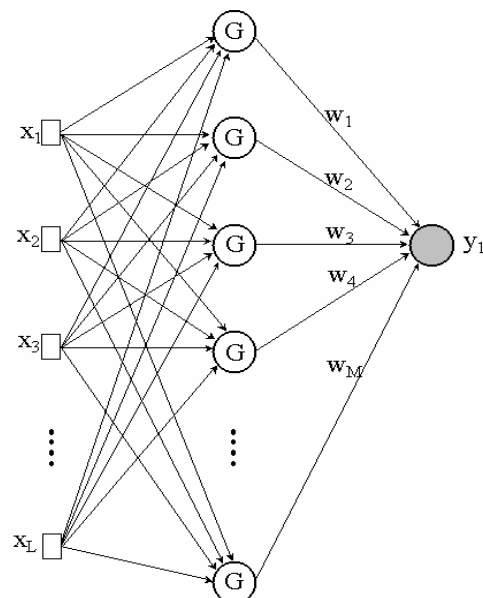




ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

## Διπλωματική Εργασία

Επιτάχυνση Διαφορικού Εξελικτικού Αλγορίθμου με τη χρήση  
Τεχνητών Νευρωνικών Δικτύων – Εφαρμογή στη  
βελτιστοποίηση αεροτομών



Ζωγράφος Ελευθέριος

Επιβλέπων Καθηγητής  
Νικολός Ιωάννης, Λέκτορας

Χανιά 2006

“Premature Optimization is the root of all evil”

D. Knuth

Στη μνήμη του Τάκη.

# ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της προσπάθειας που αποδελτιώνεται στην ανά χείρας διπλωματική εργασία, θα ήθελα να ευχαριστήσω ιδιαίτερα την οικογένειά μου για την ποικιλότροπη στήριξη που μου παρείχε όλα αυτά τα χρόνια. Επίσης ευχαριστώ θερμά το δάσκαλό μου και επιβλέποντα κ. Ιωάννη Νικολό για την καθοδήγηση. Μέσω των εμπειριών που αποκόμισα κατά τη συνεργασία μας, μπόρεσα να δω ορισμένα πράγματα με άλλο μάτι.

Θα ήθελα να ευχαριστήσω τέλος τα μέλη της εξεταστικής επιτροπής για την τιμή που μου έκαναν, να παραστούν στην κορυφαία ίσως στιγμή της φοιτητικής μου σταδιοδρομίας.

Σας ευχαριστώ όλους θερμά.

# ΠΕΡΙΕΧΟΜΕΝΑ

<b>ΠΕΡΙΕΧΟΜΕΝΑ</b>	<b>1</b>
<b>ΕΙΣΑΓΩΓΗ</b>	<b>3</b>
<b>ΣΥΝΤΜΗΣΕΙΣ</b>	<b>5</b>
<b>1</b>	<b>6</b>
<b>ΤΟ ΠΡΟΒΛΗΜΑ</b>	<b>6</b>
1.1. ΜΕΘΟΔΟΛΟΓΙΑ ΜΗΧΑΝΟΛΟΓΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ	7
1.1.1. Η ΔΙΑΔΙΚΑΣΙΑ	7
1.1.2. ΟΙ ΣΥΓΧΡΟΝΕΣ ΑΠΑΙΤΗΣΕΙΣ	8
1.1.3. ΣΥΝΟΨΗ ΤΩΝ ΜΕΘΟΔΩΝ ΣΧΕΔΙΑΣΜΟΥ	9
1.2. ΔΟΜΗ ΠΡΟΒΛΗΜΑΤΟΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ	10
1.2.1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ	10
1.2.2. Η ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ	11
1.3. Ο ΟΡΙΣΜΟΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ	12
<b>2</b>	<b>14</b>
<b>Η ΜΕΘΟΔΟΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ</b>	<b>14</b>
2.1. ΟΙ ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ	15
2.1.1. ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ ΚΑΙ ΒΑΣΙΚΟΙ ΤΥΠΟΙ ΕΑ	15
2.1.2. ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΚΑΙ ΕΞΕΛΙΚΤΙΚΑ ΠΡΟΓΡΑΜΜΑΤΑ	16
2.1.3. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΚΑΙ ΟΡΟΙ	16
2.2. ΟΙ ΔΙΑΦΟΡΙΚΟΙ ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ	18
2.3.1. Η ΜΕΘΟΔΟΛΟΓΙΑ ΤΟΥ DE	18
2.3.2. ΕΠΙΣΚΟΠΗΣΗ ΕΝΑΛΛΑΚΤΙΚΩΝ ΣΧΗΜΑΤΩΝ DE	18
2.3.3. ΣΥΓΚΡΙΣΗ DE ΚΑΙ ΑΝΤΑΓΩΝΙΣΤΙΚΩΝ ΜΕΘΟΔΩΝ	21
<b>3</b>	<b>24</b>
<b>ΤΟ ΠΡΟΣΕΓΓΙΣΤΙΚΟ ΜΟΝΤΕΛΟ</b>	<b>24</b>
3.1. ΠΡΟΣΕΓΓΙΣΤΙΚΑ ΜΟΝΤΕΛΑ	25
3.1.1. Η ΑΝΑΓΚΑΙΟΤΗΤΑ ΧΡΗΣΗΣ ΠΡΟΣΕΓΓΙΣΤΙΚΩΝ ΜΟΝΤΕΛΩΝ	25
3.1.2. ΕΠΙΣΚΟΠΗΣΗ ΤΕΧΝΙΚΩΝ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ ΤΟΠΙΟΥ ΚΑΤΑΛΛΗΛΟΤΗΤΑΣ	25
3.2. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ - ANN	29
3.2.1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΤΗΣ ΘΕΩΡΙΑΣ ΤΩΝ ANN	29
3.2.1.1. ΤΙ ΕΙΝΑΙ ΕΝΑ ANN	29
3.2.1.2. ΕΝΑΛΛΑΚΤΙΚΑ ΜΟΝΤΕΛΑ ΝΕΥΡΩΝΩΝ	30
3.2.1.3. ΕΝΑΛΛΑΚΤΙΚΕΣ ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΔΙΚΤΥΩΝ	33
3.2.1.4. Η ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΕΚΠΑΙΔΕΥΣΗΣ	34
3.2.1.5. ΕΦΑΡΜΟΓΕΣ ΤΩΝ ANN	36
3.2.2. ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ	36
3.2.3. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΣΥΝΑΡΤΗΣΕΩΣ ΑΚΤΙΝΙΚΗΣ ΒΑΣΗΣ – RBFN	38
3.2.4. ΔΙΚΤΥΑ ΠΟΛΛΑΠΛΩΝ ΣΤΟΙΒΑΔΩΝ ΠΡΟΣΘΙΑΣ ΤΡΟΦΟΔΟΤΗΣΗΣ - MLP	41
3.2.5. ΣΥΓΚΡΙΣΗ RBFN - MLP	43
3.3. ΣΥΝΔΥΑΣΜΟΣ RBFN ΚΑΙ DE	44
<b>4</b>	<b>47</b>
<b>Η ΕΦΑΡΜΟΓΗ</b>	<b>47</b>
4.1. Η ΦΥΣΗ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΜΕΘΟΔΟΥ	48
4.2. Η ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΥΛΟΠΟΙΗΣΗ ΤΗΣ ΜΕΘΟΔΟΥ	49

4.2.1. ΤΟ ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΟ ΛΟΓΙΣΜΙΚΟ .....	49
4.2.2. Η ΔΙΕΠΑΦΗ ΠΡΟΓΡΑΜΜΑΤΟΣ - ΧΡΗΣΤΗ .....	49
4.3. Η ΕΠΙΛΥΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ .....	51
4.3.1. Η ΑΕΡΟΤΟΜΗ ΚΑΙ Η ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΗΣ .....	51
4.3.2. ΟΡΙΣΜΟΣ ΚΑΙ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΠΡΟΒΛΗΜΑΤΟΣ.....	52
4.3.3. ΔΟΚΙΜΕΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ .....	53
<b>5.....</b>	<b>58</b>
<b>ΣΥΜΠΕΡΑΣΜΑΤΑ .....</b>	<b>58</b>
5.1. ΣΥΜΠΕΡΑΣΜΑΤΑ .....	59
5.2. ΠΡΟΤΑΣΕΙΣ.....	59
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ .....</b>	<b>61</b>
<b>ΠΑΡΑΡΤΗΜΑ.....</b>	<b>64</b>

# ΕΙΣΑΓΩΓΗ

Η παρούσα εργασία αναφέρεται στο πρόβλημα της βελτιστοποίησης αεροδυναμικών σχημάτων (*Aerodynamic Shape Optimization Problem - ASOP*), που αποτελεί ένα από τα σπουδαιότερα προβλήματα στο πεδίο της αεροναυτικής και της κατασκευής στροβιλοκινητήρων. Το αυξημένο ενδιαφέρον για το αντικείμενο, τροφοδοτείται στις μέρες μας από την σημαντική πρόοδο τόσο στην τεχνολογία των υπολογιστών όσο και από την ανάπτυξη ικανών μεθόδων αριθμητικής επίλυσης των σχετικών προβλημάτων. Η επάρκεια λοιπόν προγραμμάτων CAD, FEM και CFD μεταξύ άλλων, οδήγησε λογικά στην ανάγκη ανάπτυξης μεθοδολογιών, που θα βοηθούν τους μηχανικούς να φωτίσουν τη συμπεριφορά των υπαρχόντων συστημάτων και να εκμεταλλευτούν τη σχετική πληροφορία για την ανάπτυξη νέων, με βελτιωμένη απόδοση. Η διαχείριση αυτής ακριβώς της πληροφορίας είναι που αποτελεί το κριτήριο του διαχωρισμού των “μη έξυπνων” μεθόδων όπως η *τυχαία αναζήτηση* και η *εξαντλητική απαρίθμηση*, από τις *μεθόδους αριθμητικής βελτιστοποίησης*, που έχουν πλέον κυριαρχήσει στις μοντέρνες μηχανολογικές εφαρμογές.

Η ιδιότητα του βέλτιστου ορίζεται σε σχέση με τις εκάστοτε απαιτήσεις, οι οποίες αποδίδονται με τη μορφή *συναρτήσεων κόστους (Cost Functions)* ή *καταλληλότητας (Fitness Functions)*. *Μεθοδολογίες βελτιστοποίησης ενός ή πολλαπλών στόχων (Single / Multi-objective)* καθώς και με ή χωρίς *περιορισμούς (Constrained / Unconstrained)* απαιτούνται κατά περίπτωση, με αποτέλεσμα το τελικό μέγεθος του προβλήματος, όπως θα γίνει αντιληπτό στις επόμενες σελίδες, να το κατατάσσει στην κατηγορία εκείνη των προβλημάτων, όπου η εύρεση της βέλτιστης λύσης είναι πρακτικά αδύνατη με κλασσικό μαθηματικό προγραμματισμό. Ο χώρος των υποψήφιων λύσεων είναι τεράστιος και είναι επιτακτική η ανάγκη χρησιμοποίησης ευρετικού αλγορίθμου. Οι *Εξελικτικοί Αλγόριθμοι (Evolutionary Algorithms - EA)* αποτελούν ένα ισχυρό εργαλείο για την επίλυση προβλημάτων με τεράστιο χώρο λύσεων και επιλέχθηκαν στην παρούσα εργασία για την επίλυση του παραπάνω προβλήματος.

Η επιλογή των EA αντί κάποιας άλλης ευρετικής μεθόδου επίλυσης βασίστηκε στα ακόλουθα χαρακτηριστικά τους: οι EA εμφανίζουν αξιοσημείωτη ισορροπία μεταξύ της ικανότητάς τους να εξερευνούν τον χώρο λύσεων και της δυνατότητάς τους να πιέζουν προς την εύρεση καλύτερης λύσης. Συνδυάζουν χαρακτηριστικά κατευθυνόμενης και στοχαστικής αναζήτησης και εμφανίζουν μεγάλη ευστάθεια σε σχέση με άλλες αντίστοιχες μεθόδους. Προσαρμόζονται πολύ εύκολα σε διαφορετικού τύπου προβλήματα, ενώ είναι σχετικά εύκολη η εισαγωγή πολύπλοκων περιορισμών.

Ωστόσο η βελτιστοποίηση με χρήση EA είναι ιδιαίτερα χρονοβόρα διαδικασία στην πράξη. Η ανάγκη για μεγάλο αριθμό εκτιμήσεων μέσω ενός απαιτητικού σε υπολογιστικό χρόνο επιλύτη (π.χ. CFD), καθιστά τη μέθοδο ακατάλληλη για καθημερινή χρήση. Προκειμένου λοιπόν να μειώσουμε τον αριθμό των απαιτούμενων εκτιμήσεων, που είναι κυρίως υπεύθυνες για την εκτόξευση του υπολογιστικού κόστους της μεθόδου (μιας και η αντίστοιχη υπολογιστική πολυπλοκότητα του EA αυτού καθαυτού είναι συγκριτικά αμελητέα), καταφεύγουμε στη λύση της υιοθέτησης ενός *προσεγγιστικού μοντέλου (Approximation Model)*. Η βασική ιδέα πίσω από αυτή την επιλογή είναι η αντικατάσταση των υπολογιστικά ακριβών αναλυτικών επιλύσεων με γρήγορες μη ακριβείς προσεγγίσεις τους. Στην παρούσα εργασία, τη διαδικασία αυτή αναλαμβάνει ένα *Τεχνητό Νευρωνικό Δίκτυο Συναρτήσεως Ακτινικής Βάσης (Radial Basis Function Network – RBFN)*, που εκπαιδεύεται κατάλληλα για το σκοπό αυτό.

Η συνεισφορά της παρούσας εργασίας στη λύση του προβλήματος της βελτιστοποίησης αεροδυναμικών σχημάτων έγκειται στο συνδυασμό ενός νέου πολλά υποσχόμενου τύπου EA του *Διαφορικού Εξελικτικού Αλγορίθμου (Differential Evolution – DE)* με τον συγκεκριμένο τύπο νευρωνικού δικτύου (RBFN). Στην παρούσα εργασία προτείνεται μια μεθοδολογία συνδυασμού του Διαφορικού Εξελικτικού αλγορίθμου με τον

συγκεκριμένο τύπο Τεχνητού Νευρωνικού δικτύου. Η προτεινόμενη μεθοδολογία φάνηκε, μέσα από διαφορετικές εφαρμογές, ότι επιταχύνει τη σύγκλιση του ΕΑ, μειώνοντας το υπολογιστικό κόστος της μεθόδου. Η μέθοδος που αναπτύχθηκε εφαρμόστηκε στη συνέχεια στη βελτιστοποίηση αεροτομής (για συγκεκριμένες συνθήκες ροής), με κριτήρια την αύξηση της άνωσης και τη μείωση της αντίστασης. Το δικριτήριο πρόβλημα βελτιστοποίησης μοντελοποιήθηκε ως μονοκριτήριο, με την δημιουργία μοναδικής συνάρτησης κόστους, στην οποία τα δύο κριτήρια συμμετέχουν σταθμισμένα. Για να υπολογιστεί το ζητούμενο μέτωπο Pareto, εφαρμόστηκαν διαφορετικοί συντελεστές στάθμισης σε διαφορετικά τρεξίματα του αλγορίθμου. Παράλληλα, χρησιμοποιήθηκε ένας έτοιμος πολυκριτήριος Εξελικτικός Αλγόριθμος για την αντιμετώπιση του ίδιου προβλήματος και τα αποτελέσματά του συγκρίνονται με αυτά του μονοκριτηρίου.

Η διάρθρωση των κεφαλαίων της εργασίας είναι η ακόλουθη:

- Στο 1<sup>ο</sup> Κεφάλαιο γίνεται παρουσίαση του *προβλήματος της βελτιστοποίησης αεροδυναμικών σχημάτων (Aerodynamic Shape Optimization Problem - ASOP)*. Ξεκινάμε με μια ανάλυση της διαδικασίας του μηχανολογικού σχεδιασμού και της θέσης που κατέχει σε αυτόν η βελτιστοποίηση, αναφερόμαστε στις βασικές έννοιες της διαδικασίας και αναλύουμε τον τρόπο με τον οποίο δομείται ένα πρόβλημα βελτιστοποίησης. Ακολούθως γίνεται παρουσίαση του επιλυόμενου προβλήματος.
- Το Κεφάλαιο 2 εξετάζει τη χρησιμοποιούμενη μέθοδο βελτιστοποίησης, δηλαδή ο *Εξελικτικός Αλγόριθμος*. Αρχικά παρουσιάζονται οι διάφορες κατηγορίες ΕΑ, αναπτύσσεται η σχετική ορολογία και παρατίθενται στοιχεία για τη δομή τους. Ακολούθως εστιάζουμε στους *Διαφορικούς Εξελικτικούς Αλγορίθμους*, τόσο στην αρχική μορφή με την οποία εισήχθησαν όσο και σε αυτή με την οποία εδώ χρησιμοποιούνται και τέλος τους συγκρίνουμε με τους GA, προκειμένου να τεκμηριώσουμε την ανωτερότητά τους στην αντιμετώπιση του ανά χείρας προβλήματος βελτιστοποίησης.
- Το 3<sup>ο</sup> Κεφάλαιο αποτελεί την παρουσίαση του χρησιμοποιούμενου *προσεγγιστικού μοντέλου (Surrogate Model)*. Ξεκινάμε με την διερεύνηση της αναγκαιότητας χρήσης ενός τέτοιου μοντέλου και την επισκόπηση διαφόρων εναλλακτικών μεθόδων. Εν συνεχεία, αναφερόμαστε στις βασικές έννοιες της θεωρίας των ANN και επικεντρώνουμε στα RBFN και τα MLP. Το κεφάλαιο ολοκληρώνεται με την περιγραφή της διαδικασίας ενσωμάτωσης του RBFN στον DE.
- Το Κεφάλαιο 4 περιγράφει την τελική εφαρμογή για την επίλυση του προβλήματος βελτιστοποίησης αεροτομών. Αρχικά αποσαφηνίζεται η φύση της προτεινόμενης μεθόδου, βάσει της θεωρίας που έχει αναπτυχθεί στα προηγούμενα κεφάλαια. Εν συνεχεία δίνονται τα βασικά χαρακτηριστικά της προγραμματιστικής υλοποίησης (γλώσσα προγραμματισμού, διεπαφή με το χρήστη, κ.λπ.). Τέλος παρουσιάζονται όλα τα πειραματικά δεδομένα που προέκυψαν κατά την διάρκεια των δοκιμών επίλυσης του προβλήματος.
- Τέλος, στο κεφάλαιο των Συμπερασμάτων αποδελτιώνονται τα συμπεράσματα που προέκυψαν κατά την εκπόνηση της εργασίας, και επιπλέον προτείνονται μελλοντικές ερευνητικές κατευθύνσεις, που αξίζει κατά τη γνώμη μας να διερευνηθούν.
- Το Παράρτημα στο τέλος της εργασίας, φιλοξενούνται τα δεδομένα των δοκιμών, που πραγματοποιήθηκαν με σκοπό την εξαγωγή συμπερασμάτων για την αποτελεσματικότητα διαφόρων εναλλακτικών του DE αλγόριθμου μεθόδων βελτιστοποίησης, χρησιμοποιώντας μία σειρά από γνωστές μαθηματικές συναρτήσεις, καθώς και τα συμπεράσματα που εξήχθησαν.

# ΣΥΝΤΜΗΣΕΙΣ

<b>AI</b>	Artificial Intelligence	Τεχνητή Νοημοσύνη
<b>ANM</b>	<i>Annealed Nelder &amp; Mead Strategy</i>	
<b>ANN</b>	Artificial Neural Network	Τεχνητό Νευρωνικό Δίκτυο
<b>ASA</b>	Adaptive Simulated Annealing	Προσαρμοστική Προσομοιωμένη Ανόπτηση
<b>ASOP</b>	Aerodynamic Shape Optimization Problem	Πρόβλημα Βελτιστοποίησης Αεροδυναμικών Σχημάτων
<b>BEP</b>	Back Error Propagation algorithm	Αλγόριθμος Οπισθοτροφοδότησης του Σφάλματος
<b>CAD</b>	Computer Aided Design	Σχεδίαση με χρήση Η/Υ
<b>CFD</b>	Computational Fluid Dynamics	Υπολογιστική Ρευστοδυναμική
<b>CPU</b>	Central Processing Unit	Κεντρική Επεξεργαστική Μονάδα
<b>DACE</b>	Design and Analysis of Computer Experiments	Θεωρία Σχεδιασμού και Ανάλυσης Πειραμάτων σε Υπολογιστές
<b>DE</b>	Differential Evolution	Διαφορικός Εξελικτικός Αλγόριθμος
<b>DGA</b>	Distributed Genetic Algorithms	Διανεμημένοι Γενετικοί Αλγόριθμοι
<b>EA</b>	Evolutionary Algorithm	Εξελικτικός Αλγόριθμος
<b>EP</b>	Evolution Programs	Εξελικτικά Προγράμματα
<b>ES</b>	Evolution Strategies	Εξελικτικές Στρατηγικές
<b>FEM</b>	Finite Element Method	Μέθοδος Πεπερασμένων Στοιχείων
<b>FLM</b>	Fitness Landscape Modeling	Μοντελοποίηση του Τοπίου Καταλληλότητας
<b>FMS</b>	Finite State Machines	Μηχανές Πεπερασμένης Κατάστασης
<b>GA</b>	Genetic Algorithm	Γενετικός Αλγόριθμος
<b>GP</b>	Genetic Programming	Γενετικός Προγραμματισμός
<b>HEPD</b>	Hierarchical Electronic Product Definition	Ιεραρχικός Ηλεκτρονικός Καθορισμός Προϊόντος
<b>HGA</b>	Hierarchical Genetic Algorithms	Ιεραρχικοί Γενετικοί Αλγόριθμοι
<b>MLP</b>	Multi-Layer Perceptron	Δίκτυα Πολλαπλών Στοιβάδων
<b>MOOP</b>	Multi-Objective Optimization Problem	Πρόσθιας Τροφοδότησης Πρόβλημα Βελτιστοποίησης Πολλαπλών Στόχων
<b>POF</b>	Pareto Optimal Front	Βέλτιστο κατά Pareto Μέτωπο
<b>RBF</b>	Radial Basis Function	Συνάρτηση Ακτινικής Βάσης
<b>RBFN</b>	Radial Basis Function Network	Δίκτυο Συναρτήσεως Ακτινικής Βάσης
<b>RSM</b>	Response Surface Methodology	Μεθοδολογία της Επιφάνειας Απόκρισης
<b>SA</b>	Simulated Annealing	Προσομοιωμένη Ανόπτηση
<b>UAV</b>	Unmanned Aerial Vehicle	Μη-Επανδρωμένο Εναέριο Όχημα

# 1

## ΤΟ ΠΡΟΒΛΗΜΑ

ΣΤΟ πρώτο αυτό κεφάλαιο γίνεται παρουσίαση του *προβλήματος της βελτιστοποίησης αεροδυναμικών σχημάτων (Aerodynamic Shape Optimization Problem - ASOP)*. Ξεκινάμε με μια ανάλυση της διαδικασίας του μηχανολογικού σχεδιασμού και της θέσης που κατέχει σε αυτόν η βελτιστοποίηση. Στη συνέχεια αναφερόμαστε διεξοδικά στις βασικές έννοιες της διαδικασίας και αναλύουμε τον τρόπο με τον οποίο δομείται ένα πρόβλημα βελτιστοποίησης. Ακολούθως γίνεται παρουσίαση του επιλυόμενου προβλήματος.

## 1.1. ΜΕΘΟΔΟΛΟΓΙΑ ΜΗΧΑΝΟΛΟΓΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ

Ο *μηχανολογικός σχεδιασμός (Engineering Design)* είναι ιστορικά τόσο παλιός όσο οι πρώτες μηχανές και τα προϊόντα τους. Εξετάζοντας το πλέον πρόσφατο παρελθόν του, μπορεί κανείς να διαπιστώσει ότι η κυρίαρχη τάση αναφορικά με τη μεθοδολογία του τα τελευταία 20 χρόνια, είναι η μοντελοποίηση συγκεκριμένων πτυχών ενός συστήματος. Οδηγηθήκαμε έτσι, χάρη και στην πρόοδο των υπολογιστών, στην ανάπτυξη και χρήση εργαλείων προσομοίωσης, όπως τα λογισμικά CFD και FEM. Παρόλα αυτά ο σχεδιασμός πολύπλοκων μηχανολογικών συστημάτων στηρίζεται κυρίως στη γνώση και τη διαίσθηση του σχεδιαστή. Όσο δε η πολυπλοκότητα αυξάνει, τόσο δυσκολότερο είναι να έχει κανείς συνολική εικόνα των ανταλλαγμάτων (trade-offs) και των συμβιβασμών που απαιτούνται μεταξύ των διαφόρων πτυχών της σχεδίασης.

Επικεντρώνοντας στις αεροναυτικές εφαρμογές (Aeronautical Engineering), οφείλουμε να κάνουμε ορισμένες βασικές διαπιστώσεις για τη φύση τους. Πρόκειται για ένα από τα πεδία μηχανολογικής σχεδίασης με την υψηλότερη πολυπλοκότητα, μιας και ένα αεροσκάφος λόγω χάρη αποτελείται από μερικές χιλιάδες εξαρτήματα, με περισσότερες του ενός εκατομμυρίου παραμέτρους σχεδίασης αντίστοιχα [ ]. Παράλληλα εισάγεται μεγάλος αριθμός φυσικών και τεχνικών περιορισμών καθιστώντας το πρόβλημα ακόμα δυσκολότερο (multi-constraint). Ένα άλλο σημαντικό χαρακτηριστικό είναι το γεγονός ότι στη σχεδίαση υπεισέρχονται πολλές διαφορετικές επιστημονικές περιοχές, οι οποίες ταυτόχρονα καθορίζουν την εφικτότητα και ποιότητα της κάθε υποψήφιας λύσης.

### 1.1.1. Η ΔΙΑΔΙΚΑΣΙΑ

Προκειμένου να αντιμετωπιστεί το φαινόμενο της πολυπλοκότητας, που δρα απαγορευτικά στην άμεση εφαρμογή κάποιας μεθόδου βελτιστοποίησης, καταφεύγουμε συνήθως στην ιεραρχική προσέγγιση του μηχανολογικού συστήματος [Giles, 1997]. Πιο συγκεκριμένα ο *ιεραρχικός ηλεκτρονικός καθορισμός προϊόντος (Hierarchical Electronic Product Definition – HEPD)*, το πρώτο βήμα της διαδικασίας, έγκειται στη δημιουργία μιας βάσεως δεδομένων, στην οποία κάθε εξάρτημα περιγράφεται σε διάφορα επίπεδα λεπτομέρειας. Εν συνεχεία τα εργαλεία ανάλυσης αλληλεπιδρούν με το καταλληλότερο επίπεδο του HEPD. Το κάθε επίπεδο ορίζεται αναφορικά με τα υπόλοιπα, ούτως ώστε οι αλλαγές σε αυτό να κληροδοτούνται στα επόμενα επίπεδα.

Η ιεραρχική προσέγγιση ωστόσο δεν σταματά στην αναπαράσταση των επί μέρους τμημάτων αλλά επεκτείνεται στη διαχείριση του υψηλού υπολογιστικού κόστους ολόκληρου του συστήματος. Αυτό επιτυγχάνεται μέσω της διεξαγωγής του σχεδιασμού σε δύο επίπεδα: προηγείται ο προκαταρκτικός σχεδιασμός ολόκληρου του προϊόντος και ακολουθεί ο λεπτομερής σχεδιασμός των επιμέρους τμημάτων.

Στο πρώτο επίπεδο το αεροσκάφος λόγω χάρη, θεωρείται ως ένα σύστημα και λαμβάνονται υπόψη οι απαιτήσεις των καταναλωτών. Λαμβάνονται αποφάσεις για τις βασικές διαστάσεις και προδιαγραφές καθώς και τη μεταφορά τεχνογνωσίας και υποσυστημάτων από προηγούμενες αντίστοιχες μελέτες. Το επίπεδο μοντελοποίησης είναι ιδιαίτερα προσεγγιστικό, καθιστώντας την προσομοίωση γρήγορη, ενώ γίνεται χρήση

των χαμηλότερων επιπέδων της βάσης HEPD. Τελικά αποκομίζουμε μια συνολική ιδέα των συμβιβασμών που πρέπει να γίνουν και καταλήγουμε στις γενικές τεχνικές προδιαγραφές.

Στο επόμενο επίπεδο διάφορες ομάδες σχεδιαστών εργάζονται πάνω στα επιμέρους εξαρτήματα. Εδώ το ζητούμενο είναι η εκπλήρωση των στόχων που έχουν τεθεί στο προηγούμενο στάδιο, υπό τους αντίστοιχους πάντοτε περιορισμούς. Σε πολλές περιπτώσεις πρόκειται για βελτιστοποίηση γεωμετρικών σχημάτων, δεδομένου ότι οι μη γεωμετρικές παράμετροι του προβλήματος έχουν καθοριστεί στην προκαταρκτική σχεδίαση. Χρησιμοποιούνται πλέον εργαλεία υψηλής ακρίβειας, τα οποία ενεργούν στα υψηλότερα επίπεδα της βάσης HEPD. Το αποτέλεσμα της διαδικασίας είναι το εξάρτημα στην τελική του μορφή.

Η παραπάνω διαδικασία είναι σειριακή και εκτός εξαιρετικών περιπτώσεων οι αποφάσεις που παίρνονται στο προκαταρκτικό στάδιο δεν αλλάζουν στην πορεία. Το φαινόμενο αυτό δικαιολογείται από το γεγονός ότι ο αρχικός σχεδιασμός βασίζεται στην προϋπάρχουσα εμπειρία, οπότε είναι δύσκολο να προκύψουν σημαντικές εκπλήξεις. Διαγραμματικά η διαδικασία περιγράφεται στο Σχήμα 1.1:



**Σχήμα 1.1.** Η τρέχουσα σειριακή διαδικασία σχεδιασμού, δύο σταδίων [Giles, 1997].

### 1.1.2. ΟΙ ΣΥΓΧΡΟΝΕΣ ΑΠΑΙΤΗΣΕΙΣ

Επειδή ωστόσο ο σχεδιασμός είναι συνυφασμένος με την βιομηχανική παραγωγή, θα πρέπει να λαμβάνει υπόψη τις ανάγκες της, με σημαντικότερες αυτές για αυξημένη απόδοση και εγγυημένη μηχανολογική αρτιότητα, με ταυτόχρονη μείωση του χρόνου που απαιτεί ένας κύκλος ανάπτυξης προϊόντος. Παράλληλα η πίεση του ανταγωνισμού οδηγεί στην δημιουργία διαρκώς πολυπλοκότερων συστημάτων. Πως μπορεί λοιπόν να ανταπεξέλθει η κοινότητα των μηχανικών στις προκλήσεις αυτές;

Η διαφαινόμενη απάντηση είναι διπλή: με τη βελτίωση των χρησιμοποιούμενων εργαλείων αλλά και με την καλύτερη μεθοδολογικά οργάνωση της διαδικασίας του σχεδιασμού. Έτσι ο Giles [Giles, 1997] προβλέπει ότι η σημαντικότερη εξέλιξη στο χώρο της μηχανολογικής ανάλυσης θα προέλθει τις επόμενες δύο δεκαετίες από την συγκέντρωση των μεμονωμένων εργαλείων, σε ολοκληρωμένα συστήματα σχεδιασμού. Ο Andersson με τη σειρά του διερευνά το ζήτημα της μεθοδολογικής οργάνωσης και προτείνει μεθόδους για την βελτιστοποίησή της [Andersson, 2001], ενώ ο Pierret εξετάζει μια διαφορετική προσέγγιση του μηχανολογικού σχεδιασμού, ενσωματώνοντας στις παραδοσιακές μεθόδους στοιχεία *Τεχνητής Νοημοσύνης (Artificial Intelligence – AI)* [Pierret, 1999].

Σε μια προσπάθεια να αποδελτιώσουμε τις σύγχρονες απαιτήσεις από μια μεθοδολογία σχεδιασμού, θα σημειώναμε τα ακόλουθα [Giles, 1997, Andersson, 2001, Pierret, 1999]:

- Θα πρέπει να διασφαλίζει την μηχανολογική αρτιότητα και να συμβάλει στην αυξημένη απόδοση του νέου σχεδίου, μέσω της βελτίωσης των επιμέρους εργαλείων ανάλυσης που ενσωματώνει.
- Οφείλει να μειώνει την διάρκεια του κύκλου ανάπτυξης του νέου προϊόντος, βελτιστοποιώντας τις εμπλεκόμενες διαδικασίες.
- Πρέπει να διασφαλίζει την αποτροπή δημιουργίας προβλημάτων και καθυστερήσεων από την αλόγιστη ανθρώπινη παρέμβαση.
- Ιδιαίτερα επιθυμητή είναι η ικανότητα ενσωμάτωσης των ανθρώπινων γνώσεων και της εμπειρίας των σχεδιαστών, μιας και το σύστημα στοχεύει στο να τους επικουρεί κι όχι στο να τους αντικαταστήσει.
- Προκειμένου να διασφαλίζει την απρόσκοπτη συνεργασία των επιμέρους εργαλείων ανάλυσης καθώς και την σύζευξή τους με αποτελεσματικές μεθόδους βελτιστοποίησης, οφείλει να ενσωματώνει την κατάλληλη τεχνολογία λογισμικού και τις απαιτούμενες διεπαφές (interfaces) για την επικοινωνία των υποσυστημάτων αυτών.

Στο σημείο αυτό και υπό το φως των νέων απαιτήσεων, όπως μόλις απαριθμήθηκαν, οφείλουμε να επισημάνουμε δύο ελαττώματα της παραδοσιακής διαδικασίας του ιεραρχικού σχεδιασμού. Το πρώτο έχει να κάνει με το γεγονός ότι η διαδικασία στηρίζεται στην προϋπάρχουσα εμπειρία, οπότε έμμεσα οδηγεί σε τετριμμένες λύσεις, μη αφήνοντας χώρο για καινοτομίες. Το δεύτερο έχει την ίδια πηγή, δηλαδή την προσκόλληση στην υπάρχουσα γνώση και εκδηλώνεται με την αδυναμία ανταπόκρισης σε νέες και διαφορετικές απαιτήσεις σχεδίασης και στις μεταβαλλόμενες τεχνολογικές συνθήκες.

Λύση στα παραπάνω θα μπορούσε να αποτελέσει μια περισσότερο αλληλεπιδραστική διαδικασία σχεδιασμού, με πολλούς κύκλους σαν αυτόν του Σχήματος 1.2. Έτσι η αποκτώμενη γνώση από τον σχεδιασμό των επιμέρους τμημάτων ανανεώνει σε κάθε επανάληψη την εμπειρική βάση σχεδιασμού ολόκληρου του συστήματος, ενώ μπορούν να εισαχθούν και νέες μέθοδοι βελτιστοποίησης τόσο σε *τοπικό* όσο και σε *ολικό επίπεδο* (*local / global optimization*).



**Σχήμα 1.2.** Η μελλοντική αλληλεπιδραστική διαδικασία σχεδιασμού, δύο σταδίων [Giles, 1997].

### 1.1.3. ΣΥΝΟΨΗ ΤΩΝ ΜΕΘΟΔΩΝ ΣΧΕΔΙΑΣΜΟΥ

Ως εδώ κινηθήκαμε εξετάζοντας τα βήματα της διαδικασίας σχεδιασμού και εντοπίζοντας στον τρόπο ενσωμάτωσης σε αυτήν των εργαλείων ανάλυσης. Ωστόσο η *ανάλυση* (*analysis*), η διαδικασία δηλαδή καθορισμού της απόκρισης του συστήματος στο

περιβάλλον του, αποτελεί μέρος μόνο της διαδικασίας σχεδίασης. Ποιες είναι λοιπόν οι διαφορετικές μεθοδολογικές προσεγγίσεις του σχεδιασμού; Ο Pierret εντοπίζει τις ακόλουθες εναλλακτικές [Pierret, 1999]:

- *Αριθμητικές Μέθοδοι Βελτιστοποίησης (Numerical Optimization Methods)*:  
Πρόκειται για συνδυασμό μεθόδων ανάλυσης και αλγορίθμων αριθμητικής βελτιστοποίησης. Η γεωμετρία παραμετροποιείται με τη βοήθεια γενικών συναρτήσεων (πολυωνυμικές, καμπύλες Bezier κ.λπ.) και οι όροι της αντικειμενικής συνάρτησης, είναι συναρτήσεις της αεροδυναμικής και μηχανικής απόδοσης.
- *Μέθοδοι Αντίστροφου Σχεδιασμού (Inverse Methods)*:  
Οι εν λόγω μέθοδοι χρησιμοποιούν προδιαγεγραμμένες διανομές πίεσης ή ταχύτητας για να αναπαράγουν την γεωμετρία που τις προκαλεί, μέσα από αριθμητικές διαδικασίες.
- *Ευρετικές Μέθοδοι (Heuristic Methods)*:  
Εδώ έχουμε να κάνουμε με μεθόδους όπως τα *Έμπειρα Συστήματα (Expert Systems)* και τα *Συστήματα Ταξινόμησης (Classifier Systems)*. Για να βελτιώσουν την απόδοση του σχεδίου, τα πρώτα βασίζονται σε κανόνες που προκύπτουν από τη γνώση των ειδικών του πεδίου και τα δεύτερα σε κανόνες που αρχικά καθορίζονται από το χρήστη και στη συνέχεια αναπροσαρμόζονται με την πρόοδο της διαδικασίας.

## 1.2. ΔΟΜΗ ΠΡΟΒΛΗΜΑΤΟΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

### 1.2.1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Αρχικά εισάγουμε την έννοια της *αξίας (Value)* ενός σχεδίου. Η αξία είναι εγγενής ιδιότητα του σχεδίου και εκφράζει το βαθμό εκπλήρωσης των προσδοκιών που έχουμε για το σχέδιο αυτό. Γεννάται λογικά το ερώτημα ποιού τις προσδοκίες καλείται να εκπληρώσει το σχέδιο και πώς μπορεί αυτό να καταστεί σαφές στον σχεδιαστή. Κατά τη διάρκεια της εργασίας του ο τελευταίος πρέπει να ισορροπήσει ανάμεσα σε αντικρουόμενα χαρακτηριστικά του σχεδίου τόσο ποσοτικά όσο και ποιοτικά. Εδώ έγκειται και η σπουδαιότητα μεθοδολογιών όπως η *βελτιστοποίηση πολλαπλών στόχων (multi-objective optimization)*, που καταφέρνουν όπως θα δείξουμε στη συνέχεια να κάνουν αυτούς τους συμβιβασμούς “ορατούς”.

Στη συνέχεια εξετάζουμε τους βασικότερους τύπους *μεταβλητών σχεδίασης (design variables)*, των παραμέτρων δηλαδή του προβλήματος που μπορεί να μεταβάλλει ο σχεδιαστής [Aroga, 1989]. Έτσι έχουμε:

- *Ανεξάρτητες μεταβλητές σχεδίασης (Independent Design Variables)*: είναι οι ποσότητες που ο σχεδιαστής μεταβάλλει άμεσα, όπως λόγου χάρη οι γεωμετρικές διαστάσεις. Συνήθως αποκαλούνται απλώς μεταβλητές σχεδίασης.
- *Εξαρτημένες μεταβλητές σχεδίασης (Dependent Design Variables)*: πρόκειται για ποσότητες που ο σχεδιαστής δεν μεταβάλλει άμεσα, παρά μόνο μέσω των ανεξάρτητων μεταβλητών σχεδίασης, π.χ. η κατανάλωση καυσίμου. Αποκαλούνται χαρακτηριστικά της σχεδίασης και από αυτά εξάγεται η τελική του αξία.
- *Μεταβλητές Κατάστασης (State Variables)*: “ενδιάμεσος” τύπος μεταβλητής, μιας και δεν μπορεί να του ανατεθεί απευθείας τιμή, αλλά δεν συμμετέχει και στην διαμόρφωση της αξίας της σχεδίασης.
- *Μεταβλητές Λειτουργίας (Operating Variables)*: είναι μεταβλητές που μπορούν να τροποποιηθούν από το χρήστη από τη στιγμή που το σχέδιο έχει πραγματοποιηθεί.

- **Μεταβλητές Περιβάλλοντος (Environmental Variables):** οι παράμετροι που επηρεάζουν την απόδοση της σχεδίασης κατά τη χρήση της, όπως τα επιβαλλόμενα φορτία και οι θερμοκρασία.

Η επόμενη σημαντική έννοια της διαδικασίας, είναι η *αντικειμενική συνάρτηση (Objective Function)* ή *συνάρτηση κόστους (Cost Function)* ή *συνάρτηση καταλληλότητας (Fitness Function)*, όπως συναντάται στην βιβλιογραφία. Πρόκειται για το μηχανισμό έκφρασης των προτιμήσεων, από την αξιολόγηση του οποίου προκύπτει η αξία του σχεδίου. Λεπτομέρειες για την μορφοποίησή της δίνονται στην ενότητα 1.2.3.

Τέλος οι *περιορισμοί (Constraints)*, καθορίζουν τα όρια μέσα στα οποία μπορούν να κινηθούν οι τιμές των ανεξάρτητων μεταβλητών σχεδίασης. Η φυσική τους σημασία έγκειται στον προσδιορισμό των συνθηκών λειτουργίας του σχεδιαζόμενου συστήματος καθώς και των διαφόρων τεχνολογικών περιορισμών που αντιμετωπίζουμε κατά την κατασκευή του.

### 1.2.2. Η ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Θεωρούμε ένα σύστημα με ιδιότητες:

$$g_m, m = 0, P - 1 \quad (2.1)$$

που παίρνουν πραγματικές τιμές και αποτελούν το αντικείμενο της βελτιστοποίησης καθώς και τους περιορισμούς

$$g_m, m = P, P + C - 1 \quad (2.2)$$

οι οποίοι περιγράφουν ιδιότητες του συστήματος που δεν χρειάζεται να βελτιστοποιηθούν, αλλά απλώς να εκπληρωθούν προκαθορισμένες προδιαγραφές τους. Οι τιμές των παραπάνω ιδιοτήτων προσδιορίζονται από τις πραγματικές μεταβλητές σχεδίασης

$$x_j, j = 0, D - 1 \quad (2.3)$$

με

$$x_j \in [x_{jlow}, x_{jhigh}] \quad (2.4)$$

προκειμένου η λύση να είναι υλοποιήσιμη.

Η βελτιστοποίηση του συστήματος συνίσταται στην τροποποίηση του διανύσματος των ανεξάρτητων μεταβλητών σχεδίασης

$$\vec{x} = (x_0, x_1, \dots, x_{D-1})^T \quad (2.5)$$

έως ότου οι ιδιότητες  $g_m$  πάρουν τις βέλτιστες τιμές τους και ικανοποιηθούν οι περιορισμοί  $g_m, m \geq P$ . Ένα πρόβλημα βελτιστοποίησης μπορεί πάντα να επαναπροσδιοριστεί ως πρόβλημα ελαχιστοποίησης ως εξής:

$$\min F_m(\vec{x}) \quad (2.6)$$

όπου  $F_m(\vec{x})$  είναι η συνάρτηση που υπολογίζει την ιδιότητα  $g_m$  του συστήματος. Όλες αυτές οι συναρτήσεις μπορούν να συνδυαστούν σε μία συνάρτηση κόστους  $Z(\vec{x})$ , που συνήθως δίνεται από το σταθμισμένο άθροισμα:

$$Z(\vec{x}) = \sum_{m=0}^{P+C-1} w_m F_m(\vec{x}), w_m \geq 0 \quad (2.7)$$

Οι συντελεστές ή βάρη  $w_m$  υποδηλώνουν τη σπουδαιότητα που αποδίδεται στον κάθε στόχο και περιορισμό, ενώ κανονικοποιούν επίσης διαφορετικές φυσικές μονάδες. Το πρόβλημα λοιπόν της βελτιστοποίησης μπορεί πλέον να επαναδιατυπωθεί ως εξής:

$$\min Z(\vec{x}) \quad (2.8)$$

### 1.3. Ο ΟΡΙΣΜΟΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Το ASOP ιδωμένο ως πρόβλημα βελτιστοποίησης, φέρει τα βασικά χαρακτηριστικά των προβλημάτων σχεδιασμού αεροναυτικών εφαρμογών: υψηλή πολυπλοκότητα, μεγάλο αριθμό φυσικών και τεχνικών περιορισμών και μεγάλο αριθμό ανεξάρτητων μεταβλητών σχεδίασης. Ειδικότερα το πρόβλημα που αντιμετωπίζεται στην παρούσα εργασία, δεν είναι άλλο από το ευθύ πρόβλημα βελτιστοποίησης μιας αεροτομής, εκτεθειμένης σε χαμηλού αριθμού Reynolds ροή. Ο στόχος της όλης διαδικασίας είναι η εύρεση του διανύσματος  $x^*$  των μεταβλητών σχεδίασης που αντιπροσωπεύει αεροτομή, η οποία επιτυγχάνει βέλτιστη επίδοση κατά το προκαθορισμένο κριτήριο απόδοσης<sup>1</sup>. Στην περίπτωση μας στόχος είναι η μεγιστοποίηση της άνωσης και η ελαχιστοποίηση της οπισθέλκουσας της αεροτομής για δεδομένες συνθήκες ροής. Το πολυκριτήριο (δίκριτήριο) πρόβλημα βελτιστοποίησης, όπως αναλύεται στο κεφάλαιο 4, μοντελοποιείται ως μονοκριτήριο πρόβλημα βελτιστοποίησης με εφαρμογή διαφορετικών συνδυασμών βαρών μεταξύ των δύο κριτηρίων, έτσι ώστε να προσεγγιστεί το μέτωπο Pareto του πολυκριτηρίου προβλήματος.

Για την επίλυση του προβλήματος βελτιστοποίησης χρησιμοποιήθηκε Διαφορικός Εξελικτικός αλγόριθμος (DE). Οι ΕΑ και ειδικότερα η μέθοδος DE προτιμήθηκαν για την αντιμετώπιση του συγκεκριμένου προβλήματος μιας και παρουσιάζουν συγκεκριμένες επιθυμητές ιδιότητες όπως οι ακόλουθες [Price, Storn & Lampinen]:

- Δεν απαιτείται για την εφαρμογή τους η γνώση των παραγώγων της συνάρτησης που απαιτούνται για τους αλγόριθμους Μαθηματικής Ανάλυσης (π.χ. Μέθοδοι Συζυγών Κατευθύνσεων, Newton ή Σχεδόν Newton<sup>2</sup>).
- Αποφεύγουν την *ρητή απαρίθμηση (explicit enumeration)* στην οποία προβαίνουν μέθοδοι όπως ο *Δυναμικός Προγραμματισμός (Dynamic Programming)*, διαδικασία απαγορευτικά μεγάλου υπολογιστικού κόστους για τα συνήθως υψηλής διάστασης αεροναυτικά προβλήματα.
- Αντιμετωπίζουν αποδοτικότερα το πρόβλημα της εξερεύνησης του χώρου των λύσεων σε σχέση με τις μεθόδους τυχαίας αναζήτησης όπως ο Τυχαίος περίπατος ή η Προσομοιωμένη Ανόπτηση [Χριστόπουλος ,2003].

Η αξιολόγηση των υποψήφιων λύσεων γίνεται με κατάλληλους κώδικες ανάλυσης και πιο συγκεκριμένα με χρήση εργαλείων Υπολογιστικής Ρευστοδυναμικής (CFD). Ωστόσο το μέγεθος του χώρου λύσεων, που οδηγεί σε υψηλό αριθμό αξιολογήσεων για την αποδοτική λύση του προβλήματος, σε συνδυασμό με το μεγάλο υπολογιστικό κόστος των εν λόγω αναλύσεων, καθιστούν στην πράξη την εφαρμογή του DE δύσκολη υπόθεση. Αποφασιστικής σημασίας λοιπόν για την χρονικά και υπολογιστικά αποδοτική εφαρμογή του DE, είναι η αντικατάσταση του κατά το δυνατόν μεγαλύτερου αριθμού ακριβών αναλύσεων με λιγότερο ακριβείς αλλά ταχύτερες προσεγγίσεις.

Προκειμένου να επιτευχθεί ο ως άνω στόχος υιοθετήθηκε η χρήση ενός προσεγγιστικού μοντέλου και ειδικότερα ενός Τεχνητού Νευρωνικού Δικτύου. Το δίκτυο εκπαιδεύεται με δεδομένα προερχόμενα από προηγούμενες ακριβείς αναλύσεις και χρησιμοποιείται για την πρόγνωση της καταλληλότητας κάθε υποψήφιας λύσης-αεροτομής. Αποκτούμε έτσι γνώση για το ποιες λύσεις αξίζει να αξιολογηθούν περαιτέρω (δηλαδή με χρήση CFD εργαλείων) και ποιες όχι, εξοικονομώντας κατ' επέκταση υπολογιστικό χρόνο.

Απαραίτητη προϋπόθεση για την αντιμετώπιση του προβλήματος είναι η γεωμετρική αναπαράσταση της αεροτομής. Για το σκοπό αυτό έγινε χρήση καμπυλών B-Spline μιας και η συγκεκριμένη μεθοδολογία επιτρέπει τη δημιουργία λεπτομερών γεωμετρικών σχημάτων με ταυτόχρονη ελαχιστοποίηση των απαιτούμενων παραμέτρων σχεδίασης καθώς και ελεγχόμενη ομαλότητα της καμπύλης [Piegl & Tiller, 1997], [Farin,

<sup>1</sup> Η αναλυτική μοντελοποίηση και λύση του προβλήματος, είναι το αντικείμενο του Κεφαλαίου 4.

<sup>2</sup> Ο ενδιαφερόμενος για τις συγκεκριμένες μεθόδους παραπέμπετε στους [Παπαγεωργίου, 1999] και [Κουϊκόγλου, 2002]

1988]. Κάθε αεροτομή προσεγγίζεται με μία καμπύλη B-Spline, τα σημεία ελέγχου της οποίας χρησιμοποιούνται ως ανεξάρτητες μεταβλητές σχεδίασης. Περισσότερα για το θέμα περιγράφονται στο κεφάλαιο 4.

# 2

## Η ΜΕΘΟΔΟΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

ΣΤΟ παρόν κεφάλαιο εξετάζεται διεξοδικά η χρησιμοποιούμενη μέθοδος βελτιστοποίησης, δηλαδή ο *Εξελικτικός Αλγόριθμος*. Αρχικά παρουσιάζονται οι διάφορες κατηγορίες ΕΑ, αναπτύσσεται η σχετική ορολογία και παρατίθενται στοιχεία για τη δομή τους. Ακολούθως εστιάζουμε στους *Διαφορικούς Εξελικτικούς Αλγορίθμους* τόσο στην αρχική μορφή με την οποία εισήχθησαν όσο και σε αυτή με την οποία εδώ χρησιμοποιούνται και τέλος τους συγκρίνουμε με τους GA, προκειμένου να τεκμηριώσουμε την ανωτερότητά τους στην αντιμετώπιση του ανά χείρας προβλήματος βελτιστοποίησης.

## 2.1. ΟΙ ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

### 2.1.1. ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ ΚΑΙ ΒΑΣΙΚΟΙ ΤΥΠΟΙ ΕΑ

Τις τελευταίες τέσσερις δεκαετίες αυξάνει διαρκώς το ενδιαφέρον των ερευνητών για μεθόδους βελτιστοποίησης που βασίζουν τη λειτουργία τους στη μίμηση των διαδικασιών της φυσικής εξέλιξης. Οι ΕΑ ξεκίνησαν στη δεκαετία του '60 από διαφορετικούς ερευνητές και σε διαφορετικές μορφές, οι οποίες όμως είχαν ως κοινό παρονομαστή τη χρησιμοποίηση ενός πληθυσμού πιθανών λύσεων, οι οποίες εξελισσόμενες με τη χρήση τεχνικών δανεισμένων από τη γενετική, κατέληγαν σε (σχεδόν) βέλτιστες τιμές. Οι ΕΑ δε χρησιμοποιήθηκαν από όλους τους ερευνητές μόνο ως εργαλεία βελτιστοποίησης, αλλά και ως μέθοδοι εξομοίωσης της φυσικής διαδικασίας της εξέλιξης και της συμπεριφοράς ζωντανών οργανισμών.

Το 1965 ο *Ingo Rechenberg* εισήγαγε τις *Εξελικτικές Στρατηγικές* (*Evolutionstrategie* στην αυθεντική του μορφή ή *Evolution Strategies - ES*), ως μεθόδους βελτιστοποίησης των παραμέτρων αεροδυναμικών σωμάτων κατά τη διάρκεια πειραμάτων σε αεροδυναμική σήραγγα. (Οι εξελικτικές στρατηγικές χρησιμοποιούν κωδικοποίηση πραγματικών αριθμών και μόνο τη μετάλλαξη ως τεχνική διαφοροποίησης των χρωμοσωμάτων, ενώ προσομοιάζουν με τεχνικές *hill-climbing*). Ο χώρος των εξελικτικών στρατηγικών παρέμεινε ένας ενεργός χώρος έρευνας και αναπτύχθηκε ανεξάρτητα από αυτόν των γενετικών αλγορίθμων, αν και πρόσφατα έχουν αρχίσει να αλληλεπιδρούν [Rechenberg, 1973].

Την ίδια χρονική περίοδο (1966), οι *Fogel, Owens και Walsh* ανέπτυξαν τον *Εξελικτικό Προγραμματισμό*, με σκοπό να διερευνήσουν τη δυνατότητα εξέλιξης τεχνητής νοημοσύνης, με την έννοια η μηχανή να μπορεί να προβλέπει αλλαγές στο περιβάλλον της και να αντιδρά κατάλληλα. Το περιβάλλον προσομοιάζονταν με μία σειρά συμβόλων, ενώ η μηχανή θα έπρεπε να δίνει σαν έξοδο ένα σύμβολο που να αντιπροσωπεύει την αντίδρασή της στο δυναμικό περιβάλλον. Οι εξελισσόμενοι πληθυσμοί είναι *Finite State Machines (FMS)*, η δομή των οποίων μπορεί να μεταβληθεί μέσα από την εξελικτική διαδικασία. Η ιδέα ήταν οι μηχανές να εξελίσσονται μαζί με το περιβάλλον τους, ώστε να προσαρμόζονται σε αυτό [Fogel, Owens & Walsh, 1966].

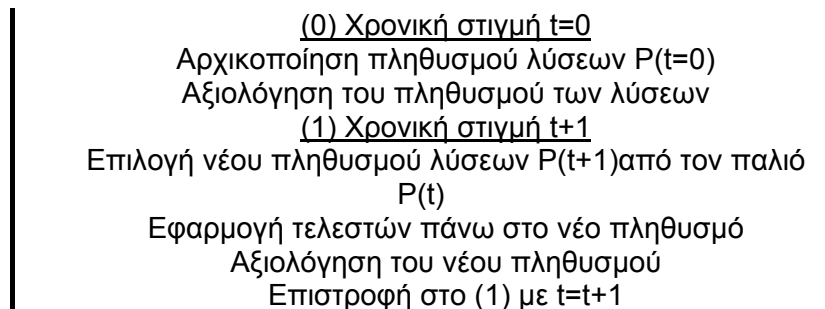
Ανάμεσα στο 1960 και το 1970 ο *John Holland* με την πρωτοποριακή δουλειά του εισήγαγε τους *Γενετικούς Αλγόριθμους* που εν συνεχεία αναπτύχθηκαν από τον ίδιο και τους φοιτητές του στο πανεπιστήμιο του Michigan. Σε αντίθεση με τις προαναφερθείσες τεχνικές, ο Holland είχε ως στόχο όχι το σχεδιασμό αλγορίθμων που να επιλύουν συγκεκριμένα προβλήματα, αλλά περισσότερο να εξετάσει κατά τρόπο γενικό το φαινόμενο της προσαρμογής, όπως αυτό παρατηρείται στη φύση και να αναπτύξει τρόπους, έτσι ώστε οι μηχανισμοί της φυσικής προσαρμογής να προσαρμοσθούν σε υπολογιστικά συστήματα. Ο ίδιος παρουσιάζει τους γενετικούς αλγορίθμους σαν μια αφηρημένη έννοια που πηγάζει από τη βιολογική εξέλιξη και εκθέτει το θεωρητικό περίγραμμα της προσαρμογής κάτω από τους γενετικούς αλγορίθμους [Holland, 1992].

Ο Γενετικός Προγραμματισμός (*Genetic Programming - GP*) που αναπτύχθηκε πρόσφατα (1992) από τον Koza, προτείνει ότι είναι καλύτερα να γίνεται εξερεύνηση των πιθανών διαφορετικών προγραμμάτων υπολογιστή που μπορούν να λύσουν ένα δεδομένο πρόβλημα. Έτσι, διαφορετικά προγράμματα ανταγωνίζονται μεταξύ τους για τη λύση του προβλήματος, με τα πιο ασθενή να πεθαίνουν και τα πιο αποτελεσματικά να εξελίσσονται [Koza, 1992].

### 2.1.2. ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΚΑΙ ΕΞΕΛΙΚΤΙΚΑ ΠΡΟΓΡΑΜΜΑΤΑ

Εκτός από τις παραπάνω γενικές κατηγορίες ΕΑ, υπάρχει μία πληθώρα τροποποιήσεων και υποκατηγοριών, ανάλογα με το πρόβλημα που αντιμετωπίζεται. Η ανοικτή μορφή των εξελικτικών αλγορίθμων επιτρέπει την ελεύθερη τροποποίηση των επί μέρους λειτουργιών τους, κάτι που περιορίζεται μόνο από τη φαντασία του εκάστοτε ερευνητή. Έτσι προκειμένου να περιγράψει το σύνολο των βασιζόμενων στην εξέλιξη συστημάτων, ο Michalewicz εισάγει τον όρο *Εξελικτικά Προγράμματα (Evolution Programs - EP)*, που περιλαμβάνει όλα τα προαναφερθέντα συστήματα.

Η βασική δομή ενός EP είναι αυτή του Σχήματος 2.1 [Michalewicz, 1999].



**Σχήμα 2.1.** Η δομή ενός Εξελικτικού Προγράμματος.

Πρέπει να τονιστεί εδώ η θεμελιώδης διαφοροποίηση μεταξύ ΕΑ και EP: ουσιαστικά ένα EP αποτελεί υπερσύνολο του ΕΑ, συνδυάζοντας τον τελευταίο με τις κατάλληλες δομές δεδομένων για την επίλυση ενός συγκεκριμένου προβλήματος. Πρόκειται δηλαδή για το επόμενο βήμα στην αντιμετώπιση του προβλήματος, μια περισσότερο εντοπισμένη στο πρόβλημα (*problem specific*) εφαρμογή. Έμμεσα στο σημείο αυτό εισάγεται κι ένα από τα σημαντικότερα ζητήματα της εφαρμογής της προτεινόμενης μεθόδου, ο καθορισμός δηλαδή των κατάλληλων δομών δεδομένων για την κωδικοποίηση του προβλήματος.

### 2.1.3. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΚΑΙ ΟΡΟΙ

Προκειμένου να εισαχθεί ο αναγνώστης στην μελέτη της δομής και της λειτουργίας των ΕΑ, απαιτείται ο ορισμός μερικών βασικών εννοιών και όρων που χρησιμοποιούνται στη σχετική ανάλυση. Αξίζει να σημειωθεί ότι σε μεγάλο βαθμό η χρησιμοποιούμενη ορολογία είναι δανεισμένη από το χώρο της βιολογίας. Για την καλύτερη λοιπόν εποπτεία της αναλογίας ανάμεσα στους όρους των δύο πεδίων, δημιουργήθηκε ο Πίνακας 2.1.

Όρος	Βιολογία	ΕΑ
<b>Χρωμόσωμα</b> ( <i>chromosome</i> )	Είναι μια αλληλουχία χαρακτήρων του DNA και αποτελεί το “σχέδιο δράσης” με το οποίο θα αναπτυχθεί ο οργανισμός.	Μία υποψήφια λύση: είναι ουσιαστικά η ακολουθία των τιμών των ανεξάρτητων μεταβλητών σχεδίασης του προβλήματος.
<b>Γονίδιο</b> ( <i>gene</i> )	Αποτελούν τους θεμέλιους λίθους του DNA, και καθένα από αυτά κωδικοποιεί μια συγκεκριμένη πρωτεΐνη.	Η κωδικοποίηση της κάθε ανεξάρτητης μεταβλητής σχεδίασης του προβλήματος.
<b>Αλληλόμορφα</b> ( <i>alleles</i> )	Αποτελούν τις διαφορετικές τιμές που μπορεί να πάρει κάθε γονίδιο.	Αποτελούν τις διαφορετικές τιμές που μπορεί να πάρει κάθε γονίδιο.
<b>Γονιδίωμα</b> ( <i>genome</i> )	Η συνολική συλλογή από γενετικό υλικό (όλα τα χρωμοσώματα μαζί) του οργανισμού.	Το σύνολο των γονιδίων ενός χρωμοσώματος (μιας υποψήφιας λύσης).
<b>Γονότυπος</b> ( <i>genotype</i> )	Το συγκεκριμένο σύνολο γονιδίων που περιέχονται σε ένα γονιδίωμα.	Οι συγκεκριμένες τιμές των ανεξάρτητων μεταβλητών σχεδίασης, όπως έχουν υλοποιηθεί με την εφαρμογή της συγκεκριμένης κωδικοποίησης.
<b>Φαινότυπος</b> ( <i>phenotype</i> )	Τα σωματικά και πνευματικά χαρακτηριστικά, όπως χρώμα ματιών, ύψος, μέγεθος εγκεφάλου, ευφυΐα κ.λπ.	Είναι η υλοποίηση της λύσης στο συγκεκριμένο πρόβλημα με βάση τις συγκεκριμένες τιμές των παραμέτρων του γονότυπου (π.χ. μια αεροτομή).
<b>Διπλοειδείς</b> ( <i>diploids</i> )	Οι οργανισμοί στους οποίους τα χρωματοσώματα παρατάσσονται σε ζευγάρια.	Δεν συνηθίζονται στους ΕΑ.
<b>Απλοειδείς</b> ( <i>haploids</i> )	Οι οργανισμοί στους οποίους τα χρωματοσώματα παρατάσσονται μόνα τους.	Οι υποψήφιες λύσεις που κωδικοποιούνται με ένα μόνο χρωμόσωμα.
<b>Διασταύρωση</b> ( <i>crossover</i> )	Σε κάθε γονέα ανταλλάσσονται γονίδια μεταξύ των ζευγαριών των χρωμοσωμάτων για να σχηματίσουν τον <i>γαμέτη</i> ( <i>gamete</i> ) -ένα απλό χρωμόσωμα- και έπειτα οι γαμέτες των δύο γονέων ζευγαρώνουν και δημιουργούν ένα πλήρες σύνολο χρωμοσωμάτων.	Συμβαίνει όταν ανταλλάσσεται ένα ολόκληρο τμήμα του χρωμοσώματος μεταξύ δύο πιθανών λύσεων, δίνοντας δύο νέες λύσεις.
<b>Μετάλλαξη</b> ( <i>mutation</i> )	Η διαδικασία κατά την οποία τα νουκλεοτίδια (πρωταρχικά στοιχεία του DNA) αλλάζουν από γονέα σε απόγονο εξαιτίας κάποιων λαθών που συνήθως λαμβάνουν χώρα.	Η διαδικασία κατά την οποία αλλάζει η τιμή ενός μόνο γονιδίου-παραμέτρου της λύσης, οπότε προκύπτει μια διαφοροποιημένη λύση.

**Πίνακας 2.1.** Αναλογία της ορολογίας Βιολογίας – Εξελικτικών Αλγορίθμων.

Για περισσότερες πληροφορίες πάνω στους διάφορους τύπους ΕΑ ο ενδιαφερόμενος αναγνώστης παραπέμπετε στη βιβλιογραφία [Goldberg, 1989], [Mitchell, 1996].

## 2.2. ΟΙ ΔΙΑΦΟΡΙΚΟΙ ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

### 2.3.1. Η ΜΕΘΟΔΟΛΟΓΙΑ ΤΟΥ DE

Στην περίπτωση του DE [Storn & Price, 1995], έχουμε να κάνουμε με έναν παράλληλο αλγόριθμο αναζήτησης για μη γραμμικά προβλήματα, ο οποίος χρησιμοποιεί σε κάθε γενιά  $NP$  διανύσματα παραμέτρων της μορφής:

$$\vec{x}_{j,G}, j = 0, NP - 1 \quad (2.1)$$

όπου το  $G$  αριθμεί τις γενιές. Ο αρχικός πληθυσμός δημιουργείται τυχαία αν δεν έχουμε κάποια πληροφορία για το σύστημα, υποθέτοντας ομοιόμορφη κατανομή πιθανότητας για οποιονδήποτε τυχαίο παράγοντα πρέπει να “γεννήσουμε”, εκτός αν ρητά δηλώνεται κάτι διαφορετικό. Αν πάλι διαθέτουμε κάποια προκαταρκτική λύση, τότε εκκινούμε προσθέτοντας ομαλά κατανεμημένες τυχαίες αποκλίσεις στην ονομαστική αυτή λύση  $\vec{x}_{nom,0}$ . Σε κάθε περίπτωση το πλήθος  $NP$  δεν μεταβάλλεται.

Η βασική ιδέα πίσω από τους DE είναι ο καθορισμός μιας διαδικασίας δημιουργίας δοκιμαστικών διανυσμάτων μεταβλητών π.χ.:  $\vec{x}_{trial,G}$ . Ο στόχος αυτός επιτυγχάνεται με την πρόσθεση του σταθμισμένου διανύσματος της διαφοράς δύο μελών του πληθυσμού, σε ένα τρίτο μέλος. Εάν το προκύπτον διάνυσμα έχει μικρότερη τιμή καταλληλότητας από το προκαθορισμένο μέλος του πληθυσμού, έστω  $\vec{x}_{old,G}$ , με το οποίο συγκρίνεται (υποθέτοντας πρόβλημα ελαχιστοποίησης), τότε αντικαθιστά το τελευταίο στο νέο πληθυσμό. Το  $\vec{x}_{old,G}$  μπορεί, αλλά δεν είναι υποχρεωτικό να είναι μέλος της διαδικασίας σχηματισμού του δοκιμαστικού διανύσματος. Για κάθε γενιά του αλγορίθμου, αξιολογείται επιπλέον η καλύτερη λύση  $\vec{x}_{best,G}$ , προκειμένου να παρακολουθούμε την πρόοδο της διαδικασίας βελτιστοποίησης.

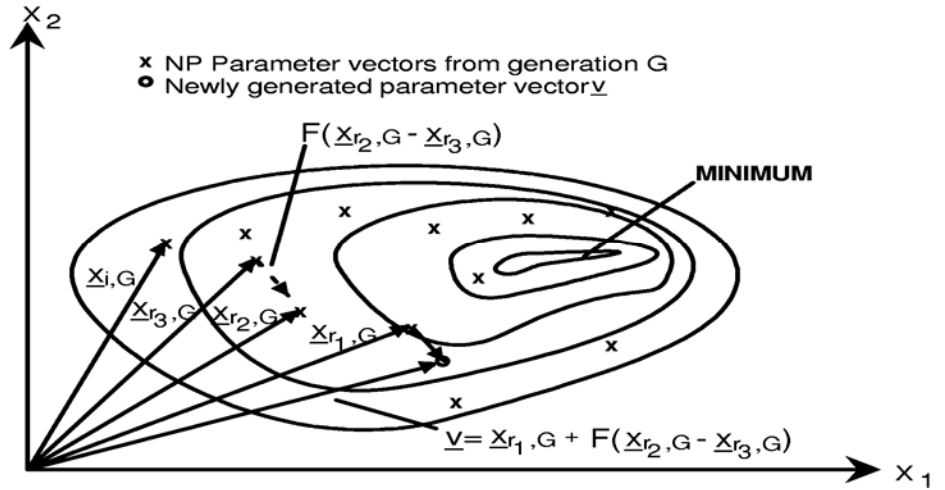
Η παραπάνω διαδικασία εξαγωγής πληροφορίας σχετικής με την απόσταση και την κατεύθυνση στο χώρο των λύσεων, με σκοπό την δημιουργία τυχαίων αποκλίσεων, οδηγεί στη δημιουργία ενός προσαρμοστικού σχήματος με πολύ καλές ιδιότητες σύγκλισης. Εν συνεχεία παρουσιάζεται μια επισκόπηση κάποιων βασικών εναλλακτικών σχημάτων.

### 2.3.2. ΕΠΙΣΚΟΠΗΣΗ ΕΝΑΛΛΑΚΤΙΚΩΝ ΣΧΗΜΑΤΩΝ DE

- Το πρώτο υπό εξέταση σχήμα οφείλεται στους Storn και Price και λειτουργεί ως εξής [Storn & Price, 1995]: όσον αφορά τη διαδικασία της μετάλλαξης, για κάθε διάνυσμα  $\vec{x}_{j,G}, j = 0, NP - 1$ , ένα δοκιμαστικό διάνυσμα  $\vec{v}_{j,G+1}$  δημιουργείται σύμφωνα με τον τύπο:

$$\vec{v}_{j,G+1} = \vec{x}_{r_1,G} + F(\vec{x}_{r_2,G} - \vec{x}_{r_3,G}), \quad (2.2)$$

με τα  $r_1, r_2, r_3 \in [0, NP - 1]$  διαφορετικούς μεταξύ τους τυχαία εκλεγμένους ακέραιους και  $F > 0$ , μια πραγματική σταθερά καθοριζόμενη από το χρήστη. Τα διάφορα διανύσματα που εμπλέκονται στην διαδικασία φαίνονται στο Σχήμα 2.2, όπου  $\vec{x} \equiv \underline{x}$ .

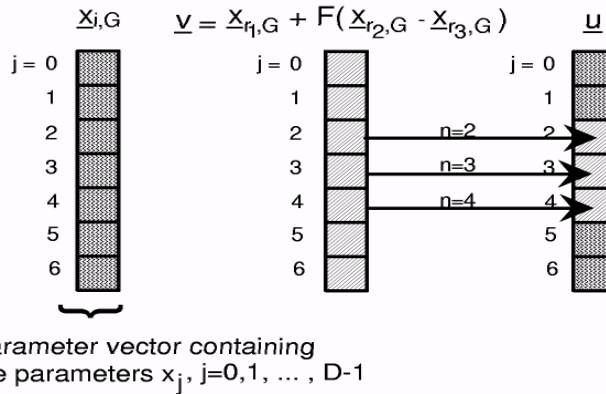


**Σχήμα 2.2.** Η διαδικασία γέννησης του  $\vec{v}_{j,G+1}$ , σε δυσδιάστατο παράδειγμα [Storn & Price, 1995].

Εν συνεχεία το *διαταραγμένο διάνυσμα* (*perturbed vector*) και το αντίστοιχο μέλος του τρέχοντος πληθυσμού υπόκεινται σε διασταύρωση, από την οποία προκύπτει ο πληθυσμός των δοκιμαστικών διανυσμάτων  $\vec{u}_{j,G+1} = (u_{1,j,G+1}, u_{2,j,G+1}, \dots, u_{n,j,G+1})$  ως εξής:

$$u_{i,j,G+1} = \begin{cases} v_{i,j,G+1}, & \text{if } rand_j \leq C_r \quad \forall j = k \\ x_{i,j,G}, & \text{άλλως} \end{cases} \quad (2.3)$$

Εδώ  $i = 1, \dots, n$ ,  $k \in \{1, \dots, n\}$  ο δείκτης της τυχαίας παραμέτρου που επιλέγεται για κάθε  $j$ . Ο συντελεστής διασταύρωσης  $C_r \in [0, 1]$ , είναι η δεύτερη παράμετρος ελέγχου του DE που εισάγεται από το χρήστη. Η διαδικασία απεικονίζεται στο Σχήμα 2.3.



**Σχήμα 2.3.** Η διαδικασία της διασταύρωσης [Storn & Price, 1995].

Τελικά η επιλογή των μελών του νέου πληθυσμού από τον τρέχοντα πληθυσμό και το αντίστοιχο δοκιμαστικό διάνυσμα γίνεται σύμφωνα με τον κανόνα:

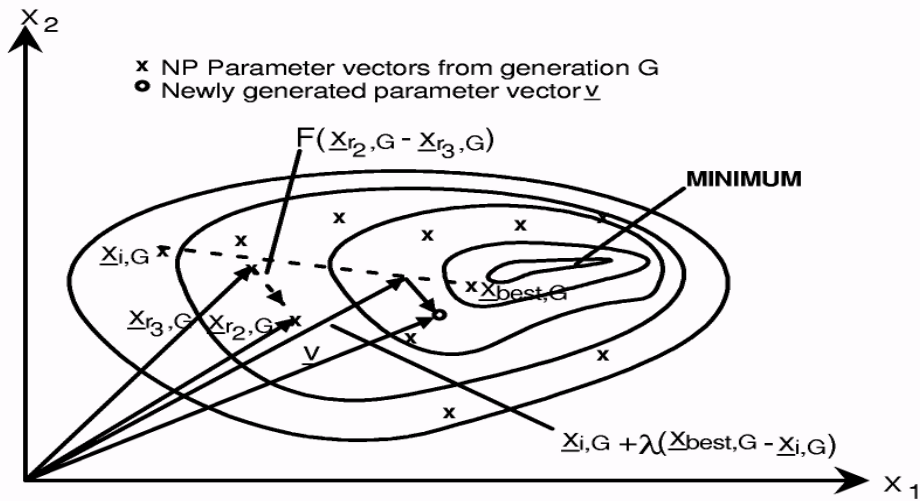
$$\vec{x}_{j,G+1} = \begin{cases} \vec{u}_{j,G+1}, & \text{if } Z(\vec{u}_{j,G+1}) \leq Z(\vec{x}_{j,G}) \\ \vec{x}_{j,G}, & \text{άλλως} \end{cases} \quad (2.4)$$

επιλύοντας πάντα πρόβλημα της μορφής  $\min Z(\vec{x})$ .

- Το δεύτερο εναλλακτικό σχήμα [Storn & Price, 1995] λειτουργεί όπως και το πρώτο με την διαφορά ότι το διάνυσμα  $\vec{v}_{j,G+1}$  δημιουργείται ως εξής:

$$\vec{v}_{j,G+1} = \vec{x}_{j,G} + \lambda (\vec{x}_{best,G} - \vec{x}_{j,G}) + F(\vec{x}_{r_2,G} - \vec{x}_{r_3,G}), \quad (2.5)$$

Εισάγεται έτσι μια νέα μεταβλητή ελέγχου  $\lambda$ , η βασική ιδέα πίσω από την οποία είναι, η ενίσχυση της *απληστίας* (*greediness*) του αλγορίθμου, ενσωματώνοντας την τρέχουσα καλύτερη λύση. Η διαφοροποίηση αυτή απεικονίζεται στο Σχήμα 2,4, ενώ η υπόλοιπη διαδικασία παραμένει απaráλλακτη.



**Σχήμα 2.4.** Η διαδικασία γέννησης του  $\vec{v}_{j,G+1}$ , σε δυοδιάστατο παράδειγμα [Storn & Price, 1995].

Οι τρεις επόμενες εναλλακτικές είναι αποτέλεσμα της εργασίας των *Fan, Lampinen* και *Dulikravnich* [Fan, Lampinen & Dulikravnich, 2003], πίσω από την οποία η βασική ιδέα είναι η επιλογή ενός *δότη* (*donor*) που να περιέχει την τοπική πληροφορία των μελών της τριπλέτας μετάλλαξης. Προσδοκούμε με την επιλογή αυτή να προσφέρουμε ένα καλύτερο σημείο εκκίνησης για τη μετάλλαξη, που με τη σειρά της θα αποδώσει καλύτερα κατανεμημένα δοκιμαστικά διανύσματα.

Τα νέα σχήματα μπορούν να περιγραφούν ως κυρτός συνδυασμός των μελών της τριπλέτας μετάλλαξης από τη γενική σχέση [Fan, Lampinen & Dulikravnich, 2003]

$$d = \sum_{i=1}^3 w_i \vec{x}_{r_i,G}, \quad (2.6)$$

όπου για τα βάρη ισχύει:  $w_i \geq 0$ , και  $\sum_{i=1}^3 w_i = 1$ . Μπορούμε πλέον να διακρίνουμε τα τρία νέα σχήματα ως εξής:

- Η περίπτωση του αριθμητικού μέσου:

$$d_1 = \frac{1}{3} \sum_{i=1}^3 \vec{x}_{r_i,G} \quad (2.7)$$

- Η περίπτωση του σταθμισμένου μέσου κατά κατανομή Gauss:

$$d_2 = \sum_{i=1}^3 \left( \lambda_i / \sum_{j=1}^3 \lambda_j \right) \bar{x}_{r_i,G}, \quad \lambda_j = randn_j, \quad (2.8)$$

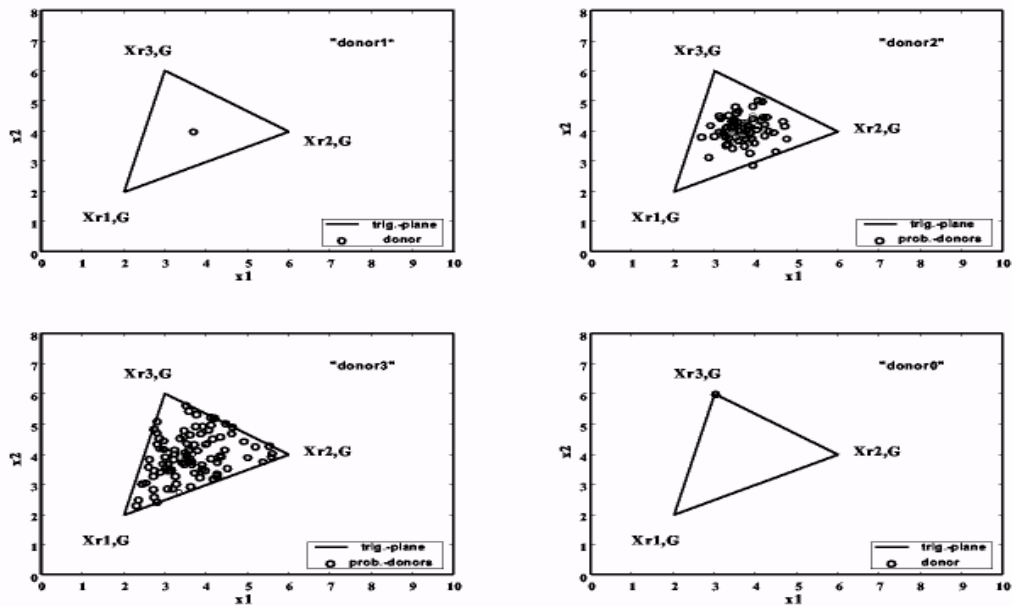
όπου η  $randn_j$  συμβολίζει τις τυχαίες τιμές που ικανοποιούν την κατανομή Gauss ως  $R(randn_j) = \exp\left[\left(-\left(randn_j - c\right)^2 / (2\sigma^2)\right)\right]$ , με  $c = 0.5$  και  $\sigma = 0.2$

- Η περίπτωση του σταθμισμένου μέσου κατά ομοιόμορφη κατανομή:

$$d_3 = \sum_{i=1}^3 \left( \lambda_i / \sum_{j=1}^3 \lambda_j \right) \bar{x}_{r_i,G}, \quad \lambda_j = rand_j[0,1], \quad (2.9)$$

όπου η  $rand_j$  συμβολίζει μια ομοιόμορφα κατανεμημένη τιμή μεταξύ  $[0.0, 1.0]$ .

Η περίπτωση ενός διδιάστατου προβλήματος φαίνεται στο Σχήμα 2.5. Μπορούμε να πούμε ότι στο χώρο των λύσεων η τριπλέτα των επιλεγμένων ατόμων σχηματίζει ένα τριγωνομετρικό υπερεπίπεδο, με κορυφές τα άτομα. Οι δότες που καθορίζονται από κάθε σχήμα θα βρίσκονται με υψηλή πιθανότητα εντός αυτού του υπερεπιπέδου. Από τη Σχέση 2.7, είναι προφανές ότι όσο μεγαλύτερο το βάρος ενός μέλους της τριπλέτας, τόσο κοντύτερά του προσελκύει το δότη.



**Σχήμα 2.5.** Γεωμετρική αναπαράσταση των σχημάτων επιλογής δότη για δυσδιάστατο πρόβλημα. Ως “donor0”, εμφανίζεται το πρώτο προταθέν σχήμα [Fan, Lampinen & Dulikravich, 2003].

### 2.3.3. ΣΥΓΚΡΙΣΗ DE ΚΑΙ ΑΝΤΑΓΩΝΙΣΤΙΚΩΝ ΜΕΘΟΔΩΝ

Ιδιαίτερο βέβαια πρακτικό ενδιαφέρον παρουσιάζει η διαφορά στην απόδοση των εναλλακτικών μεθόδων κατά την επίλυση ενός προβλήματος βελτιστοποίησης. Οι *Storn* και *Price* τεκμηριώνουν την ανωτερότητα του DE έναντι τόσο της *Annealed Nelder & Mead*

(ANM)<sup>3</sup> στρατηγικής, όσο και απέναντι στην Προσαρμοστική Προσομοιωμένη Ανόπτηση (Adaptive Simulated Annealing - ASA) [Storn & Price, 1995].

Σε ότι έχει να κάνει με τους GA, οι δύο μέθοδοι μοιράζονται μια σειρά από τα κοινά χαρακτηριστικά των ΕΑ. Εξερευνούν το χώρο των λύσεων εκτιμώντας έναν πληθυσμό υποψήφιων λύσεων ανά γενεά, για προκαθορισμένο αριθμό γενεών ή ως την εκπλήρωση κάποιου κριτηρίου σύγκλισης, βασιζόμενες στην κατάλληλη μοντελοποίηση του προβλήματος. Ως προς την υλοποίησή τους μπορούμε να εντοπίσουμε κοινά σημεία όπως η ύπαρξη τελεστή μετάλλαξης και τελεστή διασταύρωσης, κι η χρήση ενός σχήματος επιλογής, αλλά και σημαντικές διαφορές ως προς τον τρόπο που τα παραπάνω εφαρμόζονται.

Για την εξαγωγή ενός συμπεράσματος για την υπεροχή κάποιας εκ των δύο, προβήκαμε στην σύγκρισή τους με βάση μια σειρά γνωστών μαθηματικών συναρτήσεων. Ειδικότερα χρησιμοποιήθηκαν οι ακόλουθες τρεις συναρτήσεις:

- Η συνάρτηση του Ackley στην γενικευμένη εκδοχή της από τον Schwefel:

$$ff_1(\vec{x}) = -a \cdot \exp\left(-b \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(c \cdot x_i)\right) + a + e ; -32.768 \leq x_i \leq 32.768 \quad (2.10)$$

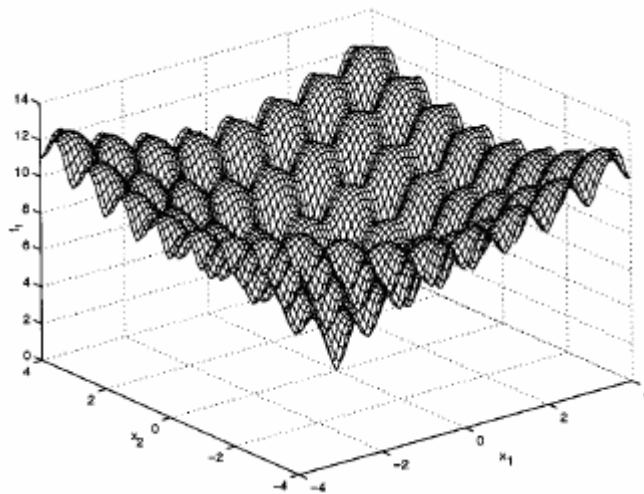
$$a=20 ; b=0.2 ; c=2\pi ; e=2.71828 ; \min(f_1)=f_1(0,\dots,0)=0$$

- Η γενικευμένη συνάρτηση της σφαίρας:

$$ff_0(x) = \sum_{i=1}^n x_i^2 \quad (2.11)$$

- Η συνάρτηση του Griewangk:

$$ff_4(x) = \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i+1}}\right) + 1 \quad (2.12)$$



Σχήμα 2.6. Η συνάρτηση του Ackley στις δύο διαστάσεις.

<sup>3</sup> Η μέθοδος καθώς και κώδικας υλοποίησής της βρίσκεται στην παράγραφο 10.4 του [Press et all, 1992]

Για κάθε μία από τις άνωθεν συναρτήσεις χρησιμοποιήθηκαν διάφοροι συνδυασμοί πλήθους μεταβλητών σχεδίασης και μεγέθους πληθυσμού, φροντίζοντας ωστόσο οι συγκρίσεις να γίνονται για ίδιο αριθμό αξιολογήσεων της μαθηματικής συνάρτησης. Τέλος ο DE αλγόριθμος συγκρίνεται τόσο με έναν κλασσικό GA όσο και με έναν *Πολλαπλών Υποπληθυσμών Γενετικό Αλγόριθμο (Multi-Deme GA)*. Τα δεδομένα των δοκιμών καθώς και ο σχολιασμός τους παρουσιάζονται στο Παράρτημα που βρίσκεται στις τελευταίες σελίδες και συνεισφέρουν στην τεκμηρίωση της ανωτερότητας του DE έναντι του κλασσικού GA. Στην παρούσα εργασία χρησιμοποιήθηκε το πρώτο σχήμα DE (σχέση 2.2) και το τρίτο σχήμα ( $d_3$ ) επιλογής δότη (σχέση 2.9).

Επιπλέον συγκριτικά στοιχεία για την απόδοση του DE έναντι άλλων γνωστών αλγορίθμων βελτιστοποίησης μπορεί να βρει κανείς στο τρίτο κεφάλαιο του [Price, Storn & Lampinen]. Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι ενώ ο DE δεν είναι πάντοτε ο γρηγορότερος αλγόριθμος, είναι από τους πλέον ευσταθείς. Η ικανότητά του να επηρεάζεται ελάχιστα από την αρχικοποίηση του πληθυσμού και η εύκολη ρύθμιση των λίγων μεταβλητών ελέγχου του, τον καθιστούν ιδανικό για την αντιμετώπιση καινοφανών προβλημάτων για τα οποία υπάρχει καθόλου ή περιορισμένη γνώση.

Τέλος μια σύγκριση των εναλλακτικών σχημάτων DE, παρατίθεται στο [Fan, Lampinen & Dulikravich, 2003]. Και σε αυτή την περίπτωση η σύγκριση γίνεται με βάση τις μαθηματικές συναρτήσεις του *De Jong* και άλλων. Το αποτέλεσμα είναι ενδεικτικό της υπεροχής των τριών προτεινόμενων από τους εν λόγω συγγραφείς σχημάτων και ιδιαίτερα των  $d_2$  και  $d_3$  (Παράγραφος 2.3.2), από τα οποία και επιλέξαμε τελικά για την εφαρμογή μας το σχήμα  $d_3$ , όπως προαναφέρθηκε.

# 3

## ΤΟ ΠΡΟΣΕΓΓΙΣΤΙΚΟ ΜΟΝΤΕΛΟ

Το τρίτο κεφάλαιο αποτελεί την παρουσίαση του χρησιμοποιούμενου προσεγγιστικού μοντέλου (*Surrogate Model*). Ξεκινάμε με την διερεύνηση της αναγκαιότητας χρήσης ενός τέτοιου μοντέλου και την επισκόπηση διαφόρων εναλλακτικών μεθόδων. Εν συνεχεία, αναφερόμαστε στις βασικές έννοιες της θεωρίας των ANN και παρουσιάζουμε διεξοδικά τις δύο εκδοχές τους που χρησιμοποιήθηκαν στα πλαίσια της ανά χείρας εργασίας: τα RBFN και τα MLP. Το κεφάλαιο ολοκληρώνεται με την περιγραφή της διαδικασίας ενσωμάτωσης του RBFN στον DE.

### 3.1. ΠΡΟΣΕΓΓΙΣΤΙΚΑ ΜΟΝΤΕΛΑ

#### 3.1.1. Η ΑΝΑΓΚΑΙΟΤΗΤΑ ΧΡΗΣΗΣ ΠΡΟΣΕΓΓΙΣΤΙΚΩΝ ΜΟΝΤΕΛΩΝ

Το σημαντικότερο πρόβλημα κατά την πρακτική εφαρμογή των ΕΑ είναι, όπως είδαμε και στο προηγούμενο κεφάλαιο, το υψηλό υπολογιστικό τους κόστος. Το τελευταίο οφείλεται στον μεγάλο αριθμό αξιολογήσεων των υποψηφίων λύσεων με υπολογιστικά “ακριβά” εργαλεία, γεγονός που καθιστά τους κλασσικούς ΕΑ ακατάλληλους για καθημερινές βιομηχανικές εφαρμογές. Γίνεται λοιπόν προφανής η ανάγκη εξεύρεσης τρόπων μείωσης του υπολογιστικού κόστους ή/και επιτάχυνσης της διαδικασίας περάτωσης μιας σχεδιομελέτης. Μερικοί τέτοιοι τρόποι είναι [Giannakoglou, 2002]:

- Ο επιμερισμός του φόρτου εργασίας μεταξύ επικοινωνούντων επεξεργαστών.
- Ο κατακερματισμός του συνολικού πληθυσμού, σε ημιαπομονωμένους υποπληθυσμούς (*demes*), όπως στην περίπτωση των *Διανεμημένων Γενετικών Αλγορίθμων* (*Distributed Genetic Algorithms – DGAs*)
- Η χρήση *ιεραρχικών μοντέλων* (*hierarchical models*) επικοινωνίας μεταξύ των υποπληθυσμών, οι οποίοι δομούνται δενδροειδώς. Τέτοια λύση αποτελούν οι *Ιεραρχικοί Γενετικοί Αλγόριθμοι* (*Hierarchical Genetic Algorithms – HGAs*).
- Η χρήση βελτιωμένων, πιο κατάλληλων εξελικτικών τελεστών και ο λιγότερο απαιτητικός υπολογιστικός χειρισμός των περιορισμών.
- Η υβριδοποίηση των ΕΑ με διάφορες παραδοσιακές τεχνικές που εκμεταλλεύονται την κλίση (*gradient based*).
- Η αντικατάσταση των ακριβών υπολογιστικά αναλυτικών εκτιμήσεων της αντικειμενικής συνάρτησης, με προσεγγιστικές αλλά ταχύτερες εκτιμήσεις. Η ουσία αυτής της μεθόδευσης έγκειται στην ανάπτυξη προσεγγιστικών μοντέλων, που αντικαθιστούν τα ακριβά εργαλεία επίλυσης σε μεγάλο αριθμό περιπτώσεων.

Στην επόμενη ενότητα παρουσιάζεται μια σειρά από προσεγγιστικές τεχνικές που εκμεταλλεύονται την πληροφορία που προέρχεται από τη *μοντελοποίηση του τοπίου καταλληλότητας* (*Fitness Landscape Modeling*). Αντίστοιχα η χρησιμοποιούμενη ορολογία, είναι προσαρμοσμένη στην οπτική του χώρου λύσεων, με τον όρο *σημείο* να ισοδυναμεί με πιθανή λύση.

#### 3.1.2. ΕΠΙΣΚΟΠΗΣΗ ΤΕΧΝΙΚΩΝ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ ΤΟΠΙΟΥ ΚΑΤΑΛΛΗΛΟΤΗΤΑΣ

Το βασικό πρόβλημα που καλούμαστε εδώ να αντιμετωπίσουμε μπορεί να διατυπωθεί ως εξής [Torczon & Trosset, 1998]:

“Με δεδομένο έναν περιορισμένο αριθμό υπολογιστικά ακριβών εκτιμήσεων της συνάρτησης καταλληλότητας για την επίλυση ενός μηχανολογικού προβλήματος βελτιστοποίησης, βρείτε τον καλύτερο

συμβιβασμό ανάμεσα στην υπολογιστική δαπάνη προσέγγισης και τη δαπάνη εκτίμησης.”

Η προτεινόμενη από τους άνωθεν συγγραφείς στρατηγική για την επίλυση του προβλήματος αυτού, έγκειται σε πρώτη φάση, στην εκτίμηση της πραγματικής αντικειμενικής συνάρτησης για έναν αριθμό σημείων. Εν συνεχεία κατασκευάζεται ένα προσεγγιστικό μοντέλο και μέσω αυτού γίνεται η εκτίμηση για ένα νέο σημείο. Από την σύγκριση με την τιμή της αναλυτικής εκτίμησης προκύπτουν τα δεδομένα για την ενημέρωση του προσεγγιστικού μοντέλου και η διαδικασία επαναλαμβάνεται. Μερικές από τις σημαντικότερες τεχνικές που συναντώνται στη βιβλιογραφία αναλύονται ακολούθως.

- Τα ANN, φαίνεται να βρίσκουν εφαρμογή σε σειρά προβλημάτων όπως ο σχεδιασμός αεροτομών, η επιβεβαίωση δεδομένων πτήσεων καθώς επίσης η πρόγνωση, διάγνωση και κατηγοριοποίηση βλαβών. Δεδομένου ότι στα ANN αφιερώνεται το υπόλοιπο του κεφαλαίου δεν θα επεκταθούμε εδώ περισσότερο.
- Η μεθοδολογία της *Επιφάνειας Απόκρισης (Response Surface Methodology – RSM)*:

Δεδομένου ενός διανύσματος  $x$  που επηρεάζει την απόκριση  $y$ , η σχέση μεταξύ τους είναι:

$$y = f(x) + e \quad (3.1)$$

όπου το  $e$  παριστά ένα τυχαίο σφάλμα κατανομημένο κατά  $N(0, s)$ . Η πραγματική συνάρτηση της επιφάνειας απόκρισης  $f(x)$  είναι συνήθως άγνωστη οπότε θεωρούμε μια προσέγγισή της έστω  $g(x)$ , οπότε οι προβλεπόμενες τιμές είναι:

$$\hat{y} = g(x) \quad (3.2)$$

Στις περισσότερες περιπτώσεις οι προσεγγιστικές συναρτήσεις είναι χαμηλής τάξης πολυώνυμα. Όταν η καμπυλότητα είναι ομαλή ένα πρώτης τάξης πολυώνυμο μπορεί να χρησιμοποιηθεί ενώ για εντονότερη καμπυλότητα ένα δεύτερης τάξης πολυώνυμο της μορφής:

$$\hat{y} = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^k \sum_{j=1, i < j}^k \beta_{ij} x_i x_j \quad (3.3)$$

Οι παράμετροι  $\beta$  καθορίζονται συνήθως μέσω παλινδρόμησης ελαχίστων τετραγώνων, προσαρμόζοντας την προσεγγιστική επιφάνεια απόκρισης σε υπάρχοντα δεδομένα ή μέσω gradient μεθόδων. Ο αναγνώστης παραπέμπεται στο [Shyy et al.] για μια επισκόπηση των εν λόγω τεχνικών.

- Η χρήση της μεθοδολογίας *Kriging*:

Σύμφωνα με την μέθοδο αυτή η άγνωστη συνάρτηση  $y(x)$  αποτελεί το άθροισμα δύο όρων: της μακροπρόθεσμα αναμενόμενης τιμής της πραγματικής συνάρτησης καταλληλότητας (drift function)  $f(x)$  και της βραχυπρόθεσμης επίδρασης κάθε σημείου στο μοντέλο (covariance function)  $Z(x)$ .

$$y(x) = f(x) + Z(x) \quad (3.4)$$

Ο πίνακας συμμεταβλητότητας της τελευταίας δίνεται από τη σχέση:

$$\text{Cov} [Z(x^i) Z(x^j)] = \sigma^2 R[R(x^i x^j)] \quad (3.5)$$

Όπου  $R$  είναι ο πίνακας συσχέτισεων και  $R(x^i x^j)$  είναι η συνάρτηση συσχέτισης ανάμεσα σε δυο εκλεγμένα σημεία  $x^i$  και  $x^j$ . Ο  $R$  είναι ένας τετραγωνικός συμμετρικός πίνακας με μονάδες κατά μήκος της διαγωνίου ενώ η  $R(x^i x^j)$  καθορίζεται από το χρήστη.

Οι προβλεπόμενες εκτιμήσεις  $\hat{y}$  της απόκρισης  $y(x)$  σε μη δοκιμασμένες τιμές του δίνονται από τη σχέση:

$$\hat{y} = \hat{\beta} + r^T(x) R^{-1} (y - f \hat{\beta}) \quad (3.6)$$

όπου  $y$  είναι το διάνυσμα που περιέχει τις τιμές της απόκρισης σε κάθε σημείο του δείγματος και  $f$  είναι ένα διάνυσμα γεμάτο με μονάδες όταν η  $f(x)$  λαμβάνεται σταθερή.  $r^T(x)$  είναι το διάνυσμα συσχέτισης ανάμεσα σε ένα μη δοκιμασμένο  $x$  και τα δειγματοποιημένα δεδομένα, ενώ δίνεται από την:

$$r^T(x) = [R(x, x^1) R(x, x^2) \dots R(x, x^{ns})]^T \quad (3.7)$$

Τέλος το  $\hat{\beta}$  δίνεται από την εξίσωση:

$$\hat{\beta} = (f^T R^{-1} f)^{-1} f^T R^{-1} y \quad (3.8)$$

Μολονότι απαιτούνται υπολογιστικά δαπανηρές αντιστροφές πινάκων για τον υπολογισμό του εξαγομένου, υπάρχει η δυνατότητα δημιουργίας τόσο ολικών (*global*) όσο και τοπικών (*local*) μοντέλων, ανάλογα με τη στρατηγική δειγματοληψίας [Giannakoglou, 2002]. Ο ενδιαφερόμενος παραπέμπετε επίσης στο [Martin & Simpson, 2003] για μια μελέτη διαφορετικών τύπων *Kriging*.

- **Διαδικασίες Gauss (Gaussian Processes - GPs):**

Οι *GPs* [Varnik, 1998] καθορίζουν ένα πιθανοθεωρητικό μοντέλο σε ένα σύνολο δεδομένων, ενώ μπορούν να προσεγγίσουν οποιαδήποτε συνάρτηση. Δουλεύουν ως εξής [Buche, Schraudolph & Koumoutsakos, 2005]:

- Αρχικά ένα πιθανοθεωρητικό μοντέλο κατασκευάζεται για ένα σύνολο δεδομένων.
- Το μοντέλο επεκτείνεται, προκειμένου να προβλέψει τη μέση τιμή και την τυπική απόκλιση της τιμής της συνάρτησης σε νέα σημεία εκτός του αρχικού συνόλου.
- Η αναμενόμενη τιμή της τιμής της συνάρτησης μεγιστοποιείται για όλα τα σημεία του συνόλου, μέσω της προσαρμογής των “υπερπαραμέτρων” (hyperparameters)

Η μαθηματική έκφραση των παραπάνω για  $N$  πλήθος δεδομένων σημείων  $X_N$  και τις αντίστοιχες τιμές  $t_N$  της συνάρτησης είναι η εξής:

$$\hat{t}_{N+1} = k^T C_N^{-1} t_N \quad (3.9)$$

$$\sigma_{t_{N+1}}^2 = \kappa - k^T C_N^{-1} k \quad (3.10)$$

$$L = -\frac{1}{2} (\log \det C_N + t_N^T C_N^{-1} t_N + N \log 2\pi) \quad (3.11)$$

$$\frac{\partial L}{\partial \theta} = \frac{1}{2} (t_N^T \Gamma_N C_N^{-1} t_N - \text{trace}(\Gamma_N)) \quad (3.12)$$

όπου,

$$\Gamma_N = C_N^{-1} \frac{\partial C_N}{\partial \theta} \quad (3.13)$$

και  $k$  είναι το διάνυσμα που περιέχει τις μεταβλητότητες ανάμεσα στα γνωστά και το καινούργιο σημείο, με  $k$  τη διακύμανση του νέου σημείου.  $C_N$  είναι ο πίνακας συμμεταβλητότητας της κατανομής Gauss. Οι εξισώσεις (3.9) και (3.10), προβλέπουν τη μέση τιμή και την τυπική απόκλιση ενός νέου σημείου, ενώ οι (3.11) και (3.12) βελτιστοποιούν τις υπερπαραμέτρους. Το σημαντικότερο μειονέκτημα των *GP*s είναι το υψηλό υπολογιστικό τους κόστος.

- *Support Vector Machines – SVM:*

Εμπνευσμένα από τη Στατιστική Θεωρία Εκμάθησης (Statistical Learning Theory) τα SVM συγκρινόμενα με άλλα αντίστοιχα μοντέλα όπως τα ANN, εμφανίζουν το πλεονέκτημα της απουσίας τοπικών ακροτάτων κατά την εκμάθηση. Επιπλέον το σφάλμα γενίκευσης δεν εξαρτάται από τη διάσταση του χώρου.

Για ένα σύνολο δειγμάτων  $(x_i, y_i)$ ,  $i=1, \dots, ns$ , η κατασκευή του μοντέλου είναι ισοδύναμη με τη βελτιστοποίηση της συνάρτησης:

$$L = \|w\|^2 + C \cdot \frac{1}{ns} \sum_{i=1}^{ns} \max\{|y_i - f(x_i)| - e\} \quad (3.14)$$

όπου,  $e$  είναι το ανεκτό σφάλμα,  $C$  μια σταθερά και  $f$

$$f(x) = wx + b \quad (3.15)$$

η υπό εκτίμηση συνάρτηση.

Ο Μη Γραμμικός Προγραμματισμός μπορεί να λύσει το παρακάτω ισοδύναμο πρόβλημα:

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \cdot \frac{1}{ns} \sum_{i=1}^{ns} (\lambda_i + \lambda_i^*) \\ \text{υπό συνθήκες:} \\ & ((w \cdot x_i) + b) - y_i \leq e + \lambda_i \\ & y_i - ((w \cdot x_i) + b) \leq e + \lambda_i^* \\ & \lambda_i^*, \lambda_i \geq 0, i = 1, \dots, ns \end{aligned} \quad (3.16)$$

- Συστήματα Επαγωγικής Εκμάθησης (inductive learning system)

Ένα Συστήματα Επαγωγικής Εκμάθησης, επάγει κανόνες τύπου Συνθήκης – Δράσης (Condition - Action rules), από δεδομένα υποδείγματα. Οι κανόνες αυτοί διαχωρίζουν τα δεδομένα σε διακριτές κατηγορίες και συνήθως συνδυάζονται σε μορφή δένδρων απόφασης για καλύτερη εμποπτεία. Η μέθοδος είναι κυρίως κατάλληλη για δεδομένα με διακριτές (μη συνεχείς) τιμές, που δίνονται στη μορφή  $(x_i, y_i)$ ,  $i=1, \dots, ns$ , με  $x_i$  το διάνυσμα των μεταβλητών σχεδίασης και  $y_i$  την αντίστοιχη

τιμή εξόδου. Η κατασκευή του μοντέλου γίνεται ως ακολούθως [Evans & Fisher, 1994]:

- ο Αρχικά συλλέγονται τα δεδομένα
- ο Ο αλγόριθμος εκπαίδευσης φτιάχνει ένα δένδρο απόφασης
  - Πρώτον, επιλέγοντας την “καλύτερα διαχωριζόμενη” μεταβλητή
  - Δεύτερον, αναδρομικά καλώντας τα υποσύνολα δεδομένων που προκύπτουν σε κάθε επανάληψη.

Είναι γενικά αποδοτικότερο να κατασκευάζονται δένδρα επιλέγοντας χαρακτηριστικά που ελαχιστοποιούν την ποσότητα πληροφορίας που απαιτείται για την ένταξη των δεδομένων σε κάποια υποκατηγορία και τα απαιτούμενα μαθηματικά για το σκοπό αυτό δίνονται στο [Evans & Fisher, 1994].

Η τεχνική που ακολουθήθηκε στην παρούσα εργασία, κάνει χρήση ANN (συγκεκριμένα ενός RBFN ANN) για μια χαμηλού υπολογιστικού κόστους αρχική εκτίμηση της καταλληλότητας των υποψήφιων λύσεων. Οι περισσότερα υποσχόμενες λύσεις υπόκεινται στη συνέχεια σε αναλυτική αξιολόγηση με τον εκάστοτε επιλύτη. Η διαδικασία συνεχίζεται με το ANN να εκπαιδεύεται ξανά για κάθε νέα γενεά, αντλώντας υποδείγματα από μια βάση δεδομένων που ενημερώνεται διαρκώς με τις καλύτερες των λύσεων. Στη συνέχεια ακολουθεί μια διεξοδική ανάλυση της βασικής θεωρίας των ANN.

Αξίζει στο σημείο αυτό να αναφερθεί ότι ο συνδυασμός Νευρωνικών Δικτύων και Εξελικτικών Αλγορίθμων δεν αποτελεί κάτι νέο. Συνηθέστερα οι ΕΑ χρησιμοποιούνται για τον προσδιορισμό της βέλτιστης αρχιτεκτονικής ενός ANN, που με τη σειρά του καλείται να επιλύσει κάποιο συγκεκριμένο πρόβλημα. Πληθώρα τέτοιων εφαρμογών περιγράφονται στη διεθνή βιβλιογραφία [Schmitz, 1999], [Mitchell, 1996], [Patel, Honavar & Balakrishnan].

## 3.2. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ - ANN

### 3.2.1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΤΗΣ ΘΕΩΡΙΑΣ ΤΩΝ ANN

Στη συγκεκριμένη ενότητα θα ορίσουμε τι είναι ένα νευρωνικό δίκτυο και θα δούμε τα διάφορα μοντέλα αναπαράστασής του. Εν συνεχεία θα αναλύσουμε τα δομικά του μέρη και τους βασικούς κανόνες που διέπουν τη μετάδοση σημάτων και τη διεξαγωγή υπολογισμών.

#### 3.2.1.1. ΤΙ ΕΙΝΑΙ ΕΝΑ ANN

Ένα ANN, αναφερόμενο επίσης στη διεθνή βιβλιογραφία ως *νευρο-υπολογιστής (neurocomputer)* ή *διασυνδεδεμένο δίκτυο (connectionist network)* ή *παράλληλος διανεμημένος επεξεργαστής (parallel distributed processor)*, είναι μια μηχανή σχεδιασμένη να μοντελοποιεί τον τρόπο με τον οποίο ο ανθρώπινος εγκέφαλος διεκπεραιώνει μια συγκεκριμένη εργασία ή λειτουργία. Θεωρώντας τον εγκέφαλο ως έναν πολύπλοκο, μη γραμμικό και παράλληλο υπολογιστή με δυνατότητα να αυτοοργανώνει τα δομικά του μέρη, τους νευρώνες και εστιάζοντας το ενδιαφέρον μας στα ANN που πραγματοποιούν χρήσιμους υπολογισμούς μέσω μιας διαδικασίας εκμάθησης, μπορούμε να δώσουμε για τα τελευταία τον εξής ορισμό [Haykin, 1999]:

“Ένα ANN είναι ένας μαζικά παράλληλος διανεμημένος επεξεργαστής, αποτελούμενος από απλές μονάδες επεξεργασίας, διαθέτουσες μια φυσική ροπή στην αποθήκευση εμπειρικής γνώσης και

την επαναδιάθεση της για χρήση. Ομοιάζει δε στον εγκέφαλο κατά δύο έννοιες:

1. Η γνώση του δικτύου αποκτάται από το περιβάλλον του μέσω μιας διαδικασίας εκμάθησης.
2. Η ισχύς των συνδέσεων μεταξύ των νευρώνων, γνωστή ως συναπτικά βάρη, αποτελεί το μέσο αποθήκευσης της αποκτώμενης γνώσης.”

Το δίκτυο συνήθως υλοποιείται με τη χρήση ηλεκτρονικών εξαρτημάτων ή προσομοιώνεται σε ψηφιακό υπολογιστή με χρήση κατάλληλου λογισμικού. Η διαδικασία που χρησιμοποιείται τότε για την πραγματοποίηση της εκμάθησης, καλείται *αλγόριθμος εκμάθησης (learning algorithm)*.

Στην πράξη ένα ANN δεν μπορεί να λειτουργήσει ανεξάρτητα και συνήθως αποτελεί τμήμα μίας ολοκληρωμένης εφαρμογής, στα πλαίσια της οποίας καλείται να εκτελέσει μια λειτουργία αρμόζουσα στις εγγενείς ικανότητές του. Μερικές από τις σημαντικότερες ιδιότητες των ANN πέρα από την υπολογιστική τους ισχύ και την ικανότητά τους να γενικεύουν μαθαίνοντας, είναι οι εξής [Haykin, 1999]:

- Η *μη γραμμικότητα (nonlinearity)*, ιδιότητα που οφείλεται στην διασύνδεση μη γραμμικών νευρώνων και είναι ιδιαίτερα χρήσιμη όταν η είσοδος του δικτύου είναι μη γραμμικής φύσεως (περίπτωση αναγνώρισης φωνής).
- Η *αντιστοίχιση εισόδου – εξόδου (input – output mapping)*, όπως στην περίπτωση της *κατηγοριοποίησης προτύπων (pattern classification)*, όπου στο δίκτυο η είσοδος που εκφράζει ένα φυσικό αντικείμενο ή γεγονός, αποδίδεται σε μια συγκεκριμένη κλάση αντικειμένων ή γεγονότων.
- Η *προσαρμοστικότητα (adaptivity)* είναι η ιδιότητα του δικτύου να αναπροσαρμόζει τα βάρη των συναπτικών του συνδέσεων προκειμένου να ανταποκριθεί στις μεταβολές στο περιβάλλον του.
- Η ικανότητα *τεκμηρίωσης της απόκρισης (evidential response)* στα πλαίσια της κατηγοριοποίησης προτύπων, έγκειται στη δυνατότητα σχεδιασμού του ANN έτσι ώστε να παρέχει πληροφόρηση για την εμπιστοσύνη στη ληφθείσα απόφαση.
- Η γνώση αναπαρίσταται από ένα ANN μέσω της δομής και του τρόπου ενεργοποίησής του. Δεδομένου λοιπόν ότι ο κάθε νευρώνας επηρεάζεται από τη λειτουργία ολόκληρου του δικτύου, τα ANN έχουν μια εγγενή ικανότητα διαχείρισης της *συναφούς πληροφορίας (contextual information)*.
- Η *αντοχή σε σφάλμα (fault tolerance)* αποτελεί σημαντικό πλεονέκτημα των ANN που οφείλεται στον τρόπο που κατανέμεται η πληροφορία μέσα στο δίκτυο. Σαν αποτέλεσμα του γεγονότος αυτού αποφεύγεται η κατάρρευση του συστήματος ενώ η απόδοση του δεν μειώνεται σημαντικά παρά για μεγάλο αριθμό αστοχούντων νευρώνων.
- Η παράλληλη επεξεργαστική φύση των ANN τα καθιστά ιδανικά για *κατά VLSI (very-large-scale-integration) εφαρμογή*, μεθοδολογία βάσει της οποίας εξαιρετικά πολύπλοκες συμπεριφορές αντιμετωπίζονται με αυστηρά ιεραρχικό τρόπο.

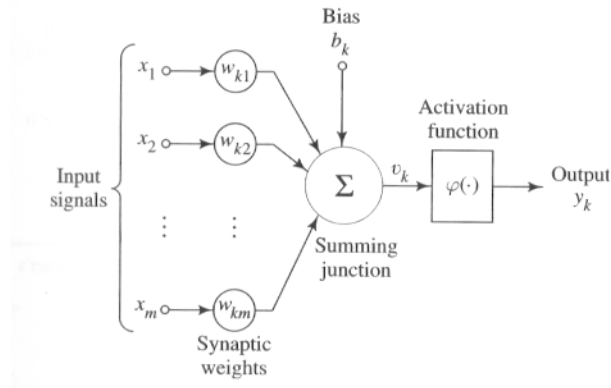
### 3.2.1.2. ΕΝΑΛΛΑΚΤΙΚΑ ΜΟΝΤΕΛΑ ΝΕΥΡΩΝΩΝ

Ο νευρώνας αποτελεί τη στοιχειώδη υπολογιστική μονάδα ενός ANN. Τα τρία βασικά συστατικά μέρη του είναι:

- Ένα σύνολο από συνάψεις, καθεμία από τις οποίες έχει το δικό της βάρος. Σαν αποτέλεσμα αυτού, κάθε είσοδος, έστω  $x_j$ , της σύναψης  $j$ , συνδέεται στον νευρώνα  $k$  πολλαπλασιαζόμενη με το βάρος  $w_{kj}$  με τα βάρη να παίρνουν και αρνητικές τιμές.

- Ένα γραμμικό συνθέτη (*linear combiner*), που υπολογίζει το σταθμισμένο άθροισμα των εισόδων.
- Μια συνάρτηση ενεργοποίησης (*activation function*), που σκοπό έχει τον περιορισμό του εύρους της εξόδου, συνήθως στα διαστήματα  $[0, 1]$  ή  $[-1, 1]$ .

Το περιγραφέν μοντέλο νευρώνα είναι αυτό του Σχήματος 3.1. Στο ίδιο σχήμα εμφανίζεται και ο παράγοντας προδιάθεσης (*bias*)  $b_k$ , που επιδρά αυξομειώνοντας την είσοδο ανάλογα με τον αν είναι θετικός ή αρνητικός αντίστοιχα.



**Σχήμα 3.1.** Μη γραμμικό μοντέλο νευρώνα [Haykin, 1999].

Μαθηματικά το αναλυόμενο μοντέλο περιγράφεται ως εξής:

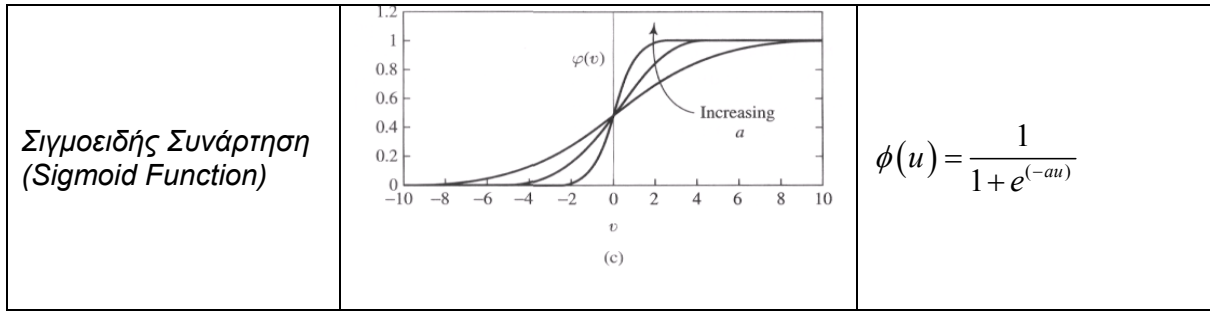
$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.17)$$

και

$$y_k = \phi(u_k + b_k) \quad (3.18)$$

όπου,  $x_1, x_2, \dots, x_m$  οι είσοδοι,  $w_{k1}, w_{k2}, \dots, w_{km}$  τα βάρη του νευρώνα  $k$ ,  $y_k$  η έξοδός του και  $\phi(\cdot)$  η συνάρτηση ενεργοποίησης. Τρεις εναλλακτικές επιλογές για την τελευταία φαίνονται στο στον Πίνακα 3.1.

Συνάρτηση Κατωφλιού (Threshold Function)	<p>(a)</p>	$\phi(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases}$
Τμηματικά Γραμμική Συνάρτηση (Piecewise- Linear Function)	<p>(b)</p>	$\phi(u) = \begin{cases} 1, & u \geq +\frac{1}{2} \\ u, & +\frac{1}{2} > u > -\frac{1}{2} \\ 0, & u \leq -\frac{1}{2} \end{cases}$



**Πίνακας 3.1.** Τρεις γνωστές συναρτήσεις ενεργοποίησης.

Χρησιμοποιώντας όρους γραφημάτων ροής σήματος (*signal flow graphs*)<sup>4</sup>, μπορούμε να μοντελοποιήσουμε το δίκτυο του Σχήματος 3.1. όπως φαίνεται στο Σχήμα 3.2. Ο τρίτος και τελευταίος εναλλακτικός τρόπος αναπαράσταση ενός δικτύου είναι η περιγραφή μέσω του *διαγράμματος της αρχιτεκτονικής του* (*architectural graph*), όπως φαίνεται στο Σχήμα 3.3.

Η ύπαρξη *ανάδρασης* (*feedback*) σε ένα δυναμικό σύστημα ορίζεται ως η κατάσταση κατά την οποία η είσοδος του συστήματος επηρεάζεται από την έξοδο του συστήματος ή κάποιου υποσυστήματος αυτού. Χαρακτηριστικό παράδειγμα τέτοιου δικτύου είναι αυτό του Σχήματος 3.4, όπου ο τελεστής A χαρακτηρίζει τον κλάδο της πρόσθιας τροφοδότησης και ο B τον κλάδο της οπισθοτροφοδότησης. Μαθηματικά η διαδικασία περιγράφεται ως εξής:

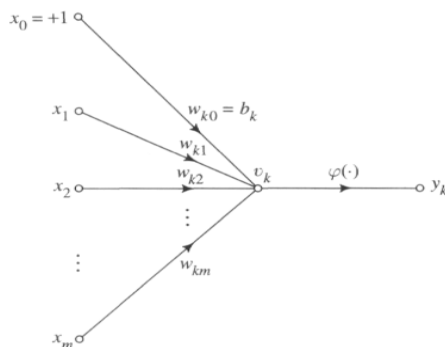
$$y_k(n) = A[x'_j(n)] \quad (3.19)$$

$$x'_j(n) = x_j(n) + B[y_k(n)] \quad (3.20)$$

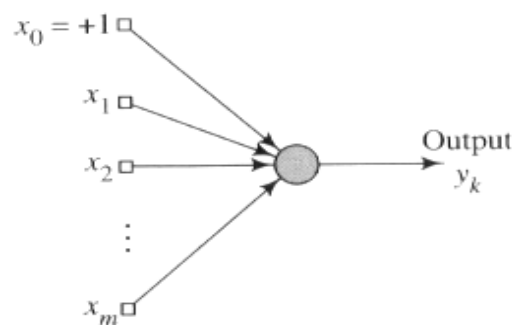
ΟΠΟΤΕ, ΤΕΛΙΚΑ:

$$y_k(n) = \frac{A}{1 - AB} [x_j(n)] \quad (3.21)$$

όπου, η ποσότητα  $\frac{A}{1 - AB}$  αποτελεί τον λεγόμενο *τελεστή κλειστού βρόγχου* (*closed-loop operator*).



**Σχήμα 3.2.** Γράφημα ροής σήματος ενός νευρώνα [Haykin, 1999].



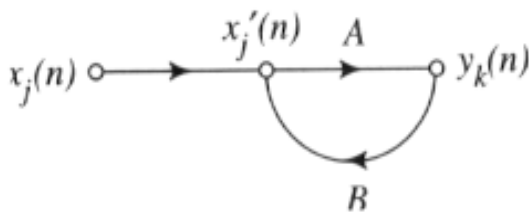
**Σχήμα 3.3.** Διάγραμμα αρχιτεκτονικής νευρώνα [Haykin, 1999].

<sup>4</sup> Για τους βασικούς κανόνες που διέπουν τη δημιουργία γραφημάτων ροής σήματος ο ενδιαφερόμενος αναγνώστης παραπέμπεται στο [Haykin, 1999].

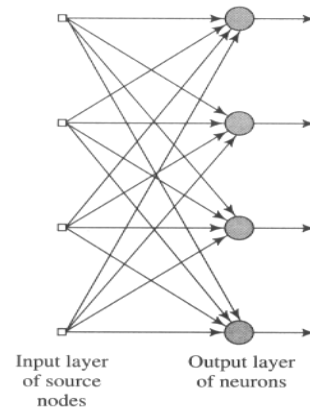
### 3.2.1.3. ΕΝΑΛΛΑΚΤΙΚΕΣ ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΔΙΚΤΥΩΝ

Τρεις είναι οι θεμελιωδώς διαφορετικές αρχιτεκτονικές δικτύου που υπάρχουν [Haykin, 1999]:

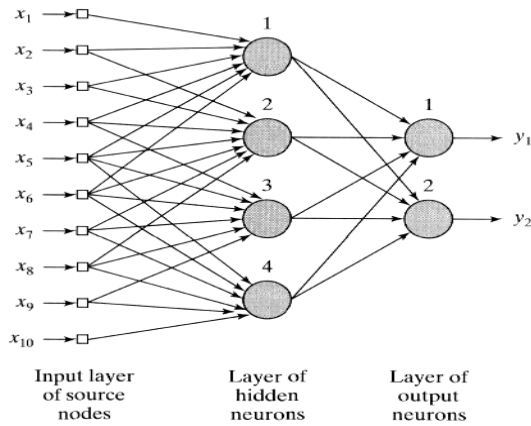
- Τα *Μονής Στοιβάδας Πρόσθιας Τροφοδότησης Δίκτυα* (*Single-Layer Feedforward Networks*) που αποτελούνται από μια στοιβάδα νευρώνων οι οποίοι προσλαμβάνουν την είσοδο (*input layer*) και την μεταβιβάζουν προς τη *στοιβάδα εξόδου* (*output layer*), αλλά όχι και το αντίστροφο (Σχήμα 3.5). Ο χαρακτηρισμός ως “μονής στοιβάδας” αναφέρεται στο γεγονός ότι υπολογισμοί γίνονται μόνο στη στοιβάδα εξόδου, με αποτέλεσμα εκείνη της εισόδου να μην μετράται.
- Τα *Πολλαπλής Στοιβάδας Πρόσθιας Τροφοδότησης Δίκτυα* (*Multilayer Feedforward Networks*) διαφέρουν λόγω της ύπαρξης ενός ή περισσοτέρων *κρυφών στοιβάδων* (*hidden layers*) υπολογιστικών νευρώνων. Δίνεται έτσι στο δίκτυο η δυνατότητα εξαγωγής στατιστικών υψηλότερης τάξης, ιδιότητα ιδιαίτερα χρήσιμη για δίκτυα με μεγάλου μεγέθους στοιβάδα εισόδου. Το δίκτυο του Σχήματος 3.6, χαρακτηρίζεται επίσης ως *πλήρως συνδεδεμένο* (*fully connected*) μιας και κάθε κόμβος μίας στοιβάδας συνδέεται με όλους τους κόμβους της προηγούμενης. Σε διαφορετική περίπτωση το δίκτυο ονομάζεται *μερικώς συνδεδεμένο* (*partially connected*).
- Τα *Αναδρομικά Δίκτυα* (*Recurrent Networks*) διαφοροποιούνται χάρη στην ύπαρξη τουλάχιστον ενός βρόγχου ανάδρασης ή ακόμη και *αυτοανατροφοδότησης* (*self-feedback*) στην περίπτωση που η έξοδος ενός νευρώνα τροφοδοτείται πίσω στην είσοδό του. Η δυναμική συμπεριφορά που αποκτά το δίκτυο έχει σαφή επίδραση στην μαθησιακή του ικανότητα. Παράδειγμα τέτοιου δικτύου αποτελεί αυτό του Σχήματος 3.7.



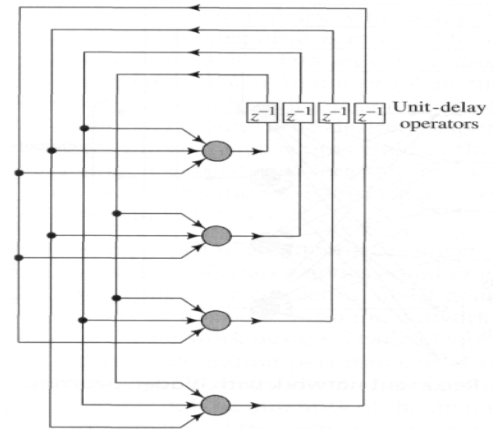
**Σχήμα 3.4.** Γράφημα ροής σήματος ενός συστήματος με ένα βρόγχο ανάδρασης [Haykin, 1999].



**Σχήμα 3.5.** Δίκτυο Πρόσθιας Τροφοδότησης Μονής Στοιβάδας [Haykin, 1999].



**Σχήμα 3.6.** Δίκτυο Πρόσθιας Τροφοδότησης Πολλαπλής Στοιβάδας [Haykin, 1999].



**Σχήμα 3.7.** Αναδρομικό Δίκτυο δίχως αυτοανατροφοδότηση και κρυφή στοιβάδα [Haykin, 1999].

#### 3.2.1.4. Η ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΕΚΠΑΙΔΕΥΣΗΣ

Η ικανότητα ενός ANN να μαθαίνει είναι πρωταρχικής σημασίας για τη χρήση του. Αν και δεν είναι εύκολο να δοθεί σαφής ορισμός για την εκμάθηση, εδώ θα υιοθετήσουμε τον ακόλουθο [Haykin, 1999]:

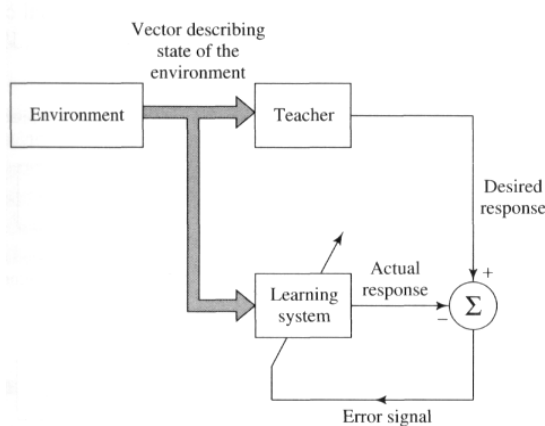
“Εκμάθηση είναι η διαδικασία κατά την οποία ένα ANN προσαρμόζει τις ελεύθερες παραμέτρους του μέσω μιας διαδικασίας διέγερσης από το περιβάλλον στο οποίο βρίσκεται. Ο τύπος της εκμάθησης καθορίζεται από τον τρόπο που γίνονται οι αλλαγές των παραμέτρων. ”

Ένα προκαθορισμένο σύνολο βημάτων για την επίλυση ενός προβλήματος εκμάθησης ονομάζεται *αλγόριθμος εκμάθησης (learning algorithm)*. Οι *κανόνες εκμάθησης (learning rules)*, περιγράφουν πώς ένα ANN μαθαίνει, οι σημαντικότεροι δε εξ αυτών είναι [Haykin, 1999]:

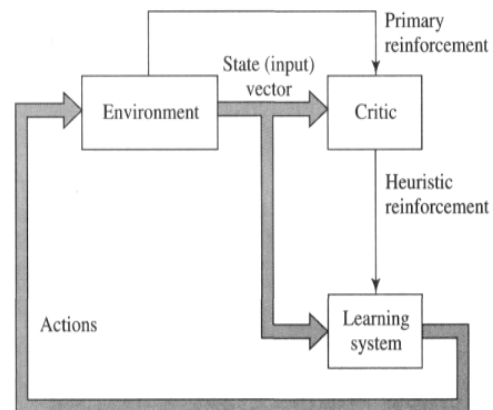
- Ο *κανόνας διόρθωσης σφάλματος (error-correction learning)*, που στοχεύει στην προσαρμογή των παραμέτρων του ANN ώστε η έξοδός του, να βρίσκεται όσο πιο κοντά γίνεται σε μια προκαθορισμένη επιθυμητή τιμή.
- Η *βασισμένη στη μνήμη εκμάθηση (memory-based learning)*, λειτουργεί με την ρητή απομνημόνευση των δεδομένων της εκπαίδευσης.
- Ο κανόνας της *Hebbian εκμάθησης (Hebbian learning)*, εμπνέεται από τα χώρο της νευροβιολογίας και πριμοδοτεί τις συνάψεις τα άκρα των οποίων ενεργοποιούνται ταυτόχρονα.
- Η *ανταγωνιστική εκμάθηση (competitive learning)* είναι μέθοδος νευροβιολογικής προέλευσης και στηρίζεται στον ανταγωνισμό μεταξύ των νευρώνων της στοιβάδας εξόδου για το ποιος θα ενεργοποιηθεί.
- Η *κατά Boltzmann εκμάθηση (Boltzmann learning)* εμπνέεται από ιδέες της στατιστικής μηχανικής.

Εξετάζοντας τώρα τη διαδικασία εκπαίδευσης ενός ANN από τη σκοπιά της αλληλεπίδρασης με το περιβάλλον του, κάνουμε λόγο για *παραδείγματα εκμάθησης (learning paradigms)*, που διακρίνονται ως εξής:

- Στην *επιβλεπόμενη εκμάθηση (supervised learning)* ή *εκμάθηση μετά δασκάλου (learning with a teacher)* όπου ο επιβλέπων έχει γνώση του περιβάλλοντος εκφρασμένη ως ένα σύνολο υποδειγμάτων εισόδου-εξόδου. Η προσαρμογή των παραμέτρων του ANN γίνεται υπό την επίδραση τόσο της εισόδου όσο και του σφάλματος, που εκφράζει την διαφορά μεταξύ επιθυμητής και πραγματικής απόκρισης του δικτύου. Η διαδικασία επαναλαμβάνεται έως ότου το ANN μάθει να “μιμείται” τον επιβλέποντα επαρκώς, σύμφωνα με κάποιο προκαθορισμένο στατιστικό κριτήριο (Σχήμα 3.8).
- Στην εκμάθηση χωρίς δάσκαλο (*learning without a teacher*), όπου δεν υπάρχει επιβλέπων ήτοι το ANN δεν έχει να μάθει κάποιο σύνολο υποδειγμάτων εισόδου-εξόδου. Δυο υποπεριπτώσεις διακρίνονται εδώ:
  - Η *ενισχυόμενη εκμάθηση (reinforcement learning)* ή *νευροδυναμικός προγραμματισμός (neurodynamic programming)*, συνίσταται στην εκμάθηση μέσω διαρκούς αλληλεπίδρασης του δικτύου με το περιβάλλον του, με σκοπό την ελαχιστοποίηση ενός βαθμωτού δείκτη απόδοσης. Η διαδικασία παρουσιάζεται στο Σχήμα 3.9.

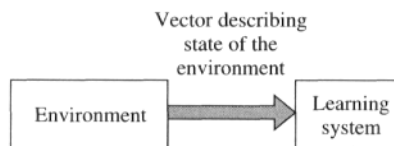


**Σχήμα 3.8.** Διάγραμμα ροής επιβλεπόμενης μάθησης [Haykin, 1999].



**Σχήμα 3.9.** Διάγραμμα ροής ενισχυόμενης μάθησης [Haykin, 1999].

- Η *μη επιβλεπόμενη εκμάθηση (unsupervised learning)*, προϋποθέτει την ύπαρξη ενός μέτρου της ποιότητας της αναπαράστασης που το ANN καλείται να μάθει. Η βελτιστοποίηση των ελεύθερων παραμέτρων του δικτύου γίνεται με βάση αυτό ακριβώς το μέτρο, όπως δείχνει το Σχήμα 3.10.



**Σχήμα 3.10.** Διάγραμμα ροής μη επιβλεπόμενης μάθησης [Haykin, 1999].

Κλείνοντας τη συγκεκριμένη ενότητα θα πρέπει να τονίσουμε ότι απώτερος σκοπός της διαδικασίας εκπαίδευσης, είναι η δημιουργία ενός δικτύου που να κάνει την καλύτερη δυνατή *γενίκευση (generalization)*, δηλαδή να εκτελεί την προγραμματισμένη για αυτό εργασία αποτελεσματικά σε άγνωστα δεδομένα, πέρα από εκείνα που χρησιμοποιήθηκαν για την εκπαίδευσή του.

### 3.2.1.5. ΕΦΑΡΜΟΓΕΣ ΤΩΝ ANN

Έξι βασικές εφαρμογές των ANN, μεγάλου πρακτικού ενδιαφέροντος διακρίνει ο Haykin [Haykin, 1999]:

- Η *συσχέτιση προτύπων (pattern association)*, που συνίσταται στην απόδοση μιας τιμής εισόδου σε μια αντίστοιχη τιμή εξόδου ως εξής:  $x_k \rightarrow y_k$ ,  $k=1,2,\dots,n$ , όπου  $n$  ο αριθμός των αποθηκευμένων στο δίκτυο προτύπων.
- Η *αναγνώριση προτύπων (pattern recognition)*, κατά την οποία ένα εισερχόμενο σήμα ή πρότυπο, καταχωρείται σε μια εκ των διατεταγμένων κλάσεων.
- Η *προσέγγιση συναρτήσεων (function approximation)* όπως η  $d = F(\cdot)$ , με την συνάρτηση-στόχο  $f(\cdot)$  να θεωρείται άγνωστη, γίνεται δεδομένου ενός συνόλου υποδειγμάτων  $\mathfrak{S} = \{(x_i, d_i)\}_{i=1}^N$ . Ουσιαστικά αποζητάμε ένα δίκτυο η  $F(\cdot)$  του οποίου να είναι όσο πιο κοντά γίνεται στην πραγματική  $f(\cdot)$ , δηλαδή:  $\|F(x) - f(x)\| < \epsilon$ , όπου  $\epsilon$  είναι ένας μικρός θετικός αριθμός.
- Ο *έλεγχος (control)* μίας διαδικασίας ή ενός συστήματος που επιθυμούμε να διατηρήσουμε σε ελεγχόμενη κατάσταση.
- Με χρήση των ANN ως φίλτρων μπορούμε να εκτελέσουμε τρεις διαφορετικές εργασίες, ήτοι:
  - *Φιλτράρισμα (filtering)*, όταν προσπαθούμε να εξάγουμε πληροφορία για κάποια ποσότητα σε δεδομένη χρονική στιγμή  $t$ , από δεδομένα που περιλαμβάνουν μέχρι και τη στιγμή αυτή.
  - *Εξομάλυνση (smoothing)*, όταν έχουμε δεδομένα για την εξαγωγή πληροφορίας και μετά τη στιγμή  $t$ . Αναμένουμε λοιπόν η εξομάλυνση να είναι στατιστικά πιο ακριβής από το φιλτράρισμα.
  - *Πρόβλεψη (prediction)*, όταν με δεδομένα μέχρι και τη χρονική στιγμή  $t$ , προσπαθούμε να εξάγουμε πληροφορία για την ποσότητα που μας ενδιαφέρει σε μια μελλοντική χρονική στιγμή έστω  $t'$ .
- *Διαμόρφωση δέσμης σήματος (beamforming)*, μια ειδική μορφή φιλτραρίσματος που αποσκοπεί στο διαχωρισμό της πηγής που μας ενδιαφέρει από τον θόρυβο του περιβάλλοντος της.

### 3.2.2. ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

Εδώ παρουσιάζεται μια σύντομη, σε μορφή πίνακα, ανασκόπηση της μοντέρνας εποχής των νευρωνικών δικτύων. Παρατίθενται τα σημαντικότερα γεγονότα που σημάδεψαν την εξέλιξη του επιστημονικού αυτού πεδίου τον τελευταίο μισό αιώνα με τις αντίστοιχες ημερομηνίες [Haykin, 1999, Callan, 1999].

1943	Η πρωτοποριακή εργασία των McCulloch και Pitts σηματοδοτεί την απαρχή της μοντέρνας περιόδου εξέλιξης των NN. Απέδειξαν ότι ένα δίκτυο αποτελούμενο από επαρκή αριθμό απλών, ταυτόχρονα ενεργών νευρώνων, συνεργαζόμενων μέσω συναπτικών συνδέσεων, δύναται να υπολογίσει οποιαδήποτε υπολογίσιμη συνάρτηση.
1948	Ο Wiener στο βιβλίο του Cybernetics, γίνεται ο πρώτος που επισημαίνει τη διασύνδεση μεταξύ στατιστικής μηχανικής και συστημάτων εκμάθησης.
1949	Ο Hebb στο The Organization of Behavior, εισάγει τον πρώτο αναλυτικό κανόνα εκμάθησης για την συναπτική προσαρμογή: «η σπουδαιότητα μιας συναπτικής σύνδεσης αυξάνει με τον αριθμό των ενεργοποιήσεων μέσω αυτής, των

	νευρώνων που συνδέει».
1954	Ο Minsky στη διδακτορική του διατριβή αποδελτιώνει μεγάλο μέρος αυτού που σήμερα αποτελεί το γνωστικό αντικείμενο των NN. Η ιδέα των μη γραμμικών προσαρμοστικών φίλτρων προτείνεται από τον Gabor. Κατά την υλοποίησή της χρησιμοποιείται μια διαδικασία εκμάθησης με ταυτόχρονη παρουσίαση στο σύστημα δειγμάτων στοχαστικών διαδικασιών και των επιθυμητών αποτελεσμάτων.
1956	Σε μία εργασία των Rochester, Holland, Haibt και Duda γίνεται η πρώτη χρήση προσομοίωσης μέσω υπολογιστή για να ελεγχθεί η θεωρία εκμάθησης του Hebb. Ο Uttley δείχνει ότι ένα NN με μεταβλητές συνάψεις μπορεί να μάθει να κατηγοριοποιεί δυαδικά πρότυπα στις αντίστοιχες κλάσεις. Το πρόβλημα του σχεδιασμού αξιόπιστων δικτύων βασιζόμενων στους αναξιόπιστους νευρώνες επιλύεται από τον von Neumann με χρήση της ιδέας της υπερεπάρκειας.
1958	Εισαγωγή του perceptron από τον Rosenblatt και της αντίστοιχης καινοτόμου μεθόδου εκμάθησης με επίβλεψη.
1960	Οι Widrow και Hoff εισάγουν τον αλγόριθμο ελαχίστων μέσων τετραγώνων (Least Mean-Square, LMS) και με τη χρήση του δημιουργούν τα Adaline(adaptive linear element) και Madaline(multiple-adaline).
1961	Ο Minsky διατυπώνει το credit assignment problem.
1963	Οι Winograd και Cowan προτείνουν την ιδέα της κατανεμημένης αναπαράστασης: ένας αυξημένος αριθμός στοιχείων αναπαριστά συλλογικά μια ιδέα, αυξάνοντας έτσι την ευστάθεια και την παραλληλότητα.
1967	Στο βιβλίο Computation: Finite and Infinite Machines του Minsky τα αποτελέσματα των McCulloch και Pitts τοποθετούνται στα πλαίσια της θεωρίας υπολογισμού και αυτομάτων.
1969	Σε μια εργασία των Willshaw, Buneman και Longuet-Higgins εμφανίζονται δυο ενδιαφέροντα μοντέλα δικτύων: ένα απλό οπτικό σύστημα που υλοποιεί μια correlation memory και ένα αντίστοιχο με την οπτική μνήμη NN. Το βιβλίο των Minsky και Papert θέτει σαφείς περιορισμούς στο τι μπορούν να κάνουν τα SLPs (Single Layer Perceptrons).
1976	Οι Willshaw και von der Malsburg δημοσιεύουν την πρώτη εργασία σχετικά με τη δημιουργία αυτοοργανούμενων χαρτών (self-organised maps).
1980	Μια καινούργια μέθοδος αυτοοργάνωσης, η adaptive resonance theory (ART) προτείνεται από τον Grossberg
1982	Ο Hopfield άλλαξε ριζικά το τοπίο, αναλύοντας τη σχέση μεταξύ των αναδρομικών νευρωνικών και αντίστοιχων μοντέλων της στατιστικής φυσικής. Για πρώτη φορά στην εργασία του αυτή αναλύει διεξοδικά την ιδέα της αποθήκευσης πληροφορίας σε δυναμικά σταθερά δίκτυα.
1983	Το ζήτημα των self-organised maps επαναπροσδιορίζεται από τον Kohonen, η εργασία του οποίου αποτελεί εφεξής μέτρο σύγκρισης στον τομέα αυτό. Την ίδια χρονιά προτείνεται και ο αλγόριθμος της Προσομοιωμένης Ανόπτησης.
1985	Οι Ackley, Hinton και Sejnowski προτείνουν την στοχαστική Μηχανή Boltzmann, την πρώτη επιτυχημένη υλοποίηση MLNN(Multi Layer NN). Καταδεικνύουν έτσι ότι η εικασία των Minsky και Papert (1969) ήταν θεμελιωμένη εσφαλμένα.
1986	Προτείνεται από τους Rumelhart, Hinton και Williams ο δημοφιλέστερος όπως απεδείχθη αλγόριθμος εκμάθησης, για εφαρμογές μηχανικής εκμάθησης. Δεν είναι άλλος από τον αλγόριθμο Back-propagation, η βασική ιδέα του οποίου πρωτοεισηχθεί από τους Bryson και Ho το 1969.
1988	Ο Linsker περιγράφει μια νέα αρχή αυτοοργάνωσης, κάνοντας χρήση της θεωρίας πληροφορίας του Shannon. Οι Broomhead και Lowe εισάγουν τα RBFN μια εναλλακτική των MLP μεθοδολογία η βασική ιδέα της οποίας πηγάζει από τη μέθοδο των συναρτήσεων δυναμικού (potential functions).
1990	Η θεωρία κανονικοποίησης του Tikhonov εφαρμόζεται στα RBFN από τους Girosi

	και Poggio.
1995	Η μεθοδολογία των SVM εισάγεται από τους Vapnik και Cortes. Πρόκειται για υπολογιστικά ισχυρά δίκτυα επιβλεπόμενης εκμάθησης με εφαρμογές στην αναγνώριση προτύπων την παλινδρόμηση και την εκτίμηση πυκνότητας.

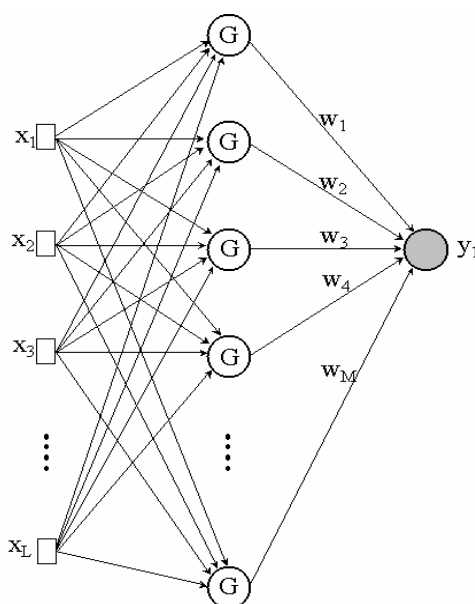
**Πίνακας 3.2.** Ορόσημα της εξέλιξης των νευρωνικών δικτύων.

Συμπληρωματικά στα παραπάνω οφείλουμε να τονίσουμε τη στενή αλληλεπιδραστική σχέση μεταξύ της επιστήμης των NN και της νευρολογίας, οι ρίζες της οποίας βρίσκονται στην πρωτοποριακή δουλειά του Ramon y Cajal πίσω στα 1911, όταν και πρότεινε την ιδέα των νευρώνων ως δομικών στοιχείων του εγκεφάλου. Έκτοτε η ανταλλαγή ιδεών και εξηγητικών σχημάτων είναι διαρκής. Στενή είναι επίσης η σχέση των NN και με άλλα επιστημονικά πεδία όπως η ψυχολογία, τα μαθηματικά η φυσική και η μηχανολογία.

Άξια μνείας είναι επίσης η διακύμανση του ενδιαφέροντος γύρω από τα NN. Περίοδοι αυξημένων προσδοκιών και ερευνητικής δραστηριότητας όπως η δεκαετία του '80, διαδέχθηκαν άλλες ουσιαστικής εγκατάλειψης με χαρακτηριστικότερο παράδειγμα τα χρόνια του '70, οπότε και οι δυνατότητες των NN αμφισβητήθηκαν σοβαρά. Σήμερα η μεγαλύτερη βεβαιότητα γύρω από τα NN είναι η αύξηση των εφαρμογών τους και η διεύρυνση της θεωρητικής τους τεκμηρίωσης, μέσα από την απάντηση σε ερωτήματα όπως αυτό για την ύπαρξη κεντρικού ρόλου για το χάος, στην μελέτη των NN.

### 3.2.3. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΣΥΝΑΡΤΗΣΕΩΣ ΑΚΤΙΝΙΚΗΣ ΒΑΣΗΣ – RBFN

Τα RBFN συνιστούν μια από τις διάφορες δυνατές προσεγγίσεις στο πρόβλημα της σχεδίασης νευρωνικών δικτύων, βλέποντάς το ως ισοδύναμο του *προβλήματος προσέγγισης μιας καμπύλης (curve fitting or approximation)*, σε έναν υψηλής διάστασης χώρο. Επομένως εδώ η εκμάθηση ισοδυναμεί με την εύρεση μιας υπερεπιφάνειας σε έναν πολυδιάστατο χώρο που δίνει την καλύτερη προσέγγιση των *δεδομένων εκπαίδευσης (training data)*, με το κριτήριο της “καλύτερης προσέγγισης”, να μετράται κατά κάποια στατιστική έννοια. Αντίστοιχα η γενίκευση συνίσταται στη χρήση αυτής της επιφάνειας για την παρεμβολή στα *δεδομένα δοκιμής (test data)*.



**Σχήμα 3.11.** Νευρωνικό Δίκτυο Συναρτήσεως Ακτινικής Βάσης - RBFN

Ένα τυπικό RBFN φαίνεται στο Σχήμα 3.11. Αποτελείται συνήθως από τρεις στοιβάδες, εισόδου, κρυφή και εξόδου, με τις συνδέσεις μεταξύ των δυο τελευταίων να αποτελούν και τις μοναδικές μεταβαλλόμενες παραμέτρους. Ουσιαστικά προβαίνει σε μια μη γραμμική αντιστοίχιση από τον χώρο της εισόδου (input space), στον κρυμμένο χώρο (hidden space)  $R^L \rightarrow R^M$ , ακολουθούμενη από μια γραμμική αντιστοίχιση από τον κρυμμένο στο χώρο της εξόδου (output space)  $R^M \rightarrow R^1$ . Μαθηματικά αυτό δικαιολογείται με βάση το θεώρημα του Cover<sup>5</sup>, σύμφωνα με το οποίο ένα πρόβλημα κατηγοριοποίησης προτύπων είναι πιθανότερο να είναι γραμμικώς διαχωρίσιμο αν μεταφερθεί σε υψηλότερης διάστασης χώρο. Επιπλέον η διάσταση του κρυμμένου χώρου είναι ευθέως ανάλογη της ικανότητας του δικτύου για ομαλή αντιστοίχιση εισόδου-εξόδου κι επομένως για ακριβέστερη προσέγγιση [Cover, 1965].

Στα RBFN, οι κρυμμένοι κόμβοι σχετίζονται με τα λεγόμενα κέντρα των συναρτήσεων ακτινικής βάσης. Στο παράδειγμά μας (Σχήμα 3.11), τα  $M$  κέντρα μπορούν να ιδωθούν ως  $L$ -διάστατα διανύσματα, οπότε στο εξής θα συμβολίζονται ως:

$$\vec{c}^{(m)}, m = 1, M \quad (3.22)$$

Η συνήθης προσέγγιση που οδηγεί στη δημιουργία των λεγομένων *Δικτύων Κανονικοποίησης* (*Regularization Networks*), είναι η επιλογή ως κέντρων, των διανυσμάτων εισόδου του *συνόλου εκπαίδευσης* (*training set*), οπότε  $M = T$ , όπου  $T$  το μέγεθος του συνόλου εκπαίδευσης. Η επιλογή αυτή εκφράζεται ως ακολούθως:

$$\vec{c}_l^{(t)} = \vec{x}_l^{(t)}, t = 1, T, l = 1, L \quad (3.23)$$

οδηγεί ωστόσο σε πολύ μεγάλη ανάγκη για μνήμη και χρόνο CPU καθώς αυξάνει το  $T$ . Επιπλέον, αν το σύνολο εκπαίδευσης δεν δειγματοληπτεί επαρκώς το πεδίο των τιμών, μπορεί να οδηγηθούμε σε *υπερπροσαρμογή* (*overfitting*) ή/και κακή γενίκευση.

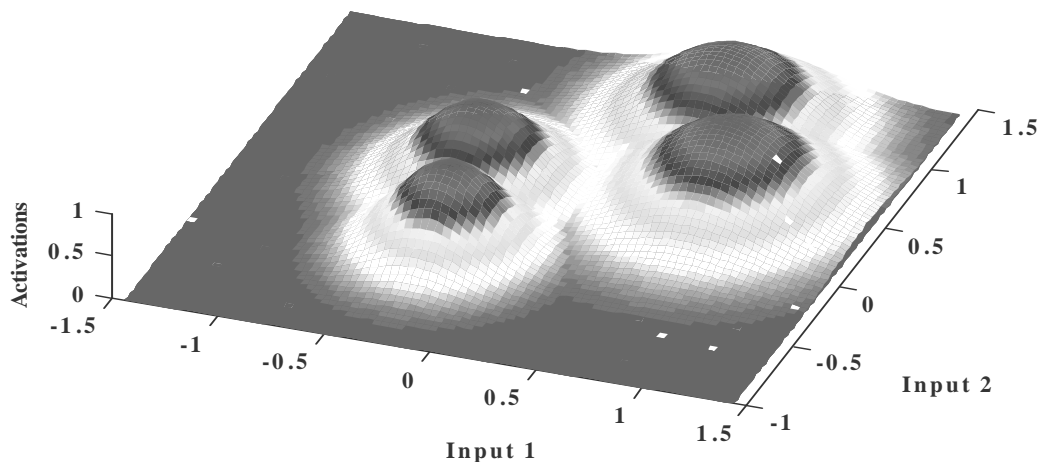
Προκειμένου να ξεπεράσουμε το συγκεκριμένο σκόπελο όπου αυτό είναι απαραίτητο (σε δίκτυα δηλαδή που ξεπερνούν τους 20-30 κρυφούς κόμβους), επιλέγουμε τα κέντρα μας να είναι ένα υποσύνολο του  $T$  ( $T > M$ ). Για την επιλογή αυτών των κέντρων χρησιμοποιήθηκε γεννήτρια τυχαίων αριθμών, με τη βοήθεια της οποίας ένα ποσοστό των διανυσμάτων του συνόλου εκπαίδευσης (συνήθως τα 2/3 αυτών), επιλέγονται ως κέντρα των ακτινικών συναρτήσεων.

Θεωρώντας ωστόσο για τους σκοπούς της παρουσίασής μας ότι  $M = T$ , δίχως βλάβη της γενικότητας, ο  $m$ -ιοστός κρυφός κόμβος παίρνει τιμή:

$$g_m^{(t)} = \Phi \left( \left\| \vec{x}^{(t)} - \vec{c}^{(t)} \right\|, r_n \right) \quad (3.24)$$

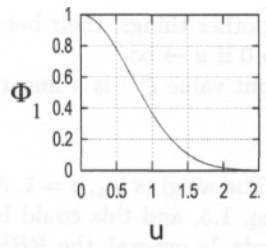
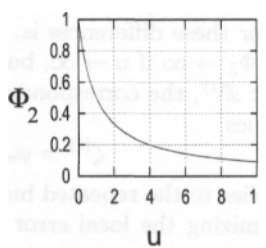
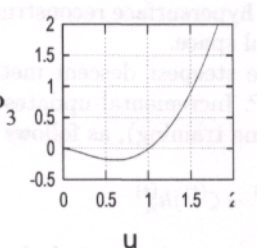
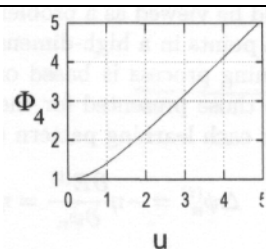
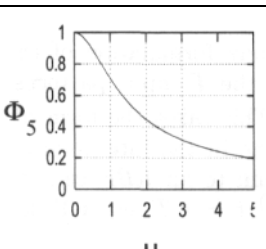
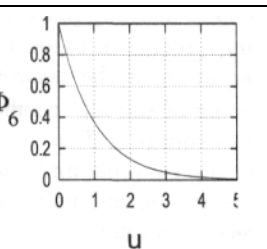
όπου,  $\vec{x}^{(t)}$  είναι το διάνυσμα εισόδου,  $\Phi(\cdot)$  η συνάρτηση ενεργοποίησης που εκτελεί την αντιστοίχιση  $R^M \rightarrow R^1$  και  $r_n$  η παράμετρος που ρυθμίζει το εύρος της RBF (Σχήμα 3.12).

<sup>5</sup> Περισσότερα για το εν λόγω θεώρημα μπορεί να βρει ο αναγνώστης στο [Haykin, 1999]



**Σχήμα 3.12.** Οι ενεργοποιήσεις τεσσάρων κέντρων συναρτήσεων ακτινικής βάσης για την περίπτωση διδιάστατου χώρου λύσεων [Schmitz, 1999].

Μερικές εναλλακτικές επιλογές για τη συνάρτηση  $\Phi(\cdot)$  και η μορφή τους φαίνονται στον Πίνακα 3.3 [Giannakoglou, 2001].

		
$\Phi_1(u, r_n) = \exp(-u^2 / r_n^2)$	$\Phi_2(u, r_n) = (u+1)^{-1}$	$\Phi_3(u, r_n) = u^2 \ln( u )$
		
$\Phi_4(u, r_n) = \sqrt{u^2 + r_n^2}$	$\Phi_5(u, r_n) = \left(\sqrt{u^2 + r_n^2}\right)^{-1}$	$\Phi_6(u, r_n) = \exp(-u / r_n)$

**Πίνακας 3.3.** Έξι πιθανές συναρτήσεις ενεργοποίησης για RBFN.

Η συμπεριφορά των παραπάνω συναρτήσεων είναι πολύ διαφορετική καθώς το  $u \rightarrow \infty$ , ενώ για παράδειγμα η  $\Phi_3 \rightarrow \infty$ , η  $\Phi_6 \rightarrow 0$ .

Με δεδομένη την είσοδο  $\vec{x}^{(t)}$ , η αντίστοιχη έξοδος του δικτύου είναι

$$y^{(t)} = w_m \cdot g_m^{(t)} \quad (3.25)$$

όπου  $w_m, m=1, M$  τα βάρη των συναπτικών συνδέσεων που πρέπει να αναπροσαρμοστούν έτσι ώστε να ελαχιστοποιηθεί το σφάλμα  $E^{(t)}$ , όπως είδαμε στην παράγραφο 3.2.1.5. για την διαδικασία προσέγγισης συναρτήσεων.

Αναφορικά με τη διαδικασία εκπαίδευσης του δικτύου έχουμε δύο επιλογές:

- Να ακολουθήσουμε την επαναληπτική διαδικασία εκπαίδευσης (*iterative learning process*), που στηρίζεται στην σταδιακή μείωση του σφάλματος  $E^{(t)}$ . Η αναπροσαρμογή των βαρών γίνεται σύμφωνα με τον τύπο:

$$w_m^{new} = w_m^{old} + \Delta w_m^{(t)} \quad (3.26)$$

πριν από την παρουσίαση κάθε νέου υποδείγματος, με

$$\Delta w_m^{(t)} = -\eta \frac{\partial E^{(t)}}{\partial w_m} = \eta (d^{(t)} - y^{(t)}) g_m^{(t)} \quad (3.27)$$

όπου,  $d^{(t)}$  είναι η επιθυμητή έξοδος-στόχος και  $y^{(t)}$  η πραγματική του δικτύου και  $\eta$  ο ρυθμός εκμάθησης (*learning rate*).

- Να ακολουθήσουμε την ευθεία διαδικασία εκπαίδευσης (*direct learning process*), που στηρίζεται στην μέσω πινάκων μορφοποίηση των βασικών εξισώσεων του προβλήματος. Έτσι από τον υπολογισμό των ενεργοποιήσεων κάθε κόμβου προκύπτει ένας πίνακας  $G$  μεγέθους  $(T \times M)$ , όπου υπενθυμίζουμε πως  $T$  είναι το μέγεθος του συνόλου εκπαίδευσης και  $M$  το πλήθος των κρυφών κόμβων. Συνεπώς κάθε γραμμή του πίνακα αντιστοιχεί σε ένα υπόδειγμα εκμάθησης και κάθε στήλη στο κέντρο μίας συνάρτησης. Η Εξίσωση (3.9) που δίνει την σχέση της γραμμικής αντιστοίχισης  $R^M \rightarrow R^1$ , παίρνει τώρα τη μορφή

$$G\bar{W} = \bar{D} \quad (3.28)$$

με τα διανύσματα των βαρών και της επιθυμητής εξόδου να είναι αντίστοιχα:

$$\begin{aligned} \bar{W} &= (w_1, w_2, \dots, w_M)^T \\ \bar{D} &= (d^{(1)}, d^{(2)}, \dots, d^{(T)})^T \end{aligned} \quad (3.29)$$

### 3.2.4. ΔΙΚΤΥΑ ΠΟΛΛΑΠΛΩΝ ΣΤΟΙΒΑΔΩΝ ΠΡΟΣΘΙΑΣ ΤΡΟΦΟΔΟΤΗΣΗΣ - MLP

Τα MLP αποτελούν μια διαφορετική προσέγγιση στο πρόβλημα της σχεδίασης νευρωνικών δικτύων υπό επίβλεψη. Ο αλγόριθμος οπισθοτροφοδότησης του σφάλματος (*Back Error Propagation algorithm - BEP*)<sup>6</sup> που εδώ χρησιμοποιείται για την εκπαίδευση μπορεί να εκληφθεί ως εφαρμογή της αναδρομικής τεχνικής γνωστής ως *στοχαστική προσέγγιση (stochastic approximation)*.

Δανειζόμενοι τους συμβολισμούς που υιοθετήσαμε στην περίπτωση των RBFN και για την απλή περίπτωση ενός MLP με τρεις στοιβάδες και μοναδική έξοδο, όμοιο με αυτό του Σχήματος 3.11, το δίκτυο ουσιαστικά προβαίνει στην μη γραμμική αντιστοίχιση

<sup>6</sup> Για ένα κατατοπιστικό διάγραμμα ροής του αλγορίθμου ο αναγνώστης παραπέμπεται στο [Callan, 1999].

$R^L \rightarrow R^M$  και ακολούθως στην επίσης μη γραμμική αντιστοίχιση  $R^M \rightarrow R^1$ . Κάθε φορά που ένα υπόδειγμα  $\vec{x}^{(t)}$  παρουσιάζεται στο δίκτυο, το σήμα προωθείται στα εξής δυο στάδια:

$$g_m^{(t)} = a(w_{lm}x_l^{(t)}), m=1, M, l=1, L \quad (3.30)$$

και

$$y^{(t)} = a(w_m g_m^{(t)}), m=1, M \quad (3.31)$$

όπου,  $a(\cdot)$  είναι η μη γραμμική συνάρτηση ενεργοποίησης που συνήθως επιλέγεται να είναι η σιγμοειδής συνάρτηση:

$$a(u) = \frac{1}{1 + e^{-u}} \quad (3.32)$$

Στόχος της διαδικασίας εκπαίδευσης είναι η αναπροσαρμογή των βαρών, που σκοπό έχει την σταδιακή μείωση του σφάλματος, που ορίζεται ως η διαφορά μεταξύ της επιθυμητής εξόδου-στόχου  $d^{(t)}$  και της πραγματικής εξόδου του δικτύου  $y^{(t)}$ . Χρησιμοποιώντας λοιπόν ένα σύνολο ζευγών εισόδου-εξόδου  $(\vec{x}^{(t)}, \vec{d}^{(t)})$ , με το σφάλμα για κάθε υπόδειγμα να είναι:

$$E^{(t)} = \frac{1}{2} (\|y^{(t)} - d^{(t)}\|)^2 \quad (3.33)$$

οι εναλλακτικές που έχουμε είναι δυο:

- Η *εκπαίδευση κατά "παρτίδα"* (*batch training*), οπότε τα βάρη του νευρωνικού αναπροσαρμόζονται αφού έχουν παρουσιαστεί όλα τα υποδείγματα του συνόλου εκπαίδευσης.
- Η *online εκπαίδευση* (*online training*), με τα βάρη να αναπροσαρμόζονται μετά την παρουσίαση κάθε υποδείγματος και πριν την παρουσίαση του επόμενου.

Στη δεύτερη περίπτωση που συνήθως αποδεικνύεται αποτελεσματικότερη, ο ΒΕΡ αλγόριθμος βασίζεται στην αρνητική κλίση του σφάλματος και στον ρυθμό εκμάθησης  $\eta$  που επιλέγει ο χρήστης. Τα βάρη αναπροσαρμόζονται λοιπόν ως εξής:

$$\begin{aligned} w_{lm}^{new} &= w_{lm}^{old} + \Delta w_{lm}^{(t)} \\ w_m^{new} &= w_m^{old} + \Delta w_m^{(t)} \end{aligned} \quad (3.34)$$

όπου

$$\begin{aligned} \Delta w_{lm}^{(t)} &= -\eta \frac{\partial E^{(t)}}{\partial w_{lm}} = \eta (d^{(t)} - y^{(t)}) a'(w_m g_m^{(t)}) w_m a'(w_{lm} x_l^{(t)}) x_l^{(t)} \\ \Delta w_m^{(t)} &= -\eta \frac{\partial E^{(t)}}{\partial w_m} = \eta (d^{(t)} - y^{(t)}) a'(w_m g_m^{(t)}) g_m^{(t)} \end{aligned} \quad (3.35)$$

με:

$$a'(u) = a(u)(1 - a(u)) \quad (3.36)$$

Από την επιλογή του ρυθμού  $\eta$  εκμάθησης εξαρτάται σε μεγάλο βαθμό η ταχύτητα σύγκλισης του αλγορίθμου, αλλά κυρίως η ικανότητά του για επαρκή μείωση του σφάλματος, δίχως εγκλωβισμό σε κάποιο τοπικό ακρότατο. Προκειμένου να

χρησιμοποιηθούν μεγαλύτερα  $\eta$  εισάγεται επικουρικά ένας όρος “αναθεώρησης με κερτημένη” (*momentum updating*), έστω  $\omega$  που παίρνει τιμές στο  $(0,1]$  κι επιδρά ως εξής:

$$\Delta w_{lm}^{(t)}|_{new} = \eta \left( d^{(t)} - y^{(t)} \right) a' \left( w_m g_m^{(t)} \right) w_m a' \left( w_{lm} x_l^{(t)} \right) x_l^{(t)} + \omega \cdot \Delta w_{lm}^{(t)}|_{old} \quad (3.37)$$

με τιμές του  $\omega$  κοντά στη μονάδα να συνιστώνται.

### 3.2.5. ΣΥΓΚΡΙΣΗ RBFN - MLP

Δεδομένου ότι τόσο τα RBFN όσο και τα MLP είναι μη γραμμικά δίκτυα πρόσθιας τροφοδότησης που χαρακτηρίζονται μάλιστα ως γενικοί προσεγγιστές (*universal approximators*), μπορούν να χρησιμοποιηθούν για τους ίδιους σκοπούς. Τίθενται λοιπόν τα ερωτήματα: α) ποιες είναι οι διαφορές τους και β) πιο από τα δυο είναι καταλληλότερο.

Αναφορικά με το πρώτο ερώτημα μπορούμε να εντοπίσουμε τα εξής [Haykin, 1999]:

- Το RBFN στη βασική του μορφή έχει μόνο μία κρυμμένη στοιβάδα ενώ το MLP μία ή περισσότερες.
- Οι νευρώνες των κρυμμένων στοιβάδων και της εξόδου μοιράζονται στα MLP το ίδιο νευρωνικό μοντέλο σε αντίθεση με τα RBFN, όπου οι νευρώνες αυτοί είναι διαφορετικής φύσης κι εξυπηρετούν διαφορετικούς σκοπούς.
- Η κρυμμένη στοιβάδα στα RBFN είναι μη γραμμική σε αντίθεση με τη στοιβάδα εξόδου που είναι γραμμική. Στους MLP ταξινομητές υποδειγμάτων πάλι, όλες οι στοιβάδες πλην της εισόδου είναι μη γραμμικές με εξαίρεση την περίπτωση προβλημάτων μη γραμμικής παλινδρόμησης, όπου προτιμάται μια γραμμική στοιβάδα εξόδου.
- Το όρισμα της συνάρτησης ενεργοποίησης στα RBFN υπολογίζει την Ευκλείδεια νόρμα, δηλαδή απόσταση, μεταξύ του διανύσματος εισόδου και του κέντρου του κόμβου. Η αντίστοιχη συνάρτηση στα MLP υπολογίζει το εσωτερικό γινόμενο και το διάνυσμα των συναπτικών βαρών.
- Τα MLP δημιουργούν ολικές προσεγγίσεις της εκάστοτε αντιστοίχισης εισόδου-εξόδου. Τα RBFN από την πλευρά τους, εκμεταλλευόμενα εκθετικά μειούμενες, τοπικές, μη γραμμικότητες όπως οι συναρτήσεις Gauss, κατασκευάζουν τοπικές προσεγγίσεις.

Συγκρίνοντας τα δύο είδη δικτύων από άποψη απόδοσης [Shyy et al.] το βασικό συμπέρασμα που προέκυψε είναι το ακόλουθο: Ενώ τα MLP χρειάζονται λιγότερους νευρώνες από τα αντίστοιχα RBFN προκειμένου να πετύχουν ίδια απόδοση, τα RBFN εκπαιδεύονται πιο γρήγορα λόγω της γραμμικής τους φύσης (αποφεύγεται η μη γραμμική παλινδρόμηση). Είναι επομένως πιο αποτελεσματικά στην εκπαίδευση με μεγάλα *training sets*. Παράλληλα ο Haykin [Haykin, 1999]<sup>7</sup>, κάνει ορισμένες συγκριτικές παρατηρήσεις για την απόδοση των δύο τύπων δικτύων σε συγκεκριμένες εφαρμογές.

Η απάντηση στο ερώτημα ποιο είδος δικτύου είναι καταλληλότερο για την δική μας εφαρμογή καθορίζεται σε μεγάλο βαθμό από τη στρατηγική επιλογή μας να υιοθετήσουμε την τακτική της τοπικής προσέγγισης (*local approximation*). Όπως λοιπόν προτείνουν διάφοροι συγγραφείς, η αποδοτικότερη επιλογή δικτύου για τη συγκεκριμένη προσέγγιση φαίνεται να είναι αυτή του RBFN. Ειδικότερα στο [Giannakoglou et al.] προτείνεται η χρήση μικρού αριθμού υποδειγμάτων εκπαίδευσης (γύρω στα 20 υποδείγματα), με σκοπό την δημιουργία μικρών τοπικής εμβέλειας RBFN. Το κέρδος από την προσέγγιση αυτή είναι το αμελητέο κόστος εκπαίδευσης του δικτύου με την ευθεία διαδικασία εκπαίδευσης

<sup>7</sup> Κεφάλαιο 5, σελ 303-304.

να χρησιμοποιείται, λόγω του μικρού μεγέθους του προς αντιστροφή πίνακα (το ζήτημα παρουσιάστηκε στην παράγραφο 3.2.3.).

### 3.3. ΣΥΝΔΥΑΣΜΟΣ RBFN ΚΑΙ DE

Όπως έχει ήδη αναφερθεί η τεχνική που ακολουθήθηκε, κάνει χρήση ANN για μια χαμηλού υπολογιστικού κόστους αρχική εκτίμηση της καταλληλότητας των υποψηφίων λύσεων. Οι περισσότερες υποσχόμενες λύσεις υπόκεινται στη συνέχεια σε αναλυτική αξιολόγηση με τον εκάστοτε επιλύτη. Η διαδικασία συνεχίζεται με το ANN να εκπαιδεύεται ξανά για κάθε νέα γενιά, αντλώντας υποδείγματα από μια βάση δεδομένων που ενημερώνεται διαρκώς με τις καλύτερες των λύσεων [Nikolos, 2004].

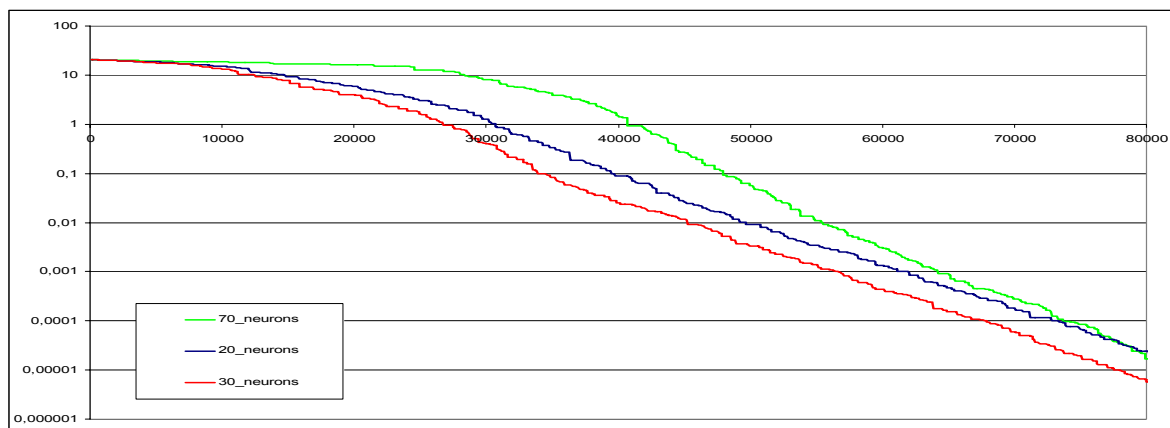
Ακολουθεί περιγραφή της συγκεκριμένης διαδικασίας σε μορφή μετακώδικα:

- Σε κάθε γενιά του αλγορίθμου, ένα μικρό ποσοστό των υποψηφίων λύσεων επιλεγμένο με ομοιόμορφη συνάρτηση πιθανότητας, αξιολογείται αναλυτικά. Οι υπόλοιπες:
  - Κάθε δοκιμαστικό διάνυσμα  $U_{i,G+1}$  αξιολογείται αρχικά με το ANN:
    - Αν εκτιμηθεί ως χειρότερο από το αντίστοιχο μέλος του τρέχοντα πληθυσμού  
Τότε καμία περαιτέρω αξιολόγηση δεν γίνεται: το τρέχον μέλος μεταφέρεται στην επόμενη γενιά και το δοκιμαστικό διάνυσμα εγκαταλείπεται.
    - Αν εκτιμηθεί ως καλύτερο από το αντίστοιχο μέλος του τρέχοντα πληθυσμού  
Τότε πραγματοποιείται μία ακριβής αξιολόγηση του δοκιμαστικού διανύσματος, ακολουθούμενη από νέα σύγκριση των δύο διανυσμάτων:
      - Αν το δοκιμαστικό διάνυσμα εξακολουθεί να είναι καλύτερο του τρέχοντος  
Τότε το δοκιμαστικό διάνυσμα περνά στην επόμενη γενιά.
      - Αλλιώς το τρέχον διάνυσμα είναι αυτό που περνά στην επόμενη γενιά.

Σημειώνεται ότι σύμφωνα με την παραπάνω διαδικασία, μόνο ακριβώς αξιολογημένες υποψήφιες λύσεις περνάνε στην επόμενη γενιά, η οποία ως εκ τούτου αποτελείται από αναλυτικά αξιολογημένα μέλη, γεγονός που ενισχύει την ευστάθεια του αλγορίθμου.

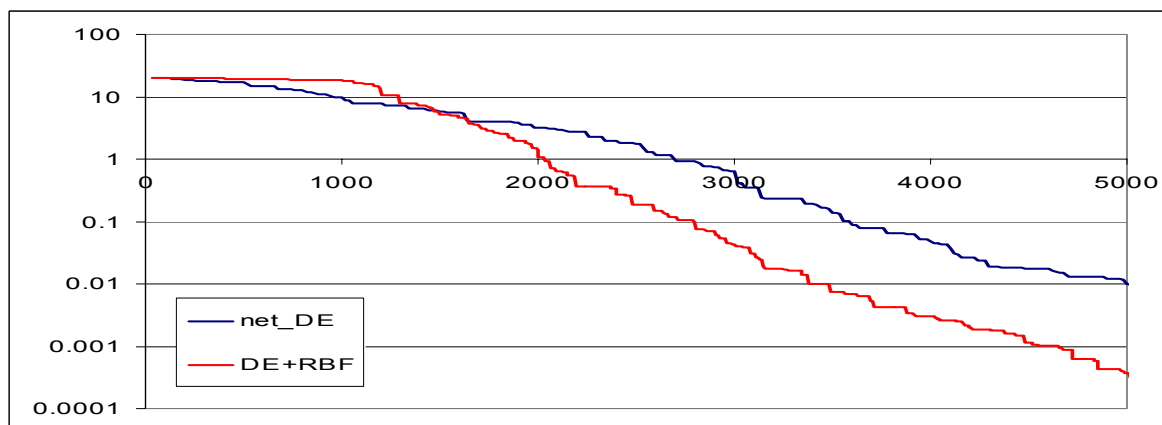
Τα αποτελέσματα της κάθε αξιολόγησης (ακριβούς ή μη) καθώς και το αντίστοιχο χρωμόσωμα, αποθηκεύονται σε μια βάση δεδομένων. Από αυτές οι καλύτερες με βάση την τιμή καταλληλότητάς τους, χρησιμοποιούνται για την επανεκπαίδευση του RBFN. Το αποτέλεσμα είναι να έχουμε ένα μοντέλο τοπικής πρόβλεψης το οποίο εξελίσσεται με τον πληθυσμό του DE από γενιά σε γενιά. Η εξοικονόμηση αναλυτικών αξιολογήσεων (μέσω πχ. CFD) και επομένως χρόνου, δεν γίνεται σε βάρος των καλύτερων λύσεων που αξιολογούνται πάντα αναλυτικά.

Αναφορικά με το μέγεθος του χρησιμοποιούμενου δικτύου, ο χρησιμοποιούμενος αριθμός των κρυμμένων κόμβων κυμαίνεται γύρω στο 30. Η επιλογή αυτή εξυπηρετεί το σκοπό της δημιουργίας ενός τοπικής πρόβλεψης μοντέλου, διατηρώντας παράλληλα το σφάλμα γενίκευσης σε χαμηλά επίπεδα. Ταυτόχρονα και όπως έδειξαν και οι συγγραφείς του [Giannakoglou et al.] μπορεί να επιλεγεί  $M = T$  όπως συζητήθηκε στην Παράγραφο 3.2.3, γεγονός που επιτρέπει την αποτελεσματική υπολογιστικά αντιστροφή του πίνακα  $G$ . Για του λόγου το αληθές παρουσιάζεται ένα ενδεικτικό διάγραμμα σύγκλισης για διαφορετικά μεγέθη δικτύων και τη συνάρτηση του Ackley ( $ff_1$ ) στο Σχήμα 3.13.

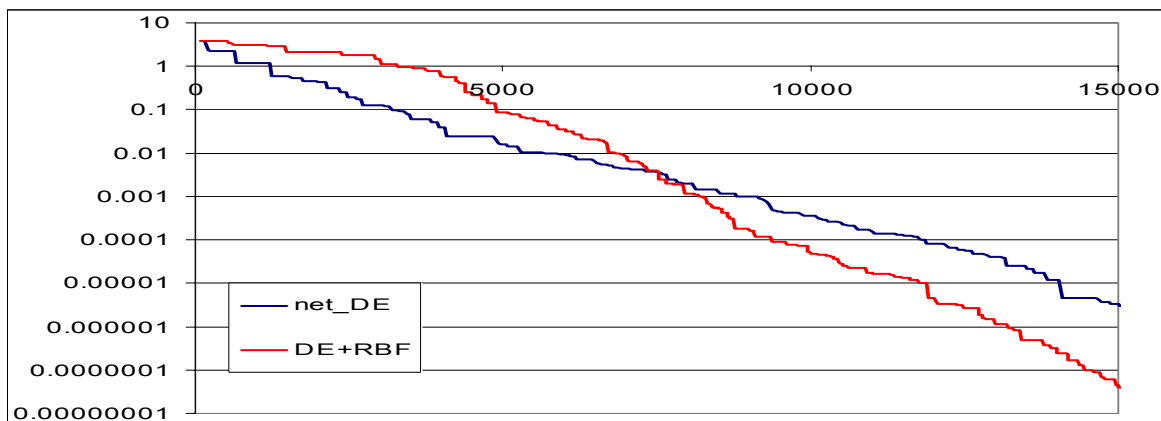


**Σχήμα 3.13.** Η περίπτωση της συνάρτησης του Ackley ( $ff_1$ ) για διάφορα μεγέθη RBFN (20, 30 και 70 νευρώνες).

Προκειμένου να ελέγξουμε την απόδοση του συνδυασμού του νευρωνικού δικτύου με τον Διαφορικό Εξελικτικό αλγόριθμο, πραγματοποιήσαμε μια σειρά δοκιμών, χρησιμοποιώντας τις μαθηματικές συναρτήσεις που εισήγαμε στην ενότητα 2.3.3. Ορισμένα ενδεικτικά διαγράμματα σύγκλισης του τροποποιημένου πλέον αλγορίθμου φαίνονται στα Σχήματα 3.14 ως 3.16. Στον οριζόντιο άξονα δίδεται ο αριθμός των αξιολογήσεων που πραγματοποιήθηκαν με το «ακριβές» μοντέλο.



**Σχήμα 3.14.** Η περίπτωση της συνάρτησης του Ackley ( $ff_1$ ). Η μπλε γραμμή αντιστοιχεί στον DE χωρίς προσεγγιστικό μοντέλο ενώ η κόκκινη στο συνδυασμό του με το RBFN.



Σχήμα 3.15. Η περίπτωση της συνάρτησης σφαίρας ( $ff_0$ ).



Σχήμα 3.16. Η περίπτωση της συνάρτησης του Griewangk ( $ff_4$ ).

Τα ανωτέρω αποτελέσματα επιβεβαιώνουν την αποτελεσματικότητα του συνδυασμού DE με το RBFN, αφού οδηγούν είτε στην επίτευξη ίδιας ποιότητας σύγκλισης για μικρότερο αριθμό αξιολογήσεων της συνάρτησης καταλληλότητας, είτε στην επίτευξη “βαθύτερης” σύγκλισης για ίδιο αριθμό αξιολογήσεων.

# 4

## Η ΕΦΑΡΜΟΓΗ

ΣΤΟ παρόν κεφάλαιο περιγράφεται η τελική εφαρμογή για την επίλυση του προβλήματος βελτιστοποίησης αεροτομών. Αρχικά αποσαφηνίζεται η φύση της προτεινόμενης μεθόδου, βάσει της θεωρίας που έχει αναπτυχθεί στα προηγούμενα κεφάλαια. Εν συνεχεία δίνονται τα βασικά χαρακτηριστικά της προγραμματιστικής υλοποίησης (γλώσσα προγραμματισμού, διεπαφή με το χρήστη, κ.λπ.). Τέλος παρουσιάζονται όλα τα πειραματικά δεδομένα που προέκυψαν κατά την διάρκεια των δοκιμών επίλυσης του προβλήματος.

#### 4.1. Η ΦΥΣΗ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΜΕΘΟΔΟΥ

Έχοντας δει στο 1<sup>ο</sup> Κεφάλαιο μια επισκόπηση των εναλλακτικών μεθόδων βελτιστοποίησης κι έχοντας εν συνεχεία αναλύσει τη θεωρία τόσο των ΕΑ όσο και των ANN, είμαστε πλέον σε θέση να περιγράψουμε την προτεινόμενη εφαρμογή και να αποσαφηνίσουμε τη φύση της. Εξετάζοντας λοιπόν τη μέθοδο από τη σκοπιά της κάθε εμπλεκόμενης θεωρίας, έχουμε:

- Η εδώ προτεινόμενη μέθοδος ανήκει στο 2<sup>ο</sup> ιεραρχικό επίπεδο, σύμφωνα με την ταξινόμηση του *Giles* (Παράγραφος 1.1.1), αφού για δεδομένες συνθήκες, ροής εν προκειμένω, καθορισμένες στο πρώτο ιεραρχικό επίπεδο σχεδιασμού, ζητάμε την βέλτιστη λύση.
- Πρόκειται για Αριθμητική Μέθοδο Βελτιστοποίησης (Παράγραφος 1.1.3), δηλαδή για συνδυασμό μεθόδων ανάλυσης (CFD) και αλγορίθμων αριθμητικής βελτιστοποίησης. Η γεωμετρία παραμετροποιείται όπως θα δούμε, με τη βοήθεια γενικών συναρτήσεων (καμπύλες B-Splines) και οι όροι της αντικειμενικής συνάρτησης, είναι συναρτήσεις της αεροδυναμικής απόδοσης.
- Το πρόβλημα βελτιστοποίησης δομείται ως πρόβλημα ελαχιστοποίησης με περιορισμούς (Παράγραφος 1.2.2).
- Πρόκειται για Μέθοδο Προσέγγισης Συναρτήσεως, όπως αυτές περιγράφονται από τον *Pierret* [Pierret, 1999], δεδομένου ότι συνδυάζει έναν αλγόριθμο βελτιστοποίησης με εργαλεία της τεχνητής νοημοσύνης για την αντιμετώπιση των προβλημάτων ανάλυσης. Η κεντρική ιδέα έγκειται στην δημιουργία προσεγγίσεων των μεθόδων ανάλυσης, οι οποίες να είναι πιο αποδοτικές και να ανακαλύπτουν την “κρυμμένη” συναρτησιακή σχέση.
- Χρησιμοποιείται τοπικού χαρακτήρα προσεγγιστικό μοντέλο μιας και το νευρωνικό δίκτυο που διαδραματίζει το ρόλο αυτό δημιουργείται (ήτοι εκπαιδεύεται) ξεχωριστά για κάθε υποψήφια λύση που αξιολογεί. Αντίθετα ένα ολικό προσεγγιστικό μοντέλο θα συνίστατο από ένα και μόνο νευρωνικό, που εκπαιδεύομενο μια φορά, θα αξιολογούσε ολόκληρο τον πληθυσμό λύσεων.
- Πρόκειται για πιθανοθεωρητικής φύσεως μέθοδο αναζήτησης, που διατηρώντας έναν πληθυσμό ατόμων που αντιπροσωπεύουν σημεία του χώρου αναζήτησης των πιθανών λύσεων περιπλανάται στο χώρο αυτό με τυχαίο τρόπο. Εν τέλει και ύστερα από ικανό αριθμό επαναλήψεων, η διαδικασία συγκλίνει σε μια σχεδόν βέλτιστη λύση.

## 4.2. Η ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΥΛΟΠΟΙΗΣΗ ΤΗΣ ΜΕΘΟΔΟΥ

### 4.2.1. ΤΟ ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΟ ΛΟΓΙΣΜΙΚΟ

Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε είναι η Microsoft Visual Basic 6 (VB6). Η επιλογή αυτή έγινε με γνώμονα την ευκολία με την οποία μπορούν να δημιουργηθούν πολύπλοκες διεπαφές (interfaces) με τη συγκεκριμένη γλώσσα, καθώς και η ευκολία δημιουργίας και διαχείρισης αρθρωτών (modular) εφαρμογών. Η πρώτη ιδιότητα είναι ιδιαίτερα χρήσιμη αν αναλογιστούμε την πληθώρα παραμέτρων της μεθόδου, ενώ η δεύτερη βοηθά τη συνεργασία πολλών ανεξάρτητων τμημάτων κώδικα π.χ. νευρωνικό δίκτυο, DE, γεννήτρια τυχαίων αριθμών, κ.λπ.

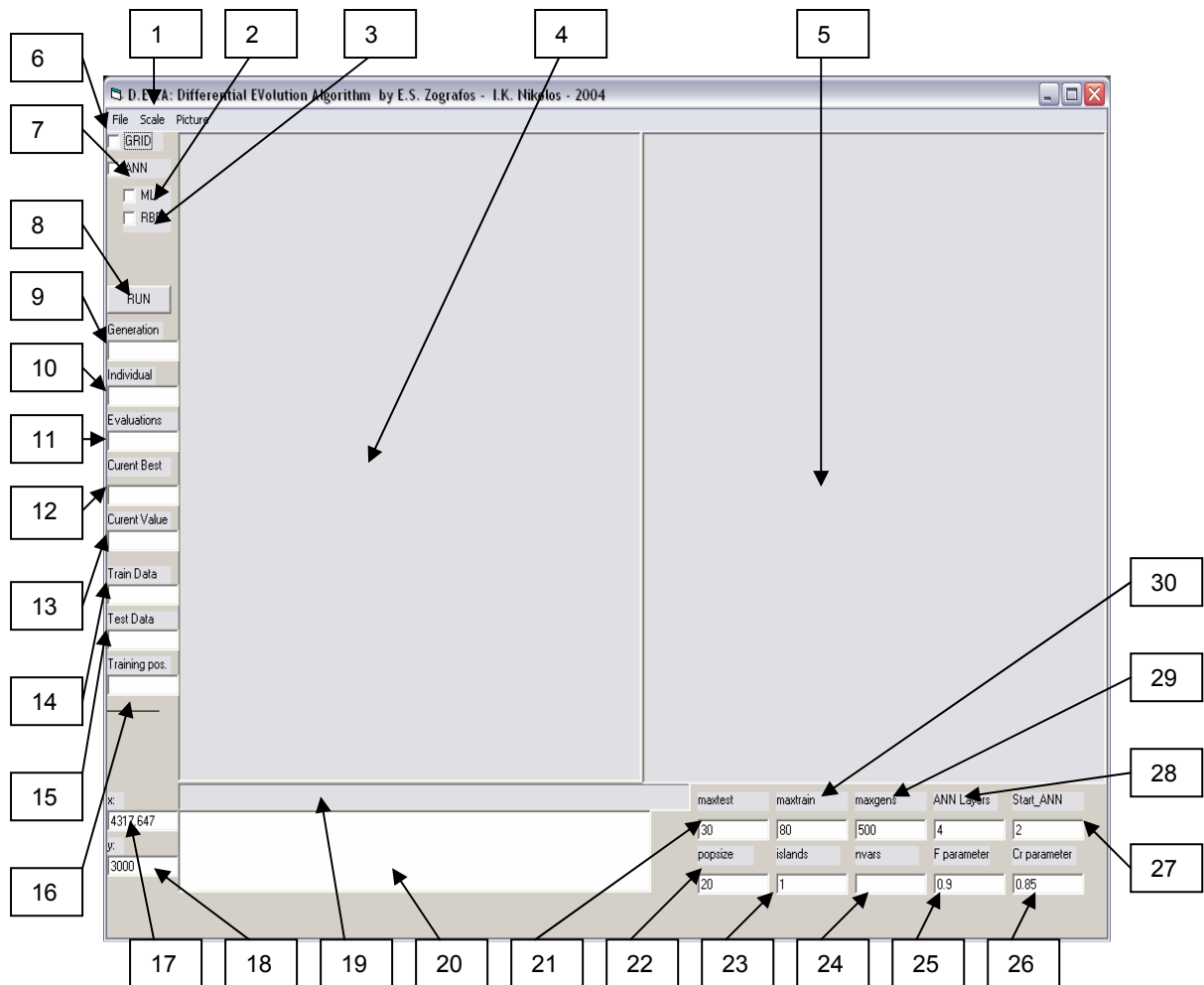
Ένα ιδιαίτερα χρήσιμο χαρακτηριστικό της εφαρμογής αποτελεί η ικανότητα συνδυασμού της με εξωτερικό λογισμικό. Αυτό μπορεί να είναι οποιοδήποτε πρόγραμμα που καλείται να υπολογίσει την εκάστοτε αντικειμενική συνάρτηση του υπό βελτιστοποίηση προβλήματος. Μπορεί επομένως να μιλάμε για λογισμικό που υπολογίζει οικονομικούς δείκτες, φυσικά μεγέθη ή μεταβλητές μηχανολογικών προβλημάτων, ανάλογα με την εφαρμογή.

Το ρόλο του εξωτερικού επιλύτη κλήθηκε εδώ να παίξει ο κώδικας XFOIL του M. Drela, Πρόκειται για λογισμικό υπολογιστικής ρευστομηχανικής το οποίο επιτρέπει την επίλυση της διδιάστατης ροής γύρω από αεροτομές, κάνοντας συνδυασμένη χρήση μιας μεθόδου «panel» για την επίλυση της δυναμικής ροής και μεθόδου επίλυσης διδιάστατων οριακών στρωμάτων. Το λογισμικό διατίθεται δωρεάν στο διαδίκτυο για ακαδημαϊκή χρήση και έχει τη δυνατότητα να λαμβάνει υπόψη τη συμπίεστικότητα της ροής, χωρίς όμως να μπορεί να υπολογίσει κύματα ασυνέχειας. Επιπλέον έχει τη δυνατότητα υπολογισμού της ροής και για χαμηλούς αριθμούς Reynolds [Drela, 1989].

### 4.2.2. Η ΔΙΕΠΑΦΗ ΠΡΟΓΡΑΜΜΑΤΟΣ - ΧΡΗΣΤΗ

Με εξαίρεση κάποια δεδομένα που το πρόγραμμα “διαβάζει” από τα αρχεία που λαμβάνει ως είσοδο, η επικοινωνία με το χρήστη γίνεται μέσω της βασικής φόρμας του προγράμματος. Η τελευταία, που εμφανίζεται στο Σχήμα 4.1, παρέχει στο χρήστη τη δυνατότητα τροποποίησης των παραμέτρων της εφαρμογής. Πιο συγκεκριμένα, μπορούν να μεταβληθούν τόσο οι παράμετροι του Διαφορικού Εξελικτικού αλγορίθμου (μέγεθος πληθυσμού, πλήθος γενεών, κ.λπ.), όσο και εκείνες του νευρωνικού δικτύου (μέγεθος συνόλου εκπαίδευσης και δοκιμής, πλήθος στοιβάδων νευρώνων, κ.λπ.).

Εξίσου σημαντικές είναι και οι πληροφορίες που παρέχονται στο χρήστη σχετικά με την εξέλιξη της διαδικασίας. Αναλυτικότερα εμφανίζονται οι τιμές για την γενιά στην οποία βρίσκεται ο αλγόριθμος, το πλήθος των αξιολογήσεων στις οποίες έχει προβεί, την τρέχουσα και την καλύτερη τιμή της αντικειμενικής συνάρτησης καθώς και την εξέλιξη της εκπαίδευσης του νευρωνικού δικτύου. Τέλος παρουσιάζεται διαγραμματικά η πορεία σύγκλισης του αλγορίθμου. Για την καλύτερη εποπτεία των δυνατοτήτων που παρέχει το interface στο χρήστη, δημιουργήθηκε ο Πίνακας 4.1, όπου εμφανίζονται όλα τα χειριστήρια και οι λειτουργίες που επιτελούν.



**Σχήμα 4.1.** Η φόρμα επικοινωνίας του προγράμματος με το χρήστη.

1	Μπάρα μενού για διαχείριση αρχείων και εικόνων.
2	Check box επιλογής του MLP.
3	Check box επιλογής του RBFN.
4	Παράθυρο αναπαράστασης της σύγκλισης του αλγορίθμου σε δεκαδική κλίμακα.
5	Παράθυρο αναπαράστασης της σύγκλισης του αλγορίθμου σε λογαριθμική κλίμακα.
6	Check box εμφάνισης πλέγματος στα δύο παράθυρα αναπαράστασης σύγκλισης.
7	Check box επιλογής χρήσης ANN σε συνδυασμό με τον DE.
8	Κουμπί εκκίνησης της διαδικασίας.
9	Εμφάνιση του αριθμού της τρέχουσας γενεάς.
10	Εμφάνιση του αριθμού του τρέχοντος αξιολογούμενου ατόμου.
11	Εμφάνιση του πλήθους των πραγματοποιημένων αναλυτικών αξιολογήσεων.
12	Η τιμή της τρέχουσας καλύτερης λύσης.
13	Η τιμή της τρέχουσας αξιολογούμενης λύσης.
14	Το πλήθος των μελών του συνόλου εκπαίδευσης που παρουσιάζεται στο ANN.
15	Το πλήθος των μελών του συνόλου δοκιμής του ANN.
16	Αριθμός επανάληψης κατά τη διαδικασία εκπαίδευσης ενός MLP ANN.
17	Η συντεταγμένη του κέρσορα στον X άξονα του παραθύρου.

18	Η συντεταγμένη του κέρσορα στον Υ άξονα του παραθύρου.
19	Η μπάρα προόδου της διαδικασίας.
20	Παράθυρο εμφάνισης του τρέχοντος χρησιμοποιούμενου εξωτερικού επιλύτη.
21	Μέγιστο επιτρεπόμενο μέγεθος του συνόλου δοκιμής του ANN.
22	Μέγεθος πληθυσμού του DE.
23	Πλήθος “νησιών” <sup>8</sup>
24	Πλήθος μεταβλητών του ανά χείρας προβλήματος.
25	Η τιμή της παραμέτρου F του DE, που καθορίζεται από το χρήστη.
26	Η τιμή της παραμέτρου Cr του DE, που καθορίζεται από το χρήστη.
27	Αριθμός γενεάς του DE κατά την οποία ξεκινά η λειτουργία του ANN.
28	Πλήθος στοιβάδων νευρώνων του MLP.
29	Μέγιστος αριθμός γενεών που διαρκεί η διαδικασία.
30	Μέγιστο επιτρεπόμενο μέγεθος του συνόλου εκπαίδευσης του ANN.

**Πίνακας 4.1.** Οι λειτουργίες της φόρμας και οι πληροφορίες που παρέχει.

### 4.3. Η ΕΠΙΛΥΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

#### 4.3.1. Η ΑΕΡΟΤΟΜΗ ΚΑΙ Η ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΗΣ

Το πρόβλημα που παρουσιάζεται στη συγκεκριμένη ενότητα, δεν είναι άλλο από το ευθύ πρόβλημα βελτιστοποίησης μιας αεροτομής, εκτεθειμένης σε ροή χαμηλού αριθμού Reynolds. Για την γεωμετρική αναπαράσταση της αεροτομής, έγινε χρήση καμπυλών B-Spline, μιας και η συγκεκριμένη μεθοδολογία επιτρέπει τη δημιουργία λεπτομερών γεωμετρικών σχημάτων με ταυτόχρονη ελαχιστοποίηση των απαιτούμενων παραμέτρων καθώς και ελεγχόμενη ομαλότητα της καμπύλης.

Η *παραμετρική σχεδίαση (parametric design)* συνδέεται με τη χρησιμοποίηση λογισμικού CAD για την περιγραφή του αντικειμένου, με τη χρήση συγκεκριμένων αριθμητικών παραμέτρων. Οι γεωμετρικές διαστάσεις του προϊόντος περιγράφονται με συναρτήσεις, οι οποίες τις συνδέουν με τις βασικές παραμέτρους σχεδίασης. Η σπουδαιότητα της παραμετρικής αναπαράστασης γεωμετριών μέσω καμπυλών και επιφανειών Bezier, B-Spline και NURBS, έγκειται στα εξής δύο σημεία:

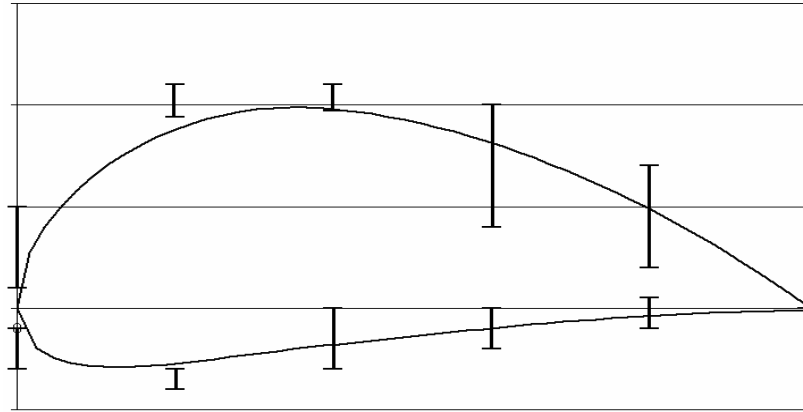
- Παρέχει την ευκολία της ταχύτατης μεταβολής της γεωμετρίας του αντικειμένου με την αλλαγή της τιμής κάποιας ή κάποιων παραμέτρων.
- Η παραμετρική περιγραφή επιτρέπει τη συνεργασία με λογισμικό βελτιστοποίησης για την γρήγορη παραγωγή και αξιολόγηση διαφορετικών εναλλακτικών σχεδιάσεων, με τη χρησιμοποίηση των λιγότερων δυνατών παραμέτρων.

Εν προκειμένω οι καμπύλες B-Spline [Piegl & Tiller, 1997] αποτέλεσαν την επιλογή μας, δεδομένου ότι χρειάζονται μικρό αριθμό παραμέτρων, τις συντεταγμένες των *σημείων ελέγχου (control points)*, προκειμένου να αποδώσουν περίπλοκες καμπύλες. Το κάθε σημείο ελέγχου έχει τοπική επίδραση στο σχήμα της καμπύλης, πράγμα που σημαίνει ότι μικρές μετακινήσεις της θέσης του επιφέρουν αλλαγές στην καμπύλη μόνο στη “γειτονιά” του κινούμενου σημείου ελέγχου.

Στην συγκεκριμένη εφαρμογή το περίγραμμα της αεροτομής αναπαρίσταται με μία και μόνο ανοιχτή καμπύλη B-Spline, με  $2N$  σημεία ελέγχου, όπου  $N$  ο αριθμός των σημείων ελέγχου σε κάθε πλευρά της αεροτομής. Το πρώτο και το τελευταίο σημείο ελέγχου, είναι πακτωμένα στην *ακμή εκφυγής (trailing edge)* της αεροτομής. Οι  $x$

<sup>8</sup> Στην περίπτωση παράλληλης επεξεργασίας σε πολλά CPU, ο πληθυσμός μπορεί να διαχωριστεί και απομονωθεί σε πολλά “νησιά” (islands), δυνατότητα που δεν παρέχεται από τον υπάρχοντα κώδικα.

συντεταγμένες των υπολοίπων σημείων ελέγχου, είναι επίσης αμετάβλητες, γεγονός που μας αφήνει τελικά με  $2N-2$  ελεύθερες παραμέτρους, που αντιστοιχούν στις  $y$  συντεταγμένες των μη σταθερών σημείων ελέγχου. Τελικά σε όλα τα σημεία ελέγχου εκτός των σταθερών, επιτρέπεται η κίνηση κατά μήκος κάθετων διαστημάτων, τα όρια των οποίων εμφανίζονται στο Σχήμα 4.2.



**Σχήμα 4.2.** Τα όρια κίνησης των  $y$  συντεταγμένων των μη σταθερών σημείων ελέγχου.

Για τις ανάγκες του επιλυόμενου προβλήματος, κάθε αεροτομή περιγράφεται από 12 σημεία ελέγχου κι επομένως 10 μεταβλητές σχεδίασης, που αποτελούν και το χρωμόσωμα της κάθε υποψήφιας λύσης - αεροτομής.

#### 4.3.2. ΟΡΙΣΜΟΣ ΚΑΙ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΠΡΟΒΛΗΜΑΤΟΣ

Το πρόβλημα ορίζεται ως η βελτιστοποίηση μορφής μιας υπάρχουσας αεροτομής, προκειμένου να μεγιστοποιηθεί η άνωσή της (lift) και να ελαχιστοποιηθεί η αντίστασή της, υπό προκαθορισμένες συνθήκες λειτουργίας. Η αεροτομή αναφοράς είναι η γνωστή NACA 4412 (Σχήμα 4.2). Οι περιπτώσεις που εξετάστηκαν είναι δύο και περιγράφονται από τις ακόλουθες συνθήκες ροής:

- $\alpha_\infty = 3^\circ$ ,  $Re = 480000$ ,  $M_\infty = 0.2$
- $\alpha_\infty = 2^\circ$ ,  $Re = 480000$ ,  $M_\infty = 0.2$

όπου,  $\alpha_\infty$  είναι η γωνία ροής του ελεύθερου ρεύματος,  $Re$  είναι ο αριθμός Reynolds του ελεύθερου ρεύματος βασισμένος στο μήκος χορδής και  $M_\infty$  είναι ο αριθμός Mach του ελεύθερου ρεύματος. Οι συνθήκες ροής αντιστοιχούν σε εκείνες που αντιμετωπίζει ένα μικρό UAV στη φάση απογείωσής του.

Προκειμένου να αρχικοποιηθεί η διαδικασία βελτιστοποίησης, τα χαμηλότερα και ανώτερα όρια κάθε μεταβλητής σχεδίασης ( $y$  συντεταγμένη του αντίστοιχου σημείου ελέγχου) καθορίστηκαν, διαμορφώνοντας ένα διάστημα για κάθε μεταβλητή. Για να ικανοποιηθεί η απαίτηση το μέγιστο πάχος της βελτιστοποιημένης αεροτομής να είναι όσο το δυνατό πιο κοντά σε αυτό της αρχικής, τα διαστήματα που αντιστοιχούν στα σημεία ελέγχου κοντά στην περιοχή μέγιστου πάχους, καθορίστηκαν με το εσωτερικό όριό τους πολύ κοντά στα αντίστοιχα σημεία ελέγχου της αεροτομής αναφοράς (Σχήμα 4.2).

Το πρόβλημα ορίζεται ως πρόβλημα ελαχιστοποίησης και η αντικειμενική συνάρτηση λαμβάνεται ως εξής:

$$f = C_D + \frac{\beta}{C_L} \quad (4.1)$$

προκειμένου να ελαχιστοποιηθεί η αντίσταση και να μεγιστοποιηθεί η άνωση. Η παράμετρος  $\beta$  χρησιμοποιείται προκειμένου να καθοριστεί η αναλογικά η σημασία της αντίστασης και της άνωσης στη διαδικασία βελτιστοποίησης. Με την επιλογή διαφορετικών τιμών, ο μονού-στόχου ΕΑ (single-objective EA) μπορεί να χρησιμοποιηθεί για να λύσει το δύο-στόχων πρόβλημα βελτιστοποίησης και για να παράξει το αντίστοιχο POF. Στην περίπτωση μας χρησιμοποιήθηκαν οι ακόλουθες οκτώ τιμές του  $\beta$ : 0.001, 0.002, 0.003, 0.005, 0,01, 0,1, 1 και 10.

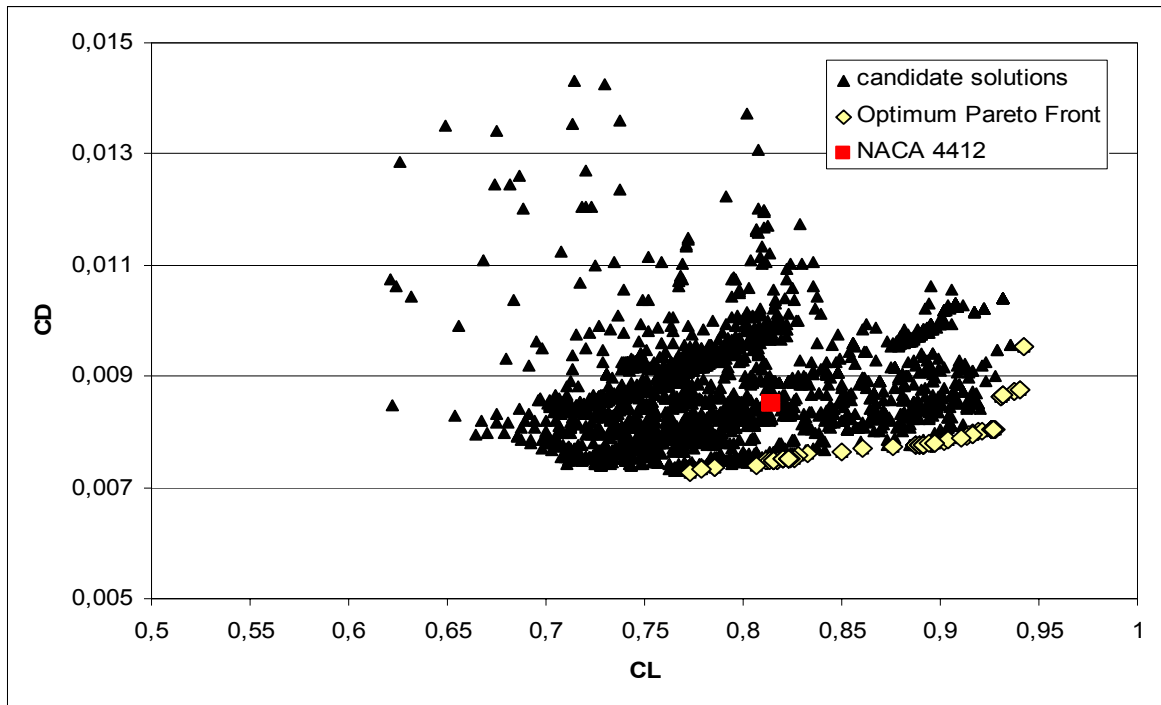
#### 4.3.3. ΔΟΚΙΜΕΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Για κάθε τιμή του  $\beta$  ο αλγόριθμος βελτιστοποίησης έτρεξε με και χωρίς το RBFN. Οι υπολογισμένοι συντελεστές αντίστασης και άνωσης παρατίθενται στον Πίνακα 4.2 για γωνία προσβολής  $3^\circ$ . Τα τρεξίματα χωρίς το RBFN εκτελέστηκαν με ένα μέγεθος πληθυσμού ίσο με 20 και για 250 γενεές, ενώ τα τρεξίματα με το RBFN εκτελέστηκαν με το ίδιο μέγεθος πληθυσμού αλλά για 500 γενεές. Για όλες τις δοκιμές οι συντελεστές  $F$  και  $Cr$  ήταν ίσα με 0,9 και 0,85 αντίστοιχα. Το μέτωπο Pareto, καθώς και όλες οι ενδιάμεσες λύσεις, όπως προέκυψαν από τον πολυκριτήριο ΕΑ, δίδονται στο Σχήμα 4.3. Στο Σχήμα 4.4 συγκρίνονται τα μέτωπα Pareto που προέκυψαν με τον πολυκριτήριο ΕΑ και με την παρούσα μεθοδολογία, ενώ δίδεται και η θέση που αντιστοιχείς την αεροτομή αναφοράς NACA 4412.

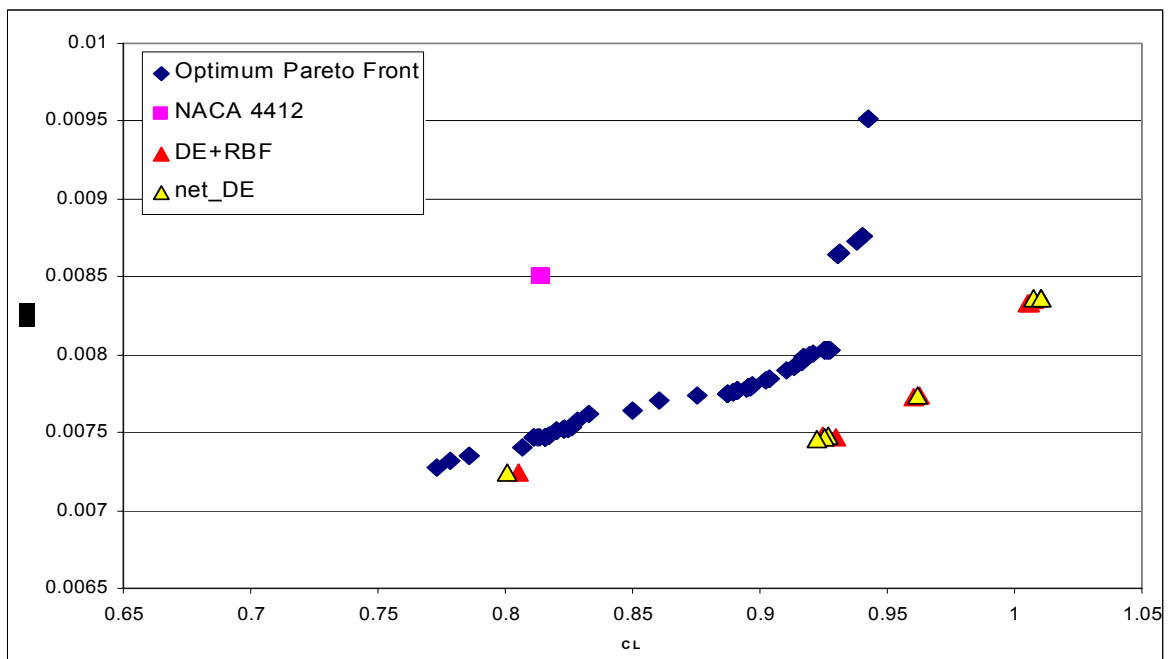
$\alpha_\infty = 3^\circ$				
$\beta$	No ANN		With RBFN	
	$C_L$	$C_D$	$C_L$	$C_D$
0.001	0.7798	0.00721	0.8192	0.00728
0.002	0.9219	0.00746	0.9204	0.00748
0.003	0.9387	0.00754	0.922	0.00746
0.005	0.9401	0.00755	0.9618	0.00774
0.01	1.036	0.0081	1.038	0.0081
0.1	1.0772	0.0086	1.077	0.0086
1	1.0739	0.00862	1.0749	0.00857
10	1.0754	0.00868	1.0357	0.00819

**Πίνακας 4.2.** Αποτελέσματα υπολογισμού των  $C_L$ ,  $C_D$  για τις δύο εξεταζόμενες περιπτώσεις για γωνία προσβολής  $3^\circ$ .

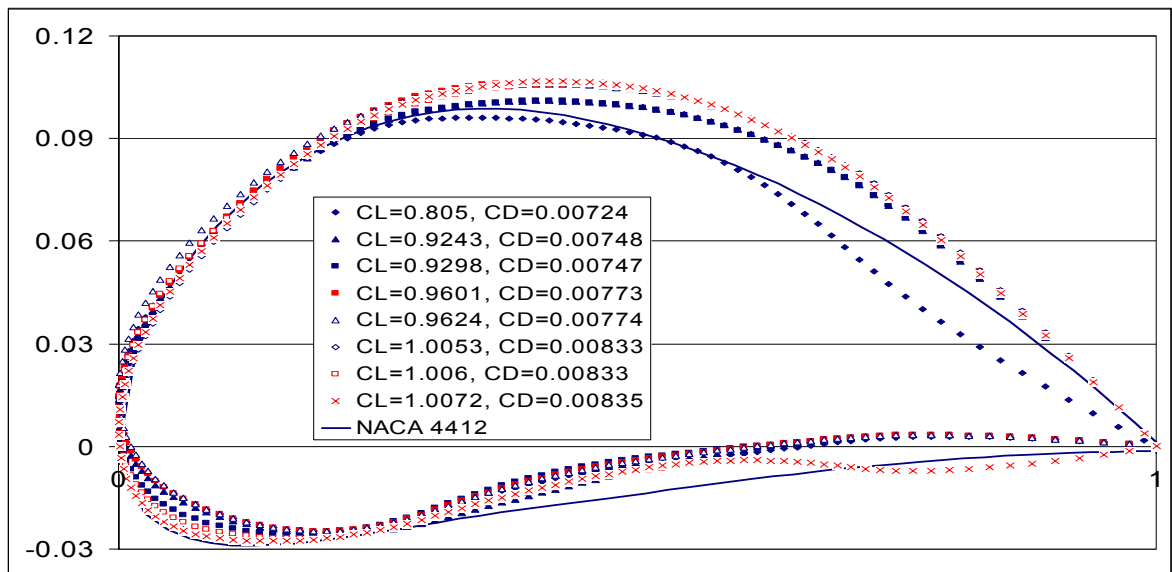
Τέλος στο Σχήμα 4.5, εμφανίζονται οι διάφορες γεωμετρίες αεροτομών που αντιστοιχούν στις λύσεις που ανήκουν στο μέτωπο Pareto, που υπολογίστηκαν χρησιμοποιώντας τον ΔΕ. Εργαζόμενοι με πανομοιότυπο τρόπο παίρνουμε το αντίστοιχο μέτωπο Pareto για την περίπτωση όπου  $\alpha_\infty = 2^\circ$  και το οποίο εικονίζεται στο Σχήμα 4.6, ενώ οι γεωμετρίες αεροτομών που αντιστοιχούν στις λύσεις, φαίνονται στο Σχήμα 4.7.



**Σχήμα 4.3.** Το πλήθος υποψήφιων λύσεων του multi-objective EA και το αντίστοιχο βέλτιστο κατά Pareto μέτωπο, όπως παρήχθη με τη χρήση Πολυκριτηρίου EA [Χριστόπουλος, 2003] (γωνία προσβολής 3°).

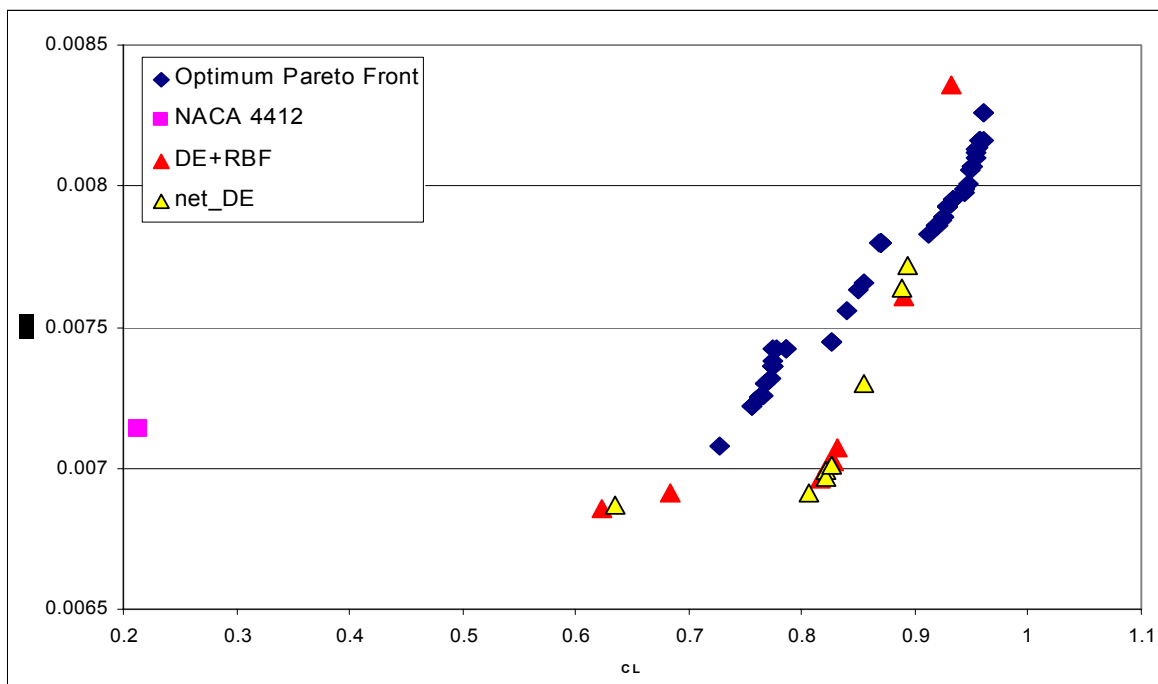


**Σχήμα 4.4.** Το βέλτιστο κατά Pareto μέτωπο του multi-objective EA και το αντίστοιχο βέλτιστο κατά Pareto μέτωπο του single-objective EA με και χωρίς το RBFN (γωνία προσβολής 3°).

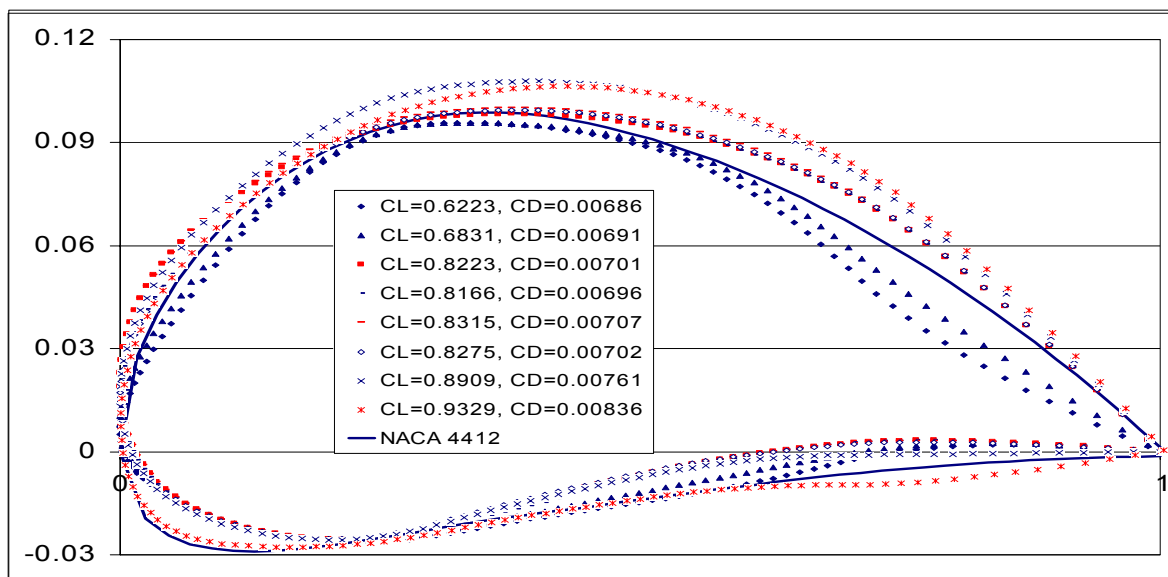


**Σχήμα 4.5.** Οι διάφορες γεωμετρίες αεροτομών που αντιστοιχούν στις λύσεις που ανήκουν στο μέτωπο Pareto, που υπολογίστηκαν χρησιμοποιώντας τον single-objective DE (γωνία προσβολής  $3^\circ$ ).

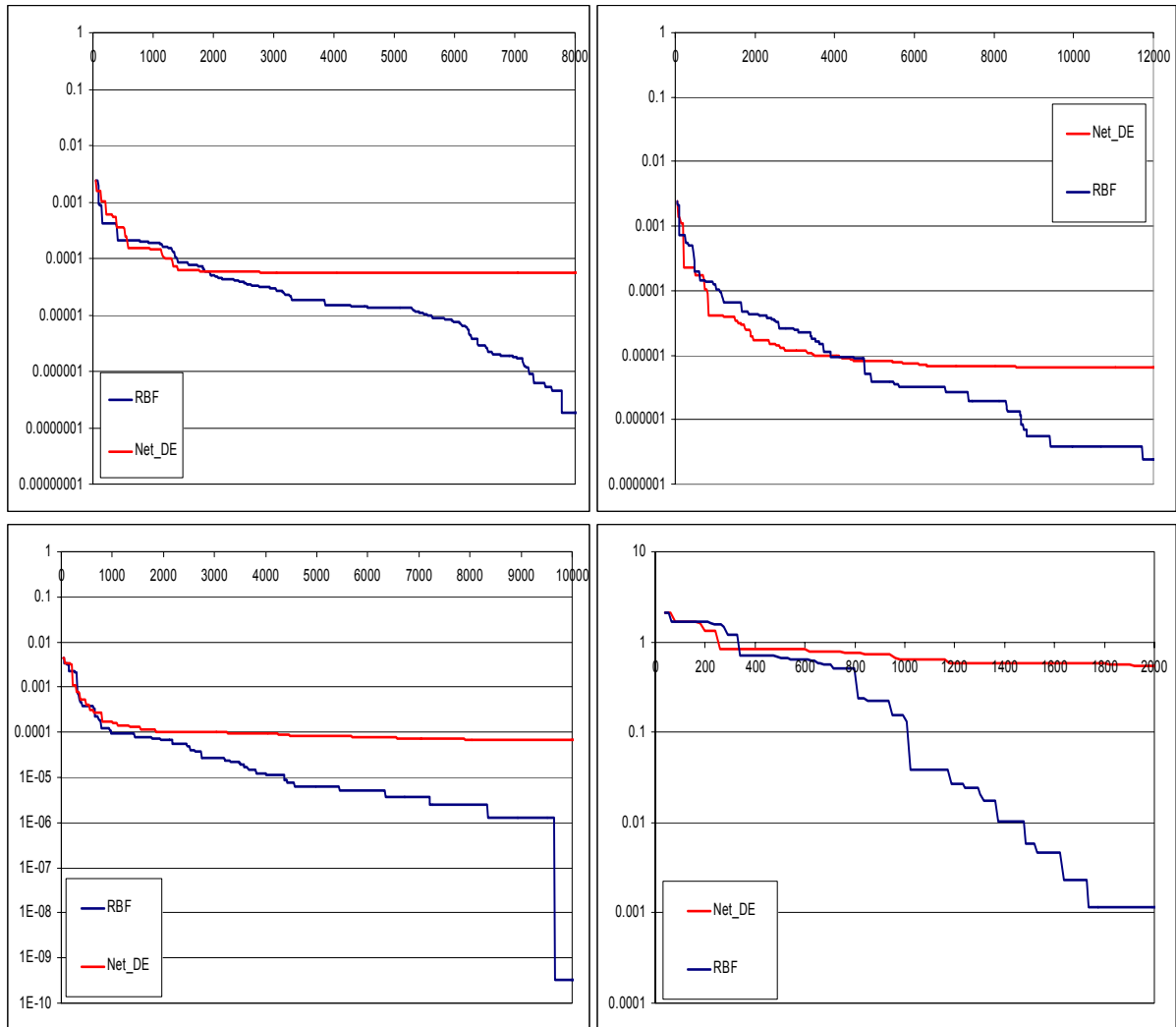
Τα αποτελέσματα του υπολογισμού, με ή χωρίς τη χρήση του μοντέλου προσέγγισης, παρουσίασαν σημαντική βελτίωση σε σχέση με την αεροτομή αναφοράς, για όλες τις τιμές. Βελτίωση επιτεύχθηκε επίσης, έναντι του μετώπου Pareto που παράγεται από πολλαπλών-στόχων ΕΑ, για τα ίδια όρια των παραμέτρων ελέγχου. Εντούτοις, το μέτωπο Pareto που υπολογίστηκε με τον πολλαπλών-στόχων ΕΑ αλγόριθμο, χρειάστηκε πολύ λιγότερες αξιολογήσεις (3000), έναντι στις αξιολογήσεις που απαιτούνται για τα 8 διαφορετικά τρεξίματα του μονού-στόχου αλγορίθμου (8x5000 για το απλό DE αλγόριθμο και σχεδόν ίδιος αριθμός για τον αλγόριθμο με το RBFN). Ορισμένα ενδεικτικά διαγράμματα για την πορεία σύγκλισης του αλγορίθμου κατά την εκτέλεση των δοκιμών, παρατίθενται στο Σχήμα 4.8



**Σχήμα 4.6.** Το βέλτιστο κατά Pareto μέτωπο του multi-objective EA και το αντίστοιχο βέλτιστο κατά Pareto μέτωπο του single-objective EA με και χωρίς το RBFN (γωνία προσβολής 2°).



**Σχήμα 4.7.** Οι διάφορες γεωμετρίες αεροτομών που αντιστοιχούν στις λύσεις που ανήκουν στο μέτωπο Pareto, που υπολογίστηκαν χρησιμοποιώντας τον single-objective DE (γωνία προσβολής 2°).



**Σχήμα 4.8.** Ενδεικτικά διαγράμματα σύγκλισης του single-objective EA με και χωρίς τη χρήση του RBFN.

# 5

## ΣΥΜΠΕΡΑΣΜΑΤΑ

ΣΤΟΝ επίλογο πλέον αυτής της εργασίας, αποδελτιώνονται τα συμπεράσματα που προέκυψαν κατά την εκπόνησή της, και επιπλέον προτείνονται μελλοντικές ερευνητικές κατευθύνσεις που αξίζει κατά τη γνώμη μας να διερευνηθούν.

## 5.1. ΣΥΜΠΕΡΑΣΜΑΤΑ

Μέσα από την πληθώρα δοκιμών που πραγματοποιήθηκαν τόσο στις μαθηματικές συναρτήσεις που χρησιμοποιήθηκαν, όσο και για την επίλυση του προβλήματος της βελτιστοποίησης αεροτομών, μπορούμε να πούμε πως δικαιώθηκαν οι θεωρητικές θέσεις που αποτέλεσαν οδηγό για την ανά χείρας εργασία. Πιο συγκεκριμένα:

- Ο Διαφορικός Εξελικτικός Αλγόριθμος δικαίωσε την φήμη των Εξελικτικών Αλγορίθμων γενικότερα, αποδεικνύοντας πως τα καταφέρνει αρκετά καλά, απέναντι σε μια πληθώρα ετερόκλητων προβλημάτων.
- Αναφορικά με τη σύγκριση DE και GA, μπορούμε να επισημάνουμε ότι στη συντριπτική πλειοψηφία των περιπτώσεων που εξετάστηκαν ο DE αποδείχθηκε αποτελεσματικότερος του GA, ιδιαίτερα για προβλήματα με μεγάλο πλήθος ανεξάρτητων μεταβλητών, οπότε και η ικανότητα σύγκλισης του τελευταίου είναι ιδιαίτερα περιορισμένη.
- Για μικρό και μέσο πλήθος μεταβλητών σχεδίασης (30 ως 60) ο DE τα καταφέρνει καλύτερα και από τον Multi-Deme GA. Σταδιακά ωστόσο και όσο ο αριθμός των ανεξάρτητων μεταβλητών αυξάνει, η κατάσταση αντιστρέφεται.
- Ο GA χρειάζεται σε κάθε περίπτωση μεγαλύτερα μεγέθη πληθυσμού από τον DE, που αρκείται σε σχετικά μικρό μέγεθος πληθυσμού συνήθως ελαφρά μεγαλύτερο ή ακόμη και ίσο του αριθμού των μεταβλητών σχεδίασης για την επίτευξη της καλύτερης απόδοσής του.
- Χαρακτηριστική τέλος είναι σε κάθε περίπτωση η ομαλότερη πορεία σύγκλισης του DE έναντι των GA, κάτι που οφείλεται στον τρόπο που κινείται στο χώρο των λύσεων αλλά και ικανότητα διατήρησης ενός ικανοποιητικού ρυθμού σύγκλισης καθώς αυξάνει το πλήθος των ανεξάρτητων μεταβλητών.
- Τα αποτελέσματα των δοκιμών επιβεβαιώνουν την αποτελεσματικότητα του συνδυασμού DE με το RBFN, αφού οδηγούν είτε στην επίτευξη ίδιας ποιότητας σύγκλισης για μικρότερο αριθμό αξιολογήσεων της συνάρτησης καταλληλότητας, είτε στην επίτευξη “βαθύτερης” σύγκλισης για ίδιο αριθμό αξιολογήσεων
- Η εφαρμογή στη βελτιστοποίηση αεροτομών κατέδειξε ότι ένας μονού-στόχου ΕΑ με κατάλληλη μορφοποίηση του επιλυόμενου προβλήματος, μπορεί να δώσει συγκρίσιμα αποτελέσματα με αυτά ενός πολλαπλών-στόχων ΕΑ, με τα ανταλλάγματα μεταξύ ποιότητας λύσεως και υπολογιστικού κόστους να παραμένουν στη διακριτική ευχέρεια του μηχανικού.

## 5.2. ΠΡΟΤΑΣΕΙΣ

Πέρα από τα όποια συμπεράσματα προέκυψαν από την ενασχόληση με την βελτιστοποίηση με χρήση Εξελικτικών Αλγορίθμων και Νευρωνικών Δικτύων και τις απορίες που λύθηκαν κατά τη διάρκειά της, κάποιες νέες δημιουργήθηκαν και εντοπίστηκαν μια σειρά ζητημάτων που θα μπορούσαν να αποτελέσουν αντικείμενο μελλοντικής μελέτης. Τα σπουδαιότερα από αυτά κατά την γνώμη του γράφοντα είναι τα ακόλουθα:

- Όσον αφορά τον DE αλγόριθμο, ιδιαίτερο ενδιαφέρον θα είχε ενδεχόμενη προσπάθεια παραλληλοποίησής του, με σκοπό την ταχύτερη διεξαγωγή των δοκιμών και της επίλυσης των προβλημάτων με ταυτόχρονη χρήση πολλών υπολογιστών. Εξάλλου το χρονοβόρο της διαδικασίας αποτελεί το μεγαλύτερο ίσως μειονέκτημα κατά την πρακτική καθημερινή εφαρμογή της μεθόδου.
- Κρίνοντας από τις καλές επιδόσεις του Multi-Deme GA αλγορίθμου, αλλά και από τη σαφή ανωτερότητα του DE απέναντι στον κλασσικό GA, η διερεύνηση των δυνατοτήτων ενός Multi-Deme DE αλγορίθμου κρίνεται εξόχως ενδιαφέρουσα.
- Αναφορικά με τη χρήση του RBFN θα μπορούσαν να πραγματοποιηθούν περισσότερες συγκριτικές δοκιμές για τις διάφορες προτεινόμενες συναρτήσεις ενεργοποίησης [Giannakoglou, 2001], με σκοπό την απόκτηση εμπειρίας στην επιλογή της κατά περίπτωση καταλληλότερης.
- Η υιοθέτηση τεχνικών clustering για την επιλογή των κέντρων των συναρτήσεων ακτινικής βάσης [Haykin, 1999], αποτελεί ενδιαφέρουσα επέκταση της παρούσας εργασίας. Με τον τρόπο αυτό βελτιώνεται η απόδοση των μεγάλου μεγέθους δικτύων που απαιτούνται, καθώς η διάσταση των προβλημάτων βελτιστοποίησης (πλήθος ανεξάρτητων μεταβλητών) αυξάνει.
- Η εφαρμογή κάποιας μεθοδολογίας για τον έλεγχο της ποιότητας της απόδοσης του προσεγγιστικού μοντέλου είναι ένα χρήσιμο επόμενο βήμα, μιας και ένα κακής απόδοσης μοντέλο μπορεί να δρα παραπλανητικά για τον ΕΑ. Αν και ήδη λάβαμε σχετική μέριμνα αξιολογώντας ακριβώς ένα ποσοστό του πληθυσμού ανά γενιά (το 10%), κάποια περισσότερο ολοκληρωμένη προσέγγιση όπως αυτή των [Jin et al.] θα μπορούσε να υιοθετηθεί.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- [Andersson, 2001] Andersson J., *Multiobjective Optimization in Engineering Design; Applications to Fluid Power Systems*, Ph.D. Thesis, Division of Fluid and Mechanical Engineering Systems Department of Mechanical Engineering Linköpings universitet, 2001.
- [Arora, 1989] Arora J.S., *Introduction to Optimum Design*, McGraw-Hill, 1989.
- [Buche, Schraudolph & Koumoutsakos, 2005] Buche D., Schraudolph N.N., Koumoutsakos P., Accelerating Evolutionary Algorithms with Gaussian Process Fitness Function Models, *IEEE Trans. on Systems, Man and Cybernetics.*, Vol. 35, No. 2, pp. 183–194, May 2005.
- [Callan, 1999] Callan R. *The essence of Neural Networks*. Prentice Hall Europe, 1999.
- [Cover, 1965] Cover T.M., Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, *IEEE Transactions on Electronic Computers*, vol. EC-14, pp. 326-334, 1965.
- [Drela, 1989] Drela M., XFOIL: An Analysis and Design System for Low Reynolds Number Airfoils, *Conference on Low Reynolds Number Airfoil Aerodynamics*, University of Notre Dame, (1989).
- [Evans & Fisher, 1994] Evans B., Fisher D., Overcoming process delays with decision tree induction, *Expert, IEEE* Vol. 9, issue 1, pp. 60–66, February 1994.
- [Fan, Lampinen & Dulikravich, 2003] Hui-Yuan Fan, Jouni Lampinen and George S. Dulikravich. Improvements to mutation donor formulation of Differential Evolution, *Proceedings of the International Congress on Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, EUROGEN 2003*, 2003.
- [Farin, 1988] Farin G., *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press Inc., 1988.
- [Fogel, Owens & Walsh, 1966] Fogel L.J., Owens A.J., and Walsh M.J., *Artificial Intelligence through Simulated Evolution*, New York: John Wiley, 1966
- [Giannakoglou & Giotis, 2001] Giannakoglou K.C., Giotis A.P., Low-Cost GAs Assisted by Artificial Neural Networks (Κεφάλαιο από το βιβλίο που επρόκειτο να εκδοθεί).
- [Giannakoglou et al.] Giannakoglou K, Giotis A, Karakasis M., Low-cost genetic optimization based on inexact pre-evaluations and the sensitivity analysis of design parameters. *J Inverse Problems Eng.*, Vol. 9, pp. 389-412, 2001.
- [Giannakoglou, 2002] Giannakoglou K.C., Design of optimal aerodynamic shapes using stochastic optimization methods and computational intelligence, *Progress in Aerospace Sciences*, Vol. 38, pp. 43-76, 2002.

- [Giles, 1997] M.B. Giles, Aerospace design: a complex task, in: *Von Karman Institute for Fluid Dynamics, Lecture Series 1997-05; Inverse design and optimisation methods*, 1997.
- [Goldberg, 1989] Goldberg D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley Longman, Inc, 1989.
- [Haykin, 1999] Haykin, S., *Neural Networks, A Comprehensive Foundation*. Second Edition, Prentice Hall, 1999.
- [Holland, 1992] Holland J. H., *Adaption in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*, MIT Press, 1992
- [Jin et al.] Jin Y., Olhofer M., and Sendhoff B., A Framework for Evolutionary Optimization with Approximate Fitness Functions, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 5, pp. 481-494, 2002.
- [Κουϊκόγλου, 2002] Κουϊκόγλου Β. *Προγραμματισμός Παραγωγής*, Πανεπιστημιακές Σημειώσεις, Χανιά, 2002.
- [Koza, 1992] Koza J.R., *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, The MIT Press, 1992.
- [Martin & Simpson, 2003] Martin J. D., Simpson T. W., A Study on the Use of Kriging Models to Approximate Deterministic Computer Models, *Proceedings of DETC'03 ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Chicago, Illinois USA, September 2-6, 2003.
- [Michalewicz, 1999] Michalewicz Z. *Genetic Algorithms + Data Structures = Evolution Programs*, Third, Revised and Extended Edition, Springer, 1999.
- [Mitchell, 1996] Mitchell M., *An Introduction to Genetic Algorithms*, MIT Press, 1996
- [Nikolos, 2004] Nikolos, I. K., Inverse Design of Aerodynamic Shapes using Differential Evolution coupled with Artificial Neural Network, *Proceedings of the ERCOFTAC Conference in Design Optimization: Methods and Applications*, Athens, Mar. 31-Apr. 2, 2004.
- [Παπαγεωργίου, 1999] Παπαγεωργίου Μ. *Μη γραμμικός Προγραμματισμός*, Πανεπιστημιακές Σημειώσεις, Χανιά, 1999.
- [Patel, Honavar & Balakrishnan] Patel M., Honavar V. & Balakrishnan K. (Eds.), *Advances in the Evolutionary Synthesis of Intelligent Agents*, The MIT Press, 2001.
- [Piegl & Tiller, 1997] Piegl L., Tiller W., *The NURBS Book*. Springer, 1997.
- [Pierret, 1999] Pierret S., *Designing turbomachinery blades by means of the function approximation concept based on Artificial Neural Networks, Genetic Algorithms, and the Navier-Stokes equations*, Ph.D. Thesis, Von Karman Institute for Fluid Dynamics, 1999.
- [Press et all, 1992] Press. W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., *Numerical Recipes in C*, Cambridge University Press, 1992.

- [Price, Storn & Lampinen] Price K.V., Storn R.M., Lampinen J.A., *Differential Evolution. A Practical Approach to Global Optimization*, Springer, 2005.
- [Rechenberg, 1973] Rechenberg I., *Evolutionstrategie*, Stuttgart: Frommann-Holzboog, 1973
- [Schmitz, 1999] Schmitz. G.P.J., *Combinatorial Evolution of Feedforward Neural Network Models for Chemical Processes*, Ph.D. Thesis, University of Stellenbosch, 1999.
- [Shyy et al.] Shyy W., Papila N., Vaidyanathan R., Tucker K., Global design optimization for aerodynamics and rocket propulsion components. *Progress in Aerospace Science*, Vol. 37, pp. 59-118, 2001.
- [Storn & Price, 1995] Storn. R., Price K.. Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Space, *Technical Report TR-95-012, ICSI*, 1995.
- [Torczon & Trosset, 1998] Torczon V., Trosset M.W. Using Approximations to Accelerate Engineering Design Optimization, *NASA/CR-1998-208460, ICASE Report No. 98-33*, 1998.
- [Vapnik, 1998] Vapnik V., *Statistical Learning Theory*, Wiley, 1998.
- [Χριστόπουλος, 2003], *Βελτιστοποίηση της γεωμετρίας στροβίλου αξονικής ροής με χρήση Εξελικτικών Αλγορίθμων*, Μεταπτυχιακή Διατριβή, Πολυτεχνείο Κρήτης, Τμήμα Μηχανικών Παραγωγής και Διοίκησης, 2003.

# ΠΑΡΑΡΤΗΜΑ

ΣΤΟ παρόν παράρτημα περιέχονται τα δεδομένα των δοκιμών που πραγματοποιήθηκαν με σκοπό την εξαγωγή συμπερασμάτων για την αποτελεσματικότητα διαφόρων εναλλακτικών μεθόδων βελτιστοποίησης και πιο συγκεκριμένα του DE αλγόριθμου, ενός κλασσικού GA όσο και ενός *Πολλαπλών Υποπληθυσμών Γενετικού Αλγόριθμου (Multi-Deme GA)* καθώς και τα συμπεράσματα που εξήχθησαν.

Για κάθε μία από τις μαθηματικές συναρτήσεις της ενότητας 2.3.3. χρησιμοποιήθηκαν διάφοροι συνδυασμοί πλήθους μεταβλητών σχεδίασης και μεγέθους πληθυσμού, φροντίζοντας ωστόσο οι συγκρίσεις να γίνονται για ίδιο αριθμό αξιολογήσεων της μαθηματικής συνάρτησης. Για τη διευκόλυνση του αναγνώστη υπενθυμίζουμε ότι οι τρεις χρησιμοποιούμενες συναρτήσεις είναι οι:

- Η συνάρτηση του *Ackley* στην γενικευμένη εκδοχή της από τον *Schwefel*:

$$ff_1(\vec{x}) = -a \cdot \exp\left(-b \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(c \cdot x_i)\right) + a + e \quad ; \quad -32.768 \leq x_i \leq 32.768$$

$$a = 20 \quad ; \quad b = 0.2 \quad ; \quad c = 2\pi \quad ; \quad e = 2.71828 \quad ; \quad \min(f_1) = f_1(0, \dots, 0) = 0$$

- Η γενικευμένη συνάρτηση της σφαίρας:

$$ff_0(x) = \sum_{i=1}^n x_i^2$$

- Η συνάρτηση του *Griewangk*:

$$ff_4(x) = \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i+1}}\right) + 1$$

Στις σελίδες που ακολουθούν παρουσιάζονται αρχικά πίνακες που περιέχουν για κάθε μια συνάρτηση δεδομένα όπως ο χρησιμοποιούμενος κάθε φορά αλγόριθμος, το πλήθος των ανεξάρτητων μεταβλητών (nvars), το μέγεθος του χρησιμοποιούμενου πληθυσμού (popsize), ο μέγιστος αριθμός γενεών για τις οποίες επαναλαμβάνεται η διαδικασία (maxgens), η τιμή της καταλληλότητας που προκύπτει για κάθε δοκιμή (fitness) και τέλος ο συνολικός αριθμός αξιολογήσεων της συνάρτησης που πραγματοποιήθηκαν (evaluations).

Εν συνεχεία παρουσιάζονται τα καλύτερα διαγράμματα σύγκλισης που πέτυχαν οι τρεις συγκρινόμενοι αλγόριθμοι σε κάθε περίπτωση (Σχήματα 1 ως 3). Για την καλύτερη κατανόησή τους διευκρινίζεται ότι:

- Οι τίτλοι των διαγραμμάτων είναι της μορφής:

αξιολογούμενη συνάρτηση \_ πλήθος ανεξάρτητων μεταβλητών

πχ. ff1\_30nvars, σημαίνει ότι αξιολογείται η συνάρτηση ff1 για 30 ανεξάρτητες μεταβλητές.

- Στα υπομνήματα των διαγραμμάτων υιοθετήθηκε η μορφή:

*χρησιμοποιούμενος αλγόριθμος \_ αξιολογούμενη συνάρτηση \_ πλήθος ανεξάρτητων μεταβλητών \_ πλήθος υποπληθυσμών x μέγεθος υποπληθυσμού*

πχ. DE\_ff1\_30\_100, σημαίνει ότι με τη χρήση του DE αλγορίθμου αξιολογείται η συνάρτηση ff1 για 30 ανεξάρτητες μεταβλητές και μέγεθος πληθυσμού ίσο με 100, ενώ GA\_ff1\_30\_2x50, σημαίνει ότι με τη χρήση του GA αλγορίθμου αξιολογείται η συνάρτηση ff1 για 30 ανεξάρτητες μεταβλητές και 2 υποπληθυσμούς μεγέθους 50 (πρόκειται για την περίπτωση του *Multi-Deme GA*).

- Στον κατακόρυφο άξονα βρίσκεται η τιμή καταλληλότητας της εκάστοτε συνάρτησης ενώ στον οριζόντιο ο αριθμός των αξιολογήσεων.

Τέλος στα Σχήματα 4 ως 6 παρατηρούμε πώς μεταβάλλεται ο ρυθμός σύγκλησης του αλγορίθμου κι επομένως η δυσκολία του προβλήματος ανάλογα με το μέγεθός του (πλήθος μεταβλητών σχεδίασης) για κάθε συνάρτηση και κάθε αλγόριθμο.

ff 1				
Δοκιμή DE	nvars	popsize	maxgens	fitness
evaluations=150000	30	25	1500	3.0257772076E+00
	30	30	5000	1.8284590375E-06
	30	40	3750	1.8300000000E-06
	30	50	3000	1.8284589700E-06
	30	100	1500	2.6149214285E-05
evaluations=150000	60	50	3000	1.3404194594E+00
	60	60	2500	1.0524962604E-04
	60	80	1875	1.1107713578E-03
	60	100	1500	2.1314633681E-02
evaluations=250000	120	75	2000	3.7002313063E+00
	120	120	2000	1.2701205294E+00
	120	160	1570	4.8557343811E+00
	120	200	1250	4.6797086404E+00
	120	250	800	9.0590988919E+00
evaluations=300000	250	250	1200	9.8118029221E+00
	250	300	1000	1.2852458390E+01

Δοκιμή GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	30	5000	4.7608246529E+00
	30	40	3750	5.2504524116E+00
	30	50	3000	3.6977700064E+00
	30	100	1500	1.2470648963E+01
evaluations=150000	60	50	3000	6.0482426007E+00
	60	60	2500	1.3238244921E+01
	60	80	1875	1.3183092942E+01
	60	100	1500	1.1799396430E+01
evaluations=250000	120	120	2000	1.2665890141E+01
	120	160	1250	2.1530558059E+01
	120	200	1250	2.0852402438E+01
	120	250	800	1.9853532333E+01
evaluations=300000	250	250	1200	1.7045310525E+01

Δοκιμή Multi-Deme GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	2x15	5000	5.2504524116E+00
	30	2x20	3750	2.9374294243E-04
	30	5x10	3000	6.6392383617E-03
	30	5x20	1500	1.0481116782E-03
evaluations=150000	60	50	3000	
	60	4x15	2500	1.0347668852E-02
	60	4x20	1875	1.0383411623E+01
	60	5x20	1500	7.4686201140E-03
evaluations=250000	120	4x30	2000	1.6955857418E+00
	120	4x40	1250	1.4579357100E+01
	120	4x50	1250	3.5394528322E+00
	120	5x50	800	4.1906177120E+00
evaluations=300000	250	5x50	1200	6.2902630449E+00
	250	5x60	1000	6.7636897005E+00

Πίνακας 1. Τα δεδομένα για την συνάρτηση ff1.

ff 0				
Δοκιμή DE	nvars	popsize	maxgens	fitness
evaluations=150000	30	25	3000	5.6569064501E-27
	30	30	3302	3.9163963124E-31
	30	40	3000	8.0453359514E-30
	30	50	3000	1.0066569342E-29
	30	100	1500	1.2173547540E-12
evaluations=150000	60	50	3000	1.6900868680E-13
	60	60	2500	1.2325923128E-11
	60	80	1875	1.3214210093E-08
	60	100	1500	4.8847352235E-07
evaluations=250000	120	120	2100	3.6970955382E-04
	120	160	1570	6.2182008066E-03
	120	200	1250	3.1535003756E-02
	120	250	1000	2.6147767084E-01
evaluations=300000	250	250	1200	1.9160099344E+00
	250	300	1000	3.6465586545E+00

Δοκιμή GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	30	5000	1.1207720217E-01
	30	40	3750	4.1911750159E-06
	30	50	3000	8.4939229801E-05
	30	100	1500	1.4187655641E-01
evaluations=150000	60	50	3000	1.6498083311E+00
	60	60	2500	5.4855953309E-01
	60	80	1875	4.0410940753E-01
	60	100	1500	2.4187379691E-01
evaluations=250000	120	120	2100	4.1128830004E+00
	120	160	1570	1.2803250037E+01
	120	200	1250	1.5357355729E+01
	120	250	1000	7.3187586439E+00
evaluations=300000	250	250	1200	1.6847742603E+01

Δοκιμή Multi-Deme GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	2x20	3750	3.6241887527E-06
	30	5x10	3000	4.1744155326E-07
	30	5x20	1500	4.7774628694E-10
evaluations=150000	60	5x10	3000	6.8409003317E-06
	60	4x15	2500	3.8720191558E-06
	60	4x20	1875	7.2195809009E-01
	60	5x20	1500	1.9051348008E-06
evaluations=250000	120	3x40	2100	7.3485848293E-05
	120	4x40	1570	1.1656213720E-04
	120	4x50	1250	3.6529114777E-04
	120	5x50	1000	1.7109607163E-01
evaluations=300000	250	5x50	1200	1.7724363276E-02
	250	5x60	1000	3.8980717728E-02
	250	4x40	1570	1.20E-02

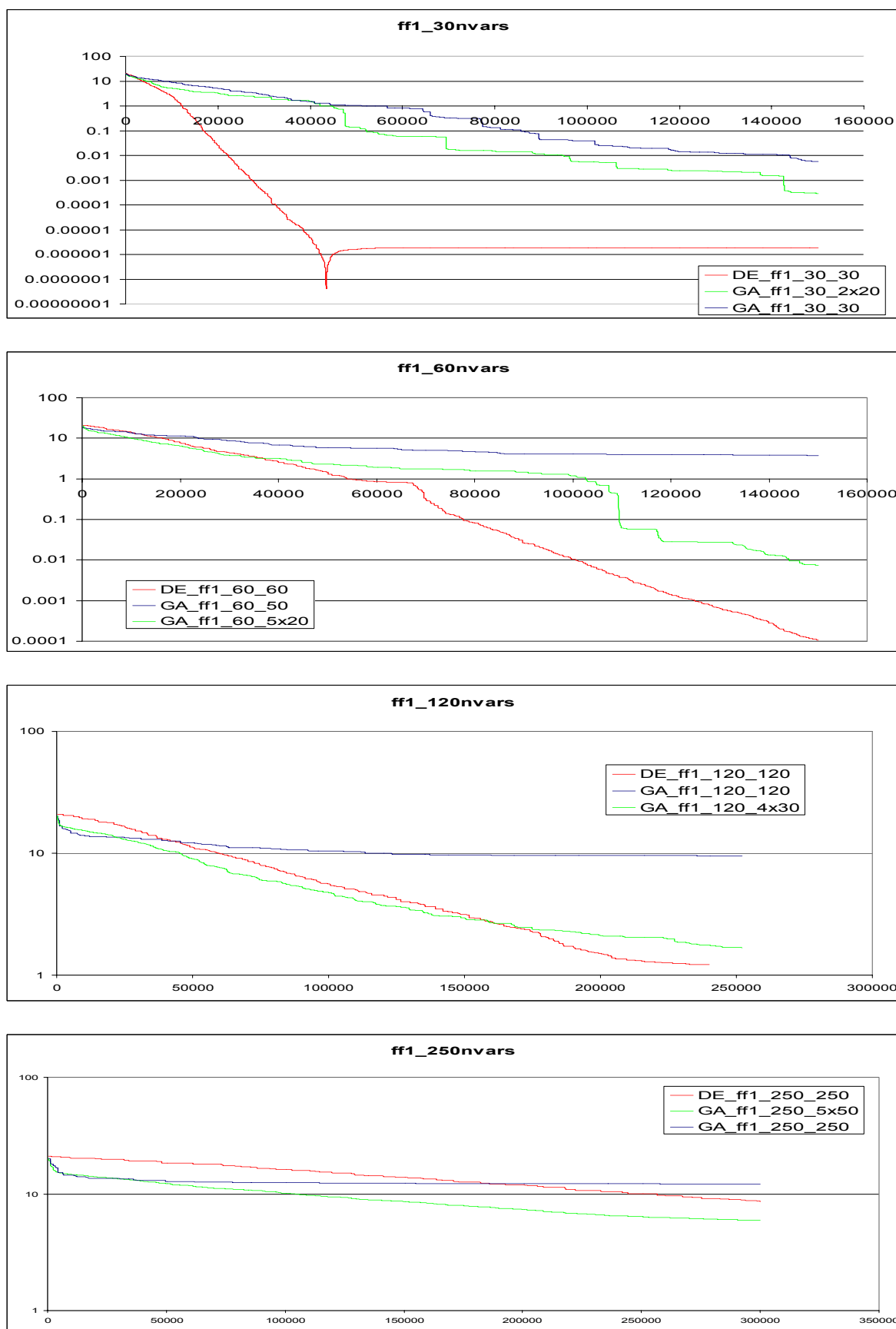
Πίνακας 2. Τα δεδομένα για την συνάρτηση ff0.

ff_4				
Δοκιμή DE	nvars	popsize	maxgens	fitness
evaluations=150000	30	40	1928	1.1100000000E-16
	30	50	2001	1.1102230246E-16
	30	100	1800	1.5904833006E-11
evaluations=150000	60	50	942	3.8300695948E-11
	60	60	3000	4.3880488132E-10
	60	80	2250	5.6480414035E-07
	60	100	1800	1.4918417428E-04
evaluations=250000	120	75		
	120	120	2100	3.0763155700E-01
	120	160	1570	1.1150448740E+00
	120	200	1250	1.6492832360E+00
evaluations=300000	120	250	1000	4.5243674390E+00
	250	250	1200	4.7925262572E+01
	250	300	1000	6.6246814050E+01

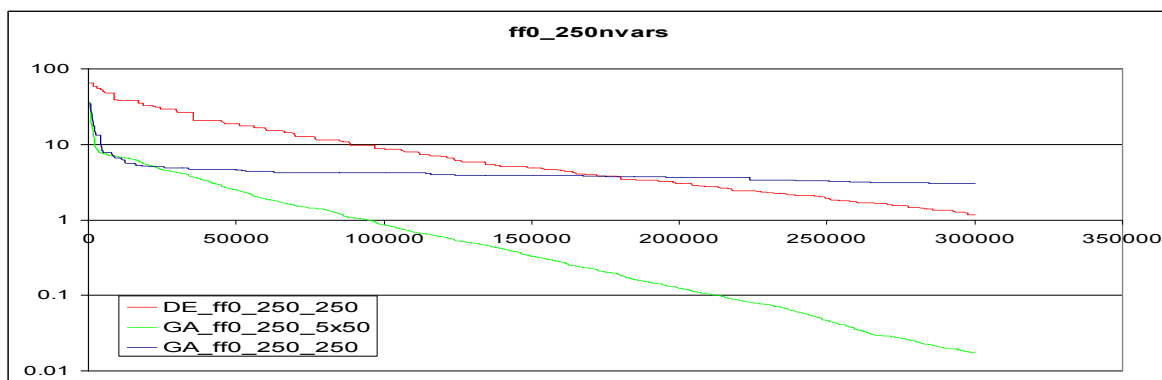
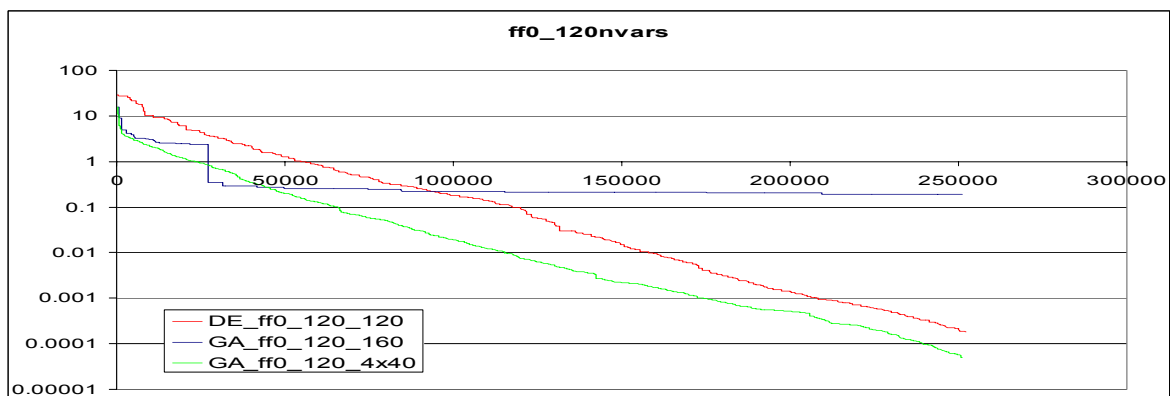
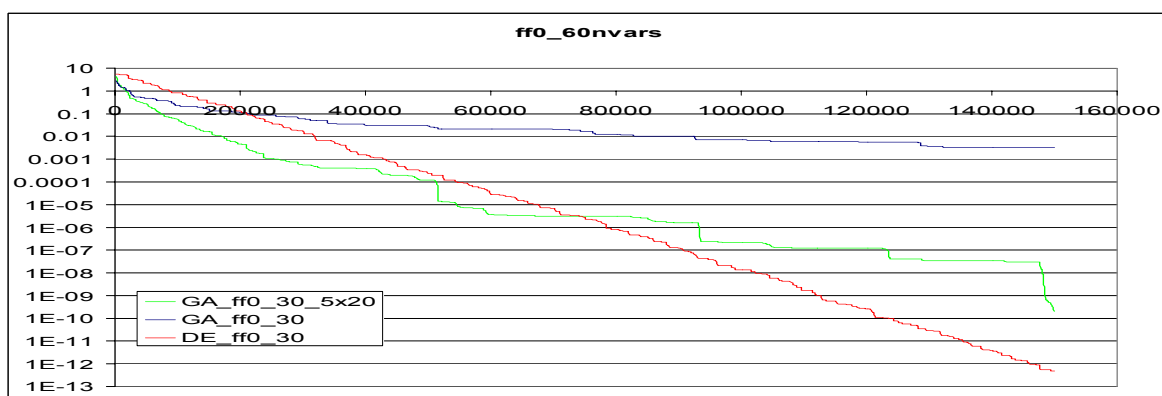
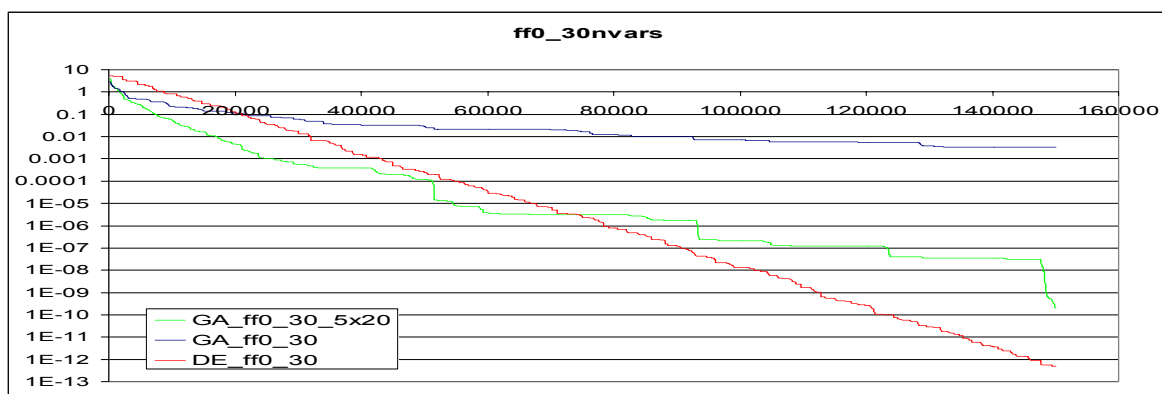
Δοκιμή GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	40	4500	4.4016214379E+00
	30	50	3600	1.5192210622E-01
	30	100	1800	4.9228897447E+02
evaluations=150000	60	50	942	6.3554756805E+02
	60	60	3000	1.3632639804E+02
	60	80	2250	1.4442881211E-04
	60	100	1800	6.3554756805E+02
evaluations=250000	120	120	2100	2.2576907745E+02
	120	160	1570	3.9165075979E+02
	120	200	1250	4.0159517812E+02
	120	250	1000	1.0331969624E+03
evaluations=300000	250	250	1200	4.7020091128E+02

Δοκιμή Multi-Deme GA	nvars	popsize	maxgens	fitness
evaluations=150000	30	2x20	4500	4.9101037191E-02
	30	2x25	3600	1.4307973474E+00
	30	5x20	1800	2.4603032700E-02
evaluations=150000	60	2x25	942	1.4626188917E+00
	60	3x20	3000	7.9544871779E+00
	60	4x20	2250	5.0906103932E+02
	60	5x20	1800	8.5568064608E-04
evaluations=250000	120	3x40	2100	1.7061778573E+00
	120	4x40	1570	8.5109481237E-01
	120	4x50	1250	2.9170571648E+01
	120	5x50	1000	5.3932736729E-01
evaluations=300000	250	5x50	1200	1.5260269221E+00
	250	5x60	1000	2.5618519433E+00

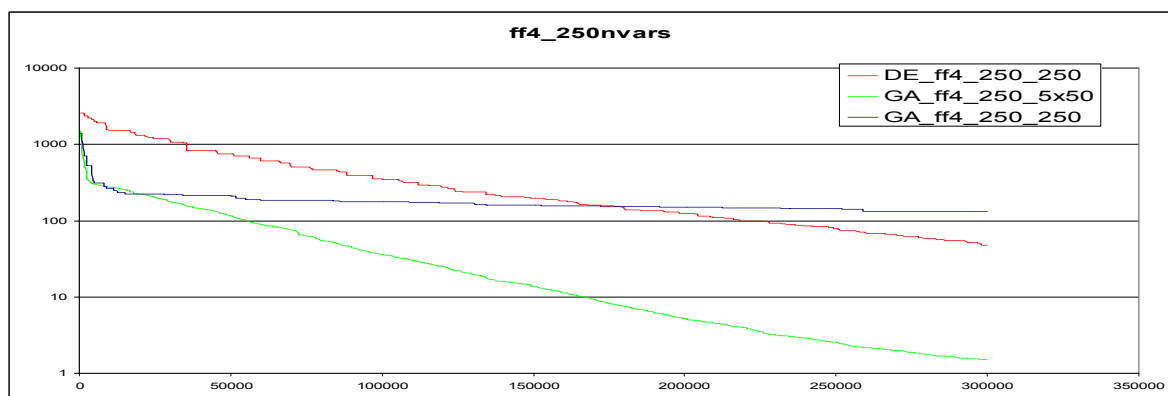
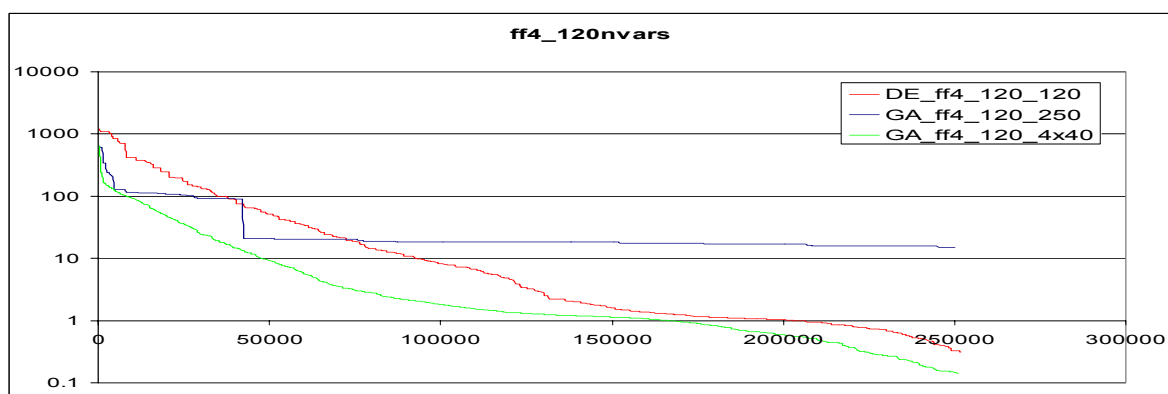
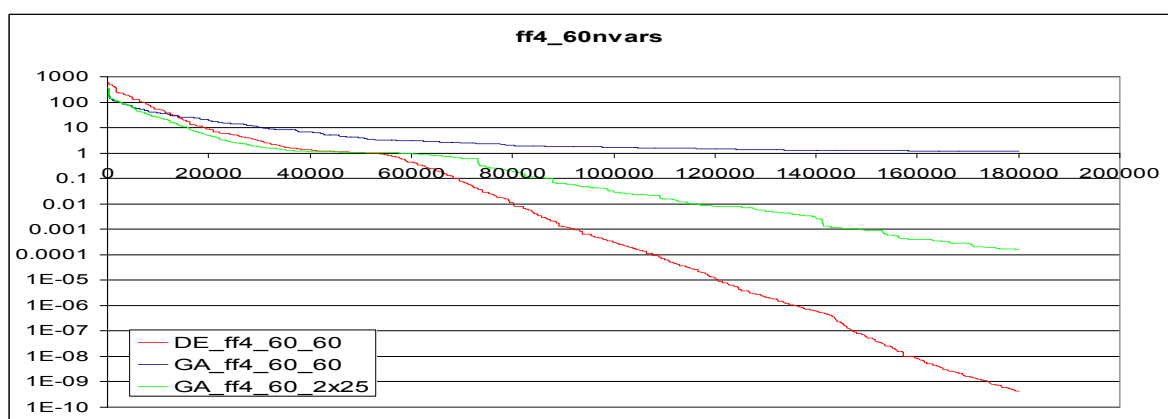
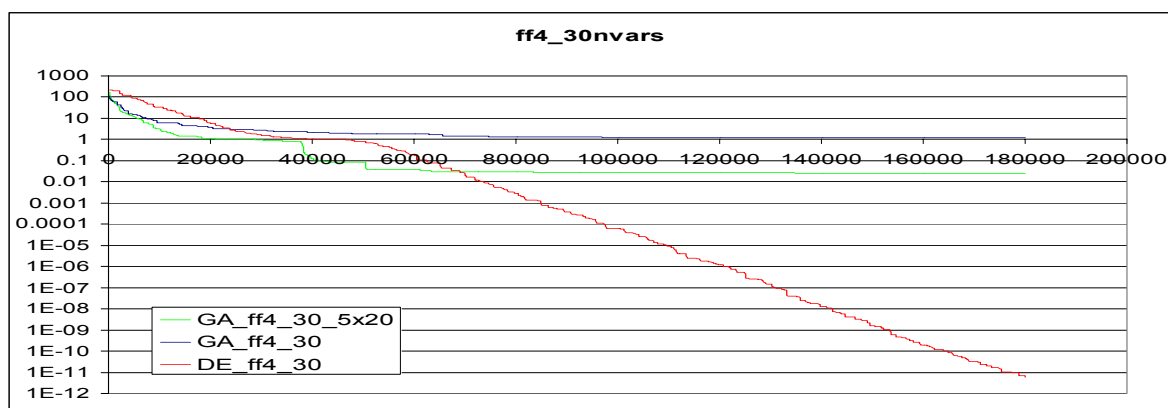
Πίνακας 3. Τα δεδομένα για την συνάρτηση ff4.



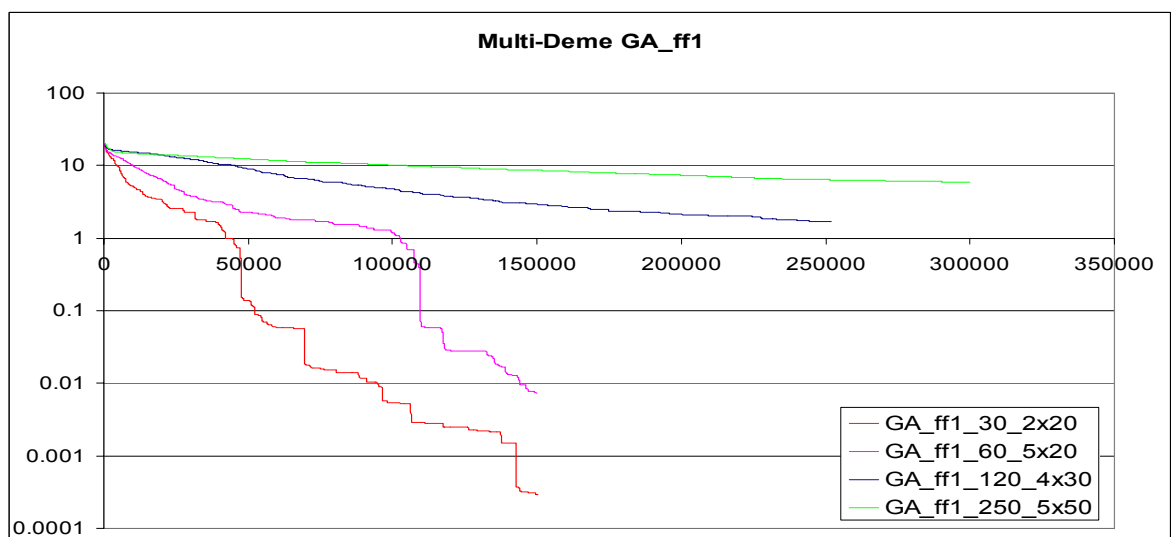
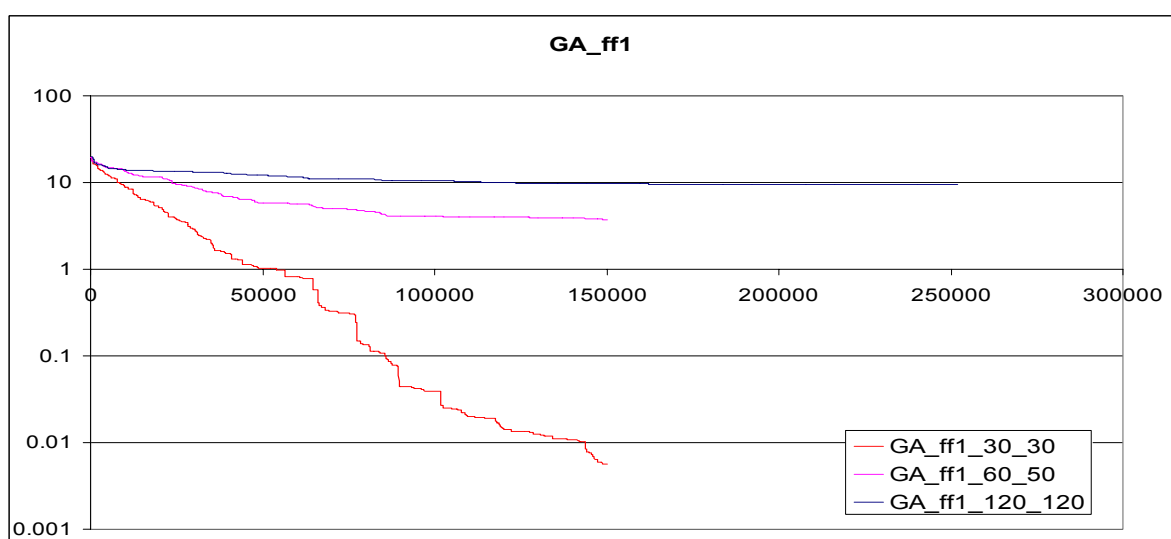
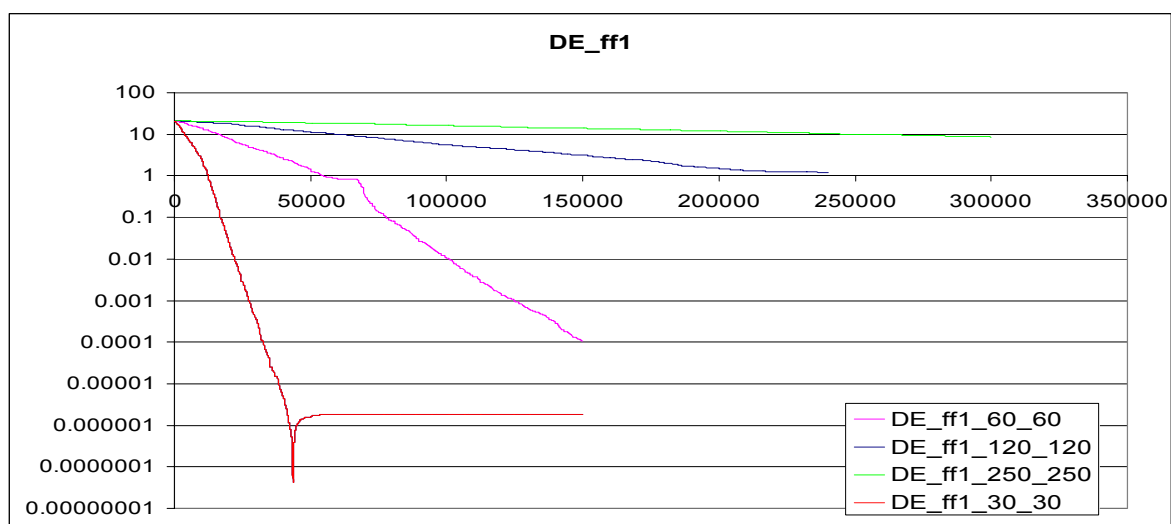
**Σχήμα 1.** Διαγράμματα σύγκλισης που πέτυχαν οι συγκρινόμενοι αλγόριθμοι για διαφορετικά πλήθη ανεξάρτητων μεταβλητών της συνάρτησης  $ff1$ .



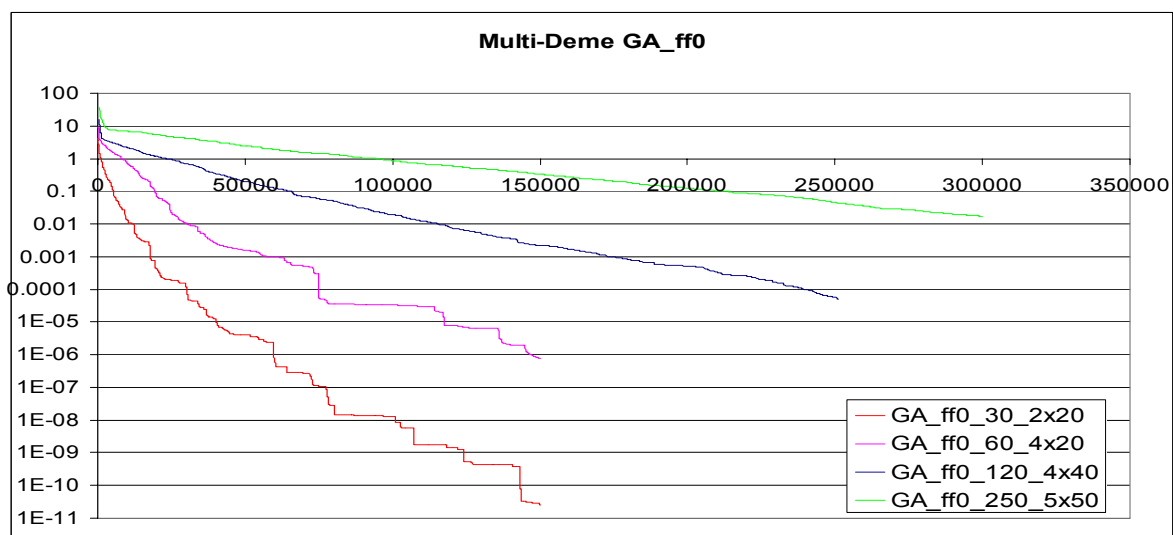
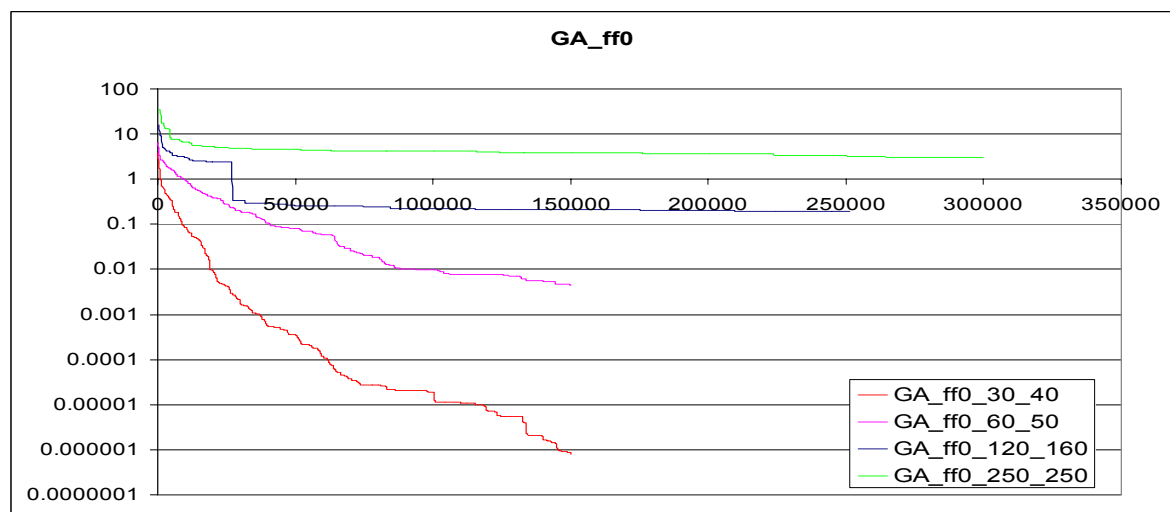
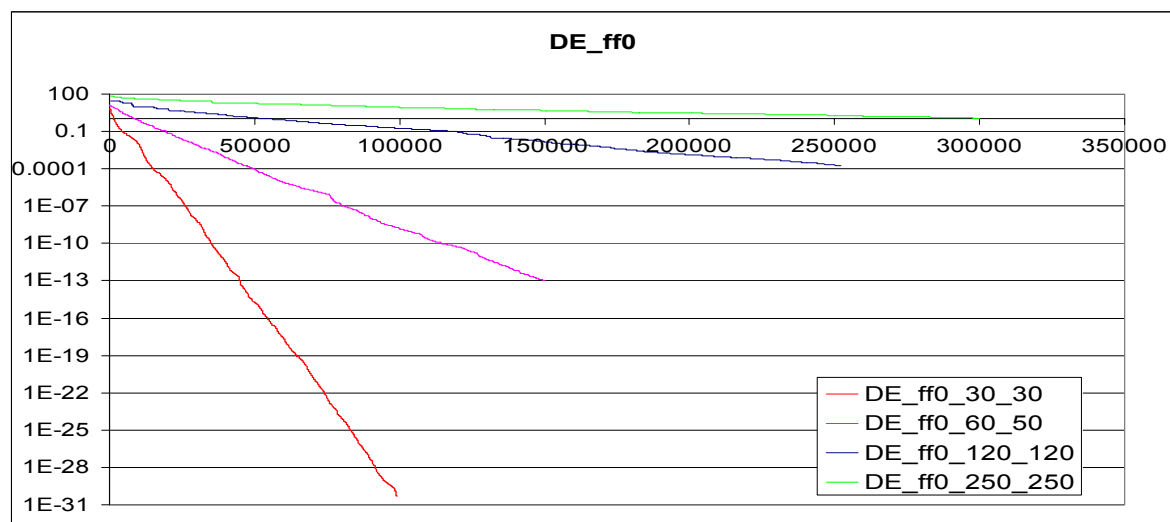
**Σχήμα 2.** Διαγράμματα σύγκλισης που πέτυχαν οι συγκρινόμενοι αλγόριθμοι για διαφορετικά πλήθη ανεξάρτητων μεταβλητών της συνάρτησης *ff0*.



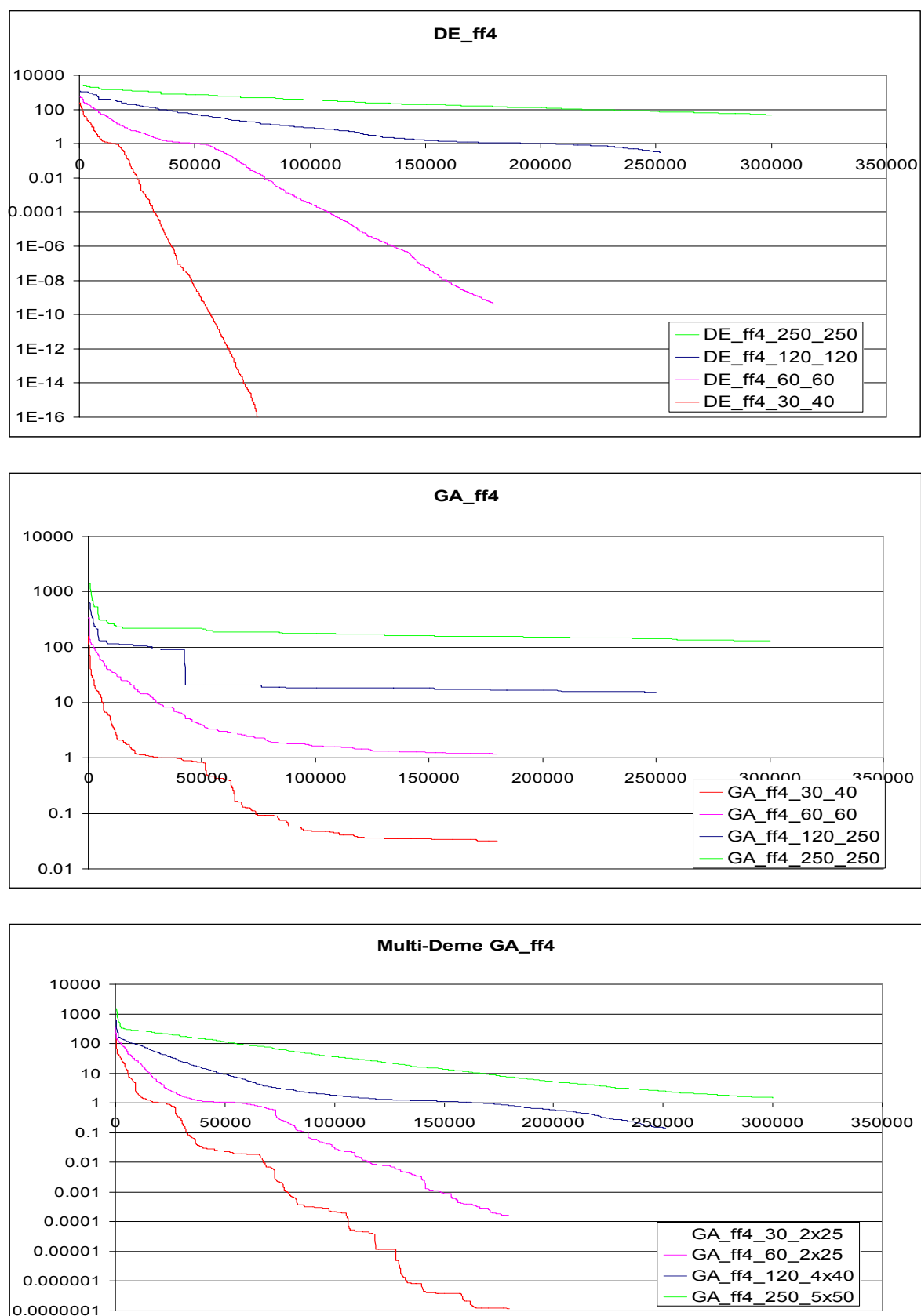
**Σχήμα 3.** Διαγράμματα σύγκλισης που πέτυχαν οι συγκρινόμενοι αλγόριθμοι για διαφορετικά πλήθη ανεξάρτητων μεταβλητών της συνάρτησης *ff4*.



**Σχήμα 4.** Η μεταβολή της σύγκλισης των συγκρινόμενων αλγορίθμων για διάφορα πλήθη ανεξάρτητων μεταβλητών της συνάρτησης *ff1*.



**Σχήμα 5.** Η μεταβολή της σύγκλισης των συγκρινόμενων αλγορίθμων για διάφορα πλήθη ανεξάρτητων μεταβλητών της συνάρτησης  $ff0$ .



**Σχήμα 6.** Η μεταβολή της σύγκλισης των συγκρινόμενων αλγορίθμων για διάφορα πλήθη ανεξάρτητων μεταβλητών της συνάρτησης *ff4*.

Υπενθυμίζοντας ότι τα προβλήματα έχουν μορφοποιηθεί ως προβλήματα ελαχιστοποίησης κι επομένως επιθυμητές είναι οι μικρότερες δυνατές τιμές καταλληλότητας, συμπερασματικά, μπορούμε να επισημάνουμε τα εξής αναφορικά με τη σύγκριση DE και GA:

- Στη συντριπτική πλειοψηφία των περιπτώσεων ο DE αποδεικνύεται αποτελεσματικότερος του GA.
- Για μικρό και μέσο πλήθος μεταβλητών σχεδίασης (30 ως 60) ο DE τα καταφέρνει καλύτερα και από τον Multi-Deme GA. Σταδιακά ωστόσο και όσο ο αριθμός των ανεξάρτητων μεταβλητών αυξάνει, η κατάσταση αντιστρέφεται με τον DE να αποδίδει λιγότερο ή περισσότερο χειρότερα από τον Multi-Deme GA ανάλογα με την αξιολογούμενη συνάρτηση.
- Η δυσκολία διατήρησης ενός ικανοποιητικού ρυθμού σύγκλισης καθώς αυξάνει το πλήθος των ανεξάρτητων μεταβλητών, είναι και για τις τρεις μεθοδολογίες παραπάνω από εμφανής.
- Ο DE φαίνεται να αρκείται σε σχετικά μικρό μέγεθος πληθυσμού για την επίτευξη της καλύτερης απόδοσής του. Το πλήθος αυτό είναι συνήθως ελαφρά μεγαλύτερο του αριθμού των μεταβλητών σχεδίασης για μικρό και μέσο πλήθος των τελευταίων (30 ως 60 μεταβλητές) ή ακόμη και ίσο για μεγαλύτερου μεγέθους προβλήματα.
- Αντίθετα ο GA χρειάζεται σε κάθε περίπτωση μεγαλύτερα μεγέθη πληθυσμού από τον DE, ενώ στην Multi-Deme εκδοχή του αποδοτικότερη φαίνεται να είναι η χρήση λιγότερων και μεγαλύτερου μεγέθους υποπληθυσμών.
- Χαρακτηριστική τέλος είναι σε κάθε περίπτωση η ομαλότερη πορεία σύγκλισης του DE έναντι των GA, κάτι που οφείλεται στον τρόπο που κινείται στο χώρο των λύσεων (κίνηση σε μια γειτονιά λύσεων κι όχι οπουδήποτε στο χώρο), με αποτέλεσμα την αποφυγή έντονων μεταβολών της τιμής καταλληλότητας από γενιά σε γενιά.