ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ &
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

‘Συγκριτική αξιολόγηση μεθόδων καθολικής βελτιστοποίησης για την παραμετροποίηση και ανάλυση βιο-οπτικών σημάτων διαγνωστικού ενδιαφέροντος.’

Κλαπάκη Στέλλα

ΕΠΙΤΡΟΠΗ

Αναπλ. Καθηγητής Μπάλας Κωνσταντίνος (επιβλέπων)
Αναπλ. Καθηγητής Μπλέτσας Άγγελος
Καθηγητής Σταυρακάκης Γεώργιος

ΧΑΝΙΑ, 2014

ΠΕΡΙΕΧΟΜΕΝΑ

Technical University of Crete
School of Electronics and Computers Engineering

Diploma Thesis

"Comparative evaluation of global optimization methods for the parameterization and analysis of bio-optical signals with diagnostic importance"

Klapaki Stella
Supervisor: Assoc. Professor Costas Balas

Abstract

Dynamic Spectral Imaging (DSI) is a novel imaging modality based on the in vivo measurement of spatial-spectral and dynamic optical characteristics induced by biomarkers. A compartmental pharmacokinetic model of the cervical neoplastic epithelium has been developed, with the aid of which the experimental data are fitted, in an attempt to correlate the in vivo measured dynamic back-scattering (DBS) signal with the actual functional and structural characteristics of the tissue. The eight identified well correlating with neoplasia parameters, were reduced to five with the aid of Global Sensitivity Analysis. In this paper we used five state-of-the-art global optimization techniques, equally constrained in a parallel working software environment in order to verify whether the solution of the inverse problem is substantially unique. The statistical results confirmed that for the reduced set, the unidentified experimental optical data converge more optimally, in comparison with the case where the full set is used. Moreover the differential evolution's algorithm error is under 10% for the five most identifiable parameters, fact that sets this one as the best performing algorithm of our study and the results quite promising on our way towards the unique estimation of those parameters, indicating a good accuracy and precision.

1 ΕΙΣΑΓΩΓΗ

Μέσα από αυτή την εργασία σκοπό έχουμε να βρούμε το κατάλληλο σετ παραμέτρων των αλγορίθμων καθολικής βελτιστοποίησης ούτως ώστε αυτοί να λύνουν με τον καλύτερο δυνατό τρόπο που τους επιτρέπει η δομή τους το πρόβλημά μας, δηλαδή την εκτίμηση των παραμέτρων μέσω της σύγκλισης πειραματικών και προσομοιωμένων καμπυλών.

Το πρόβλημα που αντιμετωπίζουμε είναι αυτό της εκτίμησης παραμέτρων σε ένα μοντέλο συστήματος συνήθεις διαφορικών εξισώσεων. Τέτοιου είδους προβλήματα συνήθως ως στόχο έχουν την ελαχιστοποίηση μιας αντικειμενικής συνάρτησης η οποία λειτουργεί σαν μέτρο της εξόδου του μοντέλου αναπαράστασης σε σχέση με τα παρατηρημένα δεδομένα.

Το μοντέλο μας αποδίδει στην έξοδό του πραγματικές τιμές στις παραμέτρους, μία από τις οποίες είναι διακριτή και όχι συνεχής (η παράμετρος n που αποδίδεται στον αριθμό των κυτταρικών στρωμάτων, παίρνει ακέραιες τιμές). Το μοντέλο δεν είναι ντετερμινιστικό οπότε και δεν μπορεί να λυθεί σε πολυωνυμικό χρόνο. Ακόμη είναι μη-γραμμικό, αφού οι μαθηματικές διαφορικές εξισώσεις που το απαρτίζουν δεν αναπαριστούν γραμμικό συνδυασμό των παραμέτρων του.

Τα προαναφερθέντα χαρακτηριστικά του προβλήματός μας μας στρέφουν μακριά από τις κλασσικές μεθόδους, οι οποίες χρησιμοποιούν παραγώγους. Κάτι τέτοιο τις καθιστά ακατάλληλες για προβλήματα με μη-γραμμικές συναρτήσεις κόστους οι οποίες παρουσιάζουν μεγάλο αριθμό ελαχίστων.

Έτσι οι μέθοδοι ολικής βελτιστοποίησης κρίθηκαν οι πλέον κατάλληλες για την περίπτωσή μας, ώστε να επιτύχουμε σύγκλιση, αλλά και κατάλληλες τιμές στις παραμέτρους (αντιστοίχιση με τις πειραματικές). Αυτού του τύπου οι μέθοδοι ελέγχουν όλο το δειγματικό χώρο (σε όλες τις διαστάσεις/παραμέτρους) με στόχο την εύρεση περιοχών αυξημένης πιθανότητας εντοπισμού του ολικού βέλτιστου. Για την επίτευξη αυτού του στόχου χρησιμοποιούν συνδυασμούς διαφόρων τεχνικών.

Αυτές οι μέθοδοι έχουν ένα αριθμό από εσωτερικές παραμέτρους που ρυθμίζουν την απόδοση του εκάστοτε αλγόριθμου δεδομένου των ιδιαιτεροτήτων του εκάστοτε προβλήματος. Στόχος μας ήταν να επιλεχθούν οι πλέον κατάλληλες τιμές για αυτές τις παραμέτρους, μέσα από τη βιβλιογραφία και μέσα από τις δοκιμές ώστε να επιτευχθεί η καλύτερη δυνατή απόδοση του αλγόριθμου βάση του μοντέλου μας.

Στο κεφάλαιο 1, της εργασίας αυτής κάνουμε αναφορά στα δυναμικά μοντέλα βιολογικών διεργασιών. Πιο συγκεκριμένα αναφερόμαστε στις διάφορες κατηγορίες που προκύπτουν από τη δυσκολία και την πολυπλοκότητα αναπαράστασης των σύνθετων φυσιολογικών διεργασιών από μαθηματικά μοντέλα.

Στο κεφάλαιο 2, περιγράφουμε το πρόβλημα. Ο καρκίνος του τραχήλου, το φαινόμενο της λεύκανσης σε συνδυασμό με το προηγούμενο κεφάλαιο μας επιτρέπουν την κατανόηση της μεταφοράς υλικών ανάμεσα σε διαμερίσματα και πώς αυτό σχετίζεται με τη λεύκανση του ιστού σε περίπτωση νεοπλασίας μετά την έγχυση διαλύματος οξικού οξέος στον τράχηλο. Ακολουθεί η μαθηματική περιγραφή του μοντέλου μας.

Στο κεφάλαιο 3, αναφερόμαστε στο αντίστροφο πρόβλημα, αυτό της εκτίμησης των τιμών ενός συνόλου βιολογικών παραμέτρων, χρησιμοποιώντας σαν είσοδο μετρήσεις δυναμικών οπτικών σημάτων. Συγκεκριμένα ένα δοκιμαστικό σήμα προσεγγίζεται από το μοντέλο και επιστρέφεται ένα σύνολο παραμέτρων. Αναλύθηκαν επίσης οι 5 μέθοδοι ολικής βελτιστοποίησης που επιστρατεύσαμε: Controlled Random Search (CRS), Differential Evolution (DE), Direct _gblSolve(DIR), Particle Swarm Optimization (PSO) και Shuffled Complex Evolution(SCE). Αποδεκτή λύση για ένα τέτοιο πρόβλημα μπορεί να θεωρηθεί μόνο εάν για τον εκάστοτε αλγόριθμο το πρόβλημα συγκλίνει ικανοποιητικά και οι τιμές των παραμέτρων προβλέπονται με ακρίβεια και συνέπεια.

Στο κεφάλαιο 4, εμφανίζουμε αναλυτικά τα αποτελέσματα της έρευνάς μας με κριτήριο το χρόνο, τον αριθμό των function evaluations και το μέσο Normalized Root Mean Square Deviation (NRMSD). Λαμβάνοντας υπ' όψιν τα αποτελέσματα μιας a priori έρευνας που έγινε με στόχο τη μείωση του θορύβου, χρησιμοποιήσαμε το μοντέλο μας με έξοδο 8 και 5 παραμέτρους. Στόχος μας να επιβεβαιώσουμε την έρευνα αυτή με στατιστικά αποτελέσματα.

Στο κεφάλαιο 5, παρουσιάζουμε τα συμπεράσματα της έρευνας ως προς την επιβεβαίωση της a priori έρευνας που ισχυρίζεται τον εντοπισμό των 5 πλέον ευδιάκριτων παραμέτρων(σε σύγκριση με τις αρχικές 8). Έπειτα έχουμε ενδείξεις ότι η παράμετρος a προκαλεί θόρυβο, κάτι που είχε προκύψει και από την a priori έρευνα. Στη συνέχεια η παράμετρος gtj εμφανίζεται ως η πλέον αξιόπιστα προβλέψιμη. Τέλος ο αλγόριθμος DE στο σύνολο του εμφανίζει τα βέλτιστα αποτελέσματα για 5 παραμέτρους. Στη συνέχεια επιχειρούμε εξήγηση των αποτελεσμάτων μας με τη βοήθεια πορισμάτων από τη βιβλιογραφία. Τέλος κάνουμε λόγο για μελλοντικές επεκτάσεις της έρευνας.

ΚΕΦΑΛΑΙΟ 1

1.1 ΕΙΣΑΓΩΓΗ [28][29][30]

Οι φυσιολογικές διεργασίες εμφανίζουν σύνθετη δυναμική, ενώ περιέχουν παραμέτρους που δεν μπορούν να μετρηθούν ή να προβλεφθούν με ευκολία. Η πολυπλοκότητα λοιπόν των μοντέλων είναι ανάλογη της πολυπλοκότητας των φυσιολογικών διεργασιών. Ένα διαμερισματικό μοντέλο είναι ένας τύπος μαθηματικού μοντέλου που χρησιμοποιείται για να περιγράψει ένα φυσιολογικό σύστημα. Οι παράμετροι που χρησιμοποιούνται σε ένα τέτοιο μοντέλο χαρακτηρίζονται ως lumped parameters οι οποίες αποτελούν απλούστευση του μαθηματικού μοντέλου ενός φυσικού συστήματος, με την οποία οι μεταβλητές οι οποίες είναι κατανεμημένες στο χώρο αναπαρίστανται ως διαβαθμισμένες τιμές. Π.χ. η θερμοκρασία ενός χώρου αντικαθίσταται από τις μέσες θερμοκρασίες. Μια τέτοια προσέγγιση μπορεί να είναι λογική να το σύστημα είναι αρκετά ομοιογενές.

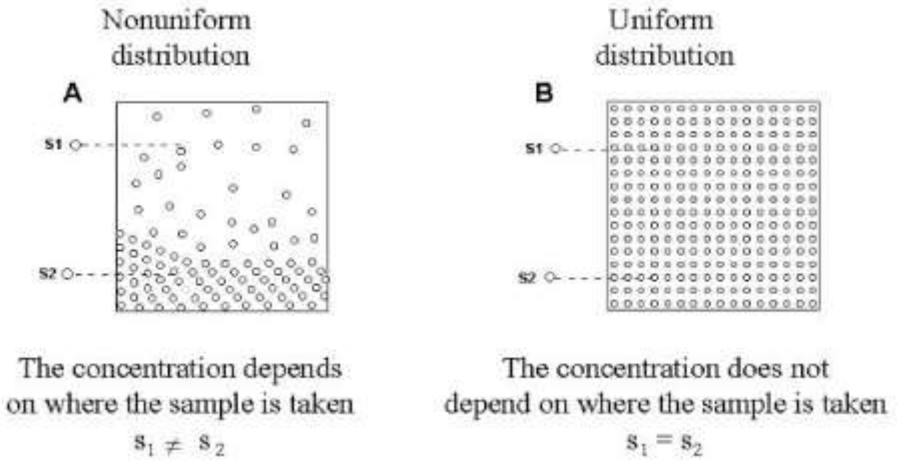
Στη βιβλιογραφία συναντάμε διάφορες κατηγορίες μοντέλων: τα αδρά αιτιοκρατικά μοντέλα (γραμμικά και μη γραμμικά), τα κατανεμημένα μοντέλα, τα στοχαστικά μοντέλα, καθώς και τα αμιγή διαμερισματικά και τα μοντέλα συστημάτων ελέγχου.

1.2 ΔΙΑΜΕΡΙΣΜΑΤΙΚΑ ΜΟΝΤΕΛΑ

Τα διαμερισματικά μοντέλα χρησιμοποιούνται κυρίως για τη μελέτη της κινητικής ουσιών σε φυσιολογικά συστήματα. Οι ουσίες αυτές μπορεί να είναι είτε εξωγενείς (φάρμακα ή ουσίες επισήμανσης) ή ενδογενείς (ορμόνες κ.λπ.). Τα διαμερισματικά μοντέλα περιγράφουν διαδικασίες παραγωγής, κατανομής, μεταφοράς, χρησιμοποίησης των ουσιών αυτών, καθώς και μηχανισμούς αλληλεπίδρασης με, και ελέγχου από, άλλες ουσίες.

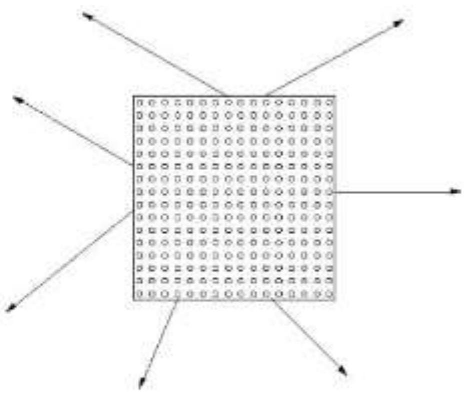
Ως *διαμέρισμα (compartment)* ορίζεται η ποσότητα μιας ουσίας, η οποία συμπεριφέρεται σα να είναι καλά αναμεμειγμένη και κινητικά ομοιογενής. Ένα *διαμερισματικό μοντέλο (compartmental model)* αποτελείται από πεπερασμένο αριθμό διαμερισμάτων με καθορισμένο τρόπο σύνδεσης μεταξύ τους. Οι συνδέσεις των διαμερισμάτων αντιπροσωπεύουν τη ροή ουσίας από ένα διαμέρισμα σε άλλο, η οποία θεωρείται ότι εξαρτάται γραμμικά ή μη γραμμικά από τη μάζα ή τη συγκέντρωση της ουσίας στο διαμέρισμα-αφετηρία.

Στη συνέχεια ορίζονται οι έννοιες της καλής ανάμιξης και της κινητικής ομοιογένειας. Καλή ανάμιξη σημαίνει ότι αν ληφθούν δύο δείγματα από ένα διαμέρισμα την ίδια χρονική στιγμή, τότε θα μετρηθεί η ίδια συγκέντρωση της ουσίας που μελετάται και άρα τα δείγματα θα είναι το ίδιο αντιπροσωπευτικά**(σχήμα4).** Επομένως η έννοια της καλής ανάμιξης αντιστοιχεί στην ομοιογένεια της πληροφορίας που περιέχεται σε κάθε απλό διαμέρισμα.



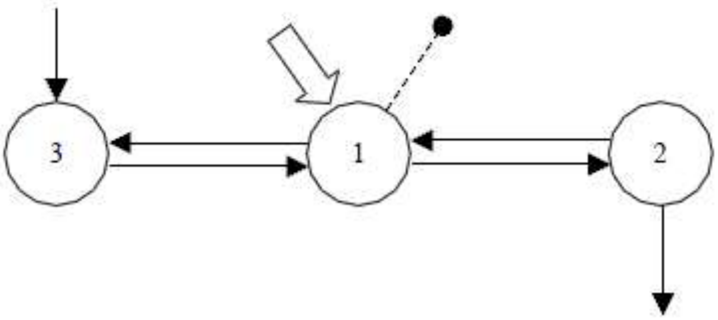
Σχήμα 1.

Κινητική ομοιογένεια σημαίνει ότι κάθε σωματίδιο στο διαμέρισμα έχει την ίδια πιθανότητα να ακολουθήσει τις οδούς για εγκατάλειψη του **διαμερίσματος** (σχήμα 2), δηλαδή να εισέλθει σε μεταβολικές διαδικασίες που σχετίζονται με μεταφορά ή χρησιμοποίηση.



Σχήμα 2.

Η έννοια του διαμερίσματος: κατανεμημένη ουσία με συναφή χαρακτηριστικά σε ομοιογενή συλλογή που συμπεριφέρονται με τον ίδιο τρόπο και επιτρέπει τη μοντελοποίηση ενός πολύπλοκου φυσιολογικού συστήματος με τη βοήθεια ενός πεπερασμένου αριθμού διαμερισμάτων και μεταβολικών οδών. Ο αριθμός των απαιτούμενων διαμερισμάτων εξαρτάται από το σύστημα που μελετάται και το πείραμα που πραγματοποιείται. Τα διαμερισματικά μοντέλα προσφέρουν βαθύτερη κατανόηση της δομής των συστημάτων που προσομοιώνουν, ενώ η ακρίβειά τους εξαρτάται από τις υποθέσεις που ενσωματώνονται κατά τη δόμησή τους. Ένα παράδειγμα διαμερισματικού μοντέλου εικονίζεται στο σχήμα 3, όπου για την αναπαράσταση των διαμερισμάτων χρησιμοποιούνται οι κύκλοι, ενώ για τις συνδέσεις τα βέλη.



Σχήμα 3. *Μοντέλο τριών διαμερισμάτων. Η μελετώμενη ουσία εισέρχεται στο διαμέρισμα (3) και εξέρχεται από το διαμέρισμα (2) προς το εξωτερικό περιβάλλον. Ανταλλαγή ουσίας έχουμε μεταξύ των διαμερισμάτων (3)-(1) και (1)-(2). Η ανταλλαγή παριστάνεται με βέλη. Το διαμέρισμα (1) είναι προσπελάσιμο. Αυτό συμβολίζεται με το μεγάλο βέλος (είσοδος ελέγχου) και τη διακεκομμένη γραμμή (μετρήσεις)*

Οι συνδέσεις των διαμερισμάτων αντιπροσωπεύουν τη ροή ουσίας από το ένα διαμέρισμα στο άλλο, η οποία θεωρείται ότι εξαρτάται γραμμικά ή μη από τη μάζα ή τη συγκέντρωση ουσίας στο διαμέρισμα-αφετηρία. Η μεταφορά υλικού ανάμεσα σε δύο διαμερίσματα πραγματοποιείται είτε λόγω φυσική μεταφοράς από μια περιοχή στην άλλη, είτε από μια χημική αντίδραση. Αυτό σημαίνει ότι στο σύστημα έχει επέλθει μια αλλαγή στις προκείμενες ποσότητες του συστήματος μέσω διαδρομών, οι οποίες βασίζονται στις σταθερές μεταβλητές του συστήματος. Μια αλλαγή σε ένα τέτοιο σύστημα κατά την πάροδο του χρόνου συνήθως εκφράζεται με διαφορικές εξισώσεις. Έτσι, η διαμερισματική ανάλυση είναι ένας γεωμετρικός τρόπος για να εκπροσωπηθεί ένα σύστημα διαφορικών εξισώσεων.

Στην πλειοψηφία των περιπτώσεων τα μαθηματικά μοντέλα που περιγράφουν τις διαδικασίες μεταβολισμού είναι ντετερμινιστικά, δηλ. δεν περιλαμβάνουν στοιχεία τυχαιότητας. Συχνά ένα συγκεντρωμένο μοντέλο είναι επαρκές για την περιγραφή της δυναμικής ενός μεταβολικού συστήματος. Αυτό σημαίνει ότι και τα κατανεμημένα, συνενώνονται και αντιμετωπίζονται σαν μία ομοιογενής οντότητα.

Τα **συγκεντρωμένα/** διαμερισματικά μοντέλα συνήθως αναπαριστώνται από ένα σύνολο πρώτης τάξης συνήθων εξισώσεων, το οποίο σύνολο μπορεί να είναι γραμμικό ή μη γραμμικό. Κάθε μοντέλο μπορεί να χαρακτηριστεί ως γραμμικό ή μη. Αυτό το σύνολο των διαφορικών εξισώσεων πρώτης τάξης που είναι γνωστό ως το σύνολο των εξισώσεων μεταβλητής κατάστασης, όπου η μάζα της ουσίας σε κάθε ένα από τα διαμερίσματα είναι μία μεταβλητή κατάσταση.

$$\dot{x} = f(x(t,p), u(t), t; p), \qquad x_0 = x(t_o, p) \qquad (1)$$

$$y = g(x(t,p); p), \qquad t_0 \leq t \leq T \qquad (2)$$

όπου $\dot{x} = \frac{dx}{dt}$; $x = [x_1\, x_2 .. x_n\,]^T$ (T δηλώνει μεταφορά) είναι το διάνυσμα των μεταβλητών κατάστασης (π.χ. μάζα ή συγκέντρωση); $u = [u_1\, u_2 .. u_r\,]^T$δηλώνει το διάνυσμα εισόδου; $y = [y_1\, y_2 .. y_m\,]^T$είναι το διάνυσμα m εξόδων ή μεταβλητών μέτρησης; $t = [t_o, T]$ δηλώνει το διάστημα παρατήρησης; $p = [p_1\, p_2 .. p_p\,]^T$ είναι το διάνυσμα άγνωστων σταθερών παραμέτρων (π.χ. παράμετροι μεταφορών); f καθορίζει τη γενικά μη γραμμική δομή του μοντέλου βασιζόμενο σε εκ των προτέρων φυσικοχημική γνώση που αφορά την μεταβλητή κατάστασης x με το διάνυσμα εισόδου u όπου οι τιμές των σχέσεων καθορίζονται από τις παραμέτρους p; g καθορίζει τη διαδικασία μέτρησης, δηλ. συνδέει το διάνυσμα εξόδου y με το διάνυσμα x όπου πάλι οι ειδικές σχέσεις ορίζονται από τις κατάλληλες παραμέτρους του p διανύσματος. Εάν οι παράμετροι p είναι συνεχείς, το σύστημα λέγεται ότι είναι χρονικά αναλλοίωτο.

1.3 ΓΡΑΜΜΙΚΑ ΜΟΝΤΕΛΑ

Σε πολλές περιπτώσεις, η υιοθέτηση ενός γραμμικού δυναμικού μοντέλου για ένα μεταβολικό σύστημα είναι απολύτως επαρκής. Σε τέτοια γραμμικά μοντέλα οι μεταβλητές κατάστασης και τα παράγωγά τους εμφανίζονται σε γραμμικό συνδυασμό, οπότε το θεώρημα επαλληλίας εφαρμόζεται. Έτσι η συνολική απόκριση συστήματος αρκετών εισόδων είναι ουσιαστικά το άθροισμα των αποκρίσεων των επιμέρους εισόδων.

Σε τέτοια γραμμικά μοντέλα οι εξισώσεις (1) και (2) μετατρέπονται σε:

$$\dot{x} = A(p)x(t) + B(p)u(t), \qquad x_0 = x(t_0, p) \qquad (3)$$

$$y = C(p)x(t), \qquad t_0 \le t \le T \qquad (4)$$

όπου A,B και C είναι σταθεροί πίνακες συντελεστών κατάλληλων διαστάσεων.

Το γραμμικό μοντέλο είναι επίσης κατάλληλο για περιπτώσεις όπου εφαρμόζεται μικρή διαταραχή-γραμμικοποίηση γύρω από το σημείο λειτουργίας. Τότε οι δυναμικές εξισώσεις είναι γραμμικές ακόμη και αν η εγγενής δυναμική ενός τέτοιου συστήματος είναι μη γραμμική.

Μη γραμμικά μοντέλα

Συχνά, μικρή διαταραχή γύρω από το σημείο λειτουργίας αποκαλύπτει ελάχιστα στοιχεία σχετικά με τη συμπεριφορά του συστήματος. Για να επιτευχθεί αυτό πρέπει να εφαρμοστούν μεγάλα σήματα διαταραχής. Σε τέτοιου είδους περιπτώσεις τα γραμμικά μοντέλα δεν χρησιμοποιούνται, εκτός κι αν η φύση του συστήματος είναι γραμμική. Έτσι, για λύσεις ακριβείας σε μη γραμμικά μοντέλα, απαιτούνται πολύπλοκες τεχνικές.

1.4 ΚΑΤΑΝΕΜΗΜΕΝΑ ΜΟΝΤΕΛΑ

Σε πολλές περιπτώσεις η εκπροσώπηση των μεταβολικών συστημάτων από συγκεντρωμένα διαμερισματικά μοντέλα είναι επαρκής. Αυτό προϋποθέτει ότι υπάρχει ομογενής κατανομή του υλικού μέσα στο χώρο. Σε μοντέλα διανομής/ κατανεμημένα δεν μπορούν να γίνουν υποθέσεις σχετικά με την ομοιογένεια. Αυτό οδηγεί σε μαθηματική διατύπωση με όρους μερικών διαφορικών εξισώσεων. Ένα γενικό εισαγωγικό πλαίσιο για τη μοντελοποίηση της κινητικής των μεταβολικών διεργασιών σε κατανεμημένα συστήματα δίνεται στο [29].

1.5 ΣΤΟΧΑΣΤΙΚΑ ΜΟΝΤΕΛΑ

Τα ντετερμινιστικά μοντέλα και τα εισόδου/εξόδου πειράματα αναγνώρισης χρησιμοποιούνταν και χρησιμοποιούνται με επιτυχία για την ποσοτική περιγραφή πολλών μεταβολικών συστημάτων. Σε ορισμένες περιπτώσεις, ειδικά σε κυτταρικό επίπεδο, ένα καθαρά ντετερμινιστικό μοντέλο δεν είναι επαρκές, αφού δεν λαμβάνει υπ' όψιν την επίδραση του τυχαίου. Πιθανολογικές ή τυχαίες επιδράσεις που είναι ιδιαίτερα σημαντικές στη μοντελοποίηση ενός συστήματος μεταβολισμού είναι εκείνες που εξελίσσονται με την πάροδο του χρόνου. Η στοχαστικότητα συμπεριλαβμάνεται στα βασικά στοιχεία του διαμερισματικού μοντέλου, χηλαδή στις μεταβλητές κατάστασης ή στις παραμέτρους ρυθμού μεταφοράς και οφείλεται σε παράγοντες όπως: μεταβολή παραμέτρων μοντέλου σε σχέση με το χρόνο, επιδράσεις περιβάλλοντος κ.λπ. Τα στοχαστικά μοντέλα χρησιμοποιούνται στο μεταβολικό και ανοσοποιητικό σύστημα και στη μοντελοποίηση φαινομένων σε κυτταρικό επίπεδο.

1.6 ΑΜΙΓΗ ΔΙΑΜΕΡΙΣΜΑΤΙΚΑ ΜΟΝΤΕΛΑ

Τα αμιγή μοντέλα διαμερισμάτων αποτελούν μια κατηγορία συγκεντρωμένων παραμετρικών μοντέλων που βρίσκουν ιδιαίτερη εφαρμογή στις μελέτες του μεταβολικού συστήματος. Τα μοντέλα περιγράφουν διαδικασίες οι οποίες δεν εμπλέκουν ενεργό ορμονικό έλεγχο, δηλαδή διαδικασίες που ρυθμίζονται μόνο από τη χημική, αποθηκευτική και μεταφορική διαδικασία. Γενικά, τέτοια μοντέλα περιγράφουν μεταβολικές διαδικασίες στις οποίες η ροή ουσίας από το ένα διαμέρισμα στο άλλο μπορεί να θεωρηθεί ότι εξαρτάται γραμμικά ή μη γραμμικά από τη μάζα ή τη συγκέντρωση του υλικού μόνο στο διαμέρισμα-αφετηρία.

Η διατύπωση των μαθηματικών εξισώσεων για τα αμιγή διαμερισματικά μοντέλα πραγματοποιείται σε δύο στάδια: Στο πρώτο στάδιο διατυπώνονται οι εξισώσεις για τη μεταβολή της μάζας κάθε διαμερίσματος, οι οποίες αποτελούνται από όρους που περιγράφουν ροή ουσίας σε αυτό και ροή από αυτό σε άλλους χώρους. Η μορφή των εξισώσεων που περιγράφουν τη δυναμική του *i*-διαμερίσματος ενός συστήματος *n* διαμερισμάτων είναι:

$$\dot{Q}_\iota(t) = R_{\iota 0} + \sum_{\substack{j=1 \\ j\neq 1}}^{n} R_{\iota j}(Q_\iota) - R_{0\iota}(Q_\iota), \qquad \iota = 1,2,\dots,n \tag{5}$$

όπου

Q_i : η ποσότητα του υλικού στο διαμέρισμα i.

R_{ij} : η ροή του υλικού στο διαμέρισμα i από το διαμέρισμα j εξαρτώμενη μόνο από την ποσότητα Q_i

R_{0i} : η έξοδος του υλικού από το διαμέρισμα i από το εξωτερικό περιβάλλον

R_{i0} : η είσοδος του υλικού από το διαμέρισμα i από το εξωτερικό περιβάλλον

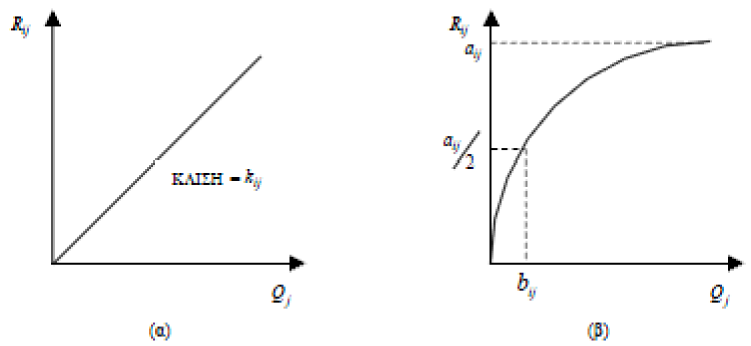
Το δεύτερο στάδιο απαιτεί τον προσδιορισμό της λειτουργικής εξάρτησης κάθε παραμέτρου
ροής. Η φύση αυτής της εξάρτησης μπορεί να είναι γραμμική ή μη, και η τιμή της παραμέτρου μπορεί να προέρχεται είτε από *a priori* γνώση ή μετά από παραμετρική εκτίμηση. Η πιο συχνή μορφή εξάρτησης είναι η γραμμική εξάρτηση και η Michaelis-Menten. Η μαθηματική περιγραφή δίνεται στη συνέχεια:

Γραμμική εξάρτηση: $R_{ij} = k_{ij} * Q_i$

όπου k_{ij} είναι σταθερά που καθορίζει τον *κλασματικό συντελεστή μεταφοράς* και προσδιορίζει
το ρυθμό της μεταφοράς ουσίας στο διαμέρισμα i από το διαμέρισμα j.

Δυναμική Michaelis – Menten: $\boldsymbol{R_{ij} = \frac{a_{ij}Q_j}{b_{ij}+Q_j}}$

όπου a_{ij} είναι η μέγιστη τιμή (τιμή κορεσμού) της ροής R_{ij} και b_{ij} είναι η τιμή της ποσότητας Q_j, στην οποία η R_{ij} είναι ίση με το μισό της μέγιστης τιμής της.



Σχήμα 6. Γραφική αναπαράσταση της λειτουργικής εξάρτησης της ροής R_{ij} προς το διαμέρισμα i : (α) γραμμική εξάρτηση, (β) δυναμική Michaelis-Menten.

ΚΕΦΑΛΑΙΟ 2

2.1 ΚΑΡΚΙΝΟΣ ΤΟΥ ΤΡΑΧΗΛΟΥ ΤΗΣ ΜΗΤΡΑΣ

Ο καρκίνος του τραχήλου της μήτρας αναπτύσσεται ως αποτέλεσμα εμμένουσας λοίμωξης από έναν ιό, τον ιό των ανθρώπινων θηλωμάτων (HPV). Ο HPV είναι ένας ιός που προκαλεί πολύ συχνά λοιμώξεις και μεταδίδεται εύκολα. Ο καρκίνος του τραχήλου της μήτρας είναι ο δεύτερος συχνότερος καρκίνος σε γυναίκες κάτω των 45 ετών, ενώ αποτελεί την τρίτη αιτία θανάτου από καρκίνο στις γυναίκες παγκοσμίως, μετά τον καρκίνο του μαστού και του πνεύμονα. Οι δύο κύριοι τύποι καρκίνου του τραχήλου της μήτρας είναι το Πλακώδες Καρκίνωμα, που συναντάται συχνότερα, και το Αδενοκαρκίνωμα.

Η εξέλιξη της λοίμωξης με HPV

Η εμμένουσα λοίμωξη με ογκογόνο τύπο του ιού μπορεί να προκαλέσει την ανάπτυξη παθολογικών και προκαρκινικών αλλοιώσεων στα κύτταρα του τραχήλου της μήτρας, γνωστών ως τραχηλικών ενδοεπιθηλιακών νεοπλασιών (CIN) οι οποίες με τον καιρό μπορεί να εξελιχθούν σε καρκινικές. Οι αλλοιώσεις κατηγοριοποιούνται ανάλογα με την πιθανότητα που έχουν να εξελιχθούν σε διηθητικό καρκίνο, σε CIN1, CIN2 και CIN3. Η εξέλιξη σε καρκίνο του τραχήλου της μήτρας είναι γενικά ασυμπτωτική ενώ η χρονική διάρκεια της εξέλιξης είναι συνήθως μεγάλη.

Η πρόληψη έχει σαν αποτέλεσμα την αντιμετώπιση του καρκίνου στο πρώτο στάδιο το οποίο θεραπεύεται πλήρως. Η θεραπεία του καρκίνου του τραχήλου της μήτρας σε πιο προχωρημένο στάδιο είναι χειρουργική ή με ακτινοβολία.

2.2 ΛΕΥΚΑΝΣΗ

Το φαινόμενο της λεύκανσης παρουσιάζεται σε μη φυσιολογικά λύτταρα όταν γίνει σε αυτά έγχυση οξικού οξέος. Πιο αναλυτικά, τα αποτελέσματα της έρευνας έδειξαν ότι τα δυναμικά οπτικά χαρακτηριστικά της λεύκανσης σχετίζονται άμεσα με την πυκνότητα κατάληψης του επιθηλίου από νεοπλασματικά κύτταρα, καθώς και από τις διαταραχές στον μεταβολισμό των κυττάρων αυτών.

Το οξικό οξύ (AA), διαχέεται στον επιθηλιακό ιστό και ιοντίζεται, κάτι που έχει ως αποτέλεσμα να παράγεται άλας και κατιόν υδρογόνου:

$$CH_3COOH \leftrightarrow CH_3COO^- + H^+$$

Το να διατηρηθεί το pH ουδέτερο (7-7,5) μέσα στο κύτταρο, είναι η σημαντικότερη δραστηριότητα του κυττάρου. Αντιδρώντας σε μία έγχυση οξικού οξέος, το κύτταρο αρχικά επιστρατεύει διάφορους σχετικά γρήγορους και αντιστρεπτούς μηχανισμούς που καταναλώνουν ή αποβάλουν τα παραγόμενα από τη διάσταση κατιόντα υδρογόνου, έτσι ώστε να αποκατασταθεί το pH. Επομένως, το pH επιστρέφει σταδιακά στην κανονική του τιμή.

Έχει βρεθεί πως η ένταση της λεύκανσης αυξάνεται ανάλογα με τη σοβαρότητα της νεοπλασίας (I,II,III). Επίσης, όσο μεγαλύτερο βαθμό νεοπλασίας έχουμε, τόσα περισσότερα νεοπλασματικά κύτταρα έχουμε. Άρα, η ένταση της λεύκανσης εξαρτάται από το ποσοστό των νεοπλασματικών κυττάρων που υπάρχουν στο επιθήλιο.

Τα νεοπλασματικά κύτταρα ζουν σε ένα χαμηλότερο εξωκυττάρικό pH συγκριτικά με τα υγιή κύτταρα, ενώ το αντίστοιχο εσωκυτταρικό είναι το ίδιο και διατηρείται σε ουδέτερα επίπεδα. Το οξικό οξύ, λοιπόν, στα νεοπλασματικά κύτταρα λόγω χαμηλής τιμής του pH παραμένει σε μεγαλύτερο βαθμό αδιάστατο απ' ότι στο περιβάλλον των φυσιολογικών κυττάρων και έτσι μπορεί επιλεκτικά να περάσει, με παθητική διάχυση την κυτταρική μεμβράνη των παθολογικών κυττάρων.

Η νεοπλασία προκαλεί πολλές δομικές και λειτουργικές διαφοροποιήσεις στους ιστούς. Χαρακτηριστικά, μία ανισορροπία του pH εμφανίζεται μεταξύ εσωκυττάριου και εξωκυττάριου χώρου, κυρίως λόγω της αυξημένης παραγωγής γαλακτικού οξέος από τα καρκινικά κύτταρα. Έτσι, ενώ το κυτταροπλασματικό ph διατηρείται ουδέτερο, το εξωκυτταρικό μειώνεται. Αυτή η κατάσταση επιτρέπει τη συγκέντρωση ασθενών οξέων εσωκυτταρικά, όπου και ιοντίζονται εκτενώς. Την ίδια στιγμή, λόγω του ότι η κυτταρική μεμβράνη είναι αδιαπέραστη σε φορτισμένα σωματίδια, τα ιόντα παγιδεύονται μέσα στο κύτταρο ('ion trapping') όπου και αλληλεπιδρούν με το περιβάλλον. Για το λόγο αυτό, η συγκέντρωση του AA εσωκυτταρικά αυξάνεται 'ανώμαλα'. Την ίδια στιγμή, η διάχυσή του μέσα στον ιστό και η διάστασή του σε ιόντα Ac^{-} και H^{+}, προκαλούν μορφολογικές αλλαγές στις πρωτεΐνες του πυρήνα. Κάτι τέτοιο μεταβάλλει το δείκτη διάθλασης του πυρήνα, μετατρέποντάς το επιθήλιο από διαφανές σε αδιαφανές ενεργοποιώντας το in vivo, κλινικά μετρήσιμο, σκεδαζόμενο οπτικό σήμα. Όπως προαναφέρθηκε, για να διαχειριστεί το επιπλέον φορτίο από ιόντα, το κύτταρο επιστρατεύει μικρής και μεγάλης διάρκειας μηχανισμούς και το φαινόμενο αντιστρέφεται. Επειδή όμως, το τραχηλικό επιθήλιο είναι στρωματοποιημένο, το AA διαχέεται επίσης κατά μήκος του ιστού cross-section (Εγκάρσια τομή), που σημαίνει ότι αυτές οι διαδικασίες επαναλαμβάνονται στα παρακάτω νεοπλαστικά επιθηλιακά επίπεδα, αυξάνοντας τον αριθμό των σκεδαζόμενων φωτονίων. Τελικά, το AA καταναλώνεται ή αποβάλλεται στο στρώμα.

2.3 ΕΞΗΓΗΣΗ ΤΟΥ ΒΙΟΛΟΓΙΚΟΥ ΜΟΝΤΕΛΟΥ

Τα επιθήλια είναι μια ομάδα ιστών η οποία με σπάνιες εξαιρέσεις, καλύπτει όλη την επιφάνεια του σώματος, τις κοιλότητες και τα αγγεία. Επίσης, τα επιθήλια, λειτουργούν ως σύνδεσμος μεταξύ διαφορετικών βιολογικών διαμερισμάτων. Με αυτόν τον τρόπο παίρνουν μέρος σε πολλές διαφορετικές δραστηριότητες όπως στην επιλεκτική διάχυση, στην απορρόφηση ή στην έκκριση και στη φυσική προστασία. Κατηγοριοποιούνται σύμφωνα με το κυτταρικό τους σχήμα, τον αριθμό των επιπέδων, το κυτταρικό περίβλημα και την περιεκτικότητα σε δερατίνη. Επίσης, όλα τα επιθήλια έχουν μία μεμβράνη στο κάτω μέρος τους, της οποίας το πάχος δεν είναι σταθερό, η οποία τα διαχωρίζει από τους ιστούς που βρίσκονται πιο κάτω από αυτά.

Ο τράχηλος της μήτρας είναι το κομμάτι της μήτρας το οποίο βρίσκεται στο πάνω μέρος του κόλπου. Καλύπτεται από δύο ειδών επιθηλιακούς ιστούς: από το στρωματοποιημένο λεπιδώδες επιθήλιο και από το κιονοειδές επιθήλιο.

Η παραπάνω εικόνα, παρουσιάζει τη βασική δομή του τραχηλικού επιθηλίου μαζί με τις δομικές αλλαγές που προκαλούνται κατά τη διάρκεια της εξέλιξης της ασθένειας. Στον υγιή ιστό, ο ρυθμός παραγωγής και πολλαπλασιασμού των κυττάρων είναι ελεγχόμενος, ενώ εκεί που παρουσιάζεται όγκος, οι μηχανισμοί ελέγχου διαταράσσονται και τα κύτταρα μπορεί να μεγαλώνουν με αφύσικα γρήγορους ρυθμούς ή να υποβάλλονται σε αφύσικες διαδικασίες διαφοροποίησης. CIN είναι ένα συνεχές φάσμα αλλαγής, σύμφωνα με το οποίο το επιθήλιο μπορεί να γίνει προοδευτικά πιο ανώμαλο στην εμφάνιση ή νεοπλασματικό. Επίσης, η ανάπτυξη νεοπλασίας συνδέεται με την έλλειψη διαφοροποίσης και διαστρωμάτωσης το οποίο έχει σαν αποτέλεσμα να ομογενοποιείται ο ιστός σε ολο το πάχος, όσο αναφορά την κυτταρική κατανομή μεγέθους και το σχήμα. Το μέγεθος των κυττάρων μειώνεται, ενώ το μέγεθος του πυρήνα αυξάνεται, έχοντας σαν αποτέλεσμα την αύξηση της αναλογίας κυττάρου-πλάσματος και του εξωκυττάριου χώρου. Δεν υπάρχουν ποσοτικές περιγραφές των βαθμών CIN και επίσης, η ιστολογική διάγνωση είναι σε μεγάλο βαθμό υποκειμενική. Γενικά, η ιστιολογική διαβάθμιση του CIN σχετίζεται σε μικρό βαθμό με το τμήμα του συνολικού πάχους του επιθηλίου που έχει προσβληθεί από νεοπλασία. Για παράδειγμα, ο ιστός ο οποίος θα διαγνωστεί ότι εμφανίζει CIN I θα είναι φυσιολογικός στα δύο τρίτα του πάχους του στο πάνω μέρος του, αλλά μεγεθυμένα κύτταρα μπορεί να είναι ορατά στο ένα τρίτο που μένει. Τα χαμηλότερα δύο τρίτα του επιθηλίου που παρουσιάζει CIN II μπορεί να παρουσιάζουν ανωμαλίες, ενώ στην περίπτωση του CIN III, μπορεί να παρουσιάζονται ανωμαλίες στην κυτταρική δομή σε όλο το πάχος του επιθηλίου.

ΣΦΙΧΤΟΙ ΔΕΣΜΟΙ ΚΑΙ ΧΑΜΟΔΕΣΜΟΙ.

Μεμονωμένα κύτταρα συνθέτουν επιθήλια και άλλοι οργανωμένοι ιστοί κολλάνε μεταξύ τους και στο περιβάλλον εξωκυτταρικό στρώμα, χρησιμοποιώντας συγκεκριμένου τύπου δομές οι οποίες ονομάζονται δεσμοί. Αυτοί οι δεσμοί παίζουν σημαντικό ρόλο στον έλεγχο της κίνησης των ιόντων και μικρών μορίων μεταξύ των κυττάρων.

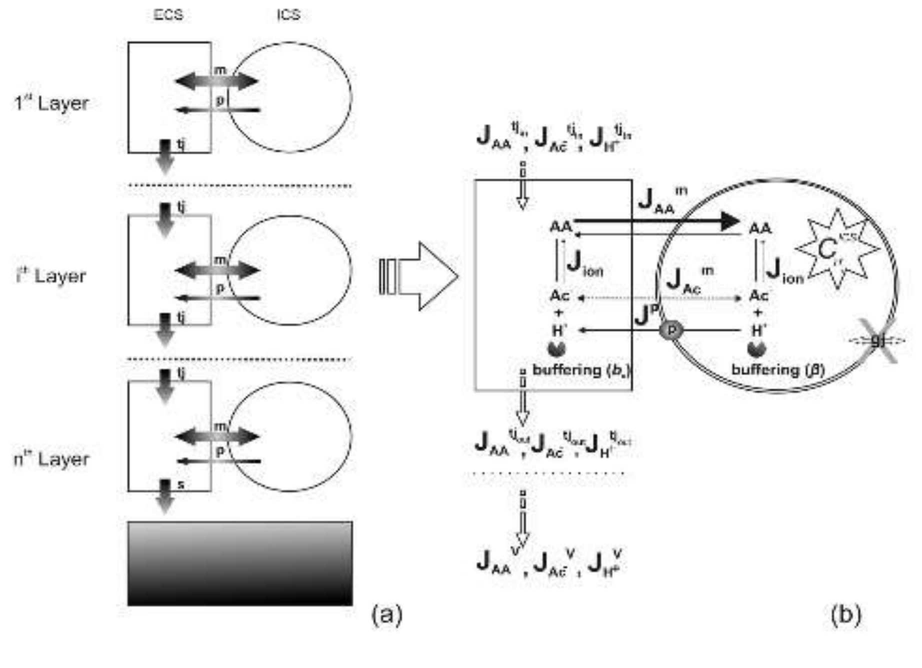
Οι σφιχτοί δεσμοί συνδέουν διπλανά επιθηλιακά κύτταρα και ελέγχουν τις παρακυτταρικές διαδρομές μεταφοράς., Οι πόροι των σφιχτών δεσμών είναι δυναμικοί και συμπεριφέρονται ως κανάλια ιόντων. Έχει βρεθεί ότι το νεοπλαστικό επιθήλιο παρουσιάζει διαρρέοντα βιολογικά εμπόδια, οι σφιχτοί δεσμοί μειώνονται και τα δεσμοσώματα χαλαρώνουν κατά την καρκινογένεση. Σαν αποτέλεσμα, η δομή του επιθηλίου παρουσιάζει κενά, επιτρέποντας μεγαλύτερη διάχυση ιόντων και μορίων.

Οι χασμοδεσμοί ελέγχουν τις εσωκυτταρικές διαδρομές μεταφοράς μέσα στο επιθήλιο. Αυτοί οι δεσμοί βοηθούν στην ενσωμάτωση των μεταβολικών δραστηριοτήτων όλων των κυττάρων σε ένα ιστό (μεταβολική συνεργασία) επιτρέποντας την απευθείας διέλευση ιόντων και μικρών μορίων από το κυτταρόπλασμα του ενός κυττάρου σε ένα άλλο.

ΔΙΑΜΕΡΙΣΜΑΤΙΚΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΤΡΑΧΗΛΙΚΟΥ ΕΠΙΘΗΛΙΟΥ

Το τραχηλικό επιθήλιο, μπορεί να περιγραφεί σαν ένα πολυεπίπεδο επιθήλιο, που αλλάζει και δομικά και λειτουργικά όταν παρουσιαστεί CIN. Λογικά υποθέτουμε ότι κάθε κύτταρο συνιστά ένα σύστημα με δύο διαμερίσματα. Το πρώτο διαμέρισμα σχετίζεται με το κύτταρο αυτό καθεαυτό (εσωκυτταρικό) και το δεύτερο διαμέρισμα με τον αντίστοιχο εξωκυττάριο χώρο (εξωκυτταρικό). Ακολουθώντας μία ομοιόμορφη κατανομή κυττάρων σε ένα διαπλαστικό επίπεδο και έχοντας υπ’ όψιν ότι η ανάπτυξη του CIN είναι ανάλογη του αριθμού των επιπέδων, μοντελοποιούμε μαθηματικά μία στοίβα από συστήματα που αποτελούνται από κύτταρα με δύο διαμερίσματα, ο ακριβής αριθμός των οποίων προκύπτει από τα επίπεδα που καθορίζονται από το βαθμό CIN.

Όπως φαίνεται στην εικόνα 8**,** πραγματοποιείται ανταλλαγή μάζας με παθητική διάχυση, μεταξύ των εσωκυττάριων και εξωκυττάριων διαμερισμάτων, διαμέσου της κυτταρικής μεμβράνης (m) και ενεργή ροή πρωτονίων (p). Η επικοινωνία μεταξύ των επιπέδων πραγματοποιείται με παθητική ηλεκροοσμωτική διάχυση, μέσω της μεμβράνης των σφιχτών δεσμών (tj), στις ανώτερες και κατώτερες περιοχές του κάθε εξωκυττάριου χώρου. Όταν τα μόρια περάσουν από το τελευταίο επίπεδο, διαχέονται στο στρώμα (s). Στην εικόνα 7, παρουσιάζεται μία πιο λεπτομερής ανάλυση ενός συστήματος από ένα επίπεδο, δύο διαμερισμάτων. Στο σχήμα αυτό περιγράφονται όλες οι ροές του συστήματος αυτού (J).



Σχήμα 7.

Οξικό οξύ, Ac^- και πρωτόνια (H^+), εισέρχονται (in) στον εξωκυττάριο χώρο του i-στού επιπέδου, αφού περάσουν από τους σφιχτούς δεσμούς ($J_{AA_i}^{tj_{in}}$,$J_{Ac_i}^{tj_{in}}$,$J_{II_i}^{tj_{in}}$). Εδώ το οξικό οξύ διίσταται (J_{ion}), ανάλογα με το εξυκυτταρικό pH (pH_{ES}) και την ικανότητα κατανάλωσης υδρογόνου εξωκυτταρικά (b_{ECS}). Έπειτα, μόρια οξικού οξέος διαπερνούν παθητικά την κυτταρική μεμβράνη. Το pH στον εξωκυττάριο χώρο πέφτει και μπαίνουν σε εφαρμογή οι μηχανισμοί κατανάλωσης υδρογόνου. Τα ιόντα Ac^- και H^+ ενώνονται και δίνουν πάλι μόρια οξικού οξέος.

Το οξικό οξύ που πέρασε από τον εξωκυττάριο χώρο μέσω της κυτταρικής μεμβράνης στο σχεδόν ουδέτερο εσωκυττάριο περιβάλλον, διίσταται. Πέφτει, λοιπόν και το pH στον εσωκυττάριο χώρο και μπαίνουν σε λειτουργία οι μηχανισμοί κατανάλωσης υδρογόνου. Λόγω της διαφοράς των τιμών του pH μεταξύ εσωκυττάριου και εξωκυττάριου χώρου, παράγεται ένα χαμηλό πεδίο αγωγιμότητας το οποίο επιτρέπει διάχυση Ac^- ($J_{Ac_i}^{m}$). Λόγω της αύξησης της συγκέντρωσης H^+ εσωκυτταρικά ($C_{II_i}^{ICS}$), η οποία θεωρείται ότι συνδέεται με το φαινόμενο της λεύκανσης, μπαίνουν σε λειτουργία οι αντλίες του κυττάρου (J^P), από τις οποίες ιόντα υδρογόνου περνάνε στον εξωκυττάριο χώρο. Τα ιόντα Ac^- και H^+ που έχουν μείνει στον εσωκυττάριο χώρο, ενώνονται και δίνουν πάλι μόρια οξικού οξέος. Επίσης, ενεργοποιούνται οι αντλίες του κυττάρου που μεταφέρουν το οξικό οξύ από τον εσωκυττάριο στον εξωκυττάριο χώρο και το εσωκυττάριο pH επανέρχεται στα φυσιολογικά του επίπεδα. Από την άλλη πλευρά, το μη ιονισμένο οξικό οξύ του εξωκυττάριου χώρου, που δεν πέρασε στον εσωκυττάριο μειώνεται καθώς περνάει στα επόμενα επίπεδα. Τα μόρια και τα ιόντα που υπάρχουν εξωκυτταρικά μαζί με αυτά που βγήκαν από το κύτταρο μεταφέρονται στα επόμενα δισπλαστικά επίπεδα ($J_{AA_i}^{tj_{out}}$,$J_{Ac_i}^{tj_{out}}$,$J_{II_i}^{tj_{out}}$). Αφού συνεχιστεί η ίδια βιολογική διαδικασία και στα επόμενα επίπεδα του επιθηλίου, το οξικό οξύ καταναλώνεται κυρίως λόγω του buffering και η ποσότητα που μένει (αν μείνει κάποια), αφού φτάσει στο στρώμα, αφαιρείται από το επιθήλιο (J_{AA}^{V},J_{Ac}^{V},J_{II}^{V}).

2.3 ΕΞΗΓΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ ΜΑΘΗΜΑΤΙΚΑ

ΒΙΟΔΕΙΚΤΕΣ ΚΑΙ ΠΑΡΑΓΩΓΗ ΔΥΝΑΜΙΚΩΝ ΟΠΤΙΚΩΝ ΣΗΜΑΤΩΝ

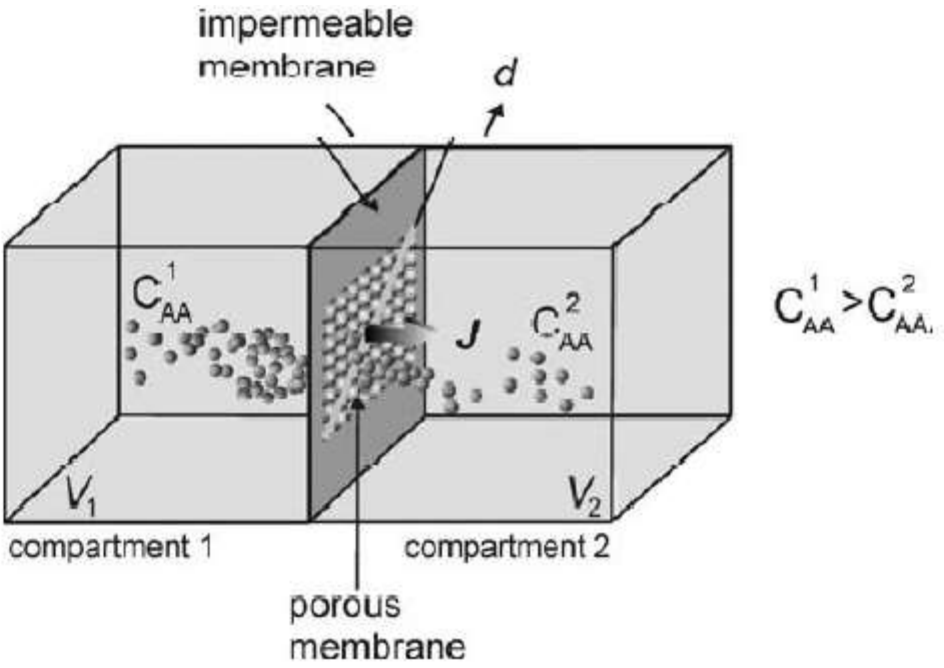
Τα τελευταία χρόνια έχει παρουσιαστεί μεγάλο ενδιαφέρον για τη χρήση οπτικών τεχνολογιών με την ανάπτυξη πρωτοπόρων εξωγενών παραγόντων, σχεδιασμένοι να εντοπίζουν συγκεκριμένα μοριακά καρκινικά και άλλων ασθενειών χαρακτηριστικά. Η οπτική μοριακή απεικόνιση και χρωμοσωματική ενδοσκόπιση, είναι οι κινητήριες δυνάμεις για την ανάπτυξη νέων και αποτελεσματικών μεθόδων που θα μπορούσαν να βελτιώσουν την ακρίβεια των in-vivo διαγνώσεων σε διάφορα όργανα. Μέσω της μέτρησης των παραγόμενων οπτικών σημάτων (χρωματικές αλλαγές) είναι δυνατό να αποκτήσουμε πολύτιμες πληροφορίες για τη δομική και λειτουργική κατάσταση του ιστού-στόχου. Με τη χρήση ενός αραιού διαλύματος οξικού οξέος ως βιοδείκτη και ένα πολυτροπικό σύστημα απεικόνισης, ειδικά σχεδιασμένο για ταυτόχρονη δυναμική και οπτική παρουσίαση επιχειρείται μη-επεμβατική διάγνωση επιθηλιακών νεοπλασιών. Αρχικά έχει εντοπιστεί μία ομάδα δομικών και λειτουργικών βιολογικών παραμέτρων που σχετίζονται με την ανάπτυξη νεοπλασίας. Βάση αυτών ένα διαμερισματικό μοντέλο έχει αναπτυχθεί για όλα τα επιθηλιακά φαινόμενα μεταφοράς και προβλέπει τις τα δυναμικά οπτικά χαρακτηριστικά κάθε πιθανού συνδυασμού τιμών των βιολογικών παραμέτρων.

Θεωρητικά οι επιθηλιακοί ιστοί εμφανίζονται σε κυτταρικούς σχηματισμούς ενός ή περισσοτέρων στρωμάτων, ιστοί που έχουν την τάση να ομογενοποιούνται κατά την ανάπτυξη της νεοπλασίας. Κάθε στρώμα κυττάρων θεωρείται μια οντότητα από ανεξάρτητα, κινητικά ομοιογενή τμήματα: ένα εξωκυττάριο (ES) και ένα εσωκυττάριο (IS). Η διάχυση μεταξύ των τμημάτων και των στρωμάτων πραγματοποιείται πορώδους μεμβράνης του κυττάρου και των σφιχτών δεσμών, αντίστοιχα. Υποθέτοντας ότι η κινητική του AA βρίσκεται σε ισορροπία διαρκώς, η σταθερή κατάσταση (ss) παθητικής διάχυσης μεταξύ δύο διαμερισμάτων με όγκους V_1, V_2 χωρισμένων με μία πορώδη μεμβράνη (p_m) επιφάνειας $A_{pm,}$ διέπεται από τον νόμο των Fick και Goldman-Hodgik-Katz:

$$J_{I\to II}^{u} = P_u \left(C_{II}^{u} - C_{I}^{u}\right)$$

$$J_{I\to II}^{ch} = \frac{ln\varepsilon}{\varepsilon - 1} P_{ch}\left(C_{I}^{ch} - C_{II}^{ch}\varepsilon\right)$$

Όπου J η ροή της αφόρτιστης (u) ή φορτισμένης (ch) διαλυμένης ουσίας από το διαμέρισμα I στο II, $P_s = \frac{D_s k_s}{d}$ περιγράφει τη διαπερατότητα της διαλυμένης ουσίας S μέσω της πορώδους μεμβράνης που χωρίζει δύο διπλανά διαμερίσματα με το D_s που δηλώνει το συντελεστή διάχυσης της S μέσω αυτής της μεμβράνης, το λάδι k: νερό partition coefficient και d το μήκος του πόρου, C η συγκέντρωση και ε=exp(-z_sFΔV/RT), όπου z το σθένος του ιόντος, F η σταθερά Faraday, ΔV η τάση διαμεμβράνης, R η σταθερά των αερίων και T η θερμοκρασία.



Σχήμα 8.

Δραστική εκροή πρωτονίων, από την άλλη, ενεργοποιείται από οποιαδήποτε μη φυσιολογική διαταραχή ενδοκυτταρικού pH και η προκληθείσα εκροή διαμορφώνεται δυναμικά και από το ενδοκυτταρικό και εξωκυτταρικό pH. Η δραστική εκροή πρωτονίων θεωρείται ως τμηματικά γραμμική συνάρτηση των pH_{IS} και pH_{ES} :

$$J_p = \begin{cases} \dfrac{pH_{IS}}{6.7-6} K_p, & pH_i < 6.7 \\ K_p, & pH_i = 6.7 \\ \left[1 - \dfrac{pH_{IS} - 6.7}{pH_{IS,0} - 6.7}\right] K_p, & pH_i > 6.7 \end{cases}$$

Η ροή J_s για όλους τους μηχανισμούς εκροής πρωτονίων του κυττάρου, $pH_{IS,0}$ το αρχικό pH_{IS} πριν την εφαρμογή του ΑΑ, 6,7 το κατώφλι pH_{IS} στο οποίο ο ρυθμός εκροής μεγιστοποιείται και K_p η μέγιστη εκροή κατά μία εκθετική συνάρτηση τουpH_{ES}. Στη συνέχεια, χρησιμοποιώντας τις (1)-(3) το ακόλουθο διαφορικό σύστημα προκύπτει για το κάθε νεοπλασματικό στρώμα:

$$\dot{C}_{IS}^{TA} = a^{-1}\left(J_m^{AA_i} + J_m^{Ac^-}\right),$$

$$\dot{C}_{IS}^{H_i^+} = -ln10^{C_{IS}^{H_i^+}} (\alpha\beta_{IS})^{-1} \left(q_{IS}^i J_m^{AA_i} - w_{IS}^i J_m^{Ac^-} - J_p^{H_i^+}\right),$$

$$\dot{C}_{ES}^{TA_i} = b^{-1} J_m^{AA_i} + b^{-1} J_m^{Ac^-} + a^{-1} g_{TJ} J_{TJ}^{AA_i} + a^{-1} g_{TJ} J_{TJ}^{Ac^-} \qquad \text{και}$$

$$\dot{C}_{ES}^{H_i^+} = -ln10^{C_{ES}^{H_i^+}} \beta_{ES}^{-1} [q_{ES}^i\left(b^{-1} J_m^{AA_i} + a^{-1} g_{TJ} J_{TJ}^{AA_i}\right) -$$

$$-w_{ES}^i(b^{-1}J_m^{Ac^-} + a^{-1} g_{TJ} J_{TJ}^{Ac^-}) + b^{-1} J_p^{H_i^+} + a^{-1} g_{TJ} J_{TJ}^{AH_i^+}],$$

Όπου η τελεία (·) συμβολίζει την παράγωγο του χρόνου, TA η συνολική συγκέντρωση AA σε ιονισμένη και μη μορφή, i το i-στο νεοπλασματικό στρώμα(i=1,2,…N), α και β οι γραμμικές διαστάσεις των IS και ES όπου έχει θεωρηθεί κυβική και ορθογώνια γεωμετρική δομή διαμερίσματος αντίστοιχα, β η ρυθμιστική δύναμη, q και w δυναμικές σταθερές ιονισμού του AA συμπεριλαμβανομένης της αυτό-καύσης του (self-burning effect), g_{TJ} εκφράζει τις γεωμετρικές διαστάσεις των πόρων των σφιχτών δεσμών και J_{TJ} η συνολική παθητική παρακυτταρική ροή που παρουσιάζεται μεταξύ των διαδοχικών στρωμάτων μέσω των σφιχτών δεσμών. Εδώ θα πρέπει να σημειωθεί ότι για τις εξισώσεις των στρωμάτων (3) και (4) καταργούνται και για το τελευταίο στρώμα η παρακυτταρική εκροή αντικαθίσταται από τον όρο $K_V C$, όπου K_V η διαπερατότητα στο όριο μεταξύ του ιστού του επιθηλίου και του στρώματος. Το σύστημα διαφορικών εξισώσεων που αναπτύχθηκε λύνεται αριθμητικά με τη βοήθεια μικτών-ακεραίων, μη-γραμμικών αλγόριθμων στο περιβάλλον της Matlab. Οι αλγόριθμοι έχουν σαν είσοδο τις βιολογικές περιοχές τιμών των παραμέτρων και σαν έξοδο την ενδοκυτταρική AA και H^+παροδική συγκέντρωση σε δεδομένες χρονικές στιγμές και του ES και του IS. Όπως έχει ήδη αναφερθεί, η παροδική αλλαγή στην ενδοκυττάρια συγκέντρωση ορίζει τα δυναμικά χαρακτηριστικά του οπτικού αποτελέσματος και για το λόγο αυτό μόνο η συγκεκριμένη έξοδος του μοντέλου χρησιμοποιείται για τη σύγκριση προβλέψεων με πειραματικά δεδομένα.

ΠΑΡΑΜΕΤΡΟΣ	ΜΟΝΑΔΑ ΜΕΤΡΗΣΗΣ	ΠΕΡΙΓΡΑΦΗ
pHECS		pH του εξωκυττάριου χώρου
pHICS		pH του εσωκυττάριου χώρου
ICS volume	μm^3	Όγκος εσωκυττάριου χώρου
ECS volume	μm^3	Όγκος εξωκυττάριου χώρου
Cell area (Acell)	μm^2	Εμβαδό κυτταρικής μεμβράνης
TJ area (Atj)		Εμβαδό μεμβράνης σφιχτών δεσμών (tight junctions)
PAA^m	μm/s	Ταχύτητα με την οποία το AA διαπερνά την μεμβράνη
PAc m	pm/s	Ταχύτητα με την οποία το Ac- διαπερνά την μεμβράνη
PAA tj	μm/s	Ταχύτητα με την οποία το AA διαπερνά τους σφιχτούς δεσμούς (tight junctions)
PAc tj	μm/s	Ταχύτητα με την οποία το Ac- διαπερνά τους σφιχτούς δεσμούς

		(tight junctions)
PH tj	μm/s	Ταχύτητα με την οποία το H+ διαπερνά τους σφιχτούς δεσμούς (tight junctions)
Intracellular buffering (*β*)	mM	Δύναμη με την οποία καταναλώνονται τα H+ στον εσωκυττάριο χώρο
Pumping rate (*kp*)	μM/s	Ρυθμός αποβολής ιόντων υδρογόνου από το κυτταρόπλασμα μέσω των αντλιών
ΔVm , ΔVtj	mV	Τάση ηλεκτρικού πεδίου μεταξύ των δύο πλευρών της μεμβράνης και αντίστοιχα για τους σφιχτούς δεσμούς (tight junctions)
AA ionization constant at 37oCn (K*a*)		AA ionization constant at 37oC (K*a*)
F	C.mol-1	Σταθερά του Faraday
R	J.mol-1K-1	Σταθερά αερίου
T	K	Θερμοκρασία
zAc		Σθένος του Ac
zH		Σθένος του H

Πίνακας 1.

Προκειμένου να μειωθεί ο αριθμός των παραμέτρων χρησιμοποιήθηκαν οι: η Global Sensitivity Analysis (GSA), parameter identifiability και parameter estimability μέθοδοι. Το περιβάλλον θεωρήθηκε εξιδανικευμένο, in silico, τα δεδομένα χωρίς θόρυβο και το μοντέλο ιδανικό και πλήρες. Πιο συγκεκριμένα η GSA παρέχει μία εκτίμηση της ευαισθησίας της εξόδου του μοντέλου για διάφορες διακυμάνσεις της εισόδου. Στη συνέχεια χρησιμοποιήθηκαν τα αποτελέσματα της GSA σαν είσοδο για την parameter identifiability όπου και κατατάσσει μία παράμετρο σαν αναγνωρίσιμη όταν η ευαισθησία της δεν είναι μηδενική ή ισότιμη με κάποιας άλλης παραμέτρου. Σαν είσοδο τα αποτελέσματα της GSA χρησιμοποιήθηκαν για την parameter estimability, όπου από αυτή τη μέθοδο καθορίσαμε το επίπεδο των αλληλεπιδράσεων ή/και συσχετίσεων μεταξύ των παραμέτρων του σετ.

Global Sensitivity Analysis

Στην GSA κάθε παράμετρος εισόδου αντιμετωπίζεται σαν τυχαία μεταβλητή που ακολουθεί μία κατανομή στην οποία τα δεδομένα δειγματοληπτούνται τυχαία. Σε μια τέτοια βάση, ισχυρά στατιστικά συμπεράσματα μπορούν να αντληθούν από την «αποσύνθεση» της διακύμανσης της εξόδου V(Y) σε όρους αύξουσας διάστασης:

$$V(Y) = \sum_i V + \sum_i \sum_j V_{ij} + \sum_i \sum_j \sum_l V_{ijl} + \cdots + V_{1,2,\ldots,k}$$

όπου, δεδομένου ενός k-διαστάσεων (παράγοντες) χώρο(**πινακας),** $V_{i,j,..,k}$ η διακύμανση της εξόδου (V) των υπό όρων προσδοκία (E), λόγω της μοναδικής παραμέτρου διακύμανσης (V_1) ή λόγω των συνδυασμένων διακυμάνσεων πολλαπλών παραμέτρων. Στην πρώτη περίπτωση $V_1 = V(E(Y|X_i))$, ενώ στη δεύτερη περίπτωση και δύο αλληλεπιδρούσες παραμέτρους $V_{ij} = V\left(E\left(Y\middle|X_i, X_j\right)\right) - V_i - V_j$. Με βάση το συμβολισμό, το μέτρο:

$$S_i = \frac{V_i}{V(Y)}$$

που ονομάζεται ευαισθησία πρώτης-τάξης (FOS) επίδραση του παράγοντα X_i στην έξοδο Y. Αυτό ορίζει, ποσοτικά, την συμβολή του παράγοντα X_i στη διακύμανση εξόδου, μόνο και αυτό αποτελεί βάση για την κατάταξη της παραμέτρου (παράγοντας ρύθμισης προτεραιότητας). Ένα άλλο μέτρο ευαισθησίας που αντιπροσωπεύει την επίδραση μιας δεδομένης παραμέτρου στην έξοδο του μοντέλου, λαμβάνοντας υπ' όψιν το μέτρο ευαισθησίας της πρώτης-τάξης και των μεγαλύτερων-τάξεων (TOS) δείκτης και δίνεται από τον τύπο:

$$S_{Ti} = S_i + S_{ij} + \cdots + S_{ij\ldots k}$$

Όπου $i<j<\ldots<k$. Αν αυτό το χαρακτηριστικό είναι, για μια δοσμένη παράμετρο, μικρότερο από ένα κατώτερο όριο, τότε η επίδρασή της στην έξοδο του μοντέλου αναμένεται να είναι φυσιολογική υποδεικνύοντας ότι αυτός ο παράγοντας δεν έχει καμία επίδραση στην διακύμανση της εξόδου. Σε αυτή την περίπτωση φαίνεται να μπορεί να μειωθεί το σετ των παραμέτρων του μοντέλου.

IDENTIFIABILITY OF MODEL PARAMETERS

Ο γενικός σκοπός της identiafiability analysis είναι να ορίσει ποια από τις παραμέτρους του μοντέλου θα μπορεί ενδεχομένως να προκύψει/αποκτηθεί από τα διαθέσιμα δεδομένα εισόδου/εξόδου και ποια θα είναι αδύνατο να εκτιμηθεί. Επίσημα, αυτό μπορεί να γίνει μόνο με ξεκάθαρο ορισμό των μεταξύ σχέσεων των παραμέτρων του μοντέλου και της εξόδου του. Ωστόσο, για υπολογιστικά μοντέλα όπως και αυτό, αυτό είναι ανέφικτο στις περισσότερες περιπτώσεις. Η GSA μπορεί να χρησιμοποιηθεί για την αριθμητική διερεύνηση του κατά πόσο οι συνδυασμοί των παραμέτρων του μοντέλου προκαλούν την εκάστοτε μετρούμενη έξοδο. Από την εξίσωση **()κ΄ξλξδσλξφς,** έστω ότι $S_{Ti}|_t$ χαρακτηρίζει το γενικό δείκτη ευαισθησίας της παραμέτρου i τη χρονική στιγμή t, όπου i= 1…k και ο πίνακας:

$$\mathbf{S}_D = \begin{bmatrix} S_{T_1}\big|_{t_1} & \cdots & S_{T_k}\big|_{t_1} \\ \vdots & \ddots & \vdots \\ S_{T_1}\big|_{t_s} & \cdots & S_{T_k}\big|_{t_s} \end{bmatrix}$$

ο δυναμικός GSA πίνακας της εξόδου του μοντέλου για κάθε μία από τις j παραμέτρους σε διακριτούς χρόνους $t_1 \ldots t_s$. Στη συνέχεια από την πίνακας συσχετισμένων παραμέτρων M μπορεί να υπολογιστεί με τη βοήθεια της κατάταξης του Spearman των συντελεστών συσχέτισης, που παρέχει την αντιστοιχία μεταξύ δύο στηλών. Η παράμετρος με συσχέτιση με όλες τις άλλες παραμέτρους μεταξύ -1 και 1 θεωρείται αναγνωρίσιμη, ενώ η παράμετρος που έχει τουλάχιστον μία συσχέτιση ακριβώς -1 ή 1 με μία άλλη παράμετρο δεν είναι αναγνωρίσιμη. Με άλλα λόγια, παράμετροι που επηρεάζουν την έξοδο του μοντέλου με ακριβώς τον ίδιο ή αντίθετο τρόπο δεν μπορούν να προσδιοριστούν. Έτσι το αρχικό σετ παραμέτρων, p μπορεί να μειωθεί στο προσδιορίσιμο σύνολο παραμέτρων p_I, μήκους I, με το να υπολογιστεί κατ’ επανάληψη το M και να απομακρύνεται μία μη προσδιορίσιμη παράμετρος κάθε φορά, μέχρι ότου να μην απομείνουν μη προσδιορίσιμες παράμετροι.

PARAMETER ESTIMABILITY.

Η δυνατότητα για την ουσιαστική εκτίμηση μιας παραμέτρου του μοντέλου αυξάνεται όταν η παράμετρος αυτή επηρεάζει την έξοδο του μοντέλου άμεσα κυρίως και όχι έμμεσα, μέσω των αλληλεπιδράσεών της με τις άλλες παραμέτρους. Ένας τρόπος να εκφραστεί είναι με τον ποσοτικό καθορισμό του estimability, σαν λόγος του πρώτου δείκτη ευαισθησίας μιας δεδομένης παραμέτρου S_i προς το άθροισμα του ολικού δείκτη ευαισθησίας ολόκληρου του σετ παραμέτρων. Όπως έχει ήδη αναφερθεί τα μέτρα ευαισθησίας αναμένεται να είναι συναρτήσει του χρόνου, τα οποία επίσης αναμένεται να εφαρμόζονται στον δείκτη ή λόγο estimability (ER). Αυτός ο δείκτης ορίζεται ως:

$$ER(t) = \frac{S_i(t)}{\sum_{j=1}^{k} ST_j(t)}$$

Μπορούμε εύκολα να συμπεράνουμε ότι το έργο της εκτίμησης μιας αναγνωρίσιμης παραμέτρου είναι απλοποιημένο όταν ο FOS δείκτης της βρίσκεται σε υψηλή θέση κατάταξης και όταν επηρεάζει άμεσα την έξοδο του μοντέλου, πράγμα που συμβαίνει όταν ο ER δείκτης πλησιάζει τη μέγιστη τιμή του.

σ.σ. Προς απλοποίηση θεωρήσαμε ομοιόμορφη κατανομή στη διακύμανση των παραμέτρων .

PARAMETERS	**DESCRIPTION**	**LOWEST VALUE**	**HIGHEST VALUE**
a	IS γραμμική διάσταση	10	20
b	ES γραμμική διάσταση	0.1	0.8
β_{ES}	Δύναμη με την οποία καταναλώνονται τα H+ στον εξωκυττάριο χώρο	0.035	0.05
β_{IS}	Δύναμη με την οποία καταναλώνονται τα H+ στον εξωκυττάριο χώρο	0.01	0.0018
n		1	10
pH_{ES}	τιμή pH του ES	6	7
pH_{IS}	τιμή pH του IS	7	7.4
g_{TJ}	διαστάσεις πόρων σφιχτών δεσμών	0.1	10

Πίνακας 2.

Οι οχτώ παράμετροι θεωρούνται το πλήρες σετ, ενώ οι τρεις (χωρίς βέλη) απορρίπτονται λόγω από την παραπάνω μελέτη. Οι πέντε παράμετροι που απομένουν είναι δομικοί και λειτουργικοί και είναι άμεσα συνυφασμένοι με την ύπαρξη και εξέλιξη του καρκίνου.

ΚΕΦΑΛΑΙΟ 3

3.1 ΑΝΤΙΣΤΡΟΦΟ ΠΡΟΒΛΗΜΑ

Αντίστροφο πρόβλημα είναι ο όρος που χρησιμοποιείται για να περιγράψει ένα πρόβλημα που σαν στόχο έχει να μετατρέψει τις μετρήσεις σε πληροφορία σχετικά με ένα φυσικό πρόβλημα ή σύστημα. Το πρόβλημα αυτής της εργασίας, είναι η εξαγωγή πληροφορίας σχετικά με ένα δεδομένο σύνολο βιολογικών παραμέτρων από δυναμικά οπτικά δεδομένα, σε μορφή καμπυλών που εκφράζουν τη συγκέντρωση υδρογόνου στο κυτταρόπλασμα του επιθηλίου της μήτρας. Ένα μαθηματικό μοντέλο προσομοιώνει τα δυναμικά χαρακτηριστικά της καμπύλης έχοντας ως παραμέτρους εισόδου τα λειτουργικά και δομικά χαρακτηριστικά του ιστού. Ο τελικός στόχος είναι δυνατότητα ανάκτησης τιμών για τις βιολογικές παραμέτρους, δομικές και λειτουργικές, χρησιμοποιώντας σαν είσοδο στο μοντέλο μια δυναμική καμπύλη. Είναι σημαντικό να μπορούμε να εκτιμάμε με ακρίβεια τα παραπάνω χαρακτηριστικά, γιατί καταφέρνουμε καθ' αυτόν τον τρόπο, για πρώτη φορά, in silico, μη επεμβατική και άμεση διάγνωση ύπαρξης και εξέλιξης νεοπλασίας.

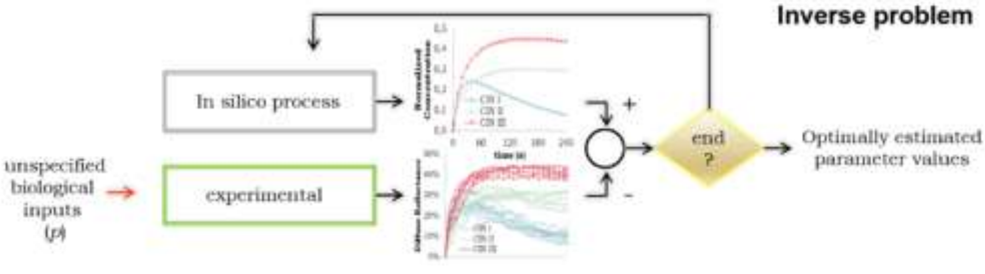
ΘΕΩΡΙΑ ΕΛΕΓΧΟΥ

Η θεωρία ελέγχου είναι ένας κλάδος της εφαρμοσμένης μηχανικής και των μαθηματικών, ο οποίος εξετάζει τη συμπεριφορά των δυναμικών συστημάτων.

IN SILICO PROCESS

Το μαθηματικό μοντέλο που περιγράφηκε στο κεφάλαιο 2, το οποίο προσομοιώνει την πειραματική διαδικασία, έχει σαν έξοδο την κανονικοποιημένη συγκέντρωση του H^+ ως προς το χρόνο, για ένα δεδομένο σετ βιολογικών παραμέτρων που λειτουργεί σαν είσοδος.

Στόχος είναι η ικανοποιητική σύγκλιση της in silico εξόδου με τα αποτελέσματα της πειραματικής διαδικασίας. Πετυχαίνοντας αυτή τη σύγκλιση, θα μπορούν να εξαχθούν από το μαθηματικό μοντέλο οι γνωστές σε μας από την πειραματική διαδικασία τιμές των βιολογικών παραμέτρων. Ικανοποιητική ακρίβεια και σύγκλιση είναι το ζητούμενο αυτής της εργασίας.



Σχήμα 7.

Στόχος μας είναι η έξοδος του μοντέλου να πλησιάσει τα πειραματικά μας δεδομένα όσο το δυνατόν περισσότερο.

Αυτό λοιπόν που πρέπει να κάνουμε είναι να ελαχιστοποιήσουμε τη μη γραμμική αντικειμενική συνάρτηση του συστήματος

$$J = \left|\left| y^{exp}(p,t) - y^{M}\left(p^{trial}, t\right)\right|\right|$$

,

όπου $f\left(\frac{dx}{dt}, x, p^{trial}, u\right)$

$$t_0 = 0$$

$$x(t_0) = x_0$$

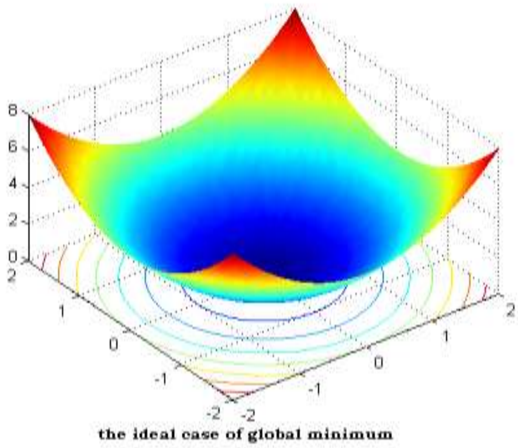
$$t \in [0,240]$$

$$p_L \leq p^{trial} \leq p_U$$

$$u = u_0,$$

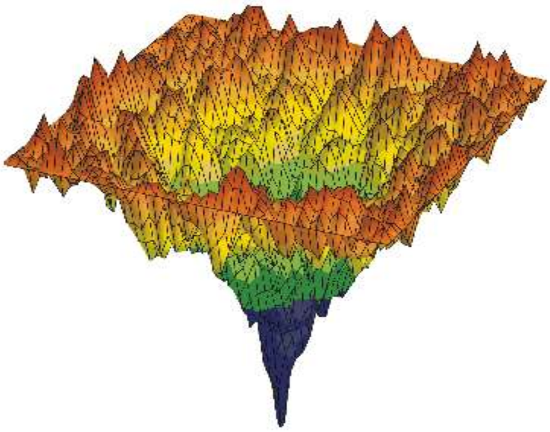
όπου f είναι η δυναμική περιγραφή του μοντέλου, x είναι οι διαφορικές μεταβλητές, p^{trial} το τελευταίο σετ p παραμέτρων που χρησιμοποιήθηκε για εκτίμηση, u η αρχική ποσότητα AA και p_L, p_U είναι τα κατώτερα και ανώτερα όρια των τιμών των παραμέτρων.

Στην ιδανική περίπτωση η αντικειμενική συνάρτηση του συστήματός μας θα έχει ένα και μοναδικό ολικό ελάχιστο.



Σχήμα 8.

Εμείς όμως δεν έχουμε να αντιμετωπίσουμε την ιδανική περίπτωση, έχουμε να διαχειριστούμε μία συνάρτηση με αρκετά τοπικά ελάχιστα. Αυτό που πρέπει εμείς να βρούμε ουσιαστικά είναι το ελάχιστο όλων των ελαχίστων. Διαφορετικά, αν δεν μπορέσουμε να βρούμε ένα ολικό ελάχιστο, θα πάρουμε την τιμή που βρίσκεται όσο το δυνατόν πιο κοντά σε αυτή του ολικού ελαχίστου. Στη ρεαλιστική περίπτωση, η συνάρτηση του συστήματός μας θα έχει την παρακάτω μορφή.



global minimum at the case of reality

Σχήμα 9.

FITTING

Στόχος μας είναι να βρεθεί μία καμπύλη ίδια με αυτή της πειραματικής διαδικασίας. Για να πετύχουμε fitting των καμπυλών αυτών, καθώς το μοντέλο που έχουμε να διαχειριστούμε είναι αρκετά πολύπλοκο, είναι πολύ σημαντικό να διατηρήσουμε τις μη γραμμικότητες που εμπεριέχονται σε αυτό.

3.2 ΑΛΓΟΡΙΘΜΟΙ ΟΛΙΚΗΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ (Global optimization)

3.2.a ΕΙΣΑΓΩΓΗ [6] [21][22][23][24]

Μία από τις βασικότερες αρχές του κόσμου μας είναι η αναζήτηση για το καλύτερο και αυτή η ιδέα αφορά τους περισσότερους τομείς στη ζωή. Στόχος είναι να βρούμε το βέλτιστο καταβάλλοντας όμως την ελάχιστη δυνατή προσπάθεια. Κάτι τέτοιο ισχύει και σε αυτή την περίπτωση. Εδώ αντιμετωπίζουμε το πρόβλημα υπολογισμού των παραμέτρων (αντίστροφο πρόβλημα) ενός μη γραμμικού δυναμικού συστήματος με στόχο την ελαχιστοποίηση της συνάρτησης κόστους. Αυτή η συνάρτηση λειτουργεί σαν κριτήριο του κατά πόσο το σετ παραμέτρων που υπολογίζονται από το μοντέλο για δεδομένες πειραματικές τιμές είναι και το πλέον κατάλληλο. Ολική βελτιστοποίηση λοιπόν, είναι το έργο της εύρεσης -από όλον τον δειγματικό χώρο- του καλύτερου σετ παραμέτρων, αυτού που σαν αποτέλεσμα δίνει τη βέλτιστη τιμή σε μια αντικειμενική συνάρτηση.

ΑΝΤΙΚΕΙΜΕΝΙΚΗ ΣΥΝΑΡΤΗΣΗ

Πιο συγκεκριμένα, μία αντικειμενική συνάρτηση $f : X \rightarrow Y \; με \; Y \subseteq R$ είναι μία μαθηματική συνάρτηση που υπόκειται σε βελτιστοποίηση.

Η περιοχή X της f καλείται δειγματικός χώρος και μπορεί να αναπαριστά κάθε τύπο στοιχείων όπως αριθμούς, λίστες, κατασκευαστικά σχέδια κ.α. Οι αντικειμενικές συναρτήσεις δεν είναι απαραίτητα αμιγώς μαθηματικές εκφράσεις, αλλά μπορούν να είναι περίπλοκοι αλγόριθμοι. Η ολική βελτιστοποίηση συμπεριλαμβάνει όλες τις πιθανές τεχνικές που μπορούν να χρησιμοποιηθούν ώστε να βρεθούν τα καλύτερα στοιχεία x^* στο X λαμβάνοντας πάντα υπ' όψιν τα κριτήρια $f \in F$. Έτσι:

(Global Minimum)Ένα ολικό ελάχιστο $\check{x} \in X$ μίας (αντικειμενικής) συνάρτησης $f : X \rightarrow R$ είναι ένα στοιχείο εισόδου με $f(\check{x}) \leq f(x) \; \forall x \; \in X$.

Γενικά λοιπόν οι μέθοδοι βελτιστοποίησης υπόκεινται σε κάποια όρια και υπακούουν σε ένα αριθμό από περιορισμούς, προκειμένου να ελαχιστοποιήσουν, όπως στην περίπτωσή μας,(ή μεγιστοποιήσουν) την αντικειμενική συνάρτηση.

Αυτά τα προβλήματα αντιμετωπίζουν γενικά δυσκολία στην εύρεση του βέλτιστου ενώ αν προσπαθήσουμε μέσω των τυπικών μεθόδων για βελτιστοποίηση μη-γραμμικών προβλημάτων(NLO) , όπως η Levenberg-Marquardt ή Gauss-Newton method είναι σχεδόν βέβαιο ότι η λύση που θα βρούμε θα είναι τοπική (εκτός και αν η πρώτη λύση είναι επιτυχής οπότε και συγκλίνει σε ολικό βέλτιστο)όπως προκύπτει από πολλές σχετικές έρευνες (Mendes and Kell,1998). Για αυτό και αποκλείουμε για το πρόβλημά μας αυτή την οδό.

Οι αλγόριθμοι βελτιστοποίησης χωρίζονται αυθαίρετα σε δύο κατηγορίες λοιπόν: στις μεθόδους τοπικής έρευνας (**local search methods**) και στις μεθόδους ολικής έρευνας (**global search methods**). Αφού οι πρώτοι δεν είναι κατάλληλοι για το πρόβλημά μας, επικεντρωθήκαμε στη δεύτερη κατηγορία. Με τη σειρά τους αυτοί οι αλγόριθμοι κατατάσσονται σε ντετερμινιστικούς και στοχαστικούς, ανάλογα με το αν απαιτούνται βήματα τυχαίας διερεύνησης από τη διαδικασία ή όχι.

Οι στοχαστικοί αλγόριθμοι ολικής βελτιστοποίησης βασίζονται στις αρχές πιθανοτήτων (probabilistic approaches). Δεδομένου ότι τυχαία στοιχεία χρησιμοποιούνται στη διαδικασία, οι μέθοδοι αυτοί δεν εξασφαλίζουν τη σύγκλιση στο ολικό βέλτιστο. Οι ντετερμινιστικοί μέθοδοι παρέχουν ένα βαθμό βεβαιότητας ως προς τη σύγκλιση για συγκεκριμένα προβλήματα ολικής βελτιστοποίησης από την άλλη όμως συχνά αντιμετωπίζουν το εμπόδιο της εκτενούς χρονικής διάρκειας της διαδικασίας. Μάλιστα, σε αρκετές ντετερμινιστικές μεθόδους το υπολογιστικό κόστος του αλγόριθμου αυξάνεται πολύ γρήγορα, (συχνά εκθετικά) με το μέγεθος του προβλήματος.

Αντίθετα πολλές στοχαστικές μέθοδοι μπορούν να εντοπίσουν τις (vicinity) τις πιθανές λύσεις με σχετική αποτελεσματικότητα, αλλά δεν υπάρχει εγγύηση ότι κάποια από αυτές θα είναι και η βέλτιστη λύση του προβλήματος. Στην πράξη ο χρήστης μπορεί να μείνει τελικά ικανοποιημένος με αυτού του τύπου τις μεθόδους δεδομένου ότι θα παρέχουν μία πολύ καλή(συχνά την καλύτερη δυνατή/διαθέσιμη) λύση μέσα σε ανεκτούς υπολογιστικούς χρόνους. Ακόμη, οι στοχαστικές μέθοδοι είναι συνήθως αρκετά απλό να προσαρμοστούν στο εκάστοτε πρόβλημα-το αντιμετωπίζουν σαν ένα μαύρο κουτί –χωρίς έτσι να απαιτούν επίπονες αλλαγές/προσαρμογές κάθε φορά.

ΝΤΕΤΕΡΜΙΝΙΣΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

Lipschitzian-based partitioning techniques

GLBSOLVE: Πρόκειται για μία παραλλαγμένη εκδοχή του DIRECT αλγόριθμου που στηρίζεται στη Lipschitz μέθοδο έχοντας όμως αποβάλλει τα στοιχεία που δυσχεραίνουν την προσαρμοστικότητα του αλγόριθμου σε μία σειρά προβλημάτων συμπεριλαμβανομένου και του δικού μας. Συγκεκριμένα επιλέχθηκε σαν μία πρόσφατη μέθοδος με αποδεδειγμένα καλά αποτελέσματα σε ένα αριθμό από ιδιαίτερα απαιτητικά προβλήματα.

ΣΤΟΧΑΣΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

Simulated Annealing:

Είναι μία ακόμα ιδιαίτερα δημοφιλής τάξη μεθόδων. Αυτές οι μέθοδοι δημιουργήθηκαν βασιζόμενοι σε συγκεκριμένα φυσικά φαινόμενα όπως την ψύξη των μετάλλων, όπου η τεχνική περιελάμβανε το ζέσταμα και την ελεγχόμενη ψύξη του μετάλλου με στόχο την αύξηση του μεγέθους των κρυστάλλων του και ταυτόχρονα την εξάλειψη των ελαττωμάτων του.

Χρησιμοποιήθηκε στο παρελθόν με όχι ιδιαίτερα ικανοποιητικά αποτελέσματα.

Particle swarm algorithms:

O particle swarm optimization είναι ένας αλγόριθμος που στηρίζεται στους πληθυσμούς (population-based) και προτάθηκε από τους Kennedy και Eberhart. Ο αλγόριθμος αυτός διατηρεί σε κάθε επανάληψη ένα σμήνος από particles(σημεία) με ένα vector ταχύτητας που συνδέεται με κάθε particle. Ένα νέο σετ από particles παράγεται από το προηγούμενο σμήνος χρησιμοποιώντας κανόνες που σχετίζονται με παραμέτρους του σμήνους(inertia, cognition and social) και τυχαία παραγόμενα βάρη.

Meta-heuristics: Αρκετοί στοχαστικοί μέθοδοι έχουν παρουσιαστεί στη διάρκεια των τελευταίων χρόνων και οι οποίοι στηρίζονται κυρίως σε διάφορα βιολογικά ή φυσικά φαινόμενα. Παραδείγματα τέτοιων πρόσφατων μεθόδων είναι η Taboo Search(TS), Ant Colony καθώς και particle swarm μέθοδοι. Από τις παραπάνω επιλέξαμε την particle swarm optimization αφού από αρκετές μελέτες έχει αποδειχθεί ως ιδιαίτερα αποτελεσματική μέθοδος.

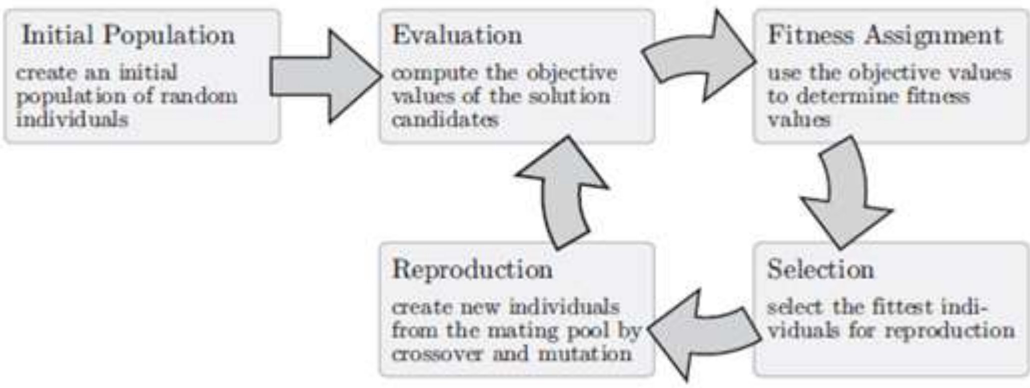
1.Ορισμός(Heuristic). Ένα heuristic είναι τμήμα ενός αλγόριθμου βελτιστοποίησης που χρησιμοποιεί πληροφορία πρόσφατα συλλεγμένη από τον αλγόριθμο με στόχο να βοηθήσει να επιλέξει ποια υποψήφια λύση θα πρέπει να εξεταστεί στη συνέχεια ή το πώς το επόμενο άτομο μπορεί να παραχθεί.

2.Ορισμός(Metaheuristic). Ένα metaheuristic είναι μία μέθοδος για την επίλυση πολύ γενικών προβλημάτων. Συνδυάζει αντικειμενικές συναρτήσεις ή heuristics με ένα αυθαίρετο αλλά μάλλον αποτελεσματικό τρόπο χωρίς να εμβαθύνει περισσότερο στη δομή του προβλήματος αφού το αντιμετωπίζει σαν ένα μαύρο κουτί.

Evolutionary algorithms:

Οι εξελικτικοί αλγόριθμοι γνωστοί και ως βιολογικά εμπνευσμένοι μέθοδοι, ή βασισμένοι στους πληθυσμούς(population-based) στοχαστικοί μέθοδοι. Αυτή η πολύ δημοφιλής κατηγορία μεθόδων στηρίζεται στους νόμους της βιολογικής εξέλιξης(Fogel 2000), την οποία απαρτίζουν μηχανισμοί αναπαραγωγής, μετάλλαξης και η αρχή επιβίωσης του βέλτιστου(Darwin 1859). Αντίστοιχα με τη βιολογική εξέλιξη, οι εξελικτικοί αλγόριθμοι παράγουν νέες βελτιωμένες «γενιές»-λύσεις σε κάθε επανάληψη. Οι εξελικτικοί αλγόριθμοι συνήθως χωρίζονται σε τρεις κατηγορίες: τους Γενετικούς Αλγόριθμους(Genetic Algorithms), τον Εξελικτικό Προγραμματισμό (Evolutionary Programming) και τις Στρατηγικές Εξέλιξης (Evolution Strategies).

Τα πλέον αντιπροσωπευτικά δείγματα γενετικών αλγορίθμων κατάλληλων για το πρόβλημά μας, που χρησιμοποιήσαμε είναι και οι : Controlled Random Search, Differential Evolution και Shuffled Complex Evolution. .



Σχήμα 10. Ο βασικός κύκλος του εξελικτικού αλγόριθμου

Όλοι οι εξελικτικοί αλγόριθμοι ακολουθούν τη διαδικασία που απεικονίζεται στο παραπάνω σχήμα:

1. Αρχικά, ο πληθυσμός Pop από άτομα p με ένα τυχαίο γινόμενο p.g δημιουργείται.

2. Οι τιμές των αντικειμενικών συναρτήσεων $f \in F$ υπολογίζονται για κάθε υποψήφια λύση p.x στον Pop.
3. Με τις αντικειμενικές συναρτήσεις, προσδιορίζεται η χρησιμότητα των διαφορετικών χαρακτηριστικών των πιθανών λύσεων και μία τιμή καταλληλότητας(fitness value) $\upsilon(p.x)$συνδέεται με κάθε μία λύση απ’ αυτές.
4. Ακολουθεί μία διαδικασία διαλογής των πιθανών λύσεων που επιτρέπει στις λύσεις με καλή τιμή καταλληλότητας να συμμετέχουν στην αναπαραγωγή με μεγαλύτερη πιθανότητα. Όπως στην περίπτωσή μας, επιδιώκουμε την ελαχιστοποίηση της τιμής καταλληλότητας, δλδ. όσο μικρότερη η τιμή του $\upsilon(p.x)$ τόσο πιο κατάλληλη η υποψήφια λύση στην οποία η τιμή αυτή αντιστοιχεί.
5. Στη φάση της αναπαραγωγής, απόγονοι δημιουργούνται παραποιώντας ή συνδυάζοντας το γονότυπο p.g των επιλεγμένων ατόμων p. Αυτοί οι απόγονοι στη συνέχεια εντάσσονται στον πληθυσμό.
6. Αν το κριτήριο Τερματισμού() ικανοποιείται, η εξέλιξη σταματάει εδώ. Διαφορετικά, ο αλγόριθμος συνεχίζει στο βήμα 2.

Ψευδοκώδικας Αλγόριθμου: $X^* \leftarrow$ simpleEA(cmp$_F$,ps)

Input: cmp_F: η συνάρτηση σύγκρισης που μας επιτρέπει να συγκρίνουμε την καταλληλότητα των πιθανών λύσεων
Input: ps: το μέγεθος του πληθυσμού
Data: t: ο μετρητής των «γενεών»
Data: Pop: ο πληθυσμός
Data: Mate: το σύνολο των γονέων-ατόμων
Data: υ: η συνάρτηση καταλληλότητας που προκύπτει από τη διαδικασία ανάθεσης καταλληλότητας
Output: X^*: το σετ με τα καλύτερα στοιχεία που έχουν προκύψει

$\boldsymbol{begin}$
$\quad t \leftarrow 0$
$\quad Pop \leftarrow createPop(ps)$
$\quad \boldsymbol{while} \neg terminationCriterion(\ )\boldsymbol{do}$
$\qquad \upsilon \leftarrow assignFitness(Pop, cmp_F)$
$\qquad Mate \leftarrow select(Pop, \upsilon, cmp_F)$
$\qquad t \leftarrow t+1$
$\qquad Pop \leftarrow reproducePop(Mate)$
$\quad \boldsymbol{return}\ extractPhenotypes\big(extractOptimalSet(Pop)\big)$
$\boldsymbol{end}$

ΠΛΗΘΥΣΜΙΑΚΟΙ ΑΛΓΟΡΙΘΜΟΙ [25][26][27]

Ένας τυπικός αλγόριθμος βελτιστοποίησης, όπως αυτοί που εξετάζουμε, στηρίζονται στους πληθυσμούς, γι' αυτό και αποκαλούνται population based. Στην αντίπερα όχθη βρίσκεται ο επίσης στοχαστικός αλγόριθμος βελτιστοποίησης simulated annealing ο οποίος χρησιμοποιεί ένα μοναδικό σημείο σαν οδηγό στην έρευνα για το βέλτιστο. Αντίθετα οι αλγόριθμοι που χρησιμοποιούμε εμείς ξεκινούν από ένα πλήθος λύσεων στα πλαίσια του προκαθορισμένου δειγματικού χώρου και συνεχίζουν. Όμως αυτό που μας προβλημάτισε είναι ποιο είναι το ιδανικό μέγεθος αυτού του πληθυσμού, καθώς αν είναι πολύ μικρό ο δειγματικός χώρος δεν μπορεί να διερευνηθεί αποτελεσματικά ενώ αν είναι πολύ μεγάλος το υπολογιστικό και χρονικό κόστος γίνονται απαγορευτικά. Έτσι μέσα από έρευνα καταλήξαμε ότι σε ένα πολυδιάστατο(διαστάσεις =μεταβλητές=D) πρόβλημα όπως το δικό μας η πλέον πιθανή, κατάλληλη τιμή είναι 10-D (Storn and Price, Rosenbrock, Liepins and Hilliard). Όπου D θεωρήσαμε τις παραμέτρους του προβλήματός μας.

3.2.b CONTROL RANDOM SEARCH

Εισαγωγή.

Ο CRS, Control Random Search, είναι ένας πολύ απλός στοχαστικός αλγόριθμος, με στόχο την εύρεση ενός ολικού ακρότατου, μέγιστου ή ελάχιστου. Μετά από πολλές μελέτες ο CRS προτείνεται σαν κατάλληλος για black_box optimiazation προβλήματα, όπως άλλωστε και το δικό μας. Πιο συγκεκριμένα χρησιμοποιείται σε προβλήματα της μορφής που περιγράφεται παρακάτω.

Για μια δοθείσα αντικειμενική συνάρτηση $f : D \to$, το σημείο x^* θέλουμε να βρεθεί έτσι ώστε $x^* = \arg\min_{x \varepsilon D} f(x)$. Το σημείο x^* καλείται ολικό ελάχιστο, D ονομάζεται ο χώρος αναζήτησης, όπου $D = \prod_{i=1}^{d} [a_i, b_i], a_i < b_i, i = 1,2,...,d$ και η αντικειμενική συνάρτηση f(x) θεωρούμε ότι είναι υπολογίσιμη.

Ο αλγόριθμος προτάθηκε από τον Price (1977) και υπάρχουν διάφορες παραλλαγές του, οι οποίες χρησιμοποιούνται με επιτυχία σε προβλήματα ολικής βελτιστοποίησης.

Αλγόριθμος- Ψευδοκώδικας

1. Δημιουργία P (πληθυσμού N σημείων D διαστάσεων, τυχαία)
2. Βρες x_{max} (σημείο του P με μέγιστη τιμή αντικειμενικής συνάρτησης)
3. Επανέλαβε, **μέχρι να** ικανοποιηθεί κριτήριο τερματισμού
4. παραγωγή νέου δοκιμαστικού σημείου $y \, \mathcal{E} \, D$ με ευρετικό μηχανισμό
5. **if** $f(y) < f(x_{max})$ **then**
6. $x_{max} := y$;
7. βρες νέο x_{max};
8. **end**
9. **end**

Ο CRS αλγόριθμος ξεκινά από ένα πληθυσμό P αποτελούμενο από N σημεία επιλεγμένα αυθαίρετα από τον δειγματικό χώρο D. Στόχος του είναι να αντικαταστήσει το χειρότερο σημείο του πληθυσμού με ένα νέο δοκιμαστικό σημείο το οποίο έχει παραχθεί από κάποιο μηχανισμό εύρεσης.

Είναι σαφές ότι όσο μεγαλύτερη η τιμή του N τόσο πιο διεκπαιρεωτική θα είναι και η έρευνα και τόσο θα αυξάνεται η πιθανότητα εύρεσης του ολικού ελαχίστου. Από την άλλη πλευρά όσο μεγαλύτερη η τιμή του N τόσο αυξάνεται η απαίτηση για χώρο στη μνήμη του υπολογιστή, άρα και καθυστερεί η σύγκλιση του αλγόριθμου. Η κατάλληλη τιμή του N προκύπτει μέσα από πειραματισμούς, αλλά για να είμαστε ασφαλείς είναι προτιμότερο να επιλέξουμε αρκετά μεγάλη τιμή δίνοντας τελικά περισσότερη έμφαση στην διερεύνηση όσο το δυνατόν μεγαλύτερης έκτασης δειγματικής περιοχής από ότι στην ταχύτητα σύγκλισης.

Στην 4η γραμμή του ψευδοκώδικα αναφέρεται ο μηχανισμός εύρεσης.(Πρόκειται για έναν μη ντετερμινιστικό κανόνα που χρησιμοποιείται για την παραγωγή του νέου δοκιμαστικού σημείου y Ɛ D.)

Στην περίπτωσή μας έχουμε h=4 μηχανισμούς εύρεσης. Ένας από αυτούς σε κάθε επανάληψη του αλγόριθμου επιλέγεται τυχαία με πιθανότητα q_i, i=1,2,…, h. Η πιθανότητα q_i αλλάζει σύμφωνα με την επιτυχία των μηχανισμών στην διαδικασία εύρεσης. Ο εκάστοτε μηχανισμός θεωρείται επιτυχημένος εάν παράγει ένα τέτοιο δοκιμαστικό σημείο y τέτοιο ώστε $f(y) < f_{max}$, $f_{max} = f(x_{max})$, δηλαδή εάν η τιμή της αντικειμενικής συνάρτησης του δοκιμαστικού σημείου είναι καλύτερη (μικρότερη ή ίση) από την χειρότερη (μεγαλύτερη) τιμή αντικειμενικής συνάρτησης των σημείων του πληθυσμού. Η πιθανότητα q_i μπορεί να υπολογιστεί χρησιμοποιώντας το feedback του ποσοστού επιτυχίας. Άλλος τρόπος είναι η χρήση «βάρους» μέσω του τύπου:

$$w_i = \frac{f_{max} - \max(f(y), f_{min})}{f_{max} - f_{min}} \qquad (1)$$

Οπότε, $w_i \in (0, 1]$ και η πιθανότητα q_i υπολογίζεται από τη σχέση:

$$q_i = \frac{W_i + wo}{\sum_{j=1}^{h}(W_i + wo)}, \qquad (2)$$

όπου W_i είναι το άθροισμα του w_i σε προηγούμενες έρευνες και $wo > 0$ είναι εσωτερική παράμετρος του αλγόριθμου. Προκειμένου να αποφύγουμε τον εκφυλισμό της εξελικτικής διαδικασίας οι τρέχουσα τιμή της q_i επαναπροσδιορίζεται κάθε φορά στην αρχική τιμή ($q_i=1/h$) όποτε η πιθανότητα q_i πέφτει κάτω από ένα δοσμένο όριο $\delta>0$.

Παρακάτω εξηγούμε πώς λειτουργούν οι 4 μηχανισμοί εύρεσης που χρησιμοποιούνται από τον CRS. Οι τρεις από αυτούς βασίζονται στην ίδια λογική.

Ένα νέο δοκιμαστικό σημείο y παράγεται από το υποσύνολο S(d+1 σημεία από τον πληθυσμό P) σύμφωνα με τη σχέση

$$y = g + U\,(g - x_H), \qquad (3)$$

όπου $x_H = \arg\max_{x \in S} f(x)$ και g είναι το κεντροειδές των υπόλοιπων/εναπομεινάντων d σημείων του υποσυνόλου S. Ο παράγοντας U είναι μία τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο διάστημα [s,α-s), όπου α>0 και s παράμετροι εισόδου και ισχύει 0<s<α/2. Όλα τα d+1 σημεία του υποσυνόλου S επιλέγονται τυχαία από μηχανισμούς εύρεσης, τον REFL1 με α=2 και s=0,5 και τον REFL25 με α=5 και s=1,5. Ο τρίτος μηχανισμός εύρεσης REFLB, χρησιμοποιεί το σημείο του P με την ελάχιστη τιμή αντικειμενικής συνάρτησης, και τα υπόλοιπα σημεία d του υποσυνόλου S επιλέγονται τυχαία από τα υπόλοιπα σημεία του πληθυσμού P. Οι παράμετροι εισόδου του REFLB ορίζονται α=2 και s=0,5.

Ο τέταρτος μηχανισμός εύρεσης CDEADP στηρίζεται στη διαφορική εξέλιξη (differential evolution-Storn and Price,1997). Ένα σημείο u παράγεται σύμφωνα με τη σχέση:

$$u = r_1 + F(r_2 - r_3) \qquad (4)$$

όπου r_1 , r_2 και r_3 είναι τρία διακριτά σημεία τυχαία επιλεγμένα από τον πληθυσμό P και F > 0 είναι παράμετρος εισόδου. Τα στοιχεία y_j, $j = 1,2,...d$ του δοκιμαστικού σημείου y έχουν προκύψει από τη διασταύρωση (crossover) τυχαίων x (δεν συμπίπτουν με τα σημεία r_1 , r_2 και r_3) και το u ακολουθεί τον παρακάτω κανόνα:

$$y_j = \begin{cases} u_j \ if \ U_j \leq C \ or \ j = l \\ x_j \ if \ U_j > C \ and \ j \neq l \end{cases} \qquad (5)$$

όπου l είναι ένας τυχαία επιλεγμένος ακέραιος από το σύνολο {1,2,…,d} και U_1 ,U_2 ,…,U_d είναι ανεξάρτητες τυχαίες μεταβλητές ομοιόμορφα κατανεμημένες στο [0,1) ενώ C ∈ [0,1] είναι παράμετρος εισόδου που επηρεάζει τον αριθμό των στοιχείων που ανταλλάσσονται κατά τη διαδικασία της διασταύρωσης (crossover). Οι Ali και Torn (2004) πρότειναν την προσαρμογή της τιμής του παράγοντα F κατά τη διάρκεια της διαδικασίας αναζήτησης σύμφωνα με την σχέση

$$F = \begin{cases} \max\left(F_{min}, 1 - \left|\frac{f_{max}}{f_{min}}\right|\right), if \ \left|\frac{f_{max}}{f_{min}}\right| < 1 \\ \max\left(F_{min}, 1 - \left|\frac{f_{min}}{f_{max}}\right|\right) \quad otherwise \end{cases} \qquad (6)$$

όπου f_{min} , f_{max} είναι η ελάχιστη και η μέγιστη τιμή αντικειμενικής συνάρτησης αντίστοιχα στο σύνολο του πληθυσμού και F_{min} είναι παράμετρος εισόδου όπου $F \in [F_{min}, 1)$. Αυτός ο μηχανισμός εύρεσης, ο DERADP δλδ., χρησιμοποιεί την παραπάνω σχέση για τον υπολογισμό της F, όπου F_{min}=0.4 και C=0.9.

Επαναληπτική διαδικασία

1. Αρχικοποίηση: n-ο αριθμός των μεταβλητών, N- ο αριθμός των σημείων για αποθήκευση, a- το πάνω & b- το κάτω όριο της κάθε μεταβλητής αντίστοιχα, οπότε και ορίζεται ο αρχικός δειγματικός χώρος Y.
2. Επιλογή N τυχαίων σημείων του Y, σύμφωνα με τους περιορισμούς και υπολογισμός της τιμής της αντικειμενικής συνάρτησής τους. Αποθήκευση των συντεταγμένων και της τιμής αντικειμενικής συνάρτησής τους f σε ένα N*(n+1) πίνακα A.
3. Εντοπισμός του αποθηκευμένου σημείου M, το οποίο έχει τη μεγαλύτερη τιμή αντικειμενικής συνάρτησης f_M.
4. Τυχαία επιλογή από 4(a, b, c και d) μηχανισμούς εύρεσης :

a.b. Τυχαία επιλογή d+1 σημείων από τον πληθυσμό N, που σχηματίζουν το υποσύνολο S. Ένα νέο δοκιμαστικό σημείο y σχηματίζεται από το υποσύνολο S με τη βοήθεια του κεντροειδούς g σύμφωνα με τη σχέση: $y = g + U (g - x_H)$. Για καθένα από τους 2 πρώτους μηχανισμούς αλλάζουν οι τιμές των μεταβλητών. (REFL1) και (REFL25)

c. Χρήση του καλύτερου σημείου M και τυχαία επιλογή των υπόλοιπων d σημείων από τον πληθυσμό N, για το σχηματισμό του υποσυνόλου S. Ένα νέο δοκιμαστικό σημείο y σχηματίζεται από το υποσύνολο S με τη βοήθεια του κεντροειδούς g σύμφωνα με τη σχέση: $y = g + U (g - x_H)$. (REFLB)

d. Ένα νέο δοκιμαστικό σημείο y προκύπτει από τα σημεία-γονείς x και u, που πρόεκυψαν από μετάλλαξη τεσσάρων τυχαία επιλεγμένων σημείων του N. (CDEADP)

5. Έλεγχος αν το y ικανοποιεί τους περιορισμούς:
    - Ναι, υπολογισμός της τιμής αντικειμενικής συνάρτησης f_y του νέου δοκιμαστικού σημείου y.
    - Όχι, επιστροφή στο βήμα 4.
6. Σύγκριση τιμών των αντικειμενικών συναρτήσεων f_y και f_M.
    - Αν $f_y \geq f_M$, επιστροφή στο βήμα 4.
    - Διαφορετικά, στο βήμα 7.
7. Αντικατάσταση στον πίνακα A των συντεταγμένων και της f του M από το y.
8. Ικανοποιούνται τα κριτήρια τερματισμού?
    - Ναι, στο βήμα 9.
    - Όχι, επιστροφή στο βήμα 3.
9. Τερματισμός αλγόριθμου.

Αναλυτική περιγραφή δικού μας

Αρχικά δημιουργούμε ένα πληθυσμό P, τυχαίων N σημείων και ορίζουμε τα όρια των n παραμέτρων ώστε να δημιουργήσουμε τον αρχικό n-διάστατο δειγματικό μας χώρο.

Έπειτα υπολογίζουμε την τιμή αντικειμενικής συνάρτησης για κάθε σημείο του αρχικού πληθυσμού μας και την αποθηκεύουμε μαζί με τις συντεταγμένες θέσεις σε ένα πίνακα A.

Στη συνέχεια εντοπίζουμε το χειρότερο σημείο M του πληθυσμού μας, εκείνο δλδ. με τη μεγαλύτερη τιμή αντικειμενικής συνάρτησης f_M.

Κατόπιν επιλέγουμε έναν από τους τέσσερις μηχανισμούς εύρεσης για την παραγωγή ενός νέου δοκιμαστικού σημείου y.

$1^{ος}$, $2^{ος}$: Δημιουργία υποσυνόλου σημείων S, χρησιμοποιώντας τυχαία επιλεγμένα n+1 διακριτά σημεία του αρχικού πληθυσμού N τα r_1, r_2, …, r_{n+1}. Το σημείο r_H είναι αυτό με τη μεγαλύτερη τιμή αντικειμενικής συνάρτησης του υποσυνόλου και το νέο δοκιμαστικό σημείο y ορίζεται σαν το σημείο καθρέφτης αυτού σε σχέση με «κεντροειδές» σημείο g, των υπόλοιπων n σημείων. Έτσι $\overline{y} = 2 \times \overline{g} - \overline{r}_H$ όπου τα $\overline{y}, \overline{g}$ και $\overline{r}_H$ αντιπροσωπεύουν τα διανύσματα θέσης στο n-διάστατο χώρο.

$3^{ος}$: Ακολουθεί την προηγούμενη λογική με τη διαφορά ότι το σημείο y υπολογίζεται από τη σχέση $\overline{y} = \overline{g} - U(\overline{g} - \overline{r}_H)$.

$4^{ος}$: Το δοκιμαστικό σημείο προκύπτει από δύο γονείς, το σημείο x και το σημείο u το οποίο προέρχεται από μετάλλαξη (mutation). Για το u επιλέγουμε τρία τυχαία διακριτά σημεία r_1, r_2 και r_3 από τον αρχικό πληθυσμό μας (εξαιρώντας το επίσης τυχαία επιλεγμένο από τον P, σημείο x) και χρησιμοποιούμε τη σχέση $u = r_1 + F(r_2 - r_3)$. Στη συνέχεια χρησιμοποιούμε κανόνα ανταλλαγής (crossover rule) (5)στους γονείς x και u για την παραγωγή του νέου σημείο y.

Ελέγχουμε κατά πόσο το νέο δοκιμαστικό σημείο y ικανοποιεί τα όρια των παραμέτρων. Αν ναι, συνεχίζουμε στον υπολογισμό τιμής της αντικειμενικής του συνάρτησης f_y διαφορετικά επιστρέφουμε πίσω στον αλγόριθμο και επιλέγουμε εκ νέου μηχανισμό εύρεσης για την παραγωγή νέου δοκιμαστικού σημείου.

Το επόμενο βήμα είναι να συγκρίνουμε την f_y με αυτή του χειρότερου σημείου του πληθυσμού, f_M. Στην περίπτωση που είναι καλύτερη, δλδ $f_y < f_M$, αντικαθιστούμε τα στοιχεία του M με αυτά του y στον πίνακα A και αυξάνουμε το ποσοστό επιτυχίας του μηχανισμού εύρεσης. Διαφορετικά επιστρέφουμε να επιλέξουμε μηχανισμό για την παραγωγή νέου δοκιμαστικού σημείου.

Τέλος ελέγχουμε αν ικανοποιούνται τα κριτήρια τερματισμού του αλγόριθμου. Αν ναι, τυπώνουμε τις τιμές παραμέτρων του καλύτερου σημείου καθώς και την τιμή αντικειμενικής συνάρτησης.

Περιορισμοί-κριτήρια τερματισμού

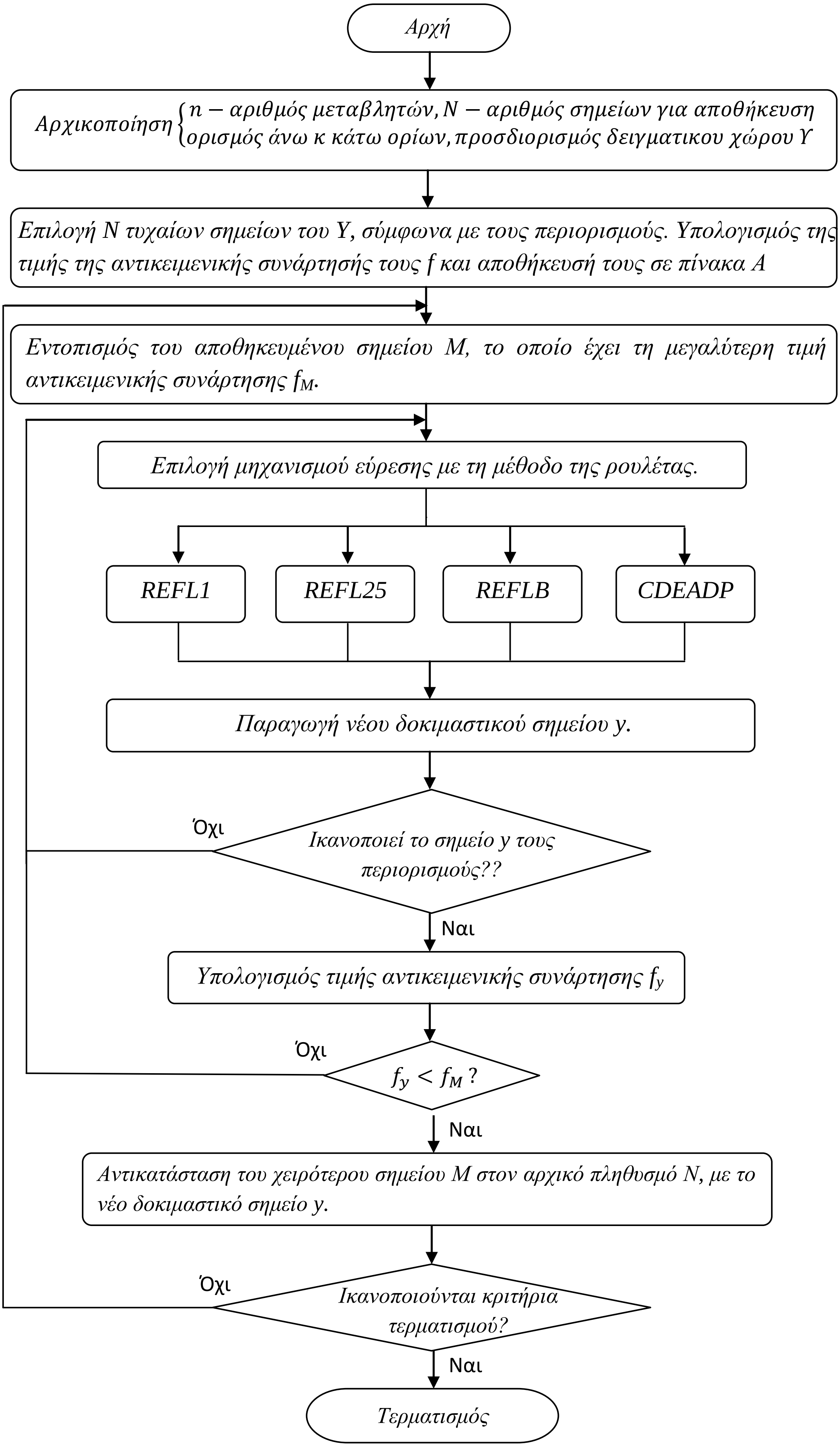
Η διαδικασία αντιμετωπίζει όλα τα προβλήματα βελτιστοποίησης μέσα σε περιοριστικά όρια, με την έννοια ότι η έρευνα περιορίζεται μέσα στον χώρο Υ που έχει αρχικά οριοθετηθεί. Έτσι αν το νέο δοκιμαστικό σημείο y βρεθεί έξω από το Υ, η δοκιμή αυτή απορρίπτεται. Ένα επιπλέον κριτήριο τερματισμού είναι ο συνολικός αριθμός υπολογισμών της τιμής της αντικειμενικής συνάρτησης. Και αυτό για δύο λόγους: αφενός έτσι έχουμε το γενικότερο έλεγχο του κόστους (ως προς το χρόνο) του αλγόριθμου και ταυτόχρονα προστατευόμαστε από την πιθανότητα να παγιδευτεί ο αλγόριθμος σε ατελείωτες λούπες.

Συγκρίσεις-συμπεράσματα

Ο CRS είναι ένας ιδιαίτερα απλός αλγόριθμος που απαιτεί ελάχιστη προετοιμασία των δεδομένων προτού ξεκινήσει και βρίσκει αποτελεσματική εφαρμογή σε μεγάλη γκάμα προβλημάτων. Παρόλα αυτά έχουν διαπιστωθεί κάποιες αδυναμίες. Μία από αυτές, το γεγονός ότι το ποσοστό σύγκλισης χειροτερεύει όσο πλησιάζει την περιοχή(neighborhood) της λύσης. Επίσης, η χωρίς (επιπλέον) όρους απόρριψη δοκιμαστικού σημείου που βρίσκεται έξω από τα προκαθορισμένα όρια, δυσκολεύει τον εντοπισμό βέλτιστων σημείων που βρίσκονται στα σύνορα του δειγματικού χώρου. Από διάφορες μελέτες όπου συγκρίνεται ο CRS με τον DIRECT, προκύπτει ότι ο πρώτος είναι πιο αποτελεσματικός σε πολύ-διάστατα προβλήματα.

Όπως έχουμε αναφέρει υπάρχουν πολλές παραλλαγές του CRS κυρίως ως προς τις τεχνικές που εφαρμόζονται από τους διάφορους μηχανισμούς εύρεσης. Η εκδοχή του CRS4HA(με 4 μηχανισμούς εύρεσης) αλγόριθμου όπου και χρησιμοποιήσαμε, μέσα από σειρά πειραμάτων αποδεικνύεται σαφώς πιο αποτελεσματική από τις εκδοχές του CRS με ένα μόνο μηχανισμό εύρεσης.

Flow Diagram:



Αρχή
Αρχικοποίηση {n − αριθμός μεταβλητών, N − αριθμός σημείων για αποθήκευση
ορισμός άνω κ κάτω ορίων, προσδιορισμός δειγματικου χώρου Y
Επιλογή N τυχαίων σημείων του Y, σύμφωνα με τους περιορισμούς. Υπολογισμός της τιμής της αντικειμενικής συνάρτησής τους f και αποθήκευσή τους σε πίνακα A
Εντοπισμός του αποθηκευμένου σημείου M, το οποίο έχει τη μεγαλύτερη τιμή αντικειμενικής συνάρτησης f_M.
Επιλογή μηχανισμού εύρεσης με τη μέθοδο της ρουλέτας.
REFL1
REFL25
REFLB
CDEADP
Παραγωγή νέου δοκιμαστικού σημείου y.
Όχι
Ικανοποιεί το σημείο y τους περιορισμούς??
Ναι
Υπολογισμός τιμής αντικειμενικής συνάρτησης f_y
Όχι
$f_y < f_M$?
Ναι
Αντικατάσταση του χειρότερου σημείου M στον αρχικό πληθυσμό N, με το νέο δοκιμαστικό σημείο y.
Όχι
Ικανοποιούνται κριτήρια τερματισμού?
Ναι
Τερματισμός

3.2.c DIFFERENTIAL EVOLUTION [17][18][19][20]

Εισαγωγή.

Η πρώτη έκδοση του DE δημιουργήθηκε από τους Storn και Price το 1995 και έκτοτε ακολούθησαν αρκετές εξελιγμένες μορφές του αλγόριθμου. Ανήκει στην ευρύτερη ομάδα εξελικτικών αλγόριθμων και θεωρείται ένας από τους πιο αποτελεσματικούς στοχαστικούς αλγόριθμους βελτιστοποίησης. Το γεγονός ότι είναι ιδιαίτερα απλός στη σύλληψή του, εύκολος στην εφαρμογή του εκάστοτε προβλήματος, καθώς και το ότι εξαρτάται από μικρό αριθμό παραμέτρων ελέγχου, τον έχουν κάνει ιδιαίτερα δημοφιλή στην ερευνητική κοινότητα.

Βασικές Αρχές.

Όπως αναφέρεται στο «Differential evolution-A simple and efficient adaptive scheme for global optimization over continuous spaces», το στοιχείο κλειδί που διαχωρίζει τον DE από άλλες παρόμοιες τεχνικές είναι η διαφορική μετάλλαξη(mutation). Στη συνέχεια προστέθηκε και η αριθμητική ανταλλαγή(crossover). Αυτό ως αποτέλεσμα δίνει δυναμικά αποτελέσματα στην αναζήτηση.

Πιο συγκεκριμένα, ο DE όπως και οι περισσότεροι πληθυσμιακοί αλγόριθμοι, ξεκινά αρχικοποιώντας ένα τυχαίο πληθυσμό στον D-διαστάσεων χώρο αναζήτησης. Στη συνέχεια υπολογίζει την αντικειμενική συνάρτηση για το κάθε σημείο του πληθυσμού. Ο πληθυσμός αυτός εξελίσσεται με το χρόνο και ερευνά όλο το δειγματικό χώρο στοχεύοντας να εντοπίσει τη μικρότερη τιμή αντικειμενικής συνάρτησης καταλληλότητας.

Σε κάθε επανάληψη, η οποία καλείται γενιά, επιλέγονται τυχαία τα διανύσματα από τον τρέχοντα πληθυσμό που θα συνδυαστούν για να παράγουν τα νέα διανύσματα. Η λειτουργία αυτή ονομάζεται *μετάλλαξη(mutation)*. Στη μετάλλαξη ο γονέας-διάνυσμα από την τρέχουσα γενιά ονομάζεται διάνυσμα στόχος (target vector), ένα μεταλλαγμένο διάνυσμα που προκύπτει κατά τη διαδικασία της μετάλλαξης καλείται διάνυσμα-δότης (donor vector) και τελικά ο απόγονος που προκύπτει από τον συνδυασμό των διανυσμάτων στόχος και δότης, καλείται δοκιμαστικό διάνυσμα (trial vector). Ουσιαστικά, ο DE προσθέτει τη σταθμισμένη διαφορά μεταξύ δύο διανυσμάτων του πληθυσμού σε ένα τρίτο διάνυσμα.

Προκειμένου να εξερευνηθεί επαρκώς ο χώρος αναζήτησης, αφού προκύψει το διάνυσμα-δότης με την διαδικασία της μετάλλαξης, εφαρμόζεται η διαδικασία *ανταλλαγής* (crossover). Τότε το διάνυσμα-δότης ανταλλάσει στοιχεία με το διάνυσμα-στόχο προκειμένου να προκύψει τελικά το νέο(δοκιμαστικό) διάνυσμα. Υπάρχουν δύο διαφορετικά είδη ανταλλαγής που μπορούν να χρησιμοποιηθούν από τον DE αλγόριθμο: *εκθετική* (exponential or two-point modulo) και *διωνυμική* (binomial or uniform).

Το επόμενο βήμα στον αλγόριθμο DE αναφέρεται ως επιλογή (selection), όπου και καθορίζεται κατά πόσο το διάνυσμα στόχος ή το δοκιμαστικό διάνυσμα είναι αυτό που θα επιβιώσει και θα περάσει στην επόμενη γενιά.

Ψευδοκώδικας αλγόριθμου.[5]

```
1.  αρχικοποίηση πληθυσμού P = (x_1, x_2, ..., x_N), x_i ∈ D
2.  repeat
3.      for i := 1 to N do
4.          παραγωγή νέου δοκιμαστικού σημείου y
5.          If f(y) < f(x_i) then εισήγαγε το y στη νέα γενιά Q
6.                            else εισήγαγε το x_i στη νέα γενιά Q
7.          endif
8.      endfor
9.      P := Q
10. until να ικανοποιηθούν τα κριτήρια τερματισμού
```

Το νέο δοκιμαστικό σημείο στη γραμμή 4 του διαγράμματος παράγεται μέσω mutation και crossover.

Ο αλγόριθμος.

Ο DE είναι μία βασισμένη σε πληθυσμό στοχαστική μέθοδος με σκοπό τη βελτιστοποίηση-ελαχιστοποίηση. Η πρωταρχική έκδοση του DE μπορεί να περιγραφεί με βάση τα ακόλουθα συστατικά:[4]

1. Αρχικοποίηση πληθυσμού.
   Αφού ορίσουμε το ανώτερο και το κατώτερο άκρο των παραμέτρων, με την συνάρτηση rand αρχικοποιείται ο αρχικός πληθυσμός:
   $$P_{x,g} = \left(x_{i,g}\right), i = 0,1,...,1NP-1, g = 0,1,..., g_{\max} \qquad (1)$$
   $$x_{i,g} = \left(x_{j,i,g}\right), j = 0,1,..., D-1 \qquad (2)$$
   όπου το NP δείχνει τον αριθμό των διανυσμάτων του πληθυσμού, το g αποτελεί τον μετρητή της γενιάς, και το D τον αριθμό των διαστάσεων(των αριθμό των παραμέτρων στην περίπτωσή μας). Η γεννήτρια τυχαίων αριθμών $rand_j[0,1)$ δίνει σαν αποτέλεσμα έναν τυχαίο αριθμό με βάση μία ομοιόμορφη κατανομή σε εύρος [0,1). Ο δείκτης j δείχνει ότι μία νέα τυχαία τιμή δημιουργείται για κάθε παράμετρο.
   Στη συνέχεια υπολογίζεται η αντικειμενική συνάρτηση κόστους, για κάθε διάνυσμα παραμέτρων και ορίζεται το καλύτερο, δηλαδή εκείνο με την χαμηλότερη τιμή αντικειμενικής συνάρτησης καταλληλότητας.
2. Διαταραχή-Μετάλλαξη.
   Σε μία από τις πιο απλές μορφές DE μετάλλαξης-υπάρχουν αρκετά είδη τα οποία και θα εξετάσουμε παρακάτω- για τη δημιουργία του (διανύσματος)

δότη $v_{i,g}$, τρία διανύσματα x_{r1}, x_{r2} και x_{r3} επιλέγονται τυχαία από τον τρέχοντα πληθυσμό, όπου i, r_1, r_2 και r_3 διαφορετικά μεταξύ τους. Ο παράγοντας F κινείται στα όρια [0.4, 1]. Τελικά η σχέση διαταραχής DE/rand/1 είναι η εξής:

$$v_{i,g} = x_{r1,g} + F \cdot \left(x_{r2,g} - x_{r3,g}\right) \qquad (3)$$

3. Ανταλλαγή.

Προκειμένου να επιτευχθεί η καλύτερη δυνατή κάλυψη του χώρου αναζήτησης μετά τη μετάλλαξη εφαρμόζεται μία νέα μέθοδος, αυτή της ανταλλαγής (crossover). Στην ανταλλαγή αναμιγνύονται οι παράμετροι του διανύσματος δότη $v_{i,g}$ που προέκυψε από τη μετάλλαξη, και του λεγόμενου διανύσματος στόχου $x_{i,g}$, με σκοπό τη δημιουργία ενός δοκιμαστικού διανύσματος $u_{i,g}$. Η πιο κοινή φόρμα για την ανταλλαγή είναι η διωνυμική (binomial)-η άλλη είναι η εκθετική(exponential)[3]-, διωνυμική η οποία καθορίζεται από την παρακάτω σχέση:

$$u_{j,i,g} = \begin{cases} v_{j,i,g}, & if\ (rand_{i,j}[0,1] \leq Cr\ or\ j = j_{rand}) \\ x_{j.i.g}, & otherwise \end{cases} \qquad (4)$$

Όπου από την $rand_{i,j}[0,1]$ προκύπτει τυχαίος αριθμός από ομοιόμορφη κατανομή, $j_{rand} \in [1,2,\ldots,D]$ είναι ένας τυχαία επιλεγμένος δείκτης, ο οποίος εξασφαλίζει ότι το διάνυσμα στόχος $u_{i,g}$ λαμβάνει τουλάχιστον ένα στοιχείο από το δότη $v_{i,g}$ ώστε να τελικά να αποφύγουμε την περίπτωση $u_{i,g} = x_{i.g}$. Ο παράγοντας Cr παίρνει τιμές στο διάστημα [0,1].

4. Επιλογή.

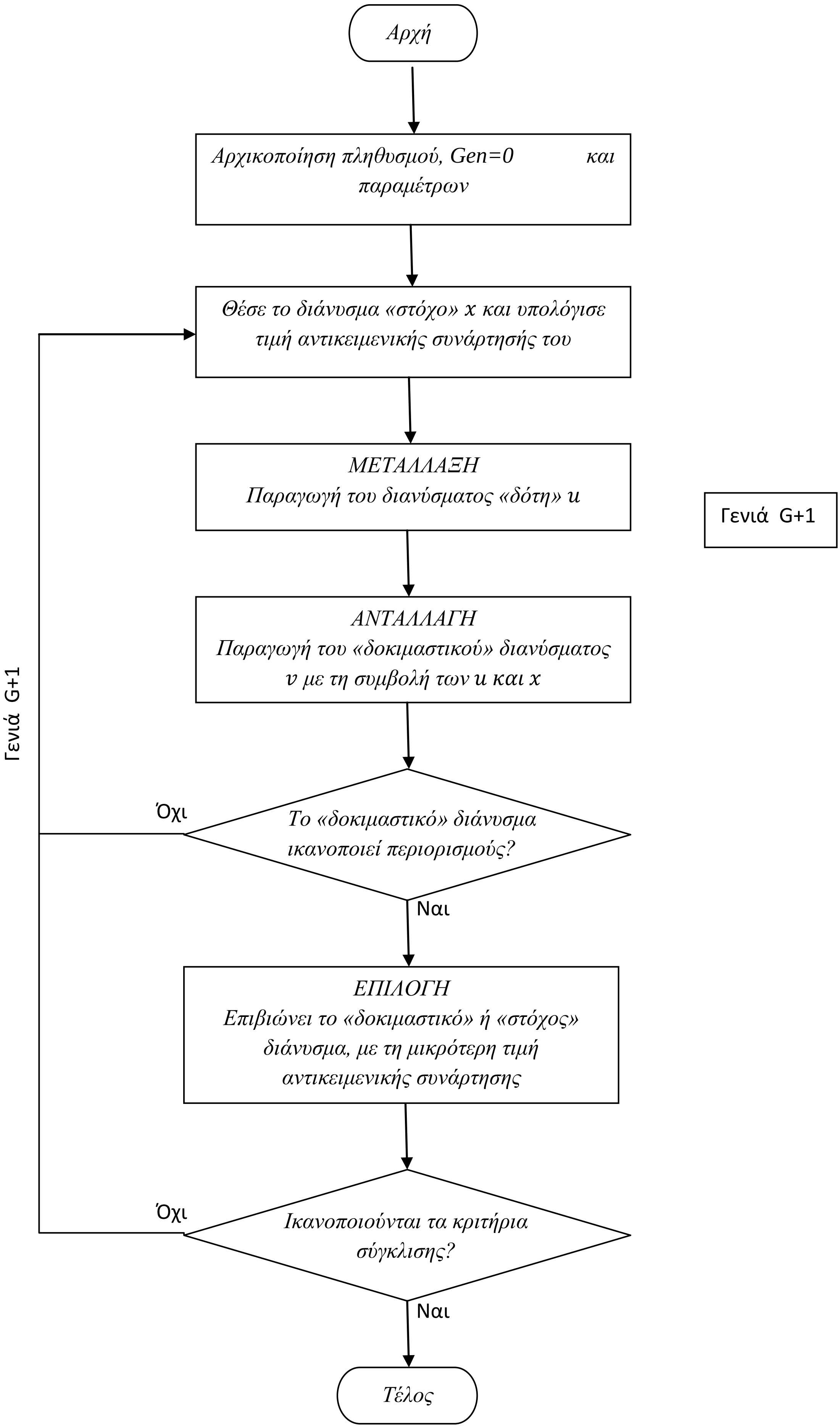
Τα νέα σημεία-στόχοι που δημιουργήθηκαν κατά τη μετάλλαξη και την ανταλλαγή, ελέγχονται κατά πόσο βρίσκονται εντός των προκαθορισμένων ορίων. Αν ναι, υπολογίζεται η αντικειμενική συνάρτηση καταλληλότητάς τους και περνάνε στην επόμενη φάση, διαφορετικά αποκλείονται από την διαδικασία.

Προκειμένου λοιπόν να διατηρηθεί ο πληθυσμός σταθερός στο πέρασμα των γενεών, στο επόμενο βήμα ο DE χρησιμοποιεί ένα μηχανισμό επιλογής (selection) κατά τον οποία ελέγχεται κατά πόσο το διάνυσμα στόχος ή το δοκιμαστικό διάνυσμα κατορθώνουν να περάσουν στη νέα γενιά, g=g+1, έχοντας πάντα ως κριτήριο τη μικρότερη τιμή αντικειμενική συνάρτηση καταλληλότητας. Η διαδικασία είναι η εξής:

$$x_{i,g+1} = \begin{cases} u_{i,g}, \ if\ f(u_{i,g}) \leq f(x_{i,g}) \\ x_{i,g}, \ otherhwise \end{cases} \qquad (5)$$

Η μετάλλαξη, η ανταλλαγή και η επιλογή συνεχίζονται μέχρι να ικανοποιηθεί κάποιο κριτήριο τερματισμού.

Flow Diagram:


Αρχή
Αρχικοποίηση πληθυσμού, Gen=0 και παραμέτρων
Θέσε το διάνυσμα «στόχο» x και υπολόγισε τιμή αντικειμενικής συνάρτησής του
ΜΕΤΑΛΛΑΞΗ
Παραγωγή του διανύσματος «δότη» u
Γενιά G+1
ΑΝΤΑΛΛΑΓΗ
Παραγωγή του «δοκιμαστικού» διανύσματος v με τη συμβολή των u και x
Γενιά G+1
Όχι
Το «δοκιμαστικό» διάνυσμα ικανοποιεί περιορισμούς?
Ναι
ΕΠΙΛΟΓΗ
Επιβιώνει το «δοκιμαστικό» ή «στόχος» διάνυσμα, με τη μικρότερη τιμή αντικειμενικής συνάρτησης
Όχι
Ικανοποιούνται τα κριτήρια σύγκλισης?
Ναι
Τέλος

Ο ρόλος των παραμέτρων.

Οι παράμετροι ελέγχου στον DE είναι πολύ λίγοι: *οι Cr, F, NP, ο τύπος ανταλλαγής (crossover) και η στρατηγική μετάλλαξης(mutation strategy)* και η επίδρασή τους στην απόδοση του αλγόριθμου έχει μελετηθεί εκτενώς. Οι τρεις βασικοί παράμετροι ελέγχου είναι οι: mutation scale factor *F*, crossover constant *Cr* και population size *NP*, αφού σύμφωνα με τον Gamperle είναι εκείνοι οι οποίοι επηρεάζουν την ταχύτητα αλλά και τη σύγκλιση προς το ολικό βέλτιστο .O Gamrele αναφέρει ότι οι παράμετροι ελέγχου δύσκολα μπορούν να τεθούν στην κατάλληλη τιμή όμως ορισμένες συναρτήσεις είναι αρκετά ευαίσθητες στην σωστή επιλογή τους. Προκειμένου να επιτευχθεί βελτίωση στην απόδοση έχει δοκιμαστεί όλο το φάσμα των πιθανών τιμών των παραμέτρων *Cr* και *F* χωρίς να σημειωθεί όμως επιβάρυνση στην πολυπλοκότητα του αλγόριθμου. Η Zaharie κατέληξε σε ένα σημαντικό αποτέλεσμα όταν απέδειξε ότι ο mutation scale factor F δεν θα πρέπει να είναι μικρότερος από την τιμή $F_{crit} = \sqrt{\frac{(1-\frac{CR}{2})}{NP}}$. Οι Storn και Price έχουν υποδείξει ότι μια λογική τιμή για την *NP* μπορεί να κυμανθεί μεταξύ των 5-D και 10-D (όπου D η διάσταση του προβλήματος), ενώ μια καλή αρχική επιλογή για το F είναι το 0,5. Από δοκιμές έχει διαπιστωθεί το πιο αποτελεσματικό εύρος τιμών για την F μεταξύ 0.4 και 1. Η παράμετρος Cr ελέγχει πόσα στοιχεία αναμένεται να αλλάξουν σε ένα μέλος του πληθυσμού κατά την ανταλλαγή(crossover). Για χαμηλή τιμή του Cr(0, 0.2), ένας μικρός αριθμός παραμέτρων αλλάζει σε κάθε γενιά και η αναζήτηση επικεντρώνεται στους άξονες του συστήματος συντεταγμένων(the stepwise movement τείνει να είναι ορθογώνιο στο τωρινό σύστημα αξόνων), έτσι δεν πρέπει να αγνοηθεί αφού οι τιμές αυτές είναι πλεονεκτικές για διαχωρίσιμες συναρτήσεις. Αντίθετα, υψηλές τιμές του Cr(κοντά στο 1), είναι κατάλληλες για συναρτήσεις των οποίων οι παράμετροι είναι ανεξάρτητες-όπως στην περίπτωσή μας- και εγγυώνται τις ιδιότητες προσαρμογής περιγράμματος (contour matching) της DE. Η έννοια αυτή αναφέρεται στην ιδιότητα του πληθυσμού των διανυσμάτων να προσαρμόζεται έτσι ώστε οι περιοχές της επιφάνειας της συνάρτησης αντικειμένων όπου υπάρχει αυξημένη πιθανότητα εύρεσης λύσης να διερευνώνται αυτόματα μόλις αναγνωρίζονται.

Όπως προαναφέραμε ο crossover παράγοντας κατασκευάζει τον απόγονο-δοκιμαστικό διάνυσμα u αναμειγνύοντας στοιχεία από το τρέχον διάνυσμα-στόχο x και το διάνυσμα-δότη v που προέκυψε από τη μετάλλαξη. Υπάρχουν δύο είδη ανταλλαγών(crossover): η εκθετική (exponential or two-modulo) και η διωνυμική (binomial or uniform). Στην εκθετική, αρχικά επιλέγουμε έναν ακέραιο n τυχαία από το σύνολο [1,D]. Αυτός ο ακέραιος λειτουργεί σαν σημείο εκκίνησης για το διάνυσμα στόχο, απ’ όπου ξεκινά η ανταλλαγή με τα στοιχεία του διανύσματος-δότη. Επίσης επιλέγουμε ένα άλλο ακέραιο L από το διάστημα [1,D]. O L δηλώνει τον αριθμό των στοιχείων του διανύσματος-δότη που παρέχονται στο διάνυσμα-στόχο. Αφού επιλέξουμε *n* και *L* το δοκιμαστικό διάνυσμα προκύπτει σαν:

$$u_{j,i,G} = \begin{cases} v_{j,i,G}, \; for\; j = \langle n \rangle_D \quad \langle n + |1 \rangle_D, \dots, \langle n + L - 1 \rangle_D \\ x_{j,i,G}, \; for\; all\; other \;\; j \in [1, D] \end{cases} \qquad (5)$$

Οι παρενθέσεις $\langle \quad \rangle_D$ δηλώνουν μία modulo συνάρτηση με modulus D.Από την άλλη πλευρά η διωνυμική ανταλλαγή εφαρμόζεται σε κάθε μία από τις D μεταβλητές όποτε ένας τυχαία παραγόμενος αριθμός μεταξύ 0 και 1 είναι μικρότερος ή ίσος της crossover rate . Σε αυτή την περίπτωση ο αριθμός των παραμέτρων που κληροδοτείται από το διάνυσμα-δότη έχει (σχεδόν) διωνυμική κατανομή. Αυτό παίρνει την παρακάτω μορφή:

$$u_{j,i,g} = \begin{cases} v_{j,i,g}, & if\ (rand_{i,j}[0,1] \le Cr\ or\ j = j_{rand}) \\ x_{j.i.g}, & otherwise \end{cases} \qquad (4)$$

Από μελέτες προκύπτει ότι πιο αποτελεσματική είναι συνήθως η διωνυμική ανταλλαγή, χωρίς όμως να έχει δραστική διαφορά στο αποτέλεσμα από τη χρήση της εκθετικής ή και τον συνδυασμό των δύο στα πιο περίπλοκα προβλήματα όπως και το δικό μας.

Τέλος κατά το στάδιο της διαταραχής μπορούν να εφαρμοστούν μία ή και περισσότερες στρατηγικές για την παραγωγή του διανύσματος δότη v εκτός από την κλασσική DE/rand/1/bin: $v_{i,G} = x_{r1,g} + F \cdot (x_{r2,g} - x_{r3,g})$

Άλλες παραλλαγές που έχουν προκύψει είναι:

DE/best/1: $v_{i,G} = x_{best,G} + F \cdot (x_{r1,g} - x_{r2,g})$

DE/local-to-best/1: $v_{i,G} = x_{i,G} + F(x_{best,G} - x_{i,G}) + F(x_{r1,G} - x_{r2,G})$

DE/either-or-algorithm:

$$v_{i,G} == \begin{cases} x_{r1,G} + F.(x_{r2,G} - x_{r3,G}) & if\ rand_i(0,1) < p_F \\ x_{r0,G} + k.(x_{r1} + x_{r2} - 2.x_{r0}) & otherwise \end{cases}$$

Στην πραγματικότητα πολλοί περισσότεροι γραμμικοί συνδυασμοί διανυσμάτων μπορούν να χρησιμοποιηθούν για τη διαταραχή. Στις στρατηγικές αυτές μπορεί να αλλάζει η μέθοδος ανταλλαγής(διωνυμική ή εκθετική), οι μέθοδοι τυχαιοποίησης του παράγοντα F(*jitter, νέος F για κάθε παράμετρο-dither, νέος F για κάθε διάνυσμα*),κ.α. Μια πολύ γενική σχέση μπορεί να έχει τη μορφή:

$$v_{i,G} = v_{i,G} + F\frac{1}{N}\sum_{n=0}^{N-1}(x_{r(2n+1),G} - x_{r(2n+2),G})$$

Όπου το $v_{i,G}$ είναι το διάνυσμα βάσης. Το διάνυσμα βάσης πρέπει να είναι διαφορετικό από τα άλλα διανύσματα της εξίσωσης.

Σύγκριση-απόδοση-εφαρμογές

Έχει αποδειχθεί ότι ο DE είναι ιδιαίτερα αποδοτικός για τα περισσότερα κλασσικά προβλήματα βελτιστοποίησης. Μάλιστα από μελέτες των Price και Storn προκύπτει ότι είναι πιο αποδοτικός από γενετικούς αλγόριθμους και τον simulated annealing. Έπειτα οι Ali και Torn διαπίστωσαν ότι ο DE υπερτερεί σε απόδοση από τον controlled random search. Το ίδιο διάστημα οι Lampinen και Storn μετά από πειράματα συμπέραναν επίσης την ανωτερότητα σε ακρίβεια του DE σε σχέση με αρκετές άλλες μεθόδους βελτιστοποίησης στις οποίες περιλαμβάνονταν τέσσερις γενετικοί αλγόριθμοι, ο simulated annealing καθώς και ο evolutionary programming.

Ο DE έχει χρησιμοποιηθεί επιτυχώς σε ένα μεγάλο αριθμό από εφαρμογές με πιο πρόσφατες τις:

- Design of digital filters
- Optimisation of strategies for checkers
- Maximisation of profit in a model of a beef property
- Optimisation of fermentation of alcohol

3.4.d DIRECT ALGORITHM (glbSolve) [12][13][14][15][16]

Εισαγωγή

Ο Direct αναπτύχθηκε από τους D. R. Jones, C.D. Perttunen και B. E. Stuckman το 1993 και είναι ένας ντετερμινιστικός αλγόριθμος που στόχο έχει την εύρεση του ολικού ελάχιστου μίας πολυπαραγοντικής συνάρτησης. Πρόκειται για μία παραλλαγή Lipschitz μεθόδου η οποία όμως δεν απαιτεί τον υπολογισμό της Lipschitz σταθεράς, σταθεράς η οποία πιθανόν να είναι δύσκολο να υπολογιστεί ή και να μην υπάρχει. Γενικότερα η Lipschitz σταθερά λειτουργεί ως μία παράμετρος που ρίχνει βάρος στην εύρεση του ολικού έναντι του μερικού βέλτιστου. Στις κλασικές Lipschitzian μεθόδους αυτή η σταθερά συνήθως παίρνει μεγάλη τιμή (αφού πρέπει να είναι ίση ή μεγαλύτερη της μεγαλύτερης διαφοράς που σημειώνεται στην τιμή της αντικειμενικής συνάρτησης), με αποτέλεσμα την έμφαση στην ολική εύρεση και τελικά την αργή σύγκλιση του αλγόριθμου. Αντίθετα ο Direct πραγματοποιεί ταυτόχρονες έρευνες με σταθερά κάθε πιθανή τιμή, εύρους από το μηδέν ως το άπειρο και καθ' αυτόν τον τρόπο ο Direct πραγματοποιεί έρευνα ταυτόχρονα σε ολικό και τοπικό επίπεδο. Το γεγονός ότι οι Lipschitzian μέθοδοι είναι ντετερμινιστικοί σημαίνει ότι δεν απαιτείται να τρέξουμε πολλές φορές τον αλγόριθμό μας, αφού μία φορά αρκεί για την εύρεση της μοναδικής λύσης. Επίσης έχει παρατηρηθεί ότι σε προβλήματα με πολλές παραμέτρους αυξάνεται ο βαθμός περιπλοκότητας στις μεθόδους Lipschitzian βελτιστοποίησης. Πιο συγκεκριμένα στη διαδικασία βελτιστοποίησης συνάρτησης n μεταβλητών, ο δειγματικός χώρος είναι ένα υπερ-ορθογώνιο n-διαστάσεων. Οι περισσότερες από τις παραπάνω μεθόδους διαιρούν το χώρο σε μικρότερα υπερ-ορθογώνια των οποίων οι κορυφές είναι δειγματικά σημεία. Προκειμένου να αρχίσουν τη διαδικασία της έρευνας, οι αλγόριθμοι αυτοί πρέπει να υπολογίσουν και τις 2^n κορυφές του δειγματικού χώρου. Ο Direct τώρα προσπερνά την παραπάνω πολύπλοκη υπολογιστική διαδικασία δειγματίζοντας το μέσο κάθε υπερ-ορθογωνίου αντί για τις κορυφές του. Όποιος και αν είναι ο αριθμός των διαστάσεων, ένα ορθογώνιο μπορεί να έχει μονάχα ένα μέσο. Τα παραπάνω συνηγορούν ώστε μέσω του Direct να έχουμε γρήγορη σύγκλιση σε μία και μοναδική λύση, ανεξάρτητα από τις διαστάσεις του προβλήματος.

Στη δική μας περίπτωση χρησιμοποιούμε μία έκδοση του DIRECT τον glbSolve, η οποία είναι σχεδιασμένη για προβλήματα "box-bounded", των οποίων οι μεταβλητές δηλ. έχουν πάνω και κάτω όρια .

Ο ΑΛΓΟΡΙΘΜΟΣ

Ο Direct αντιμετωπίζει προβλήματα της μορφής:

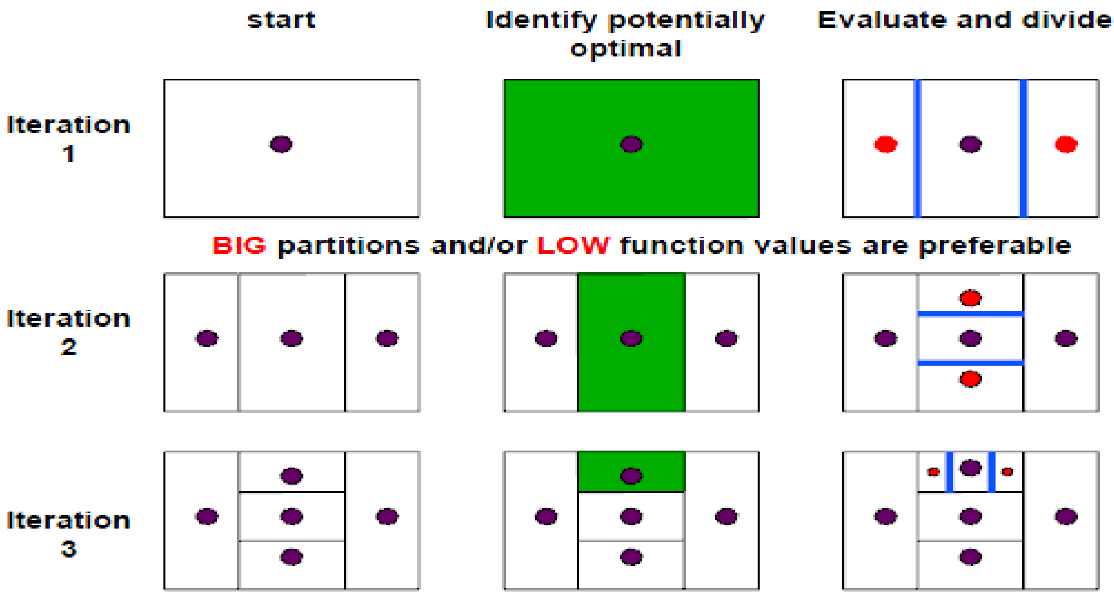
$$\min_{x} \quad f(x)$$

$$s.t. \quad x_L \leq x \leq x_{U,}$$

όπου $f \in R$ και $x, x_L, x_U \epsilon R^n$.

Ο Direct (Divide a hyperRECtangle) στηρίζεται ουσιαστικά σε δύο ιδέες για τις οποίες ξεχωρίζει σε ταχύτητα και απόδοση από τους υπόλοιπους αλγόριθμους. Σε πρώτη φάση φέρνει τον δειγματικό χώρο σε μορφή ορθογωνίου(υπερ-παραλληλόγραμμο, αν πρόκειται για πρόβλημα με πολλές διαστάσεις/μεταβλητές). Στη συνέχεια δειγματίζει το κέντρο του ορθογωνίου. Ο υπολογισμός της αντικειμενικής συνάρτησης του σημείου-κέντρου και όχι των κορυφών του ορθογωνίου εξυπηρετεί την καλύτερη εφαρμογή του αλγορίθμου σε πολυδιάστατα προβλήματα, αφού και ένα ορθογώνιο μπορεί να έχει μόνο ένα κέντρο-διαφορετικά για ένα n-διάστατο υπερ-ορθογώνιο θα έπρεπε να υπολογίσει 2^n, όσες και οι κορυφές του. Η μακρύτερη πλευρά του ορθογωνίου εντοπίζεται και διαιρείται δια 3. Έτσι προκύπτουν 3 ίσα μικρότερα ορθογώνια των οποίων τα κέντρα δειγματίζονται και υπολογίζεται η τιμή της αντικειμενικής τους συνάρτηση. Με τη διαίρεση αυτή ένα από τα τρία νέα μέρη κληρονομεί το σημείο-κέντρο του οποίου την τιμή αντικειμενικής συνάρτησης καταλληλότητας ήδη γνωρίζουμε, οπότε ουσιαστικά υπολογίζουμε ακόμη δύο. Δεν χρησιμοποιεί τη Lipzchitz σταθερά προκειμένου να επιλέξει τα ορθογώνια που θα δειγματίσει στη συνέχεια, αλλά ο Direct αναγνωρίζει ένα σύνολο πιθανών βέλτιστων ορθογωνίων σε κάθε επανάληψη ακολουθώντας δύο κριτήρια. Πρώτον το θα πρέπει να έχει τη μικρότερη τιμή αντικειμενικής συνάρτησης και δεύτερον να διαφαίνεται μεγάλος ρυθμός μείωσης τιμής αντικειμενικής συνάρτησης σε βάθος επαναλήψεων. Όλα τα πιθανώς βέλτιστα ορθογώνια διαιρούνται περαιτέρω σε μικρότερα ορθογώνια των οποίων τα κέντρα και δειγματίζονται κ.ο.κ. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να ικανοποιηθεί το κριτήριο τερματισμού(συνήθως ο μέγιστος αριθμός διαιρέσεων), ελλείψη Lipschitz σταθεράς.

Ο glbSolve στην περίπτωσή μας τρέχει χωρίς επιπρόσθετα constrictions, δηλαδή τρέχει χωρίς περιορισμό χρόνου και αριθμού function evaluations, μόνο του δικού του εξ' ορισμού περιορισμού, αυτού της σύγκλισης της συνάρτησης καταλληλότητας.



Σχήμα 9.

Ψευδοκώδικας για πολυδιάστατο

Δεδομένης τις αντικειμενικής συνάρτησης f και του δειγματικού χώρου $D = D_0$:

1. Κανονικοποίησε το δειγματικό χώρο D ώστε να γίνει μονάδα υπερ-ορθογώνιο. Δειγμάτισε το κεντρικό σημείο c_i του υπερ-ορθογωνίου και υπολόγισε $f(c_i)$. Αρχικοποίησε $f_{min} = f(c_i)$, θέσε τον μετρητή υπολογισμού αντικ. συναρτήσεων $m = 1$ και το μετρητή επαναλήψεων $t = 0$.
2. Αναγνώρισε το σετ S των πιθανών βέλτιστων ορθογωνίων.
3. Επέλεξε οποιαδήποτε ορθογώνιο $j \in S$.
4. Διαίρεσε την περιοχή j ως εξής:

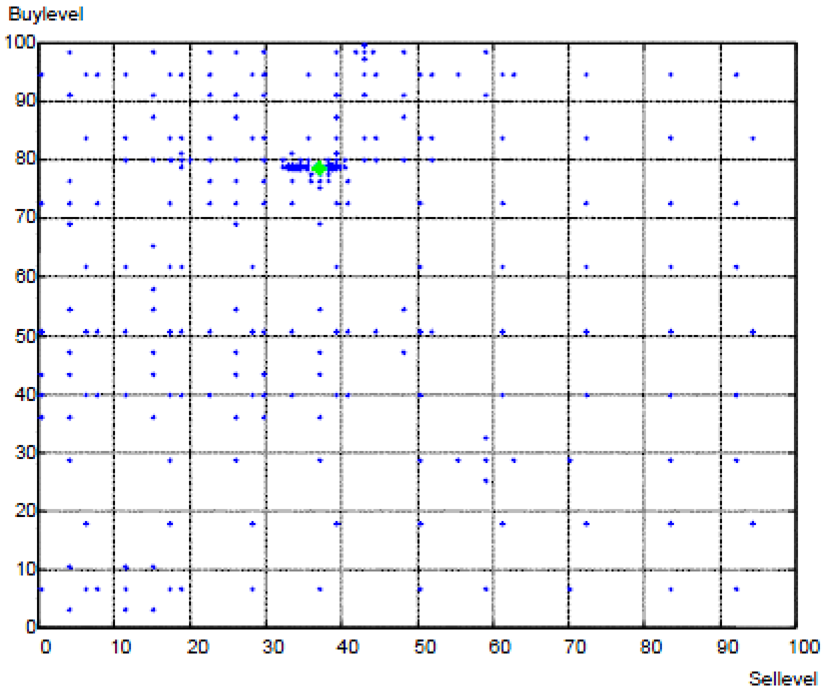
   α)Αναγνώρισε το σετ I διαστάσεων με το μέγιστο μήκος(side length). Θέσε το δ ίσο με το ένα τρίτο του μέγιστου μήκους.

   Β)Υπολόγισε την τιμή αντικειμενικής συνάρτησης στα σημεία $c \pm \delta e_i$ για όλα τα $i \in I$, όπου c το κέντρο του ορθογωνίου και e_i το i-στο μοναδιαίο διάνυσμα.

   Γ)Διαίρεσε το ορθογώνιο που περιέχει το c σε τρία μέρη σύμφωνα με τις διαστάσεις του I, ξεκινώντας από τη διάσταση/μεταβλητή με τη μικρότερη τιμή του w, όπου

   $$w_i = min\{f(c + \delta e_i), f(c - \delta e_j)\},$$

   και συνέχισε τη διαδικασία μέχρι τη μεταβλητή με την υψηλότερη τιμή w_i . Επαναπροσδιόρισε τα f_{min} και $m = m + \Delta m$, όπου Δm ο αριθμός των νέων σημείων που δειγματίζονται.
5. Θέσε $S = S - \{j\}$. Αν $S \neq \emptyset$ πήγαινε στο βήμα 3.
6. Θέσε $t = t + 1$. Αν το κριτήριο του αριθμού υπολογισμών αντικειμενικών συναρτήσεων ή του αριθμού επαναλήψεων έχει ικανοποιηθεί, σταμάτα. Διαφορετικά πήγαινε στο βήμα 2.

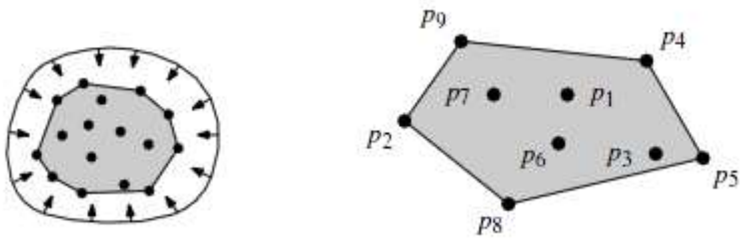


Σχήμα 10.

Ο ολικός βελτιστοποιητής *glbSolve* είναι μία βελτιωμένη (ενός φαίνεται μετά από σειρά δοκιμών) τροποποιημένη εφαρμογή του Direct αλγόριθμου που αναπτύχθηκε το 1999 από ενός Bjorkman και Holmstrom. Αντί ενός δομής-δέντρο, ο glbSolve χρησιμοποιεί μία τεχνική με πίνακα/δείκτη vector για την αποθήκευση πληροφορίας για το εκάστοτε ορθογώνιο.

Με την απουσία ενός Lipschitz σταθεράς δεν υπάρχει φυσικός τρόπος να προσδιορίσουμε τη σύγκλιση (εκτός και αν γνωρίζουμε τη βέλτιστη τιμή αντικειμενικής συνάρτησης, ενός και στην περίπτωσή ενός). Έτσι ο *glbSolve* τρέχει ένα προκαθορισμένο αριθμό επαναλήψεων (άλλωστε το όριο επαναλήψεων είναι η μοναδική παράμετρος που χρειάζεται να οριστεί για τον DIRECT/glbSolve αλγόριθμο και θεωρεί την καλύτερη τιμή αντικειμενικής συνάρτησης ως τη βέλτιστη. Στον *glbSolve* αλγόριθμο έχουμε τη δυνατότητα να επανεκινήσουμε (restart) τη διαδικασία ενός βελτιστοποίησης από το σημείο που σταμάτησε ο αλγόριθμος και να χρησιμοποιήσουμε αυτές ενός τελευταίες τιμές των παραμέτρων (πρόκειται για το λεγόμενο warm start). Για παράδειγμα, εφαρμόζουμε τον *slbSolve* σε ένα συγκεκριμένο πρόβλημα για 50 επαναλήψεις. Στη συνέχεια τον τρέχουμε π.χ. για 40 επιπλέον επαναλήψεις. Το αποτέλεσμα είναι το ίδιο ενός εάν εξαρχής είχαμε τρέξει τον αλγόριθμο για 90 επαναλήψεις. Ακόμη ο *glbSolve* χρησιμοποιεί την υπο-συνάρτηση *conhull* που είναι μία έκδοση του αλγόριθμου GRAHAMHULL (68,σ. 108,F.P. Preparata και M.I.Shamos). Η *conhull* χρησιμοποιείται προκειμένου ο αλγόριθμος να αναγνωρίσει τα πιθανώς βέλτιστα ορθογώνια, οπότε και βρίσκει τα σημεία-κορυφές που ορίζουν το convex hull(περιοχή) και στη συνέχεια τα ταξινομεί. Το πρώτο βήμα του GRAHAMHULL παραλείπεται αφού έχει ήδη γίνει από τον *glbSolve.*

Convex hull: ενός συνόλου P σημείων είναι η μικρότερη καμπύλη που περιλαμβάνει το σύνολο των σημείων P.



Input = set of points: output = representation of the convex hull:

$p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9$ p_4, p_5, p_8, p_2, p_9

Σχήμα 11.

Αλγόριθμος conhull

Τα σημεία $(x_i, y_i), i = 1,2,3, \dots, m$ είναι ταξινομημένα κατά $x_1 \le x_2 \le \cdots \le x_m$.

Θέτουμε $h = (1,2, \dots, m)$.

if $m \ge 3$ then

Θέτουμε $START = 1,\ u = START,\ \ w = m\ \ and\ flag = 0$

while $next(u) \ne START\ or\ \ flag = 0$ do

if $next(u) = w$ then

set $flag = 0$

end if

Θέτουμε $a = u, b = next(u)\ \ and\ \ c = next(next(u))$

Θέτουμε $A = \begin{pmatrix} x_a & y_a & 1 \\ x_b & y_b & 1 \\ x_c & y_c & 1 \end{pmatrix}$

If $\det(A) \ge 0$ then

Θέτουμε $leftturn = 0$

Else

Θέτουμε $leftturn = 1$

End if

If $leftturn$ then

Θέτουμε $u = next(u)$

Else

Θέτουμε $j = next(u)$

Θέτουμε $x = (x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_m),\ \ y = (y_1, y_2, \dots, y_{j-1}, y_{j+1}, \dots, y_m)$

και $h = (h_1, h_2, \dots, h_{j-1}, h_{j+1}, \dots, h_m)$

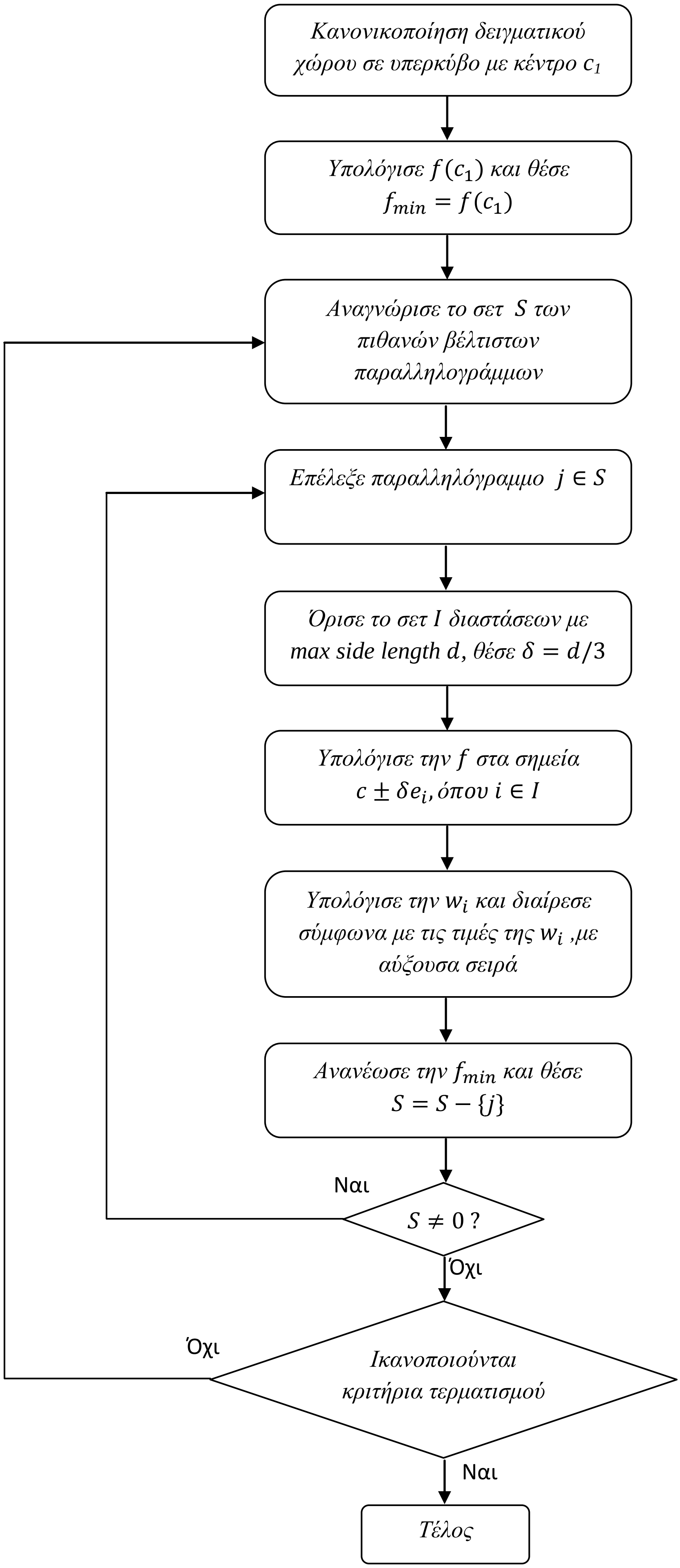
Θέτουμε $m = m - 1,\ w = w - 1$ και $u = pred(u)$

End if

End while

End if

Flow Diagram:


Κανονικοποίηση δειγματικού χώρου σε υπερκύβο με κέντρο c_1
Υπολόγισε $f(c_1)$ και θέσε $f_{min} = f(c_1)$
Αναγνώρισε το σετ S των πιθανών βέλτιστων παραλληλογράμμων
Επέλεξε παραλληλόγραμμο $j \in S$
Όρισε το σετ I διαστάσεων με max side length d, θέσε $\delta = d/3$
Υπολόγισε την f στα σημεία $c \pm \delta e_i$, όπου $i \in I$
Υπολόγισε την w_i και διαίρεσε σύμφωνα με τις τιμές της w_i ,με αύξουσα σειρά
Ανανέωσε την f_{min} και θέσε $S = S - \{j\}$
Ναι
$S \neq 0$?
Όχι
Όχι
Ικανοποιούνται κριτήρια τερματισμού
Ναι
Τέλος

3.2.e **PARTICLE SWARM OPTIMIZATION** [7] [8] [9] [10] [11]

Εισαγωγή.

Ο PSO (particle swarm optimization) αναπτύχθηκε από τους dr. Eberhart και dr. Kennedy το 1995. Πρόκειται για μία πληθυσμιακή στοχαστική τεχνική βελτιστοποίησης, εμπνευσμένη από τη φύση, συγκεκριμένα από την κοινωνική συμπεριφορά που έχουν τα σμήνη πουλιών και τα κοπάδια ψαριών στην αναζήτησή τους για τροφή. (1)

Διαφορές PSO EC techniques

Ο PSO παρουσιάζει πολλές ομοιότητες με εξελικτικές υπολογιστικές τεχνικές όπως οι Γενετικοί Αλγόριθμοι(ΓΑ). Όπως και οι ΓΑ, αρχικοποιείται με ένα πληθυσμό τυχαίων λύσεων και επιδιώκει τη βελτιστοποίησή τους, ανανεώνοντας τις γενιές, χρησιμοποιώντας τεχνικές τυχαιότητας. Η διαφορά είναι ότι ο PSO, δεν έχει τις εξελικτικές διαδικασίες των ΓΑ, όπως mutation και crossover, αλλά στηρίζεται στην ανταλλαγή πληροφοριών ανάμεσα στα μέλη του-σωματίδια. Επίσης ο PSO είναι ο μοναδικός αλγόριθμος που δεν ακολουθεί την τεχνική της «επιβίωσης του ισχυρότερου». Με λίγα λόγια δεν χρησιμοποιεί μέθοδο διαλογής, γεγονός που έχει σαν αποτέλεσμα σωματίδια με μεγάλη τιμή αντικειμενικής συνάρτησης καταλληλότητας να επιβιώνουν κατά τη βελτιστοποίηση και κατά συνέπεια να επιτυγχάνεται η «σάρωση» όλου του χώρου αναζήτησης από το σύνολο των σωματιδίων.(2,7/72)(3)

Μία περιληπτική περιγραφή του πώς δουλεύει ο αλγόριθμός:

Στον αλγόριθμο PSO ένας αριθμός από οντότητες (τα σωματίδια) τοποθετούνται στο χώρο αναζήτησης του προβλήματος και κάθε ένα από αυτά αξιολογεί τη θέση του βάσει της αντικειμενικής συνάρτησης καταλληλότητας. Στη συνέχεια, κάθε σωματίδιο αποφασίζει πώς θα κινηθεί λαμβάνοντας υπ’ όψιν τη δική του τρέχουσα θέση, τη δική του προηγούμενη καλύτερη θέση και τις θέσεις κάποιων άλλων σωματιδίων του σμήνους. Η επόμενη επανάληψη γίνεται όταν όλα τα σωματίδια του σμήνους έχουν μετακινηθεί. Ως αποτέλεσμα, το σμήνος μετακινείται όλο μαζί, σαν ένα σμήνος πουλιών που αναζητεί τροφή και ολοένα πλησιάζει πιο κοντά στο στόχο.

Βασικές έννοιες.

Τα μέλη-πιθανές λύσεις ονομάζονται σωματίδια (particles) και το σύνολο των δυνατών λύσεων σε κάθε επανάληψη αποτελεί το σμήνος σωματιδίων(swarm).

Η αρχικοποίηση του πληθυσμού γίνεται τυχαία στο χώρο αναζήτησης του προβλήματος. Στη συνέχεια, σε κάθε βήμα ανανεώνεται η θέση του εκάστοτε σωματιδίου χρησιμοποιώντας:

- την καλύτερη προσωπική του θέση «ατομικό βέλτιστο»(personal, pbest), -τιμή που διατηρείται στη μνήμη- αξιολογώντας την μέσω της αντικειμενικής συνάρτησης καταλληλότητας.
- την καλύτερη θέση που έχει επιτευχθεί μέχρι στιγμής από οποιοδήποτε μέλος στο σύνολο του σμήνους «ολικό βέλτιστο»(global best, gbest), επίσης αξιολογούμενη μέσω της αντικειμενικής συνάρτησης -και επίσης αποθηκεύεται στη μνήμη. Στην περίπτωση που το σωματίδιο i θεωρεί μόνο ένα μέρος του πληθυσμού ως τοπολογικούς γείτονες, το «ολικό βέλτιστο» καλείται «τοπικό βέλτιστο»(local best, lbest). Έτσι προκύπτουν οι δύο εκδοχές pso, ολική και τοπική. Εμάς θα μας απασχολήσει η εκδοχή του ολικού βέλτιστου.(2,4/27),(3)

Θεωρώντας ότι ο χώρος αναζήτησής μας είναι D-διαστάσεων , τότε η θέση του i-στου σωματιδίου-πιθανής λύσης, του σμήνους N (swarm) δυνατών λύσεων, μπορεί να παρασταθεί με τη βοήθεια ενός D-διάστατου διανύσματος $X_i = (x_{i1}, x_{i2}, ..., x_{iD})^T$, όπου x_{id}^T είναι η θέση του i-οστού σωματιδίου, στην T-στη επανάληψη, ως προς την d μεταβλητή(διάσταση).

Η ταχύτητα(velocity) του σωματιδίου παριστάνεται από ένα άλλο D-διάστατο διάνυσμα $V_i = (v_{i1}, v_{i2}, ..., v_{iD})^T$,όπου v_{id}^T είναι η ταχύτητα του i-οστού σωματιδίου, στην T-στη επανάληψη, ως προς την d μεταβλητή(διάσταση).

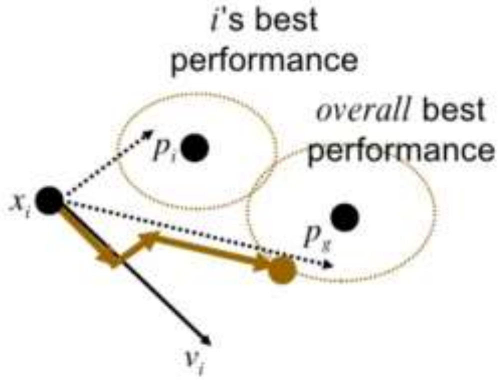
Ορίζοντας σαν δείκτη του καλύτερου σωματιδίου σε όλο το σμήνος το g (δηλ., το g-στο σωματίδιο έχει την καλύτερη αντικειμενική συνάρτηση καταλληλότητας) ,το κάθε σωματίδιο του σμήνους βρίσκει τον προσανατολισμό του-«πετάει» στο χώρο αναζήτησης χρησιμοποιώντας τις εξής δύο εξισώσεις:

$$v_{id}^{t+1} = v_{id}^{t} + c_1 r_1^t \left(p_{id}^t - x_{id}^t\right) + c_2 r_2^t \left(p_{gd}^t - x_{id}^t\right) \qquad (1)$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \qquad (2)$$

όπου d=1, 2, …, D; και D οι διαστάσεις του προβλήματός μας

όπου i=1, 2, …, N; και N είναι το μέγεθος του σμήνους(σχετίζεται με διαστάσεις του εκάστοτε προβλήματος συνήθως)



Σχήμα 12.

Οι c1 και c2 είναι δύο θετικές σταθερές επιτάχυνσης $0 < c_1, c_2 \leq 2$, που επιτρέπουν σε ένα σωματίδιο να ρυθμίσει την επιρροή που έχει η προσωπική καλύτερη θέση και η καλύτερη θέση του σμήνους αντίστοιχα, στην εξίσωση μεταβολής της ταχύτητας. Πιο συγκεκριμένα:

c1 ο συντελεστής *γνωστικότητας(cognitive),* ο οποίος λαμβάνει υπόψη την προσωπική εμπειρία της σωματιδίου και ρυθμίζει την κατεύθυνση του σωματιδίου ως της την προσωπική του καλύτερη θέση pbest.

C2 ο συντελεστής *κοινωνικότητας(social),* ο οποίος εκφράζει την επιρροή που έχει ένα σωματίδιο από το καλύτερο σωματίδιο του σμήνους και ρυθμίζει την κατεύθυνση του σωματιδίου ως της την ολικά καλύτερη θέση αυτή του gbest.

R_1, r_2 είναι τυχαίοι αριθμοί ομοιόμορφα κατανεμημένοι στο διάστημα [0,1], προσδίδουν ένα βαθμό τυχαιότητας, άρα και στοχαστικότητας στον τρόπο μετακίνησης του σμήνους;

Και n=1, 2, …, προσδιορίζει τον αριθμό των επαναλήψεων. (2)

Ο βασικός αλγόριθμος για την εύρεση του «ολικού βέλτιστου»:

Το πρώτο βήμα του PSO είναι, της και της περισσότερους αλγόριθμους, η αρχικοποίηση των μεταβλητών του. Πιο συγκεκριμένα της έχουμε προαναφέρει εκτός από την ρύθμιση των μεταβλητών ελέγχου του κάθε αλγόριθμου ξεχωριστά, επιδιώξαμε όλοι της οι αλγόριθμοι να τρέχουν υπό της της συνθήκες(αρχικός πληθυσμός, αριθμός αντικειμενικών συναρτήσεων καταλληλότητας, όριο σύγκλισης). Επομένως(1):

1) Αρχικοποίηση του πληθυσμού των σωματιδίων (particles) N=80 με τυχαία θέση (1) και ταχύτητα (2).
2) Για κάθε particle, υπολογίζουμε την τιμή της αντικειμενικής του συνάρτησης.

3) Συγκρίνουμε την τιμή της αντικειμενικής συνάρτησης που μόλις υπολογίσαμε, με την καλύτερη τιμή που έχει παρουσιάσει μέχρι στιγμής το συγκεκριμένο σωματίδιο(pbest). Εάν η τωρινή τιμή της αντικειμενικής συνάρτησης είναι καλύτερη(μικρότερη) της pbest, τότε αντικαθιστούμε την τιμή της pbest με την τελευταία υπολογισθείσα τιμή αντικειμενικής συνάρτησης.
4) Συγκρίνουμε την τιμή αντικειμενικής συνάρτησης με την καλύτερη τιμή που έχουν παρουσιάσει στο σύνολό της όλα τα σωματίδια μέχρι στιγμής(gbest). Εάν η τωρινή τιμή της αντικειμενικής συνάρτησης του συγκεκριμένου σωματιδίου είναι καλύτερη της gbest, τότε αντικαθιστούμε την gbest με αυτή και ανανεώνουμε τον δείκτη g.
5) Ανανεώνουμε της σχέσεις ταχύτητας(1) και θέσης(2) αντίστοιχα με τα νέα δεδομένα.
6) Επιστροφή στο βήμα (2) και αύξηση του μετρητή $t = t + 1$ μέχρι κάποιο κριτήριο τερματισμού να ικανοποιηθεί.

Σαν κριτήρια τερματισμού λειτουργούν η επίτευξη επιθυμητής τιμής αντικειμενικής συνάρτησης, ο μέγιστος αριθμός επαναλήψεων(γενεών) και η υπέρβαση του CPU χρόνου.

```
Begin

Αρχικοποίηση πληθυσμού d-διαστάσεων PSO: S

Do{

    for each particle i ε [1…s ]:?

        Calculate fitness value f ( xi )

        If   f  ( xi ) < f ( pi )  then pi := xi  ;

        If   f ( pi )  < f ( pg )   then pg := pi  ;

    end

    for i:=1 to S do?

        for d:=1 to n do?
```

$$u_{id}^{n1} = u_{id}^{n} + cr_1^n \left(p_{id}^n - x_{id}^n\right) + cr_2^n \left(p_{gd}^n - x_{id}^n\right) \quad (1)$$

$$x_{id}^{n+1} = x_{id}^{n} + u_{id}^{n+1} \quad (2)$$

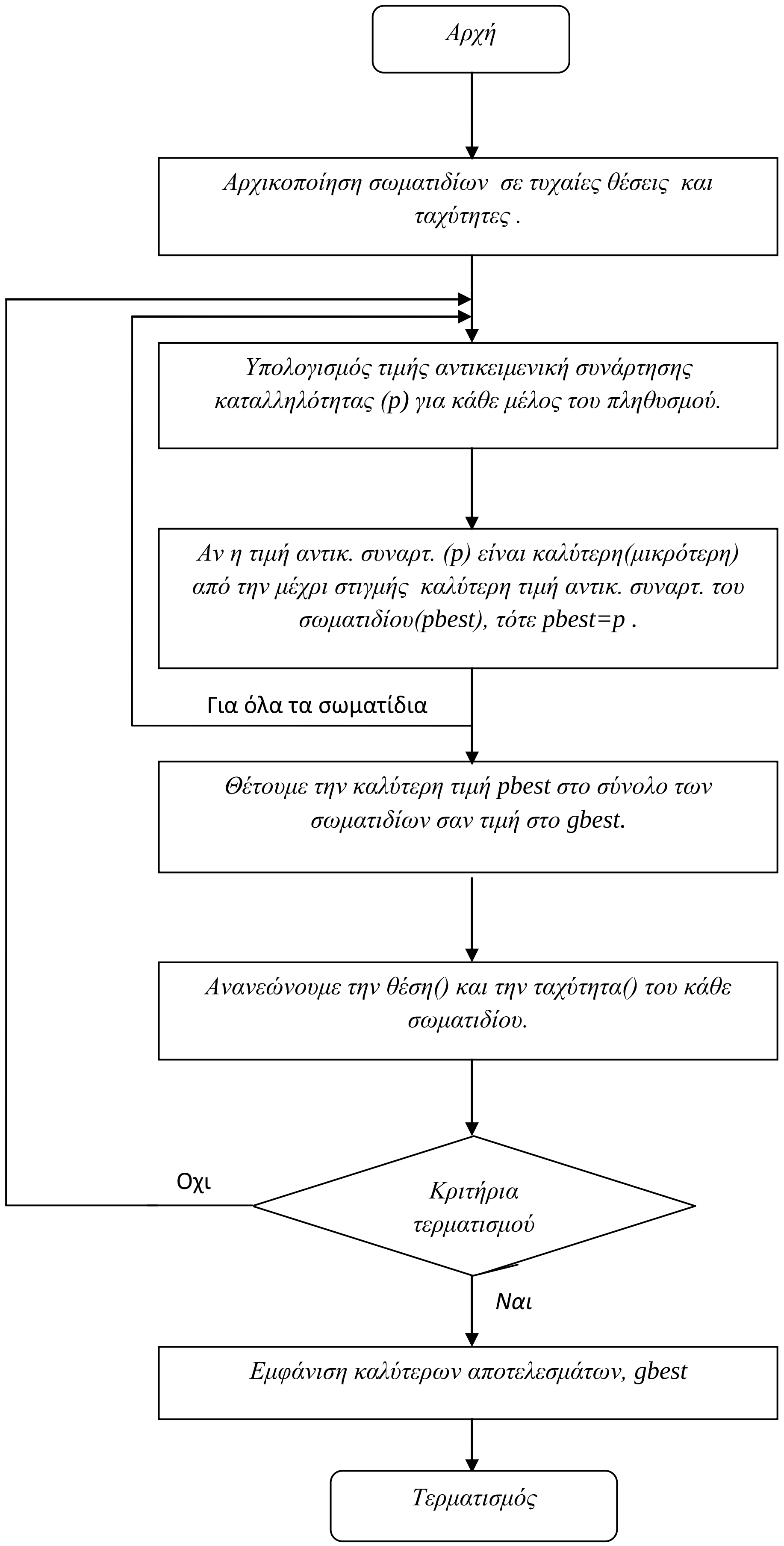
```
        end?

    end

    }While (termination criterion is not met)
```

Figure. Ο ψευδοκώδικας του βασικού PSO αλγόριθμου

Flow Diagram:



Γειτονιές και τοπολογίες

Οι εξισώσεις (1) και (2) χρησιμοποιούνται στη βασική έκδοση του PSO αλγόριθμου. Η κατάλληλη ρύθμιση των σταθερών επιτάχυνσης είναι σημαντική γιατί της είδαμε επηρεάζει την ταχύτητα σύγκλισης του αλγόριθμου και την εύρεση τοπικών βέλτιστων. Για c2=0, τα σωματίδια λειτουργούν ανεξάρτητα και πραγματοποιούν τοπική αναζήτηση(local search). Ενώ για c1=0, όλο το σμήνος λειτουργεί σαν ένα. Συνήθως οι σταθερές επιτάχυνσης c1 και c2 τίθονται ίσες ώστε να επιτευχθεί ισορροπία μεταξύ εξερεύνησης και εκμετάλλευσης στο σμήνος, αν και κατά βάση η σχέση μεταξύ των δύο τιμών εξαρτάται από το εκάστοτε πρόβλημα (Ενγκελ 2006). Στην αρχική έκδοση του PSO χρησιμοποιούνταν οι τιμές $c_1 = c_2 = 2$, της μετά από έρευνα αποδείχθηκε ότι καθ' αυτόν τον τρόπο συχνά τα σωματίδια ξέφευγαν από τον χώρο αναζήτησης.[2]

Παράμετρος περιορισμού της ταχύτητας.

Σημαντικό μέτρο για την εξακρίβωση της αποτελεσματικότητας του αλγόριθμου είναι η σχέση μεταξύ της ικανότητας αναζήτησης (exploration) και εκμετάλλευσης (exploitation). Η ικανότητα αναζήτησης ορίζει τη δυνατότητα διερεύνησης όλων των διαστάσεων του χώρου αναζήτησης και αναγνώρισης των πιο ελπιδοφόρων τμημάτων του χώρου, της διακρίνεται αδυναμία στην εύρεση ικανοποιητικού βέλτιστου σε αυτά τα ελπιδοφόρα τμήματα. Η εκμετάλλευση από την άλλη πλευρά, δηλώνει την τάση του σωματιδίου να επικεντρώνεται στην εύρεση τοπικού βέλτιστου και να μην αλληλεπιδρά και να μοιράζεται πληροφορίες με τα υπόλοιπα σωματίδια του σμήνους. Καθ' αυτόν τον τρόπο ο χώρος αναζήτησης δεν διερευνάται αποτελεσματικά και προκύπτουν μη αξιόπιστα αποτελέσματα. Για τον PSO το κλειδί είναι η ισορροπία ανάμεσα στην αναζήτηση και την εκμετάλλευση. Έτσι μεγάλη τιμή στον c2 π.χ., σημαίνει μεγαλύτερη βαρύτητα στο ολικό βέλτιστο(global best position)- βαρύτητα στην εξερεύνηση. Από την άλλη πλευρά, αν η γνωστική παράμετρος c1 ξεπερνά κατά πολύ την κοινωνική c2, βαρύτητα δίνεται στην τοπική εκμετάλλευση από τα σωματίδια.

Από τη στιγμή που δεν υπάρχει μηχανισμός που να ελέγχει την ταχύτητα του σωματιδίου, ήταν αναγκαία η χρήση μιας της παραμέτρου, της $V_{\max}$, που ορίζει το όριο της ταχύτητας, το οποίο εάν ξεπεραστεί η ταχύτητα τελικά ισούται με $V_{\max}$. Της αποδείχτηκε, η σωστή ρύθμιση της της παραμέτρου ήταν ιδιαίτερα σημαντική για την απόδοση του αλγόριθμου, καθώς για πολύ μεγάλες τιμές είναι πιθανό τα σωματίδια να προσπεράσουν καλές λύσεις, ενώ για ιδιαίτερα μικρές τιμές της $V_{\max}$ έχουν ως αποτέλεσμα τη μη αποτελεσματική εξερεύνηση του χώρου αναζήτησης. Το πρόβλημα αντιμετωπίζεται με την προσθήκη της παραμέτρου του βάρους αδράνειας w, $0.4 < w < 1.2$, και χρησιμοποιείται για να καθορίσει την επίδραση που έχει η προηγούμενες τιμές ταχύτητας στη νέα ταχύτητα, και ταυτόχρονα επιφέρει ισορροπία της ιδιότητες εκμετάλλευσης και εξερεύνησης των σωματιδίων. Η εξίσωση ταχύτητας(1) , μετασχηματίζεται ως εξής (Eberhart and Shi,98):

$$v_{id}^{n+1} = w * v_{id}^{n} + c_1 r_1^n \left(p_{id}^n - x_{id}^n\right) + c_2 r_2^n \left(p_{gd}^n - x_{id}^n\right) \qquad (3)$$

Ο Clerc ανέπτυξε μια παραλλαγή του PSO παρόμοια με αυτή του συντελεστή αδράνειας, για να ισοσταθμίσει την ικανότητα διερεύνησης/εκμετάλλευσης του αλγόριθμου. Η παραλλαγή αυτή βασίζεται σε έναν συντελεστή περιορισμού (constriction coefficient) στην εξίσωση ανανέωσης της ταχύτητας των σωματιδίων (Clerc et al,2002). Έτσι η εξίσωση ταχύτητας εκφράζεται ως:

$$v_{id}^{n+1} = \chi\left(v_{id}^{n} + c_1 r_1^n \left(p_{id}^n - x_{id}^n\right) + c_2 r_2^n \left(p_{gd}^n - x_{id}^n\right) \right) \qquad (4)$$

Όπου

$$\chi = \frac{2\kappa}{\left|2 - \phi - \sqrt{\phi^2 - 4\phi}\right|} \qquad (5)$$

με $\phi = \phi_1 + \phi_2 > 4$, $\phi_1 = c_1 * r_1$, $\phi_2 = c_2 * r_2$, και κ ε [0,1].

Όταν χρησιμοποιείται αυτή η μέθοδος, της και στην περίπτωσή της, τότε η τιμή του συντελεστή φ ισούται με 4.1, το χ ισούται με 0.7298 και οι συντελεστές c1 με 2,8 και c2 με 1,3. Έχει αποδειχτεί μέσα από έρευνες ότι αν φ≥4 και $\kappa \in [0,1]$ τότε το σμήνος συγκλίνει σε ένα σταθερό σημείο, ενώ η τιμή που προτείνεται για το χ είναι μεταξύ μηδέν και 1, δλδ., $\chi \in [0,1]$. Σημαντική είναι της η επίδραση του συντελεστή κ στη νέα μέθοδο, αφού όταν η τιμή του πλησιάζει στο μηδέν τότε παρατηρείται εκμετάλλευση και γρήγορη σύγκλιση του σμήνους, ενώ όταν η τιμή του είναι κοντά στο ένα παρατηρείται εξερεύνηση του χώρου έρευνας και αργή σύγκλιση του σμήνους, της και στο πρόβλημά της.

Το μέγεθος του σμήνους.

Της περισσότερες εφαρμογές του PSO αλγόριθμου, ο αριθμός των σωματιδίων κυμαίνεται στα 20-60. Οι Van den Bergh και Engelbrecht έχουν δείξει ότι παρά το γεγονός ότι υπάρχει μια μικρή βελτίωση στην βέλτιστη τιμή με την αύξηση του μεγέθους του σμήνους, ένα πολύ μεγάλο σμήνος θα αυξήσει το υπολογιστικό κόστος των αντικειμενικών συναρτήσεων κάτι που πιθανόν θα οδηγήσει σε σφάλμα το πρόγραμμα. Οι Eberhart και Shi υποστήριξαν ότι το μέγεθος του σμήνους ουσιαστικά δεν επηρεάζει την απόδοση της PSO μεθόδου. Εμείς τελικά ακολουθήσαμε την άποψη των Parsopoulos και Vrahatis που συγκλίνει με τη γενική έρευνα που είχαμε κάνει και ορίσει της αρχικές συνθήκες ότι το μέγεθος του σμήνους θα πρέπει να είναι ίσο με 10D για όλα τα προβλήματα(π.χ. για ένα 10-διάστατο πρόβλημα, είναι ίσο με 100).

3.2.f Shuffled Complex Evolution

Ο στόχος του SCE αλγόριθμου που αναπτύχθηκε το 1992 από τον Duan είναι να βρει το μοναδικό βέλτιστο σετ παραμέτρων μέσα στο δειγματικό χώρο, με κριτήριο την τιμή της αντικειμενικής συνάρτησης καταλληλότητας.

Η SCE-UA μέθοδος βασίζεται σε τέσσερις «ιδέες»:

- Συνδυασμό ντετερμινιστικών και τυχαίων προσεγγίσεων.
- Συστηματική εξέλιξη της συνόλου σημείων με στόχο την ολική βελτίωση.
- Ανταγωνιστική εξέλιξη. (Competitive evolution)
- Μετατόπιση συνόλων. (Complex shuffling)

Η χρήση ντετερμινιστικών προσεγγίσεων επιτρέπει στον SCE αλγόριθμο να χρησιμοποιήσει αποδοτικά πληροφορίες για την εξερεύνηση του χώρου, ενώ με της τυχαίες προσεγγίσεις επιτυγχάνει ευελιξία και αποδοτικότητα. Η αναζήτηση ξεκινά με ένα τυχαία επιλεγμένο σύνολο σημείων από το δειγματικό χώρο Ω. Η μετατόπιση συνόλων επιτρέπει την πρόσβαση της πληροφορίες κάθε συνόλου καθ’όλη τη διάρκεια της έρευνας βοηθώντας έτσι ώστε να μην παγιδεύεται ο αλγόριθμος σε τοπικά βέλτιστες λύσεις. Οι στρατηγικές συστηματικής εξέλιξης εξασφαλίζουν την αναζήτηση του χώρου βάσει της αντικειμενικής συνάρτησης, ενώ η εφαρμογή της ανταγωνιστικής εξέλιξης ενισχύει την σύγκλιση στο ολικό max/min. Με λίγα λόγια ο συνδυασμός των παραπάνω τεχνικών διασφαλίζει ότι η πληροφορία που περιέχεται στο δείγμα αξιοποιείται πλήρως, καθώς και ότι η πληροφορία αυτή δεν εκφυλίζεται. Αυτό καθιστά τον SCE ιδιαίτερα αποτελεσματικό και ευέλικτο για μεγάλο αριθμό προβλημάτων.

<u>Σχέση SCE-CRS</u>

Η SCE μέθοδος συνδυάζει τα πλεονεκτήματα του CRS (Price,) αλγόριθμου με τη στρατηγική του competitive evolution (Holland,1992) και του complex shuffling (Duan, Sorooshian, 1993).

Ο CRS αλγόριθμος είναι το σημείο αφετηρίας καθώς έχει κάποια από τα χαρακτηριστικά που προαναφέρθηκαν και είναι ιδιαίτερα εύκολο να προσαρμοστεί και να χρησιμοποιηθεί στην εκάστοτε περίσταση και να είναι αποτελεσματικός στην σύγκλιση στο ολικό βέλτιστο.

Της ταυτόχρονα έχει και αρκετές αδυναμίες: η αργή σύγκλιση, το γεγονός ότι εύκολα εγκλωβίζεται σε τοπικά ακρότατα, της και το γεγονός ότι ο πληθυσμός τείνει να μειώνεται τάχιστα αφού σε κάθε επανάληψη αντικαθίσταται το χειρότερο σημείο ολόκληρου του πληθυσμού. Πειράματα έδειξαν ότι είναι πιθανόν πολύ σύντομα ο πληθυσμός να εμφανίζει επαναλαμβανόμενα σημεία, με αποτέλεσμα τον εκφυλισμό του και τον πρόωρο τερματισμό του. Η δομή του SCE αλγορίθμου εξαλείφει τα παραπάνω προβλήματα.

Σύντομη περιγραφή

Ο αλγόριθμος ξεκινά με έναν τυχαία επιλεγμένο πληθυσμό σημείων στον προκαθορισμένο δειγματικό χώρο. Τα σημεία ταξινομούνται σε σειρά αύξουσας τιμής αντικειμενικής συνάρτησης καταλληλότητας, έτσι ώστε το πρώτο σημείο να έχει τη μικρότερη τιμή και το τελευταίο τη μεγαλύτερη. Ο τυχαίος της πληθυσμός διαχωρίζεται σε αρκετά σύνολα. Κάθε σύνολο επιτρέπεται να εξελιχθεί ανεξάρτητα και να διερευνήσει το δειγματικό χώρο σε διάφορες κατευθύνσεις. Κάθε μεμονωμένο σημείο του συνόλου έχει τη δυνατότητα να συμμετάσχει στη διαδικασία παραγωγής της νέου σημείου. Από κάθε σύνολο, κάποια σημεία επιλέγονται να συνθέσουν ένα υποσύνολο, όπου εφαρμόζεται η Nelder and Mead Simplex Method(NMSM, 1965) με στόχο την ολική βελτιστοποίηση. Τα «καλύτερα» σημεία έχουν μεγαλύτερες πιθανότητες να επιλεγούν για την παραγωγή νέων απογόνων. Η NMSM εφαρμόζει reflection και contraction βήματα με στόχο την παραγωγή της πιο βελτιωμένου απόγονου. Αυτό το νέο σημείο/απόγονος αντικαθιστά το σημείο με τη χειρότερη επίδοση στο σύνολο. Τα σημεία των συνόλων που έχουν χρησιμοποιηθεί/εξελιχθεί συγκεντρώνονται και ταξινομούνται εκ νέου, μετατοπίζονται και ξαναμοιράζονται σε νέα σύνολα προκειμένου να διαμοιράσουν την καινούρια πληροφορία. Η διαδικασία αυτή επαναλαμβάνεται μέχρι κάποιο κριτήριο τερματισμού να ικανοποιηθεί.

Μέθοδος

NGS = αριθμός συνόλων στον αρχικό πληθυσμό

NPG = αριθμός σημείων σε κάθε σύνολο

NPT = αριθμός σημείων σε ολόκληρο τον αρχικό πληθυσμό, NPT=NGS*NPG

NPS = αριθμός σημείων σε κάθε υπο-σύνολο

MINGS = ελάχιστος αριθμός συνόλων που απαιτείται

1. Αρχικοποίηση του δείγματος: ο αλγόριθμος παράγει ένα αρχικό πληθυσμό NPT σημείων ακολουθώντας της περιορισμούς που του έχουμε θέσει και στη συνέχεια υπολογίζει την τιμή της αντικειμενικής συνάρτησης στο εκάστοτε σημείο.
   Στην πορεία όπου αλλάξαμε της συνθήκες των πειραμάτων ο αλγόριθμος ξεκινάει με ένα standard αρχικό πληθυσμό ο οποίος είναι κοινός για όλους της αλγόριθμους.
2. Ταξινόμηση: ταξινόμηση των NPT σημείων κατά αύξουσα σειρά, έτσι ώστε το πρώτο σημείο να έχει τη μικρότερη τιμή αντικειμενικής συνάρτησης(το καλύτερο) και το τελευταίο τη μεγαλύτερη τιμή αντικειμενικής συνάρτησης(το χειρότερο).

3. Διαχωρισμός σε σύνολα: ταξινόμηση των NPT σημείων σε NGS σύνολα, καθένα από τα οποία περιλαμβάνει NPG σημεία. Τα σύνολα χωρίζονται με τέτοιο τρόπο ώστε το πρώτο σύνολο να περιλαμβάνει κάθε NGS*(k+1)+1 ταξινομημένο σημείο, το δεύτερο σύνολο περιλαμβάνει κάθε NGS*(k+1)+2 ταξινομημένο σημείο κ.ο.κ, όπου k = 1,2,…, NPG.
4. Ανάπτυξη κάθε συνόλου: ανάπτυξη κάθε συνόλου ανεξάρτητα ακολουθώντας τα βήματα ανάπτυξης του CCE αλγόριθμου.
5. Αναδιάταξη συνόλων: συνδυασμός των σημείων των ανεπτυγμένων πλέον συνόλων σε ένα μοναδικό δειγματικό πληθυσμό. Αναδιάταξη των σημείων του πληθυσμού κατά αύξουσα τιμή αντικειμενικής συνάρτησης.
6. Έλεγχος σύγκλισης: Εάν κάποιο από τα κριτήρια ικανοποιείται ο αλγόριθμος τερματίζεται διαφορετικά συνεχίζει στο βήμα 4 (ή στο 7).
7. (Νέα προσθήκη) Έλεγχος μείωσης αριθμού συνόλων: Αν MINGS<NGS, απομάκρυνση του συνόλου με τα τελευταία ταξινομημένα σημεία. Θέτουμε NGS= NGS-1 και NPT=NGS*NPG και επιστροφή στο βήμα 4. Αν MINGS=NGS, επιστροφή στο βήμα 4.

Σημείωση: Το βήμα 7, όπου αφού ολοκληρωθεί της κύκλος του αλγόριθμου και δεν ικανοποιείται κάποιο από τα κριτήρια τερματισμού, αφαιρείται το τελευταίο σύνολο, αυτό δλδ. Με τα χειρότερα σημεία. Η διαδικασία αυτή συνεχίζεται μέχρι να έχει απομείνει ο ελάχιστος αριθμός συνόλων MINGS.

Δοκιμάσαμε την παραπάνω τεχνική στο πρόβλημά της, αλλά δεν διαπιστώσαμε βελτίωση στην έξοδο του αλγόριθμου, οπότε και την απορρίψαμε.

1. NPS σημεία επιλέγονται σύμφωνα με μία κατανομή πιθανοτήτων τα οποία απαρτίζουν το υπο-σύνολο το οποίο πρόκειται να αναπτυχθεί. Λόγω της κατανομής τα καλύτερα σημεία- αυτά με τη μικρότερη αντικειμενική συνάρτηση- έχουν μεγαλύτερη πιθανότητα να επιλεγούν να συνθέσουν το υπο-σύνολο, από ότι τα χειρότερα σημεία, αυτά δλδ. Με τη μεγαλύτερη αντικειμενική συνάρτηση.
2. Υπάρχουν 3 τύποι ανάπτυξης των υποσυνόλων: μέσω reflection, contraction και mutation. Το «reflecting» βήμα επιτυγχάνεται της την ανάκλαση του χειρότερου σημείου του υπο-συνόλου μέσω του «κεντροϊδούς» των υπόλοιπων στοιχείων του υπο-συνόλου. Εάν το ανακλώμενο σημείο έχει μικρότερη τιμή αντικειμενικής συνάρτησης από το χειρότερο σημείο, το χειρότερο σημείο εξαιρείται και αντικαθίσταται από το νέο σημείο. Στην αντίθετη περίπτωση όπου το νέο σημείο δεν είναι βελτιωμένο πάμε στο βήμα 4. Τώρα στην περίπτωση που το νέο σημείο βρίσκεται έξω από το δειγματικό χώρο πάμε στο βήμα 3.
3. Ένα νέο σημείο που βρίσκεται ανάμεσα στο χειρότερο σημείο και το «κεντροειδές» των υπολοίπων, παράγεται κατά το «contraction» βήμα. Εάν το νέο σημείο και πάλι δεν έχει μικρότερη τιμή αντικειμενικής συνάρτησης από το χειρότερο σημείο, πάμε στο βήμα 4 .

4. Κατά το «mutation» βήμα παράγεται τυχαία ένα σημείο σύμφωνα με της περιορισμούς για να αντικαταστήσει το χειρότερο σημείο του υπο-συνόλου.
5. Επανέλαβε τα βήματα 2-4 α φορές, $\alpha \geq 1$ όπου α ο αριθμός των απογόνων που κάθε υποσύνολο παράγει(και αντικαθιστά) πριν επιστρέψει στο σύνολο.
6. Επανέλαβε τα βήματα 1-4 β φορές, $\beta \geq 1$ όσες και τα βήματα εξέλιξης του κάθε συνόλου πριν αυτά αναδιαταχθούν.

Παράμετροι:

Οι παράμετροι του αλγόριθμου είναι: m: αριθμός σημείων συνόλου, q: αριθμός σημείων υποσυνόλου, p: αριθμός συνόλων, $p_{min:}$ ελάχιστος αριθμός συνόλων που απαιτούνται στον πληθυσμό, α: αριθμός συνεχόμενων απογόνων που παράγονται από κάθε υποσύνολο, β: αριθμός βημάτων εξέλιξης από το κάθε σύνολο.

Θεωρητικά ο αριθμός σημείων σε κάθε σύνολο, m, μπορεί να πάρει οποιαδήποτε τιμή μεγαλύτερη ή ίση του δύο. Ο αριθμός των σημείων σε κάθε υποσύνολο q, μπορεί να κυμαίνεται μεταξύ 2 και m. Ο αριθμός των απογόνων α, μπορεί να είναι οποιοσδήποτε αριθμός ίσος ή μεγαλύτερος του 1. Αν α=1 μόνο της από της αρχικούς γονείς θα αντικατασταθεί, όσο της μεγαλύτερο α τόσο δίνεται βαρύτητα στην εύρεση τοπικού ελαχίστου. Ο αριθμός των βημάτων εξέλιξης β μπορεί να είναι οποιοσδήποτε θετικός ακέραιος. Αν ο β είναι μικρός τα σύνολα αναδιατάσσονται συχνά, αλλά δεν ενισχύεται η ανεξάρτητη διερεύνηση του χώρου των παραμέτρων. Εάν το β είναι μεγάλο, κάθε σύνολο θα συρρικνωθεί γρήγορα σε μία μικρή ομάδα σημείων και η αποτελεσματικότητα της ολικής αναζήτησης μπορεί να χαθεί. Η παράμετρος p_{min} κυμαίνεται μεταξύ του ένα και του p. Όσο η διερευνητική διαδικασία εξελίσσεται τόσο ο πληθυσμός συγκλίνει σε μικρότερο χώρο και λιγότερα σημεία απαιτούνται για να συνθέσουν ένα επαρκή πληθυσμό.

Κατάλληλες τιμές για της παραπάνω παραμέτρους προτείνονται από τον Duan(1994) : NPG(m)= (2n+1), NPS(q)=(n+1), α=1και β=(2n+1), όπου n: αριθμός διαστάσεων του προβλήματος. Έτσι σαν αποτέλεσμα έχουμε ότι ο χρήστης οφείλει να ρυθμίσει μόνο α) τον αριθμό των συνόλων NGS(p) και β) το μέγεθος του πληθυσμού s, που της ορίζει τον αριθμό των σημείων σε κάθε σύνολο. Ο Vrugt(2003) συστήνει τη χρήση μεγάλων πληθυσμών προκειμένου να εξασφαλίσουμε καλύτερη διερεύνηση του δειγματικού χώρου, χωρίς της υπερβολές αφού μπορεί να οδηγήσουν σε χρονοβόρες και μη περεταίρω αποδοτικές διεργασίες.

Κριτήρια τερματισμού:

Σε περίπτωση που ο αλγόριθμος έχει υπολογίσει ένα μέγιστο αριθμό αντικειμενικών συναρτήσεων ή αν μετά από ένα αριθμό επαναλήψεων το ποσοστό μεταβολής στην τιμή της αντικειμενικής συνάρτησης είναι μικρότερο από ένα προκαθορισμένο αριθμό(tolerance), τότε ο αλγόριθμός της τερματίζεται.

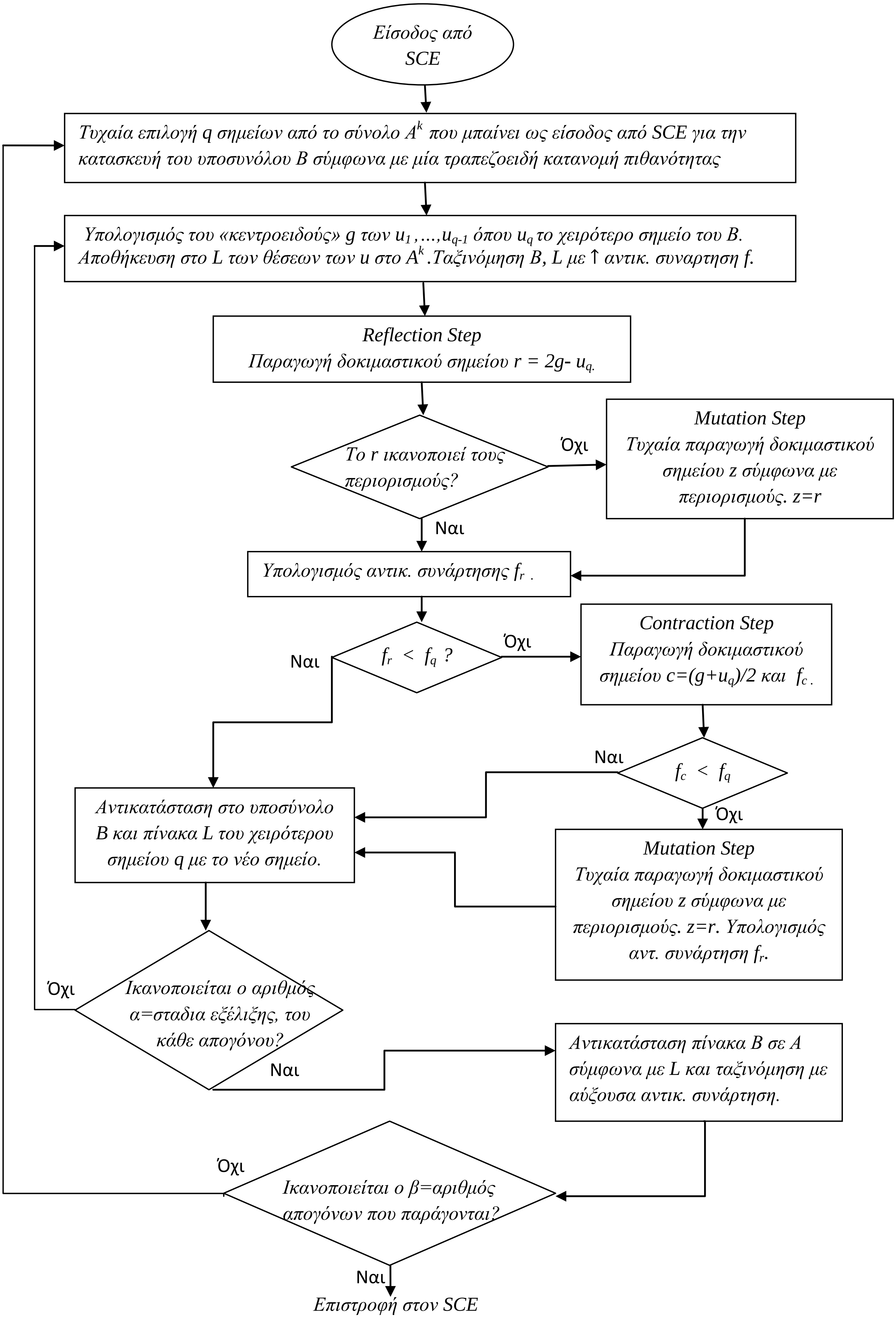
Ανάλυση αλγόριθμου

Η προσέγγιση του SCE αλγόριθμου ακολουθεί την ιδέα της φυσικής εξέλιξης. Της αρκετά μεγάλος αριθμός σημείων θα εξασφαλίσει τη συλλογή πληροφορίας σχετικά με τον αριθμό, τη θέση και το μέγεθος των περιοχών μεγαλύτερου ενδιαφέροντος. Ο αρχικός δειγματικός πληθυσμός NPT συνίσταται από αρκετές κοινωνίες, καθεμία από της οποίες αναπτύσσεται/εξελίσσεται ανεξάρτητα και αυτό επιτρέπει την καλύτερη κάλυψη του δειγματικού χώρου. Μετά από ένα αριθμό γενεών, οι κοινωνίες αναγκάζονται να αναμειχθούν με αποτέλεσμα να προκύπτουν νέες μέσω του shuffling. Αποτέλεσμα της παραπάνω διαδικασίας είναι να ενισχύεται η επιβίωση αφού διαμοιράζεται η πληροφορία που συλλέγεται ανεξάρτητα από την κάθε κοινωνία.

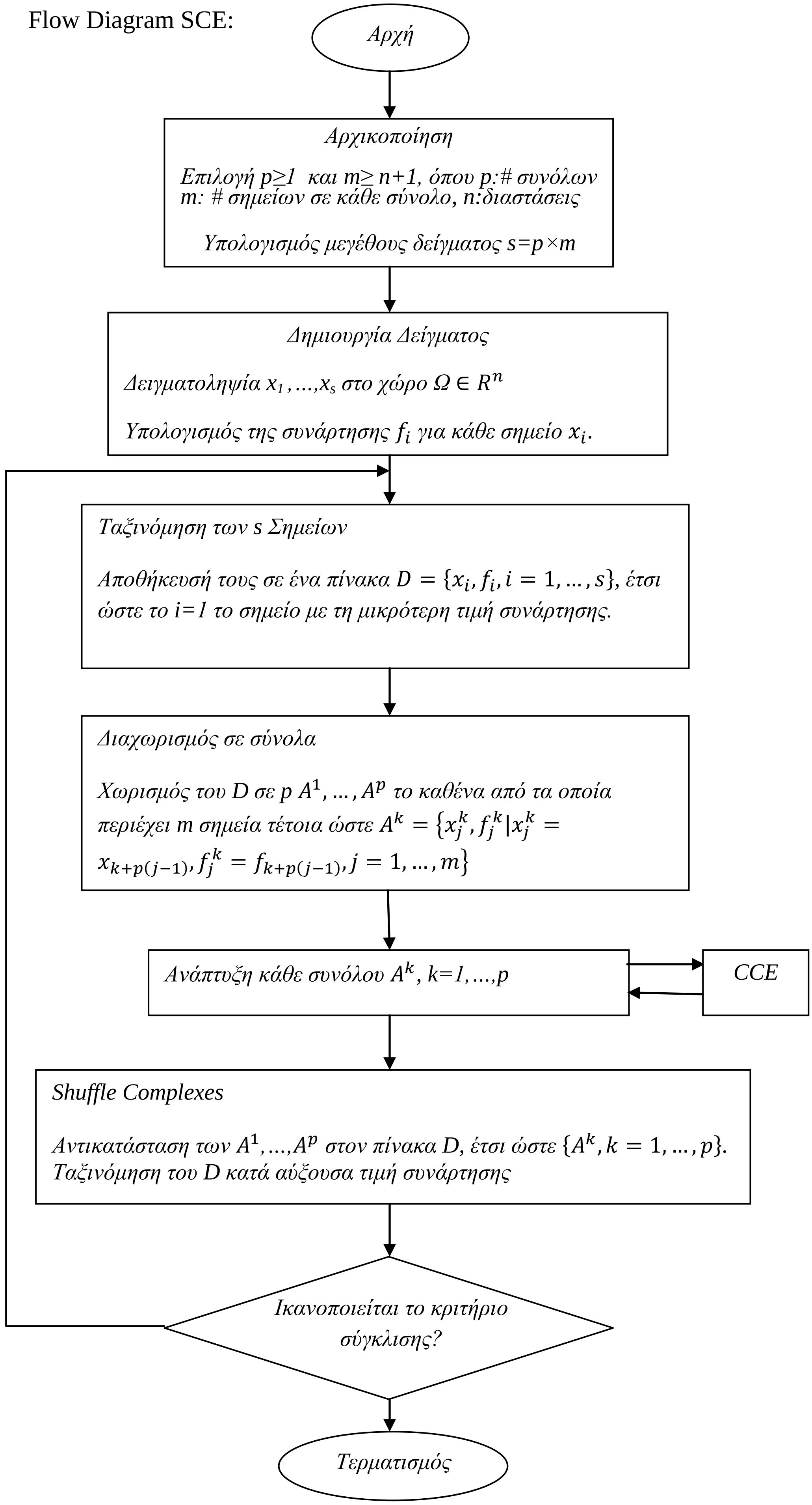
Κάθε μέλος της κοινωνίας/σύνολο είναι της εν δυνάμει γονέας με την ικανότητα να συμμετάσχει στη διαδικασία της αναπαραγωγής. Ένα υποσύνολο επιλεγμένο από το σύνολο είναι σαν ένα ζευγάρι γονέων, μόνο που το υποσύνολο μπορεί να αποτελείται από περισσότερα από δύο μέλη. Προκειμένου να εξασφαλίσουμε την ανταγωνιστικότητα της διαδικασίας εξέλιξης, απαιτείται η αύξηση πιθανότητας συμμετοχής των καλύτερων σημείων/μελών στην παραγωγή απογόνων. Επιπρόσθετα, το γεγονός ότι νέοι απόγονοι εμφανίζονται σε τυχαία σημεία στο δειγματικό χώρο κάτω από δεδομένες συνθήκες, διασφαλίζει τη διερεύνηση όλου του χώρου και τον μη εγκλωβισμό σε «ανέλπιδες» περιοχές. Της βοηθά στο να αυξηθεί η ποσότητα της πληροφορίας που αποθηκεύεται σε κάθε δείγμα. Το γεγονός ότι κάθε της απόγονος αντικαθιστά το χειρότερο σημείο του εκάστοτε υπο-συνόλου, και όχι ολόκληρου του πληθυσμού, διασφαλίζει ότι κάθε γονέας τουλάχιστον μία φορά θα συνεισφέρει στην διαδικασία εξέλιξης πριν τελικά απορριφθεί ή αντικατασταθεί. Κατ' αυτόν τον τρόπο καμία πληροφορία του δείγματος δεν αγνοείται.

[1] [2] [3] [4] [5]

Flow Diagram CCE:


Είσοδος από SCE
Τυχαία επιλογή q σημείων από το σύνολο A^k που μπαίνει ως είσοδος από SCE για την κατασκευή του υποσυνόλου B σύμφωνα με μία τραπεζοειδή κατανομή πιθανότητας
Υπολογισμός του «κεντροειδούς» g των u1 ,…,uq-1 όπου uq το χειρότερο σημείο του B. Αποθήκευση στο L των θέσεων των u στο A^k .Ταξινόμηση B, L με ↑ αντικ. συναρτηση f.
Reflection Step
Παραγωγή δοκιμαστικού σημείου r = 2g- uq.
Το r ικανοποιεί τους περιορισμούς?
Όχι
Mutation Step
Τυχαία παραγωγή δοκιμαστικού σημείου z σύμφωνα με περιορισμούς. z=r
Ναι
Υπολογισμός αντικ. συνάρτησης fr .
fr < fq ?
Ναι
Όχι
Contraction Step
Παραγωγή δοκιμαστικού σημείου c=(g+uq)/2 και fc .
fc < fq
Ναι
Όχι
Αντικατάσταση στο υποσύνολο B και πίνακα L του χειρότερου σημείου q με το νέο σημείο.
Mutation Step
Τυχαία παραγωγή δοκιμαστικού σημείου z σύμφωνα με περιορισμούς. z=r. Υπολογισμός αντ. συνάρτηση fr.
Ικανοποιείται ο αριθμός α=σταδια εξέλιξης, του κάθε απογόνου?
Όχι
Ναι
Αντικατάσταση πίνακα B σε A σύμφωνα με L και ταξινόμηση με αύξουσα αντικ. συνάρτηση.
Ικανοποιείται ο β=αριθμός απογόνων που παράγονται?
Όχι
Ναι
Επιστροφή στον SCE

Flow Diagram SCE:



ΚΕΦΑΛΑΙΟ 4

4.4 ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΣΩΤΕΡΙΚΩΝ ΠΑΡΑΜΕΤΡΩΝ.

Για καθένα από της παραπάνω αλγόριθμους έγινε η αντίστοιχη εκτενής έρευνα στη βιβλιογραφία. Στόχος να ρυθμιστούν, κατάλληλα για το πρόβλημά της, οι παράμετροι του κάθε αλγόριθμου ξεχωριστά,. Σε περιπτώσεις όπου αυτό δεν ήταν ξεκάθαρο από της διάφορες σχετικές μελέτες, εμείς μέσα από σειρά δοκιμών καταλήξαμε της πλέον πιθανόν κατάλληλες τιμές παραμέτρων για εμάς. Πιο συγκεκριμένα τέτοιου είδους δοκιμές χρειάστηκε να κάνουμε στην περίπτωση των SCE και DE αλγόριθμων. Ακολουθούν τα αποτελέσματα:

- Στον SCE η μέθοδος Mings εξετάστηκε για ngs=5 και ngs=8 και απορρίφθηκε, αφού τα αποτελέσματα απογοητευτικά.
- Στον DE ελέχθηκαν οι τιμές για της παραμέτρους
  F_weight = 0.4, 0.7, 0.9
  F_cr = 0.8, 0.95
  Strategy = 1, 2, 3, 4, 5
  χρησιμοποιώντας 2 πειραματικές καμπύλες. Δεν αποδείχθηκε συσχέτιση των τιμών των παραμέτρων με την επίδοση του αλγόριθμου.

4.5 ΚΟΙΝΕΣ ΣΥΝΘΗΚΕΣ.

Στη συνέχεια υπό κοινές συνθήκες/περιορισμούς εφαρμόσαμε της παραπάνω αλγόριθμους με στόχο να βρούμε της τιμές για τα δύο σετ παραμέτρων(5 και 8 αντίστοιχα), για δύο διαφορετικά σετ περιορισμών. Στην κάθε περίπτωση, τρέξαμε συνολικά 32 φορές τον κάθε αλγόριθμο, δλδ. Τέσσερις φορές για το ίδιο σετ πειραματικών τιμών- 8 καμπυλών.

Set_constr 1:

- 7200 function evaluations
- 3 hours (180’)
- convergence of objective function

Set_constr 2:

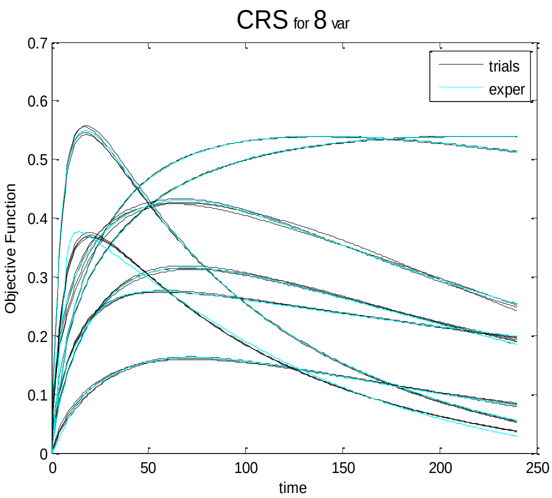
- 3 hours(180’)
- 0.001 convergence of objective function

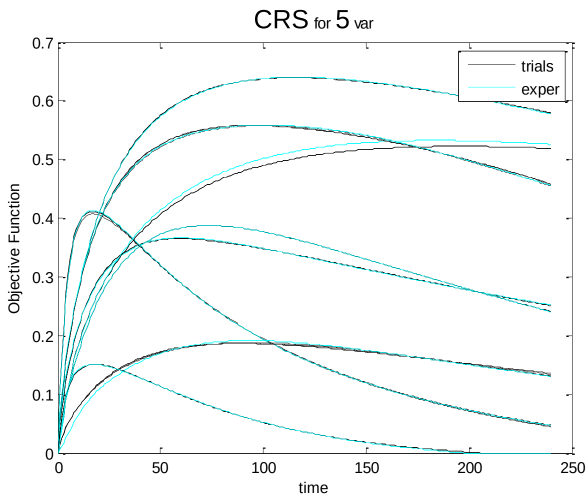
Όλοι οι αλγόριθμοι έτρεξαν στο ίδιο προγραμματιστικό περιβάλλον προκειμένου να επιτευχθεί αξιοπιστο στατιστικό αποτέλεσμα. Χρησιμοποιήθηκε i7 8-thread CPU στα 2.66GHz ανά πυρήνα με 8GB RAM μνήμη σε Matlab 2012a (7 πυρήνες).
Για κάθε πειραματική καμπύλη θεωρούμε γνωστές της τιμές των παραμέτρων που την σχηματίζουν. Εφαρμόζουμε λοιπόν της αλγόριθμους με στόχο να εντοπίσουμε τον πλέον αποδοτικό και χρησιμοποιούμε τα εξής κριτήρια:

- Τη σύγκλιση παραμέτρων
- Τη σύγκλιση καμπύλης
- Το χρόνο προσομοίωσης/αριθμό συναρτήσεων κόστους

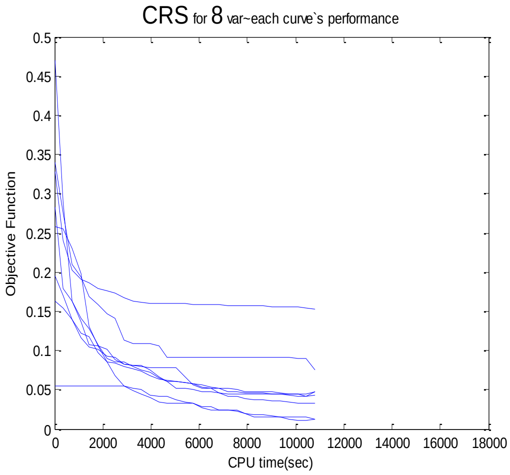
4.6 ΣΥΓΚΛΙΣΗ ΚΑΜΠΥΛΗΣ ΚΑΙ ΧΡΟΝΟΣ

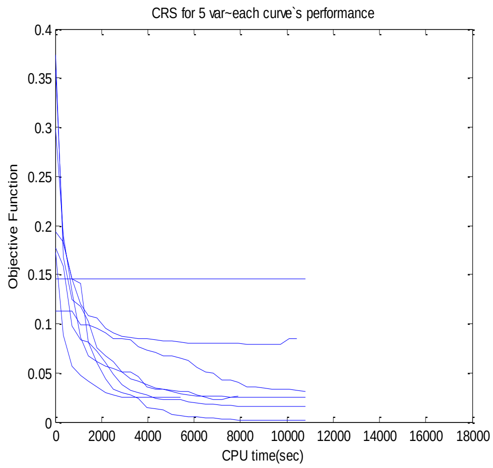
Αποτελέσματα για CRS





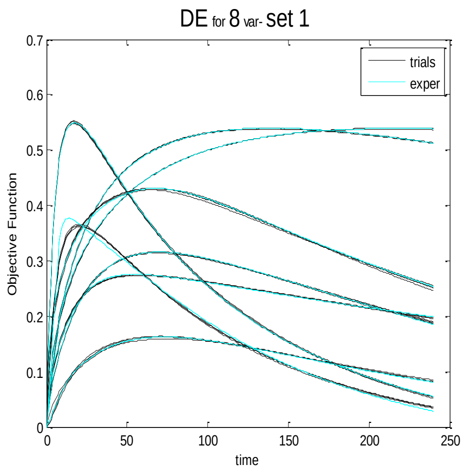
Οι παραπάνω γραφικές απεικονίζουν το fitting των πειραματικών δεδομένων με τη βέλτιστη καμπύλη του μοντέλου για 8 και 5 παραμέτρους αντίστοιχα. Όπου για τον CRS διαπιστώνουμε καλύτερη σύγκλιση της 5 παραμέτρους.

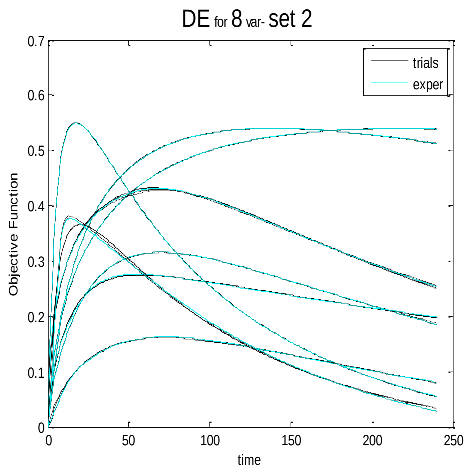


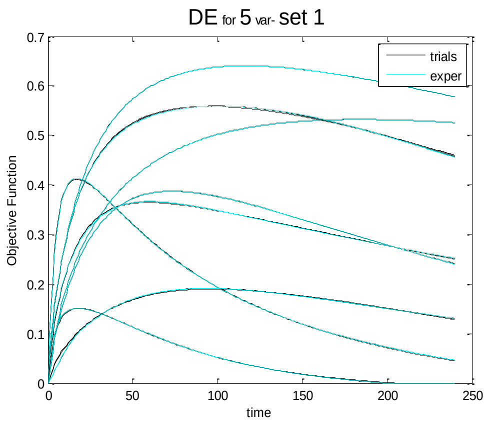


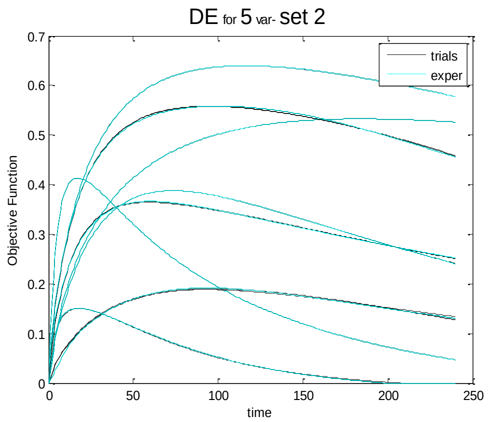
Στην προκειμένη απεικονίζεται η σύγκλιση της κάθε της από της 8 πειραματικές καμπύλες μέσα στο χρόνο. Διαπιστώνουμε καλύτερη σύγκλιση της 5 παραμέτρους.

Αποτελέσματα για DE

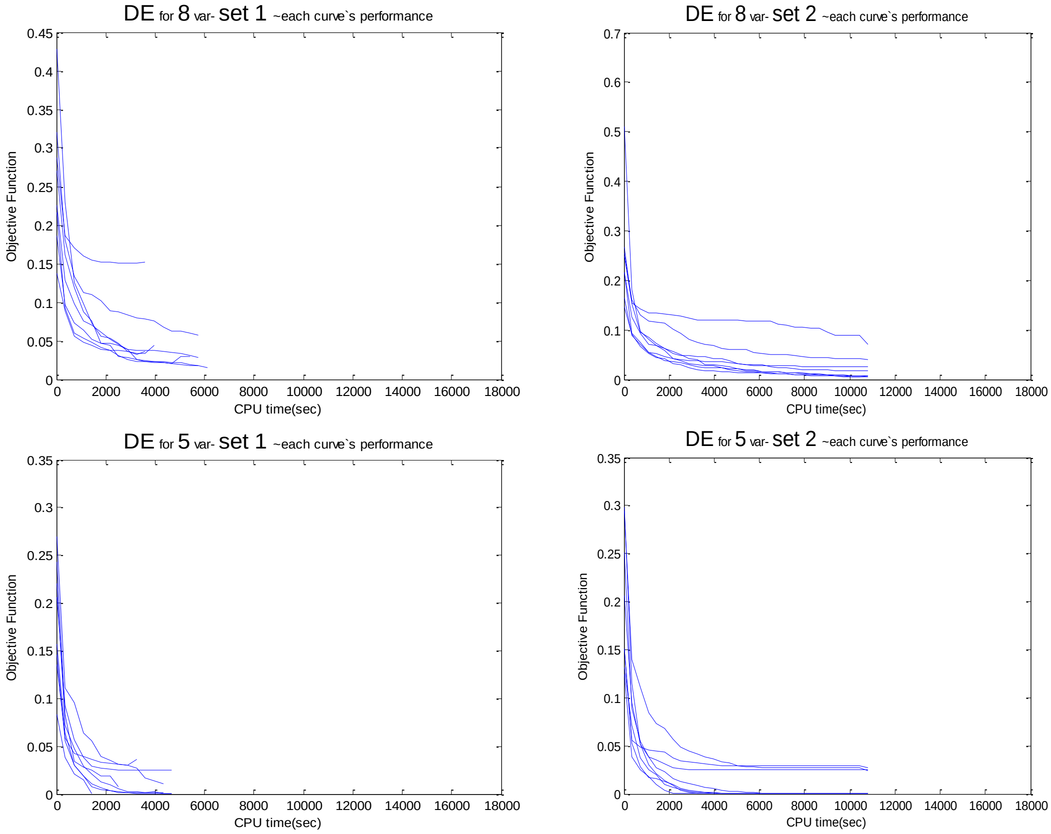






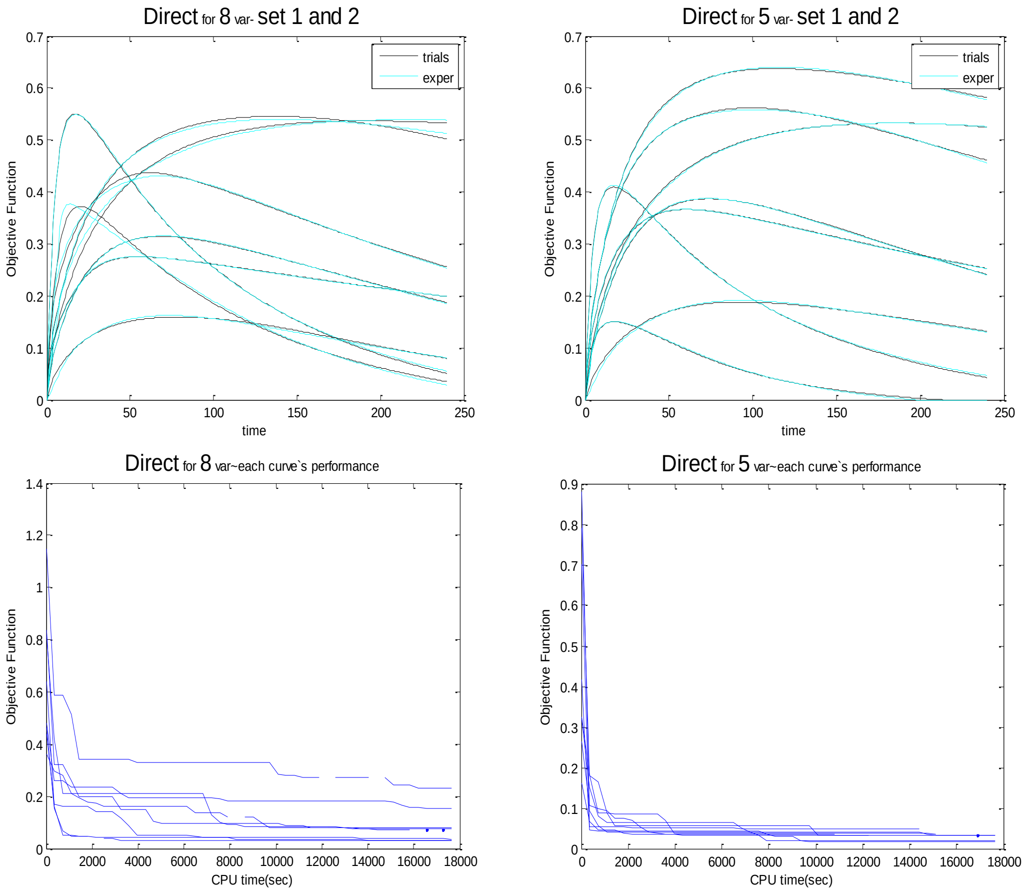


Καλύτερη σύγκλιση στο $2^{ο}$ σετ για 8 και για 5 παραμέτρους.



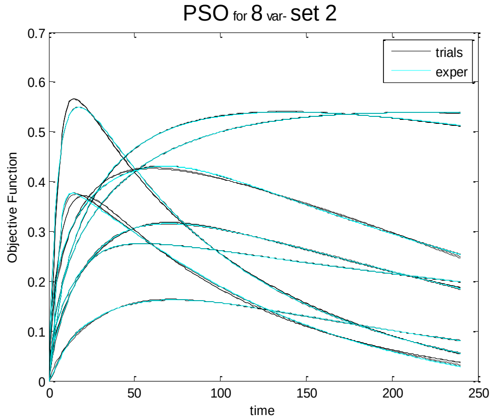
Ταχύτερη και καλύτερη σύγκλιση παρατηρείται για 5 παραμέτρους και ειδικότερα για το $2^{ο}$ σετ περιορισμών.

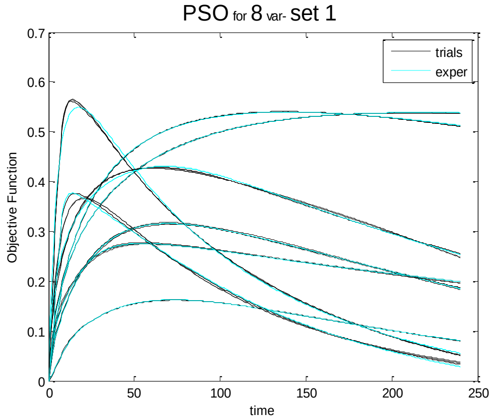
Αποτελέσματα για DIR

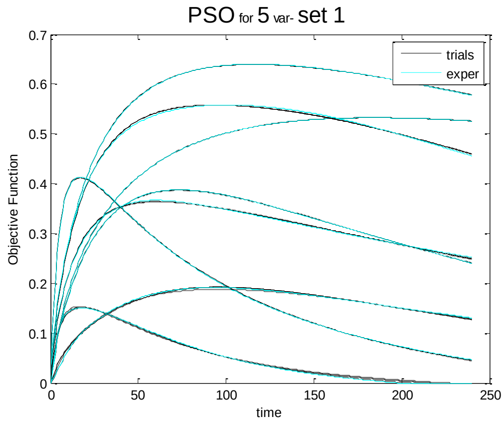


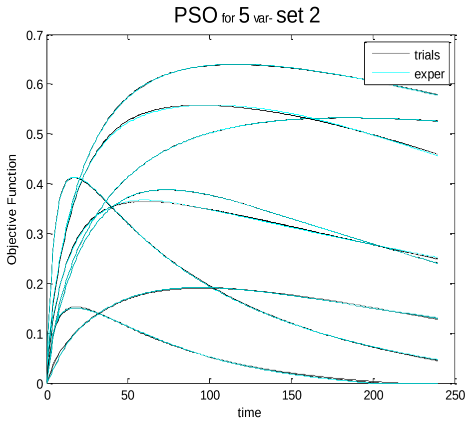
Είναι εμφανής η ταχύτερη και βέλτιστη σύγκλιση για τον ντερμινιστικό αλγόριθμο για 5 παραμέτρους.

Αποτελέσματα για PSO

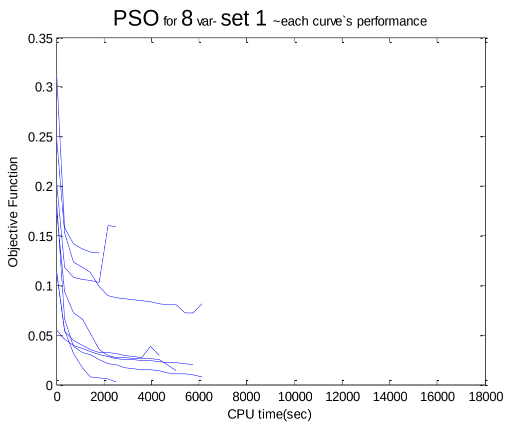


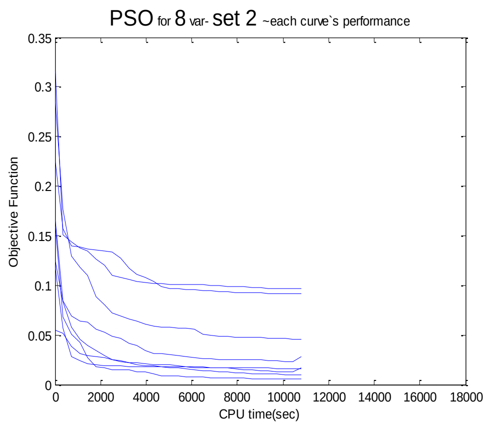


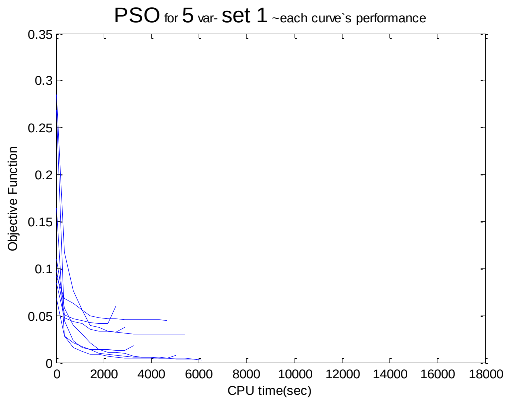


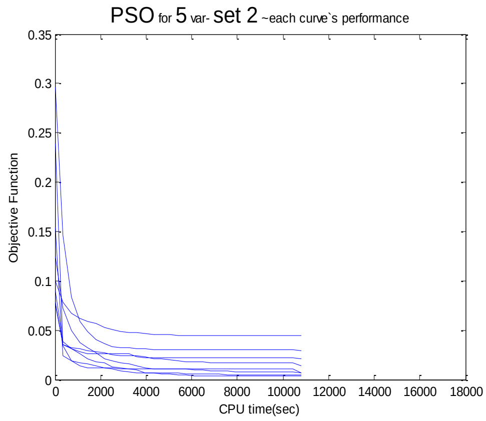


Και ο PSO παρουσιάζει καλύτερη σύγκλιση για 5 παραμέτρους.



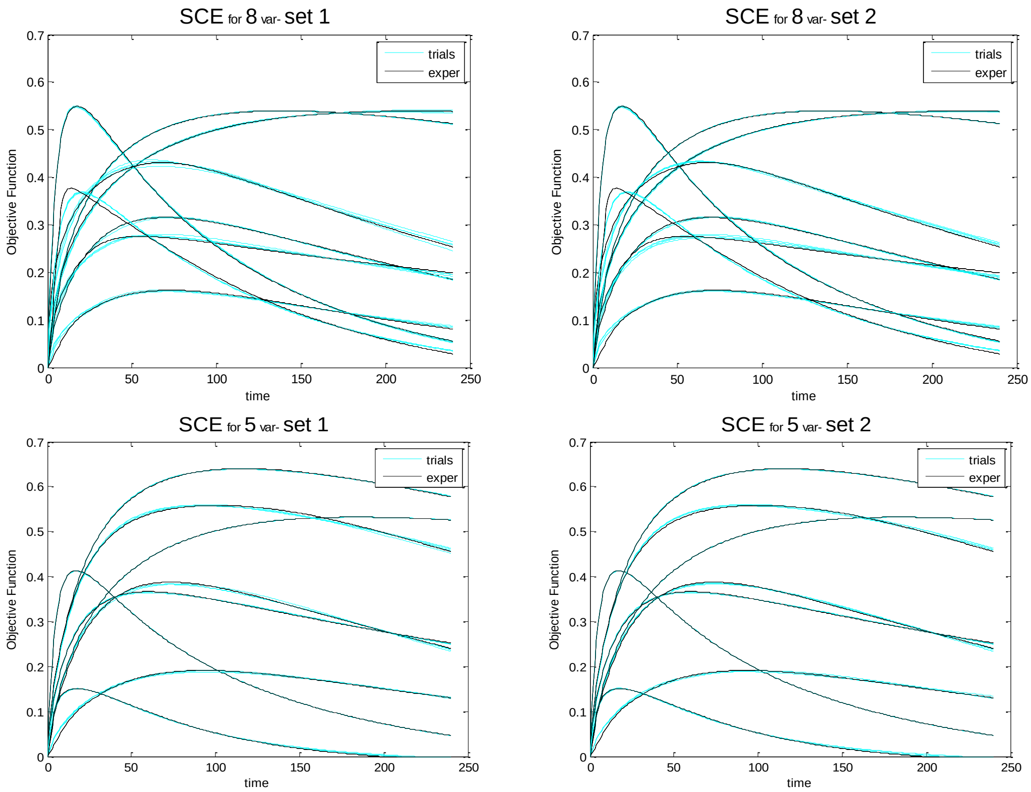




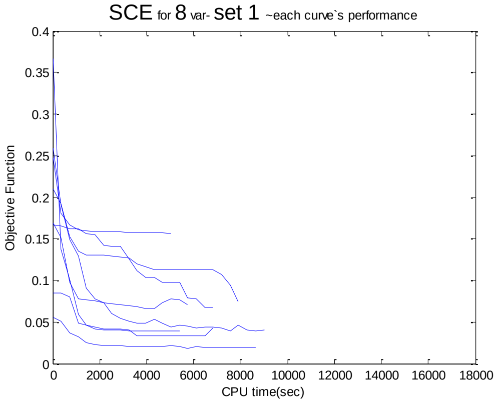


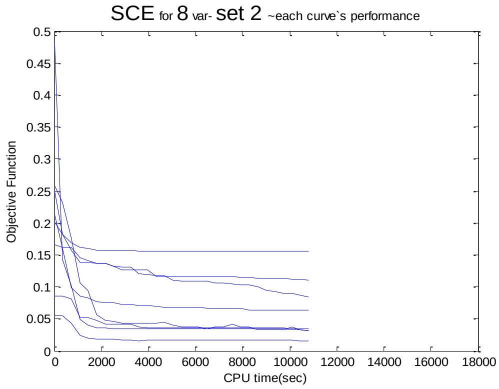
Και σε αυτή την περίπτωση, του PSO διαπιστώνουμε γρηγορότερη και βέλτιστη σύγκλιση για 5 παραμέτρους και για το $2^{ο}$ σετ περιορισμών.

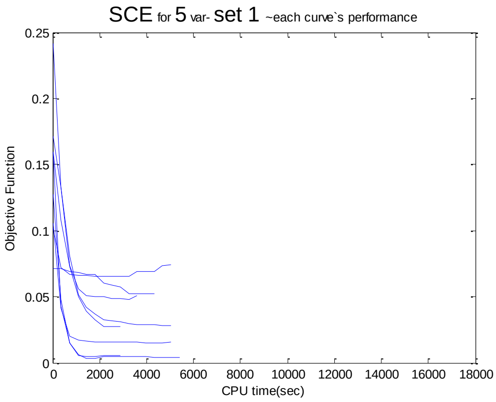
Αποτελέσματα για SCE

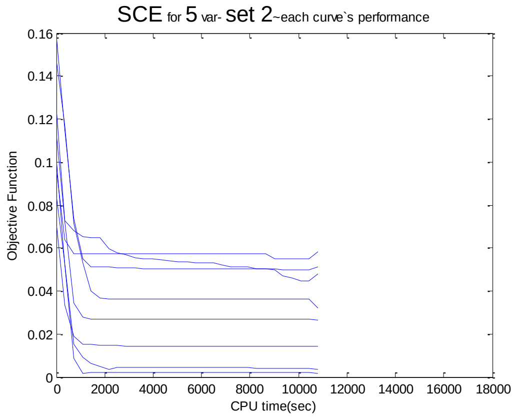


Και τέλος για τον SCE επιβεβαιώνεται για ακόμα μία φορά η βέλτιστη σύγκλιση για της 5 παραμέτρους.









4.7 NRMSD

Συγκρίναμε τα αποτελέσματά με βάση το μέσο Normalized Root Mean Square Deviation που δίνει κάθε αλγόριθμος για την κάθε παράμετρο.

Το RMSD είναι μία ποσότητα που χρησιμοποιείται συχνά και αντιπροσωπεύει της διαφορές μεταξύ των τιμών που προβλέπονται από ένα μοντέλο και των τιμών που πραγματικά παρατηρούνται από το πρόβλημα που μοντελοποιείται. Στη δική της περίπτρωση, έχουμ της διαφορές μεταξύ των τιμών των παραμέτρων που περιμένουμε να πάρουμε και των τιμών που πραγματικά πήραμε της 32 φορές που τρέξαμε τον κάθε αλγόριθμο για το κάθε σετ περιορισμών.

NRMSD είναι η ποσότητα RMSD της το εύρος των τιμών, δηλαδή η διαφορά του ανώτερου και κατώτερου ορίου της εκάστοτε παραμέτρου. Η τιμή αυτή εκφράζεται σαν ποσοστό.

Συγκεντρωτικά αποτελέσματα.

Της παρακάτω πίνακες βλέπουμε συγκεντρωτικά την τιμή της NRMSD για 8 και 5 παραμέτρους για το $1^{ο}$ και $2^{ο}$ αντίστοιχα σετ περιορισμών.

Comparison Of Algorithms- Set 1

NRMSD	a%	b%	β_{ES}%	β_{IS}%	n%	pHe %	pHi %	g_{TJ}%
CRS	24.5	22	22.4	23.7	27.5	21.2	27.1	15.5
DE	31.2	24.3	25.4	34.3	29.6	20	28.4	22.5
DIRECT	28.8	22	31	25.9	21.4	10.9	17.8	19.9
PSO	21.4	18.1	31.4	27.3	15.12	22.8	26.7	8.2
SCE	21.3	17.3	26.1	32.1	18.8	9.8	22	14.2

Πίνακας 3.

NRMSD	a%	b%	n%	pHe%	g_{TJ}%
CRS	20.4	13.4	24.8	16.4	8.6
DE	18.3	11.5	16.9	12.6	6.2
DIRECT	36.6	16.4	14.2	13.9	8.6
PSO	27.6	17.7	20.2	14.7	7.8
SCE	24.9	16.3	19.6	16.9	4.4

Πίνακας 4.

Comparison Of Algorithms- Set 2

NRMSD	a%	b%	β_{ES}%	β_{IS}%	n%	pHe %	pHi %	g_{TJ}%
CRS	24.5	22	22.4	23.7	27.5	21.2	27.1	15.5
DE	25.9	23.1	18.8	34.9	21.08	19.6	19	17.3
DIRECT	28.8	22	31	25.9	21.4	10.9	17.8	19.9
PSO	21.9	20	32	29	17.2	23	25.2	8
SCE	19	12.1	25.9	26.2	20.8	13.2	23.3	14.2

Πίνακας 5.

NRMSD	a%	b%	n%	pHe%	g_{TJ}%
CRS	20.4	13.4	24.8	16.4	8.6
DE	9.8	3.6	18.2	14.5	3.6
DIRECT	36.6	16.4	14.2	13.9	8.6
PSO	29.3	15.5	16.6	14.6	5.7
SCE	19.9	13.4	21.4	16.4	4.7

Πίνακας 6.

Ενώ το σύνολο των παραμέτρων για το κάθε σετ φαίνεται συγκεντρωτικά παρακάτω.

Comparison Of Algorithms

NRMSD	Set 1	Set 2	Set 1	Set 2
	8_7200	8_3	5_7200	5_3
CRS	23.8%	23.8%	16.7%	16.7%
DE	26.9%	22.5%	13.1%	9.9%
DIRECT	22.2%	22.2%	18%	18%
PSO	21.4%	22.2%	17.6%	16.3%
SCE	20.2%	19.4%	16.4%	15.1%

Πίνακας 7.

Διαπιστώνουμε ότι η καλύτερη απόδοση αλγόριθμου στο σύνολο είναι αυτή του DE για 5 παραμέτρους και για 3 ώρες.

ΚΕΦΑΛΑΙΟ 6

ΣΥΜΠΕΡΑΣΜΑΤΑ-ΒΙΒΛΙΟΓΡΑΦΙΑ

[31-45]

Τα τελευταία χρόνια έχουν πραγματοποιηθεί αρκετές έρευνες πάνω στους αλγόριθμους βελτιστοποίησης πάνω σε μεγάλο αριθμό προβλημάτων. Εμείς στην προσπάθειά μας να υπολογίσουμε τις παραμέτρους του μη γραμμικού μας συστήματος καταλήξαμε να χρησιμοποιήσουμε ορισμένους από τους πλέον διαδεδομένους αλγόριθμους βελτιστοποίησης και κατόπιν επιδιώξαμε να επιβεβαιώσουμε τα αποτελέσματά μας, μέσα από τη βιβλιογραφία.

Στην έρευνα μας στους CRS, DE, DIRECT, PSO και SCE αλγόριθμους βελτιστοποίησης, ο πλέον αποτελεσματικός ήταν ο DE. Ακολουθούν οι υπόλοιποι στοχαστικοί με μικρή διαφορά απόδοσης μεταξύ τους. Ενώ ο μοναδικός ντετερμινιστικός που χρησιμοποιήσαμε, ο DIRECT, ήταν απογοητευτικός.

Στις διάφορες έρευνες μέσα από τη βιβλιογραφία, κριτήρια για την απόδοση των αλγόριθμων αποτελούν, το κατά πόσο πλησιάζει ο αλγόριθμος το ολικό ελάχιστο και η ταχύτητα σύγκλισης του αλγόριθμου(άλλοτε με κριτήριο το χρόνο και άλλοτε με κριτήριο τον αριθμό αντικειμενικών συναρτήσεων που υπολογίζονται). Τέλος καθοριστικός παράγοντας καταλληλότητας ενός αλγόριθμου είναι ο όσο το δυνατόν μικρότερος αριθμός παραμέτρων του, ώστε να μπορεί εύκολα να προσαρμοστεί στις συνθήκες του εκάστοτε προβλήματος.

Ακόμη το πρόβλημα εντοπισμού του ολικού βέλτιστου σε ένα χώρο με πολλά τοπικά βέλτιστα είναι ένα κλασικό πρόβλημα για όλα τα συστήματα που μπορούν να προσαρμοστούν και να μάθουν, όπως δηλ. λειτουργούν αυτού του τύπου οι αλγόριθμοι. Αυτό αποτελεί και ένα από τα βασικά προβλήματα που συνάντησαν οι αλγόριθμοι και στο δικό μας πρόβλημα, το να παγιδευτούν δηλ. σε ένα τοπικό βέλτιστο(ελάχιστο).

Επίσης γεγονός αποτελεί ότι όσο λιγότεροι οι παράμετροι/διαστάσεις του προβλήματος τόσο καλύτερα και γρηγορότερα τα αποτελέσματα του εκάστοτε αλγόριθμου, γεγονός που συμφωνεί και με τα δικά μας πορίσματα.

Οι στοχαστικές μέθοδοι στηρίζονται εξ' ορισμού σε probabilistic approaches. Δεδομένου λοιπόν ότι στην διαδικασία εμπλέκεται το στοιχείο της τυχαιότητας, αυτές οι μέθοδοι δεν μπορούν να εγγυηθούν αξιοπιστία στη σύγκλιση στο ολικό βέλτιστο. Τις περισσότερες φορές παρουσιάζουν χαμηλούς ρυθμούς σύγκλισης ειδικότερα στα τελευταία στάδια της έρευνας. Αυτό μπορεί να έχει σαν αποτέλεσμα εκτενείς υπολογιστικούς χρόνους ειδικότερα αν απαιτείται μεγάλη ακρίβεια.

Όλοι οι στοχαστικοί αλγόριθμοι που χρησιμοποιήσαμε βασίζονται στους πληθυσμούς. Αυτό σημαίνει ότι ξεκινούν με ένα σύνολο N σημείων που απαρτίζουν ένα πληθυσμό S (και υπολογίζουν την τιμή αντικειμενικής συνάρτησης για το καθένα απ' αυτά). Στη συνέχεια με διάφορες στρατηγικές, οι διάφοροι αλγόριθμοι εξελίσσουν αυτά τα σημεία, από γενιά σε γενιά, μέχρι κάποιο κριτήριο τερματισμού να ικανοποιείται και τελικά να εντοπιστεί το ολικό ελάχιστο.

Ο CRS αλγόριθμος αντικαθιστά το χειρότερο σημείο του S (αυτό με τη μικρότερη τιμή αντικειμενικής συνάρτησης) σε κάθε γενιά. Αντίθετα ο DE σε κάθε γενιά προσπαθεί να αντικαταστήσει όλα τα σημεία στον S, μέσω των mutation και crossover διαδικασιών όπου νέα δοκιμαστικά σημεία παράγονται και αντικαθιστούν αυτά τις προηγούμενης γενιάς (εάν η τιμή της αντικειμενικής τους συνάρτησης είναι μικρότερη, άρα και καλύτερη). Καθ' αυτόν τον τρόπο ο DE έχει τη δυνατότητα να διερευνά καλύτερα το δειγματικό χώρο, γεγονός που τον καθιστά πιο αποτελεσματικό.

Σε άλλες μελέτες ο DE αποδεικνύεται ιδιαίτερα αποτελεσματικός για προβλήματα ρεαλιστικής φύσης, όπου ο υπολογισμός της αντικειμενικής συνάρτησης συχνά απαιτεί αρκετό χρόνο. Ειδικότερα, όταν εξετάστηκε μαζί με τον PSO σε ένα μεγάλο αριθμό χαρακτηριστικών επιστημονικών προβλημάτων ο DE ξεπέρασε σε επίδοση, ακρίβεια και ταχύτητα, τον PSO σχεδόν σε όλα. Ο PSO χαρακτηρίζεται σαν ένας από τους πλέον υποσχόμενους αλγόριθμους βελτιστοποίησης και ακολουθεί από κοντά σε επίδοση τον DE, αφού και οι δύο είναι σχεδιασμένοι για μη γραμμικά, πολυδιάστατα προβλήματα και δεν εγκλωβίζονται σε τοπικά ελάχιστα. Όμως αυτό που τον διαφοροποιεί από τον DE τελικά, είναι ότι δεν χρησιμοποιεί διαδικασίες mutation και crossover, καθώς επίσης και το γεγονός ότι ο πληθυσμός κινείται σαν ένα σώμα μέσα στο δειγματικό χώρο και δεν εξελίσσεται από γενιά σε γενιά.

Τέλος αυτό που διαφοροποιεί τον DE από τους υπόλοιπους στοχαστικούς, βασισμένους σε πληθυσμούς αλγόριθμους είναι ότι είναι ιδιαίτερα απλός, έχει μόνο δύο παραμέτρους ελέγχου. Έτσι είναι πιο εύκολο να προσαρμοστεί/ρυθμιστεί στις συνθήκες και ανάγκες του εκάστοτε προβλήματος.

Ο Direct ο μοναδικός ντετερμινιστικός αλγόριθμος που εξετάσαμε δεν έφερε ικανοποιητικά αποτελέσματα. Οι ντετερμινιστικοί μέθοδοι είναι εκείνες που εξασφαλίζουν συνήθως ένα βαθμό βεβαιότητας ως προς τη σύγκλιση στο ολικό βέλτιστο για δεδομένα προβλήματα βελτιστοποίησης, παρ' όλα αυτά πρέπει να σημειωθεί ότι κανένας αλγόριθμος δεν μπορεί να επιλύσει τέτοιου είδους προβλήματα σε καθορισμένα πλαίσια χρόνου. Ο Direct εντοπίζει γρήγορα τις περιοχές με μεγάλη πιθανότητα εύρεσης του ολικού βέλτιστου. Όμως η βασισμένη σε πλέγμα επαναληπτική μέθοδος που χρησιμοποιεί οδηγεί σε αργή σύγκλιση όταν ο αλγόριθμος βρεθεί κοντά στο βέλτιστο, ειδικότερα όταν πρόκειται για μη γραμμικά ή με περιορισμούς παραμέτρους προβλήματα (η δομή του πλέγματος δυσκολεύει την έρευνα αν το ολικό βέλτιστο βρίσκεται πάνω στο πλέγμα/ορθογώνιο). Γενικότερα ενώ μέσα από μελέτες έχει αποδειχθεί η αποτελεσματικότητά του για γραμμικά προβλήματα, δεν έχει μελετηθεί εκτενώς ως προς μη-γραμμικά, περιορισμένα προβλήματα, όπως το δικό μας. Ο Jones άλλωστε προτείνει τη χρήση και αλγόριθμου τοπικής βελτιστοποίησης (ίσως στοχαστικής φύσης ώστε να εισαχθεί στη διαδικασία και παράγοντας τυχαιότητας) μαζί με τον Direct,δηλ. υβριδισμό, για καλύτερα και ταχύτερα αποτελέσματα.

Όπως προκύπτει λοιπόν από τη βιβλιογραφία και επιβεβαιώνεται στην περίπτωσή μας, οι στοχαστικές μέθοδοι παρέχουν μία καλή λύση στο πρόβλημα και ακόμα και αν δεν είναι η ιδανική, μπορεί να είναι η πλέον ικανοποιητική και μέσα σε ανεκτά όρια χρόνου.

ΣΥΜΠΕΡΑΣΜΑΤΑ-ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

Τα αποτελέσματα της έρευνάς μας επιβεβαιώνουν την a priori έρευνα που ισχυρίζεται ότι οι τρεις παράμετροι (pH_{IS} , β_{ES} , β_{IS}) λειτουργούν ως θόρυβος στο μοντέλο μας, με αποτέλεσμα όταν τις απομονώσουμε να έχουμε πολύ καλύτερα αποτελέσματα και σε σύγκλιση καμπύλης και στον υπολογισμό των παραμέτρων. Ακόμη η παράμετρος a παρουσιάζει σε κάθε περίπτωση τα χειρότερα αποτελέσματα και αυτό συμφωνεί με το γεγονός ότι η a priori έρευνα την έχει θέσει υπό αμφισβήτηση.

Επίσης οι αλγόριθμοι διαπιστώνουμε ότι για πέντε παραμέτρους γίνονται όλοι γρηγορότεροι και τερματίζουν λόγω του αριθμού των function evaluations, σαφώς νωρίτερα. Παρά το γεγονός ότι τα function evaluations ολοκληρώνονται σε λιγότερο από το μισό χρόνο (1,5 h.) οι αλγόριθμοι εμφανίζουν μικρή βελτίωση στην απόδοσή τους στο διπλάσιο χρόνο, με εξαίρεση τον DE για 5 παραμέτρους.

Ο DE παρουσιάζει τα καλύτερα αποτελέσματα με κριτήριο το μέσο NRMSD για πέντε παραμέτρους και ξεφεύγει από την απόδοση των υπόλοιπων αλγόριθμων. Μάλιστα το 9,9% (κάτω από 10%) στις 3 hours για 32 επαναλήψεις, είναι ικανοποιητικό ως προς την ακρίβεια και την επαναληψιμότητα.

Ακολουθεί ο αλγόριθμος SCE, PSO και CRS με μικρή διαφορά στη μεταξύ τους απόδοση.

Ο DIR ο μοναδικός ντετερμινιστικός αλγόριθμος παρουσιάζει και τα χειρότερα με απόσταση αποτελέσματα.

Η παράμετρος g_{TJ} εμφανίζει τα καλύτερα ποσοστά. Γεγονός σημαντικό γιατί ως λειτουργική παράμετρος, δεν έχουμε άλλη μέθοδο μέτρησής της.

Σαν μελλοντική εργασία θα μπορούσαμε να επιδιώξουμε να ασχοληθούμε με περισσότερους στοχαστικούς αλγόριθμους τους οποίους θα συνδυάζαμε με κάποιο ντετερμινιστικό (υβριδικός) αφού σύμφωνα με τη βιβλιογραφία είναι το πλέον υποσχόμενο σενάριο στο χώρο. Θα πρέπει επίσης να ληφθεί σοβαρά υπ' όψιν η εξαίρεση και της μεταβλητής a έπειτα από αυτά τα αποτελέσματα. Θα μπορούσαμε επίσης να τρέξουμε το μοντέλο χρησιμοποιώντας κλινικά δεδομένα για τη δημιουργία χάρτη των βιολογικών παραμέτρων σε πραγματικό χρόνο. Ακόμη θα μπορούσαμε να προσθέσουμε συνθήκες για να ελέγχουμε τη σύγκλιση των καμπυλών σε συγκεκριμένα/κρίσιμα σημεία της καμπύλης (βάρη). Θα μπορούσε ακόμη να μελετηθεί κατά πόσο αυξάνεται η επίδραση μίας παραμέτρου σε σχέση με τις υπόλοιπες σε αυτά τα κρίσιμα σημεία της καμπύλης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Shuffled Complex Evolution Approach for effective and efficient global minimization, Q.Y. Duan, V.K.Gupta and S. Sorooshian, 1993.

[2] Program OPT3 Shuffled Complex Evolution Concept Represantation, Search of Two Parameter Space, Duan, 1991-1993.

[3] Automatic Calibration Of The Pitman Model Using The Shuefled Complex Evolution Method, J G Ndiritu.

[4] Optimal Design Of Water Distribution Network Using Shuffled Complex Evolution, Shie-Yui Liong and Md. Atiquzzaman

[5]Optimal Use Of The SCE-UA Global Optimization Method For Calibrating Watershed Models,Duan Sorooshian Gupta,1994

[6] Ιωάννης Γ. Τσούλος (2008), Γεννετικοί Αλγόριθμοι, Ιωάννινα

[7](Eberhart, R.C., & Kennedy, J.: A new optimizer using particle swarm theory. In Proceedings of the sixth international symposium on micro machine and human science, Nayoya, Japan, Piscataway:IEEE (1995), 39 – 45.)

[8](Recent approaches to global optimization problems
through Particle Swarm Optimization, K.E. PARSOPOULOS1 and M.N. VRAHATIS2)

[9](file:///C:/Users/User/Desktop/PSO/tutorials.php.htm)

[10](http://www.hydrol-earth-syst-sci-discuss.net/8/9707/2011/hessd-8-9707-2011.pdf)

[11]icon(http://www.slideshare.net/Soheyl/effectiveness-and-efficiency-of-particle-swarm-optimization-technique-in-inverse-heat-conduction-analysis)

[12]Averaging Estimators for Kernel Regressions,Chu-An Liu

[13]Lipschitzian Optimization Without the Lipschitz Constant, D.R. Jones, C.D. Perttunen, B.E. Stuckman

[14](icon) Optimization without algebraic models: algorithms, software and applications, Nick Sahinidis

[15](icon+)CS6100: Topics in Design and analysis of algorithms, convex hull, John Augustine

[16](icon+)Global Otimization Using the Direct Algorithm in Matlab, M. Bjorkman and K. Holmstrom

[17]R. Storn, and K. Price (1996), Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces

[18](http://mathworld.wolfram.com/DifferentialEvolution.html)

[19](http://invenio.lib.auth.gr/record/114654/files/pappas.pdf)

[20](Adaptive Differential Evolution and Exponential Crossover,Josef Tvrdik)

[21](http://www.cs.sandia.gov/opt/survey/intro.html)

[22](http://en.wikipedia.org/wiki/Global_optimization)

[23]Derivative-free optimization: A reviw of algorithms and comparison of software implementations, Rios and Sahinidis

[24]Wikipedia.org/wiki/Simulated_annealing

[25]Swagatam, Differential Evolution: A survey of the State of Art

[26]Parsopoulos and Vrahatis, Recent approaches to global optimization problems through Particle Swarm Optimization

[27]Rios and Sahinidis, Derivative-free optimization:A review of algorithms and comparison of software implementations

[28] Costas Balas, George Papoutsoglou, and Antonis Potirakis (2008), In Vivo Molecular Imaging of Cervical Neoplasia Using Acetic Acid as Biomarker, Technical Educational Institute of Crete, Chania

[29] Costas Balas (2001), A Novel Optical Imaging Method for the Early Detection, Quantitative Grading, and Mapping of Cancerous and Precancerous Lesions of Cervix

[30] Costas Balas, George Papoutsoglou, Estimation of Neoplasia-Related Biological Parameters Through Modeling and Sensitivity Analysis of Optical Molecular Imaging Data, Technical Educational Institute of Crete, Chania

[31]A hybrid approach for efficient and robust parameter estimation in biochemical pathways. Maria Rodriguez-Fernandez a, Pedro Mendesb, Julio R. Banga

[32]Genetic and other global optimization algorithms - comparison and use in calibration problems

[33]Genetic Algorithm – an Approach to Solve Global Optimization Problems .PRATIBHA BAJPAI

[34]Experiences in using evolutionary and non-evolutionary optimization methods in models calibration. Dimitri P. Solomatine

[35]Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces by Rainer Storn

[36]Population Set Based Global Optimization Algorithms: Some Modi_cations and Numerical Studies_ M. M. Ali

[37]A Comparative Study of Differential Evolution, Particle Swarm Optimization, and Evolutionary Algorithms on Numerical Benchmark Problems

[38]Lecture 16 Differential Evolution And Particle Swarm Optimisation

[39]Evaluation of Differential Evolution and Particle Swarm Optimization Algorithms at Training of Neural Network for Stock Prediction

[40]A hybrid approach for efficient and robust parameter estimation in biochemical pathways

[41]Maria Rodriguez-Fernandez a, Pedro Mendesb, Julio R. Banga

[42] The Deterministic Global Optimization Algorithm DIRECT, Layne T. Watson and Jian He

[43] Direct using local search on surrogates, Thomas Hemker, Christian Werner

[44] Advanced Global Optimisation Tools for Mission Analysis and Design, Pierluigi Di Lizia, Gianmarco Radice

[45]Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods