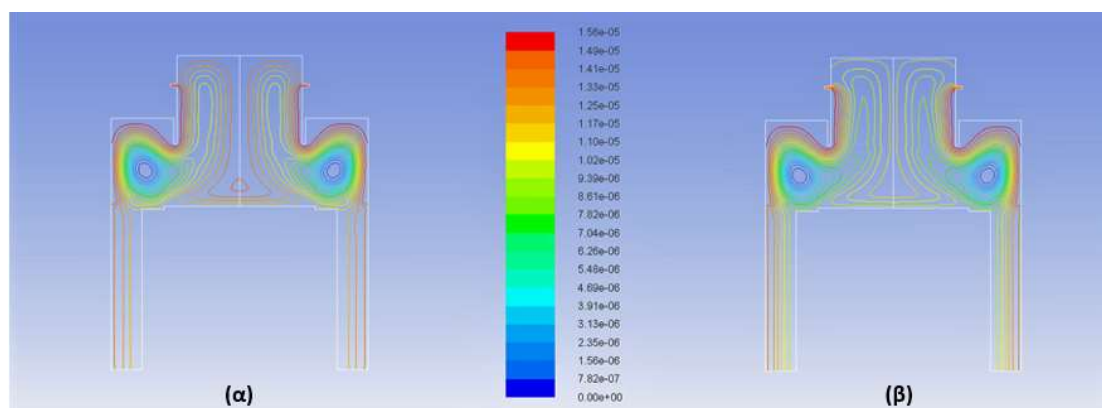
Α/Α	Μόνιμη Κατάσταση		Σφάλμα (%)
	Παροχή mf ($\cdot 10^{-5}$) kg/s	Κλάδος	
MK_1.1	1,5	(Κάτω)	2,38
MK_1.2	3,5	Κάτω	0,85
MK_1.3	5,5	Άνω	8,41
MK_1.4	6,5	Άνω	10,15

Πίνακας 5. 6: Αποκλίσεις δοκιμαστικών επιβολών από την μόνιμη κατάσταση. Σε παρένθεση βρίσκεται ο κλάδος που είναι εκτός της περιοχής πολλαπλότητας. Διερεύνηση με χρήση προσομοιώσεων επιβολών που αποτελούνται από 120 snapshots.

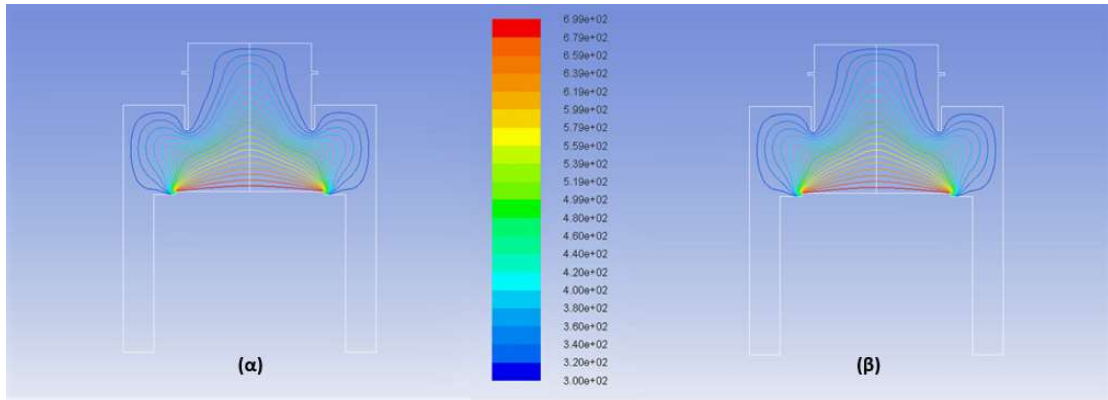
Στα διαγράμματα 5.6 έως 5.13, παρουσιάζονται οι κατανομές των θερμοκρασιών και των ταχυτήτων στον αντιδραστήρα χημικής απόθεσης από ατμό, για καθεμία από τις μόνιμες καταστάσεις που υπολογίστηκαν. Το αριστερό διάγραμμα παρουσιάζει την κατανομή που έχει προκύψει από την προσομοίωση μέσω του Fluent, ενώ το δεξιό την κατανομή που έχει προβλεφθεί από το νευρωνικό δίκτυο που εκπαιδεύτηκε στην παράγραφο 5.3.A (iii).



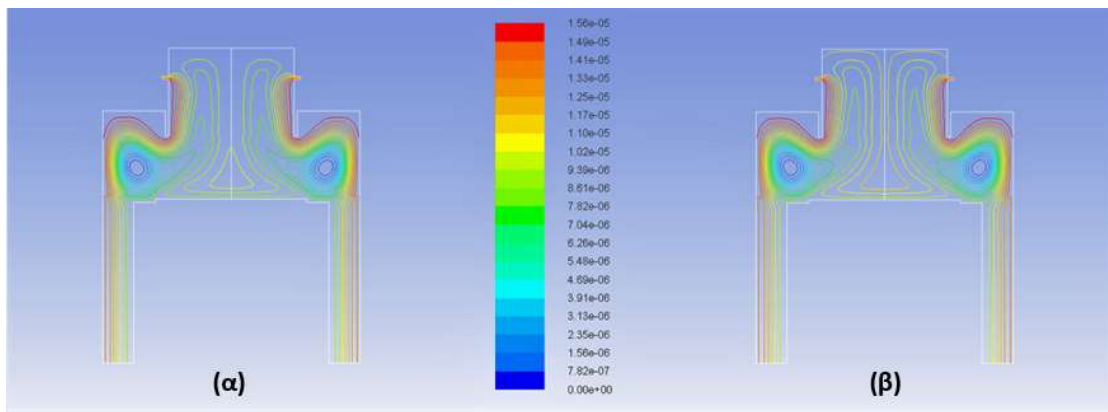
Διάγραμμα 5. 6: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=1,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



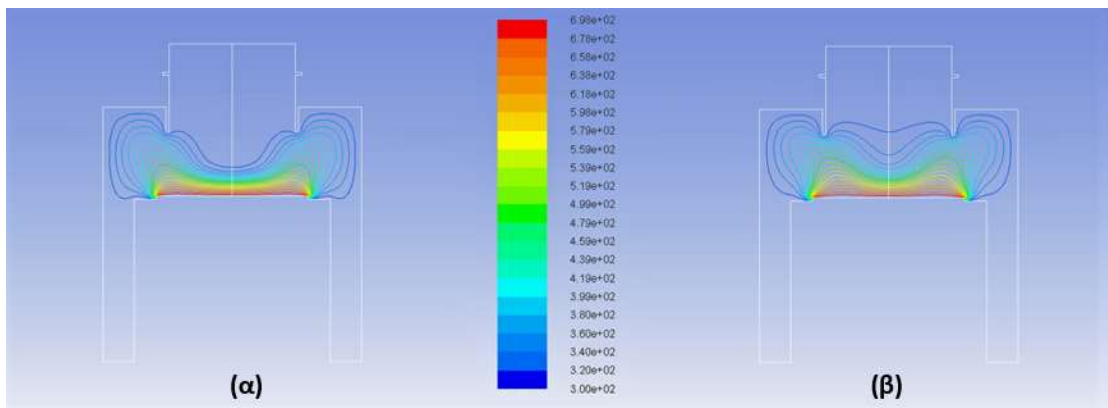
Διάγραμμα 5. 7: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=1,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN



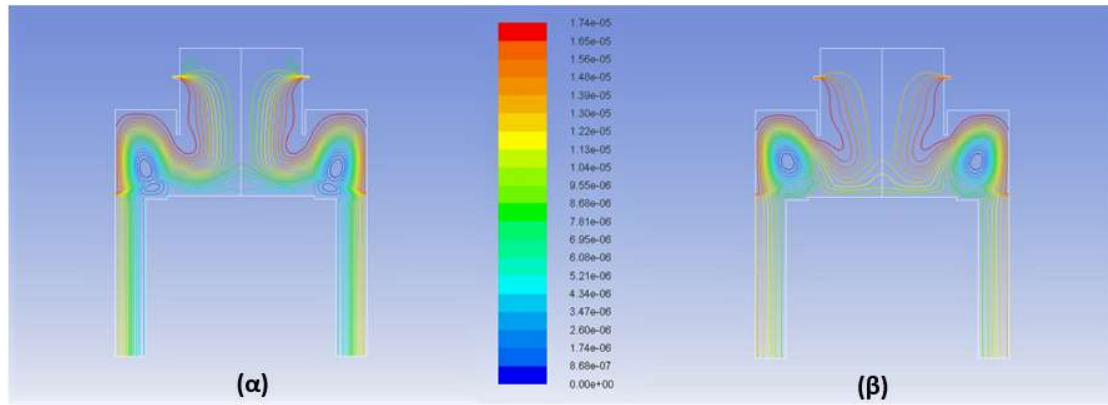
Διάγραμμα 5. 8: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=3,5 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται α) CFD, β) ANN



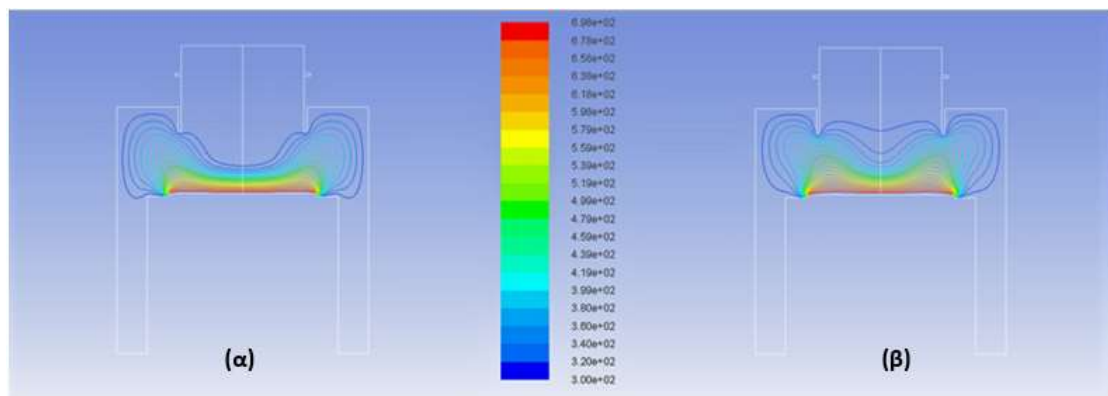
Διάγραμμα 5. 9: Κατανομή ταχύτητας στον αντιδραστήρα για την μόνιμη κατάσταση με τιμή παροχής $m_f=3,5 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται α) CFD, β) ANN



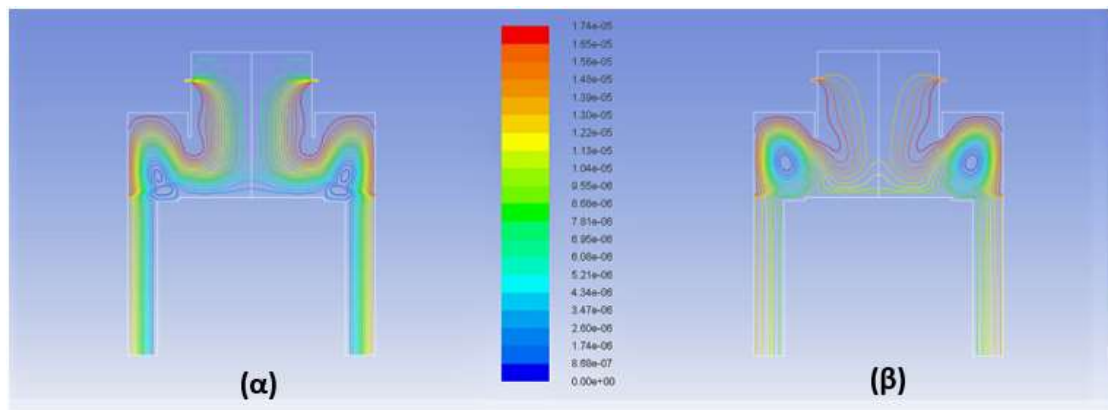
Διάγραμμα 5. 10: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=5,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN



Διάγραμμα 5. 11: Κατανομή ταχύτητας στον αντιδραστήρα για την μόνιμη κατάσταση με τιμή παροχής $m_f=5,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN



Διάγραμμα 5. 12: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=6,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN



Διάγραμμα 5. 13: Κατανομές ταχύτητας στον αντιδραστήρα για την μόνιμη κατάσταση με τιμή παροχής $m_f=6,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN

Από τα παραπάνω διαγράμματα φαίνεται ότι οι κατανομές της θερμοκρασίας και της ταχύτητας στον κάτω κλάδο, Διαγράμματα 5.6 – 5.9, έχουν προσεγγιστεί ικανοποιητικά. Το νευρωνικό δίκτυο αποτυγχάνει να περιγράψει με μεγάλη ακρίβεια τον άνω ευσταθή κλάδο, Διαγράμματα 5.10 – 5.13, κι αυτό θα μπορούσε να προβλεφθεί από τις μεγάλες τιμές σφαλμάτων που παρουσιάστηκαν στον Πίνακα 5.5.

B) 2η Διερεύνηση

i) Επιβολές

Με στόχο να μειωθεί η απόκλιση μεταξύ των μόνιμων καταστάσεων της προσομοίωσης και εκείνης που υπολογίζεται από το μοντέλο μειωμένης τάξης, αυξήθηκε ο αριθμός των επιβολών που χρησιμοποιούνται για την κατασκευή του μοντέλου.

Για την αύξηση των επιβολών υπολογίστηκαν αρχικά 15 νέες μόνιμες καταστάσεις από τις οποίες έπειτα πραγματοποιήθηκαν οι επιβολές στο σύστημα. Οι καταστάσεις αυτές αποτέλεσαν τις αρχικές καταστάσεις από τις οποίες ξεκινούσε το σύστημα, ενώ όλες οι επιβολές είχαν την ίδια τελική μόνιμη κατάσταση με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s. Ο υπολογισμός των νέων μόνιμων καταστάσεων υλοποιήθηκε με χρήση δεδομένων από τις ήδη υπάρχουσες μόνιμες καταστάσεις που είχαν υπολογισθεί μέσω του λογισμικού Fluent (Παράγραφος 5.1).

Συγκεκριμένα, μια αρχική κατάσταση περιγράφεται αν για δεδομένη χρονική στιγμή είναι γνωστά, για όλα τα κελιά του αντιδραστήρα (15066 κελιά, Παράγραφος 2.4), η πίεση, η θερμοκρασία και οι δύο συνιστώσες, ακτινική και αξονική, της ταχύτητας. Δηλαδή ένα διάνυσμα κατάστασης με 60264 στοιχεία, το οποίο αποτελεί την πρώτη στήλη του πίνακα τροχιάς (trajectory) της εν λόγω επιβολής. Η τυχαία επιλογή των αρχικών καταστάσεων μέσα σε έναν δεδομένο χώρο λύσεων (παράγραφος 3.8) που εφαρμόζεται στην παρούσα εργασία, καθιστά εφικτό τον υπολογισμό των νέων μόνιμων καταστάσεων με χρήση των μέσων όρων των διανυσμάτων κατάστασης των ήδη υπάρχουσών μόνιμων καταστάσεων. Με αυτόν τον τρόπο υπολογίζονται καταστάσεις που αποτελούν διαταραγμένες θέσεις των μόνιμων καταστάσεων που υπολογίζονται μέσω της προσομοίωσης, είναι όμως εντός του εύρους από το οποίο επιλέγονται οι τυχαίες αρχικές μόνιμες καταστάσεις του αντιδραστήρα και μπορούν έτσι να χρησιμοποιηθούν για την κατασκευή του μοντέλου.

Με αυτόν τον τρόπο διατηρούνται ίδιες όλες οι συνοριακές συνθήκες του αντιδραστήρα, ενώ ο υπολογισμός των μόνιμων καταστάσεων είναι πολύ πιο σύντομος.

Ο νέος πίνακας επιβολών που προκύπτει παρουσιάζεται παρακάτω. Με N# συμβολίζονται οι νέες επιβολές που υπολογίστηκαν με τον παραπάνω αναφερόμενο τρόπο. Υπενθυμίζεται ότι όλες οι επιβολές έχουν ως τελική κατάσταση την τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s και για την κατασκευή του μοντέλου χρησιμοποιούνται μόνο τα 120 πρώτα snapshots της τροχιάς που ακολουθεί η κάθε επιβολή.

Αξίζει να σημειωθεί ότι η επιβολή που αρχίζει από την μόνιμη κατάσταση με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s και έχει ως τελική κατάσταση την ίδια τιμή μαζικής παροχής, είτε είναι στον κάτω (Επιβολή N2, Πίνακας 5.7) είτε στον άνω κλάδο (Επιβολή N3), μπορεί να πραγματοποιηθεί καθώς οι αρχικές αυτές καταστάσεις είναι υπολογισμένες με τον τρόπο που περιγράφηκε παραπάνω, και συνεπώς τα διανύσματα των μόνιμων καταστάσεων τους διαφέρουν από αυτά που υπολογίζονται μέσω του λογισμικού Fluent, για τις αντίστοιχες μόνιμες καταστάσεις. Όπως αναφέρθηκε οι αρχικές καταστάσεις που υπολογίστηκαν αποτελούν

διαταραγμένες καταστάσεις της μόνιμης κατάστασης με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s.

A/A	Αρχική κατάσταση	
	Παροχή $m_0 (10^{-5})$ kg/s	Κλάδος
1	1,5	(Κάτω)
N1	2	(Κάτω)
N8	2,25	(Κάτω)
2	2,5	Κάτω
3	2,5	Άνω
N9	2,75	Κάτω
N2	3	Κάτω
N3	3	Άνω
N10	3,25	Κάτω
N11	3,25	Άνω
4	3,5	Κάτω
5	3,5	Άνω
N12	3,75	Κάτω
N13	3,75	Άνω
N4	4	Κάτω
N5	4	Άνω
N14	4,25	Άνω
6	4,5	Κάτω
7	4,5	Άνω
N15	4,75	Άνω
N6	5	Άνω
8	5,5	(Άνω)
N7	6	(Άνω)
9	6,5	(Άνω)

Πίνακας 5. 7: Βηματικές επιβολές στο σύστημα ανοιχτού βρόχου. Ο κλάδος βρίσκεται σε παρένθεση αν η μόνιμη κατάσταση βρίσκεται εκτός της περιοχής πολλαπλότητας. Διερεύνηση με χρήση 24 επιβολών που αποτελούνται από 120 snapshots.

ii) Κατασκευή PCA Βάσης

Για την κατασκευή της PCA βάσης με τη μέθοδο των στιγμιοτύπων, χρησιμοποιήθηκαν ως τροχιές οι δυναμικές όλων των επιβολών, εκτός της Επιβολής 5. Η νέα βάση που προκύπτει από τη χρήση του αλγόριθμου του Σχήματος 3.3 έχει κλάσμα περιεχόμενης πληροφορίας $\rho=98$ και καλείται Z4.

Στον παρακάτω πίνακα παρουσιάζεται το σχετικό πληροφοριακό περιεχόμενο για κάθε προσέγγιση PCA των τεσσάρων ιδιοσυναρτήσεων από τις οποίες αποτελείται, μέχρι τη βάση Z4.

Διάσταση προσέγγισης PCA, d	1	2	3	4
Σχετικό πληροφοριακό περιεχόμενο, I(d)	0,7481	0,9268	0,9819	0,9876

Πίνακας 6. 8: Σχετικό πληροφοριακό περιεχόμενο προσεγγίσεων PCA για διαφορετικές διαστάσεις βάσης. Διερεύνηση που χρησιμοποιεί 24 επιβολές οι οποίες αποτελούνται από 120 snapshots.

Παρατηρείται ότι το πληροφοριακό περιεχόμενο των επιβολών, που περιέχει το κάθε ιδιοδιάνυσμα, αυξάνεται συγκριτικά με την προηγούμενη διερεύνηση.

Η αξιολόγηση της βάσης πραγματοποιήθηκε με τις ίδιες επιβολές όπως και στην προηγούμενη διερεύνηση, και τα σχετικά σφάλματα ήταν της τάξης μεγέθους $10^{-2} - 10^{-4}$, που είναι ικανοποιητικά για την παρούσα εργασία

iii) Εκπαίδευση Νευρωνικού Δικτύου

Το νευρωνικό δίκτυο εκπαιδεύτηκε με χρήση της βάσης Z4, και για την εκπαίδευση του συγκεκριμένου δικτύου, χρησιμοποιήθηκαν 22 από τις 24 επιβολές, με τις Επιβολές 3 και 6 (Πίνακας 5.7) να αποτελούν τις εξωτερικές επιβολές. Το δίκτυο αυτό έχει καθυστέρηση $d=1$ και $L=10$ κρυφούς νευρώνες.

Με την αξιολόγηση που έγινε κατά την πρόβλεψη γνωστών μονίμων καταστάσεων με χρήση του νευρωνικού δικτύου που εκπαιδεύτηκε, παρατηρήθηκε ότι οι συντελεστές σφαλμάτων του δικτύου κυμαίνονται στην ίδια τάξη μεγέθους όπως και οι συντελεστές των προηγούμενων νευρωνικών δικτύων.

iv) Υπολογισμός μονίμων καταστάσεων

Με χρήση του μοντέλου μειωμένης τάξης που κατασκευάστηκε στην διερεύνηση αυτή, έγινε εκ νέου ο υπολογισμός των τεσσάρων μονίμων καταστάσεων.

Ο υπολογισμός των αποκλίσεων της μόνιμης κατάστασης που υπολογίζεται από το νευρωνικό δίκτυο από την μόνιμη κατάσταση που υπολογίζεται από την προσομοίωση, πραγματοποιείται με χρήση του Πηλίκου 5.2. Στον Πίνακα 5.9 παρουσιάζονται οι τέσσερις μόνιμες καταστάσεις που προσεγγίστηκαν και το επί τις εκατό σφάλμα που προέκυψε για καθεμία. Για διευκόλυνση στη σύγκριση των δύο υπολογιστικών πειραμάτων, ο πίνακας περιλαμβάνει και μια στήλη με τα σφάλματα της προηγούμενης διερεύνησης.

A/A	Μόνιμη Κατάσταση		Διερευνήσεις	
	Παροχή mf ($\cdot 10^{-5}$) kg/s	Κλάδος	1η	2η
			9 επιβολές 120 timesteps Σφάλμα (%)	24 επιβολές 120 timesteps Σφάλμα(%)
Π1	1,5	(Κάτω)	2,38	2,42
Π2	3,5	Κάτω	0,85	0,81
Π3	5,5	Άνω	8,41	6,86
Π4	6,5	Άνω	10,15	8,68

Πίνακας 5. 9: Αποκλίσεις δοκιμαστικών επιβολών από την μόνιμη κατάσταση. Διερεύνηση με χρήση 24 επιβολών που αποτελούνται από 120 snapshots.

Συγκρίνοντας τις αποκλίσεις που προκύπτουν από κάθε προσπάθεια επίτευξης μόνιμης κατάστασης, με διαφορά τον αριθμό των επιβολών που χρησιμοποιούνται κάθε φορά για την κατασκευή του μοντέλου μειωμένης τάξης, παρατηρείται ότι ενώ στον κάτω κλάδο τα αποτελέσματα είναι πολύ κοντά, στον άνω κλάδο οι τιμές των σφαλμάτων μειώνονται σημαντικά. Για τον λόγο αυτό, στα επόμενα υπολογιστικά πειράματα που θα πραγματοποιηθούν επιλέγεται ο αριθμός των επιβολών που χρησιμοποιούνται για την κατασκευή του μοντέλου μειωμένης τάξης να είναι 24.

Η αύξηση του αριθμού των επιβολών που πραγματοποιήθηκε για να αυξηθεί η ακρίβεια του μοντέλου, είχε ως αποτέλεσμα και την αύξηση του συνολικού αριθμού των δεδομένων που χρησιμοποιούνται για την κατασκευή του. Για να ικανοποιούνται και τα κριτήρια του χαμηλού υπολογιστικού κόστους, της χαμηλής απαίτησης σε υπολογιστική μνήμη και της μεγάλης ταχύτητας υπολογισμών γίνεται μια περαιτέρω μελέτη του συστήματος, με μείωση του χρονικού διαστήματος που αφήνονται οι επιβολές να εξελιχθούν. Αναπτύσσεται, συνεπώς, μια ακόμη διερεύνηση στην οποία η κατασκευή της βάσης, η εκπαίδευση του νευρωνικού δικτύου και η πρόβλεψη γνωστών μονίμων καταστάσεων από το νευρωνικό δίκτυο, πραγματοποιούνται χρησιμοποιώντας και τις 24 επιβολές, με τη διαφορά ότι ο αριθμός των timesteps που εκτελεί η κάθε επιβολή μειώνεται στο μισό, δηλαδή 60.

Γ) 3η Διερεύνηση

i) Επιβολές

Οι επιβολές που χρησιμοποιούνται σε αυτή την διερεύνηση, έχουν όλες ως τελική κατάσταση την ίδια τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s αλλά εκτελούν μόνο τα πρώτα 60 timesteps της τροχιάς τους. Στην κατανόηση αυτού, βοηθάει το Διάγραμμα 5.1, στο οποίο φαίνεται ποιο κομμάτι των δεδομένων από τις τυχαίες τροχιές που εκτελούν κάποιες επιβολές, χρησιμοποιείται για την κατασκευή του μοντέλου μειωμένης τάξης. Στην προσέγγιση αυτή μειώνεται ο αριθμός των «κόκκινων τετραγώνων» στην αρχή της τροχιάς κάθε επιβολής.

Ο Πίνακας 5.7 είναι αντιπροσωπευτικός και για αυτή την διερεύνηση. Υπενθυμίζεται ότι ο αριθμός των επαναληπτικών βημάτων που εκτελεί κάθε προσομοίωση επιβολής σε ένα timestep μέσω του Fluent, παραμένει 50 σε όλη την εξέλιξη της εργασίας.

ii) Κατασκευή PCA Βάσης

Η βάση, όπως και σε όλα τα προηγούμενα πειράματα, κατασκευάστηκε με τη μέθοδο των στιγμιοτύπων, με χρήση του αλγορίθμου του Σχήματος 3.3 και έχει κλάσμα περιεχόμενης πληροφορίας $\rho=98$. Για την κατασκευή της χρησιμοποιήθηκαν ως τροχιές οι δυναμικές των 23 από τις 24 επιβολές, με την Επιβολή 5 να αποτελεί την εξωτερική επιβολή.

Στον παρακάτω πίνακα παρουσιάζεται το σχετικό πληροφοριακό περιεχόμενο για κάθε προσέγγιση PCA των τεσσάρων ιδιοσυναρτήσεων από τις οποίες αποτελείται, μέχρι τη βάση Z4.

Διάσταση προσέγγισης PCA, d	1	2	3	4
Σχετικό πληροφοριακό περιεχόμενο, I(d)	0,7513	0,9210	0,9777	0,9856

Πίνακας 5. 10: Σχετικό πληροφοριακό περιεχόμενο προσεγγίσεων PCA για διαφορετικές διαστάσεις βάσης. Διερεύνηση με χρήση 24 επιβολών οι οποίες αποτελούνται από 60 snapshots.

Παρατηρείται μια μικρή μείωση του πληροφοριακού περιεχομένου των ιδιοδιανυσμάτων συγκριτικά με το προηγούμενο πείραμα. Η μείωση αυτή είναι αναμενόμενη εφόσον μειώνεται ο αριθμός των snapshots των επιβολών που χρησιμοποιούνται για την κατασκευή της βάσης και συνεπώς ο αριθμός των δεδομένων από τα οποία αντλούνται πληροφορίες για το συνολικό σύστημα.

Η αξιολόγηση της βάσης πραγματοποιήθηκε με τις ίδιες επιβολές όπως και στις προηγούμενες διερευνήσεις, και τα σχετικά σφάλματα ήταν της τάξης μεγέθους $10^{-2} - 10^{-4}$, που είναι ικανοποιητικά για την παρούσα εργασία.

iii) Εκπαίδευση Νευρωνικού Δικτύου

Το νευρωνικό δίκτυο εκπαιδεύτηκε με χρήση της βάσης Z4, και για την εκπαίδευση του συγκεκριμένου δικτύου, χρησιμοποιήθηκαν 22 από τις 24 επιβολές, με τις ίδιες επιβολές όπως και στο νευρωνικό δίκτυο της Παραγράφου 5.3.B, να αποτελούν τις εξωτερικές. Το δίκτυο αυτό έχει $d=1$ και $L=10$ κρυφούς νευρώνες.

Πραγματοποιήθηκε αξιολόγηση του δικτύου που εκπαιδεύτηκε, μέσω πρόβλεψης γνωστών μονίμων καταστάσεων μέσω αυτού, και παρατηρήθηκε ότι οι συντελεστές σφαλμάτων του δικτύου κυμαίνονται στην ίδια τάξη μεγέθους όπως και οι συντελεστές των προηγούμενων νευρωνικών δικτύων.

iv) Υπολογισμός μονίμων καταστάσεων

Με χρήση της βάσης Z4 και του νευρωνικού δικτύου που κατασκευάστηκαν στην διερεύνηση αυτή, έγινε εκ νέου ο υπολογισμός τεσσάρων μονίμων καταστάσεων.

Ο υπολογισμός των αποκλίσεων της μόνιμης κατάστασης που υπολογίζεται από το νευρωνικό δίκτυο από την μόνιμη κατάσταση που υπολογίζεται από την προσομοίωση, πραγματοποιείται με χρήση του Πηλίκου 5.2. Στον Πίνακα 5.11 παρουσιάζονται οι τέσσερις καταστάσεις που προσεγγίστηκαν και το επί τις εκατό σφάλμα που προέκυψε για καθεμία.

Στην διερεύνηση αυτή μελετάται η μείωση του αριθμού timesteps που εκτελεί το σύνολο των επιβολών που χρησιμοποιούνται για την κατασκευή του μοντέλου μειωμένης τάξης. Όπως φαίνεται και στον παρακάτω Πίνακα 5.1, τα αποτελέσματα που προκύπτουν από τον υπολογισμό των μόνιμων καταστάσεων σε σύγκριση με αυτά της προηγούμενης διερεύνησης, στην οποία οι επιβολές εκτελούν 120 timesteps, δεν εμφανίζουν ουσιαστική διαφορά. Για τον λόγο αυτό, για την περαιτέρω μελέτη της συμπεριφοράς του συστήματος, πραγματοποιήθηκαν δύο επιπλέον έλεγχοι. Με τον ίδιο αριθμό επιβολών, και χρησιμοποιώντας τις ίδιες επιβολές ως εσωτερικές και εξωτερικές κατά την κατασκευή της βάσης και την εκπαίδευση του νευρωνικού δικτύου, αλλά και για τον έλεγχο τους κατά τον υπολογισμό γνωστών μόνιμων καταστάσεων, υπολογίστηκαν οι ίδιες μόνιμες καταστάσεις, χρησιμοποιώντας επιβολές στις οποίες:

- α) Διπλασιάστηκε η διάρκεια των χρονικών βημάτων (timesteps) από 0,1 σε 0,2s, διατηρώντας όμως σταθερό τον συνολικό αριθμό των timesteps σε 60,
- β) Μειώθηκε ο αριθμός των timesteps που εκτελούν οι επιβολές σε 30, διατηρώντας όμως την διάρκεια τους σταθερή σε 0,1 s.

Στον Πίνακα 5.11 παρουσιάζονται τα σφάλματα των αποκλίσεων όλων των παραπάνω περιπτώσεων.

A/A	Μόνιμη Κατάσταση		Διερευνήσεις			
	Παροχή	Κλάδος	2η	3η	3α	3β
	$m_f (\cdot 10^{-5})$ kg/s		24επιβολές	24επιβολές	24επιβολές	24επιβολές
			120 timesteps	60 timesteps	60 timesteps	60 timesteps dt=0.2 s
		Σφάλμα (%)	Σφάλμα(%)	Σφάλμα(%)	Σφάλμα(%)	
MK_1	1,5	(Κάτω)	2,42	2,38	2,41	2,4
MK_2	3,5	Κάτω	0,81	0,84	0,82	0,82
MK_3	5,5	Άνω	6,86	7,00	7,02	6,41
MK_4	6,5	Άνω	8,68	8,82	8,83	8,25

Πίνακας 5. 11: Αποκλίσεις κατά τον υπολογισμό μόνιμων καταστάσεων με χρήση των 4 διαφορετικών μοντέλων μειωμένης τάξης.

Από τον πίνακα 5.11 γίνεται εμφανές ότι δεν υπάρχουν ουσιαστικές διαφορές μεταξύ των τεσσάρων αυτών περιπτώσεων που μελετήθηκαν, ενώ ο αριθμός των snapshots που χρησιμοποιούνται κυμαίνεται στο εύρος από 120 μέχρι και 30. Αυτό που μπορεί να παρατηρηθεί σε όλες τις περιπτώσεις, είναι η μεγάλη διαφορά στις αποκλίσεις του κάτω κλάδου από αυτές του άνω.

Μια εξήγηση για την διαφορά που παρατηρείται μεταξύ των τιμών των σφαλμάτων των δύο κλάδων, θα μπορούσε να είναι ότι στον άνω κλάδο οι επιβολές χρειάζονται περισσότερο χρόνο για να συγκλίνουν στη λύση και συνεπώς χρησιμοποιώντας ακριβώς τον ίδιο αριθμό snapshots με τις επιβολές του κάτω κλάδου, οι οποίες είναι πιθανό να έχουν φτάσει πιο κοντά στη σύγκλιση, δημιουργούνται μεγαλύτερα σφάλματα στην πρόβλεψη των μόνιμων καταστάσεων του άνω κλάδου. Συνεπώς για απαντηθεί όσο επαρκέστερα γίνεται στα πλαίσια της διπλωματικής εργασίας και αυτό το ερώτημα, πραγματοποιείται μια ακόμη διερεύνηση στην οποία οι επιβολές που χρησιμοποιούνται για την κατασκευή του μοντέλου δεν αποτελούνται από τον ίδιο αριθμό snapshots. Συγκεκριμένα, όλες οι επιβολές θα εξελίσσονται και πάλι προς την ίδια μόνιμη κατάσταση, αλλά το κριτήριο με το οποίο θα επιλέγεται μετά από πόσα timestep θα σταματήσει η κάθε επιβολή είναι μια δεδομένη τιμή σφάλματος μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης, κατά την τροχιά της επιβολής προς την νέα κατάσταση.

Δ) 4η Διερεύνηση

i) Επιβολές

Σε αυτή την διερεύνηση μελετήθηκαν δύο περιπτώσεις, στην πρώτη (4α) οι επιβολές αφήνονται να εξελιχθούν από μια αρχική κατάσταση προς την ίδια μόνιμη κατάσταση με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ Kg/s, μέχρι η τιμή σφάλματος μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης (σχέση 5.1) να είναι μικρότερη από 10^{-4} , ενώ στην δεύτερη (4β) μέχρι η τιμή του σφάλματος αυτού να είναι μικρότερη από 10^{-3} .

Επίσης, για υπολογιστικούς λόγους αφαιρέθηκε η επιβολή N2 του Πίνακα 6.7, η αρχική κατάσταση της οποίας είχε τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s και άνηκε στον κάτω κλάδο, ενώ είχε υπολογιστεί με την μέθοδο που υπολογίστηκαν οι επιπλέον

αρχικές μόνιμες καταστάσεις στην παράγραφο 5.3.B. Έτσι προκύπτει ένα σύνολο 23 επιβολών η κάθε μια εκ των οποίων αποτελείται από διαφορετικό αριθμό snapshots.

Επισημαίνεται ότι στην δεύτερη περίπτωση στην οποία η τιμή του σφάλματος αυξάνεται, κάποιες επιβολές του κάτω κλάδου χρειαζόταν να εκτελέσουν πολύ μικρό αριθμό timesteps για να ξεπεράσουν την δεδομένη τιμή του σφάλματος και συνεπώς αποτελούνταν από πολύ μικρό αριθμό στιγμιοτύπων. Για τον λόγο αυτό ως ελάχιστος αριθμός snapshots για μια επιβολή επιλέχθηκαν τα 10.

Στον παρακάτω πίνακα παρουσιάζονται αναλυτικά οι επιβολές και ο αριθμός snapshots που εκτελεί η κάθε επιβολή.

A/A	Αρχική κατάσταση		4α snapshots $aa < 10^{-4}$	4β snapshots $aa < 10^{-3}$
	Παροχή	Κλάδος		
	$m_0 \cdot 10^{-5}$ kg/s			
1	1,5	(Κάτω)	45	10
N1	2	(Κάτω)	38	10
N8	2,25	(Κάτω)	33	10
2	2,5	Κάτω	27	10
3	2,5	Άνω	64	10
N9	2,75	Κάτω	16	10
N3	3	Άνω	26	10
N10	3,25	Κάτω	13	10
N11	3,25	Άνω	35	10
4	3,5	Κάτω	25	10
5	3,5	Άνω	53	10
N12	3,75	Κάτω	33	10
N13	3,75	Άνω	63	10
N4	4	Κάτω	40	10
N5	4	Άνω	70	17
N14	4,25	Άνω	75	23
6	4,5	Κάτω	48	10
7	4,5	Άνω	79	27
N15	4,75	Άνω	79	27
N6	5	Άνω	79	27
8	5,5	(Άνω)	79	27
N7	6	(Άνω)	81	29
9	6,5	(Άνω)	83	30

Πίνακας 5. 12: Συνολικός αριθμός επιβολών που χρησιμοποιούνται στην τέταρτη διερεύνηση και συνολικός αριθμός snapshots που χρησιμοποιείται για κάθε επιβολή των δύο περιπτώσεων 4α και 4β.

Και στις δύο περιπτώσεις πραγματοποιήθηκε η κατασκευή βάσης, η εκπαίδευση του νευρωνικού δικτύου και ο υπολογισμός των μόνιμων καταστάσεων χρησιμοποιώντας τις ίδιες επιβολές, με τη διαφορά τον αριθμό των snapshots κάθε επιβολής, αναλόγως την περίπτωση. Παρακάτω, θα αναλυθεί εκτενέστερα η κατασκευή της βάσης και η εκπαίδευση του νευρωνικού δικτύου της δεύτερης περίπτωσης (4β), ενώ στο τέλος

της παραγράφου θα παρουσιαστούν οι μόνιμες καταστάσεις που υπολογίστηκαν και στις δύο περιπτώσεις.

ii) Κατασκευή PCA Βάσης

Η βάση κατασκευάστηκε με τη μέθοδο των στιγμιοτύπων, με χρήση του αλγορίθμου του Σχήματος 3.3 και έχει κλάσμα περιεχόμενης πληροφορίας $\rho=98$. Για την κατασκευή της χρησιμοποιήθηκαν ως τροχιές οι δυναμικές των 22 από τις 23 επιβολές, με την Επιβολή 5 να αποτελεί την εξωτερική επιβολή.

Στον παρακάτω πίνακα παρουσιάζεται το σχετικό πληροφοριακό περιεχόμενο για κάθε προσέγγιση PCA των τεσσάρων ιδιοσυναρτήσεων από τις οποίες αποτελείται, μέχρι τη βάση Z4.

Διάσταση προσέγγισης POD, d	1	2	3	4
Σχετικό πληροφοριακό περιεχόμενο, I(d)	0,7362	0,9018	0,9633	0,9802

Πίνακας 5. 13: Σχετικό πληροφοριακό περιεχόμενο προσεγγίσεων PCA για διαφορετικές διαστάσεις βάσης. Διερεύνηση που χρησιμοποιεί 23 επιβολές τα snapshots των οποίων επιλέγονται με βάση την τιμή του σφάλματος.

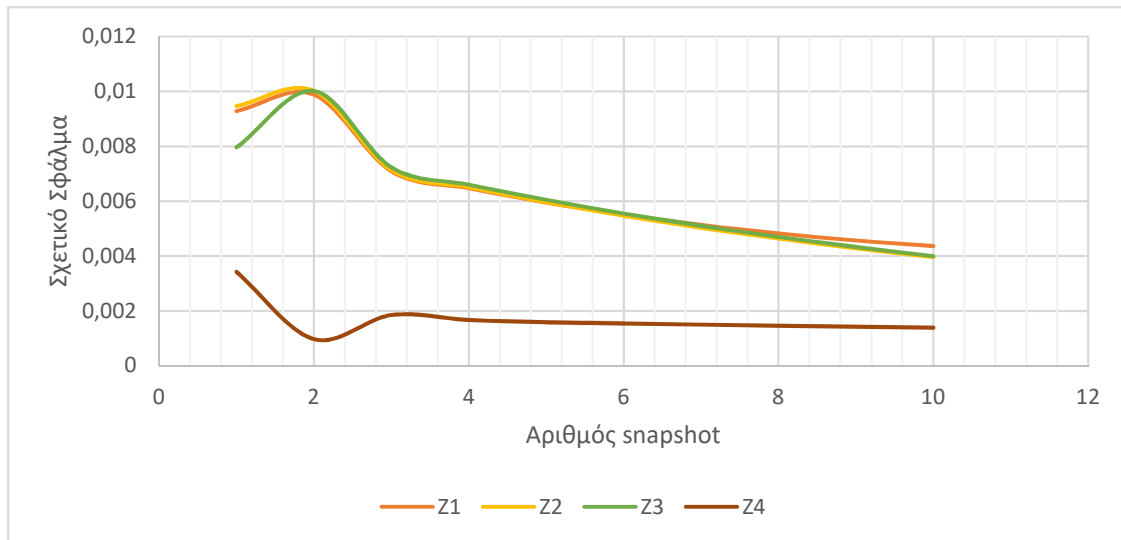
Παρατηρείται ότι το πληροφοριακό περιεχόμενο των επιβολών, που περιέχει το κάθε ιδιοδιάνυσμα, μειώνεται συγκριτικά με το προηγούμενο πείραμα. Η αξιολόγηση της βάσης πραγματοποιήθηκε με τις ίδιες επιβολές όπως και στις προηγούμενες διερευνήσεις, και τα σχετικά σφάλματα ήταν της τάξης μεγέθους $10^{-2} - 10^{-4}$, που είναι ικανοποιητικά για την παρούσα εργασία.

Στη συνέχεια αξιολογήθηκαν οι τέσσερις παραπάνω βάσεις ως προς το σφάλμα που παρουσιάζουν στην προσέγγιση των ίδιων γνωστών επιβολών όπως σε όλα τα προηγούμενα πειράματα. Τα σφάλματα αυτά υπολογίστηκαν για την εξωτερική Επιβολή 5 και για την εσωτερική Επιβολή 2 (Πίνακας 5.12) και σημειώνονται στον Πίνακα 5.14. Συγκεκριμένα τα σφάλματα που παρουσιάζονται είναι το ελάχιστο σφάλμα για κάθε βάση, το μέγιστο, καθώς και το τελικό σφάλμα δηλαδή το σφάλμα του τελευταίου snapshot.

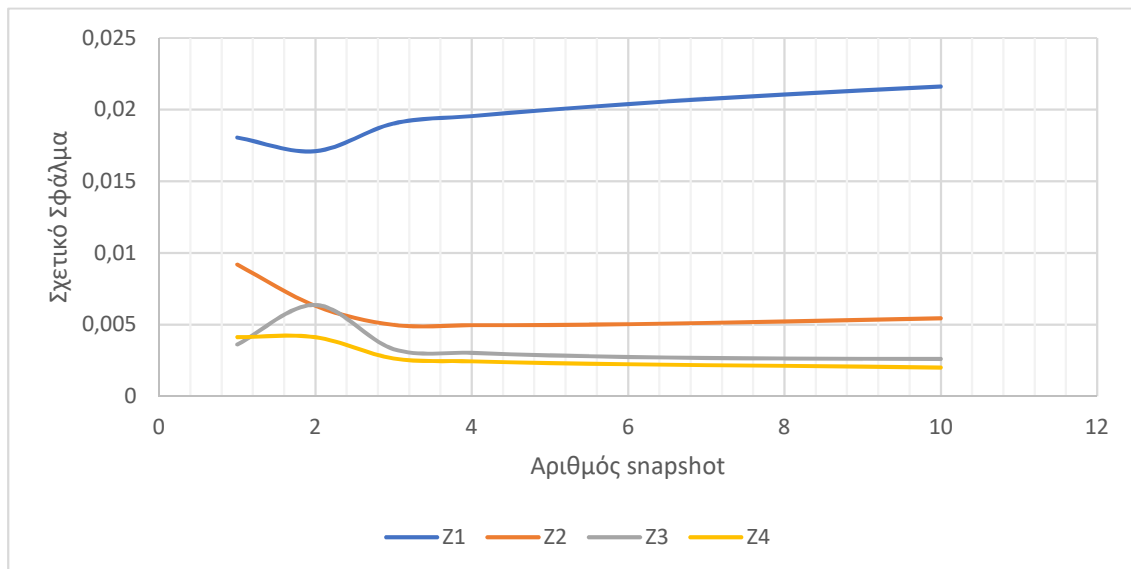
Βάσεις	Εσωτερική Επιβολή 2 $\dot{m}_0=2,5 \cdot 10^{-5}$ kg/s σε $\dot{m}_r=3 \cdot 10^{-5}$ kg/s (Κάτω κλάδος)			Εξωτερική Επιβολή 5 $\dot{m}_0=3,5 \cdot 10^{-5}$ kg/s σε $\dot{m}_r=3 \cdot 10^{-5}$ kg/s (Άνω κλάδος)		
	Min	Max	Τελικό Σφάλμα	Min	Max	Τελικό Σφάλμα
Z1	0,0044	0,0099	0,00436	0,0171	0,0216	0,0216
Z2	0,0040	0,0100	0,00396	0,0050	0,0092	0,0054
Z3	0,0040	0,0100	0,00400	0,0026	0,0064	0,0026
Z4	0,0010	0,0034	0,00139	0,0020	0,0041	0,0020

Πίνακας 5. 14: Σχετικά σφάλματα προσεγγίσεων PCA για διαφορετικών διαστάσεων βάσεις, κατά την προσέγγιση μιας εσωτερικής επιβολής και μιας εξωτερικής. Διερεύνηση με χρήση 23 επιβολών, τα snapshots των οποίων επιλέγονται με βάση την τιμή του σφάλματος.

Στα Διαγράμματα 5.14 και 5.15 παρουσιάζεται το σχετικό σφάλμα κατά την προσέγγιση των δύο επιβολών, για τις τέσσερις διαφορετικές βάσεις συναρτήσει του αριθμού των timesteps που εκτελούνται σε κάθε επιβολή.



Διάγραμμα 5. 14: Σχετικό σφάλμα κάθε βάσης για την εσωτερική επιβολή από $\dot{m}_i(t)=2,5 \cdot 10^{-5}$ kg/s (κάτω κλάδος) σε $\dot{m}_i(t)=3 \cdot 10^{-5}$ kg/s. Διερεύνηση με χρήση προσομοιώσεων 23 επιβολών, τα snapshots των οποίων επιλέγονται με βάση την τιμή του σφάλματος.



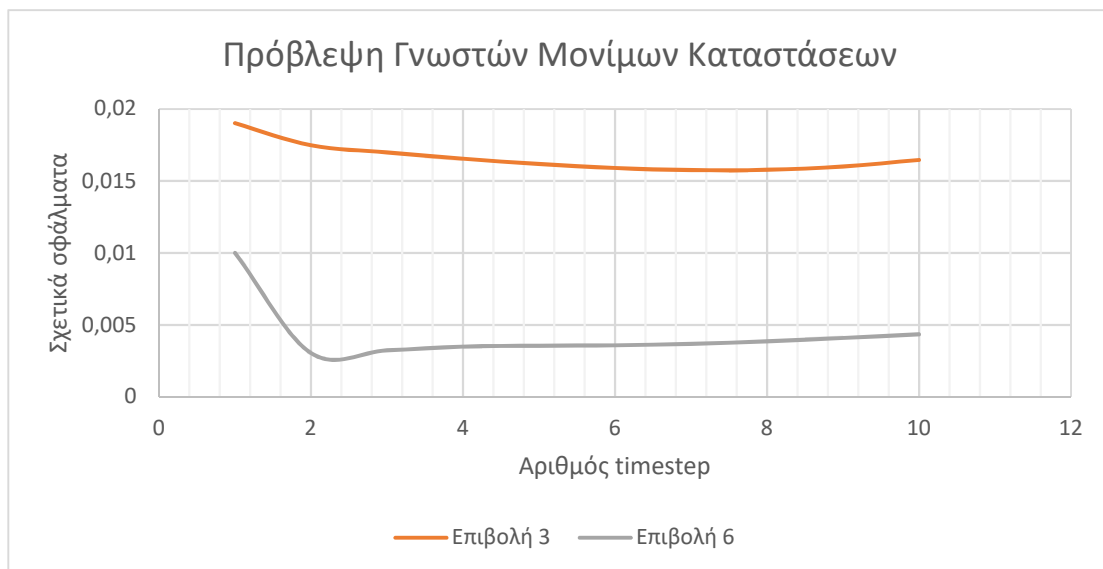
Διάγραμμα 5. 15: Σχετικό σφάλμα κάθε βάσης για την εξωτερική επιβολή από $\dot{m}_i(t)=3,5 \cdot 10^{-5}$ kg/s (Άνω κλάδος) σε $\dot{m}_i(t)=3 \cdot 10^{-5}$ kg/s. Διερεύνηση με χρήση προσομοιώσεων 23 επιβολών, τα snapshots των οποίων επιλέγονται με βάση την τιμή του σφάλματος.

Από τα παραπάνω διαγράμματα παρατηρείται ότι η βάση Z4 που αποτελείται και από τα τέσσερα ιδιοδιανύσματα και που περιέχει το μεγαλύτερο πληροφοριακό περιεχόμενο, παρουσιάζει το μικρότερο σφάλμα. Για τον λόγο αυτό, αλλά και για να είναι εφικτή η σύγκριση των τελικών αποτελεσμάτων με τα προηγούμενα πειράματα, επιλέγεται για την κατασκευή της βάσης η χρήση και των τεσσάρων ιδιοδιανυσμάτων.

iii) Εκπαίδευση Νευρωνικού Δικτύου

Το νευρωνικό δίκτυο εκπαιδεύτηκε με χρήση της βάσης Z4 που κατασκευάστηκε σε αυτή τη διερεύνηση, και για την εκπαίδευση του δικτύου χρησιμοποιήθηκαν 21 από τις 23 επιβολές. Η Επιβολή 3 και 6 (Πίνακας 5.12) δεν αποτέλεσαν επιβολές εκπαίδευσης και συνεπώς καλούνται εξωτερικές. Το δίκτυο αυτό έχει καθυστέρηση $d=1$ και το κρυφό επίπεδο αποτελείται από $L=10$ νευρώνες.

Για την αξιολόγηση του δικτύου πραγματοποιήθηκε η πρόβλεψη, μέσω αυτού, γνωστών μόνιμων καταστάσεων του πλήρους συστήματος και μελετήθηκαν τα σχετικά σφάλματα που προκύπτουν. Προσεγγίστηκαν και οι 23 επιβολές του Πίνακα 5.12, και οι τιμές των σφαλμάτων ήταν της τάξης μεγέθους $10^{-2} - 10^{-4}$ που είναι ικανοποιητικά για τα δεδομένα της εργασίας. Για λόγους οικονομίας, παρακάτω θα παρουσιαστούν τα διαγράμματα των σχετικών σφαλμάτων που προκύπτουν κατά την προσέγγιση δύο γνωστών μόνιμων καταστάσεων του συστήματος, που όμως δεν αποτελούν επιβολές εκπαίδευσης για το νευρωνικό δίκτυο. Συγκεκριμένα προβάλλονται τα σφάλματα που προκύπτουν από την πρόβλεψη της Επιβολής 3 και της Επιβολής 6 (Πίνακας 5.12), του μοντέλου πλήρους τάξης.



Διάγραμμα 5. 16: Σχετικό σφάλμα της προσέγγισης του Νευρωνικού δικτύου, ως προς την πρόβλεψη των εξωτερικών Επιβολών 3 και 6. (Πίνακας 5.12). Διερεύνηση με χρήση προσομοιώσεων 23 επιβολών, τα snapshots των οποίων επιλέγονται με βάση την τιμή του σφάλματος

iv) Υπολογισμός μόνιμων καταστάσεων

Όπως αναφέρθηκε και στην Παράγραφο 5.3.Δ, σε αυτήν την διερεύνηση μελετήθηκαν δύο διαφορετικές περιπτώσεις σχετικά με την τιμή του σφάλματος κατά την εξέλιξη μιας επιβολής, από μια εκ των γνωστών αρχικών μόνιμων καταστάσεων, προς την τελική μόνιμη κατάσταση. Με χρήση της βάσης PCA και του νευρωνικού δικτύου που κατασκευάστηκαν σε κάθε μια εκ των δύο περιπτώσεων, πραγματοποιήθηκε ο υπολογισμός των τεσσάρων μόνιμων καταστάσεων του μελετώμενου συστήματος.

Με χρήση του Πηλίκου 5.2 υπολογίστηκαν έπειτα, οι αποκλίσεις των μονίμων καταστάσεων που υπολογίζονται από το μοντέλο μειωμένης τάξης που κατασκευάστηκε, από τις μόνιμες καταστάσεις που υπολογίζονται από τον κώδικα υπολογιστικής ρευστοδυναμικής Ansys Fluent.

Επειδή τα αποτελέσματα ήταν πολύ ικανοποιητικά, υπολογίστηκαν δέκα ακόμη καταστάσεις, στο εύρος τιμών που μελετάται το σύστημα, μέσω του λογισμικού Fluent. Στη συνέχεια οι καταστάσεις αυτές προσεγγίστηκαν από το μοντέλο μειωμένης τάξης. Είναι σημαντικό να αναφερθεί ότι οι μόνιμες καταστάσεις που υπολογίστηκαν εφόσον δεν έχουν χρησιμοποιηθεί για την κατασκευή του μοντέλου αποτελούν τελείως άγνωστες καταστάσεις για το σύστημα. Για να μελετηθεί η ορθότητα της πρόβλεψης του μοντέλου στα σημεία στροφής του κλάδου, κάποιες από τις καταστάσεις επιλέχθηκε να βρίσκονται κοντά στα σημεία αυτά.

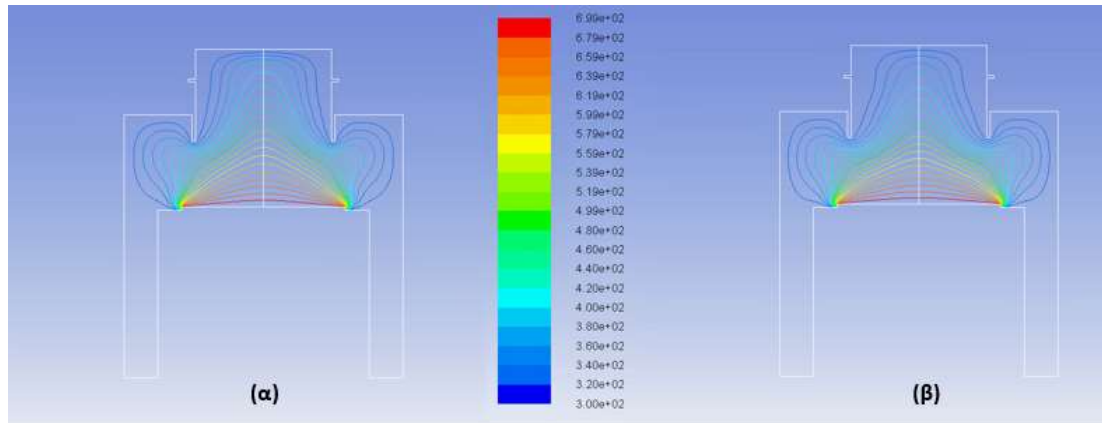
Οι αποκλίσεις αυτές παρουσιάζονται στον παρακάτω πίνακα:

A/A	Steady State ($\cdot 10^{-5}$ kg/s)	Κλάδος	Διερευνήσεις	
			4 (α) 23 επιβολές $aa < 10^{-4}$	4(β) 23 επιβολές $aa < 10^{-3}$
			Σφάλμα (%)	Σφάλμα (%)
MK_1	1,5	Κάτω	2,43	2,47
MK_2	3,5	Κάτω	0,79	0,77
MK_3	5,5	Άνω	6,68	5,88
MK_4	6,5	Άνω	8,51	7,74
MK_5	7	Άνω	9,45	8,69
MK_6	5,2	Άνω	6,07	5,26
MK_7	4	Άνω	3,19	2,38
MK_8	5,1	Άνω	5,86	5,05
MK_9	4,9	Άνω	5,42	4,63
MK_10	3,5	Άνω	1,59	0,82
MK_11	2,3	Άνω	5,43	6,15
MK_12	2,3	Κάτω	1,23	1,27
MK_13	1,9	Κάτω	1,87	1,91
MK_14	3,2	Κάτω	0,29	0,27

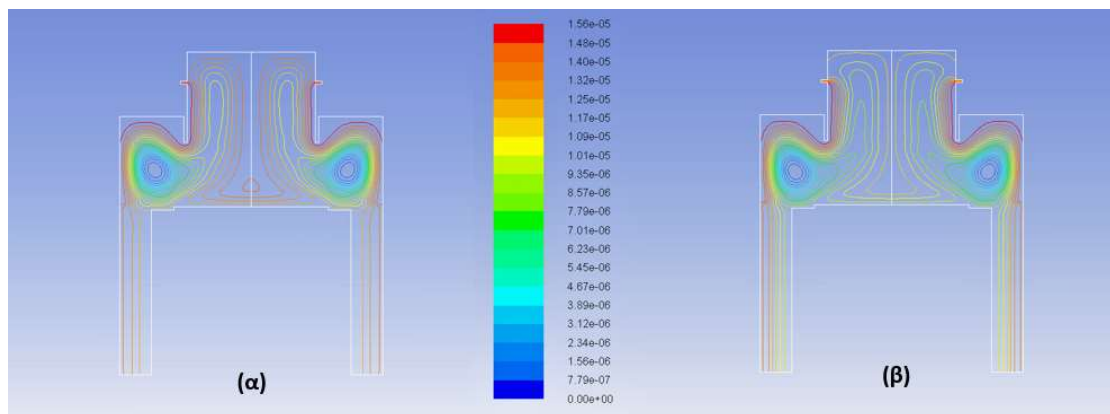
Πίνακας 5. 15: Αποκλίσεις κατά τον υπολογισμό μόνιμων καταστάσεων με χρήση των 2 διαφορετικών μοντέλων μειωμένης τάξης.

Οι αποκλίσεις μεταξύ των γνωστών μόνιμων καταστάσεων που υπολογίστηκαν, MK_1-MK_4, είναι πάρα πολύ ικανοποιητικές, καθώς το μεγαλύτερο ποσοστό των σφαλμάτων μειώθηκε συγκριτικά με τα προηγούμενα πειράματα. Η πρόβλεψη των άγνωστων μόνιμων καταστάσεων MK_5 – MK_14 επίσης είχε σφάλματα στο εύρος που κυμαίνονται οι προβλέψεις για τις γνωστές καταστάσεις του συστήματος, ακόμα και κοντά στα σημεία στροφής των δύο κλάδων.

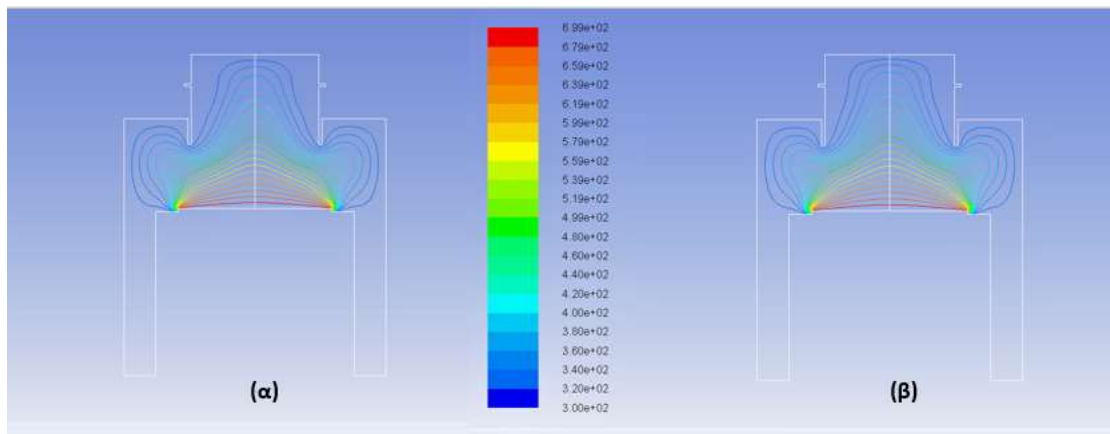
Στα διαγράμματα από 5.17 έως 5.24, παρουσιάζονται οι κατανομές των θερμοκρασιών και των ταχυτήτων στον αντιδραστήρα χημικής απόθεσης από ατμό, για τις μόνιμες καταστάσεις MK_1 – MK_4, που υπολογίστηκαν από το νευρωνικό δίκτυο της διερεύνησης 4β. Το αριστερό διάγραμμα παρουσιάζει την κατανομή που έχει προκύψει από την προσομοίωση μέσω του Fluent, ενώ το δεξιό την κατανομή που έχει προβλεφθεί από το νευρωνικό δίκτυο που εκπαιδεύτηκε στην παράγραφο αυτή.



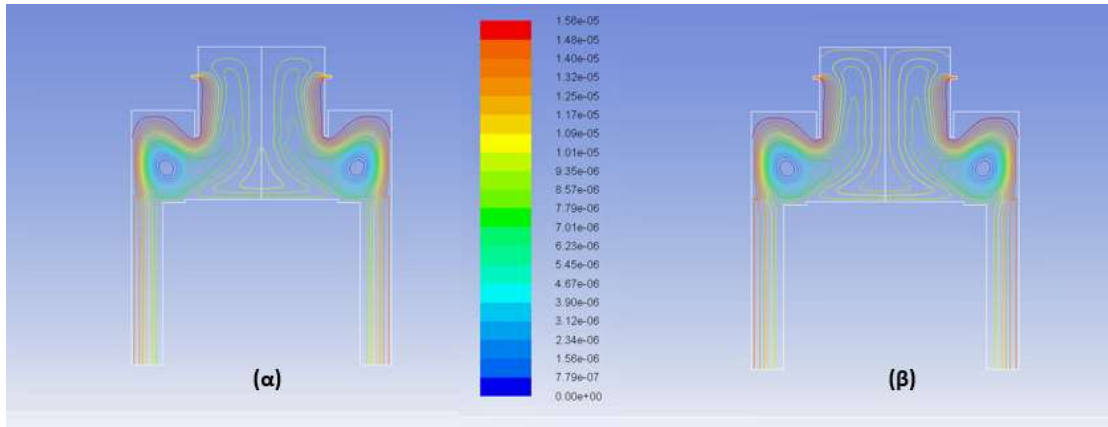
Διάγραμμα 5. 17: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=1,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



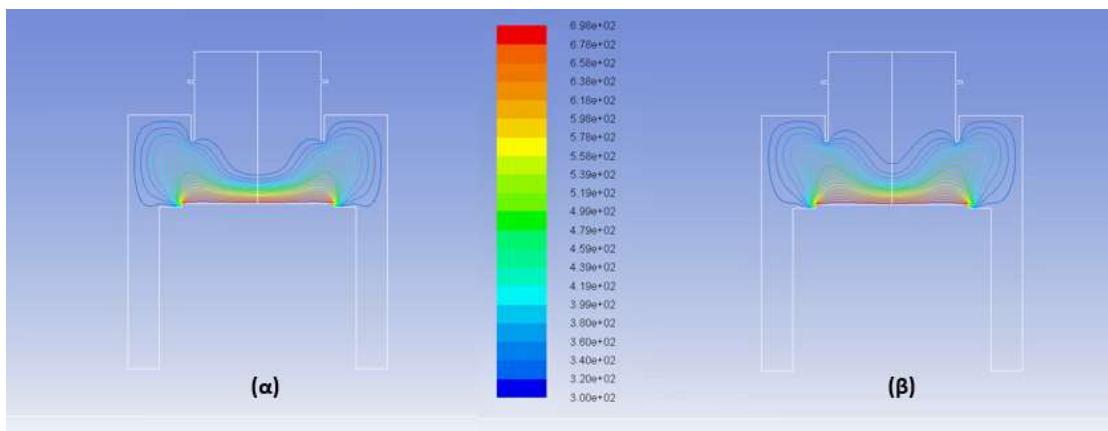
Διάγραμμα 5. 18: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=1,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



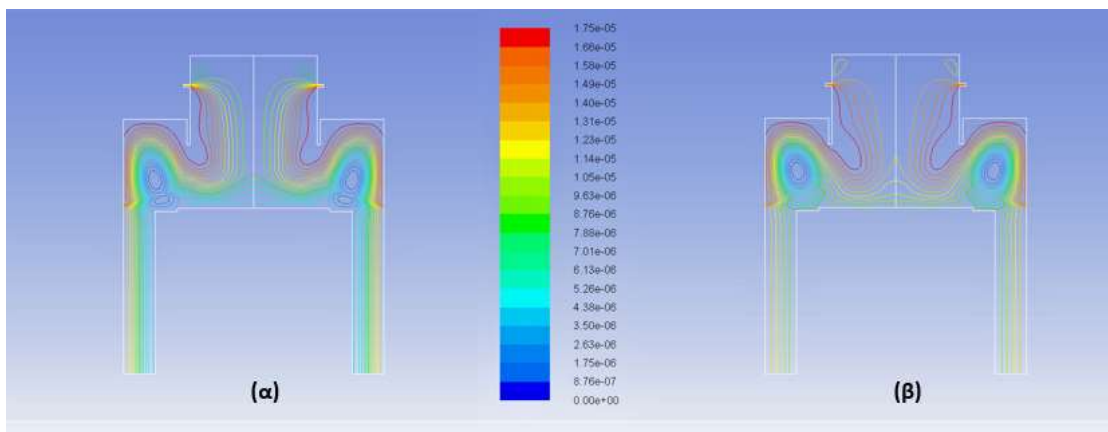
Διάγραμμα 5. 19: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=3,5 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



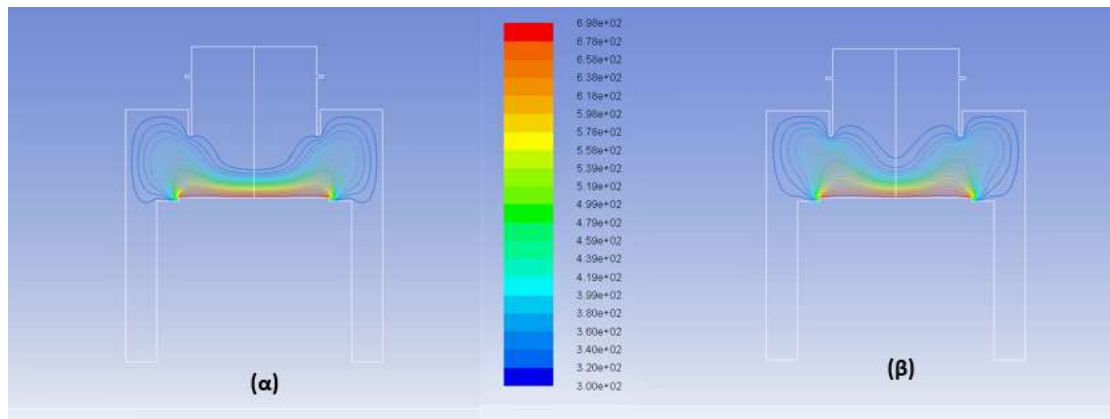
Διάγραμμα 5. 20: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=3,5 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται α) CFD, β) ANN



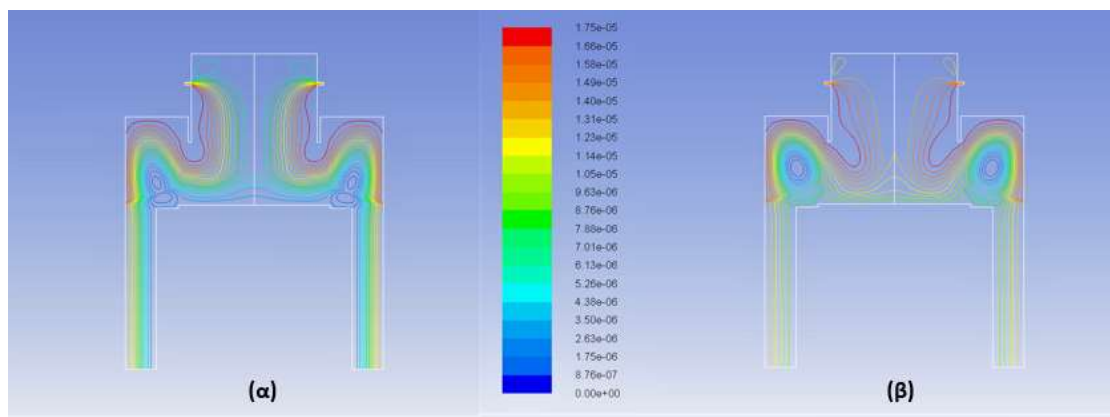
Διάγραμμα 5. 21: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=5,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



Διάγραμμα 5. 22: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=5,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN



Διάγραμμα 5. 23: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=6,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



Διάγραμμα 5. 24: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=6,5 \cdot 10^{-5}$ kg/s. Προσέγγιση που υπολογίζεται α) CFD, β) ANN

Όπως φαίνεται και από την προσομοίωση του αντιδραστήρα, οι μόνιμες καταστάσεις που προβλέφθηκαν από το μοντέλο μειωμένης τάξης έχουν μικρές διαφορές από αυτές που υπολογίζονται μέσω του λογισμικού Fluent. Τα σφάλματα που παρατηρούνται στον άνω κλάδο δεν έφτασαν τις τιμές αυτών του κάτω, αλλά μειώθηκαν σε μεγάλο βαθμό. Το γεγονός αυτό δείχνει ότι είναι πιο ορθή η μελέτη του συστήματος και συνεπώς η επιλογή του αριθμού των snapshots, μέσω μιας τιμής σφάλματος κατά την εξέλιξη των επιβολών που χρησιμοποιούνται, από την αυθαίρετη επιλογή του ίδιου αριθμού snapshots για όλες τις επιβολές των δύο κλάδων.

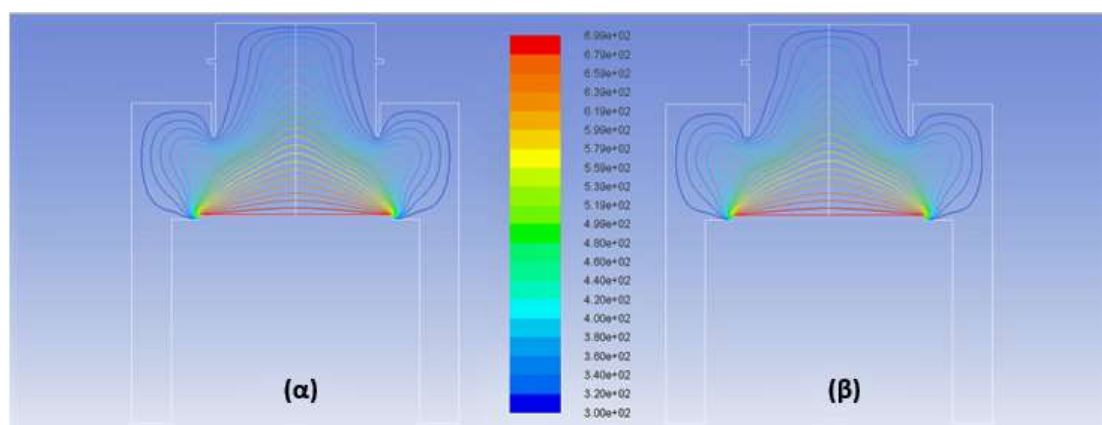
Στον παρακάτω πίνακα παρουσιάζονται συνολικά τα σφάλματα των τεσσάρων μόνιμων καταστάσεων που υπολογίζονταν σε κάθε διερεύνηση.

Διερεύνηση\MK	$1,5 \cdot 10^{-5}$ kg/s	$3,5 \cdot 10^{-5}$ kg/s	$5,5 \cdot 10^{-5}$ kg/s	$6,5 \cdot 10^{-5}$ kg/s
1	2,38	0,85	8,41	10,15
2	2,42	0,81	6,86	8,68
3	2,38	0,84	7	8,82
3i	2,41	0,82	7,02	8,83
3ii	2,4	0,82	6,41	8,25
4α	2,43	0,79	6,68	8,51
4β	2,47	0,77	5,88	7,74

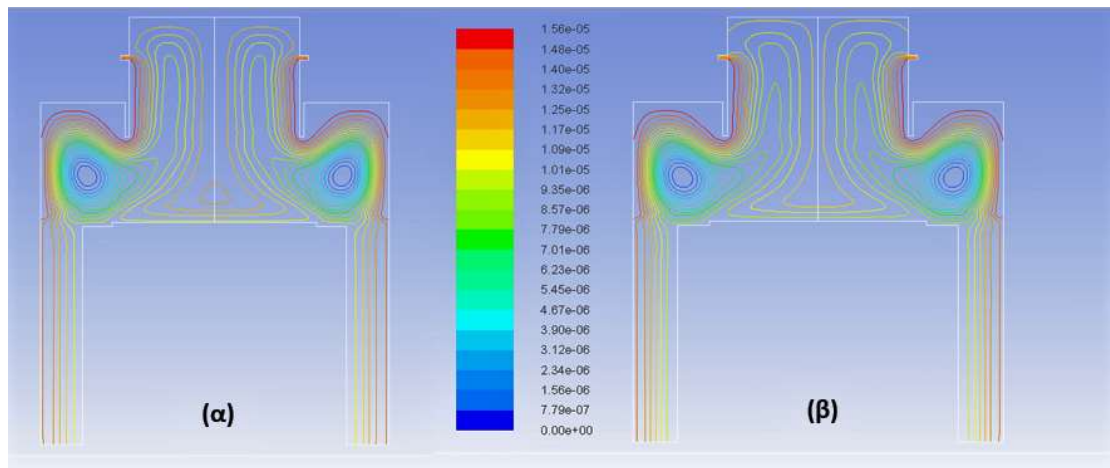
Πίνακας 5. 16: Επί τοις εκατό σφάλματα κατά την προσέγγιση των τεσσάρων γνωστών μόνιμων καταστάσεων. Κάθε στήλη του πίνακα περιέχει τα σφάλματα μιας μόνιμης κατάστασης σε καθεμία από τις διερευνήσεις που πραγματοποιήθηκαν.

Εκτός από την προσέγγιση της πρώτης μόνιμης κατάστασης κατά την οποία το σφάλμα της τελευταίας διερεύνησης αυξάνεται στο πρώτο δεκαδικό ψηφίο, όλα τα υπόλοιπα σφάλματα της τελευταίας διερεύνησης (4β) μειώνονται συγκριτικά με τις προηγούμενες.

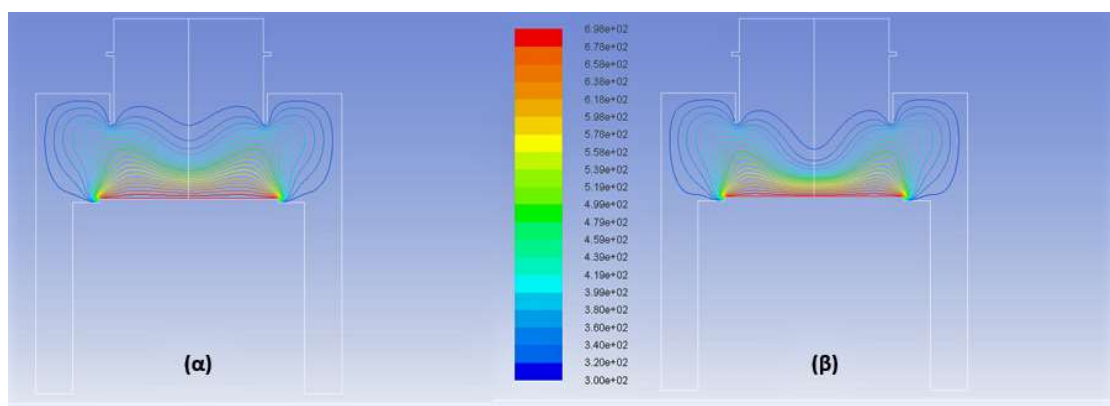
Στη συνέχεια παρουσιάζονται τα διαγράμματα των κατανομών της θερμοκρασίας και της ταχύτητας στην προσομοίωση του αντιδραστήρα, που προέκυψαν κατά την πρόβλεψη των μόνιμων καταστάσεων στο σημείο στροφής του συστήματος, με τιμή μαζικής παροχής $2,3 \cdot 10^{-5}$ kg/s, αρχικά στον κάτω κλάδο κι έπειτα στον άνω. Το αριστερό διάγραμμα παρουσιάζει την κατανομή που έχει προκύψει από την προσομοίωση μέσω του Fluent, ενώ το δεξιό την κατανομή που έχει προβλεφθεί από το μοντέλο μειωμένης τάξης που κατασκευάστηκε στην παράγραφο αυτή.



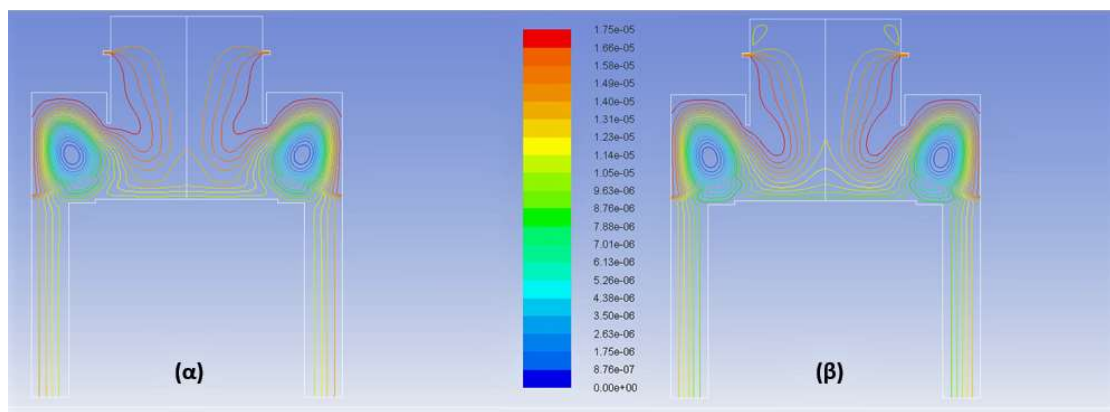
Διάγραμμα 5. 25: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=2,3 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



Διάγραμμα 5. 26: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=2,3 \cdot 10^{-5}$ kg/s (Κάτω κλάδος). Προσέγγιση που υπολογίζεται α) CFD, β) ANN



Διάγραμμα 5. 27: Κατανομή Θερμοκρασίας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=2,3 \cdot 10^{-5}$ kg/s (Ανω κλάδος). Προσέγγιση που υπολογίζεται: α) CFD, β) ANN



Διάγραμμα 5. 28: Κατανομή ταχύτητας στον αντιδραστήρα XAA, για την μόνιμη κατάσταση με τιμή παροχής $m_f=2,3 \cdot 10^{-5}$ kg/s (Ανω κλάδος). Προσέγγιση που υπολογίζεται α) CFD, β) ANN

Από τα διαγράμματα των προσομοιώσεων του αντιδραστήρα για την μόνιμη κατάσταση στην περιοχή που πραγματοποιείται η στροφή των δύο κλάδων, προκύπτει εύλογα το συμπέρασμα ότι, παρά την δυσκολία της προσέγγισης στην περιοχή αυτή λόγω της ιδιαιτερότητας που παρουσιάζει το φυσικό σύστημα, το μοντέλο μειωμένης τάξης που κατασκευάστηκε στην ενότητα αυτή μπορεί και προσεγγίζει την δυναμική του συστήματος με ικανοποιητική ακρίβεια.

Αν αθροιστεί η τελευταία στήλη του Πίνακα 5.12, που δείχνει τον αριθμό στιγμιοτύπων κάθε επιβολής της διερεύνησης (4β), προκύπτει το πλήθος των δεδομένων (Data) που χρησιμοποιήθηκε για την κατασκευή του μοντέλου. Στην διερεύνηση αυτή (4β) προκύπτει ότι ο αριθμός των δεδομένων που χρησιμοποιήθηκαν είναι 367, ενώ στην πρώτη διερεύνηση στην οποία οι επιβολές αποτελούνταν όλες από τον ίδιο αριθμό snapshots ο αριθμός δεδομένων ήταν $9(\text{επιβολές}) \times 120(\text{snapshots}) = 1080 (\text{data})$. Η επιτυχία της μεθόδου είναι φανερή καθώς με σχεδόν τριπλάσια μείωση του αριθμού δεδομένων που χρησιμοποιούνται, οι προσεγγίσεις των μόνιμων καταστάσεων με χρήση του μοντέλου ήταν καλύτερες από αυτές της πρώτης αλλά και των υπολοίπων διερευνήσεων.

Στην τελευταία διερεύνηση φανερώνεται πόσο σημαντική είναι η ορθή επιλογή των δεδομένων που χρησιμοποιούνται κατά την κατασκευή ενός Μοντέλου Μειωμένης Τάξης. Ενώ αποτελεί την διερεύνηση κατά την οποία η κατασκευή της PCA βάσης και η εκπαίδευση του νευρωνικού δικτύου πραγματοποιούνται με τον μικρότερο αριθμό δεδομένων, μπορεί και προσεγγίζει το πλήρες μελετώμενο σύστημα με την μεγαλύτερη ακρίβεια, συγκριτικά με τα υπόλοιπα υπολογιστικά πειράματα που πραγματοποιήθηκαν κατά την εξέλιξη της εργασίας.

6. Συμπεράσματα & Προτάσεις

6.1 Συμπεράσματα

Κατά την διεξαγωγή της παρούσας εργασίας, πραγματοποιήθηκε ενδελεχής μελέτη της δυναμικής συμπεριφοράς ενός αντιδραστήρα χημικής απόθεσης από ατμό, στον οποίο θεωρούνται αμελητέες οι πιθανές χημικές αντιδράσεις μεταξύ των αντιδρώντων. Η μελέτη της διεργασίας πραγματοποιήθηκε χρησιμοποιώντας αποτελέσματα προσομοιώσεων του αντιδραστήρα στον εμπορικό κώδικα υπολογιστικής ρευστοδυναμικής Ansys Fluent και κατασκευάζοντας έπειτα τέσσερα μη-γραμμικά μοντέλα μειωμένης τάξης για τους δύο ευσταθείς κλάδους του συστήματος. Τα μοντέλα μειωμένης τάξης κατασκευάστηκαν χρησιμοποιώντας μια εναλλακτική εφαρμογή της ανάλυσης κυρίων συνιστωσών και με την εκπαίδευση κατάλληλων τεχνητών νευρωνικών δικτύων με ανατροφοδότηση. Η εναλλακτική υλοποίηση της μεθόδου έγκειται στο ότι για την κατασκευή της ορθοκανονικής βάσης χρησιμοποιούνται δεδομένα από μη μόνιμες καταστάσεις. Το φυσικό σύστημα παρουσιάζει εξάρτηση τόσο από το χώρο όσο από το χρόνο, την χωρική εξάρτηση απαλείφει η ανάλυση κυρίων συνιστωσών ενώ την χρονική εξέλιξη του συστήματος προβλέπουν τα τεχνητά νευρωνικά δίκτυα.

Στόχος της εργασίας ήταν η κατασκευή ενός μοντέλου μειωμένης τάξης, που να περιγράφει με ικανοποιητική ακρίβεια την δυναμική συμπεριφορά του συστήματος και στους δύο ευσταθείς κλάδους, χρησιμοποιώντας δεδομένα που λαμβάνονται από μη-μόνιμες καταστάσεις. Η ιδέα βασίστηκε στη μελέτη που πραγματοποίησαν οι Graham και Kevrekidis [3] ως προς τις εναλλακτικές προσεγγίσεις της ανάλυσης κυρίων συνιστωσών, για την αντιμετώπιση ενός εκ των περιορισμών της μεθόδου που είναι η προσέγγιση μέσω αυτής, μόνο της δυναμικής του υποσυνόλου από το οποίο λαμβάνονται τα δεδομένα και όχι του πλήρους συστήματος. Για την επίτευξη αυτού, λαμβάνονται δεδομένα από την χρονική εξέλιξη του συστήματος καθώς μεταβαίνει από τυχαία επιλεγμένες μόνιμες καταστάσεις, εντός του εύρους στο οποίο εμφανίζεται η πολλαπλότητα των λύσεων, προς την ίδια κάθε φορά τελική κατάσταση με τιμή μαζικής παροχής $3 \cdot 10^{-5}$ kg/s (επιβολή). Με χρήση του λογισμικού Matlab το σύστημα δεν αφήνεται να συγκλίνει στη λύση, αλλά σταματάει μετά από ένα συγκεκριμένο χρονικό διάστημα.

Στα τέσσερα μοντέλα μειωμένης τάξης που κατασκευάστηκαν πραγματοποιήθηκε μελέτη ως προς τον αριθμό των επιβολών του συστήματος που χρησιμοποιούνται για την λήψη των κατάλληλων δεδομένων και του χρονικού διαστήματος στο οποίο αφήνονται οι αρχικές καταστάσεις να εξελιχθούν. Στα τρία πρώτα μοντέλα όλες οι αρχικές καταστάσεις είτε ανήκουν στον κάτω ευσταθή κλάδο είτε στον άνω αφήνονται να εξελιχθούν για το ίδιο χρονικό διάστημα, που είναι διαφορετικό για κάθε διερεύνηση. Στην τέταρτη διερεύνηση όλες οι αρχικές καταστάσεις αφήνονται να εξελιχθούν μέχρι να ξεπεράσουν μια συγκεκριμένη τιμή σφάλματος, μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης κατά την διάρκεια της τροχιάς τους. Παρατηρήθηκε ότι οι αρχικές καταστάσεις που ανήκουν στον κάτω κλάδο ξεπερνούν

πολύ πιο σύντομα την τιμή αυτή συγκριτικά με τις αρχικές καταστάσεις που ανήκουν στον άνω κλάδο.

Σε κάθε μια διερεύνηση προσεγγίστηκαν τέσσερις τυχαία επιλεγμένες, γνωστές μόνιμες καταστάσεις του συστήματος και αποδείχθηκε ότι το μοντέλο με την μεγαλύτερη ακρίβεια κατά την πρόβλεψη των καταστάσεων αυτών ήταν το τέταρτο. Για την κατασκευή του τέταρτου μοντέλου χρησιμοποιήθηκαν 23 επιβολές, οι οποίες αφέθηκαν να εξελιχθούν μέχρι η τιμή του σφάλματος μεταξύ δύο διαδοχικών διανυσμάτων κατάστασης να είναι μικρότερη από 10^{-3} . Αποτελεί το μοντέλο που χρησιμοποιεί τα λιγότερα στο σύνολο δεδομένα και όμως μπορεί και προσεγγίζει με μεγαλύτερη ακρίβεια το σύστημα. Για τον περαιτέρω έλεγχο της ακρίβειας του μοντέλου πραγματοποιήθηκε ο υπολογισμός εννέα επιπλέον άγνωστων μόνιμων καταστάσεων, κάποιες εκ των οποίων βρίσκονται κοντά στα σημεία στροφής των κλάδων. Τα αποτελέσματα είναι πολύ ικανοποιητικά καθώς τα περισσότερα έχουν τιμή στο εύρος 2-5%, ενώ μεγαλύτερο ποσοστό εμφανίζουν κοντά στα σημεία στροφής των δύο κλάδων το οποίο είναι αναμενόμενο λόγω της ιδιαιτερότητας που εμφανίζει το σύστημα στις περιοχές αυτές.

Κατέστη λοιπόν εφικτή η κατασκευή ενός μοντέλου μειωμένης τάξης, που ενώ το ίδιο έχει κατασκευαστεί από δεδομένα μη μόνιμων καταστάσεων μπορεί να προβλέπει με ικανοποιητική ακρίβεια την συμπεριφορά του πλήρους μελετώμενου συστήματος. Η μόνη γνωστή πληροφορία που χρειάζεται είναι η αρχική κατάσταση του συστήματος καθώς και η νέα τιμή της μαζικής παροχής, ενώ δεν είναι απαραίτητη η γνώση των εξισώσεων που διέπουν το σύστημα. Ένα μοντέλο μειωμένης τάξης σαν αυτό, μειώνει ακόμη περισσότερο την ταχύτητα των υπολογισμών, την απαίτηση σε υπολογιστική μνήμη καθώς και το κόστος για την μελέτη μιας πολύπλοκης διεργασίας όπως είναι η Χημική Απόθεση από Ατμό. Αποτελεί συνεπώς ένα πολύ σημαντικό και ταυτόχρονα οικονομικό εργαλείο για την πρόβλεψη της συμπεριφοράς του συστήματος σε πραγματικό χρόνο.

6.2 Προτάσεις για περαιτέρω έρευνα

Η επιτυχία της μεθόδου που εφαρμόστηκε στην παρούσα εργασία, ανοίγει πολλαπλά πεδία μελέτης για το σύστημα χημικής απόθεσης από ατμό.

Αρχικά, το μοντέλο μειωμένης τάξης που κατασκευάστηκε σε κάθε μια από τις τέσσερις διερευνήσεις, περιγράφει και τους δύο ευσταθείς κλάδους του συστήματος. Κατά τον υπολογισμό των μόνιμων καταστάσεων με χρήση του μοντέλου μειωμένης τάξης, παρατηρήθηκε ότι οι τιμές των αποκλίσεων μεταξύ των δύο ευσταθών κλάδων διαφέρουν κατά πολύ. Το καλύτερο μοντέλο ήταν εν τέλη αυτό, για την κατασκευή του οποίου δεν χρησιμοποιήθηκε ο ίδιος αριθμός δεδομένων και για τους δύο κλάδους, αλλά ο αριθμός αυτός εξαρτιόνταν από μια τιμή σφάλματος την οποία έπρεπε να ξεπεράσει το σύστημα κατά την εξέλιξη του σε κάθε επιβολή. Συνεπώς, προτείνεται να εξετασθεί περαιτέρω η τιμή αυτή του σφάλματος, που παίζει καθοριστικό ρόλο για τον αριθμό των snapshots κάθε επιβολής που χρησιμοποιείται για την κατασκευή του μοντέλου.

Παρατηρήθηκε, επίσης, ότι η πρόβλεψη που επιτυγχάνεται από το νευρωνικό δίκτυο για την δυναμική απόκριση του συστήματος στην περιοχή των σημείων στροφής του κάθε κλάδου δεν απέχουν πολύ από τις προσεγγίσεις των υπολοίπων μόνιμων καταστάσεων. Με δεδομένη την δυσκολία υπολογισμού της δυναμικής συμπεριφοράς του συστήματος στα σημεία αυτά, τα αποτελέσματα των προσεγγίσεων είναι αρκετά ικανοποιητικά. Για την περαιτέρω όμως αύξηση της ακρίβειας των προσεγγίσεων του μοντέλου μειωμένης τάξης, είναι χρήσιμο να εξετασθούν τρόποι με τους οποίους μπορεί να επιτευχθεί ακόμα μεγαλύτερη βελτίωση της πρόβλεψης στις περιοχές αυτές. Μια πιθανή λύση θα ήταν να επιλεγθούν περισσότερες αρχικές καταστάσεις κοντά στα σημεία στροφής, ώστε να αυξηθεί ο αριθμός των δεδομένων που προέρχονται από τις προβληματικές αυτές περιοχές και που χρησιμοποιούνται για την κατασκευή της βάσης και την εκπαίδευση του νευρωνικού δικτύου.

Εν συνεχεία, προτείνεται να επαναληφθεί η ίδια μελέτη για κάθε έναν ευσταθή κλάδο ξεχωριστά. Τα αποτελέσματα από την κατασκευή ενός μοντέλου μειωμένης τάξης για την περιγραφή του πλήρους συστήματος από δεδομένα μη-μόνιμων καταστάσεων ήταν πολύ ικανοποιητικά, συνεπώς τίθεται το ερώτημα πόσο καλύτερη πρόβλεψη μπορεί να πραγματοποιήσει ένα τέτοιου είδους μοντέλο μειωμένης τάξης που εστιάζει μόνο στον έναν από τους δύο κλάδους. Τα δεδομένα που λαμβάνονται για την κατασκευή του, δηλαδή, να προέρχονται από επιβολές που πραγματοποιούνται μόνο στον κάτω ευσταθή κλάδο ή μόνο στον άνω.

Επιπλέον, η ιδέα της παρούσας εργασίας, βασίστηκε στην αναπαράσταση της δυναμικής του συστήματος με την κατασκευή μιας ορθοκανονικής βάσης και την εκπαίδευση ενός νευρωνικού δικτύου, με χρήση δεδομένων τα οποία λαμβάνονται από ένα σύνολο σημείων που βρίσκονται κοντά σε τυχαία επιλεγμένες αρχικές καταστάσεις. Λόγω του μικρού χρονικού διαστήματος κατά το οποίο αφήνεται μια αρχική κατάσταση να εξελιχθεί στην εργασία αυτή δεν χρησιμοποιήθηκαν δεδομένα εκπαίδευσης από επιβολές στις οποίες πραγματοποιείται εναλλαγή των κλάδων. Προτείνεται λοιπόν η περαιτέρω μελέτη της ιδέας αυτής με τον συνδυασμό δεδομένων από αρχικές καταστάσεις που αφήνονται να εξελιχθούν για ένα σύντομο χρονικό διάστημα και δεδομένων από επιβολές στις οποίες πραγματοποιείται αλλαγή κλάδου. Σε καμία από τις δύο περιπτώσεις το σύστημα δεν αφήνεται να φτάσει σε μόνιμη κατάσταση. Στην περίπτωση των επιβολών που πραγματοποιείται αλλαγή κλάδου για την πιο ολοκληρωμένη προσέγγιση του συστήματος οι επιβολές να αποτελούνται από μεγαλύτερο αριθμό στιγμιότυπων ώστε να αποκαλυφθεί περισσότερη πληροφορία κατά την τροχιά αυτών προς την τελική κατάσταση.

Στο υπό μελέτη σύστημα η μόνη παράμετρος που μεταβάλλεται με τον χρόνο είναι η μαζική παροχή του αντιδραστήρα. Το γεγονός αυτό μπορεί να απλοποιεί κατά πολύ την μελέτη του συστήματος, απέχει όμως αρκετά από την πραγματικότητα. Η κατασκευή ενός πιο ρεαλιστικού μοντέλου μπορεί να επιτευχθεί αν μελετηθεί το ίδιο σύστημα στο οποίο όμως θα μεταβάλλονταν με το χρόνο περισσότερες παράμετροι, όπως η θερμοκρασία εισόδου, η θερμοκρασία των τοιχωμάτων και του

υποστρώματος, και η πίεση λειτουργίας. Με αυτόν τον τρόπο καθίσταται εφικτή η μελέτη της επίδρασης και αυτών των παραγόντων στα πεδία ροής και θερμοκρασίας του αντιδραστήρα, όπως και στην ποιότητα του παραγόμενου υμενίου.

Ακόμα, ένας τρόπος που προτείνεται για την δημιουργία ενός πιο ρεαλιστικού και ολοκληρωμένου μοντέλου είναι τα δεδομένα για την κατασκευή της βάσης και την εκπαίδευση του νευρωνικού δικτύου, να περιλαμβάνουν και τις χημικές αντιδράσεις που μπορεί να πραγματοποιούνται στην αέρια φάση και στην στερεά επιφάνεια. Με τον τρόπο, οι ποιοτικές προδιαγραφές της διεργασίας, όπως είναι ο ρυθμός απόθεσης και η ποιότητα του παραγόμενου υμενίου, λαμβάνονται υπόψη κατά την προσέγγιση του πλήρους συστήματος από το μοντέλο μειωμένης τάξης.

Τέλος, η ιδέα στην οποία βασίστηκε η παρούσα εργασία προέρχεται από την μελέτη που πραγματοποίησαν οι Graham et al. ως προς τις εναλλακτικές προσεγγίσεις της μεθόδου ανάλυσης κυρίων συνιστωσών. Στην μελέτη τους ανέπτυξαν δύο πειράματα, ως προς τα εναλλακτικά σύνολα και τους εναλλακτικούς μέσους όρους των συνόλων δεδομένων που χρησιμοποιούνται για την κατασκευή μιας ορθοκανονικής βάσης. Η ιδέα της εργασίας προήλθε από το πρώτο εκ των δύο πειραμάτων, συνεπώς προτείνεται να πραγματοποιηθεί μια μελέτη που να εστιάζει στην ιδέα του δεύτερου πειράματος, δηλαδή τον υπολογισμό μέσων όρων συνόλων δεδομένων για την κατασκευή του μοντέλου μειωμένης τάξης και όχι απαραίτητα μέσων όρων χρονικών διαστημάτων. Στόχος αυτού είναι η πιο σφαιρική και ολοκληρωμένη πληροφόρηση για το πλήρες μελετώμενο σύστημα.

7. 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 Engineering* **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] M. Graham, I. Kevrekidis, "Alternative approaches to the Karhuen-Loeve decomposition for model reduction and data analysis", *Department of Chemical Engineering, Princeton University*, March 1995
- [4] A. Fluent, Theory Guide, Ansys Inc 5, 2009
- [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, SpringerVerlag 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 Research, 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] Λ. Θεοδόση-Κοκκίνου, "Τεχνητά Νευρωνικά Δίκτυα και εφαρμογές στα Συστήματα Αυτόματου Ελέγχου", Διπλωματική Εργασία, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών της πολυτεχνικής Σχολής του Πανεπιστημίου Πατρών, Πάτρα, 2013
- [12] F. L. Lewis, S. Jagannathan and A. Yesildirek, Neural network control of robot manipulators and nonlinear systems, London: Taylor & Francis, 1999
- [13] M. T. Hagan, H. B. Demuth, M. H. Beale, Neural Network Design, Boston, MA: PWS Publishing, 1996
- [14] I. B. Levitan, L. K. Kaczmarek, "The Neuron. Cell and Molecular Biology", 4th edition, Oxford University Press, 2015
- [15] A. K. Jain, J. Mao, K. M. Mohiuddin, "Artificial Neural Networks: A Tutorial." *Computer* **29**(3), 31-44 , (1996)
- [16] H. T. Siegelmann, B. G. Horne, C. Lee Giles, "Computational Capabilities of Recurrent NARX Neural Networks." *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics* **27**(2), 208-215 (1997)
- [17] P. A. Gkinis, E. D. Koronaki, A. Skouteris, I. G. Aviziotis & A. G. Boudouvis. "Building a data-driven Reduced Order Model of a Chemical Vapor Deposition

- process from low-fidelity CFD simulations." *Chemical Engineering Science* **199**, 371-380 (2019)
- [18] E. D. Koronaki, P. A. Gkinis, L. Beex, S. P. A. Bordas, C. Theodoropoulos, & A. G. Boudouvis. "Classification of states and model order reduction of large scale Chemical Vapor Deposition processes with solution multiplicity." *Computers & Chemical Engineering* **121**, 148-157 (2019)
- [19] W. Xie, I. Bonis, and C. Theodoropoulos, "Linear MPC based on data-driven Artificial Neural Networks for large-scale nonlinear distributed parameter systems," in Proceedings of the 22nd European Symposium on Computer Aided Process Engineering, London, (2012)
- [20] M. F. Barone, D. J. Segalman, H. Thornquist, and I. Kalashnikova, "Galerkin Reduced Order Models for Compressible Flow with Structural Interaction," in 46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2008
- [21] Ε. Άγγλου, "Ανάπτυξη ρεαλιστικών μοντέλων μειωμένης τάξης για μη γραμμικές διεργασίες χημικής απόθεσης από ατμό", Διπλωματική Εργασία, Σχολή Χημικών Μηχανικών ΕΜΠ, Αθήνα, 2017
- [22] Α. Σκουτέρης, "Αποδοτική μείωση τάξης 3-D μοντέλου Χημικής Απόθεσης από Ατμό με τη μέθοδο Κύριων Συνιστωσών και Τεχνητά Νευρωνικά Δίκτυα", Διπλωματική Εργασία, Σχολή Χημικών Μηχανικών ΕΜΠ, Αθήνα, 2018
- [23] T. Lin, B. G. Horne, P. Tiño, C. Lee Giles, "Learning Long-Term Dependencies in NARX Recurrent Neural Networks." *IEEE Transactions on Neural Networks* **7**(6), 1329-1338 (1996)