



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Χημικών Μηχανικών

Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών

Υπολογιστική Μηχανική

Μεταπτυχιακή εργασία

Καβάγια Χρυσούλας

Διπλωματούχου Μηχανολόγου Μηχανικού Ε.Μ.Π.

**Προσομοίωση Καταστατικών Νόμων Υλικών με
Νευρωνικά Δίκτυα**

Επιβλέποντες: *Ε. Παπαδρακάκης, Καθηγητής Ε.Μ.Π.*

Ν. Λαγαρός, Λέκτορας Ε.Μ.Π.

Αθήνα 2010

Πρόλογος

Ένα από τα θέματα που απασχολεί συνεχώς τους επιστήμονες – μηχανικούς είναι η μελέτη της συμπεριφοράς των υλικών σε διάφορες καταπονήσεις και κυρίως νέων υλικών. Όπως είναι ήδη γνωστό η μηχανική συμπεριφορά των υλικών εκφράζεται ποσοτικά με τους καταστατικούς νόμους, οι οποίοι καθορίζουν τις σχέσεις τάσεων - παραμορφώσεων, δηλαδή τα μεγέθη των τάσεων που αναπτύσσονται κατά την παραμόρφωση του υλικού. Οι καταστατικοί νόμοι είναι απαραίτητοι κατά την αναλυτική ή αριθμητική επίλυση προβλημάτων που αφορούν φορτίσεις των υλικών, όπως π.χ. τον υπολογισμό των μετακινήσεων μίας δοκού όταν αυτή υποβληθεί σε εξωτερικά φορτία. Επίσης οι εξισώσεις αυτές, έχουν διάφορα χαρακτηριστικά όπως εξάρτηση από τις παραμορφώσεις (είναι η πιο απλή μορφή), από το ρυθμό των παραμορφώσεων, από την ιστορία των φορτίσεων και από τις φυσικές ιδιότητες του υλικού. Τέλος οι καταστατικές εξισώσεις δεν είναι φυσικοί νόμοι, δεν έχουν δηλαδή τη γενική εφαρμογή των φυσικών νόμων, αλλά είναι συσχετίσεις που βασίζονται είτε στα θερμοδυναμικά αξιώματα (με κάποιες πρόσθετες παραδοχές), όπως π.χ. ο νόμος της γραμμικής ελαστικότητας, είτε σε φαινομενολογικές θεωρήσεις που προκύπτουν από πειραματικές κυρίως μετρήσεις της συμπεριφοράς των υλικών.

Όπως αναφέρθηκε και προηγουμένως, οι καταστατικοί νόμοι είναι απαραίτητοι για την ανάλυση προβλημάτων που αφορούν στις φορτίσεις των υλικών ή στην καταπόνηση μιας κατασκευής γενικά. Τα μοντέλα καταστατικών νόμων διαφόρων υλικών που χρησιμοποιούνται έως τώρα, προκύπτουν από συμβατικές μεθόδους μοντελοποίησης, οι οποίες χρησιμοποιούν μαθηματικές εκφράσεις για να προσεγγίσουν την πειραματικά παρατηρούμενη συμπεριφορά των υλικών. Αυτά τα μοντέλα παρουσιάζουν αρκετά προβλήματα στη χρήση τους, είτε γιατί η συμπεριφορά κάποιου υλικού, υποβαλλόμενο σε ένα συγκεκριμένο είδος φόρτισης, είναι πολύπλοκη για να μοντελοποιηθεί (π.χ. το έδαφος ή η ανακυκλιζόμενη συμπεριφορά των υλικών), είτε γιατί η χρήση τους σε διάφορα προβλήματα ανάλυσης κατασκευών, καθιστούν πολύ χρονοβόρα την επίλυση των προβλημάτων αυτών.

Η δυσκολία που υπήρχε λοιπόν, με τη χρήση των συμβατικών καταστατικών μοντέλων, οδήγησε τους επιστήμονες στη χρήση ενός υπολογιστικού εργαλείου για την μοντελοποίηση της συμπεριφοράς των υλικών, τα τεχνητά νευρωνικά δίκτυα (ANN: artificial neural networks), τα οποία είναι εμπνευσμένα από τη βιολογία και αποτελούνται από στοιχεία που συμπεριφέρονται κατά τρόπο ανάλογο με αυτόν των βιολογικών κυττάρων. Πιο συγκεκριμένα τα νευρωνικά δίκτυα είναι παράλληλα υπολογιστικά μοντέλα που χρησιμεύουν για την αναπαράσταση γνώσης και

διαχείρισης πληροφοριών. Η δομή τους αποτελείται από ένα επίπεδο εισόδου (είσοδος της πληροφορίας), από ένα ή περισσότερα εσωτερικά επίπεδα (κρυμμένα επίπεδα) και από ένα επίπεδο εξόδου, το οποίο παρέχει την επιθυμητή για το δίκτυο απόκριση. Τα επίπεδα μεταξύ τους συνδέονται μέσω κάποιων συγκεκριμένων παραμέτρων, τα βάρη σύνδεσης. Ο λόγος που χρησιμοποιήθηκαν στην μοντελοποίηση καταστατικών νόμων είναι οι εξαιρετικές ικανότητες εκμάθησης που έχουν, οι οποίες μπορούν να χρησιμοποιηθούν στην εκμάθηση περίπλοκων μη γραμμικών εξισώσεων. Τα νευρωνικά δίκτυα λόγω των ιδιαίτερων χαρακτηριστικών τους, μπορούν να εκπαιδευτούν άμεσα, με αποτελέσματα από πειράματα, που γίνονται σε διάφορα υλικά και να αποθηκεύσουν αυτή την πληροφορία στα βάρη σύνδεσης. Με αυτόν τον τρόπο αναπαριστούν διάφορους νόμους υλικών και έτσι έχουν τη δυνατότητα να συνεργάζονται με υπολογιστικά πακέτα (ANSYS, ABAQUS) για την επίλυση διαφόρων προβλημάτων μηχανικής, μειώνοντας κατά πολύ το υπολογιστικό κόστος.

Η παρούσα μεταπτυχιακή εργασία λοιπόν, ασχολείται με τη μοντελοποίηση καταστατικών νόμων διαφόρων υλικών μέσω τεχνητών νευρωνικών δικτύων. Τα υλικά που επιλέχθηκαν είναι χάλυβας, σκυρόδεμα και έδαφος. Το νευρωνικό δίκτυο εκπαιδεύτηκε με πειραματικά δεδομένα, που προέκυψαν από τη βιβλιογραφία, για δύο είδη φόρτισης, την μονοτονική μονοαξονική φόρτιση και την ανακυκλιζόμενη μονοαξονική φόρτιση. Τα αποτελέσματα για τα τρία υλικά και για τα δύο είδη φορτίσεων είναι πολύ ικανοποιητικά, γεγονός που παραπέμπει σε περαιτέρω έρευνα ακόμα και για τριαξονικές φορτίσεις.

Ευχαριστίες

Η παρούσα διπλωματική εργασία πραγματοποιήθηκε στα πλαίσια του ΔΠΜΣ «Υπολογιστική Μηχανική» της Σχολής Χημικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου κάτω από την επίβλεψη του κ. Ε. Παπαδρακάκη, Καθηγητή της Σχολής Πολιτικών Μηχανικών του Ε.Μ.Π., τον οποίο θα ήθελα να ευχαριστήσω θερμά για την εμπιστοσύνη που μου έδειξε με την ανάθεση της μεταπτυχιακής εργασίας και την καθοδήγησή του στην εκπόνηση της εργασίας. Επίσης θα ήθελα να ευχαριστήσω τον κ. Ν. Λαγαρό, Λέκτορα της Σχολής Πολιτικών Μηχανικών του Ε.Μ.Π. για την βοήθεια που μου παρείχε όλον αυτό τον καιρό.

Καβάγια Χρυσούλα

Πίνακας Περιεχομένων

ΠΡΟΛΟΓΟΣ	I
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	V
ΚΕΦΑΛΑΙΟ 1^ο: ΚΑΤΑΣΤΑΤΙΚΟΙ ΝΟΜΟΙ	11
1.1 ΕΙΣΑΓΩΓΗ	13
1.2 ΚΑΤΑΣΤΑΤΙΚΟΙ ΝΟΜΟΙ ΧΑΛΥΒΑ	15
1.2.1 ΔΙΓΡΑΜΜΙΚΟ ΜΟΝΤΕΛΟ ΧΑΛΥΒΑ	16
1.2.2 ΥΛΙΚΟ ΜΟΝΤΕΛΟ ΜΕΝΕΓΟΤΤΟ - ΡΙΝΤΟ	19
1.2.3 ΤΡΙΣΔΙΑΣΤΑΤΟ ΚΑΤΑΣΤΑΤΙΚΟ ΜΟΝΤΕΛΟ.....	21
1.3 ΚΑΤΑΣΤΑΤΙΚΟΙ ΝΟΜΟΙ ΣΚΥΡΟΔΕΜΑΤΟΣ	22
1.3.1 ΥΛΙΚΟ ΜΟΝΤΕΛΟ ΤΩΝ ΚΕΝΤ ΚΑΙ ΡΑΚ	22
1.4 ΚΑΤΑΣΤΑΤΙΚΟΙ ΝΟΜΟΙ ΕΔΑΦΟΥΣ	27
1.4.1 ΓΡΑΜΜΙΚΗ ΙΣΟΤΡΟΠΗ ΕΛΑΣΤΙΚΟΤΗΤΑ.....	28
1.4.1.1 Η έκφραση των σχέσεων ελαστικότητας ως προς τις ολικές τάσεις	31
1.4.2 ΜΗ ΓΡΑΜΜΙΚΕΣ ΣΧΕΣΕΙΣ ΤΑΣΕΩΝ - ΠΑΡΑΜΟΡΦΩΣΕΩΝ	33
1.4.3 Η ΕΝΝΟΙΑ ΤΗΣ ΔΙΑΔΡΟΜΗΣ ΤΩΝ ΤΑΣΕΩΝ.....	36
1.4.4 ΜΟΝΟΔΙΑΣΤΑΤΗ ΣΥΜΠΙΕΣΗ ΕΔΑΦΩΝ.....	39
1.4.5 ΑΝΑΣΚΟΠΗΣΗ ΤΩΝ ΒΑΣΙΚΟΤΕΡΩΝ ΚΑΤΑΣΤΑΤΙΚΩΝ ΜΟΝΤΕΛΩΝ ΤΟΥ ΕΔΑΦΟΥΣ	47
1.4.5.1 Μοντέλο Mohr-Coulomb.....	47
1.4.5.2 Τροποποιημένο μοντέλο Cam clay.....	49
1.4.5.3 Μοντέλο Duncan – Chang (Υπερβολικό μοντέλο).....	50
1.4.5.4 Μοντέλο εδάφους με κράτυνση	52
ΚΕΦΑΛΑΙΟ 2^ο: ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	55

2.1 ΕΙΣΑΓΩΓΗ	57
2.2 ΟΙ ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΩΝ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ (ANN)	58
2.2.1 Το βιολογικό πρότυπο	59
2.2.2 Ο τεχνητός νευρώνας	60
2.2.3 Συναρτήσεις μεταφοράς	61
2.2.4 Δίκτυα προσθίας τροφοδοτήσης	63
2.2.4.1 Τεχνητό νευρωνικό δίκτυο ενός επιπέδου	64
2.2.4.2 Τεχνητό νευρωνικό δίκτυο πολλών επιπέδων.....	65
2.2.4.3 Μη γραμμικές συναρτήσεις μεταφοράς.....	66
2.2.5 Δίκτυα με επιστροφή (RECURRENT NETWORKS).....	67
2.2.6 Η εκπαίδευση των τεχνητών νευρωνικών δικτύων.....	67
2.2.6.1 Αντικείμενο της εκπαίδευσης	67
2.2.6.2 Εκπαίδευση με επίβλεψη (Supervised training)	67
2.2.6.3 Εκπαίδευση χωρίς επίβλεψη (Unsupervised training).....	68
2.2.6.4 Αλγόριθμοι εκπαίδευσης	68
2.3 Ο ΑΛΓΟΡΙΘΜΟΣ ΕΚΠΑΙΔΕΥΣΗΣ ΠΙΣΩ ΔΙΑΔΟΣΗΣ.....	69
2.3.1 Αριθμητικές μέθοδοι βελτιστού σχεδιασμού	69
2.3.1.1 Η μέθοδος της Απότομης Καθόδου (Steepest Descent method)	70
2.3.1.2 Η μέθοδος της Συζυγούς Διανυσματικής Κλίσεως (Conjugate Gradient method) .	71
2.3.1.3 Η μέθοδος Newton (Newton's method).....	72
2.3.1.4 Η μέθοδος των Marquardt-Levenberg	73
2.3.1.5 Τοπικό ελάχιστο	73
2.3.1.6 Μέγεθος βήματος	74

2.3.2 Ο ΑΛΓΟΡΙΘΜΟΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ RPROP.....	75
2.3.3 ΠΙΣΩ ΔΙΑΔΟΣΗ.....	78
2.3.3.1 Εκπαίδευση.....	79
2.3.3.2 Εμπρός πέρασμα.....	81
2.3.3.3 Πίσω πέρασμα.....	81
2.3.4 ΈΛΕΓΧΟΣ ΠΟΡΕΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ.....	85
2.3.5 Η ΠΑΡΑΜΕΤΡΟΣ BIAS.....	86
2.3.6 Ο ΟΡΟΣ MOMENTUM.....	86
2.3.7 ΠΑΡΑΛΥΣΗ ΤΟΥ ΔΙΚΤΥΟΥ.....	87
2.3.8 ΥΠΕΡΒΟΛΙΚΗ ΕΚΠΑΙΔΕΥΣΗ ΔΙΚΤΥΟΥ.....	88
2.3.9 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΔΙΚΤΥΟΥ.....	90
2.3.10 ΕΠΙΛΟΓΗ ΤΟΥ ΣΥΝΟΛΟΥ ΕΚΠΑΙΔΕΥΣΗΣ.....	91
2.4 ΑΛΛΕΣ ΤΕΧΝΙΚΕΣ ΕΚΠΑΙΔΕΥΣΗΣ	91
2.4.1 ΣΤΟΧΑΣΤΙΚΕΣ ΜΕΘΟΔΟΙ.....	91
2.4.1.1 Εκπαίδευση τύπου Boltzmann.....	94
2.4.1.2 Εκπαίδευση Cauchy.....	96
2.4.1.3 Ο αλγόριθμος της πίσω διάδοσης και εκπαίδευση Cauchy.....	97
2.4.2 ΓΕΝΕΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗ.....	98
ΚΕΦΑΛΑΙΟ 3^ο: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	99
3.1 MATERIAL MODEL FOR COMPOSITES USING NEURAL NETWORKS (1993).....	101
3.2 NEW NESTED ADAPTIVE NEURAL NETWORKS (NANN) FOR CONSTITUTIVE MODELING (1997).....	101
3.3 APPLICATION OF NEURAL NETWORK TO THE MODELING OF SOME CONSTITUTIVE LAW (1998).....	102
3.4 AN INTERFACE FOR IMPLICIT CONSTITUTIVE MODELING (1998).....	102

3.5 IMPLICIT CONSTITUTIVE MODELING FOR VISCOPLASTICITY USING NEURAL NETWORKS (1998).....	103
3.6 NEURAL NETWORK BASED CONSTITUTIVE MODEL FOR RUBBER MATERIAL (2004)	103
3.7 SELF–LEARNING ELEMENTS FOR INVERSE ESTIMATION OF THERMAL CONSTITUTIVE MODELS (2005)...	104
3.8 NEURAL NETWORK MATERIAL MODEL ENHANCEMENT: OPTIMIZATION THROUGH VISUALIZATION AND SELECTIVE DATA REMOVAL (2005).....	104
3.9 CHARACTERIZING RATE–DEPENDENT MATERIAL BEHAVIORS IN SELF–LEARNING SIMULATION (2006)..	105
3.10 NEURAL NETWORK CONSTITUTIVE MODEL FOR RATE – DEPENDENT MATERIALS (2006).....	105
3.11 CONSTITUTIVE SOIL BEHAVIOR REPRESENTATION VIA ARTIFICIAL NEURAL NETWORKS: A SHIFT FROM SOIL MODELS TO SOIL BEHAVIOR DATA (2006)	106
3.12 NOVEL APPROACH TO INTEGRATION OF NUMERICAL MODELING AND FIELD OBSERVATION FOR DEEP EXCAVATIONS (2006).....	106
3.13 CONSTITUTIVE MODELING OF Sn/Ag AND Sn/Ag/Cu SOLDER ALLOYS (2006)	107
3.14 A NEW NEURAL NETWORK – BASED MODEL FOR HYSTERETIC BEHAVIOR (2007).....	108
3.15 NONLINEAR CONSTITUTIVE MODELS FOR FRP COMPOSITES USING ARTIFICIAL NEURAL NETWORKS (2007)	109
3.16 INTEGRATION OF LABORATORY TESTING AND CONSTITUTIVE MODELING OF SOILS (2007).....	109
3.17 EXTRACTING INELASTIC METAL BEHAVIOR THROUGH INVERSE ANALYSIS: A SHIFT IN FOCUS FROM MATERIAL MODELS TO MATERIAL BEHAVIOR (2007)	110
3.18 NEURAL NETWORK BASED CONSTITUTIVE MODEL FOR ELASTOMERIC FOAMS (2007)	111
3.19 APPLICATION OF ARTIFICIAL NEURAL NETWORK FOR CONSTITUTIVE MODELING IN FINITE ELEMENT ANALYSIS (2007).....	111
3.20 DEVELOPMENT OF NEURAL NETWORK BASED HYSTERETIC MODELS FOR STEEL BEAM- COLUMN CONNECTIONS THROUGH SELF–LEARNING SIMULATION (2007)	112
3.21 NEURAL NETWORK MATERIAL MODEL ENHANCEMENT: OPTIMIZATION THROUGH SELECTIVE DATA REMOVAL (2007)	113
3.22 FIRE ANALYSIS OF STEEL FRAMES WITH THE USE OF ARTIFICIAL NEURAL NETWORKS	113

3.23 ARTIFICIAL NEURAL NETWORKS IN NUMERICAL MODELING OF COMPOSITES (2008).....	114
3.24 SELF—LEARNING FEM/NMM APPROACH TO IDENTIFICATION OF EQUIVALENT MATERIAL MODELS FOR PLANE STRESS PROBLEM (2008).....	115
3.25 COMPARISON OF TWO INVERSE ANALYSIS TECHNIQUES FOR LEARNING DEEP EXCAVATION RESPONSE (2009)	115
3.26 INTELLIGENT FINITE ELEMENT METHOD AND APPLICATION TO SIMULATION OF BEHAVIOR OF SOILS UNDER CYCLIC LOADING (2009)	116
ΚΕΦΑΛΑΙΟ 4^ο : ΕΦΑΡΜΟΓΕΣ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ.....	119
4.1 ΓΕΝΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ NN.....	121
4.2 ΜΟΝΟΤΟΝΙΚΗ ΦΟΡΤΙΣΗ	122
4.2.1 ΜΟΝΟΑΞΟΝΙΚΟΣ ΕΦΕΛΚΥΣΜΟΣ ΔΟΜΙΚΟΥ ΧΑΛΥΒΑ ΘΕΡΜΗΣ ΕΛΑΣΗΣ ΣΕ ΜΕΤΑΒΑΤΙΚΕΣ ΣΥΝΘΗΚΕΣ ΘΕΡΜΑΝΣΗΣ	122
4.2.2 ΑΞΟΝΙΚΗ ΣΥΜΠΙΕΣΗ ΟΠΛΙΣΜΕΝΟΥ ΣΚΥΡΟΔΕΜΑΤΟΣ.....	125
4.2.3 ΚΑΤΑΣΤΑΤΙΚΟΣ ΝΟΜΟΣ ΔΙΑΤΜΗΣΗΣ ΕΔΑΦΟΥΣ.....	128
4.3 ΑΝΑΚΥΚΛΙΖΟΜΕΝΗ ΦΟΡΤΙΣΗ.....	130
4.3.1 ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΗΣ ΥΣΤΕΡΗΤΙΚΗΣ ΣΥΜΠΕΡΙΦΟΡΑΣ ΤΩΝ ΥΛΙΚΩΝ	130
<i>4.3.1.1. Περιγραφή της μεθόδου μοντελοποίησης της υστερητικής συμπεριφοράς των υλικών</i>	<i>131</i>
4.3.2 ΜΟΝΟΑΞΟΝΙΚΗ ΣΥΜΠΙΕΣΗ ΑΝΟΞΕΙΔΩΤΟΥ ΧΑΛΥΒΑ 316.....	135
4.3.3 ΜΟΝΟΑΞΟΝΙΚΗ ΣΥΜΠΙΕΣΗ ΟΠΛΙΣΜΕΝΟΥ ΣΚΥΡΟΔΕΜΑΤΟΣ	137
4.3.4 ΤΡΙΑΞΟΝΙΚΗ ΔΟΚΙΜΗ ΤΟΥ ΕΔΑΦΟΥΣ ΚΑΤΩ ΑΠΟ ΚΥΚΛΙΚΗ ΑΞΟΝΙΚΗ ΦΟΡΤΙΣΗ	142
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	147

ΚΕΦΑΛΑΙΟ 1^ο

ΚΑΤΑΣΤΑΤΙΚΟΙ ΝΟΜΟΙ

1.1 Εισαγωγή

Οι εξισώσεις που χαρακτηρίζουν τις φυσικές ιδιότητες των υλικών ονομάζονται καταστατικές εξισώσεις ή καταστατικοί νόμοι. Λόγω της μεγάλης ποικιλίας των υλικών και των φορτίσεων, η μελέτη και η ανάπτυξη καταστατικών εξισώσεων είναι ίσως ένα από τα πιο ενδιαφέροντα πεδία της μηχανικής. Αν και η θεωρία της μηχανικής του συνεχούς μέσου έχει καθιερώσει κάποιες βασικές αρχές για τη συστηματική ανάπτυξη καταστατικών νόμων (Malvern 1969), πολλοί καταστατικοί νόμοι έχουν αναπτυχθεί μέσα από εμπειρικές σχέσεις που βασίζονται σε πειραματικές ενδείξεις.

Οι καταστατικές εξισώσεις ή νόμοι, κατά γενικό κανόνα, είναι σχέσεις μεταξύ τάσεων σ και παραμορφώσεων ϵ . Πιο συγκεκριμένα οι τάσεις σ εκφράζονται συναρτήσει των παραμορφώσεων ϵ , του ρυθμού των παραμορφώσεων $\dot{\epsilon}$, της ιστορίας των παραμορφώσεων και των φυσικών ιδιοτήτων του υλικού. Η πιο απλή περίπτωση καταστατικού νόμου, στον οποίο δεν υπάρχει εξάρτηση της τάσης από το ρυθμό ή την ιστορία της παραμόρφωσης, είναι ο νόμος του Hooke για τα γραμμικά ελαστικά υλικά. Αυτός εκφράζεται από τη σχέση:

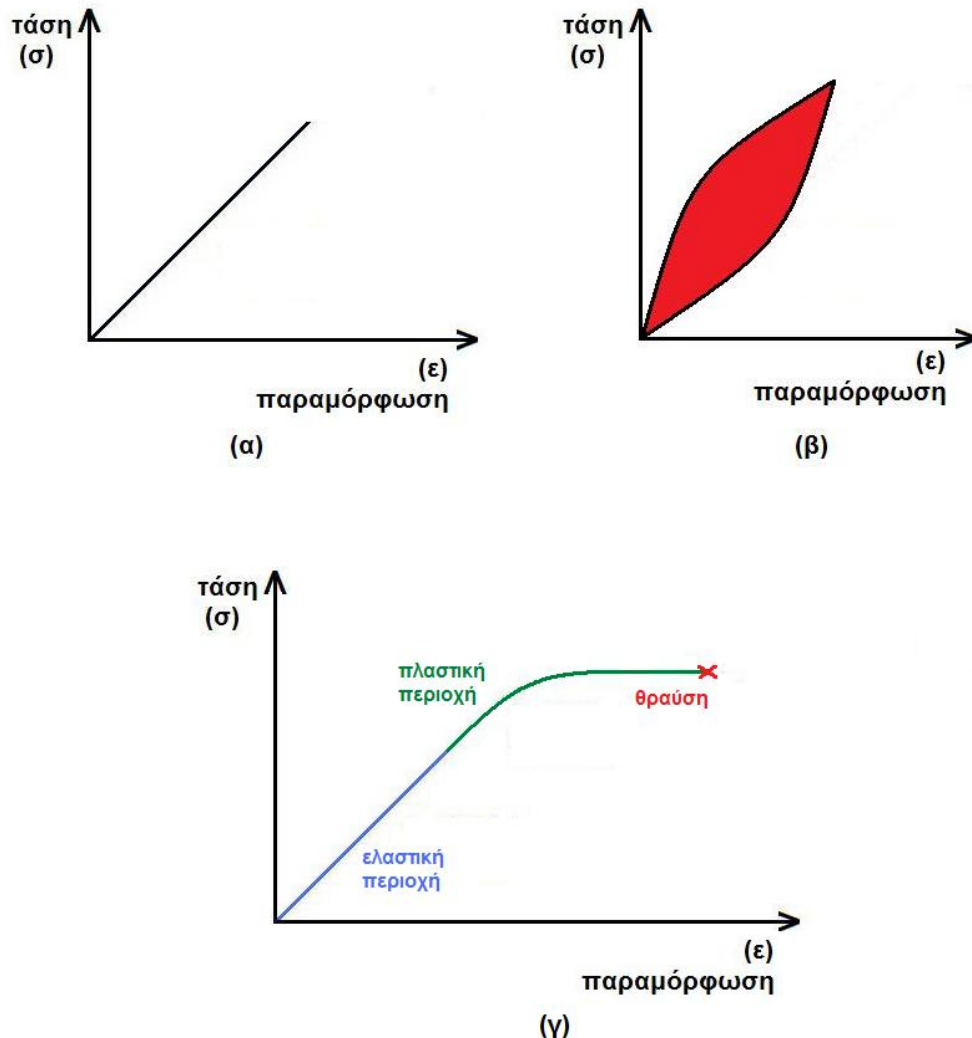
$$\sigma = E\epsilon$$

όπου E είναι ο λόγος σ/ϵ και ονομάζεται μέτρο ελαστικότητας. Για μικρές παραμορφώσεις όλα τα στερεά υλικά συμπεριφέρονται σαν γραμμικά ελαστικά υλικά [1]. Γενικά, η αντίδραση ενός υλικού κάτω από ένα συγκεκριμένο είδος καταπόνησης είναι είτε ελαστική, είτε ανελαστική, η οποία με τη σειρά της μπορεί να είναι ιξωδοελαστική (ή βισκοελαστική) ή πλαστική. Αναλυτικότερα:

- ✓ **Ελαστικά** υλικά είναι αυτά, στα οποία όταν εφαρμοσθεί μια τάση και στη συνέχεια αφαιρεθεί, η παραμόρφωση που είχε προκληθεί, αρχικά, «χάνεται» και το υλικό επανέρχεται στην αρχική του κατάσταση. Η καταστατική εξίσωση που αντιπροσωπεύει αυτά τα υλικά είναι ο νόμος του Hooke.
- ✓ **Ιξωδοελαστικά (ή βισκοελαστικά)** είναι τα υλικά, τα οποία εμφανίζουν συγχρόνως ιξώδη και ελαστικά χαρακτηριστικά, ενώ υφίστανται κάποια παραμόρφωση. Αυτό έχει ως αποτέλεσμα την εμφάνιση ενός βρόχου υστέρησης στην καμπύλη τάσης – παραμόρφωσης και την χρονική εξάρτηση της αντίδρασης (παραμόρφωσης) του υλικού σε μια δεδομένη φόρτιση.
- ✓ **Πλαστικά** υλικά είναι αυτά, στα οποία όταν εφαρμοσθεί μία τάση συμπεριφέρονται ελαστικά μέχρι μία συγκεκριμένη τιμή της τάσης αυτής και πάνω από αυτή την τιμή (όριο διαρροής) συμπεριφέρονται πλαστικά, δηλαδή

δεν μπορούν να επανέλθουν στο αρχικό τους σχήμα και η παραμόρφωση είναι μόνιμη.

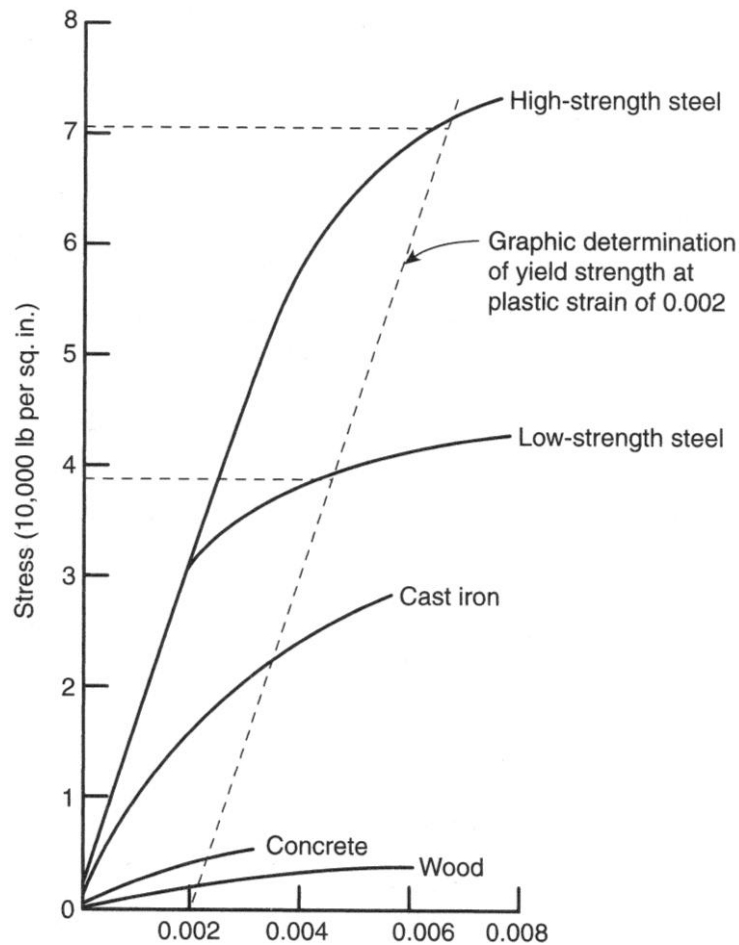
Στη συνέχεια παρουσιάζονται οι καμπύλες τάσης – παραμόρφωσης που αντιστοιχούν στην ελαστική, βισκοελαστική και πλαστική συμπεριφορά των υλικών.



Σχήμα 1.1: Καμπύλες σ - ϵ για (α) ελαστική, (β) βισκοελαστική και (γ) πλαστική συμπεριφορά των υλικών

Για τα πραγματικά υλικά ο προσδιορισμός της μηχανικής τους συμπεριφοράς γίνεται μέσα από πειραματικές δοκιμές, όπως για παράδειγμα η δοκιμή εφελκυσμού. Σε αυτή την τεχνική, ειδικά προετοιμασμένα, κυλινδρικά δείγματα φορτίζονται αξονικά σε μια μηχανή δοκιμής. Η παραμόρφωση προσδιορίζεται από την αλλαγή του μήκους, σε προσημασμένα σημεία αναφοράς πάνω στο δείγμα και συνήθως μετράται από επιμηκυσσιόμετρα. Τα δεδομένα φόρτισης, τα οποία συλλέγονται από έναν αισθητήρα φορτίου, διαιρούνται με την διατομή στην περιοχή δοκιμής για να υπολογίσει την τάση. Τα δεδομένα αξονικής τάσης – παραμόρφωσης καταγράφονται και σχηματίζουν

διάγραμμα με βάση πρότυπες πειραματικές τεχνικές. Το επόμενο σχήμα δείχνει καμπύλες μονοαξονικής τάσης – παραμόρφωσης διαφόρων δομικών υλικών.



Σχήμα 1.2: Καμπύλες τάσεων – παραμορφώσεων διαφόρων υλικών

Όπως έχει ήδη αναφερθεί η παρούσα εργασία ασχολείται με τη μοντελοποίηση καταστατικών νόμων διαφόρων υλικών, μέσω νευρωνικών δικτύων. Για το λόγο αυτό στις επόμενες ενότητες γίνεται αναφορά στους καταστατικούς νόμους που αφορούν στο χάλυβα, στο σκυρόδεμα και στο έδαφος.

1.2 Καταστατικοί νόμοι χάλυβα

Οι καταστατικές σχέσεις του χάλυβα περιλαμβάνουν:

- ✓ την ελαστική συμπεριφορά του υλικού,
- ✓ το κριτήριο διαρροής του υλικού που καθορίζει πότε διαρρέει το υλικό,
- ✓ τον κανόνα ροής που καθορίζει την συμπεριφορά του υλικού μετά την διαρροή και

- ✓ τις παραμέτρους ανακύκλισης που ορίζουν την συμπεριφορά του υλικού σε ανακύκλιση (κυκλική φόρτιση).

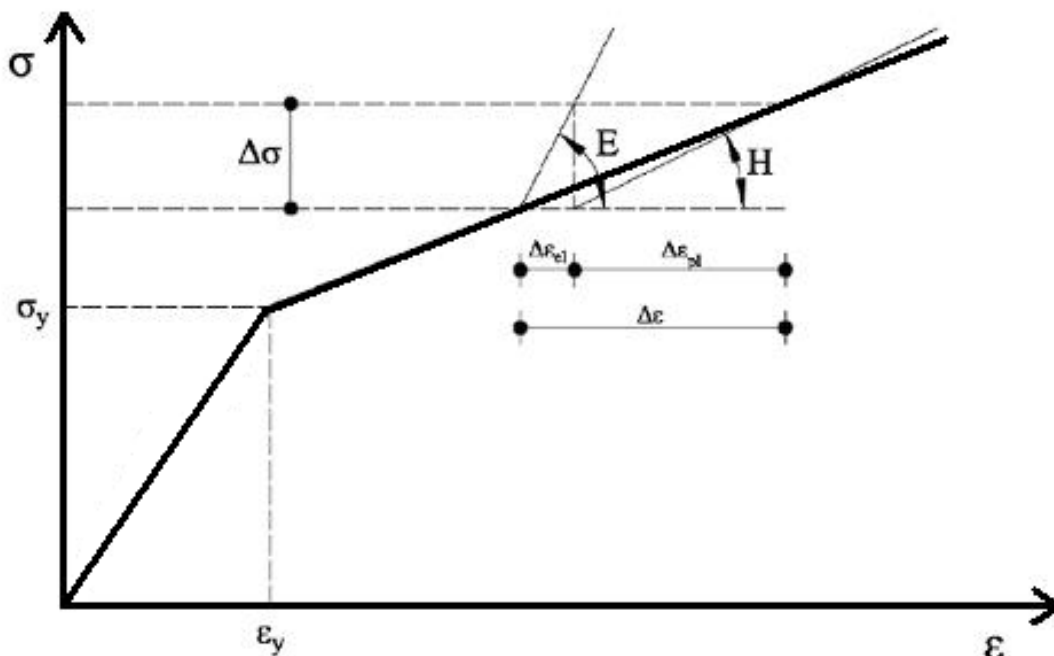
1.2.1 Διγραμμικό μοντέλο χάλυβα

Σε ένα διγραμμικό μοντέλο υλικού (σχήμα 1.3) η ελαστική συμπεριφορά του υλικού υπολογίζεται από τον νόμο του Hooke, έχοντας το αρχικό μέτρο ελαστικότητας E και το σημείο διαρροής, το οποίο καθορίζεται από την τάση διαρροής σ_y . Η συμπεριφορά του υλικού μετά το σημείο διαρροής προσδιορίζεται από μία δεύτερη κεκλιμένη γραμμή, της οποίας η κλίση προσδιορίζεται από το εφαπτομενικό μητρώο δυσκαμψίας E_T . Το εφαπτομενικό μητρώο δυσκαμψίας E_T συνδέεται με το αρχικό μέτρο ελαστικότητας, μέσω του συντελεστή σκλήρυνσης b από την σχέση:

$$E_T = b \cdot E \quad (1.1)$$

Για ένα τέλεια πλαστικό υλικό ο συντελεστής σκλήρυνσης είναι μηδέν. Εκτός όμως από τον συντελεστή σκλήρυνσης b , πολλές φορές χρησιμοποιείται και η παράμετρος σκλήρυνσης H , που συνδέει την τάση σ με την πλαστική παραμόρφωση ε_{pl} μέσα από τη σχέση:

$$\sigma = H \cdot \varepsilon_{pl} \quad (1.2)$$



Σχήμα 1.3: Διγραμμικός νόμος υλικού (για χάλυβα)

Η παράμετρος σκλήρυνσης H , το εφαπτομενικό μητρώο δυσκαμψίας E_T και ο συντελεστής σκλήρυνσης b συνδέονται με τη σχέση:

$$H = \frac{E_T}{1-b} \quad (1.3)$$

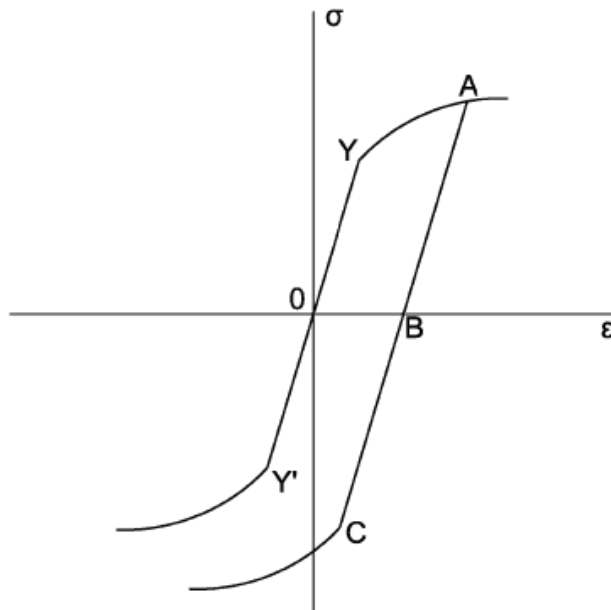
ή

$$E_T = E \cdot \left(1 - \frac{E}{E-H} \right) \quad (1.4)$$

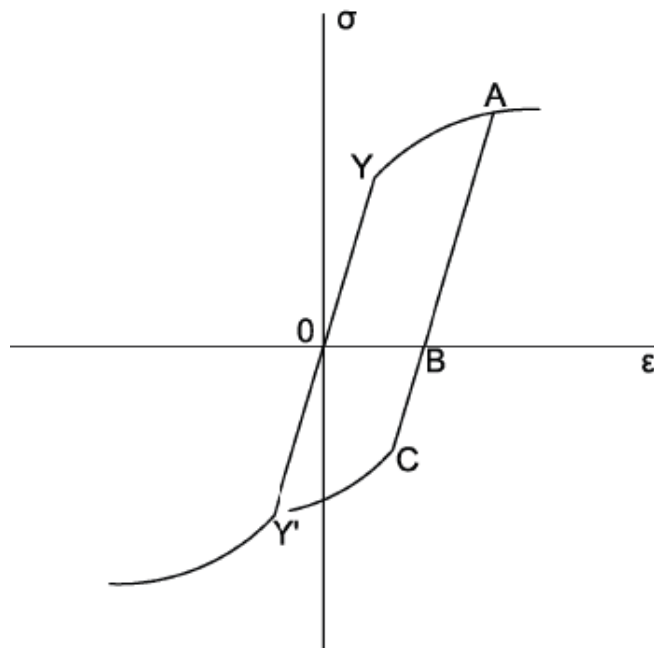
Για χαλύβδινους φορείς, όταν η τάση διαρροής σε εφελκυσμό υπερβεί την τάση διαρροής σε θλίψη, τότε η αντοχή σε διαρροή στον επόμενο κύκλο φόρτισης θα είναι διαφορετική. Αυτό είναι γνωστό ως φαινόμενο Bauschinger. Το φαινόμενο Bauschinger αναφέρεται σε μια ιδιότητα των υλικών, όπου τα χαρακτηριστικά τάσεων – παραμορφώσεων του υλικού, αλλάζουν, ως αποτέλεσμα της μικροσκοπικής κατανομής τάσεων στο υλικό. Για παράδειγμα, η αύξηση της εφελκυστικής αντοχής σε διαρροή γίνεται σε βάρος της θλιπτικής αντοχής σε διαρροή. Το φαινόμενο Bauschinger μπορεί να προσομοιωθεί με κινηματική ή ιστροπική σκλήρυνση.

Για την ιστροπική σκλήρυνση η επιφάνεια διαρροής θεωρείται ότι επεκτείνεται ισότροπα σε μέγεθος, διατηρώντας το κέντρο της, με βάση αυτή την υπόθεση. Αυτό σημαίνει πως, η πλαστική παραμόρφωση δεν επηρεάζει το σχήμα της επιφάνειας διαρροής. Αν το αρχικό σημείο διαρροής υποδηλώνεται με το γράμμα Υ στην καμπύλη τάσεων – παραμορφώσεων (σχήμα 1.4), η τάση αποφόρτισης AB, ξεκινώντας από το σημείο A είναι ίδια με την τάση διαρροής AC στη συμπίεση και είναι μεγαλύτερη από την αρχική τάση διαρροής. Η ιδέα αυτής της υπόθεσης είναι πολύ απλή για να εκφράσει την γενίκευση του φαινομένου Bauschinger, αλλά παρόλα αυτά χρησιμοποιείται ευρέως σε πολλά είδη αναλύσεων, όταν η διεύθυνση της φόρτισης αλλάζει ελαφρώς.

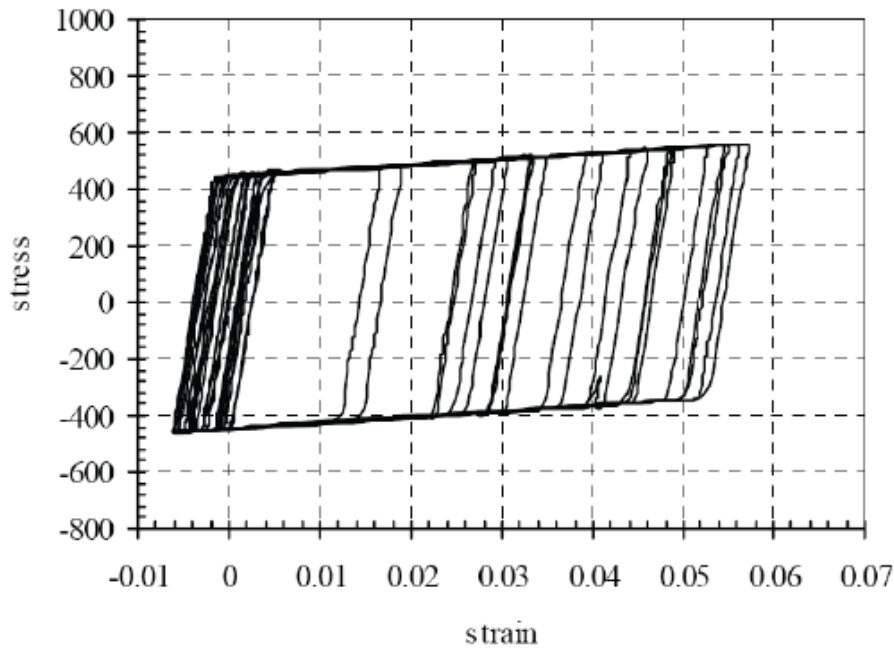
Για την κινηματική σκλήρυνση το μοντέλο θεωρεί, ότι η επιφάνεια διαρροής με αμετάβλητο σχήμα κινείται στο χώρο των τάσεων, λόγω της πλαστικής παραμόρφωσης όπως παρουσιάζεται στο σχήμα (1.5) και για το διάγραμμα τάσεων – παραμορφώσεων ισχύει ότι $YY' = AC$. Αυτό είναι το αποτέλεσμα της μετακίνησης του υστερητικού βρόχου στην καμπύλη τάσεων – παραμορφώσεων.



Σχήμα 1.4: Ισοτροπική σκλήρυνση



Σχήμα 1.5: Κινηματική σκλήρυνση



Σχήμα 1.6: Παράδειγμα κινηματικής σκλήρυνσης

1.2.2 Υλικό μοντέλο Menegotto - Pinto

Οι Menegotto και Pinto (1973) [2] πρότειναν ένα μοντέλο για τον οπλισμένο χάλυβα, για το οποίο η σχέση τάσεων – παραμορφώσεων δίνεται από την παρακάτω μη γραμμική εξίσωση.

$$\sigma^* = b\varepsilon^* + \frac{(1-b)\varepsilon^*}{(1+\varepsilon^{*R})^{1/R}} \quad (1.5)$$

όπου η πραγματική τάση και παραμόρφωση (σ^* , ε^*) είναι μία συνάρτηση του χρονικού διαστήματος φόρτισης / αποφόρτισης και δίνεται από την παρακάτω σχέση:

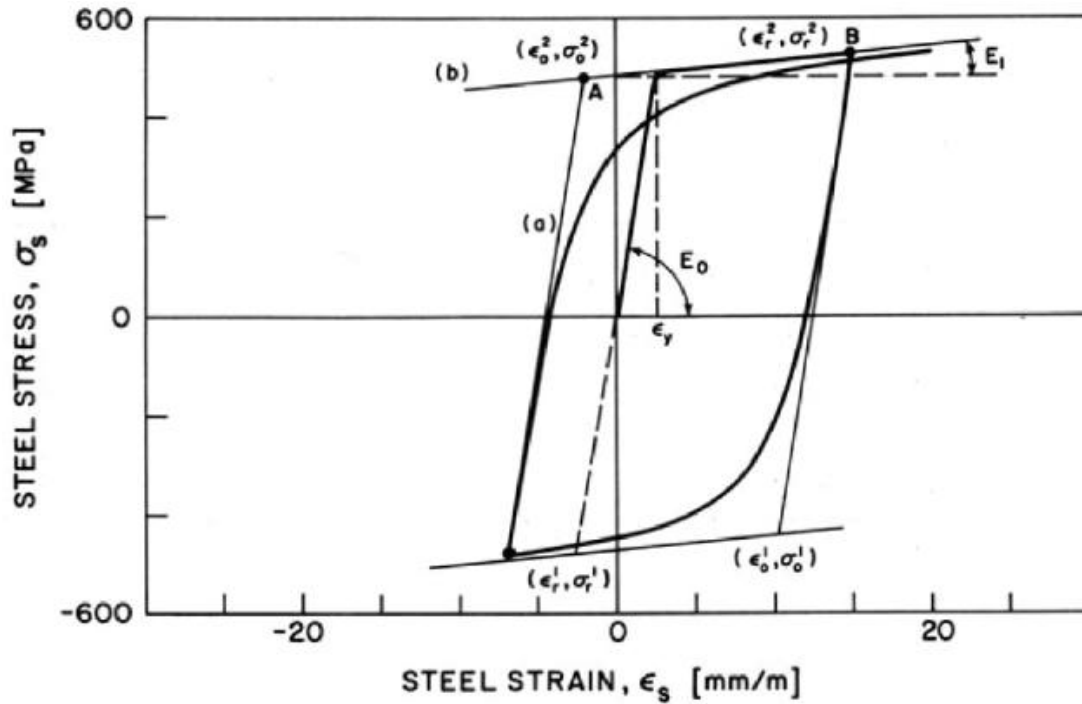
$$\varepsilon^* = \frac{\varepsilon - \varepsilon_r}{\varepsilon_o - \varepsilon_r} \quad (1.6)$$

και

$$\sigma^* = \frac{\sigma - \sigma_r}{\sigma_o - \sigma_r} \quad (1.7)$$

b είναι ο λόγος του αρχικού μέτρου ελαστικότητας \mathbf{E} προς το εφαπτομενικό μητρώο δυσκαμψίας \mathbf{E}_T και \mathbf{R} είναι μία παράμετρος που προσδιορίζει το σχήμα της καμπύλης αποφόρτισης (σχήμα 1.7). Σε αυτή την εφαρμογή θεωρείται, ότι οι καμπύλες αναφοράς (καμπύλες τάσης – παραμόρφωσης που περιβάλλουν την κυκλική απόκριση), όπως και οι καμπύλες φόρτισης – αποφόρτισης μπορεί να χαρακτηρίζονται από την εξίσωση (1.5).

Η εφαρμογή αυτή παραβλέπει επίσης και τη «στάθμη» του σημείου διαρροής. Τα τελευταία χρόνια, αρκετοί ερευνητές έχουν προτείνει υλικά μοντέλα που χρησιμοποιούν την εξίσωση Menegotto-Pinto για τον προσδιορισμό της συμπεριφοράς του ενισχυμένου χάλυβα υπό ανακυκλιζόμενη φόρτιση. Ένα τέτοιο μοντέλο είναι αυτό που προτείνουν οι Stanton και McNiven (1979) [3]. Το μοντέλο αυτό χρησιμοποιεί μια προσεγγιστική έκδοση της εξίσωσης Menegotto (εξίσωση 1.5) για τη βελτίωση της υπολογιστικής απόδοσης και θεωρεί ότι οι καμπύλες αναφοράς για το χάλυβα, ο οποίος υπόκειται σε ανακυκλιζόμενη φόρτιση ακολουθεί τη μονοτονική περιβάλλουσα. Ένα δεύτερο μοντέλο είναι αυτό που προτείνει ο Φιλίππου (1983) [4]. Το μοντέλο αυτό ενσωματώνει την εξίσωση (1.5), ακριβώς για να περιγράψει την συμπεριφορά του χάλυβα σε περίπτωση αποφόρτισης. Το μοντέλο απορρέει από την υπόθεση, ότι η καμπύλη αναφοράς, που προσδιορίζει την σχέση τάσεων – παραμορφώσεων σε ανακυκλιζόμενη φόρτιση, είναι τρι – γραμμική. Η ισοτροπική, κυκλική τάση σκλήρυνσης έχει ενσωματωθεί, μέσω μετατόπισης της καμπύλης αναφοράς ως συνάρτηση του βήματος της πλαστικής παραμόρφωσης. Πρόσφατα, η εξίσωση (1.5) έχει ενσωματωθεί σε ένα εξελιγμένο μοντέλο που προτείνουν οι Chang και Mander (1994) [5]. Το μοντέλο αυτό υποθέτει, ότι το σχήμα της καμπύλης αναφοράς ορίζεται από την μονοτονική σχέση τάσεων – παραμορφώσεων. Το μοντέλο αιτιολογεί την κυκλική παραμόρφωση σκλήρυνσης, μέσω της μεταφοράς της καμπύλης αναφοράς ως συνάρτηση της ιστορίας της παραμόρφωσης. Επιπλέον, το μοντέλο ενσωματώνει την μεταβλητότητα της αρχικής δυσκαμψίας στη φάση της αποφόρτισης, την κυκλική παραμόρφωση χαλάρωσης και προηγούμενους κύκλους φόρτισης – αποφόρτισης. Κάθε ένα από τα τέσσερα αυτά μοντέλα, προβλέπει με ικανοποιητική ακρίβεια την παρατηρούμενη συμπεριφορά του ενισχυμένου χάλυβα, που υπόκειται σε ανακυκλιζόμενη φόρτιση και είναι παρόμοια με τη συμπεριφορά κατασκευών από οπλισμένο σκυρόδεμα, οι οποίες υπόκεινται σε σεισμικά φορτία. Το μοντέλο που προτείνουν οι Chang και Mander (1994), παρέχει πρόβλεψη της συμπεριφοράς του χάλυβα με ικανοποιητική ακρίβεια, ενώ το προτεινόμενο μοντέλο από τον Φιλίππου παρέχει αφενός την πρόβλεψη της συμπεριφοράς του χάλυβα με ικανοποιητική ακρίβεια και αφετέρου σχετικά απλή εφαρμογή και βαθμονόμηση. Σύμφωνα με τον Castiglioni το προτεινόμενο μοντέλο των Menegotto – Pinto μπορεί επίσης να βαθμονομηθεί και να χρησιμοποιηθεί στις χαλύβδινες κατασκευές.



Σχήμα 1.7: Υλικό μοντέλο των Menegotto – Pinto

1.2.3 Τρισδιάστατο καταστατικό μοντέλο

Σε ένα τρισδιάστατο καταστατικό μοντέλο το καταστατικό μητρώο δίνεται από τον τύπο:

$$C_{3D} = \frac{1}{1+\nu} \begin{bmatrix} \frac{1-\nu}{1-2\nu} \frac{\sigma_x^2}{S} & & & & & & & & \text{sym.} \\ \nu \frac{\sigma_x^2 \sigma_y^2}{1-2\nu S} & \frac{1-\nu}{1-2\nu} \frac{\sigma_y^2}{S} & & & & & & & \\ \nu \frac{\sigma_x^2 \sigma_z^2}{1-2\nu S} & \nu \frac{\sigma_y^2 \sigma_z^2}{1-2\nu S} & \frac{1-\nu}{1-2\nu} \frac{\sigma_z^2}{S} & & & & & & \\ -\frac{\sigma'_x \tau_{xy}^2}{S} & -\frac{\sigma'_y \tau_{xy}^2}{S} & -\frac{\sigma'_z \tau_{xy}^2}{S} & \frac{1}{2} \frac{\tau_{xy}^2}{S} & & & & & \\ -\frac{\sigma'_x \tau_{yz}^2}{S} & -\frac{\sigma'_y \tau_{yz}^2}{S} & -\frac{\sigma'_z \tau_{yz}^2}{S} & -\frac{\tau_{xy} \tau_{yz}}{S} & \frac{1}{2} \frac{\tau_{yz}^2}{S} & & & & \\ -\frac{\sigma'_x \tau_{zx}^2}{S} & -\frac{\sigma'_y \tau_{zx}^2}{S} & -\frac{\sigma'_z \tau_{zx}^2}{S} & -\frac{\tau_{xy} \tau_{zx}}{S} & -\frac{\tau_{yz} \tau_{zx}}{S} & \frac{1}{2} \frac{\tau_{zx}^2}{S} & & & \end{bmatrix}$$

όπου:

$$S = \frac{2}{3} \bar{\sigma}^2 \left(1 + \frac{H'}{3G} \right) \quad (1.8)$$

και H' είναι η κλίση του βήματος $\delta\bar{\sigma}$ προς το βήμα της πλαστικής παραμόρφωσης $\delta\varepsilon_p$ και $\bar{\sigma}$ είναι η πραγματική τάση. Το κριτήριο von Mises δηλώνει ότι η διαρροή εμφανίζεται όταν:

$$\bar{\sigma} - \sigma_y = 0 \Leftrightarrow \sqrt{\frac{3}{2} \sigma'_{ij} \sigma'_{ij}} - \sigma_y = 0 \Leftrightarrow \sqrt{3J_2} - \sigma_y = 0 \quad (1.9)$$

Στην παραπάνω εξίσωση J_2 είναι η δεύτερη αναλλοίωτη του τανυστή της τάσης, σ_y είναι η τάση διαρροής μιας μονοαξονικής δοκιμής και σ_{ij} είναι οι αποκλίνοντες τάσεις.

1.3 Καταστατικοί νόμοι σκυροδέματος

Για να αξιολογηθεί αποτελεσματικά η γραμμική συμπεριφορά του οπλισμένου σκυροδέματος, τρία θέματα θα πρέπει να μοντελοποιηθούν με ακρίβεια:

- ✓ η συμπεριφορά του σκυροδέματος,
- ✓ η συμπεριφορά του οπλισμού και
- ✓ ο δεσμός μεταξύ του οπλισμού και του σκυροδέματος.

Αν και ο δεσμός μεταξύ του οπλισμού και του σκυροδέματος είναι πολύ σημαντικό κομμάτι για την ανάλυση των κατασκευών από οπλισμένο σκυροδέμα, η παρούσα εργασία επικεντρώνεται μόνο στην καταστατική μοντελοποίηση του συγκεκριμένου υλικού. Στις επόμενες περιπτώσεις που θα μελετηθούν γίνεται η θεώρηση, ότι ο δεσμός μεταξύ οπλισμού και σκυροδέματος είναι τέλειος.

1.3.1 Υλικό μοντέλο των Kent και Park

Η συμπεριφορά του σκυροδέματος σε μονοαξονική συμπίεση υπολογίζεται από το υλικό μοντέλο των Kent και Park, το οποίο τροποποιήθηκε στη συνέχεια από τον Scott και είναι γνωστό και ως τροποποιημένο μοντέλο Kent και Park [6]. Αν και έχουν προταθεί πιο πολύπλοκα και ίσως πιο ακριβή υλικά μοντέλα, για να περιγράψουν τη συμπεριφορά του σκυροδέματος, το τροποποιημένο μοντέλο των Kent και Park συνδυάζει την απλότητα και την ακρίβεια. Η αντοχή του σκυροδέματος σε εφελκυσμό θεωρείται αμελητέα, ενώ η σχέση τάσεων και παραμορφώσεων υπό συμπίεση, αποτελείται από τρία τμήματα:

$$\sigma = \begin{cases} Kf'_c \left[\frac{2\varepsilon}{\varepsilon_o} - \left(\frac{\varepsilon}{\varepsilon_o} \right)^2 \right] & , \varepsilon \leq \varepsilon_o \\ Kf'_c [1 - Z(\varepsilon - \varepsilon_o)] \geq 0.2Kf' & , \varepsilon_o \leq \varepsilon \leq \varepsilon_u \\ 0.2Kf' & , \varepsilon_u \leq \varepsilon \end{cases} \quad (1.10)$$

$$\varepsilon_o = 0.002K \quad (1.11)$$

$$K = 1 + \frac{\rho_s f_{yh}}{f'_c} \quad (1.12)$$

$$Z = \frac{0.5}{\frac{3 + 0.29f'_c}{145f'_c - 1000} + 0.75\rho_s \sqrt{\frac{h'}{s_h}} - 0.002K} \quad (1.13)$$

όπου ε_o είναι η παραμόρφωση του σκυροδέματος που αντιστοιχεί στην μέγιστη τάση, K είναι ένας συντελεστής ο οποίος αντιπροσωπεύει την αύξηση της αντοχής, λόγω περίσφιξης (για απερίσφικτο σκυρόδεμα θεωρείται ότι είναι ίσος με 1), Z είναι η κλίση του φθίνοντος τμήματος της καμπύλης (σχήμα 1.8), f'_c είναι η αντοχή σε συμπίεση σε MPa, f_{yh} είναι η τάση διαρροής του εγκάρσιου οπλισμού σε MPa, ρ_s είναι ο λόγος του όγκου των συνδετήρων προς τον περισφυγμένο πυρήνα του σκυροδέματος, h' είναι η διάσταση του πυρήνα του σκυροδέματος και s_h είναι η κάθετη απόσταση μεταξύ των συνδετήρων. Η τελική παραμόρφωση του περιορισμένου σκυροδέματος μπορεί να υπολογιστεί, όπως προτείνει ο Scott, από τον τύπο που ακολουθεί:

$$\varepsilon_{\max} = 0.004 + 0.9\rho_s \left(\frac{f_{yh}}{300} \right) \quad (1.14)$$

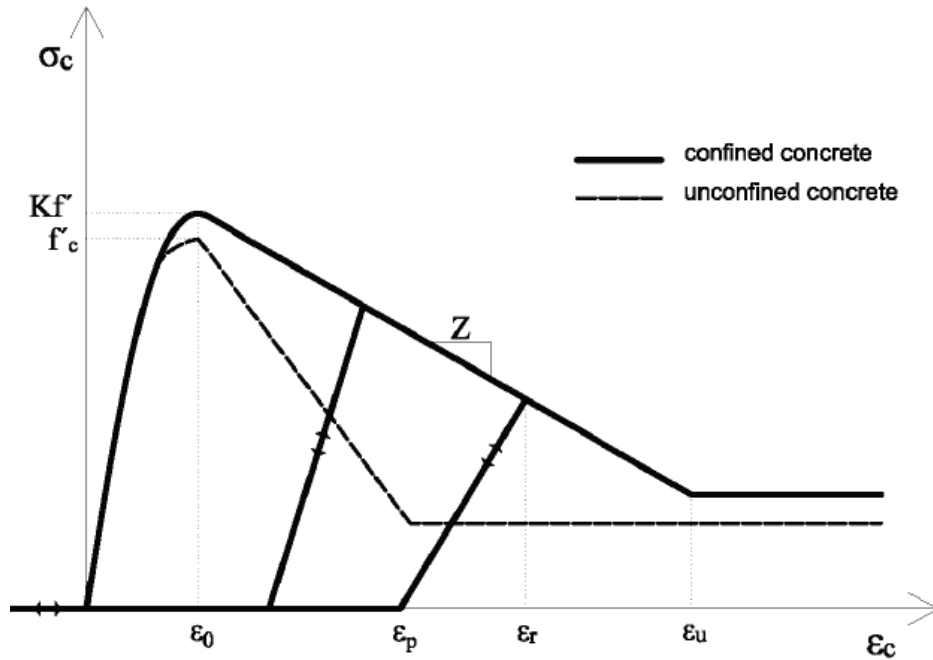
Όταν η παραμόρφωση συμπίεσης υπερβαίνει την τιμή ε_u , θεωρείται ότι είναι ίση με την τιμή 0.005 και η αντοχή γίνεται ίση με την τιμή $0.2f'_c$.

Η υστερητική συμπεριφορά διέπεται από τους παρακάτω κανόνες:

- ✓ Η σχέση μεταξύ της παραμόρφωσης ε_r της περιβάλλουσας καμπύλης στη φάση της αποφόρτισης, που ονομάζεται περιβάλλουσα παραμόρφωση και της παραμόρφωσης ε_p που είναι μηδέν και ονομάζεται πλαστική παραμόρφωση, δίνεται από τον τύπο:

$$\frac{\varepsilon_p}{\varepsilon_o} = \begin{cases} 0.145 \left(\frac{\varepsilon_r}{\varepsilon_o} \right)^2 + 0.13 \left(\frac{\varepsilon_r}{\varepsilon_o} \right) & , \frac{\varepsilon_r}{\varepsilon_o} < 2 \\ 0.707 \left(\frac{\varepsilon_r}{\varepsilon_o} - 2 \right) + 0.834 & , \frac{\varepsilon_r}{\varepsilon_o} \geq 2 \end{cases} \quad (1.15)$$

Το πρώτο σκέλος της παραπάνω εξίσωσης εισήχθη από τους Karsan και Jirsa (1996) και είναι ανέφικτο να δώσει αποδεκτά αποτελέσματα για σχετικά μεγάλες παραμορφώσεις, γι 'αυτό και στη συνέχεια προστέθηκε και το δεύτερο σκέλος.



Σχήμα 1.8: Τροποποιημένο μοντέλο των Kent – Park κατά τη διάρκεια φόρτισης και αποφόρτισης

- ✓ Η τάση του σκυροδέματος είναι ίση με μηδέν, για παραμόρφωση μικρότερη από εκείνη που αντιστοιχεί σε πλήρη αποφόρτιση, δεδομένου ότι η αντοχή σε εφελκυσμό δεν υπολογίζεται. Αυτή η περίπτωση είναι αντιπροσωπευτική της εμφάνισης ρωγμών.
- ✓ Η τάση του σκυροδέματος είναι ίση με μηδέν κατά τη διάρκεια της φόρτισης, εφόσον η παραμόρφωση είναι μικρότερη από εκείνη που αντιστοιχεί σε πλήρη αποφόρτιση ε_p . Όταν η παραμόρφωση γίνεται μεγαλύτερη από την παραπάνω τιμή, η καμπύλη φόρτισης είναι η ίδια με την καμπύλη αποφόρτισης (με αντίθετη κατεύθυνση). Αν και η πειραματική συμπεριφορά του σκυροδέματος δεν είναι συμβατή με την παραδοχή αυτή, η διαφορά από την άποψη των αποτελεσμάτων δεν είναι σημαντική. Επομένως για λόγους απλότητας χρησιμοποιείται η ίδια καμπύλη για φόρτιση και αποφόρτιση.

Το μέτρο ελαστικότητας E του σκυροδέματος υπολογίζεται ως εφαπτομένη της αντίστοιχης καμπύλης:

$$E = \begin{cases} 2Kf'_c \left(1 - \frac{\varepsilon}{\varepsilon_o} \right) & , \varepsilon \leq \varepsilon_o \\ Z & , \varepsilon_o \leq \varepsilon \leq \varepsilon_u \\ 0 & , \varepsilon_u \leq \varepsilon \end{cases} \quad (1.16)$$

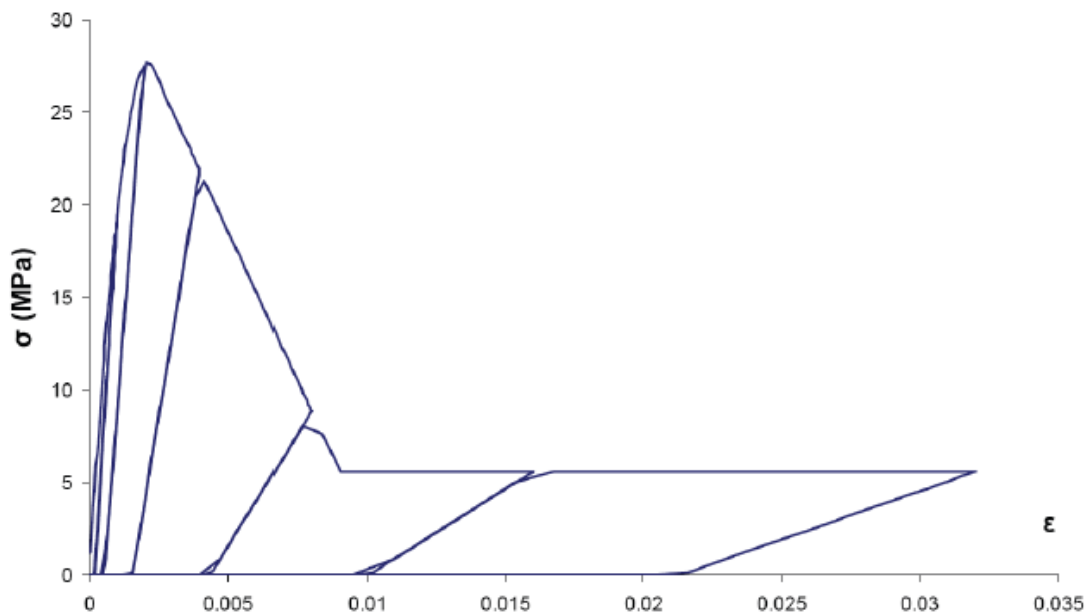
Στο πρώτο βήμα της φόρτισης το μέτρο ελαστικότητας έχει την τιμή:

$$E_o = 2 \left(\frac{Kf'_c}{\varepsilon_o} \right) \quad (1.17)$$

και κατά τη διάρκεια της αποφόρτισης δίνεται από τον τύπο:

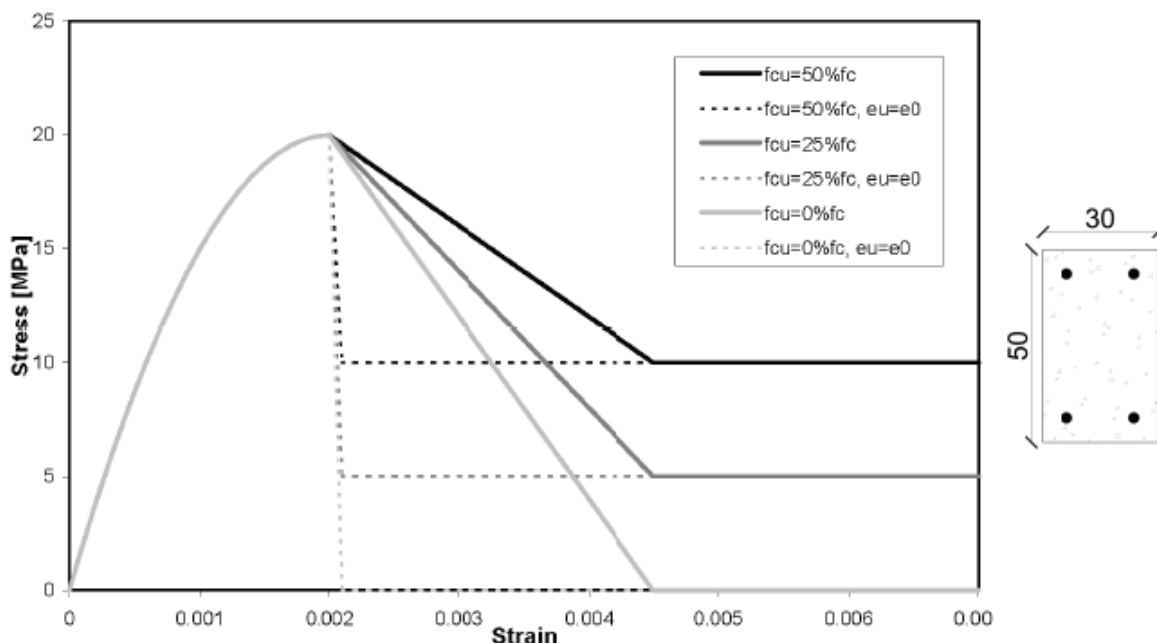
$$E = \begin{cases} \frac{\sigma_r}{\varepsilon_r - \varepsilon_p} & , \varepsilon_r - \varepsilon_p \geq \frac{\sigma_r}{E_o} \\ E_o & , \varepsilon_r - \varepsilon_p < \frac{\sigma_r}{E_o} \end{cases} \quad (1.18)$$

όπου σ_r είναι η τάση που αντιστοιχεί στο τρέχων σημείο αποφόρτισης. Η συμπεριφορά του υλικού υπό την επίδραση ενός σιγμοειδούς αυξανόμενου κυκλικού φορτίου παρουσιάζεται στο σχήμα (1.9).



Σχήμα 1.9: Τροποποιημένο μοντέλο των Kent – Park υπό κυκλική φόρτιση

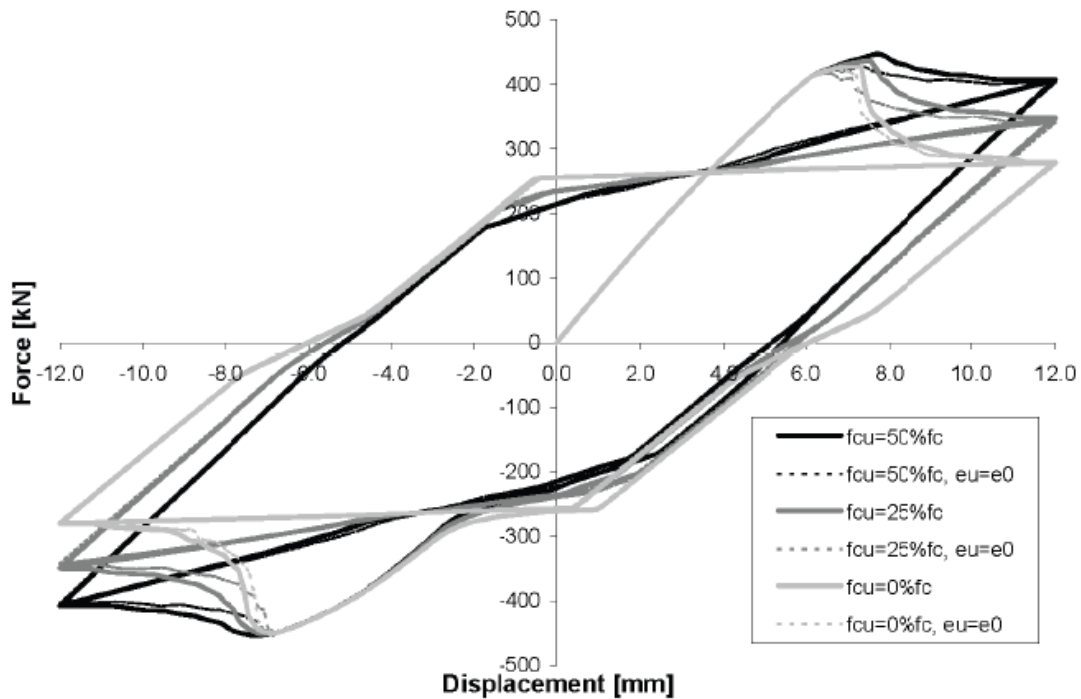
Στο παραπάνω διάγραμμα παρατηρείται, ότι στη σχέση τάσεων – παραμορφώσεων του σκυροδέματος δεν παίζει κανένα ρόλο ο αριθμός των κύκλων φόρτισης, αν και αντιπροσωπεύει την ενέργεια που απορροφάται σε κάθε κύκλο φόρτισης.



Σχήμα 1.10: (α) Διάγραμμα τάσεων – παραμορφώσεων για διαφορετικές καμπύλες αποφόρτισης, (β) διατομή του δοκιμίου

Η επίδραση της καμπύλης αποφόρτισης (ή καμπύλη μετά την μέγιστη τάση) του καταστατικού νόμου του σκυροδέματος για μονοαξονική φόρτιση αποδεικνύεται στο δοκίμιο σκυροδέματος που παρουσιάζεται στο σχήμα (1.10.β). Το δοκίμιο καταπονείται με αξονικό φορτίο στην κορυφή του, ενώ το μήκος του είναι ίσο με 1.5m, έχει ορθογώνια διατομή 50x30 cm και είναι ενισχυμένο με 4Φ20 ράβδους. Για τη μελέτη της επίδρασης της καμπύλης αποφόρτισης του υλικού, θεωρούνται τρεις τιμές της παραμένουσας τάσης f_{cu} , οι οποίες είναι το 50, 25 και 0% της μέγιστης τάσης συμπίεσης f_c' . Σε κάθε περίπτωση θεωρείται, ότι η καμπύλη αποφόρτισης έχει μεγάλη κλίση (απότομη αποφόρτιση) από την παραμένουσα τάση ($e_u = e_0$) και μετά και η καμπύλη χαλάρωσης έχει μηδενική κλίση σε σχέση με τον οριζόντιο άξονα. Επίσης στο σχήμα (1.10.α) παρουσιάζονται και καμπύλες αποφόρτισης με μικρότερη κλίση.

Η συμπεριφορά του δοκιμίου, το οποίο καταπονείται με ανακυκλιζόμενη φόρτιση, για όλες τις περιπτώσεις παραμένουσας τάσης, απεικονίζεται στο παρακάτω διάγραμμα τάσεων – παραμορφώσεων. Είναι εμφανές από το διάγραμμα ότι η παραμένουσα τάση έχει σημαντική επίδραση στην φέρουσα ικανότητα του δοκιμίου μετά τη διαρροή, ενώ η επίδραση της κλίσης της καμπύλης αποφόρτισης του νόμου του υλικού είναι σχετικά μικρή.



Σχήμα 1.11: Συμπεριφορά δοκιμίου από σκυρόδεμα υπό ανακυκλιζόμενη φόρτιση

1.4 Καταστατικοί νόμοι εδάφους

Στα εδαφικά υλικά οι καταστατικοί νόμοι θα έπρεπε να προκύψουν από τη μικροσκοπική θεώρηση των μηχανισμών παραμόρφωσης (ελαστικές και ανελαστικές παραμορφώσεις κόκκων, ολίσθηση και κύλιση μεταξύ των κόκκων κλπ.) και από τη μελέτη του τρόπου μετάδοσης των δυνάμεων μεταξύ των κόκκων. Η μέθοδος αυτή όμως, υπερβαίνει τις σημερινές υπολογιστικές δυνατότητες (εκτός ίσως από ορισμένες πολύ απλές περιπτώσεις γεωμετρίας κόκκων) λόγω κυρίως του τεράστιου αριθμού επαφών μεταξύ των κόκκων: π.χ. σε ένα κυβικό εκατοστό λεπτόκοκκης άμμου υπάρχουν περί τα πέντε εκατομμύρια σημεία επαφής μεταξύ κόκκων, ενώ στην περίπτωση των αργιλικών εδαφών ο αριθμός αυτός είναι πολύ μεγαλύτερος. Κατά συνέπεια, στην περίπτωση των εδαφικών υλικών οι σχέσεις τάσεων - παραμορφώσεων που συνήθως χρησιμοποιούνται είτε είναι όμοιες με αυτές που εφαρμόζονται σε άλλα τεχνικά υλικά (π.χ. ο νόμος της γραμμικής ελαστικότητας) είτε προκύπτουν από γενικεύσεις των αποτελεσμάτων πειραματικών μετρήσεων σε δείγματα εδαφικών υλικών (άμμων, αργίλων κλπ.).

Λόγω της πολυφασικότητας των εδαφικών υλικών και κυρίως λόγω της παρουσίας νερού στους πόρους το οποίο αλληλεπιδρά με τον εδαφικό σκελετό (όπως π.χ. κατά τη στερεοποίηση), οι καταστατικές σχέσεις των εδαφικών υλικών αναφέρονται στον εδαφικό σκελετό, δηλαδή συσχετίζουν τις παραμορφώσεις του εδαφικού σκελετού (που, όμως, ταυτίζονται με τις μακροσκοπικές παραμορφώσεις του εδάφους) με τις

ενεργές τάσεις, τις τάσεις δηλαδή που αναλαμβάνονται από τον εδαφικό σκελετό. Έτσι, στα επόμενα, όπου αναφέρονται "τάσεις", θα θεωρούνται οι ενεργές τάσεις. Τέλος, θα πρέπει να αναφερθεί ότι στις πρακτικές εφαρμογές ζητείται ο υπολογισμός των τάσεων (και των παραμορφώσεων) που προκαλούνται από κάποια εξωτερική φόρτιση (π.χ. από την κατασκευή ενός κτιρίου). Επειδή, όμως, τα εδάφη είναι συνήθως φορτισμένα και πριν από την επιβολή της συγκεκριμένης εξωτερικής φόρτισης (τυπικά υπόκεινται στις τάσεις λόγω του βάρους των υπερκείμενων εδαφικών σχηματισμών), στην πραγματικότητα ζητείται η σχέση μεταξύ της μεταβολής των ενεργών τάσεων και της αντίστοιχης μεταβολής των παραμορφώσεων που οφείλονται στην εξωτερική φόρτιση. Για τους λόγους αυτούς, οι σχέσεις τάσεων-παραμορφώσεων που περιγράφονται στα επόμενα εδάφια παρουσιάζονται με τη μορφή συσχετίσεων μεταξύ των μεταβολών των ενεργών τάσεων ($\Delta\sigma'_{ij}$) και των αντίστοιχων μεταβολών των συνιστωσών των παραμορφώσεων ($\Delta\varepsilon_{kl}$).

1.4.1 Γραμμική ισότροπη ελαστικότητα

Ο απλούστερος καταστατικός νόμος συμπεριφοράς των εδαφών περιγράφεται από τις σχέσεις της γραμμικής ισότροπης ελαστικότητας [7]. Κατά τη θεωρία αυτή, οι σχέσεις τάσεων-παραμορφώσεων εξαρτώνται από δύο θεμελιώδεις παραμέτρους (ιδιότητες του υλικού), το μέτρο Ελαστικότητας E και το λόγο του Poisson ν . Διάφορα άλλα παράγωγα μεγέθη προκύπτουν με συνδυασμό των ανωτέρω, όπως:

- ✓ Το μέτρο διάτμησης: $G = \frac{E}{2(1+\nu)}$
- ✓ Το μέτρο συμπίεσης: $K = \frac{E}{3(1-2\nu)}$
- ✓ Το μέτρο μονοδιάστατης συμπίεσης: $D = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)}$

Οι καταστατικές σχέσεις τάσεων – παραμορφώσεων της γραμμικής ισότροπης ελαστικότητας έχουν την ακόλουθη μορφή:

$$\Delta\varepsilon_{xx} = \frac{1}{E}[\Delta\sigma'_{xx} - \nu(\Delta\sigma'_{yy} + \Delta\sigma'_{zz})] \quad (1.19)$$

$$\Delta\varepsilon_{yy} = \frac{1}{E}[\Delta\sigma'_{yy} - \nu(\Delta\sigma'_{xx} + \Delta\sigma'_{zz})] \quad (1.20)$$

$$\Delta \varepsilon_{zz} = \frac{1}{E} [\Delta \sigma'_{zz} - \nu (\Delta \sigma'_{yy} + \Delta \sigma'_{xx})] \quad (1.21)$$

$$\Delta \gamma_{xy} = \frac{1}{G} \Delta \sigma'_{xy} \quad (1.22)$$

$$\Delta \gamma_{yz} = \frac{1}{G} \Delta \sigma'_{yz} \quad (1.23)$$

$$\Delta \gamma_{zx} = \frac{1}{G} \Delta \sigma'_{zx} \quad (1.24)$$

Οι συνέπειες των ανωτέρω σχέσεων αναφέρονται στα επόμενα:

(1) Με άθροιση κατά μέλη των τριών πρώτων εξισώσεων προκύπτει:

$$\Delta \varepsilon_{vol} = \frac{1}{K} \frac{\Delta \sigma'_{xx} + \Delta \sigma'_{yy} + \Delta \sigma'_{zz}}{3} \quad (1.25)$$

Η σχέση αυτή έχει σαν συνέπεια, ότι στην περίπτωση της αστράγγιστης φόρτισης κορεσμένων εδαφικών υλικών (δηλαδή στην περίπτωση που το νερό των πόρων δεν μπορεί να διαφύγει), οπότε $\Delta \varepsilon_{vol} = 0$, η μέση ορθή ενεργός τάση δεν μεταβάλλεται, δηλαδή:

$$\Delta \sigma' \equiv \frac{1}{3} (\Delta \sigma'_{xx} + \Delta \sigma'_{yy} + \Delta \sigma'_{zz}) = 0 \quad (1.26)$$

οπότε, κατά τον ορισμό των ενεργών τάσεων:

$$\Delta \sigma' \equiv \frac{1}{3} (\Delta \sigma_{xx} + \Delta \sigma_{yy} + \Delta \sigma_{zz}) = \Delta u \quad (1.27)$$

δηλαδή η αύξηση της πίεσης πόρων ισούται με τη μέση αύξηση των ορθών ολικών τάσεων.

(2) Οι σχέσεις της γραμμικής ισότροπης ελαστικότητας υποδηλώνουν την έλλειψη συσχέτισης μεταξύ των ορθών τάσεων και των διατμητικών παραμορφώσεων, καθώς και μεταξύ των διατμητικών τάσεων και των ορθών παραμορφώσεων, επειδή οι εξισώσεις (1.19-1.24) συσχετίζουν τις ορθές τάσεις μόνον με τις ορθές παραμορφώσεις και τις διατμητικές τάσεις μόνον με τις διατμητικές παραμορφώσεις. Κατά συνέπεια, η γραμμική ισότροπη ελαστικότητα προβλέπει, ότι στη δοκιμή της απλής διάτμησης, κατά την οποία αυξάνει βαθμιαία η διατμητική τάση που επιβάλλεται στο δοκίμιο χωρίς να

μεταβάλλονται οι ορθές τάσεις, οι ορθές συνιστώσες της παραμόρφωσης είναι μηδέν, οπότε δεν μεταβάλλεται ούτε ο όγκος του δοκιμίου.

Η μηχανική συμπεριφορά των εδαφικών υλικών απέχει σημαντικά από τις προβλέψεις της γραμμικής ισότροπης ελαστικότητας. Έτσι, π.χ. η καμπύλη τάσεων - παραμορφώσεων των εδαφικών υλικών δεν είναι ευθύγραμμη (δηλαδή το μέτρο ελαστικότητας δεν είναι σταθερό). Επιπλέον, η επιβολή μόνον διατμητικών τάσεων προκαλεί ορθές παραμορφώσεις (π.χ. μεταβολή του όγκου) στις στραγγιζόμενες δοκιμές και μεταβολή της πίεσης πόρων στις αστράγγιστες δοκιμές. Παρά αυτά, η υπόθεση της ισότροπης γραμμικής ελαστικότητας χρησιμοποιείται ευρύτατα στα εδαφικά υλικά. Η ευρύτατη χρήση της γραμμικής ισότροπης ελαστικότητας, αν και εκ πρώτης όψεως φαίνεται αδικαιολόγητη, οφείλεται στους εξής λόγους:

- ✓ Με την υπόθεση της γραμμικής ισότροπης ελαστικότητας έχουν επιλυθεί πολλά συνήθη προβλήματα της Εδαφομηχανικής, όπως π.χ. η κατανομή των τάσεων στο έδαφος κάτω από κυκλικό θεμέλιο, λωριδωτό θεμέλιο κλπ, για τα οποία δεν υπάρχει αναλυτική λύση με άλλους πιο πολύπλοκους καταστατικούς νόμους.
- ✓ Στα γραμμικώς ελαστικά υλικά ισχύει η αρχή της επαλληλίας, η οποία χρησιμοποιείται ευρύτατα για την ανάλυση σύνθετων φορτίσεων σε άλλες απλούστερες, για τις οποίες υπάρχει αναλυτική λύση. Έτσι, π.χ. η κατανομή των τάσεων στο έδαφος από το σύνολο των μεμονωμένων πεδίων ενός κτιρίου μπορεί να θεωρηθεί σαν άθροισμα των τάσεων που οφείλονται σε ένα έκαστο των πεδίων, περίπτωση για την οποία υπάρχει αναλυτική λύση όταν το έδαφος θεωρηθεί ισότροπο και γραμμικώς ελαστικό.
- ✓ Με τη γραμμική ελαστικότητα αποφεύγονται θεωρητικά προβλήματα, όπως η ύπαρξη μοναδικής λύσης, η πολλαπλότητα λύσεων ανάλογα με την αλληλουχία εφαρμογής των φορτίων, κακώς ορισμένα προβλήματα, καθώς επίσης και τα προβλήματα σύγκλισης των αριθμητικών μεθόδων ανάλυσης.
- ✓ Σε πολλές περιπτώσεις συνήθων φορτίσεων η επίλυση με τη θεωρία ελαστικότητας δεν απέχει σημαντικά από επιλύσεις με "ακριβέστερους" καταστατικούς νόμους, τουλάχιστον όσον αφορά τις κατανομές των τάσεων στο έδαφος. Τούτο οφείλεται στο ότι οι κατανομές των τάσεων που προκύπτουν με βάση τη θεωρία ελαστικότητας είναι, για ορισμένους τύπους φορτίσεων, ανεξάρτητες του μέτρου ελαστικότητας του εδάφους, δηλαδή οι κατανομές τάσεων δεν μεταβάλλονται για διαφορετικές τιμές του μέτρου ελαστικότητας. Βεβαίως, δεν θα ισχύει η ίδια κατανομή τάσεων, εάν το έδαφος έχει μή-

γραμμική συμπεριφορά. Στις περιπτώσεις αυτές οι κατανομές τάσεων είναι ανεξάρτητες του μέτρου ελαστικότητας, υπό την προϋπόθεση ότι το υλικό είναι ισότροπο και γραμμικώς ελαστικό και ότι το μέτρο ελαστικότητας είναι σταθερό και ομοιόμορφο σε ολόκληρη την εδαφική μάζα, δηλαδή δεν μεταβάλλεται από θέση σε θέση ούτε με τη μεταβολή της φόρτισης. Πάντως, η παραπάνω ανεξαρτησία των τάσεων από το μέτρο ελαστικότητας μειώνει, γενικά, τα σφάλματα που προκύπτουν κατά την εκτίμηση των τάσεων σε εδαφικά υλικά με θεώρηση γραμμικής ελαστικότητας, ακόμη και στις περιπτώσεις που η συμπεριφορά του εδάφους απέχει σημαντικά από την ελαστικότητα.

- ✓ Τέλος, ένας ακόμη σημαντικός λόγος της ευρείας χρήσης της θεωρίας ελαστικότητας στα εδαφικά υλικά είναι ότι μέχρι σήμερα δεν έχει επιτευχθεί η διατύπωση ενός άλλου καταστατικού νόμου, ο οποίος αποδεδειγμένα να δίνει ακριβέστερες προβλέψεις σε ένα ευρύ φάσμα προβλημάτων της Εδαφομηχανικής και για μια ποικιλία εδαφικών υλικών, παρά τις έντονες προσπάθειες των ερευνητών στον τομέα αυτό.

1.4.1.1 Η έκφραση των σχέσεων ελαστικότητας ως προς τις ολικές τάσεις

Αναφέρθηκε προηγουμένως ότι οι σχέσεις τάσεων-παραμορφώσεων ενός εδαφικού υλικού περιγράφουν τη συμπεριφορά του εδαφικού σκελετού και κατά συνέπεια συσχετίζουν τις παραμορφώσεις ($\Delta \varepsilon_{ij}$) με τις αντίστοιχες μεταβολές των ενεργών τάσεων ($\Delta \sigma'_{ij}$).

Στην περίπτωση της πλήρως στραγγιζόμενης φόρτισης, στην περίπτωση δηλαδή που η φόρτιση είναι επαρκώς αργή, ώστε οι τυχόν αναπτυσσόμενες υπερπιέσεις όρων να εκτονώνονται πλήρως ($\Delta u = 0$), το σύνολο των εξωτερικώς επιβεβλημένων φορτίων αναλαμβάνεται από τις ενεργές τάσεις, οπότε ισχύει:

$$\Delta \sigma'_{ij} = \Delta \sigma_{ij} \quad (1.28)$$

δηλαδή οι μεταβολές των ολικών και των ενεργών τάσεων είναι ίσες. Κατά συνέπεια, μπορεί να θεωρηθεί ότι οι γνωστές σχέσεις ελαστικότητας (δηλαδή με την ίδια ακριβώς μορφή) συσχετίζουν τις παραμορφώσεις με τις μεταβολές των ολικών τάσεων.

Στην περίπτωση της αστραγγιστής φόρτισης, στην περίπτωση δηλαδή που η φόρτιση είναι πολύ ταχεία ώστε το νερό των πόρων να μην προλαβαίνει να διαφύγει, το έδαφος παραμορφώνεται διατηρώντας σταθερό όγκο, ενώ ταυτόχρονα αναπτύσσονται υπερπιέσεις πόρων. Στην περίπτωση αυτή οι σχέσεις τάσεων - παραμορφώσεων δεν

μπορούν να εκφραστούν απευθείας ως προς τις μεταβολές των ολικών τάσεων. Στα επόμενα περιγράφονται οι αναγκαίοι μετασχηματισμοί, ώστε οι γνωστές σχέσεις τάσεων-παραμορφώσεων της ισότροπης γραμμικής ελαστικότητας να εκφραστούν συναρτήσει των μεταβολών των ολικών τάσεων.

Κατά την αστράγγιστη φόρτιση ενός ισότροπου ελαστικού εδάφους, η συνθήκη της μη - μεταβολής του όγκου οδηγεί στη σχέση (βλ. εξίσωση 1.26):

$$\Delta\sigma'_{xx} + \Delta\sigma'_{yy} + \Delta\sigma'_{zz} = 0 \quad (1.29)$$

ή ισοδύναμα:

$$\Delta u = \frac{1}{3}(\Delta\sigma_{xx} + \Delta\sigma_{yy} + \Delta\sigma_{zz}) \quad (1.30)$$

Εάν, στη συνέχεια, στην (τυπική) σχέση της ισότροπης ελαστικότητας που έχει τη μορφή:

$$\Delta\varepsilon_{xx} = \frac{1}{E}[\Delta\sigma'_{xx} - \nu(\Delta\sigma'_{yy} + \Delta\sigma'_{zz})] \quad (1.31)$$

αντικατασταθούν οι μεταβολές των ενεργών τάσεων ($\Delta\sigma'_{ii}$) από τις σχέσεις:

$$\Delta\sigma'_{ii} = \Delta\sigma_{ii} - \Delta u \quad (1.32)$$

προκύπτει η σχέση:

$$\Delta\varepsilon_{xx} = \frac{1}{E}[\Delta\sigma_{xx} - \nu(\Delta\sigma_{yy} + \Delta\sigma_{zz}) - (1-2\nu)\Delta u] \quad (1.33)$$

Και με χρήση της σχέσης (1.29) προκύπτει ότι:

$$\Delta\varepsilon_{xx} = \frac{2(1+\nu)}{3E}[\Delta\sigma_{xx} - \frac{1}{2}(\Delta\sigma_{yy} + \Delta\sigma_{zz})] \quad (1.34)$$

ή ισοδύναμα:

$$\Delta\varepsilon_{xx} = \frac{1}{E_u}[\Delta\sigma_{xx} - \nu_u(\Delta\sigma_{yy} + \Delta\sigma_{zz})] \quad (1.35)$$

όπου:

$$E_u \equiv \frac{3E}{2(1+\nu)} \quad (1.36)$$

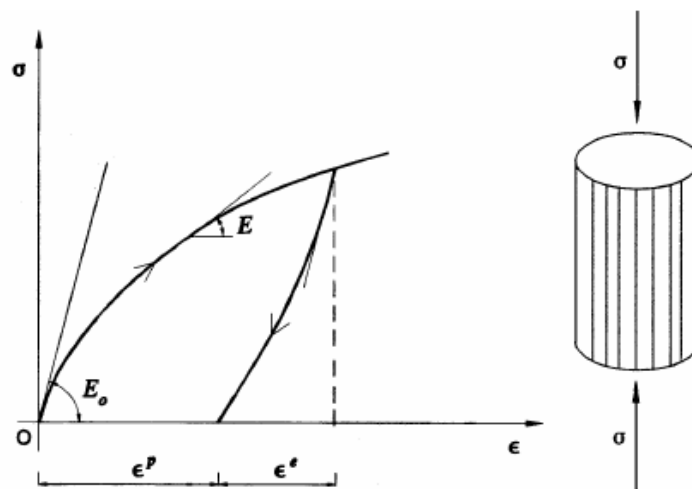
$$\nu_u \equiv \frac{1}{2}$$

και E , ν είναι οι παράμετροι ελαστικότητας του εδαφικού σκελετού. Κατά συνέπεια, στην περίπτωση της αστράγγιστης φόρτισης οι γνωστές σχέσεις ελαστικότητας μπορούν να γραφούν και ως προς τις μεταβολές των ολικών τάσεων (σχέσεις 1.35) με την ίδια μορφή που γράφονται ως προς τις μεταβολές των ενεργών τάσεων (σχέσεις 1.31), αρκεί να αντικατασταθεί το E με το E_u και το ν με το ν_u , όπου τα E_u και ν_u δίνονται από τις σχέσεις (1.36). Για μία τυπική τιμή του λόγου Poisson του εδαφικού σκελετού $\nu = 0.25$, η σχέση (1.36) δίνει: $E_u = 1.20E$, δηλαδή το «αστράγγιστο μέτρο ελαστικότητας» είναι 20% περίπου μεγαλύτερο από το μέτρο ελαστικότητας του εδαφικού σκελετού.

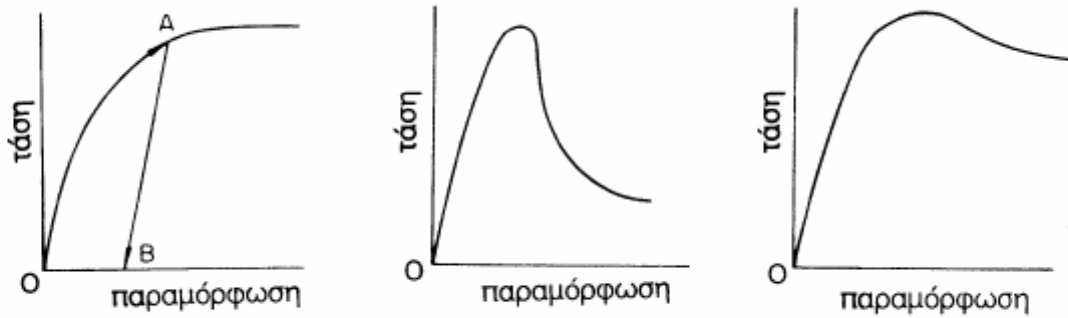
Σε ενδιάμεσες καταστάσεις μεταξύ της πλήρως στραγγιζόμενης και της αστράγγιστης φόρτισης, οι σχέσεις τάσεων-παραμορφώσεων μπορούν να γραφούν με τη μορφή των σχέσεων (1.33), χωρίς όμως να είναι δυνατόν να απαλειφθεί η υπερπίεση πόρων Δu .

1.4.2 Μη γραμμικές σχέσεις τάσεων - παραμορφώσεων

Εκτεταμένες μετρήσεις της μηχανικής συμπεριφοράς εδαφικών σχηματισμών στο εργαστήριο, αλλά και σε φυσική κλίμακα, αποδεικνύουν ότι η καμπύλη τάσεων - παραμορφώσεων των εδαφικών υλικών είναι έντονα μη - γραμμική. Επιπλέον, ένα σημαντικό τμήμα της παραμόρφωσης είναι ανελαστικό, δηλαδή δεν αναιρείται με την αφαίρεση του φορτίου (σχήμα 1.12). Τυπικές μορφές των καμπύλων τάσεων - παραμορφώσεων εδαφικών υλικών φαίνονται στο σχήμα (1.13) και δείχνουν την πολυπλοκότητα της μηχανικής συμπεριφοράς των υλικών αυτών [8].



Σχήμα 1.12: Ελαστικές και ανελαστικές παραμορφώσεις εδαφών



Σχήμα 1.13: Τυπικές μορφές καμπυλών τάσεων – παραμορφώσεων

Η πολυπλοκότητα της συμπεριφοράς των εδαφών επιτείνεται και από τα ακόλουθα χαρακτηριστικά τους:

- ✓ Η συμπεριφορά των κοκκωδών εδαφών είναι διαφορετική από τη συμπεριφορά των αργιλικών εδαφών. Αλλά, ακόμη και η συμπεριφορά ενός συγκεκριμένου τύπου εδάφους ποικίλει ανάλογα με την κατάσταση στην οποία ευρίσκεται. Έτσι, μια άμμος με χαλαρή δομή έχει διαφορετική συμπεριφορά από την ίδια άμμο σε πυκνή δομή και το ίδιο ισχύει για τη συμπεριφορά μαλακών και σιφρών αργίλων.
- ✓ Η συμπεριφορά των εδαφικών υλικών εξαρτάται από την προϊστορία των φορτίσεων, δηλαδή από την αλληλουχία των εντατικών καταστάσεων που έχουν προηγηθεί της συγκεκριμένης φόρτισης.
- ✓ Η συμπεριφορά των εδαφών εξαρτάται από τον τρόπο μεταβολής των συνιστωσών των τάσεων, δηλαδή παρουσιάζει έντονη ανισοτροπία. Έτσι, π.χ. εάν δύο όμοια εδαφικά δοκίμια (που συνίστανται από το ίδιο υλικό και έχουν υποστεί την ίδια προϊστορία φορτίσεων) υποβληθούν το μεν πρώτο στην ακόλουθη μεταβολή των τάσεων:

$$\Delta\sigma'_{xx} = a \quad , \quad \Delta\sigma'_{yy} = \Delta\sigma'_{zz} = \Delta\sigma'_{xz} = \Delta\sigma'_{yz} = \Delta\sigma'_{xy} = 0 \quad (1.37)$$

ενώ το δεύτερο σε:

$$\Delta\sigma'_{yy} = a \quad , \quad \Delta\sigma'_{xx} = \Delta\sigma'_{zz} = \Delta\sigma'_{xz} = \Delta\sigma'_{yz} = \Delta\sigma'_{xy} = 0 \quad (1.38)$$

η συμπεριφορά τους θα είναι, γενικά, διαφορετική. Η συγκεκριμένη αλληλουχία μεταβολών των τάσεων στην οποία υποβάλλεται ένα εδαφικό στοιχείο ονομάζεται διαδρομή των τάσεων και μελετάται στο επόμενο εδάφιο.

- ✓ Η συμπεριφορά των εδαφικών υλικών επηρεάζεται από την ύπαρξη δομικών δεσμών μεταξύ των κόκκων τους, λόγω π.χ. σιμέντωσης ή θιξοτροπίας. Τέτοιοι δεσμοί μπορούν να αναπτυχθούν στο μακρόχρονο διάστημα που έχει μεσολαβήσει από τη γένεσή τους μέχρι σήμερα. Συνέπεια της ανάπτυξης δεσμών μεταξύ των κόκκων των εδαφικών υλικών είναι ότι το ίδιο υλικό σε αδιατάρακτη μορφή (δηλαδή όπως συναντάται στην φύση) παρουσιάζει τελείως διαφορετική συμπεριφορά απ' ό,τι μετά από αναμόχλευση και καταστροφή των δεσμών, ακόμη και εάν μετά την αναμόχλευση υποβληθεί στην προϊστορία φορτίσεων του αδιατάρακτου υλικού.
- ✓ Τέλος, η συμπεριφορά των περισσότερων εδαφικών υλικών εξαρτάται από την ταχύτητα επιβολής της φόρτισης και παρουσιάζει χαρακτηριστικά ερπυσμού (δηλαδή παραμόρφωση υπό σταθερές ενεργές τάσεις) ανεξάρτητα από το φαινόμενο της στερεοποίησης.

Λόγω της πολυπλοκότητας της μηχανικής συμπεριφοράς των εδαφών, η επιτυχία γενικών καταστατικών νόμων στην ποσοτική πρόβλεψη της συμπεριφοράς τους είναι περιορισμένη. Έτσι, η μελέτη της συμπεριφοράς των εδαφών βασίζεται συχνά σε πειραματικές μετρήσεις που γίνονται στο εργαστήριο. Επειδή, όμως, η παραμόρφωση των εδαφών εξαρτάται από το μέγεθος και τον τρόπο της μεταβολής των τάσεων, θα πρέπει να επιβληθούν στο εργαστήριο οι ίδιες μεταβολές τάσεων που εκτιμάται ότι θα επιβληθούν και στη φύση κατά τη φόρτιση του εδάφους. Η παραπάνω απαίτηση εισάγει πρόσθετες (και πολλές φορές ανυπέβλητες) δυσκολίες στην ήδη πολύπλοκη μελέτη της μηχανικής συμπεριφοράς των εδαφών. Ειδικότερα:

- ✓ Οι μεταβολές των τάσεων που θα επιβληθούν στη φύση, γενικά, δεν είναι γνωστές. Πράγματι, ενώ τα εξωτερικώς επιβεβλημένα φορτία (π.χ. το φορτίο ενός υποστηλώματος) είναι συνήθως γνωστά, οι τάσεις που τα φορτία αυτά προκαλούν σε ένα εδαφικό στοιχείο στο εσωτερικό του εδάφους δεν είναι επακριβώς γνωστές. Ακόμη, όμως, και στην περίπτωση που οι αναπτυσσόμενες τάσεις είναι γνωστές, τα εδαφικά στοιχεία είναι άπειρα και αναγκαστικά θα πρέπει να γίνει επιλογή των εδαφικών στοιχείων που θα δοκιμασθούν στο εργαστήριο.
- ✓ Οι δυνατότητες επιβολής τάσεων στις εργαστηριακές διατάξεις είναι περιορισμένες. Έτσι, είναι πρακτικά αδύνατον να επιβληθούν στο εργαστήριο οι γενικές τριαξονικές εντατικές καταστάσεις που θα συμβούν στη φύση και συνεπώς θα πρέπει να απλουστευθούν ώστε να μπορούν να επιβληθούν από τις διαθέσιμες εργαστηριακές διατάξεις.

- ✓ Αν στα παραπάνω προστεθούν οι δυσκολίες λόγω της ανομοιομορφίας του εδάφους, της διατάραξης των δειγμάτων κατά τη δειγματοληψία και της αδυναμίας, πολλές φορές, λήψης δειγμάτων από τις θέσεις που είναι επιθυμητό, το πρόβλημα της μελέτης των μηχανικών ιδιοτήτων του εδάφους καθίσταται ιδιαίτερα πολύπλοκο.

Έτσι, κατ' ανάγκη, ο Γεωτεχνικός Μηχανικός προβαίνει σε παραδοχές, επιλογές και απλοποιήσεις, η επιτυχία των οποίων καθορίζει και την επιτυχία του Γεωτεχνικού σχεδιασμού του έργου. Οι επιτυχίες αποφάσεις στα παραπάνω προβλήματα απαιτούν γνώση των μηχανισμών που διέπουν τις συνθήκες φόρτισης, κατανόηση των αρχών της μηχανικής συμπεριφοράς των εδαφών, γνώση των αναλυτικών και υπολογιστικών δυνατοτήτων αλλά και σημαντική εμπειρία σε παρόμοια έργα και παρόμοια εδάφη. Στο τελευταίο, ίσως, οφείλεται συχνά και ο χαρακτηρισμός της Εδαφομηχανικής σαν Τέχνης μάλλον παρά σαν Επιστήμης. Από όσα αναφέρθηκαν στα προηγούμενα αλλά και από όσα ακολουθήσουν, θα καταστεί σαφές ότι στην πραγματικότητα πρόκειται για ένα συνδυασμό Επιστήμης και Τέχνης.

1.4.3 Η έννοια της διαδρομής των τάσεων

Αναφέρθηκε προηγουμένως ότι η μηχανική συμπεριφορά των εδαφών εξαρτάται από την αλληλουχία των τάσεων που τους επιβάλλονται. Η διαδρομή των τάσεων είναι ένας γραφικός τρόπος για την περιγραφή της αλληλουχίας των επιβεβλημένων τάσεων. Ας θεωρηθεί ένα εδαφικό στοιχείο, στο οποίο επιβάλλονται ολικές τάσεις σ_{xx} , σ_{yy} , σ_{zz} , σ_{xy} , σ_{xz} , και σ_{yz} και έχει αναπτυχθεί πίεση πόρων u , οπότε οι αντίστοιχες ενεργές τάσεις ορίζουν την εντατική κατάσταση του στοιχείου, που είναι:

$$\begin{aligned} \sigma'_{xx} &= \sigma_{xx} - u, & \sigma'_{yy} &= \sigma_{yy} - u, & \sigma'_{zz} &= \sigma_{zz} - u \\ \sigma'_{xy} &= \sigma_{xy}, & \sigma'_{yz} &= \sigma_{yz} \end{aligned} \quad (1.39)$$

Οι ανωτέρω τάσεις μπορούν να παρασταθούν γραφικά σε έναν εξαδιάστατο χώρο (στον οποίο οι άξονες αντιστοιχούν στις έξι συνιστώσες των τάσεων) από ένα σημείο M που αντιστοιχεί στη διατεταγμένη εξάδα $(\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz})$ και ένα σημείο M' που αντιστοιχεί στη διατεταγμένη εξάδα $(\sigma'_{xx}, \sigma'_{yy}, \sigma'_{zz}, \sigma'_{xy}, \sigma'_{xz}, \sigma'_{yz})$. Το σημείο M παριστάνει την κατάσταση των ολικών τάσεων του εδαφικού στοιχείου ενώ το σημείο M' την αντίστοιχη κατάσταση των ενεργών τάσεων. Αν οι τάσεις που ασκούνται στο εδαφικό στοιχείο μεταβληθούν, τότε τα σημεία M και M' μετακινούνται διαγράφοντας μια τροχιά (διαδρομή) που περιγράφει την αλληλουχία των εντατικών καταστάσεων στις οποίες υποβάλλεται το εδαφικό στοιχείο μεταξύ της αρχικής και της τελικής εντατικής κατάστασης. Η διαδρομή που διαγράφεται από το σημείο M των ολικών τάσεων

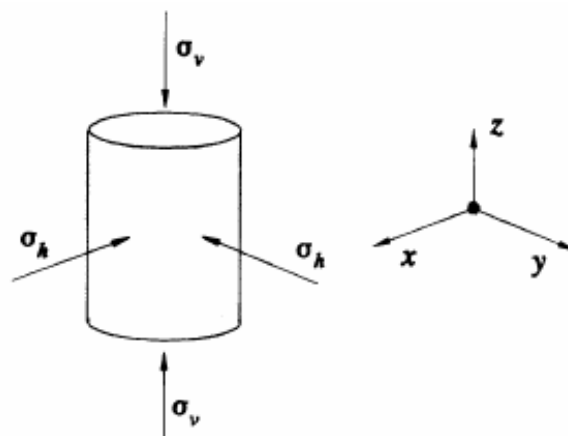
ονομάζεται διαδρομή ολικών τάσεων (ΔΟΤ), ενώ η διαδρομή που διαγράφεται από το σημείο M' των ενεργών τάσεων ονομάζεται διαδρομή ενεργών τάσεων (ΔΕΤ).

Σε ειδικές περιπτώσεις διαδρομών τάσεων όπου ορισμένες από τις συνιστώσες των τάσεων είναι μηδέν, οι διαδρομές των τάσεων μπορούν να παρασταθούν γραφικά σε χώρους με λιγότερες από έξι διαστάσεις. Μια τέτοια, αρκετά συνήθης, περίπτωση είναι η λεγόμενη "τριαξονική" ένταση (σχήμα 1.14), κατά την οποία το εδαφικό στοιχείο υποβάλλεται στην ακόλουθη αξονοσυμμετρική ένταση:

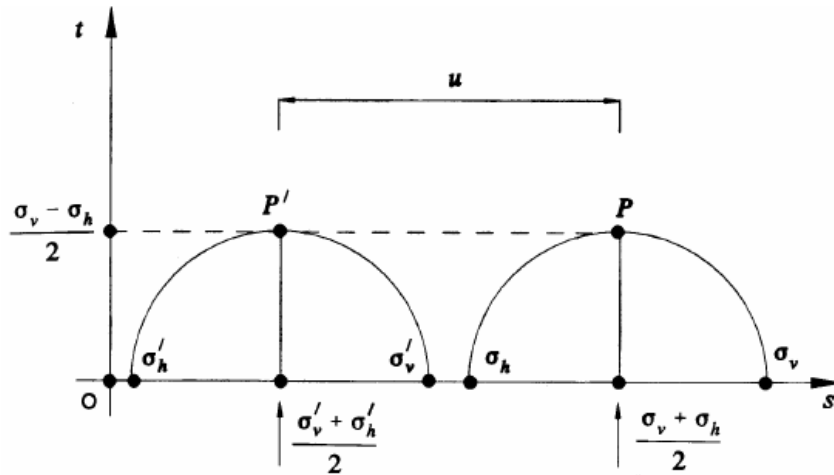
$$\begin{aligned}\sigma_{zz} &= \sigma_v \\ \sigma_{xx} &= \sigma_{yy} = \sigma_h \\ \sigma_{xy} &= \sigma_{zx} = \sigma_{yz} = 0\end{aligned}\tag{1.40}$$

οπότε οι ανεξάρτητες (μη - μηδενικές) συνιστώσες των τάσεων είναι δύο και συνεπώς οι διαδρομές τάσεων μπορούν να παρασταθούν σε ένα διδιάστατο χώρο (π.χ. σ_v , σ_h). Στην περίπτωση αυτή συχνά χρησιμοποιείται ένας διδιάστατος χώρος, ο οποίος ορίζεται από τα εντατικά μεγέθη:

$$\begin{aligned}s &= \frac{\sigma_v + \sigma_h}{2} \\ t &= \frac{\sigma_v - \sigma_h}{2}\end{aligned}\tag{1.41}$$

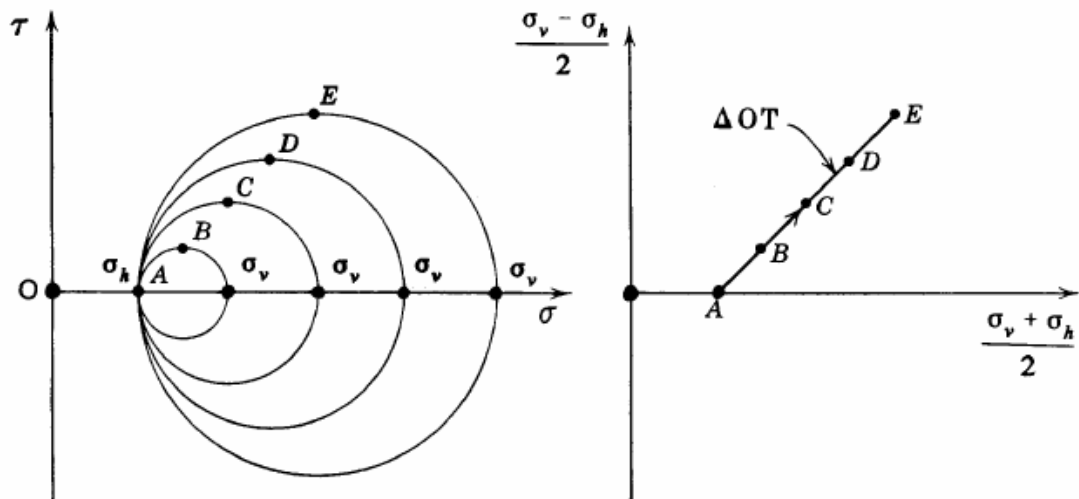


Σχήμα 1.14: Τριαξονική ένταση

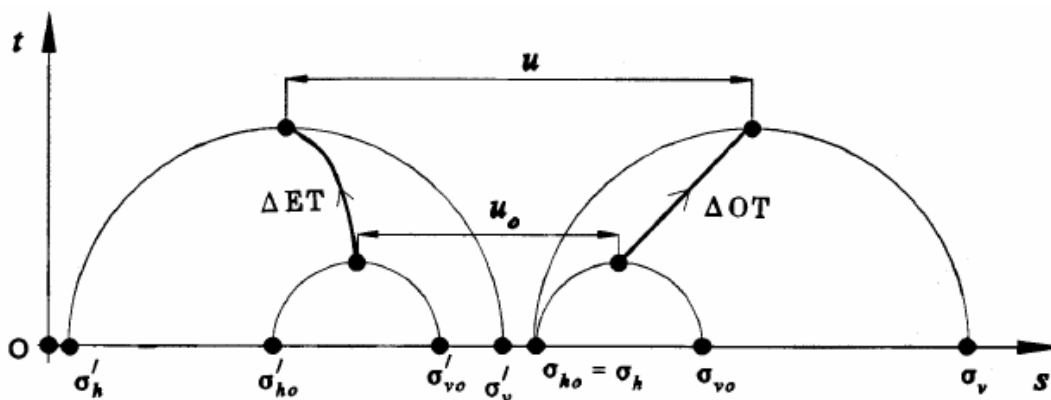


Σχήμα 1.15: Χώρος τάσεων s, t

Στο σχήμα (1.15) φαίνεται ο χώρος (s, t) και τα σημεία των ολικών τάσεων (P) και ενεργών τάσεων (P') για κάποια εντατική κατάσταση. Τα σημεία P και P' αποτελούν τις κορυφές των κύκλων Mohr που αντιστοιχούν στις ολικές και ενεργές τάσεις, βάσει του ορισμού των μεγεθών s και t . Στην ιδιότητα αυτή έγκειται και η χρήση του χώρου (s, t) αντί του (σ_v, σ_h) για τη γραφική παράσταση των διαδρομών τάσεων. Πράγματι, εάν είναι γνωστή η θέση του σημείου P , μπορεί εύκολα να σχεδιασθεί ο κύκλος Mohr των ολικών τάσεων που αντιστοιχεί στην κατάσταση αυτή (το ίδιο ισχύει και με το σημείο P'). Έτσι, η διαδρομή των ολικών τάσεων αντιστοιχεί στην τροχιά που διαγράφει η κορυφή του κύκλου Mohr των ολικών τάσεων, και η διαδρομή των ενεργών τάσεων αντιστοιχεί στην τροχιά που διαγράφει η κορυφή του κύκλου Mohr των ενεργών τάσεων. Είναι προφανές ότι οι κύκλοι των ολικών και ενεργών τάσεων έχουν ίση διάμετρο (επειδή $\sigma_v - \sigma_h = \sigma'_v - \sigma'_h$) και η (οριζόντια) απόστασή τους ισούται με την εκάστοτε πίεση πόρων, δηλαδή: $(P'P) = u$. Το σχήμα (1.16) παρουσιάζει την εξέλιξη των κύκλων Mohr σε μία "τριαξονική" εντατική κατάσταση κατά την οποία η κατακόρυφη τάση (σ_v) αυξάνει, ενώ η οριζόντια τάση (σ_h) παραμένει σταθερή. Στο ίδιο σχήμα φαίνεται και η αντίστοιχη διαδρομή των ολικών τάσεων (ΔOT). Τέλος, το σχήμα (1.17) παρουσιάζει μια τυπική διαδρομή ολικών τάσεων (ΔOT) και την αντίστοιχη διαδρομή ενεργών τάσεων (ΔET). Η απόσταση μεταξύ των δύο διαδρομών εκφράζει, ως γνωστόν, το εκάστοτε μέγεθος της πίεσης πόρων.



Σχήμα 1.16: Διαδρομές τάσεων



Σχήμα 1.17: Διαδρομές τάσεων

1.4.4 Μονοδιάστατη συμπίεση εδαφών

Μονοδιάστατη συμπίεση είναι η φόρτιση ενός εδαφικού στοιχείου κατά την οποία του επιβάλλεται ορθή παραμόρφωση στον κατακόρυφο άξονα (z) με ταυτόχρονη παρεμπόδιση των ορθών παραμορφώσεων στους άξονες (x, y) καθώς και όλων των διαμηθικών παραμορφώσεων, δηλαδή:

$$\varepsilon_{zz} \neq 0, \quad \varepsilon_{xx} = \varepsilon_{yy} = \gamma_{xy} = \gamma_{xz} = \gamma_{yz} = 0 \quad (1.42)$$

Η μονοδιάστατη συμπίεση είναι συνήθης τρόπος παραμόρφωσης των εδαφικών υλικών, επειδή π.χ. συμβαίνει κατά την ιζηματογένεση. Πράγματι, η απόθεση των εδαφικών υλικών στους πυθμένες των θαλασσών γίνεται σε οριζόντιες στρώσεις μεγάλης έκτασης, οπότε με τη συνεχιζόμενη απόθεση των υπερκείμενων ιζημάτων οι υποκείμενες εδαφικές αποθέσεις συμπιέζονται στην κατακόρυφη διεύθυνση. Επιπλέον, λόγω της μεγάλης έκτασης των αποθέσεων αυτών η παραμόρφωση στο οριζόντιο επίπεδο παρεμποδίζεται, με συνέπεια να πληρούνται οι προϋποθέσεις της

μονοδιάστατης συμπίεσης. Ο τύπος αυτός της παραμόρφωσης είναι αρκετά συνήθης και στα τεχνικά έργα, επειδή κατακόρυφες φορτίσεις μεγάλης έκτασης σε οριζόντιες εδαφικές στρώσεις οδηγούν σε μονοδιάστατη συμπίεση, όπως π.χ. επιχώματα οδοποιίας μεγάλου πλάτους, δεξαμενές υγρών καυσίμων μεγάλης διαμέτρου, κτίρια με μεγάλη επιφάνεια κάτοψης εδραζόμενα σε πλάκα θεμελίωσης (γενική κοιτόστρωση) αλλά ακόμη και θεμελιώσεις με μεμονωμένα πέδιλα σε εδαφικές στρώσεις μικρού πάχους (σε σχέση με τη διάσταση του πεδίου). Λόγω των εκτεταμένων πρακτικών εφαρμογών αλλά και της δυνατότητας επιβολής της στο εργαστήριο με απλές σχετικές διατάξεις, η μονοδιάστατη συμπίεση αποτελεί τη συνηθέστερη δοκιμή Εδαφομηχανικής.

Εάν θεωρηθεί ότι το έδαφος είναι ισότροπο και έχει γραμμική συμπεριφορά, η εφαρμογή των γνωστών σχέσεων ελαστικότητας οδηγεί στους ακόλουθους τύπους για τη μονοδιάστατη συμπίεση:

$$\begin{aligned}\Delta\sigma'_{zz} &= D \cdot \Delta\varepsilon_{zz} \\ \Delta\sigma'_{xx} &= \Delta\sigma'_{yy} = \frac{\nu}{1-\nu} \Delta\sigma'_{zz} \\ \Delta\sigma'_{xy} &= \Delta\sigma'_{yz} = \Delta\sigma'_{zx} = 0 \\ \Delta\varepsilon_{vol} &= \Delta\varepsilon_{zz} \\ \Delta\sigma' &= K \Delta\varepsilon_{zz}\end{aligned}\tag{1.43}$$

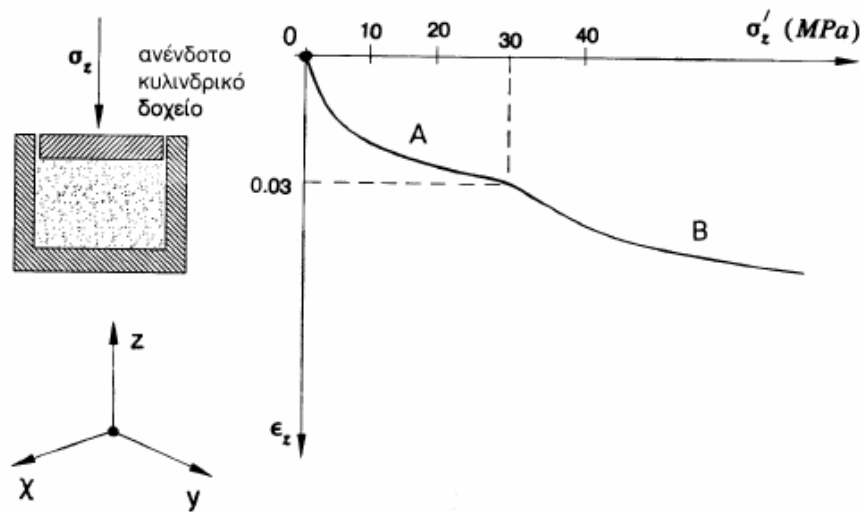
όπου D είναι το μέτρο μονοδιάστατης συμπίεσης και $\Delta\sigma'$ η μέση αύξηση των ορθών ενεργών τάσεων:

$$\Delta\sigma' \equiv \frac{1}{3} (\Delta\sigma'_{xx} + \Delta\sigma'_{yy} + \Delta\sigma'_{zz})$$

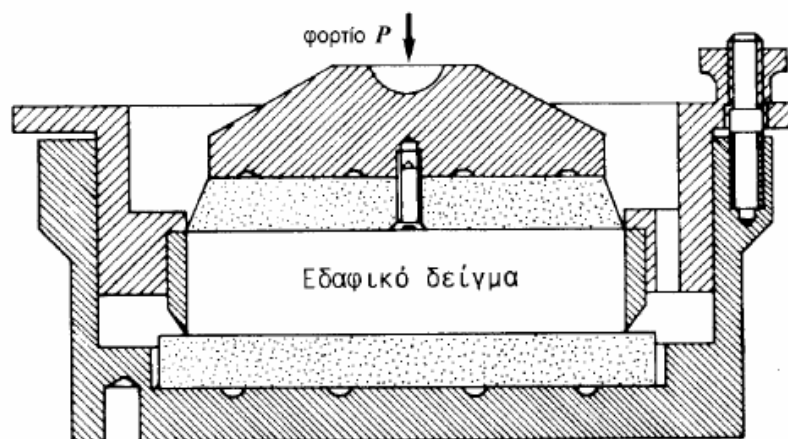
Από τις προηγούμενες σχέσεις προκύπτει ότι η γραμμική ελαστικότητα προβλέπει ευθύγραμμη σχέση μεταξύ της αξονικής τάσης (σ'_{zz}) και της αξονικής παραμόρφωσης (ε_{zz}).

Το σχήμα (1.18) παρουσιάζει τυπικά αποτελέσματα πειραματικών μετρήσεων της αξονικής ενεργού τάσης (σ'_{zz}) και της αντίστοιχης αξονικής παραμόρφωσης κατά τη μονοδιάστατη συμπίεση μιας ομοιόμορφης χονδρόκοκκης χαλαζιακής άμμου στην εργαστηριακή διάταξη που φαίνεται στο σχήμα (1.19). Η θετική διεύθυνση του άξονα της παραμόρφωσης είναι προς τα κάτω και αντιστοιχεί σε συμπίεση του εδάφους, δηλαδή σε μείωση του αρχικού ύψους του δοκιμίου. Επειδή η άμμος είναι ξηρή, οι ολικές τάσεις είναι ίσες με τις ενεργές τάσεις. Θα πρέπει, όμως, να αναφερθεί ότι η συμπεριφορά θα είναι η ίδια και στην περίπτωση που η άμμος είναι υγρή ή κορεσμένη

με νερό αλλά η φόρτιση αρκετά αργή, ώστε το νερό των πόρων να διαφεύγει ελεύθερα κατά τη συμπίεση της άμμου, χωρίς να αυξάνονται οι υδατικές πιέσεις (δηλαδή $u = 0$ κατά τη διάρκεια της φόρτισης).



Σχήμα 1.18: Μονοδιάστατη συμπίεση άμμου



Σχήμα 1.19: Εργαστηριακή διάταξη μονοαξονικής συμπίεσης

Η καμπύλη φόρτισης του σχήματος (1.18) δεν είναι ευθύγραμμη, όπως προβλέπει η γραμμική ελαστικότητα, αλλά κατά ποιοτική θεώρηση αποτελείται από τρεις κλάδους:

- ✓ Για τάσεις μέχρι περίπου 25 MPa, η καμπύλη στρέφει τα κοίλα προς τα άνω, δηλαδή η άμμος γίνεται προοδευτικά περισσότερο "δύσκαμπτη" (παρουσιάζει συμπεριφορά σκλήρυνσης ή κράτυνσης) με την αύξηση του κατακόρυφου φορτίου. Η συμπεριφορά αυτή είναι αναμενόμενη, επειδή με την αύξηση του φορτίου και την επακόλουθη συμπίεση η άμμος γίνεται πυκνότερη (το πορώδες μειώνεται) και συνεπώς λιγότερο παραμορφώσιμη. Πράγματι, όπως αναφέρθηκε στα προηγούμενα, η σχετική ολίσθηση και κύλιση μεταξύ των

κόκκων, στην οποία οφείλονται κυρίως οι παραμορφώσεις των αμμωδών εδαφών, είναι ευκολότερη στα χαλαρά απ' ότι στα πυκνά εδάφη.

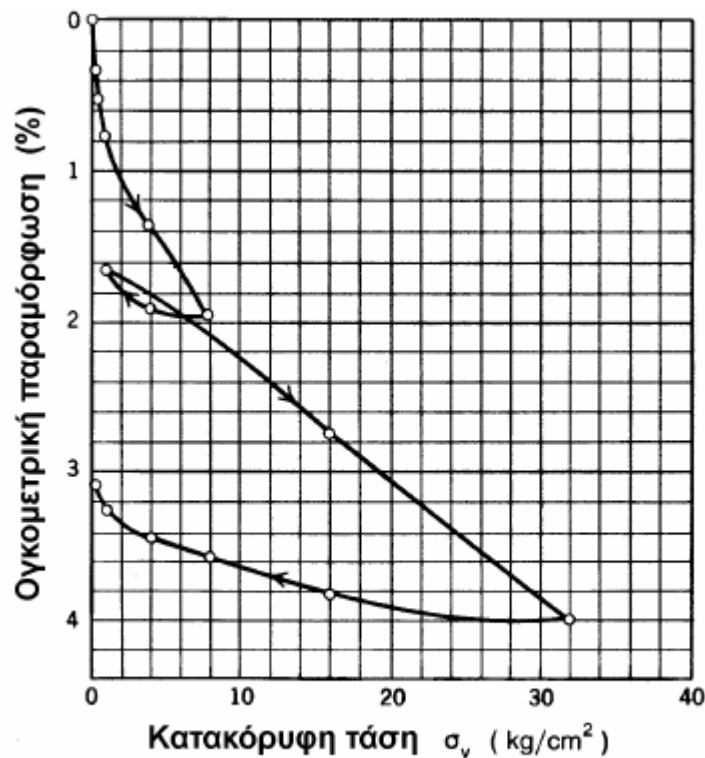
- ✓ Για τάσεις μεταξύ 25 και 35 MPa, η καμπύλη τάσεων – παραμορφώσεων παρουσιάζει ανάστροφη καμπυλότητα (στρέφει τα κοίλα προς τα κάτω), καθώς η άμμος προοδευτικά γίνεται πιο "εύκαμπτη". Το φαινόμενο αυτό, που ονομάζεται διαρροή, οφείλεται σε θρυμματισμό κόκκων και επιτρέπει νέες ολισθήσεις μεταξύ των κόκκων, δηλαδή συμπίεση της άμμου. Κατά τη φάση αυτή της συμπίεσης ακούγονται και χαρακτηριστικοί ήχοι που υποδηλώνουν τη θραύση κόκκων. Επιπλέον, κοκκομετρικές αναλύσεις πριν και μετά τη δοκιμή αποδεικνύουν ότι η διαβάθμιση της άμμου έχει μεταβληθεί με αύξηση του ποσοστού του λεπτόκοκκου κλάσματος.
- ✓ Για τάσεις μεγαλύτερες των 35 MPa, οι θραύσεις των κόκκων μειώνονται σημαντικά και εμφανίζεται πάλι συμπεριφορά κράτυνσης, με βαθμιαία αύξηση της "δυσκαμψίας" της άμμου και μείωση της συμπιεστότητας. Οι παραμορφώσεις οφείλονται και πάλι σε σχετικές ολισθήσεις μεταξύ των κόκκων, οι οποίες με τη μεταβολή της κοκκομετρικής διαβάθμισης λόγω των θραύσεων κόκκων αρχικά αυξήθηκαν αλλά προοδευτικά μειώνονται και πάλι.

Η διαδικασία που περιγράφηκε προηγουμένως λαμβάνει χώρα κατά τη μονοδιάστατη συμπίεση όλων των κοκκωδών εδαφών, ίσως χωρίς πλήρη διάκριση μεταξύ των τριών φάσεων, όπως στην ομοιόμορφη χονδρόκοκκη χαλαζιακή άμμο. Τα κύρια χαρακτηριστικά, όμως, της συμπεριφοράς, που είναι η βαθμιαία μείωση της συμπιεστότητας και η έναρξη του θρυμματισμού των κόκκων με την υπέρβαση κάποιας κρίσιμης τάσης, εμφανίζονται σε όλα τα κοκκώδη υλικά. Η κρίσιμη τάση, στην οποία εμφανίζεται θραύση κόκκων, είναι μικρότερη στις χονδρόκοκκες άμμους απ' ότι στις λεπτόκοκκες καθώς επίσης στις χαλαρές άμμους, σε άμμους με γωνιώδεις κόκκους, σε άμμους με ομοιόμορφους κόκκους και σε άμμους με κόκκους από υλικό μικρής αντοχής. Όλοι οι παραπάνω λόγοι (πλην του τελευταίου) συντελούν στη μείωση του αριθμού των επαφών μεταξύ των κόκκων (ανά μονάδα όγκου του υλικού), δηλαδή στην αύξηση της πραγματικής τάσης που ασκείται μεταξύ των κόκκων.

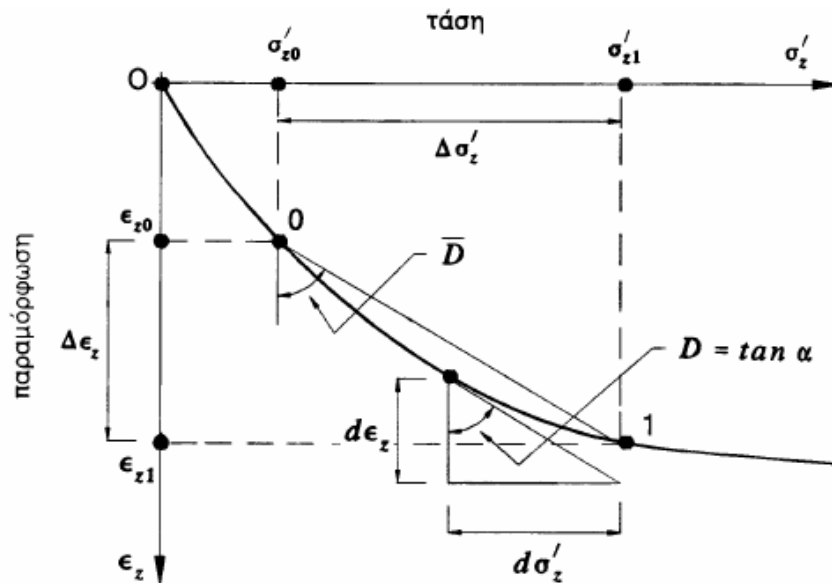
Στα περισσότερα πρακτικά προβλήματα οι τάσεις που ασκούνται μεταξύ των κόκκων είναι αρκετά μικρές, ώστε η συνεισφορά της θραύσης των κόκκων στην παραμόρφωση του εδάφους να είναι αμελητέα. Τάσεις μεγαλύτερες από την κρίσιμη εμφανίζονται σε περιπτώσεις υψηλών λιθόρριπτων φραγμάτων από εύθρυπτα υλικά και σε περιπτώσεις υποχωρήσεων του εδάφους λόγω σημαντικού και εκτεταμένου

καταβιβασμού της στάθμης των υπόγειων υδάτων (αντλήσεις για ύδρευση, άρδευση κλπ), η οποία προκαλεί αύξηση των ενεργών τάσεων.

Το σχήμα (1.20) παρουσιάζει την καμπύλη τάσεων-παραμορφώσεων μιας καλά διαβαθμισμένης άμμου κατά την αρχική συμπίεση, την επακόλουθη αποφόρτιση, την επαναφόρτιση σε μια τάση μεγαλύτερη της αρχικής και τελικά μια δεύτερη αποφόρτιση. Οι κλάδοι φόρτισης παρουσιάζουν βαθμιαία κράτυνση του υλικού με μείωση της συμπιεστότητας. Οι κλάδοι αποφόρτισης αναιρούν μικρό μόνο μέρος της συμπίεσης, δηλαδή η άμμος παρουσιάζει σημαντικές ανελαστικές (πλαστικές) παραμορφώσεις. Τούτο εξηγείται επειδή οι μηχανισμοί ολίσθησης και κύλισης μεταξύ των κόκκων, στους οποίους οφείλεται το μεγαλύτερο ποσοστό της παραμόρφωσης των κοκκωδών εδαφών, είναι μη-αντιστρεπτοί. Τέλος, οι καμπύλες τάσεων - παραμορφώσεων δείχνουν ότι ο κλάδος επαναφόρτισης του υλικού ακολουθεί πρακτικά την καμπύλη αποφόρτισης μέχρι τη μέγιστη τάση φόρτισης και κατόπιν συνεχίζει στην προέκταση του προηγούμενου κλάδου φόρτισης. Έτσι, κατά ποιοτική θεώρηση μπορεί να θεωρηθεί, ότι οι κλάδοι αποφόρτισης και επαναφόρτισης μέχρι την προηγούμενη μέγιστη τάση προκαλούν κυρίως ελαστικές παραμορφώσεις, ενώ ο κλάδος της αρχικής φόρτισης και οι κλάδοι επαναφόρτισης πέραν της μέγιστης τάσης προκαλούν τόσο ελαστικές όσο και (κυρίως) πλαστικές παραμορφώσεις.



Σχήμα 1.20: Καμπύλη τάσεων – παραμορφώσεων άμμου



Σχήμα 1.21: Ορισμός μέτρου συμπίεσης

Λόγω της έντονης καμπυλότητας του διαγράμματος τάσεων-παραμορφώσεων κατά τη μονοδιάστατη συμπίεση των κοκκωδών εδαφών, ο ορισμός του (σταθερού) μέτρου μονοδιάστατης συμπίεσης D πρέπει να αναθεωρηθεί. Το μέτρο D μπορεί, πλέον, να ορισθεί από την εφαπτόμενη της καμπύλης σε κάποια θέση (βλ. σχήμα 1.21) ως:

$$D \equiv \frac{d\sigma'_z}{d\varepsilon_z} \quad (1.44)$$

Το μέτρο D αυξάνει με την αύξηση της κατακόρυφης τάσης, καθώς η συμπιεστότητα του υλικού μειώνεται. Η έννοια του εφαπτομενικού (στιγμιαίου) μέτρου συμπίεσης έχει θεωρητική κυρίως χρησιμότητα, καθώς στις πρακτικές εφαρμογές ενδιαφέρει η συμπιεστότητα του εδαφικού υλικού μεταξύ της αρχικής κατακόρυφης τάσης (σ'_{z0}) και της τελικής ($\sigma'_{z1} = \sigma'_{z0} + \Delta\sigma'_z$). Αν οι παραπάνω τιμές των ενεργών τάσεων αντιστοιχούν σε παραμορφώσεις ε_{z0} και ε_{z1} , τότε η συμπιεστότητα του υλικού εκφράζεται από την πρόσθετη παραμόρφωση $\Delta\varepsilon_z = \varepsilon_{z1} - \varepsilon_{z0}$, δηλαδή από το μέσο μέτρο μονοδιάστατης συμπίεσης:

$$\bar{D} \equiv \frac{\Delta\sigma'_z}{\Delta\varepsilon_z} \quad (1.45)$$

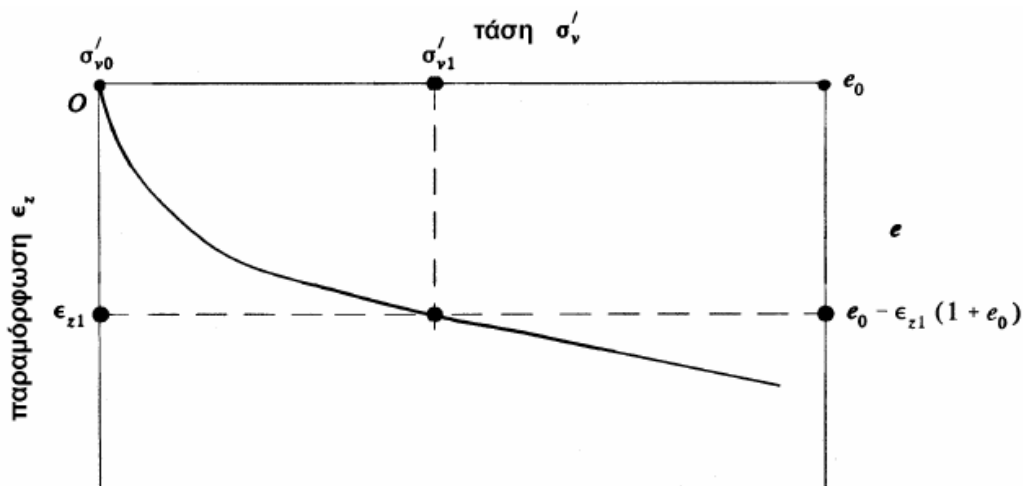
Με βάση τον ορισμό του μέσου μέτρου μονοδιάστατης συμπίεσης, η υποχώρηση δ της επιφάνειας ενός εδαφικού στρώματος πάχους H , λόγω της μεταβολής της κατακόρυφης τάσης κατά $\Delta\sigma'_z$, δίνεται από τη σχέση:

$$\delta = \Delta \varepsilon_z \cdot H = H \cdot \frac{\Delta \sigma'_z}{D} \quad (1.46)$$

Προφανώς, επειδή η τιμή του D δεν εξαρτάται μόνον από τη μεταβολή $\Delta \sigma'_z$ της κατακόρυφης τάσης αλλά και από την αρχική της τιμή σ'_{z0} , η υποχώρηση εξαρτάται και από την αρχική τάση. Κατά συνέπεια, η υποχώρηση ενός εδαφικού στρώματος είναι μικρότερη όταν η κατακόρυφη τάση αυξηθεί από 300 kPa σε 400 kPa ($\Delta \sigma'_z = 100$ kPa), παρά όταν αυξηθεί από 150 kPa σε 250 kPa ($\Delta \sigma'_z = 100$ kPa και πάλι).

Τα αποτελέσματα της δοκιμής μονοδιάστατης συμπίεσης παρουσιάζονται συχνά σε διάγραμμα που εμφανίζει το δείκτη πόρων (e) αντί της αξονικής παραμόρφωσης (βλ. σχήμα 1.22). Αποδεικνύεται ότι:

$$\Delta \varepsilon_z = \Delta \varepsilon_{vol} = -\frac{\Delta e}{1 + e_0} \quad (1.47)$$



Σχήμα 1.22: Καμπύλη τάσεων - παραμορφώσεων

Η μονοδιάστατη συμπίεση των συνεκτικών (αργιλικών) εδαφών είναι ποιοτικώς όμοια με των κοκκωδών εδαφών. Οι κυριότερες διαφορές συμπεριφοράς των δύο τύπων εδαφών κατά τη δοκιμή της μονοδιάστατης συμπίεσης είναι:

- ✓ Στα αργιλικά εδάφη δεν εμφανίζεται συνήθως το φαινόμενο της θραύσης των κόκκων, δηλαδή απουσιάζει ο Β' κλάδος της διαρροής, επειδή ο μηχανισμός παραμόρφωσης των αργιλικών εδαφών είναι διαφορετικός από αυτόν των κοκκωδών εδαφών: οι παραμορφώσεις των αργιλικών εδαφών οφείλονται κυρίως σε καμπτικές παραμορφώσεις των πλακιδίων, μεταβολή του πάχους των διπλών στρώσεων και σχετικές ολισθήσεις μεταξύ των διπλών στρώσεων.

Κατά συνέπεια, η καμπύλη τάσεων - παραμορφώσεων των αργιλικών εδαφών στρέφει διαρκώς τα κοίλα προς τα άνω με συνεχώς μειούμενη συμπιεστότητα.

- ✓ Οι άργιλοι, γενικά, έχουν μεγαλύτερο ποσοστό κενών απ' ότι οι άμμοι και κατά συνέπεια είναι περισσότερο συμπιεστές. Επιπλέον, οι άργιλοι παρουσιάζουν μεγάλες διαφορές συμπιεστότητας για διαφορετικές τιμές της φυσικής υγρασίας. Τυπικά, μία άργιλος παρουσιάζει δεκαπλάσια (ή και μεγαλύτερη) συμπιεστότητα στο όριο υδαρότητας απ' ότι στο όριο πλαστικότητας.
- ✓ Τέλος, στις αργίλους η συμπίεση απαιτεί χρόνο για να συντελεσθεί (λόγω του φαινομένου της στερεοποίησης), ενώ στις άμμους συμβαίνει πρακτικά ακαριαία. Βεβαίως, η διαφορά αυτή δεν οφείλεται σε διαφορετικά χαρακτηριστικά συμπιεστότητας του εδαφικού σκελετού αλλά στο υδρο-μηχανικό φαινόμενο της στερεοποίησης.

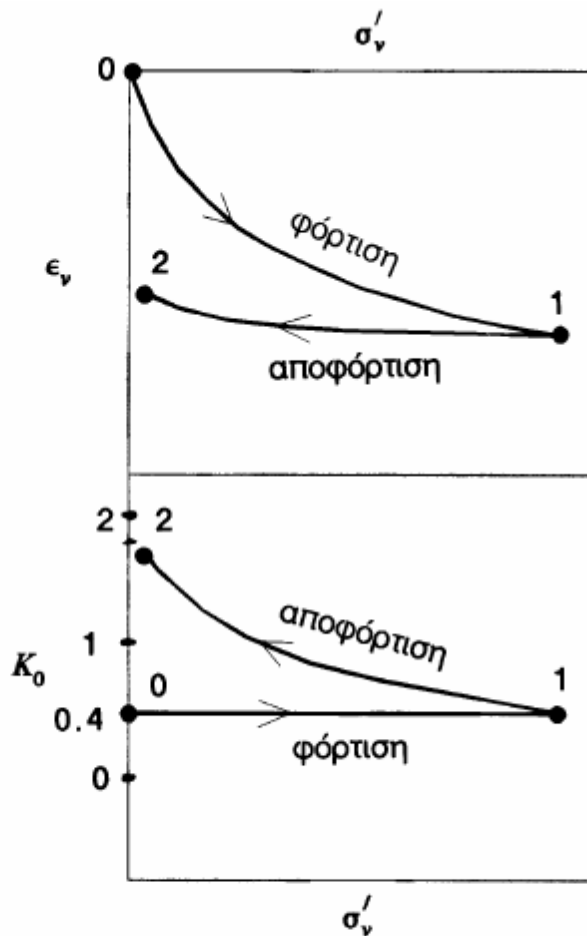
Μέχρι στιγμής, κατά τη μελέτη του φαινομένου της μονοδιάστατης στερεοποίησης των εδαφικών υλικών, εξετάσθηκε η κατακόρυφη (κύρια) τάση και η σχέση της με την αντίστοιχη κατακόρυφη παραμόρφωση. Η παραμόρφωση αυτή είναι η μόνη μη - μηδενική συνιστώσα της παραμόρφωσης. Όσον αφορά όμως τις τάσεις, εκτός από την κατακόρυφη τάση, αναπτύσσονται και οριζόντιες τάσεις στα παράπλευρα επίπεδα, όπου η παραμόρφωση παρεμποδίζεται από τα ανένδοτα τοιχώματα του δοχείου. Είναι λογικό ότι με την αύξηση της κατακόρυφης τάσης και τη συμπίεση στην κατακόρυφη διεύθυνση, το έδαφος τείνει να διογκωθεί πλευρικά. Επειδή, όμως, η πλευρική διογκωση παρεμποδίζεται, αναπτύσσονται οριζόντιες θλιπτικές τάσεις σ'_h , οι οποίες αυξάνουν με την αύξηση της κατακόρυφης τάσης σ'_v (ή σ'_z). Μεγάλος αριθμός εργαστηριακών μετρήσεων αποδεικνύει ότι ο λόγος σ'_h/σ'_v παραμένει σταθερός κατά τη φόρτιση και λαμβάνει τιμές μεταξύ 0.40 και 0.60 για διάφορους τύπους εδαφών, όπου οι μεγαλύτερες τιμές αντιστοιχούν στα αργιλικά υλικά και οι μικρότερες στα αμμώδη. Ο λόγος αυτός συμβολίζεται με:

$$K_o \equiv \frac{\sigma'_h}{\sigma'_v} \quad (1.48)$$

και ονομάζεται συντελεστής οριζόντιας ώθησης με παρεμπόδιση της πλευρικής παραμόρφωσης. Σύμφωνα με τη θεωρία της γραμμικής ισότροπης ελαστικότητας, η τιμή του K_o είναι:

$$K_o \equiv \frac{\nu}{1-\nu} \quad (1.49)$$

και λαμβάνει τιμές: $K_0 = 0.33 - 0.67$ για τιμές του λόγου Poisson $\nu = 0.25 - 0.40$. Εργαστηριακές μετρήσεις της οριζόντιας τάσης σ'_h κατά την αποφόρτιση των εδαφικών υλικών δείχνουν ότι η οριζόντια τάση μειώνεται αλλά με ρυθμό πολύ βραδύτερο από τη μείωση της κατακόρυφης τάσης, με συνέπεια ο λόγος K_0 να αυξάνει. Το σχήμα (1.23) παρουσιάζει ένα τυπικό διάγραμμα φόρτισης και αποφόρτισης και τις αντίστοιχες τιμές του λόγου K_0 . Η μικρή μείωση του σ'_h κατά την αποφόρτιση οφείλεται στο γεγονός ότι οι ολισθήσεις και κυλίσεις μεταξύ των κόκκων είναι στο μεγαλύτερό τους ποσοστό μη - αντιστρεπτές.



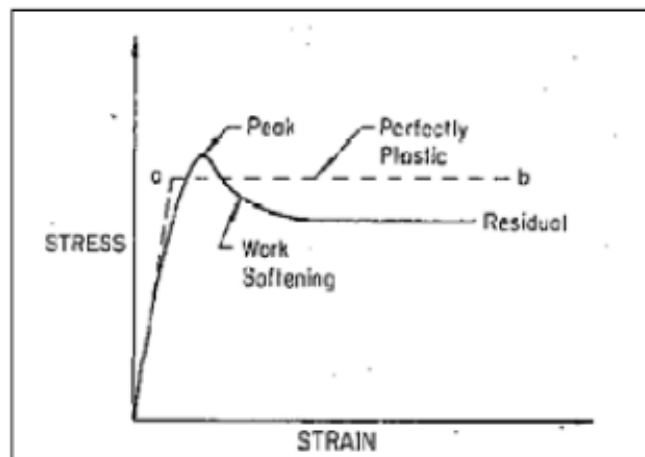
Σχήμα 1.23: Μεταβολή της οριζόντιας τάσης κατά την αποφόρτιση

1.4.5 Ανασκόπηση των βασικότερων καταστατικών μοντέλων του εδάφους

1.4.5.1 Μοντέλο Mohr-Coulomb

Το μοντέλο Mohr-Coulomb, όπως φαίνεται στο σχήμα (1.24) είναι ένα ελαστικό, τέλεια πλαστικό μοντέλο, το οποίο χρησιμοποιείται συχνά για την μοντελοποίηση της συμπεριφοράς του εδάφους. Στη γενική κατάσταση τάσεων, το συγκεκριμένο μοντέλο τάσεων - παραμορφώσεων συμπεριφέρεται γραμμικά στην ελαστική περιοχή, με δύο

παραμέτρους που προκύπτουν από τον ορισμό του νόμου Hooke (το μέτρο ελαστικότητας E και του λόγου Poisson ν). Υπάρχουν, επίσης, δύο παράμετροι που ορίζουν τα κριτήρια αστοχίας (η γωνία τριβής ϕ και η συνεκτικότητα γ) και μια παράμετρος για να περιγράψει τον κανόνα ροής (γωνία διασταλτικότητας ψ που χρησιμοποιείται για την μοντελοποίηση μιας ρεαλιστικής αναστρέψιμης μεταβολής του όγκου λόγω διάτμησης).

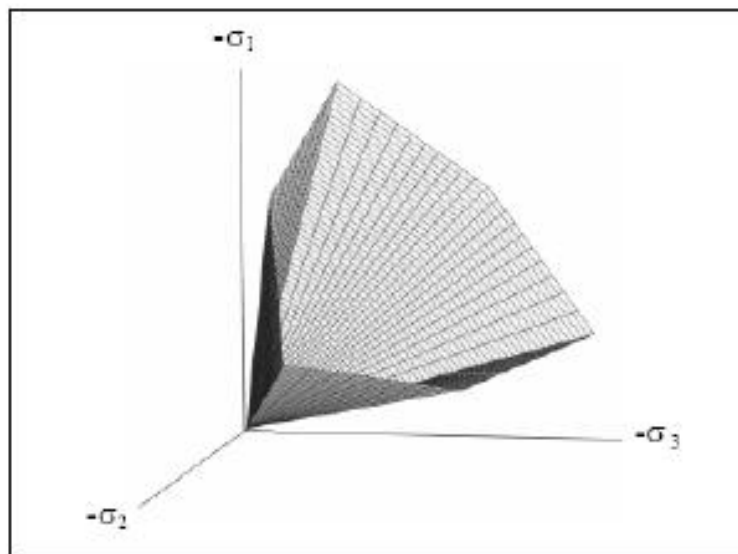


Σχήμα 1.24: Ελαστική – τέλεια πλαστική θεώρηση του μοντέλου Mohr-Coulomb

Στη συμβατική θεωρία πλαστικότητας, ο κανόνας ροής χρησιμοποιείται ως ο νόμος της εξέλιξης για το ρυθμό της πλαστικής παραμόρφωσης. Αν η συνάρτηση του πλαστικού δυναμικού είναι η ίδια με την συνάρτηση διαρροής, ο κανόνας ροής ονομάζεται σχετικός κανόνα ροής και αν είναι διαφορετική, καλείται μη σχετικός κανόνα ροής. Στην Εδαφομηχανική, ως σχετικός κανόνα ροής έχει χρησιμοποιηθεί για να περιγράψει τη συμπεριφορά της περιοχής, στην οποία είναι σημαντική η ύπαρξη αρνητικής διασταλτικότητας, το μοντέλο Cam clay για κανονικά στερεοποιημένο άργιλο. Ωστόσο, ο μη σχετικός κανόνας ροής χρησιμοποιείται συχνά για να περιγράψει τη συμπεριφορά της άμμου με τις αρνητικές και θετικές διασταλτικότητες.

Το μοντέλο Mohr-Coulomb είναι ένα απλό και εφαρμόσιμο μοντέλο στον τριδιάστατο χώρο των τάσεων (βλ. σχήμα 1.25) με δύο μόνο παραμέτρους αντοχής για να περιγράψει την πλαστική συμπεριφορά. Όσον αφορά τη συμπεριφορά της αντοχής, αυτό το μοντέλο αποδίδει καλύτερα. Οι ερευνητές έχουν δείξει μέσω πραγματικών τριαξονικών δοκιμών, ότι ο συνδυασμός των τάσεων, που προκαλούν αστοχία σε πραγματικά δείγματα εδαφικών υλικών, συμφωνούν αρκετά καλά με το εξαγωνικό σχήμα της καμπύλης αστοχίας (Goldscheider, 1984) [9]. Το μοντέλο αυτό εφαρμόζεται για την ανάλυση της σταθερότητας των φραγμάτων, των κλίσεων, των αναχωμάτων και των ρηχών θεμελίων.

Παρά το γεγονός ότι η αστοχία προσεγγίζεται γενικά καλά σε στραγγιζόμενες συνθήκες, η πραγματική διαδρομή τάσεων που ακολουθείται σε μη στραγγιζόμενα υλικά μπορεί να αποκλίνει σημαντικά από τις παρατηρήσεις. Είναι προτιμότερο να χρησιμοποιούνται παράμετροι μη στραγγιζόμενης διάτμησης σε μια μη στραγγιζόμενη ανάλυση, με γωνία τριβής ίση με το μηδέν. Η δυσκαμψία (άρα και η παραμόρφωση) πριν προσεγγίσει την τοπική διάτμηση είναι ελλιπέστατα, μοντελοποιημένη. Για τέλεια πλαστικότητα, το μοντέλο δεν περιλαμβάνει φαινόμενα κράτυνσης ή το φαινόμενο χαλάρωσης του εδάφους.



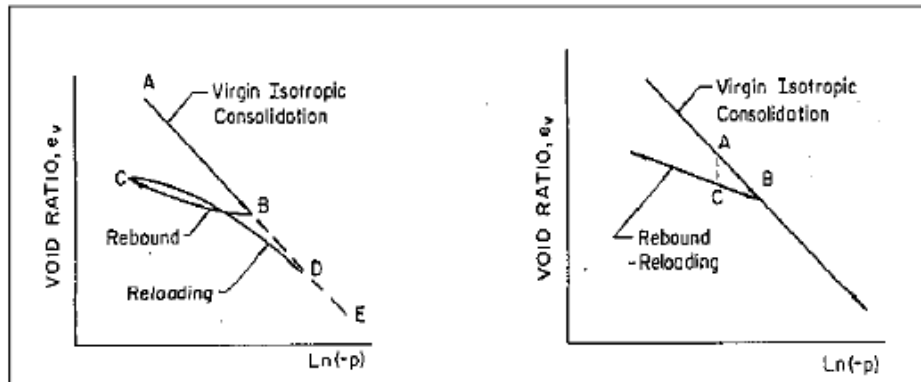
Σχήμα 1.25: Η επιφάνεια διαρροής του μοντέλου Mohr-Coulomb στο χώρο των κύριων τάσεων ($c = 0$)

1.4.5.2 Τροποποιημένο μοντέλο Cam clay

Πολλοί επιστήμονες στο παρελθόν είχαν ερευνήσει την πιθανότητα να μοντελοποιήσουν το έδαφος σαν ένα υλικό με φαινόμενα κράτυνσης. Ο Roscoe (1963) χρησιμοποίησε την θεωρία της παραμόρφωσης σκλήρυνσης της πλαστικότητας για να διατυπώσει ένα πλήρες μοντέλο τάσης – παραμόρφωσης για τον κανονικά στερεοποιημένο ή ελαφριά υπερστερεοποιημένο άργιλο στη τριαξονική δοκιμή γνωστό και ως μοντέλο Cam clay (Schofield και Wroth, 1968). Ο Burland (1965) [10] πρότεινε μια τροποποιημένη έκδοση του μοντέλου Cam clay και αυτό το μοντέλο επεκτάθηκε στη συνέχεια στην γενική κατάσταση των τριδιάστατων τάσεων.

Το τροποποιημένο μοντέλο Cam clay είναι ένα ελαστικό, πλαστικό μοντέλο παραμόρφωσης, όπου η μη γραμμική συμπεριφορά μοντελοποιείται με τη βοήθεια της πλαστικής σκλήρυνσης. Το μοντέλο βασίζεται στη βασική υπόθεση ότι υπάρχει μια λογαριθμική σχέση μεταξύ της μέσης πραγματικής τάσης p^o και του δείκτη πόρων του εδάφους e . Οι αρχικές γραμμές συμπίεσης και επανασυμπίεσης είναι γραμμικές στο

διάστημα $e - \ln(p^o)$, το οποίο είναι περισσότερο ρεαλιστικό για τους σχεδόν κανονικά στερεοποιημένους αργίλους (σχήμα 1.26). Μόνο η γραμμική ελαστική συμπεριφορά μοντελοποιείται πριν από τη διαρροή και μπορεί να οδηγή σε παράλογες τιμές του v λόγω των λογαριθμικών γραμμικών γραμμών συμπίεσης.



(a)

(b)

Σχήμα 1.26: (α) Συμπεριφορά του πραγματικού εδάφους στην υδροστατική πίεση, (β) Συμπεριφορά του ιδεατού εδάφους στην υδροστατική πίεση

Το μοντέλο αυτό είναι περισσότερο κατάλληλο για να περιγράψει την παραμόρφωση παρά την αστοχία, ειδικά για κανονικά στερεοποιημένα μαλακά εδάφη. Το μοντέλο έχει καλύτερη απόδοση σε εφαρμογές, που εμπλέκονται συνθήκες φόρτωσης, όπως αναχώματος ή θεμελίωση. Το μοντέλο περιλαμβάνει τέσσερις παραμέτρους, οι οποίες είναι, ο δείκτης ιστροπικής λογαριθμικής συμπίεσης λ , ο δείκτης διόγκωσης κ , ο λόγος Poisson για την φόρτιση και την επαναφόρτιση ν_{ur} , η σταθερά τριβής M , η τάση πριν την στερεοποίηση p_c και τον αρχικό δείκτη πόρων του εδάφους e . Η διατμητική αντοχή μπορεί να μοντελοποιηθεί μόνο με τη χρήση της σταθεράς τριβής. Στην περίπτωση της αρχικής αποστραγγιζόμενης αποκλίνουσας φόρτισης του μαλακού εδάφους, το μοντέλο προβλέπει πιο ρεαλιστικά την αποστραγγιζόμενη διατμητική αντοχή σε σύγκριση με το μοντέλο Mohr-Coulomb

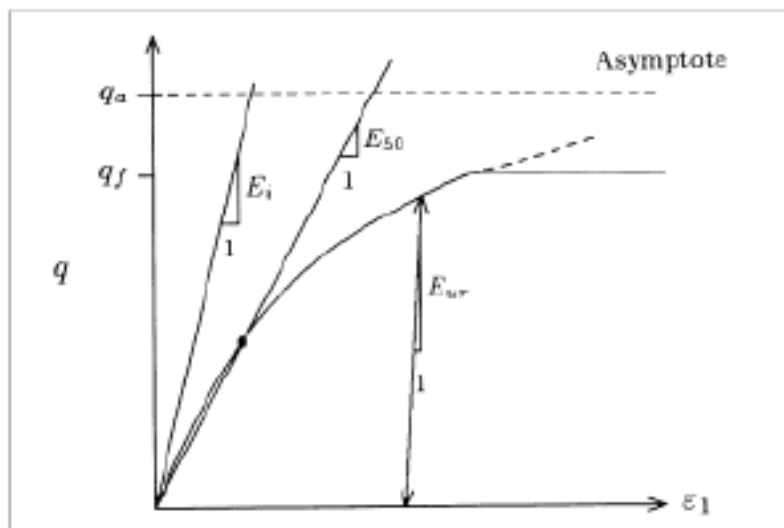
1.4.5.3 Μοντέλο Duncan – Chang (Υπερβολικό μοντέλο)

Όπως είναι γνωστό, η συμπεριφορά του εδάφους είναι εξαιρετικά μη γραμμική και αναστέλλει την εξαρτώμενη από την τάση, δυσκαμψία. Εκτός από τα μοντέλα που έχουν παρουσιαστεί ήδη, το μοντέλο Duncan-Chang μοντέλο είναι ένα στοιχειώδες μη – γραμμικό, εξαρτώμενο από την τάση, μοντέλο, το οποίο είναι επίσης γνωστό ως υπερβολικό μοντέλο (Duncan και Chang, 1970) [11].

Το μοντέλο αυτό βασίζεται σε μία καμπύλη τάσεων – παραμορφώσεων, που προκύπτει από μία στραγγιζόμενη τριαξονική δοκιμή συμπίεσης για άργιλο και για

άμμο, η οποία μπορεί να προσεγγιστεί από την καμπύλη της υπερβολής με πολύ καλή ακρίβεια (Kondner, 1936), όπως φαίνεται στο σχήμα (1.27). Βασίζεται στην ιδέα του Ohde, σύμφωνα με την οποία η δυσκαμψία του εδάφους μπορεί να προκύψει από μία παράμετρο εξαρτώμενη από την τάση, χρησιμοποιώντας έναν εκθετικό νόμο σχηματισμού. Το κριτήριο αστοχίας του μοντέλο βασίζεται στις δύο παραμέτρους αντοχής του μοντέλου Mohr- Coulomb. Το πιο σημαντικό, είναι ότι το μοντέλο αυτό περιγράφει τα τρία πιο σημαντικά χαρακτηριστικά του εδάφους, τα οποία είναι η μη γραμμικότητα, εξαρτώμενη από την τάση και η ανελαστική συμπεριφορά του συνεκτικού και μη συνεκτικού εδάφους.

Για ένα συγκεκριμένο επίπεδο της τάσης περίσφιξης, η διάκριση γίνεται μεταξύ της δυσκαμψίας E_t (εξαρτώμενη από την τάση) μιας αρχικής φόρτισης και μίας σταθερής δυσκαμψίας E_{ur} φόρτισης και αποφόρτισης. Η φόρτιση προσδιορίζεται από την συνθήκη $d(\sigma_1/\sigma_3) > 0$. Σε αυτή τη συνθήκη η πλαστική παραμόρφωση εμφανίζεται, εφ' όσον η τρέχουσα τάση είναι στην επιφάνεια διαρροής. Για να συνεχιστεί η πλαστική ροή, η τρέχουσα τάση πρέπει να παραμείνει στην επιφάνεια διαρροής. Διαφορετικά η τάση θα πρέπει να μειωθεί κάτω από την τιμή διαρροής. Σε αυτή την περίπτωση δεν απαιτείται περαιτέρω πλαστική παραμόρφωση και όλες οι στοιχειώδεις παραμορφώσεις είναι ελαστικές. Αυτή η κατάσταση $d(\sigma_1/\sigma_3) < 0$ καλείται «εκφόρτωση».



Σχήμα 1.27: Η υπερβολική σχέση τάσεων – παραμορφώσεων για την κύρια φόρτιση σε μία πρότυπη στραγγιζόμενη τριαξονική δοκιμή

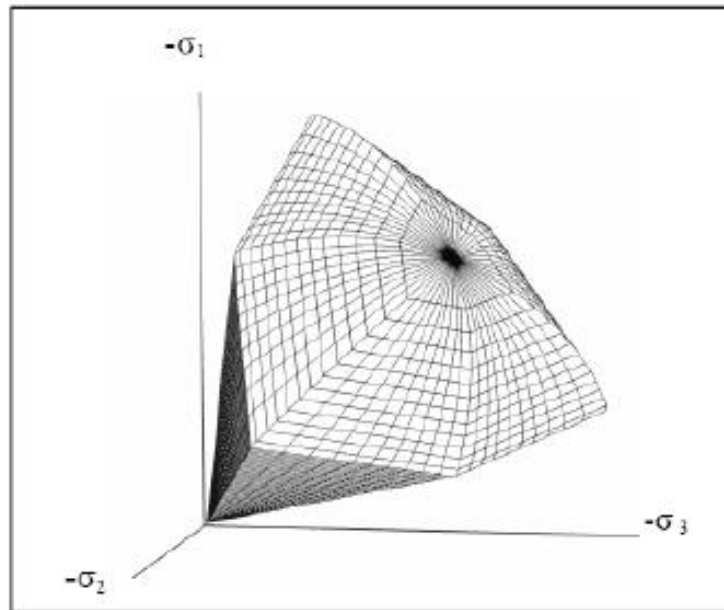
Το μοντέλο Duncan-Chang χρησιμοποιείται ευρέως, δεδομένου ότι οι εδαφολογικές παράμετροί του μπορούν να ληφθούν άμεσα από την πρότυπη τριαξονική δοκιμή. Είναι μια απλή, αλλά προφανής βελτίωση του μοντέλου Mohr-Coulomb. Από αυτή την άποψη, αυτό το μοντέλο προτιμάται περισσότερο από το μοντέλο Mohr-Coulomb. Το

κριτήριο αστοχίας του περιγράφεται με τη βοήθεια του κριτηρίου αστοχίας του μοντέλου Mohr - Coulomb, αλλά αυτό δεν διατυπώνεται κατάλληλα στα πλαίσια της πλαστικότητας. Κατά συνέπεια, η διασταλτικότητα δεν μπορεί να περιγραφεί. Αυτό το μοντέλο προσεγγίζει την εδαφολογική συμπεριφορά χρησιμοποιώντας μόνο δύο παραμέτρους δυσκαμψίας και προτιμάται πάρα πολύ για την πρακτική μοντελοποίηση.

1.4.5.4 Μοντέλο εδάφους με κράτυνση

Το μοντέλο εδάφους με κράτυνση (Brinkgreve & Vermeer, 1997; Schanz, 1998) είναι ένα πραγματικό, δεύτερης τάξης, μοντέλο για το έδαφος και χρησιμοποιείται για οποιαδήποτε εφαρμογή. Το μοντέλο περιλαμβάνει τριβή σκλήρυνσης για την μοντελοποίηση της πλαστικής διατμητικής τάσης στην αποκλίνουσα φόρτιση και κράτυνση για τη μοντελοποίηση της πλαστικής ογκομετρικής παραμόρφωσης στην αρχική συμπίεση. Ο διαχωρισμός της σκλήρυνσης μπορεί να γίνει σε δύο τύπους, την διατμητική σκλήρυνση και την σκλήρυνση της συμπίεσης. Η διατμητική σκλήρυνση χρησιμοποιείται για τη μοντελοποίηση των παραμενουσών παραμορφώσεων, λόγω της κύριας αποκλίνουσας φόρτισης. Η σκλήρυνση της συμπίεσης χρησιμοποιείται για την μοντελοποίηση παραμενουσών πλαστικών παραμορφώσεων, λόγω της κύριας συμπίεσης στην φόρτιση του συμπιεσομέτρου και στην ισοτροπική φόρτιση. Το παρών μοντέλο περιέχει και τους δύο τύπους σκλήρυνσης. Η καμπύλη της διαρροής στην τριδιάστατη κατάσταση παρουσιάζεται στο σχήμα (1.28). Το κριτήριο αστοχίας του μοντέλου προσδιορίζεται με βάση το αντίστοιχο του μοντέλου Mohr – Coulomb. Εξαιτίας των δύο τύπων σκλήρυνσης το μοντέλο είναι επίσης ακριβές για προβλήματα που περιέχουν μείωση της μέσης πραγματικής τάσης και την ίδια στιγμή μεταβολή της αντοχής σε διάτμηση. Τέτοια προβλήματα συμβαίνουν σε περιπτώσεις εκσκαφών και κατασκευής τούνελ.

Όσον αφορά την συμπεριφορά της δυσκαμψίας το μοντέλο περιλαμβάνει ένα εκθετικό νόμο σχηματισμού της εξαρτώμενης από την τάση, δυσκαμψίας, παρόμοιο με αυτό που χρησιμοποιείται στο μοντέλο Duncan-Chang. Στην πραγματικότητα το παρών μοντέλο παρουσιάζει αντιστοιχία με το μοντέλο Duncan-Chang, σχετικά με την υπερβολική καμπύλη τάσεων – παραμορφώσεων, όταν προσομοιώνει μία πρότυπη στραγγιζόμενη τριαξονική δοκιμή (σχήμα 1.27). Δεδομένου ότι το μοντέλο σκλήρυνσης του εδάφους βασίζεται στην πλαστικότητα σκλήρυνσης παρά στη μη γραμμική πλαστικότητα, υπερνικά τους περιορισμούς και τις ασυνέπειες του μοντέλου Duncan-Chang, όσον αφορά τη διασταλτικότητα και την φυσική φόρτιση.



Σχήμα 1.28: Συνολική επιφάνεια διαρροής του μοντέλου hardening soil στο χώρο των κυρίων τάσεων για μη συνεκτικό έδαφος

Μερικά από τα βασικά χαρακτηριστικά αυτού του μοντέλου είναι η εξαρτώμενη από την τάση δυσκαμψία σύμφωνα με τον εκθετικό νόμο (m), πλαστική παραμόρφωση λόγω της αρχικής αποκλίνουσας φόρτισης (E^{ef}_{50}), πλαστική παραμόρφωση λόγω της αρχικής συμπίεσης (E^{ef}_{oed}) και του κριτηρίου αστοχίας σύμφωνα με το μοντέλο Mohr-Coulomb (c , φ και ψ).

ΚΕΦΑΛΑΙΟ 2^ο

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1 Εισαγωγή

Ο ανθρώπινος εγκέφαλος εκτελεί διάφορες λειτουργίες σχετικές με την αντίληψη, όπως είναι η αναγνώριση εικόνων και ομιλίας, με αξιοσημείωτη ευκολία και αποτελεσματικότητα. Αντίθετα, τέτοιου είδους γνωστικές λειτουργίες αποτελούν έργο, εξαιρετικά δύσκολο, για τους ψηφιακούς ηλεκτρονικούς υπολογιστές, που όμως χαρακτηρίζονται από την πολύ μεγάλη ταχύτητα και ακρίβεια λειτουργίας. Τα *τεχνητά νευρωνικά δίκτυα* (ANN: *artificial neural networks*) έχουν ως στόχο την ταχεία επίλυση τέτοιων προβλημάτων μέσω μηχανισμών, των οποίων η σύλληψη βασίζεται σε βιολογικά πρότυπα.

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από τη βιολογία, αποτελούνται από στοιχεία τα οποία συμπεριφέρονται κατά τρόπο ανάλογο των πιο στοιχειωδών λειτουργιών των βιολογικών κυττάρων. Τα στοιχεία αυτά είναι οργανωμένα κατά τέτοιο τρόπο, ώστε να προσομοιάζουν την ανατομία του ανθρωπίνου εγκεφάλου. Παρά την, όχι όμως και τόσο, μεγάλη ομοιότητα τους με τον εγκέφαλο επιτυγχάνουν να προσεγγίσουν ένα μεγάλο αριθμό χαρακτηριστικών της δομής αλλά και της λειτουργίας του εγκεφάλου. Για παράδειγμα, μαθαίνουν χρησιμοποιώντας την εμπειρία, την οποία έχουν συσσωρεύσει, έχουν την ικανότητα γενίκευσης από προηγούμενα παραδείγματα σε νέα, μπορούν να επεξεργαστούν μια ομάδα δεδομένων και να ξεχωρίσουν από αυτή τα ουσιωδέστερα χαρακτηριστικά.

Ένα σύνολο τιμών αποτελούμενο από εισόδους ή από ζεύγη εισόδων - εξόδων (I/O pairs) επιδρά στο δίκτυο και βάσει του συνόλου αυτού, οι παράμετροι του δικτύου διορθώνονται έτσι ώστε να παραχθεί η επιθυμητή για το δίκτυο απόκριση. Η διαδικασία αυτή ονομάζεται *εκπαίδευση* (*training*). Υπάρχει μεγάλη ποικιλία αλγορίθμων εκπαίδευσης, ο καθένας από τους οποίους έχει τα προτερήματα και τα ελαττώματά του. Ένα καλά εκπαιδευμένο δίκτυο δεν θα πρέπει να παρουσιάζει ευαισθησία στην απόκρισή του για μικρές αλλαγές στις εισόδους του. Είναι σημαντικό να επισημανθεί ότι από την φύση τους τα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα της γενίκευσης, χωρίς όμως να χρησιμοποιούν κάτι ανάλογο της ανθρώπινης νοημοσύνης. Ορισμένα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα να αποσπούν τα ουσιώδη στοιχεία από ένα σύνολο δεδομένων εισαγωγής και να τα χρησιμοποιούν για να παράγουν κάτι που ήταν κρυμμένο μέσα σε αυτό το σύνολο.

Οι δυνατότητες και τα πλεονεκτήματα των τεχνητών νευρωνικών δικτύων μπορούν να συνοψιστούν στα παρακάτω: **(i)** εκπαίδευση σε παράλληλο υπολογιστικό περιβάλλον, **(ii)** ικανότητα εκμάθησης, **(iii)** ικανότητα αναδημιουργίας αντικειμένων όταν είναι

γνωστό ένα μόνο μέρος της περιγραφής τους, **(iv)** ικανότητα αποθήκευσης αντικειμένων και ανάκληση τους με βάση το περιεχόμενο και όχι τη διεύθυνση, *σχετικιστική μνήμη (associative memory)*, **(v)** ικανότητα προσομοίωσης συναρτήσεων.

Η χρησιμοποίηση τεχνικών τεχνητής νοημοσύνης, όπως είναι τα νευρωνικά δίκτυα έχει αρχίσει να βρίσκει εφαρμογή σε προβλήματα υπολογιστικής μηχανικής την τελευταία δεκαετία. Τέτοιου είδους εφαρμογές είναι η ανάλυση αξιοπιστίας κατασκευών, όπου στόχος είναι η πρόβλεψη των αποτελεσμάτων της ανάλυσης [13], προβλήματα βέλτιστου σχεδιασμού όπου οι απαιτούμενες τιμές των περιορισμών για κάθε νέο σχεδιασμό προέρχεται από την πρόβλεψη ενός κατάλληλα εκπαιδευμένου νευρωνικού δικτύου [14-18], έχουν επίσης εφαρμοστεί σε προβλήματα θραυστομηχανικής [19], σε προβλήματα προσαρμοστικών πεπερασμένων στοιχείων [20], καθώς επίσης και στη μοντελοποίηση καταστατικών νόμων διαφόρων υλικών, το οποίο αποτελεί και θέμα μελέτης της παρούσας εργασίας. Ένα σωστά εκπαιδευμένο νευρωνικό δίκτυο μπορεί να παράγει αποδεκτά, από πλευράς ακρίβειας, αποτελέσματα ενώ απαιτεί πάρα πολύ μικρή υπολογιστική προσπάθεια. Η ιδιότητα αυτή του NN είναι και το βασικό πλεονέκτημα αυτής της διαδικασίας. Παραδείγματος χάριν, η προσέγγιση της λύσης της ανάλυσης φορέων μέσω ενός νευρωνικού δικτύου έχει μεγάλη αξία σε περιπτώσεις χρονοβόρων αναλύσεων, όπου είναι αναγκαία μια γρήγορη εκτίμηση της συμπεριφοράς των φορέων.

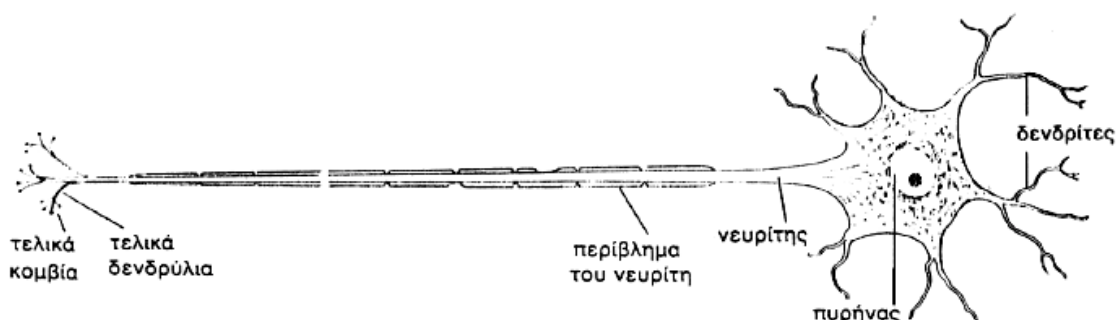
2.2 Οι βασικές αρχές των τεχνητών νευρωνικών δικτύων (ANN)

Τα τεχνητά νευρωνικά δίκτυα έχουν αναπτυχθεί σε μεγάλη ποικιλία σχηματισμών - αρχιτεκτονικών. Οι αρχιτεκτονικές στις οποίες μπορούμε να κατατάξουμε τα τεχνητά νευρωνικά δίκτυα είναι δύο: *Δίκτυα Πρόσθιας Τροφοδότησης (feed-forward neural networks)*, στα οποία η ροή των σημάτων γίνεται προς μία κατεύθυνση και *Δίκτυα Πίσω Τροφοδότησης ή Δίκτυα με Ανατροφοδότηση (feed-back or recurrent neural networks)*, στα οποία πραγματοποιείται επαναφορά της ροής σε μονάδες του δικτύου. Αν και οι δύο αυτοί τύποι αρχιτεκτονικών βασίζονται σε διαφορετική φιλοσοφία και υπακούουν σε διαφορετικές αρχές, παρ'όλα αυτά χαρακτηρίζονται από μία κοινή ιδιότητα, η οποία έγκειται στην ικανότητα να μαθαίνουν μέσω της εκπαίδευσης. Εκπαίδευση ονομάζεται η διαδικασία μέσω της οποίας τα νευρωνικά δίκτυα αποκτούν την ικανότητα να εκτελούν μια συγκεκριμένη εργασία. Η διαδικασία εκπαίδευσης μπορεί να γίνει με *επίβλεψη (supervised)* ή *χωρίς επίβλεψη (unsupervised)*.

2.2.1 Το βιολογικό πρότυπο

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από την βιολογία. Οι ερευνητές που ασχολούνται με τον σχεδιασμό των τεχνητών νευρωνικών δικτύων και των αλγορίθμων που τα διέπουν έχουν ως πρότυπο την δομή του ανθρωπίνου εγκεφάλου. Επειδή όμως η γνώση, όσον αφορά τον ανθρώπινο εγκέφαλο και τις λειτουργίες του, είναι περιορισμένη, ώστε να αποτελέσει οδηγό για τους ερευνητές, αναγκάστηκαν να προχωρήσουν πέρα από την υπάρχουσα, έως τώρα, βιολογική γνώση και να αναζητήσουν φανταστικές δομές που να προσομοιώνουν τα βιολογικά πρότυπα. Έτσι σχεδιάστηκαν δίκτυα, τα οποία είναι οργανικώς μη εφικτά ή απαιτούν ένα μεγάλο αριθμό υποθέσεων, όσον αφορά την ανατομία του εγκεφάλου και την λειτουργία του. Παρά την μικρή, έως πολλές φορές ανύπαρκτη σχέση τους με τη βιολογική δομή του εγκεφάλου, τα τεχνητά νευρωνικά δίκτυα συχνά συγκρίνονται, όσον αφορά τη δομή αλλά και τη λειτουργία του εγκεφάλου.

Το ανθρώπινο νευρικό σύστημα είναι χτισμένο από πολύπλοκα κύτταρα, τα οποία ονομάζονται *Νευρώνες* (neurons). Κατά μία εκτίμηση ο ανθρώπινος εγκέφαλος αποτελείται από 10^{11} νευρώνες που σχηματίζουν έναν τεράστιο αριθμό συνδέσεων (10^{15}) μεταξύ τους, μέσω διαδρόμων επικοινωνίας που είναι πιθανό να έχουν μήκος, μεγαλύτερο του ενός μέτρου. Κάθε νευρώνας μοιράζεται πολλά κοινά χαρακτηριστικά με άλλα κύτταρα στο σώμα, αλλά εμφανίζει ορισμένες μοναδικές ικανότητες, όπως να δέχεται, να επεξεργάζεται και να μεταφέρει ηλεκτροχημικά σήματα μέσω των νευρωνικών μονοπατιών, τα οποία αποτελούν το σύστημα επικοινωνίας του εγκεφάλου. Κάθε τέτοιο ηλεκτροχημικό σήμα ονομάζεται νευρική ώση.



Σχήμα 2.1: Βιολογικός νευρώνας

Στο σχήμα (2.1) παρουσιάζεται η δομή ενός τυπικού ζεύγους βιολογικών νευρώνων. Ένας τυπικός νευρώνας, στην πιο συνηθισμένη του μορφή, αποτελείται από το *Σώμα* και τις *Αποφυάδες* του. Οι αποφυάδες είναι δύο ειδών: ο *Νευρίτης* που είναι μια μακριά αποφυάδα, που το μήκος της μπορεί να είναι από λίγα δέκατα του χιλιοστομέτρου μέχρι πάνω από ένα μέτρο, σε μερικές περιπτώσεις, και οι *Δενδρίτες* (*Dendrites*) που

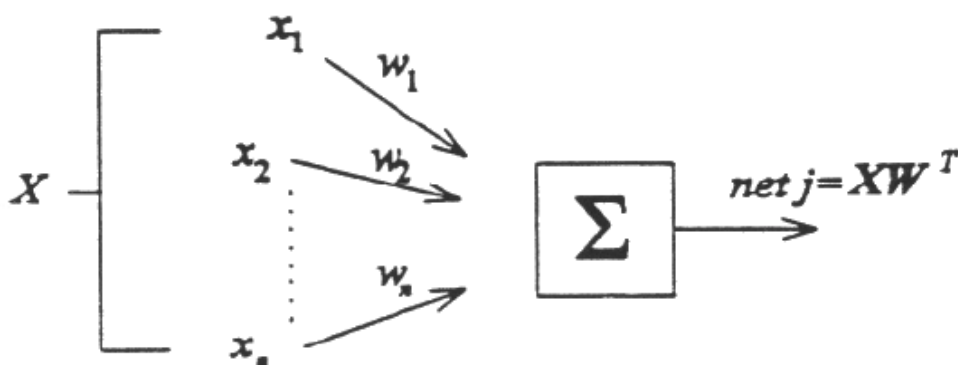
είναι πολλοί, κοντοί και έχουν διακλαδώσεις σαν δέντρο. Οι τελικές απολήξεις του νευρίτη λέγονται *τελικά δενδρύλια*, στα άκρα των οποίων υπάρχουν τα *τελικά κομβία*. Κάθε νευρώνας με τα τελικά κομβία μπορεί να έρχεται σε λειτουργική σύνδεση με τους Δενδρίτες των γειτονικών νευρώνων. Η θέση της λειτουργικής αυτής σύνδεσης, μέσω της οποίας μεταβιβάζεται η νευρική ώση, λέγεται *Σύναψη (Synapse)*. Από την μεριά της σύναψης όπου έγινε η υποδοχή, οι εισόδοι κατευθύνονται στο σώμα του κυττάρου, όπου και αθροίζονται. Ορισμένες από τις εισόδους τείνουν να διεγείρουν το κύτταρο ενώ άλλες τείνουν να αναχαιτίσουν την διέγερση του.

Για να προκληθεί νευρική ώση πρέπει το ερέθισμα να έχει τέτοια ένταση, ώστε να γίνει αναστροφή του δυναμικού ηρεμίας στον νευρώνα. Ένας νευρώνας βρίσκεται σε ηρεμία όταν δεν επιδρά πάνω του κάποιο ερέθισμα. Ερεθίσματα μικρότερης έντασης δεν προκαλούν νευρική ώση, ενώ εντονότερα ερεθίσματα δίνουν απλώς μεγαλύτερη συχνότητα νευρικών ώσεων.

2.2.2 Ο τεχνητός νευρώνας

Ο τεχνητός νευρώνας σχεδιάστηκε ώστε να μιμηθεί τα βασικά χαρακτηριστικά του βιολογικού νευρώνα. Η λειτουργία του τεχνητού νευρώνα μπορεί να περιγραφεί ως ακολούθως: Ένα σύνολο εισόδων οι οποίοι αντιστοιχούν στα ηλεκτροχημικά σήματα, που αναφέρθηκαν στην περιγραφή του βιολογικού νευρώνα και προέρχονται από γειτονικούς νευρώνες, επιδρά επί του νευρώνα. Η επιρροή που έχει κάποιο σήμα σε κάποιο νευρώνα προσεγγίζεται με τον πολλαπλασιασμό της εισόδου (σήμα) με κάποια παράμετρο η οποία ονομάζεται *συναπτικό βάρος (synaptic weight)*, το οποίο είναι ανάλογο της αντοχής της σύναψης του βιολογικού νευρώνα. Τα γινόμενα [βάρος]×[είσοδος] αθροίζονται ώστε να οριστεί το επίπεδο διέγερσης του νευρώνα.

$$NET = x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad (2.1)$$



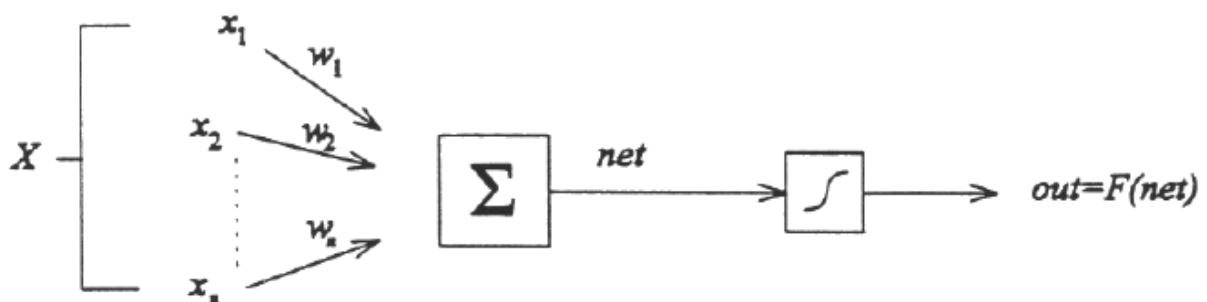
Σχήμα 2.2: Τεχνητός νευρώνας

Στο σχήμα (2.2) παρουσιάζεται ένας τεχνητός νευρώνας. Παρά την μεγάλη ποικιλία αρχιτεκτονικών δικτύων που υπάρχει, όλες βασίζονται στον σχηματισμό του τεχνητού νευρώνα που παρουσιάζεται στο σχήμα (2.2). Όπως παρατηρούμε ένα σύνολο από εισόδους x_1, x_2, \dots, x_n επιδρά στον τεχνητό νευρώνα. Οι εισοδοί αυτές μπορούν να ενσωματωθούν σε ένα διάνυσμα \mathbf{X} . Κάθε σήμα πολλαπλασιάζεται με ένα βάρος w_1, w_2, \dots, w_n πριν γίνει η άθροιση. Κάθε βάρος αντιστοιχεί στην “αντοχή” της κάθε βιολογικής σύνδεσης. Το σύνολο των βαρών σχηματίζει το διάνυσμα \mathbf{W} . Η σχέση (2.1) με την χρήση διανυσμάτων γίνεται ως εξής:

$$\text{NET} = [x_1 \quad x_2 \quad \dots \quad x_n] \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad (2.2)$$

2.2.3 Συναρτήσεις μεταφοράς

Το σήμα **NET** που παράγεται από την εξίσωση (2.2) στις περισσότερες περιπτώσεις περνά και από ένα επιπλέον στάδιο επεξεργασίας που ορίζεται από μία συνάρτηση, τη *συνάρτηση μεταφοράς* (*transfer function*), ώστε να παραχθεί τελικά η έξοδος του νευρώνα η οποία συμβολίζεται **OUT**. Η συνάρτηση μεταφοράς μπορεί να πάρει διάφορες μορφές.



Σχήμα 2.3: Τεχνητός νευρώνας με την συνάρτηση μεταφοράς

Στο σχήμα (2.3) περιγράφεται η διαδικασία υπολογισμού της εξόδου **OUT**: το σήμα **NET** συμπιέζεται από την συνάρτηση F , ώστε το τελικό σήμα **OUT** που παράγεται να μην υπερβαίνει κάποια χαμηλά όρια ανεξάρτητα της τιμής του σήματος **NET**. Η συνάρτηση μεταφοράς F ονομάζεται και *συνάρτηση σύνθλιψης* (*squashing function*). Σκοπός επίσης της συνάρτησης μεταφοράς είναι η προσομοίωση των καταστάσεων ηρεμίας (0 ή -1) και διέγερσης (1 ή +1) του βιολογικού νευρώνα. Οι παρακάτω συναρτήσεις μεταφοράς έχουν χρησιμοποιηθεί:

(α) Η απλή γραμμική συνάρτηση

$$\text{OUT} = F(\text{NET}) = k \cdot \text{NET} \quad (2.3)$$

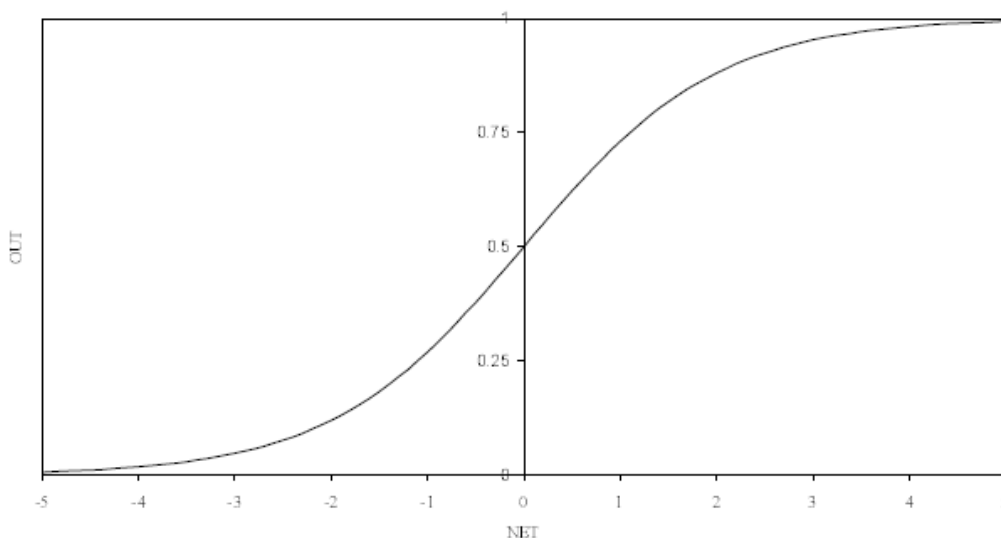
όπου k είναι ένας σταθερός αριθμός.

(β) Η συνάρτηση κατωφλίου

$$\begin{aligned} \text{OUT} &= 1 \text{ \acute{a}\nu } \text{NET} > T \\ \text{OUT} &= 0 \text{ διαφορετικά} \end{aligned} \quad (2.4)$$

όπου T είναι η σταθερή τιμή του κατωφλίου. Η συνάρτηση κατωφλίου είναι εκείνη που χρησιμοποιείται στο μοντέλο των McCulloch και Pitts [21].

(γ) Η συνάρτηση που προσομοιώνει με **μεγαλύτερη ακρίβεια** τα μη γραμμικά χαρακτηριστικά μεταφοράς των βιολογικών νευρώνων. Μια τέτοια συνάρτηση μεταφοράς που συχνά επιλέγεται είναι η **λογιστική** ή **“σιγμοειδής”** συνάρτηση (*logistic or sigmoid function*), όπως συχνά αποκαλείται, λόγω της μορφής του γραφήματος της που είναι σχήματος S, όπως φαίνεται στο σχήμα (2.4).



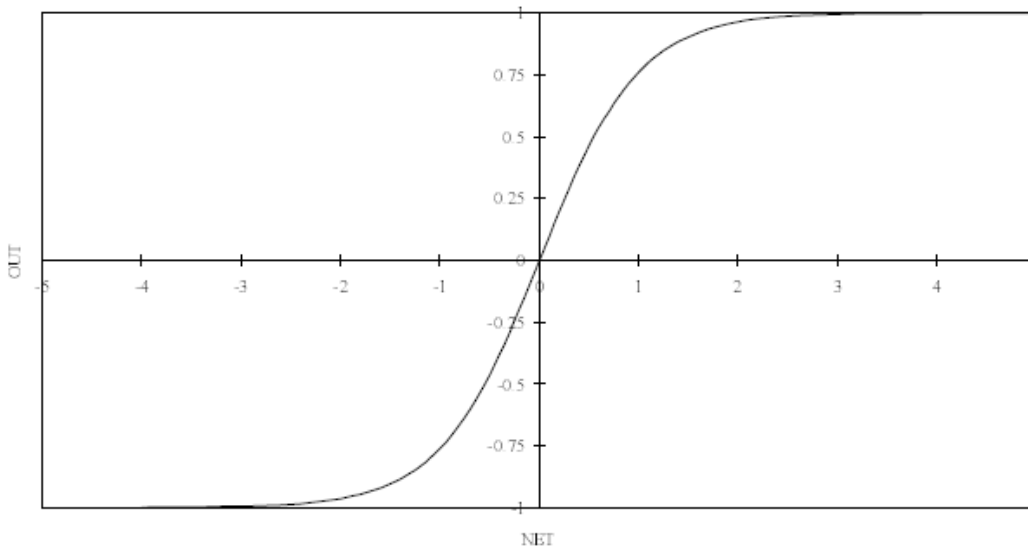
Σχήμα 2.4: Η σιγμοειδής λογιστική συνάρτηση

Το σήμα εξόδου, που θα παραχθεί από την σιγμοειδή συνάρτηση μεταφοράς, έχει τη μορφή:

$$\text{OUT} = \frac{1}{1 + e^{-\text{NET}}} = F(\text{NET}) \quad (2.5)$$

Μία δεύτερη, συχνά χρησιμοποιούμενη, συνάρτηση μεταφοράς είναι η συνάρτηση υπερβολικής εφαπτομένης. Έχει παρόμοιο γράφημα με την σιγμοειδή συνάρτηση και συχνά χρησιμοποιείται από βιολόγους για το μαθηματικό μοντέλο της δράσης των κυττάρων. Χρησιμοποιούμενη σε τεχνητό νευρωνικό δίκτυο έχει την έκφραση:

$$\text{OUT} = \tanh(\text{NET}) \quad (2.6)$$



Σχήμα 2.5: Η συνάρτηση υπερβολικής εφαπτομένης

Όπως η σιγμοειδής συνάρτηση, η συνάρτηση υπερβολικής εφαπτομένης παρουσιάζει γράφημα σχήματος S (βλέπε σχήμα 2.5) αλλά είναι αντισυμμετρική με αποτέλεσμα: **OUT** = 0 όταν **NET** = 0. Αντίθετα με τη σιγμοειδή συνάρτηση, η συνάρτηση υπερβολικής εφαπτομένης δίνει δυαδικές τιμές εξόδου, οι οποίες προσφέρονται για ορισμένα δίκτυα.

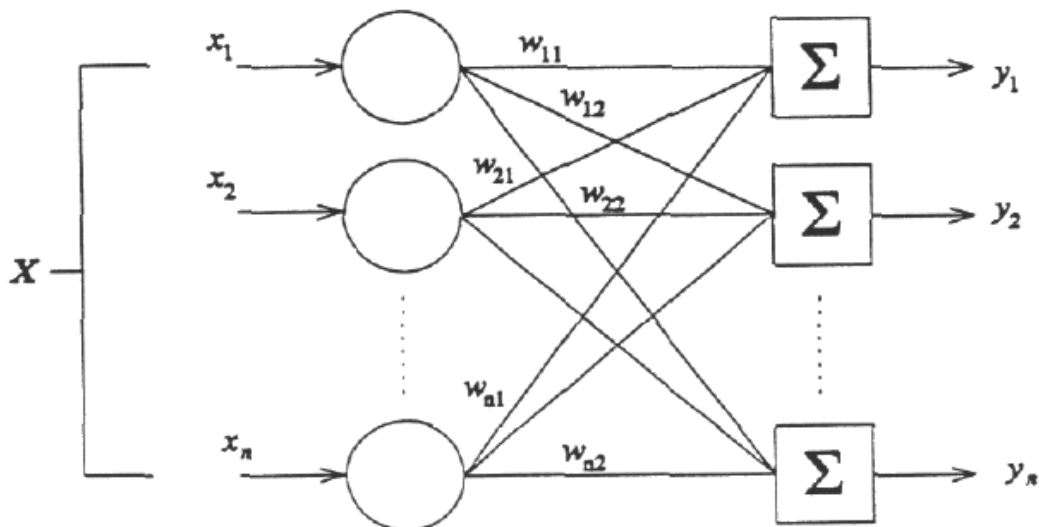
2.2.4 Δίκτυα Πρόσθιας Τροφοδότησης

Τα τεχνητά νευρωνικά δίκτυα αυτού του τύπου είναι δυνατό να αποτελούνται από ένα ή περισσότερα επίπεδα ή στοιβάδες (layers) αποτελούμενα από τεχνητούς νευρώνες. Οι τεχνητοί νευρώνες, οι οποίοι ανήκουν σε γειτονικά δίκτυα συνδέονται μεταξύ τους μέσω των συναπτικών βαρών. Αυτού του τύπου οι αρχιτεκτονικές ονομάζονται *Πρόσθιας Τροφοδότησης*, εφόσον οι έξοδοι κάθε επιπέδου τροφοδοτούν τους νευρώνες του αμέσως επόμενου επιπέδου. Τα μοντέλα Perceptron (Rosenblatt 1962 [22]) και Adaline (Windrow 1962 [23]) αποτελούν τα πρώτα μοντέλα αυτού του τύπου αρχιτεκτονικής. Όμως οι πραγματικές δυνατότητες των νευρωνικών δικτύων πρόσθιας τροφοδότησης ανακαλύφθηκαν πρόσφατα με την εμφάνιση των πολυεπίπεδων νευρωνικών δικτύων [24-26]. Ένα δίκτυο πρόσθιας τροφοδότησης μπορεί να προσομοιαστεί ως ένα σύστημα απεικόνισης ενός συνόλου εισόδων σε ένα σύνολο εξόδων. Το δίκτυο αποκτά

αυτή την ικανότητα προσαρμόζοντας κατάλληλα το σύνολο των συναπτικών βαρών μέσω της διαδικασίας εκπαίδευσης.

2.2.4.1 Τεχνητό νευρωνικό δίκτυο ενός επιπέδου

Οι υπολογιστικές δυνατότητες των νευρωνικών δικτύων προέρχονται από την σύνδεση νευρώνων μεταξύ τους σε σχηματισμό δικτύου. Το πιο απλό δίκτυο είναι εκείνο το οποίο αποτελείται από μια ομάδα νευρώνων συνδεδεμένων μεταξύ τους σε ένα επίπεδο όπως φαίνεται στο σχήμα (2.6).



Σχήμα 2.6: Νευρωνικό δίκτυο ενός επιπέδου

Οι κυκλικοί κόμβοι στα αριστερά της εικόνας χρησιμεύουν μόνο στην κατανομή των εισόδων στους νευρώνες, δεν εκτελούν κανένα υπολογισμό και δεν θεωρούνται ότι αποτελούν ένα επίπεδο. Η είσοδος κάθε νευρώνα είναι το άθροισμα των γινομένων [βάρος]×[είσοδος]. Πρακτικά τα τεχνητά και τα βιολογικά δίκτυα είναι δυνατό να μην έχουν πολλές από τις συνδέσεις τους ενεργές. Στο σχήμα (2.6) παρουσιάζεται ένα πλήρες δίκτυο για λόγους γενικότητας.

Τα βάρη του δικτύου σχηματίζουν το μητρώο w . Η διάσταση του μητρώου είναι $m \times n$, όπου m είναι το πλήθος των εισόδων και n το πλήθος των νευρώνων. Για παράδειγμα το βάρος που συνδέει την τρίτη είσοδο με τον δεύτερο νευρώνα θα είναι το $w_{3,2}$. Έτσι με αυτό τον τρόπο είναι φανερό ότι ο υπολογισμός του συνόλου των εξόδων των νευρώνων **NET**, για ένα επίπεδο είναι ένας απλός πολλαπλασιασμός μητρώων.

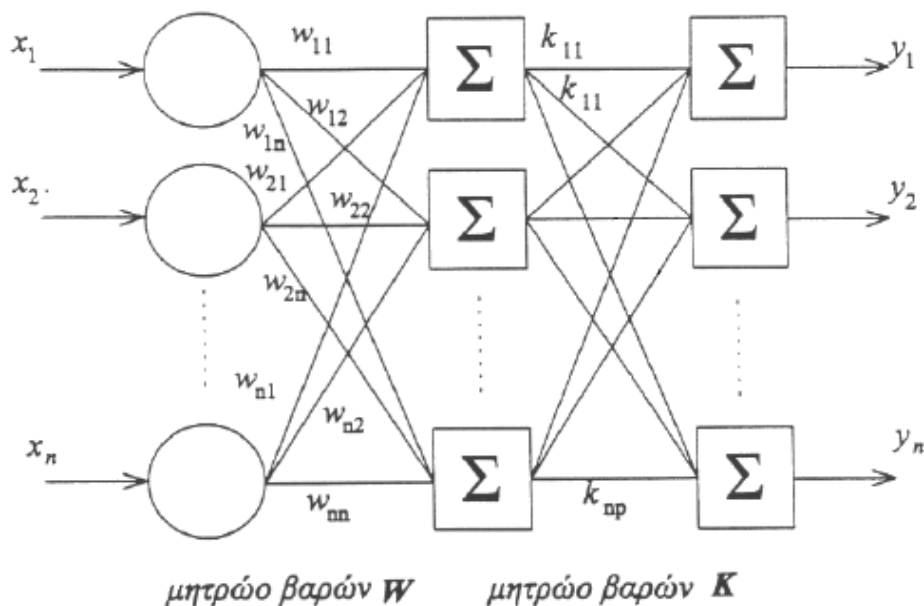
$$\text{NET} = \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix} \cdot \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{m1} & w_{m2} & \dots & w_{mn} \end{bmatrix} \quad (2.7)$$

οπότε οι έξοδοι του δικτύου είναι:

$$\text{out}_i = F\left(\sum_{j=1}^m w_{ji} \cdot x_j\right) \quad (2.8)$$

2.2.4.2 Τεχνητό νευρωνικό δίκτυο πολλών επιπέδων

Αρχικά η έρευνα επικεντρώθηκε σε νευρωνικά δίκτυα ενός επιπέδου με γραμμική ή μη - γραμμική συνάρτηση μεταφοράς. Όμως σύντομα έγινε αντιληπτό ότι τα δίκτυα ενός επιπέδου δεν είναι δυνατό να δώσουν λύση σε μια μεγάλη ποικιλία προβλημάτων. Μεγαλύτερα, πιο πολύπλοκα δίκτυα προσφέρουν μεγαλύτερες υπολογιστικές δυνατότητες. Η εισαγωγή των πολυεπίπεδων δικτύων έδωσε νέα ώθηση στον τομέα των νευρωνικών δικτύων. Ο κύριος λόγος, ο οποίος απέτρεπε την χρήση των πολυεπίπεδων δικτύων ήταν, ότι δεν υπήρχε μία τεχνική κατάλληλη για την εκπαίδευση τέτοιων δικτύων. Ο αλγόριθμος της *Πίσω Διάδοσης (Backpropagation)* που προτάθηκε από τους Rumelhart και McClelland χρησιμοποιήθηκε με επιτυχία για την εκπαίδευση των πολυεπίπεδων δικτύων.



Σχήμα 2.7: Νευρωνικό δίκτυο δύο επιπέδων

Τα πολυεπίπεδα δίκτυα μπορούν να σχηματιστούν από μία ομάδα επιπέδων κατάλληλα συνδεδεμένων μεταξύ τους. Οι έξοδοι ενός επιπέδου αποτελούν τις εισόδους του επόμενου επιπέδου. Στο σχήμα (2.7) εμφανίζεται ένα τέτοιο δίκτυο, το οποίο είναι δίκτυο με *πλήρη σύνδεση* (*fully connected network*), εφόσον οι νευρώνες ενός επιπέδου είναι συνδεδεμένοι με όλους τους νευρώνες του επόμενου και του προηγούμενου επιπέδου. Τα επίπεδα των νευρώνων που βρίσκονται μεταξύ των επιπέδων εισόδου και εξόδου ονομάζονται *κρυμμένα επίπεδα* (*hidden layers*).

2.2.4.3 Μη γραμμικές συναρτήσεις μεταφοράς

Η αύξηση των δυνατοτήτων των πολυεπίπεδων δικτύων, έναντι των δικτύων του ενός επιπέδου επιτυγχάνεται με την χρησιμοποίηση μη γραμμικής συνάρτησης μεταφοράς στην παραγωγή των εξόδων στα κρυμμένα επίπεδα. Για να γίνει κατανοητή η διαφορά ας θεωρηθεί η περίπτωση ενός δικτύου δύο επιπέδων δηλαδή ενός κρυφού και ενός επιπέδου εξόδου. Ο υπολογισμός της εξόδου ενός δικτύου δύο επιπέδων πραγματοποιείται σε δύο βήματα: **(i)** υπολογίζεται η έξοδος του πρώτου επιπέδου, που επιτυγχάνεται με τον πολλαπλασιασμό του διανύσματος εισόδου με το αντίστοιχο μητρώο συναπτικών βαρών και **(ii)** στην συνέχεια πραγματοποιείται ο υπολογισμός της εξόδου του δεύτερου επιπέδου που είναι η έξοδος του δικτύου συνολικά. Στην περίπτωση που δεν υπάρχει κάποια μη γραμμική συνάρτηση μεταφοράς πολλαπλασιάζεται η έξοδος του κρυφού επιπέδου με το δεύτερο μητρώο βαρών. Αυτό μπορεί να εκφραστεί ως εξής:

$$(\mathbf{X}^T \mathbf{W}_1) \mathbf{W}_2 \quad (2.9)$$

όπου \mathbf{X} είναι το διάνυσμα εισόδου, ενώ \mathbf{W}_1 και \mathbf{W}_2 είναι τα μητρώα βαρών του κρυφού και του επιπέδου εξόδου αντίστοιχα. Ο παραπάνω πολλαπλασιασμός μπορεί να γραφτεί, γνωρίζοντας ότι ισχύει η προσεταιριστική ιδιότητα μεταξύ των μητρώων, με την σχέση

$$\mathbf{X}^T (\mathbf{W}_1 \mathbf{W}_2) \quad (2.10)$$

Η σχέση αυτή αποδεικνύει ότι ένα δίκτυο δύο επιπέδων είναι ισοδύναμο με ένα δίκτυο ενός επιπέδου, το οποίο θα έχει ως μητρώο βαρών το γινόμενο των μητρώων βαρών του δικτύου με τα δύο επίπεδα. Έτσι, κάθε πολυεπίπεδο δίκτυο με γραμμική συνάρτηση μεταφοράς μπορεί να αντικατασταθεί από ένα ισοδύναμο δίκτυο ενός επιπέδου. Επομένως οι υπολογιστικές ικανότητες ενός δικτύου ενός επιπέδου είναι περιορισμένες. Για να επεκταθούν οι δυνατότητες των νευρωνικών δικτύων πέρα από

αυτές των δικτύων ενός επιπέδου γίνεται χρήση μη γραμμικών συναρτήσεων μεταφοράς.

2.2.5 Δίκτυα με επιστροφή (Recurrent networks)

Τα δίκτυα που είδαμε μέχρι αυτό το σημείο δεν έχουν συνδέσεις ανατροφοδότησης δηλαδή συνδέσεις μέσω βαρών που εκτείνονται από τις εξόδους ενός επιπέδου στις εισόδους του ίδιου ή του προηγούμενου επιπέδου. Τα δίκτυα με την αρχιτεκτονική αυτού του είδους ονομάζονται δίκτυα με επιστροφή. Το πιο απλό μοντέλο τέτοιας αρχιτεκτονικής είναι το δίκτυο Hopfield [27].

2.2.6 Η εκπαίδευση των τεχνητών νευρωνικών δικτύων

2.2.6.1 Αντικείμενο της εκπαίδευσης

Ένα δίκτυο εκπαιδεύεται, έτσι ώστε η παρουσίαση ενός νέου συνόλου εισόδων να παράγει ένα επιθυμητό ή τουλάχιστον συνεπές σύνολο εξόδων. Η εκπαίδευση επιτυγχάνεται με συνεχή εφαρμογή διανυσμάτων εισόδου και έλεγχο των διανυσμάτων εξόδου, ενώ οι τιμές των συναπτικών βαρών του δικτύου διορθώνονται σύμφωνα με μια προκαθορισμένη διαδικασία. Κατά τη διάρκεια της εκπαίδευσης τα βάρη του δικτύου βαθμιαίως συγκλίνουν σε σταθερές τιμές, έτσι ώστε κάθε διάνυσμα εισόδου να παράγει το αντίστοιχο επιθυμητό διάνυσμα εξόδου. Οι αλγόριθμοι εκπαίδευσης χωρίζονται σε δύο κατηγορίες: **(i)** στους αλγορίθμους με επίβλεψη και **(ii)** στους αλγορίθμους χωρίς επίβλεψη.

2.2.6.2 Εκπαίδευση με επίβλεψη (Supervised training)

Η εκπαίδευση με επίβλεψη απαιτεί την ύπαρξη ενός συνόλου ζευγών διανυσμάτων εισόδου και εξόδου, εκ των οποίων τα διανύσματα εξόδου ονομάζονται και *διανύσματα στόχοι* (*target vectors*) και αποτελούν τις επιθυμητές εξόδους του δικτύου. Ένας απλός εκπαιδευτικός τύπος ή *ζεύγος εκπαίδευσης* (*training pair*) αποτελείται από δύο διανύσματα γραμμής, τα οποία περιέχουν τις τιμές των παραμέτρων εισόδου και εξόδου, αντίστοιχα. Κάθε ένα ζεύγος εκπαίδευσης είναι υποσύνολο του συνολικού μητρώου εκπαίδευσης. Συνήθως ένα δίκτυο εκπαιδεύεται πάνω σ' ένα αριθμό τέτοιων ζευγών εκπαίδευσης. Το σύνολο όλων αυτών των ζευγών, πάνω στο οποίο εκπαιδεύεται το δίκτυο, ονομάζεται *σύνολο ζευγών εκπαίδευσης* (*training set*).

Τα στάδια της εκπαίδευσης με επίβλεψη είναι τα ακόλουθα: **(i)** Επίδραση ενός διανύσματος εισόδου στο δίκτυο. **(ii)** Υπολογισμός της εξόδου του δικτύου και σύγκριση με το αντίστοιχο διάνυσμα στόχου. **(iii)** Επιστροφή στο δίκτυο του λάθους πρόβλεψης. **(iv)** Διόρθωση των βαρών του δικτύου σύμφωνα με ένα αλγόριθμο ο

οποίος στοχεύει στη ελαχιστοποίηση του σφάλματος πρόβλεψης. Τα ζεύγη εκπαίδευσης εφαρμόζονται διαδοχικά και για κάθε ένα από αυτά υπολογίζεται το σφάλμα. Τα συναπτικά βάρη διορθώνονται με βάση το συνολικό σφάλμα του δικτύου για το σύνολο των ζευγών εκπαίδευσης. Η διαδικασία διόρθωσης των βαρών συνεχίζεται μέχρι το λάθος, για όλα τα μέλη του συνόλου εκπαίδευσης, να φτάσει σε ένα αποδεκτά χαμηλό επίπεδο.

2.2.6.3 Εκπαίδευση χωρίς επίβλεψη (Unsupervised training)

Η εκπαίδευση με επίβλεψη, παρά την εφαρμογή της σε πολλές περιπτώσεις, δεν έχει το βιολογικό ανάλογο. Είναι δύσκολο να εννοηθεί μηχανισμός εκμάθησης στον εγκέφαλο που θα συγκρίνει τις επιθυμητές με τις τρέχουσες εξόδους. Η χωρίς επίβλεψη εκπαίδευση που είναι μία καλύτερη προσέγγιση του βιολογικού αναλόγου, αναπτύχθηκε από τον Kohonen το 1984 [28] και στη συνέχεια εξελίχθηκε από πολλούς άλλους ερευνητές. Αυτού του είδους η εκπαίδευση δεν απαιτεί κάποιο διάνυσμα στόχου για την σύγκριση του με τις προβλεπόμενες τιμές. Το σύνολο των ζευγών εκπαίδευσης αποτελείται μόνο από τα διανύσματα εισόδου. Οι αλγόριθμοι εκπαίδευσης διορθώνουν τα βάρη του δικτύου ώστε να παράγουν διανύσματα εξόδων που να είναι “συνεπή”. Δηλαδή η εφαρμογή δύο παρόμοιων διανυσμάτων εισόδου να παράγουν παρόμοια διανύσματα εξόδου.

2.2.6.4 Αλγόριθμοι εκπαίδευσης

Η εργασία του D. O. Hebb [29] το 1961 αποτελεί την βάση στην οποία στηρίχθηκαν όλοι οι αλγόριθμοι που αναπτύχθηκαν αργότερα. Πριν από την εργασία του Hebb είχε αναγνωριστεί ότι, η διαδικασία εκπαίδευσης σε ένα βιολογικό σύστημα εμπεριέχει κάποιες αλλαγές στην φυσιολογία των νευρώνων, αλλά δεν υπήρχε κάποια ξεκάθαρη ιδέα για το τι διεργασίες πραγματοποιούνται μετά την φάση εκπαίδευσης. Βασισμένος σε έρευνα σε βιολογικό επίπεδο ο Hebb παρουσίασε μια τεχνική για την εκμάθηση ενός συνόλου βιολογικών νευρώνων.

Σύμφωνα με την τεχνική αυτή, τα βάρη αυξάνονται στην περίπτωση που οι νευρώνες πηγής και κατάληξης διεγείρονται ταυτόχρονα. Συνεπώς ένα τεχνητό νευρωνικό δίκτυο εκπαιδευμένο βάση του αλγορίθμου Hebb θα αυξάνει τα βάρη του δικτύου ανάλογα με το βαθμό διέγερσης των νευρώνων πηγής και κατάληξης.

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \eta \cdot OUT_i \cdot OUT_j \quad (2.11)$$

όπου: $w_{ij}^{(t)}$ είναι η τιμή του βάρους από τον νευρώνα i στον νευρώνα j πριν τη διόρθωση, $w_{ij}^{(t+1)}$ είναι η τιμή του βάρους από τον νευρώνα i στον νευρώνα j μετά τη

διόρθωση, η είναι ο συντελεστής εκπαίδευσης, \mathbf{OUT}_i είναι η έξοδος του νευρώνα i και είσοδος του νευρώνα j και \mathbf{OUT}_j είναι η έξοδος του νευρώνα j .

Γενικά η ποικιλία αλγορίθμων εκπαίδευσης που χρησιμοποιούνται σήμερα είναι μεγάλη. Υπάρχουν αλγόριθμοι για επιβλεπόμενη και αλγόριθμοι για μη επιβλεπόμενη εκπαίδευση. Στο επόμενο εδάφιο θα παρουσιαστεί ο αλγόριθμος *Πίσω Διάδοσης* (*Backpropagation*), ο οποίος αποτελεί τον βασικό αλγόριθμο επιβλεπόμενης εκπαίδευσης.

2.3 Ο Αλγόριθμος εκπαίδευσης Πίσω Διάδοσης

Στα πρώτα στάδια εφαρμογής των NN, τους ερευνητές δεν τους απασχολούσε η ανάπτυξη αλγορίθμων εκπαίδευσης πολυεπίπεδων νευρωνικών δικτύων. Όταν όμως διαπιστώθηκε ότι τα δίκτυα ενός επιπέδου είχαν περιορισμένες δυνατότητες τότε το ενδιαφέρον επεκτάθηκε και προς τα πολυεπίπεδα δίκτυα. Η ανακάλυψη του αλγορίθμου της *Πίσω Διάδοσης* (*Backpropagation*) έπαιξε σημαντικό ρόλο στην ενδυνάμωση του ενδιαφέροντος για τα τεχνητά νευρωνικά δίκτυα, γιατί ο αλγόριθμος της πίσω διάδοσης κατόρθωσε να διευρύνει το πεδίο εφαρμογής των τεχνητών νευρωνικών δικτύων. Ο αλγόριθμος πίσω διάδοσης αποτελεί μία μέθοδο εκπαίδευσης πολυεπίπεδων τεχνητών νευρωνικών δικτύων πρόσθιας τροφοδότησης με επίβλεψη. Η εκπαίδευση μπορεί να θεωρηθεί ως ένα πρόβλημα ελαχιστοποίησης χωρίς περιορισμούς όπου τα βάρη είναι οι παράμετροι σχεδιασμού, ενώ η συνάρτηση σφάλματος είναι η αντικειμενική συνάρτηση. Η μέθοδος της Απότομης Καθόδου (*Steepest Descent*) είναι η πλέον ευρέως χρησιμοποιούμενη μέθοδο που χρησιμοποιείται από τον αλγόριθμο της πίσω διάδοσης.

2.3.1 Αριθμητικές μέθοδοι βέλτιστου σχεδιασμού

Στην διεθνή βιβλιογραφία έχει παρουσιαστεί ένα μεγάλο πλήθος αλγορίθμων για την αντιμετώπιση προβλημάτων βελτιστοποίησης χωρίς περιορισμούς. Οι μέθοδοι αυτές στην πλειονότητα τους χρησιμοποιούν την διανυσματική κλίση (*gradient*) της αντικειμενικής συνάρτησης και βασίζονται στην μικρού μεγέθους διόρθωση του τρέχοντος διανύσματος των βαρών του δικτύου με τον ακόλουθο τρόπο: $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \Delta\mathbf{w}^{(k)}$, $k=0,1,2,\dots$, όπου \mathbf{k} είναι το βήμα της επανάληψης, $\mathbf{w}^{(0)}$ είναι η αρχική επιλογή, ενώ $\Delta\mathbf{w}^{(k)}$ είναι η διόρθωση της τρέχουσας κατάστασης. Ως *τρέχουσα κατάσταση του δικτύου* ονομάζεται η κατάσταση, στην οποία αντιστοιχούν οι τρέχουσες τιμές των βαρών στις οποίες αναλογεί ένα επίπεδο σφάλματος και σε ισοδυναμία με την ορολογία που χρησιμοποιείται στην επιστημονική περιοχή του βέλτιστου σχεδιασμού μπορεί να ονομαστεί ως τρέχων σχεδιασμός. Στην συνέχεια της παραγράφου θα γίνει

μία περισσότερο λεπτομερής περιγραφή των σημαντικότερων αλγορίθμων αυτής της κατηγορίας.

2.3.1.1 Η μέθοδος της Απότομης Καθόδου (Steepest Descent method)

Η μέθοδος της Απότομης Καθόδου (Steepest Descent) είναι η πρώτη που βασίστηκε στη διανυσματική κλίση και εφαρμόστηκε για τη βελτιστοποίηση προβλημάτων χωρίς περιορισμούς. Η μέθοδος προτάθηκε από τον Cauchy το 1847 [30] και βασίζεται στην εύρεση της διεύθυνσης \mathbf{d} , στην οποία η συνάρτηση σφάλματος $\mathbf{E}(\mathbf{w})$ μειώνεται ταχύτερα, σε τοπικό επίπεδο (διεύθυνση απότομης καθόδου). Στην μέθοδο αυτή η διεύθυνση της μέγιστης μείωσης της συνάρτησης σφάλματος ορίζεται με την αρνητική της κλίση στη θέση \mathbf{w} . Η μέθοδος μπορεί να περιγραφεί με τα ακόλουθα βήματα:

Βήμα 1^ο: Γίνεται εκτίμηση ενός αρχικού σχεδιασμού $\mathbf{w}^{(0)}$. Στην περίπτωση των βαρών ενός δικτύου η παραγωγή της αρχικής κατάστασης γίνεται με τυχαίο τρόπο, ο δείκτης των επαναλήψεων τίθεται στο μηδέν $\mathbf{k} = 0$, και επιλέγεται μία παράμετρος ελέγχου της σύγκλισης $\epsilon > 0$, καθώς και ένα επίπεδο ανοχής σφάλματος στην πρόβλεψη.

Βήμα 2^ο: Υπολογίζεται η κλίση της συνάρτησης σφάλματος $E(\mathbf{w}^{(k)})$ με την ακόλουθη σχέση $\mathbf{c}^{(k)} = \nabla E(\mathbf{w}^{(k)})$ και στη συνέχεια το μέτρο αυτής $\|\mathbf{c}^{(k)}\|$. Αν ισχύει $\|\mathbf{c}^{(k)}\| < \epsilon$ τότε η διαδικασία τερματίζεται και θεωρείται ως βέλτιστος σχεδιασμός ο $\mathbf{w}^* = \mathbf{w}^{(k)}$, διαφορετικά η διαδικασία συνεχίζεται με το επόμενο *Βήμα*.

Βήμα 3^ο: Υπολογίζεται η διεύθυνση της απότομης καθόδου στην θέση $\mathbf{w}^{(k)}$, $\mathbf{d}^{(k)} = -\mathbf{c}^{(k)}$.

Βήμα 4^ο: Υπολογίζεται το νέο διάνυσμα $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha_k \cdot \mathbf{d}^{(k)}$, όπου α_k αποτελεί το μήκος βήματος. Η διαδικασία επιστρέφει στο *Βήμα 2*, αφού τεθεί $k = k+1$.

Το κριτήριο τερματισμού στο *Βήμα 2* αποτελεί ένα ικανό κριτήριο βελτιστοποίησης κυρτών συναρτήσεων. Στην περίπτωση μη - κυρτών συναρτήσεων, δηλαδή συναρτήσεων με πολλά τοπικά ελάχιστα, η συνθήκη τερματισμού ικανοποιείται σε κάθε ένα από τα τοπικά ελάχιστα του χώρου αναζήτησης. Η συνάρτηση σφάλματος είναι μία συνάρτηση που γενικά παρουσιάζει πολλά τοπικά ελάχιστα, συνεπώς το κριτήριο τερματισμού $\|\mathbf{c}^{(k)}\|$, δεν είναι αποτελεσματικό. Στην περίπτωση εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου τα κριτήρια τερματισμού που χρησιμοποιούνται είναι: **(i)** το σφάλμα της τετραγωνικής ρίζας των μέσων τετραγώνων και **(ii)** ο βαθμός διόρθωσης, ο οποίος έχει σχέση με το επίπεδο σφάλματος που αποδεχόμαστε στην πρόβλεψη του δικτύου. Η πρόβλεψη αυτή έχει σχέση με το κάθε πρότυπο του συνόλου εκπαίδευσης ξεχωριστά. Λεπτομερέστερη περιγραφή των δύο κριτηρίων τερματισμού θα γίνει σε επόμενη παράγραφο του κεφαλαίου.

2.3.1.2 Η μέθοδος της Συζυγούς Διανυσματικής Κλίσεως (Conjugate Gradient method)

Η μέθοδος της Συζυγούς Διανυσματικής Κλίσεως παρουσιάστηκε από τους Fletcher και Reeves [31] και προέκυψε από μία απλή αλλά αποτελεσματική τροποποίηση της μεθόδου της Απότομης Καθόδου. Η διαφορά των δύο τεχνικών έγκειται στην ορθογωνικότητα των κατευθύνσεων των διαδοχικών επαναλήψεων. Στην περίπτωση της μεθόδου της Απότομης Καθόδου οι διευθύνσεις είναι ορθογωνικές μεταξύ τους, ενώ στην περίπτωση της μεθόδου της συζυγούς Διανυσματικής Κλίσεως οι διευθύνσεις δεν είναι ορθογωνικές. Η μέθοδος μπορεί να περιγραφεί με τα ακόλουθα βήματα:

Βήμα 1^ο: Γίνεται εκτίμηση ενός αρχικού σχεδιασμού $\mathbf{w}^{(0)}$. Στην περίπτωση των βαρών ενός δικτύου η παραγωγή της αρχικής κατάστασης γίνεται με τυχαίο τρόπο. Ο δείκτης των επαναλήψεων τίθεται στο μηδέν $\mathbf{k} = 0$, εκλέγεται μία παράμετρος ελέγχου της σύγκλισης $\epsilon > 0$, καθώς και ένα επίπεδο ανοχής σφάλματος στην πρόβλεψη. Υπολογίζεται το αρχικό διάνυσμα διεύθυνσης $\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} = -\nabla E(\mathbf{w}^{(0)})$, και το μέτρο $\|\mathbf{c}^{(k)}\|$. Στην περίπτωση που ισχύει $\|\mathbf{c}^{(k)}\| < \epsilon$, η διαδικασία τερματίζεται και θεωρείται ως βέλτιστος σχεδιασμός ο $\mathbf{w}^* = \mathbf{w}^{(k)}$, διαφορετικά η διαδικασία συνεχίζει με το Βήμα 4.

Βήμα 2^ο: Υπολογίζεται η κλίση της συνάρτησης σφάλματος $E(\mathbf{w}^{(k)})$ $\mathbf{c}^{(k)} = \nabla E(\mathbf{w}^{(k)})$ και στη συνέχεια το μέτρο $\|\mathbf{c}^{(k)}\|$. Στην περίπτωση που ισχύει $\|\mathbf{c}^{(k)}\| < \epsilon$, η διαδικασία τερματίζεται και θεωρείται ως βέλτιστος σχεδιασμός ο $\mathbf{w}^* = \mathbf{w}^{(k)}$, διαφορετικά η διαδικασία συνεχίζεται με το επόμενο Βήμα.

Βήμα 3^ο: Υπολογίζεται η νέα συζυγής διεύθυνση αναζήτησης στην θέση $\mathbf{w}^{(k)}$:

$$\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta_k \mathbf{d}^{(k-1)} \quad (2.12)$$

όπου

$$\beta_k = \begin{cases} \frac{\nabla E_{k+1} \cdot (\nabla E_{k+1} - \nabla E_k) / \mathbf{d}^{(k)} \cdot (\nabla E_k - \nabla E_{k+1})}{\nabla E_{k+1} \cdot (\nabla E_{k+1} - \nabla E_k) / \nabla E_k \cdot \nabla E_k} & \text{Hestenes-Stiefel} \\ \frac{\nabla E_{k+1} \cdot (\nabla E_{k+1} - \nabla E_k) / \nabla E_k \cdot \nabla E_k}{\nabla E_{k+1} \cdot \nabla E_{k+1} / \nabla E_k \cdot \nabla E_k} & \text{Polak-Ribiere} \\ \frac{\nabla E_{k+1} \cdot \nabla E_{k+1} / \nabla E_k \cdot \nabla E_k}{\nabla E_{k+1} \cdot \nabla E_{k+1} / \nabla E_k \cdot \nabla E_k} & \text{Fletcher-Reeves} \end{cases} \quad (2.13)$$

Βήμα 4^ο: Υπολογίζεται το νέο διάνυσμα $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha_k \mathbf{d}^{(k)}$, όπου α_k αποτελεί το μήκος βήματος. Η διαδικασία επιστρέφει στο Βήμα 2, αφού τεθεί $\mathbf{k} = \mathbf{k} + 1$.

Η διαφορά των δύο μεθόδων βρίσκεται στο Βήμα 3 με την εξίσωση (2.12). Στο βήμα αυτό, η τρέχουσα διεύθυνση της μεθόδου της απότομης καθόδου τροποποιείται με την προσθήκη ενός όρου, που αναλογεί στο βήμα διόρθωσης και χρησιμοποιείται στην

προηγούμενη επανάληψη. Παρά την, φαινομενικά απλή, τροποποίηση αυτή είναι ιδιαίτερα αποτελεσματική όσον αφορά την ταχύτητα σύγκλισης. Στην περίπτωση χρησιμοποίησης της μεθόδου για την εκπαίδευση νευρωνικού δικτύου τα κριτήρια σύγκλισης αντικαθίστανται από εκείνα που περιγράφησαν στην πρώτη μέθοδο.

2.3.1.3 Η μέθοδος Newton (Newton's method)

Στην περίπτωση των μεθόδων της απότομης καθόδου και της συζυγούς διανυσματικής κλίσεως χρησιμοποιείται πληροφορία πρώτης τάξης για την απεικόνιση της μορφής της συνάρτησης σφάλματος στην τρέχουσα θέση και του υπολογισμού της νέας διεύθυνσης αναζήτησης. Αν χρησιμοποιηθούν παράγωγοι δεύτερης τάξης γίνεται δυνατή η ακριβέστερη απεικόνιση της επιφάνειας σφάλματος και συνεπώς εντοπίζεται καλύτερη διεύθυνση αναζήτησης. Μια τέτοια τεχνική υιοθετείται στην μέθοδο Newton, στην οποία απαιτείται ο υπολογισμός του μητρώου Hessian της συνάρτησης σφάλματος σε κάθε επανάληψη. Το μητρώου Hessian είναι ένα συμμετρικό μητρώο. Απαιτούνται $n(n+1)/2$ παραγωγίσεις δεύτερης τάξης της συνάρτησης σφάλματος όπου n είναι το πλήθος των βαρών του δικτύου. Η μέθοδος μπορεί να περιγραφεί με τα ακόλουθα βήματα:

Βήμα 1^ο: Γίνεται εκτίμηση ενός αρχικού σχεδιασμού $\mathbf{w}^{(0)}$. Στην περίπτωση των βαρών ενός δικτύου η παραγωγή της αρχικής κατάστασης γίνεται με τυχαίο τρόπο. Ο δείκτης των επαναλήψεων τίθεται στο μηδέν $\mathbf{k} = 0$, εκλέγεται μία παράμετρος ελέγχου της σύγκλισης $\epsilon > 0$, καθώς και ένα επίπεδο ανοχής σφάλματος στην πρόβλεψη.

Βήμα 2^ο: Υπολογίζεται η κλίση της συνάρτησης σφάλματος $E(\mathbf{w}^{(k)})$ $\mathbf{c}^{(k)} = \nabla E(\mathbf{w}^{(k)})$, και στη συνέχεια το μέτρο $\|\mathbf{c}^{(k)}\|$. Στην περίπτωση που ισχύει $\|\mathbf{c}^{(k)}\| < \epsilon$, η διαδικασία τερματίζεται και θεωρείται ως βέλτιστος σχεδιασμός ο $\mathbf{w}^* = \mathbf{w}^{(k)}$, διαφορετικά η διαδικασία συνεχίζει με το επόμενο Βήμα.

Βήμα 3^ο: Υπολογίζεται το Hessian μητρώο από την σχέση

$$\mathbf{H}(\mathbf{w}^{(k)}) = \left[\frac{\partial^2 E(\mathbf{w}^{(k)})}{\partial w_i \partial w_j} \right], i = 1, 2, \dots, n; j = 1, 2, \dots, n \quad (2.14)$$

Βήμα 4^ο: Υπολογίζεται η νέα διεύθυνση αναζήτησης στην θέση $\mathbf{w}^{(k)}$, $\mathbf{d}^{(k)} = -\mathbf{H}^{-1} \mathbf{c}^{(k)}$. Ο υπολογισμός του διανύσματος $\mathbf{d}^{(k)}$ γίνεται απευθείας με την επίλυση ενός συστήματος γραμμικών εξισώσεων χωρίς να απαιτείται ο υπολογισμός του αντιστρόφου Hessian μητρώου.

Βήμα 5^ο: Υπολογίζεται το νέο διάνυσμα $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha_k \cdot \mathbf{d}^{(k)}$, όπου α_k αποτελεί το μήκος βήματος. Η διαδικασία επιστρέφει στο *Βήμα 2*, αφού τεθεί $\mathbf{k} = \mathbf{k}+1$.

Το πλεονέκτημα της μεθόδου είναι η τετραγωνική σύγκλιση, η οποία επιτυγχάνεται, ενώ το μειονέκτημα της μεθόδου είναι ο χρονοβόρος υπολογισμός των παραγώγων δεύτερης τάξης σε κάθε επανάληψη της μεθόδου. Για την αντιμετώπιση αυτού προτάθηκε μία τροποποίηση από τον Marquardt η οποία παρουσιάζεται στην συνέχεια.

2.3.1.4 Η μέθοδος των Marquardt-Levenberg

Ο Marquardt [32] πρότεινε μία τροποποίηση της μεθόδου Newton, σύμφωνα με την οποία η μέθοδος μακριά από την λύση συμπεριφέρεται σαν την μέθοδο της Απότομης Καθόδου, ενώ κοντά στην λύση σαν την μέθοδο Newton, η οποία παρουσιάζεται ιδιαίτερα αποτελεσματική στην περιοχή της λύσης. Η μέθοδος μπορεί να περιγραφεί με τα ακόλουθα βήματα:

Βήμα 1^ο: Γίνεται εκτίμηση ενός αρχικού σχεδιασμού $\mathbf{w}^{(0)}$. Στην περίπτωση των βαρών ενός δικτύου η παραγωγή της αρχικής κατάστασης γίνεται με τυχαίο τρόπο. Ο δείκτης των επαναλήψεων τίθεται στο μηδέν $\mathbf{k} = 0$, εκλέγεται μία παράμετρος ελέγχου της σύγκλισης $\epsilon > 0$, καθώς και ένα επίπεδο ανοχής σφάλματος στην πρόβλεψη. Επίσης επιλέγεται και μία σταθερά λ_0 με μεγάλη τιμή (π.χ. $1.0E+04$).

Βήμα 2^ο: Υπολογίζεται η κλίση της συνάρτησης σφάλματος $E(\mathbf{w}^{(k)})$ $\mathbf{c}^{(k)} = \nabla E(\mathbf{w}^{(k)})$, υπολογίζεται το μέτρο $\|\mathbf{c}^{(k)}\|$. Στην περίπτωση που ισχύει $\|\mathbf{c}^{(k)}\| < \epsilon$, η διαδικασία τερματίζεται και θεωρείται ως βέλτιστος σχεδιασμός ο $\mathbf{w}^* = \mathbf{w}^{(k)}$, διαφορετικά η διαδικασία συνεχίζει με το επόμενο *Βήμα*.

Βήμα 3^ο: Υπολογίζεται το Hessian μητρώο από την εξίσωση (2.14).

Βήμα 4^ο: Υπολογίζεται η νέα διεύθυνση αναζήτησης στην θέση $\mathbf{w}^{(k)}$, $\mathbf{d}^{(k)} = -(\mathbf{H} + \lambda_k \mathbf{I})^{-1} \mathbf{c}^{(k)}$.

Βήμα 5^ο: Στην περίπτωση που $E(\mathbf{w}^{(k)} + \mathbf{d}^{(k)}) < E(\mathbf{w}^{(k)})$ τότε η διαδικασία συνεχίζεται με το επόμενο *Βήμα*. Διαφορετικά τίθεται $\lambda_k = 2\lambda_k$ και η διαδικασία επιστρέφει στο *Βήμα 4*.

Βήμα 6^ο: Υπολογίζεται το νέο διάνυσμα $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \mathbf{d}^{(k)}$. Η διαδικασία επιστρέφει στο *Βήμα 2*, αφού τεθεί $\lambda_{k+1} = 0.5 \cdot \lambda_k$ και $\mathbf{k} = \mathbf{k}+1$.

2.3.1.5 Τοπικό ελάχιστο

Η μέθοδος της απότομης καθόδου ακολουθεί την κλίση της επιφάνειας της συνάρτησης σφάλματος προς την διεύθυνση της μέγιστης μείωσης, επιδιώκοντας με τη διόρθωση των βαρών να εντοπίσει το ελάχιστο δυνατό σφάλμα πρόβλεψης. Τόσο η μέθοδος της

απότομης καθόδου όσο και οι υπόλοιπες μέθοδοι που βασίζονται στη διανυσματική κλίση της αντικειμενικής συνάρτησης, παρουσιάζουν καλή συμπεριφορά στις περιπτώσεις που η επιφάνεια σφάλματος είναι κυρτή (convex), δηλαδή εμφανίζει μοναδικό ελάχιστο. Όμως η επιφάνεια σφάλματος ενός πολύπλοκου δικτύου εμφανίζει πολλές περιοχές τοπικών ελαχίστων και ιδιαίτερα στον χώρο των πολλών διαστάσεων, δηλαδή δικτύων με πολλούς κόμβους στα κρυμμένα επίπεδα. Οπότε ενδέχεται ο αλγόριθμος εκπαίδευσης να παγιδευτεί σε μία από τις περιοχές αυτές. Σε πολλές περιπτώσεις ένα τοπικό ελάχιστο είναι δυνατό να αποτελέσει αποδεκτή λύση για τη φάση εκπαίδευσης του δικτύου, αφού δεν είναι δυνατό να γνωρίζουμε το είδος του βέλτιστου σημείου στο οποίο έχει συγκλίνει ο αλγόριθμος της πίσω διάδοσης. Πολλές φορές επιβάλλεται, στην περίπτωση που εντοπιστεί κάποια λύση που δεν είναι αποδεκτή, να επαναλαμβάνεται η διαδικασία εκπαίδευσης για νέες τυχαίες τιμές των βαρών, αλλά και πάλι δεν υπάρχει καμία εγγύηση ότι θα εντοπιστεί μία αποδεκτή λύση.

Δυνατότητα “απόδρασης” από περιοχές τοπικών ελαχίστων έχουν οι στοχαστικές μέθοδοι εκπαίδευσης οι οποίες όμως είναι σχετικά αργές. Ο Wasserman (1988) [33] πρότεινε μία μέθοδο η οποία συνδυάζει την στοχαστική μέθοδο της *Μηχανής Cauchy* με τον αλγόριθμο της απότομης καθόδου σε μια προσπάθεια να παράγει έναν αξιόπιστο αλγόριθμο που να επιτυγχάνει τον εντοπισμό του καθολικού ελαχίστου.

2.3.1.6 Μέγεθος βήματος

Σε κάποιο από τα βήματα τόσο της μεθόδου της απότομης καθόδου όσο και των υπολοίπων αλγορίθμων βελτιστοποίησης, που αναφέρθηκαν προηγουμένως χρησιμοποιείται η σταθερά α_k , η οποία ελέγχει το πόσο μεγάλο θα είναι το βήμα διόρθωσης σε κάθε επανάληψη. Η επιλογή της παραμέτρου αυτής είναι πολύ σημαντική για την ταχύτητα σύγκλισης του αλγορίθμου εκπαίδευσης. Στην περίπτωση που το μέγεθος βήματος είναι πολύ μικρό η σύγκλιση ενδέχεται να είναι πολύ αργή. Αντίθετα στην περίπτωση που είναι πολύ μεγάλο ενδέχεται να παρουσιαστεί αριθμητική αστάθεια. Η χρήση σταθερού μήκους βήματος προκαλεί, όπως φάνηκε από αριθμητικά παραδείγματα που εξετάστηκαν, πολλά προβλήματα. Μία τιμή βήματος που συμπεριφέρεται καλά για τα πρώτα στάδια της διαδικασίας εκπαίδευσης μπορεί να οδηγήσει σε λάθος δρόμο στα επόμενα στάδια εκπαίδευσης. Για τον λόγο αυτό, καλό είναι ο χειριστής του προγράμματος εκπαίδευσης να παρακολουθεί την διαδικασία, ώστε να επεμβαίνει όταν χρειάζεται στην περίπτωση επιλογής σταθερού βήματος. Ο Wasserman το 1988 [34] περιέγραψε ένα αλγόριθμο, βάση του οποίου το μέγεθος του βήματος διορθώνεται με αυτόματο τρόπο, καθώς η διαδικασία εκπαίδευσης προχωράει. Συνήθως στην έναρξη της διαδικασίας εκπαίδευσης εκλέγεται η τιμή 0.5

και ανάλογα με την συμπεριφορά της εκπαίδευσης το μέγεθος βήματος αυξάνεται ή μειώνεται [14].

2.3.2 Ο αλγόριθμος βελτιστοποίησης RPROP

Η μέθοδος βελτιστοποίησης RPROP [35], που χρησιμοποιείται στα νευρωνικά δίκτυα *Πίσω Διάδοσης* (*Backpropagation*) για τη βελτιστοποίηση της συνάρτησης σφάλματος, αποτελεί ένα νέο αποτελεσματικό σύστημα που βοηθά τη διαδικασία εκπαίδευσης του NN, διότι εκτελεί την άμεση προσαρμογή του βήματος του βάρους και βασίζεται στις πληροφορίες της τοπικής κλίσης. Η βασική διαφορά της, από τις παραπάνω μεθόδους που αναφέρθηκαν, είναι ότι η διαδικασία προσαρμογής των βαρών δεν εμποδίζεται από την συμπεριφορά της διανυσματικής κλίσης.

Για να επιτευχθεί αυτό, εισάγεται για κάθε βάρος χωριστά, μία μεμονωμένη περιοδικά ανανεωμένη τιμή (update – value) Δ_{ij} , η οποία καθορίζει απλώς το μέγεθος του ανανεωμένου του βάρους. Αυτή η προσαρμοστικά και περιοδικά ανανεωμένη τιμή, αλλάζει συνεχώς κατά τη διάρκεια της διαδικασίας εκμάθησης, βασισμένη στην τοπική κλίση της συνάρτησης σφάλματος \mathbf{E} , σύμφωνα με τον ακόλουθο κανόνα εκμάθησης:

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ * \Delta_{ij}^{(t-1)} & , \text{ if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ \eta^- * \Delta_{ij}^{(t-1)} & , \text{ if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)} & , \text{ else} \end{cases} \quad (2.15)$$

where $0 < \eta^- < 1 < \eta^+$

Ο κανόνας προσαρμογής των τιμών των βαρών στις νέες, κάθε φορά, τιμές τους, λειτουργεί ως εξής:

Κάθε φορά που η μερική παράγωγος της συνάρτησης σφάλματος \mathbf{E} ως προς το αντίστοιχο βάρος w_{ij} αλλάζει το πρόσημό της, το οποίο υποδηλώνει ότι η τελευταία ανανεωμένη τιμή ήταν πολύ μεγάλη και ο αλγόριθμος έχει υπερπηδήσει ένα τοπικό ελάχιστο, η ανανεωμένη τιμή Δ_{ij} μειώνεται από τον παράγοντα η^- . Διαφορετικά εάν η μερική παράγωγος διατηρεί το πρόσημό της, τότε η ανανεωμένη τιμή Δ_{ij} αυξάνεται ελαφρώς προκειμένου να επιταχυνθεί η σύγκλιση στις σχεδόν επίπεδες περιοχές. Μόλις η ανανεωμένη τιμή για κάθε βάρος προσαρμοστεί, το ανανεωμένο βάρος ακολουθεί τον επόμενο κανόνα:

Εάν η παραγωγός του είναι θετική (αυξανόμενο σφάλμα), το βάρος μειώνεται από την ανανεωμένη τιμή του, ενώ αν η παράγωγος είναι αρνητική, το ανανεωμένο βάρος προστίθεται:

$$\Delta w_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)} & , \text{ if } \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ +\Delta_{ij}^{(t)} & , \text{ if } \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ 0 & , \text{ else} \end{cases} \quad (2.16)$$

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \Delta w_{ij}^{(t)}$$

Εντούτοις, υπάρχει μια εξαίρεση: Εάν η μερική παράγωγος αλλάξει πρόσημο, δηλαδή το προηγούμενο βήμα ήταν πάρα πολύ μεγάλο και το ελάχιστο χάθηκε, τότε επανέρχεται το προηγούμενο ανανεωμένο βάρος. Αυτό εξηγείται από την παρακάτω σχέση:

$$\Delta w_{ij}^{(t)} = -\Delta w_{ij}^{(t-1)} \quad , \quad \text{ if } \frac{\partial E^{(t-1)}}{\partial w_{ij}} * \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \quad (2.17)$$

Εξαιτίας της «υπαναχώρησης» του βήματος του βάρους, η παράγωγος υποτίθεται ότι θα αλλάξει το πρόσημό της μία φορά ακόμα στο επόμενο χρονικό βήμα. Προκειμένου να αποφευχθεί η διπλή απόρριψη της ανανεωμένης τιμής, δεν πρέπει να υπάρξει προσαρμογή της ανανεωμένης τιμής του βάρους στο επόμενο βήμα. Πρακτικά αυτό μπορεί να γίνει με τη ρύθμιση $\frac{\partial E^{(t-1)}}{\partial w_{ij}} := 0$ στον παραπάνω κανόνα προσαρμογής. Οι

ανανεωμένες τιμές και τα βάρη αλλάζουν κάθε φορά που όλα τα ζεύγη δεδομένων παρουσιάζονται μία φορά στο δίκτυο (εκπαίδευση μέσω επαναλήψεων). Η παραπάνω διαδικασία, για όλα τα βάρη και τις παραμέτρους bias, μπορεί να περιγραφεί με τα κάτωθι βήματα:

$$\checkmark \quad \text{Εάν} \left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) > 0 \right) \text{ τότε,}$$

$$\Delta_{ij}(t) = \text{minimum} (\Delta_{ij}(t-1) * \eta^+, \Delta_{\max})$$

$$\Delta w_{ij}(t) = -\text{sin g} \left(\frac{\partial E}{\partial w_{ij}}(t) \right) * \Delta_{ij}(t)$$

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$$

✓ Η εάν $\left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) < 0 \right)$ τότε,

$$\Delta_{ij}(t) = \text{maximum} (\Delta_{ij}(t-1) * \eta^-, \Delta_{\min})$$

$$w_{ij}(t+1) = w_{ij}(t) - \Delta_{ij}(t-1)$$

$$\frac{\partial E}{\partial w_{ij}}(t) = 0$$

✓ Διαφορετικά εάν $\left(\frac{\partial E}{\partial w_{ij}}(t-1) * \frac{\partial E}{\partial w_{ij}}(t) = 0 \right)$ τότε,

$$\Delta_{ij}(t) = -\sin g \left(\frac{\partial E}{\partial w_{ij}}(t) \right) * \Delta_{ij}(t)$$

$$w_{ij}(t+1) = w_{ij}(t) + \Delta_{ij}(t)$$

Κατά την έναρξη της διαδικασίας εκπαίδευσης όλες οι ανανεωμένες τιμές ρυθμίζονται (σετάρονται) σε μία αρχική τιμή Δ_0 . Η τιμή αυτή προσδιορίζει άμεσα το μέγεθος του πρώτου βήματος του βάρους και συνήθως επιλέγεται σε μια λογική αναλογία με το μέγεθος των αρχικών τιμών των βαρών. Μία καλή τιμή για το Δ_0 είναι $\Delta_0 = 0.1$. Επίσης ακόμα και για μεγαλύτερες ή μικρότερες τιμές της Δ_0 η γρήγορη σύγκλιση επιτυγχάνεται.

Το εύρος των ανανεωμένων τιμών περιορίζεται από μια μέγιστη τιμή $\Delta_{\max} = 50$ και από μια ελάχιστη $\Delta_{\min} = 1e^{-6}$, ώστε να αποφευχθούν προβλήματα υπερχειλίσσης / υποχειλίσσης των κυμαινόμενων μεταβλητών. Σε πολλές εφαρμογές για την τιμή του άνω ορίου Δ_{\max} επιλέγονται πολύ μικρότερες τιμές της τιμής 50, όπως 0.1 με την οποία επιτυγχάνεται πιο ομαλή μείωση του σφάλματος.

Η επιλογή των τιμών για τους παράγοντες μείωσης η^- και αύξησης η^+ γίνεται με βάση τις ακόλουθες εκτιμήσεις. Εάν εμφανιστεί η υπερπήδηση ενός τοπικού ελαχίστου, αυτό σημαίνει πως η προηγούμενη ανανεωμένη τιμή ήταν πάρα πολύ μεγάλη. Επειδή οι πληροφορίες της κλίσης δεν δείχνουν πόσο είναι το ελάχιστο που χάθηκε μία καλή σκέψη είναι η διχοτόμηση της ανανεωμένης τιμής. Επομένως $\eta^- = 0.5$. Ο παράγοντας αύξησης από την άλλη πρέπει να είναι αρκετά μεγάλος για να επιτρέπει την γρήγορη αύξηση της ανανεωμένης τιμής στις σχεδόν επίπεδες περιοχές της συνάρτησης σφάλματος, όταν στην άλλη πλευρά η διαδικασία εκμάθησης μπορεί να είστε αρκετά διαταραγμένη, εάν ένας πάρα πολύ μεγάλος παράγοντας αύξησης οδηγήσει σε διαρκείς αλλαγές της κατεύθυνσης του βήματος του βάρους. Από διάφορα παραδείγματα της βιβλιογραφίας έχει δειχθεί ότι η τιμή 1.2 είναι μία καλή τιμή για τον παράγοντα η^+ , η

οποία δίνει πολύ καλά αποτελέσματα, σύμφωνα με το πρόβλημα πάντα. Επίσης έχει αποδειχθεί ότι μικρές διαφορές στην παραπάνω τιμή δεν προκαλεί ούτε βελτίωση, ούτε και επιδείνωση των αποτελεσμάτων.

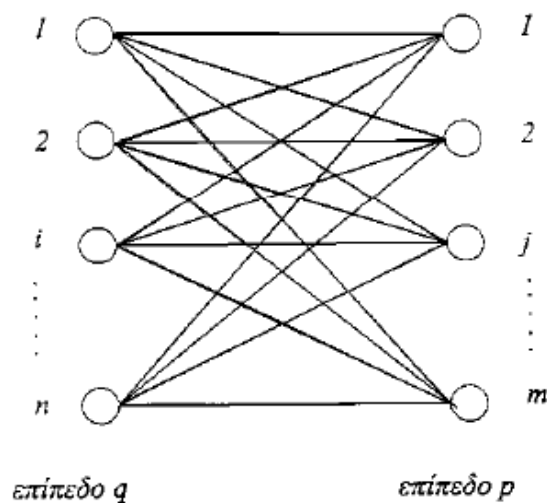
2.3.3 Πίσω Διάδοση

Στον αλγόριθμο της πίσω διάδοσης η εκπαίδευση διεξάγεται με τη διάδοση μέσω του δικτύου ενός συνόλου από εκπαιδευτικούς τύπους εισόδου. Το δίκτυο αυτό αποτελείται από ένα επίπεδο εισόδου, ένα ή περισσότερα κρυμμένα επίπεδα και ένα επίπεδο εξόδου, όπως φαίνεται στο σχήμα (2.7). Κάθε επίπεδο αποτελείται από τις αντίστοιχες μονάδες του (στοιχεία επεξεργασίας, νευρώνες ή κόμβους) και συνδέσεις βάρους.

Οι εισοδοί $x_i, i=1, 2, \dots, n$, οι οποίες γίνονται δεκτές από το επίπεδο εισόδου είναι ανάλογες των ηλεκτροχημικών συστημάτων που δέχονται οι νευρώνες στον ανθρώπινο εγκέφαλο. Στο απλούστερο μοντέλο, οι εισοδοί αυτές πολλαπλασιάζονται με τα βάρη των συνδέσεων και η ενεργή είσοδος $NET_{p,j}$, στα στοιχεία επεξεργασίας είναι το σταθμισμένο άθροισμα των αρχικών εισόδων.

$$net_{p,j} = \sum_{i=1}^n w_{p,ij} net_{q,i} + b_{p,j} \quad (2.18)$$

όπου $w_{p,ij}$ αποτελεί το βάρος σύνδεσης ή το συναπτικό βάρος του επιπέδου p , το οποίο ξεκινάει από τον νευρώνα i του επιπέδου q (πηγή) και καταλήγει στον νευρώνα j του επιπέδου p (στόχος), $NET_{q,i}$ είναι η έξοδος που παράγεται στον νευρώνα i του επιπέδου q , όπως φαίνεται στο σχήμα (2.8), και $b_{p,j}$ είναι μια καθοδηγητική παράμετρος που χρησιμοποιείται για να εξομαλυνθεί (modulate) η έξοδος του στοιχείου. Οι εισοδοί x_i αντιστοιχούν στο $NET_{q,i}$ με $i=1,2,\dots,n$ στην περίπτωση του επιπέδου εισόδου.



Σχήμα 2.8: Συνδεσμολογία μεταξύ δύο διαδοχικών επιπέδων

Στο βιολογικό πρότυπο, ένας τυπικός νευρώνας παράγει ένα σήμα εξόδου μόνο όταν το εισερχόμενο σήμα τον διεγείρει. Στα νευρωνικά δίκτυα αυτή η διαδικασία προκαλείται, όπως αναφέρθηκε σε προηγούμενη παράγραφο του κεφαλαίου, από τις συναρτήσεις μεταφοράς. Η συνάρτηση που συνήθως χρησιμοποιείται είναι η σιγμοειδής συνάρτηση μεταφοράς. Το βασικό πλεονέκτημα της σιγμοειδούς συνάρτησης μεταφοράς είναι η ικανότητά της να χειριστεί και μεγάλο και μικρό μέγεθος σήματα εισόδου. Ο προσδιορισμός των κατάλληλων βαρών και καθοδηγητικών παραμέτρων είναι ενσωματωμένος στην διαδικασία εκπαίδευσης του δικτύου. Η παράγωγος dF της σιγμοειδούς συναρτήσεως δίνεται από τη σχέση:

$$dF(\text{net}_{k,i}) = \text{out}_{k,i}(1 - \text{out}_{k,i}) \quad (2.19)$$

Η σιγμοειδής συνάρτηση συμπιέζει την ακτίνα του σήματος **NET**, έτσι ώστε το αποτέλεσμα **OUT** να βρίσκεται μεταξύ του μηδέν και του ένα. Η συνάρτηση συμπίεσης παράγει την απαιτούμενη μη γραμμικότητα στον αλγόριθμο. Υπάρχουν πολλές συναρτήσεις που θα μπορούσαν να χρησιμοποιηθούν ως συναρτήσεις μεταφοράς για τον αλγόριθμο πίσω διάδοσης. Ο αλγόριθμος απαιτεί η συνάρτηση αυτή να είναι παντού παραγωγίσιμη, γιατί η παράγωγος της συναρτήσεως μεταφοράς χρησιμοποιείται στο βήμα διόρθωσης των βαρών. Η σιγμοειδής συνάρτηση ικανοποιεί αυτήν την απαίτηση και έχει και ένα πρόσθετο πλεονέκτημα παρέχοντας μια μορφή αυτόματου ελέγχου της κλίσης της καμπύλης. Για μικρές τιμές σημάτων η κλίση της καμπύλης εισόδων/εξόδων είναι απότομη, συνεπώς παράγεται μεγάλη κλίση, ενώ η κλίση μικραίνει καθώς το μέγεθος του σήματος γίνεται μεγαλύτερο.

2.3.3.1 Εκπαίδευση

Αντικείμενο της εκπαίδευσης ενός δικτύου είναι ο προσδιορισμός των βαρών του δικτύου έτσι ώστε η εφαρμογή ενός συνόλου εισόδων να παράγει το επιθυμητό σύνολο εξόδων. Στη βασική έκδοση του αλγορίθμου πίσω διάδοσης χρησιμοποιείται ο αλγόριθμος της *απότομης καθόδου* (Steepest Descent). Η τοπική συνάρτηση σφάλματος για ένα τυχαίο p ζεύγος εκπαίδευσης μπορεί να μορφωθεί ως ακολούθως:

$$E_p = \frac{1}{2} \sum_{i=1}^{N_{\text{out}}} (\text{tar}_i^{(p)} - \text{out}_{k,i})^2 = \frac{1}{2} \sum_{i=1}^{N_{\text{out}}} (\text{err}_i^{(p)})^2 \quad (2.20)$$

ενώ η καθολική συνάρτηση σφάλματος είναι η ακόλουθη:

$$E = \sum_{p=1}^{N_p} E_p = \frac{1}{2} \sum_{p=1}^{N_p} \sum_{i=1}^{N_{\text{out}}} (\text{tar}_i^{(p)} - \text{out}_{k,i})^2 \quad (2.21)$$

όπου $\mathbf{tar}_i^{(p)}$ και $\mathbf{out}_{k,i}$ είναι τα αποτελέσματα στόχου και τα αποτελέσματα πρόβλεψης, αντιστοίχως για τον νευρώνα i του επιπέδου εξόδου \mathbf{k} , για το δείγμα εκπαίδευσης \mathbf{p} . Οι τεταγμένες του διανύσματος στόχου βρίσκονται μεταξύ 0 και 1. Κατά συνέπεια, προκειμένου οι εξισώσεις (2.20, 2.21) να παράγουν λογικά αποτελέσματα, οι έξοδοι των εκπαιδευτικών διατάξεων που επιλέγονται, πρέπει να **κανονικοποιούνται** στο ίδιο διάστημα τιμών.

Ο υπολογισμός του ελαχίστου της καθολικής συνάρτησης σφάλματος \mathbf{E} γίνεται με τις παρακάτω δύο τεχνικές: Η πρώτη τεχνική ονομάζεται *on-line* κατά την οποία τα παραδείγματα εκπαίδευσης παρουσιάζονται σειριακά, συνήθως σε τυχαία σειρά διαδοχής. Για κάθε παράδειγμα (ή για μία ομάδα παραδειγμάτων) υπολογίζεται ξεχωριστά η τιμή της τοπικής συνάρτησης σφάλματος και βάσει αυτής γίνεται η τροποποίηση των τιμών των βαρών. Η δεύτερη τεχνική ονομάζεται *batch μέθοδος εκπαίδευσης*, βάσει της οποίας ελαχιστοποιείται η καθολική συνάρτηση σφάλματος \mathbf{E} συσσωρεύοντας τις αλλαγές των βαρών για όλα τα δείγματα εκπαίδευσης πριν από την τελική διόρθωση των βαρών.

Η *on-line* τεχνική εκπαίδευσης δικτύων ακολουθεί τα παρακάτω βήματα: **(i)** Επιλέγεται με τυχαίο τρόπο το επόμενο ζεύγος εκπαίδευσης από το σύνολο ζευγών εκπαίδευσης και επιδρά το διάνυσμα εισόδου στο δίκτυο. **(ii)** Υπολογίζεται η έξοδος του δικτύου. **(iii)** Υπολογίζεται το σφάλμα πρόβλεψης μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου. **(iv)** Ελαχιστοποιείται το σφάλμα και διορθώνονται τα βάρη του δικτύου. Επαναλαμβάνονται τα βήματα (i) μέχρι (iv) για κάθε διάνυσμα στο σύνολο εκπαίδευσης, μέχρι το σφάλμα σε όλο το σύνολο να είναι κάτω από κάποιο αποδεκτό όριο.

Η *batch* τεχνική εκπαίδευσης δικτύων απαιτεί τα ακόλουθα βήματα: **(i)** Όλα τα ζεύγη εκπαίδευσης επιδρούν διαδοχικά στο δίκτυο. **(ii)** Υπολογίζεται η έξοδος του δικτύου για κάθε ένα από αυτά τα πρότυπα. **(iii)** Υπολογίζεται το σφάλμα πρόβλεψης μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου. **(iv)** Υπολογίζεται το μέγεθος της διόρθωσης για το σύνολο του σφάλματος και γίνεται ενημέρωση των βαρών. Επαναλαμβάνονται τα βήματα (i) μέχρι (iv) έως ότου το σφάλμα πρόβλεψης για όλο το σύνολο εκπαίδευσης να γίνει αποδεκτά χαμηλό.

Μπορεί να παρατηρηθεί ότι τα βήματα (i) και (ii), των δύο τεχνικών διόρθωσης, αποτελούν ένα προς τα εμπρός πέρασμα (*forward pass*), έτσι ώστε να παραχθούν οι έξοδοι του δικτύου, ενώ τα βήματα (iii) και (iv) αποτελούν το πίσω πέρασμα (*reverse*

pass), όπου το υπολογιζόμενο σφάλμα πρόβλεψης επιστρέφει πίσω στο δίκτυο όπου και χρησιμοποιείται για την διόρθωση των βαρών.

2.3.3.2 Εμπρός πέρασμα

Ένα διάνυσμα εισόδου **X** επιδρά στο δίκτυο και ένα διάνυσμα εξόδου **OUT** παράγεται. Το ζεύγος των διανυσμάτων εισόδου - στόχου, **X** και **TAR** αντίστοιχα, προέρχονται από το σύνολο εκπαίδευσης. Οι υπολογισμοί σε ένα πολυεπίπεδο δίκτυο πραγματοποιούνται σειριακά επίπεδο - επίπεδο, ξεκινώντας από το κοντινότερο επίπεδο στην είσοδο του δικτύου. Οι εισοδοί **NET** επιδρούν σε κάθε νευρώνα του πρώτου κρυφού επιπέδου και υπολογίζεται το άθροισμα των γινομένων: [βάρους]×[είσοδος]. Η συνάρτηση μεταφοράς **F** συμπιέζει το άθροισμα αυτό για να παραχθεί το σήμα **OUT**, για κάθε νευρώνα του επιπέδου. Όταν υπολογιστεί το σύνολο των εξόδων για το επίπεδο, αυτό χρησιμοποιείται ως είσοδος του επομένου επιπέδου. Αυτή η διαδικασία επαναλαμβάνεται επίπεδο - επίπεδο μέχρι να υπολογιστεί το τελικό σύνολο εξόδων του δικτύου. Στο επίπεδο εξόδου οι υπολογιζόμενες έξοδοι ή διαφορετικά οι παρατηρούμενες έξοδοι, αφαιρούνται από τις επιθυμητές εξόδους προκειμένου να παραχθεί το σφάλμα πρόβλεψης:

$$\text{err}_i^{(p)} = \text{tar}_i^{(p)} - \text{out}_{k,i} \quad (2.22)$$

2.3.3.3 Πίσω πέρασμα

Τόσο για την περίπτωση της on-line όσο και της batch τεχνικής εκπαίδευσης το μέγεθος διόρθωσης είναι ανάλογο της αρνητικής κλίσης της συνάρτησης σφάλματος.

(α) Διόρθωση βαρών του επιπέδου εξόδου

Επειδή οι επιθυμητές τιμές εξόδου είναι διαθέσιμες για κάθε νευρώνα του επιπέδου εξόδου, η διόρθωση των συναπτικών βαρών του επιπέδου εξόδου πραγματοποιείται εύκολα χρησιμοποιώντας τον κανόνα διόρθωσης *Δέλτα (Delta rule)*. Για την περίπτωση των εξόδων των κρυφών επιπέδων δεν υπάρχουν κάποιες επιθυμητές τιμές για να συγκριθούν, οπότε η εκπαίδευση είναι πολυπλοκότερη.

Αρχικά θα γίνει αναφορά στην περίπτωση του on-line σχήματος εκπαίδευσης. Σύμφωνα με το σχήμα αυτό, η διόρθωση των συναπτικών βαρών είναι ανάλογη της αρνητικής κλίσης της τοπικής συνάρτησης σφάλματος **E_p**:

$$\Delta w_{k,ji}^{t+1} = -\eta \frac{\partial E_p}{\partial w_{k,ji}} = -\eta \frac{\partial E_p}{\partial \text{net}_{k,i}} \frac{\partial \text{net}_{k,i}}{\partial w_{k,ji}} \quad (2.23)$$

Λαμβάνοντας υπόψη την εξίσωση (2.18) είναι δυνατό να οριστεί η παράμετρος δέλτα για τον νευρώνα i του επιπέδου k .

$$\delta_{k,i} = -\frac{\partial E_p}{\partial \text{net}_{k,i}} = -\frac{\partial E_p}{\partial \text{err}_i^{(p)}} \frac{\partial \text{err}_i^{(p)}}{\partial \text{net}_{k,i}} = \text{err}_i^{(p)} dF(\text{net}_{k,i}) \quad (2.24)$$

Ακολούθως οι τιμές $\delta_{k,i}$ χρησιμοποιούνται στον υπολογισμό της διόρθωσης των βαρών στο επίπεδο εξόδου k σύμφωνα με την σχέση

$$\Delta w_{k,ji}^{t+1} = \eta \delta_{k,i} \text{out}_{p,j}^{(p)} \quad (2.25)$$

$$w_{k,ji}^{t+1} = w_{k,ji}^t + \Delta w_{k,ji}^{t+1} \quad (2.26)$$

όπου η είναι ένας συντελεστής που δηλώνει τον ρυθμό εκμάθησης και που συνήθως λαμβάνεται στο διάστημα $[0.1-0.9]$ και $\text{out}_{p,j}^{(p)}$ δηλώνει την έξοδο στο αμέσως προηγούμενο επίπεδο p για το p ζεύγος εκπαίδευσης. Ο συντελεστής η είναι ανάλογος του μήκους βήματος των αλγορίθμων βελτιστοποίησης με μαθηματικό προγραμματισμό, $w_{k,ji}^{t+1}$ είναι η τιμή του βάρους από τον νευρώνα j του κρυμμένου επιπέδου p στον νευρώνα i του επιπέδου εξόδου k στο βήμα διόρθωσης $t+1$.

Στην περίπτωση του batch σχήματος εκπαίδευσης, αντίστοιχα, όπου συσσωρεύονται οι διορθώσεις των συναπτικών βαρών για όλα τα πρότυπα εκπαίδευσης, η διόρθωση των συναπτικών βαρών μπορεί να γραφτεί ως εξής:

$$\Delta w_{k,ji}^{t+1} = \sum_{p=1}^{N_p} \Delta_p w_{k,ji}^{t+1} = -\eta \frac{\partial E}{\partial w_{k,ji}} = \eta \sum_{p=1}^{N_p} \delta_{k,i}^{(p)} \text{out}_{p,j}^{(p)} \quad (2.27)$$

Κάθε βήμα της διαδικασίας batch αφού έχει διατρέξει όλα τα δείγματα εκπαίδευσης ονομάζεται Epoch.

(β) Η διόρθωση των βαρών στα κρυμμένα επίπεδα

Για τα κρυμμένα επίπεδα δεν υπάρχει διάνυσμα στόχου, οπότε δεν μπορεί να χρησιμοποιηθεί η διαδικασία εκπαίδευσης που περιγράφηκε πιο πάνω. Το γεγονός αυτό εμπόδιζε την χρήση πολυεπίπεδων δικτύων μέχρι να εμφανιστεί ο αλγόριθμος πίσω διάδοσης. Ο αλγόριθμος αυτός διορθώνει τα βάρη των κρυμμένων επιπέδων διαδίδοντας το σφάλμα προς τα πίσω, μέσω του δικτύου, επίπεδο - επίπεδο, διορθώνοντας τα βάρη σε κάθε επίπεδο. Οι εξισώσεις (2.25) και (2.26)

χρησιμοποιούνται για όλα τα επίπεδα, δηλαδή τόσο για το επίπεδο εξόδου όσο και για τα κρυμμένα επίπεδα, οπότε είναι απαραίτητο να υπολογιστεί ο παράγοντας δ για τα κρυμμένα επίπεδα χωρίς την ύπαρξη κάποιου διανύσματος στόχου.

Αρχικά θα ασχοληθούμε με το *on-line* σχήμα εκπαίδευσης. Θεωρούμε ένα νευρώνα στο κρυμμένο επίπεδο p το οποίο είναι το αμέσως προηγούμενο επίπεδο της εξόδου. Στο εμπρός πέρασμα, αυτός ο νευρώνας διαδίδει την τιμή εξόδου του, στο επίπεδο εξόδου μέσω των βαρών σύνδεσης του επιπέδου εξόδου. Σε αντιστοιχία με την εξίσωση (2.24) μπορούν να διατυπωθούν οι ακόλουθες εξισώσεις διόρθωσης των βαρών για το κρυμμένο επίπεδο p .

$$\delta_{p,j} = -\frac{\partial E_p}{\partial \text{net}_{p,j}} \quad (2.28)$$

Στην περίπτωση των κρυμμένων επιπέδων ο συντελεστής αυτός δεν είναι άμεσα υπολογίσιμος και για το λόγο αυτό θα υπολογιστεί άμεσα μέσω γνωστών ποσοτήτων. Η εξίσωση (2.28) γράφεται:

$$\delta_{p,j} = -\frac{\partial E_p}{\partial \text{net}_{p,j}} = -\frac{\partial E_p}{\partial \text{out}_{p,j}^{(p)}} \frac{\partial \text{out}_{p,j}^{(p)}}{\partial \text{net}_{p,j}} \quad (2.29)$$

Λαμβάνοντας υπόψη ότι

$$\text{out}_{p,j}^{(p)} = F(\text{net}_{p,j}) \quad (2.30)$$

και

$$\begin{aligned} -\frac{\partial E_p}{\partial \text{out}_{p,j}^{(p)}} &= -\sum_{l=1}^n \frac{\partial E_p}{\partial \text{net}_{k,l}} \frac{\partial \text{net}_{k,l}}{\partial \text{out}_{p,j}^{(p)}} = \\ &= \sum_{l=1}^n \left[-\frac{\partial E_p}{\partial \text{net}_{k,l}} \right] \frac{\partial}{\partial \text{out}_{p,j}^{(p)}} \left[\sum_{m=1}^n w_{k,ml}^{t+1} \cdot \text{net}_{p,m} \right] = \\ &= \sum_{l=1}^n \delta_{k,l} \frac{\partial}{\partial \text{out}_{p,j}^{(p)}} \left[\sum_{m=1}^n w_{k,ml}^{t+1} \cdot \text{net}_{p,m} \right] = \\ &= \sum_{l=1}^n \delta_{k,l} \cdot w_{k,jl}^{t+1} \end{aligned} \quad (2.31)$$

καθώς επίσης και ότι

$$\delta_{p,j} = -\frac{\partial E_p}{\partial \text{out}_{p,j}^{(p)}} \frac{\partial F}{\partial \text{net}_{p,j}} \quad (2.32)$$

οπότε η εξίσωση (2.32) σε συνδυασμό με την εξίσωση (2.27) εκφρασμένη για το p κρυμμένο επίπεδο προσδιορίζει τις διορθώσεις των βαρών για τα κρυμμένα επίπεδα. Η εξ. (2.29) γίνεται:

$$\delta_{p,j} = dF(\text{net}_{p,j}) \left(\sum_{l=1}^n \delta_{k,l} w_{k,jl} \right) \quad (2.33)$$

Στην περίπτωση της batch μεθόδου εκπαίδευσης οι διορθώσεις των τιμών των βαρών γίνονται με βάση την εξίσωση (2.27). Μετά τον υπολογισμό των διορθώσεων των βαρών, οι διορθωμένες τιμές των βαρών δίνονται από τη σχέση $w_{p,ij}^{t+1} = w_{p,ij}^t + \Delta w_{p,ij}^{t+1}$, και χρησιμοποιούνται για τον επόμενο κύκλο εκπαίδευσης. Η διαδικασία επαναλαμβάνεται για όλα τα ζεύγη εκπαίδευσης, έως ότου επιτευχθεί το επιθυμητό επίπεδο σφάλματος. Στην παρούσα μεταπτυχιακή εργασία η τεχνική εκπαίδευσης που ακολουθήθηκε είναι η batch τεχνική.

Παρακάτω παρουσιάζονται ορισμένα πλεονεκτήματα και μειονεκτήματα των δύο τεχνικών:

- ✓ Η on-line τεχνική εκπαίδευσης χρησιμοποιείται στην περίπτωση που όλα τα ζεύγη εκπαίδευσης δεν είναι διαθέσιμα πριν την έναρξη της διαδικασίας εκπαίδευσης, ή στην περίπτωση που είναι επιθυμητό, η παραγωγή των ζευγών εκπαίδευσης να γίνεται παράλληλα με την διαδικασία εκπαίδευσης, λόγω μεγάλου υπολογιστικού κόστους.
- ✓ Η on-line τεχνική εκπαίδευσης εκλέγεται ως καταλληλότερη στην περίπτωση μεγάλου πλήθους ζευγών εκπαίδευσης, και στην περίπτωση που η απαιτούμενη μνήμη κάνει αδύνατη την χρήση της batch τεχνικής [36].
- ✓ Η on-line τεχνική συμπεριφέρεται καλύτερα στην περίπτωση τοπικών ελαχίστων, λόγω της τυχματικότητας με την οποία εκτελείται.
- ✓ Η batch τεχνική συνδυάζεται καλύτερα με πιο σύνθετες μεθόδους βελτιστοποίησης [37].

2.3.4 Έλεγχος πορείας εκπαίδευσης

Υπάρχουν δύο κριτήρια, βάση των οποίων ελέγχεται αν το δίκτυο έχει φτάσει σε ένα αποδεκτό επίπεδο πρόβλεψης ή όχι. Το πρώτο είναι το σφάλμα της τετραγωνικής ρίζας των μέσων τετραγώνων (Root Mean Square RMS) το οποίο υπολογίζεται από τη σχέση

$$e_{\text{RMS}} = \sqrt{E} = \sqrt{\frac{1}{N_p N_{\text{out}}} \sum_{N_p} \sum_{i=1}^{N_{\text{out}}} (\text{tar}_i - \text{out}_i)^2} \quad (2.34)$$

όπου N_p είναι ο συνολικός αριθμός των ζευγών στο σύνολο εκπαίδευσης και N_{out} είναι ο αριθμός μονάδων στο επίπεδο εξόδου. Το μέγεθος e_{RMS} παρέχει ένα μέτρο της διαφοράς μεταξύ προβλεφθέντων και “ακριβών” τιμών σε κάθε κύκλο εκπαίδευσης, για το σύνολο των ζευγών εκπαίδευσης. Αναφέρεται δηλαδή στο καθολικό σφάλμα του δικτύου.

Ένα δεύτερο κριτήριο ελέγχου της εκπαίδευσης αποτελεί ο *βαθμός διόρθωσης* (*correctness ratio*) του συνόλου εκπαίδευσης [38]. Για να υπολογιστεί ο βαθμός διόρθωσης του δικτύου, οι πραγματικές έξοδοι του δικτύου συγκρίνονται με τις επιθυμητές τιμές των εξόδων. Το δίκτυο θεωρείται “ακριβές” στην περίπτωση που οι προβλεπόμενες τιμές είναι κοντά στις επιθυμητές. Ο βαθμός διόρθωσης χαρακτηρίζει την τρέχουσα κατάσταση του δικτύου και είναι το ποσοστό των ζευγών του συνόλου εκπαίδευσης για τα οποία το δίκτυο εμφανίζεται να έχει “ακριβής” πρόβλεψη. Η βέλτιστη τιμή του βαθμού διόρθωσης είναι το 100%, δηλαδή η περίπτωση που για όλα τα ζεύγη εκπαίδευσης το δίκτυο εμφανίζει “σωστή” πρόβλεψη.

Ένα ερώτημα που προκύπτει είναι ποια είναι η σχέση του e_{RMS} με τον βαθμό διόρθωσης. Δηλαδή αν για κάποια χαμηλή τιμή του σφάλματος e_{RMS} παρουσιάζεται αντίστοιχα υψηλή τιμή του βαθμού διόρθωσης. Η σχέση δεν είναι γραμμική. Είναι δυνατό να επιτευχθεί κατάσταση του δικτύου με χαμηλή τιμή σφάλματος e_{RMS} , το οποίο όμως να παρουσιάζει χαμηλή τιμή του βαθμού διόρθωσης και συνεπώς ο βαθμός εκπαίδευσης του δικτύου να μην είναι αποδεκτός. Ας θεωρήσουμε ένα δίκτυο, του οποίου τα περισσότερα πρότυπα οδηγούν σε τιμή τοπικού λάθους $\text{err}_{k,i} = \text{tar}_{k,i} - \text{out}_{k,i}$ κοντά στο μηδέν και τα υπόλοιπα κοντά στην μονάδα. Όπως αναφέρθηκε προηγουμένως, οι έξοδοι του δικτύου είναι κανονικοποιημένες στο διάστημα [0,1] οπότε στην περίπτωση που η τιμή του τοπικού σφάλματος είναι κοντά στη μονάδα η πρόβλεψη είναι λανθασμένη. Σε αυτή την περίπτωση, η τιμή του βαθμού διόρθωσης δεν μπορεί να είναι υψηλή αν και το συνολικό σφάλμα e_{RMS} εμφανίζει χαμηλή τιμή. Αν

θεωρηθεί ένα δεύτερο δίκτυο, του οποίου το τοπικό λάθος είναι $\mathbf{err}_{k,i} = \mathbf{tar}_{k,i} - \mathbf{out}_{k,i} \approx 0.1$ για όλα τα πρότυπα, τότε η τιμή του βαθμού διόρθωσης θα είναι κοντά στο 100%. Το φαινόμενο αυτό είναι συχνό για την περίπτωση μεγάλου πλήθους προτύπων εκπαίδευσης [39]. Επίσης το φαινόμενο αυτό εμφανίζεται για την περίπτωση μη κατάλληλων αρχιτεκτονικών (δηλαδή πολλοί / λίγοι νευρώνες στο κρυμμένο επίπεδο) ή στην περίπτωση λανθασμένου συνδυασμού των παραμέτρων εκπαίδευσης και momentum [40].

Για να αποφευχθεί το παραπάνω φαινόμενο είναι απαραίτητο να ελέγχεται το μέγιστο τοπικό λάθος των δειγμάτων σε κάθε κύκλο διόρθωσης των βαρών παράλληλα με το καθολικό λάθος. Οι Yu και Simmons [38] πρότειναν μια νέα τεχνική αντιμετώπισης του φαινομένου αυτού, η οποία ονομάζεται η μέθοδος του *Κατερχόμενου Έψιλον* (*descending epsilon*). Σύμφωνα με την τεχνική αυτή επιλέγεται στην αρχή της διαδικασίας εκπαίδευσης μία τιμή για την παράμετρο *έψιλον*, η οποία εκφράζει τα περιθώρια σφάλματος και υπολογίζεται στη συνέχεια το τοπικό σφάλμα για κάθε ένα δείγμα του συνόλου εκπαίδευσης. Η πίσω διάδοση του σφάλματος πραγματοποιείται μόνο για εκείνα τα ζεύγη τα οποία παρουσιάζουν τιμή σφάλματος μεγαλύτερη του *έψιλον*. Όταν επιτευχθεί για όλα τα ζεύγη μέγεθος τοπικού σφάλματος μικρότερο του “έψιλον” τότε η τρέχουσα τιμή του “έψιλον” μειώνεται και η διαδικασία συνεχίζεται μέχρι να επιτευχθεί μια χαμηλή τιμή τοπικού λάθους για όλα τα ζεύγη.

2.3.5 Η παράμετρος Bias

Είναι επιθυμητό σε πολλές περιπτώσεις να υπάρχει σε κάθε νευρώνα η καθοδηγητική παράμετρος (Bias), η χρήση της οποίας επιτρέπει την πιο γρήγορη σύγκλιση της διαδικασίας εκπαίδευσης. Η παράμετρος αυτή μπορεί εύκολα να εισαχθεί στον αλγόριθμο εκπαίδευσης προσθέτοντας ένα βάρος σε κάθε νευρώνα. Αυτό το βάρος διορθώνεται με τον ίδιο ακριβώς τρόπο όπως όλα τα άλλα βάρη του δικτύου. Η μόνη διαφορά έγκειται στο ότι οι νευρώνες αυτοί έχουν σαν είσοδο πάντα την τιμή +1.

2.3.6 Ο όρος Momentum

(1) Οι Rumelhart, Hinton και Williams [41] το 1986 πρότειναν μία μέθοδο για τη μείωση του χρόνου εκπαίδευσης του αλγορίθμου πίσω διάδοσης, όμως η προτεινόμενη μεθοδολογία αύξησε την ευστάθεια του αλγορίθμου. Η μέθοδος ονομάζεται *momentum backpropagation* και βασίζεται στην προσθήκη ενός όρου στην εξίσωση διόρθωσης των βαρών, που είναι ανάλογος κάποιου ποσοστού της διόρθωσης του προηγούμενου βήματος. Η διόρθωση των βαρών με τον όρο αυτό και για την περίπτωση του επιπέδου εξόδου \mathbf{k} γράφεται:

$$\Delta w_{k,ji}^{t+1} = \eta \delta_{k,i} \text{out}_{p,j} + a \Delta w_{k,ji}^t \quad (2.35)$$

ενώ οι τρέχουσες τιμές των βαρών δίνονται από την εξίσωση:

$$w_{k,ji}^{t+1} = w_{k,ji}^t + \Delta w_{k,ji}^{t+1} \quad (2.36)$$

όπου a είναι ο συντελεστής *momentum* και συνήθως εκλέγεται κοντά στο 0.5 κατά την έναρξη της διαδικασίας ενώ στην περίπτωση που εντοπιστεί κάποιο τοπικό ελάχιστο αυξάνει μέχρι την τιμή 0.9 [14].

(2) Οι Sejnowski και Rosenberg [42] το 1987 πρότειναν μία παρόμοια μέθοδο για τη βελτίωση του κλασικού αλγορίθμου της πίσω διάδοσης, σύμφωνα με την οποία το μέγεθος της διόρθωσης των βαρών γίνεται με τον ακόλουθο τρόπο:

$$\Delta w_{k,ji}^{t+1} = a \Delta w_{k,ji}^t + (1-a) \delta_{k,i} \text{out}_{p,j} \quad (2.37)$$

ενώ οι αλλαγές των βαρών υπολογίζονται από την σχέση:

$$w_{k,ji}^{t+1} = w_{k,ji}^t + \eta \Delta w_{k,ji}^{t+1} \quad (2.38)$$

όπου a είναι ένας συντελεστής με τιμή από το διάστημα [0, 1] και η από το ίδιο διάστημα.

Στην περίπτωση που ο συντελεστής a λάβει την τιμή 0 τότε η διόρθωση του βάρους περιλαμβάνει μηδενικό ποσοστό της προηγούμενης διόρθωσης. Αντίθετα αν ο συντελεστής λάβει την τιμή 1.0 τότε η νέα αλλαγή αγνοείται και επαναλαμβάνεται η προηγούμενη διόρθωση. Ο συντελεστής η χρησιμεύει στη ρύθμιση του μεγέθους της μείωσης αλλαγής των βαρών και είναι επιλογή του χρήστη.

2.3.7 Παράλυση του δικτύου

Κάτω από ορισμένες συνθήκες η διαδικασία εκπαίδευσης μπορεί να οδηγηθεί σε μια κατάσταση κατά την οποία οι διορθώσεις των βαρών του δικτύου να οδηγούν σε μηδενική βελτίωση της εκτίμησης του. Το φαινόμενο αυτό ονομάζεται *παράλυση του δικτύου* (*network paralysis*) και είναι δυνατό να οδηγήσει σε εκπαίδευση διάρκειας απείρου χρόνου.

Το φαινόμενο αυτό εμφανίζεται όταν ένα μεγάλο ποσοστό των νευρώνων του δικτύου έχουν τιμές βαρών μεγάλου μεγέθους και συνεπώς παράγουν μεγάλες τιμές του σήματος **NET**. Τότε και οι τιμές των εξόδων **OUT** παίρνουν μεγάλες τιμές, οι οποίες

βρίσκονται στα όρια της σιγμοειδούς συνάρτησης μεταφοράς. Στις θέσεις αυτές η παράγωγος της σιγμοειδούς συνάρτησης προσεγγίζει το μηδέν και σύμφωνα με την εξίσωση (2.25), κατά την οποία ο αλγόριθμος πίσω διάδοσης υπολογίζει το μέγεθος της αλλαγής των βαρών χρησιμοποιείται η παράγωγος της σιγμοειδούς συνάρτησης. Για τους νευρώνες που εμφανίζουν το πρόβλημα της παράλυσης, οι τιμές της παραγώγου κοντά στο μηδέν οδηγούν σε μηδενικές διορθώσεις. Σε περιπτώσεις που το φαινόμενο παρατηρείται σε μεγάλο μέρος του δικτύου τότε η διαδικασία εκπαίδευσης είναι πολύ αργή, ακινητοποιημένη.

Δυστυχώς δεν μπορεί να προβλεφθεί πότε ένα δίκτυο κατά την διαδικασία εκπαίδευσης θα παρουσιάσει το πρόβλημα της παράλυσης. Εμπειρικά έχει παρατηρηθεί ότι, σχετικά μικρού μεγέθους βήματα συνήθως δεν προκαλούν το φαινόμενο αυτό, το οποίο όμως μπορεί να οδηγήσει στην υπέρμετρη αύξηση του χρόνου της εκπαίδευσης. Για τον περιορισμό του φαινομένου αυτού, εκτός από το διάνυσμα επιθυμητών εξόδων συστήνεται να κανονικοποιούνται στο διάστημα $[0, 1]$ και τα διανύσματα εισόδων.

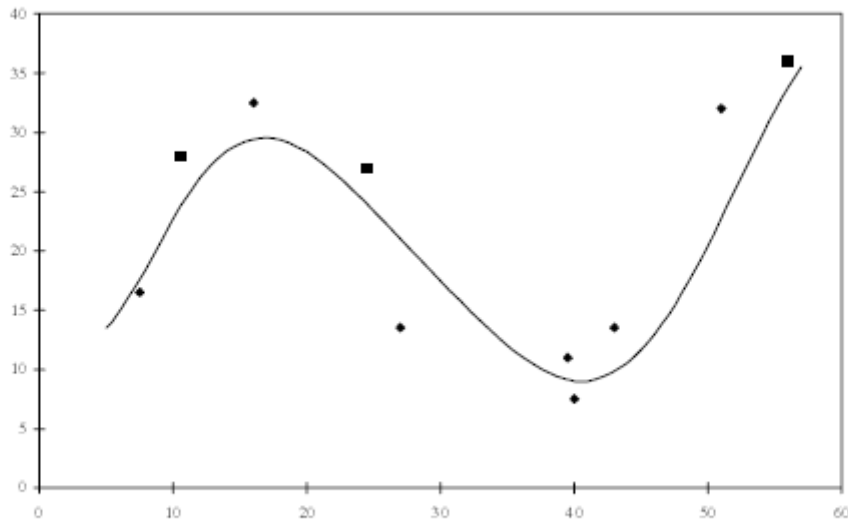
Μια τεχνική για την αντιμετώπιση του φαινομένου της παράλυσης είναι η ανίχνευση των νευρώνων, των οποίων το σήμα εξόδου **OUT** προσεγγίζει τα όρια της σιγμοειδούς. Οι τιμές των βαρών που τροφοδοτούν έναν τέτοιο νευρώνα συμπιέζονται μέσω μιας συνάρτησης ανάλογης εκείνης που παράγει το σήμα εξόδου ενός νευρώνα. Τα όρια αυτής της συνάρτησης είναι τα $(+5, -5)$ ή κάποιες άλλες κατάλληλες τιμές. Οι νέες τιμές των βαρών υπολογίζονται από την ακόλουθη σχέση:

$$w_{k,ji} = -5 + \frac{10}{1 + \exp(-w_{k,ji}/5)} \quad (2.39)$$

Η συνάρτηση αυτή, μειώνει σημαντικά το μέγεθος των βαρών μεγάλου μεγέθους, ενώ αντίθετα τα βάρη μικρού μεγέθους δεν επηρεάζονται. Έχει αποδειχθεί πειραματικά, ότι η τεχνική αυτή αντιμετωπίζει το πρόβλημα της παράλυσης χωρίς σημαντική ενόχληση της διαδικασίας εκπαίδευσης [43].

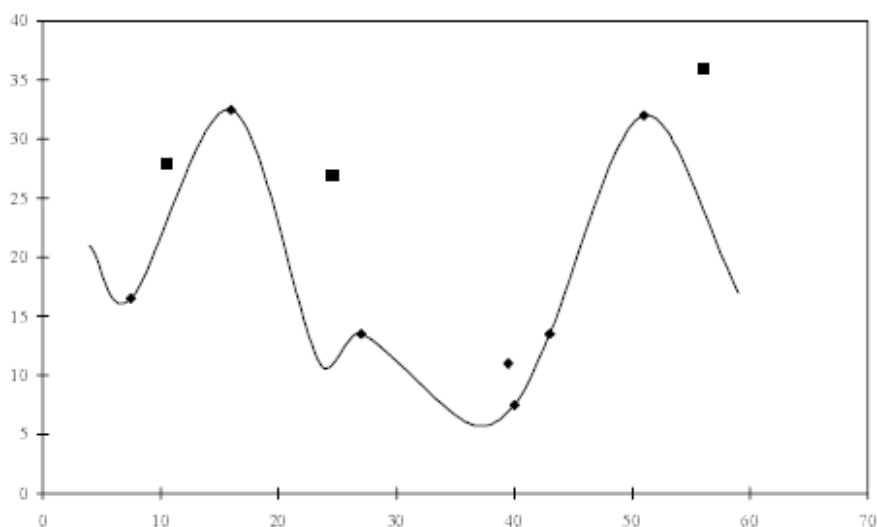
2.3.8 Υπερβολική εκπαίδευση δικτύου

Έστω ότι με τη βοήθεια ενός νευρωνικού δικτύου επιχειρείται η προσομοίωση μιας συνάρτησης. Στο σχήμα (2.9) εμφανίζεται η προσομοίωση που έχει επιτευχθεί καθώς επίσης και κάποιες κουκίδες, από τις οποίες οι μαύρες αντιπροσωπεύουν τα δείγματα που χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου, ενώ οι λευκές τα δείγματα που χρησιμοποιήθηκαν για τον έλεγχο της πρόβλεψης του δικτύου.



Σχήμα 2.9: Προσομοίωση συνάρτησης

Αν το γράφημα της προσομοιούμενης συνάρτησης διερχόταν από τα δείγματα που χρησιμοποιήθηκαν για την εκπαίδευση τότε το λάθος θα ήταν ίσο με μηδέν. Αν χρησιμοποιηθούν περισσότεροι κόμβοι στο κρυμμένο επίπεδο είναι δυνατό να επιτευχθεί μεγαλύτερη ακρίβεια, ως προς τα δείγματα εκπαίδευσης και συνεπώς προσομοίωση της εξίσωσης όπως φαίνεται στο σχήμα (2.10).



Σχήμα 2.10: Προσομοίωση συνάρτησης

Όπως παρατηρείται, παρόλο που το γράφημα ταιριάζει περισσότερο με την πρώτη προσομοίωση στα ζεύγη εκπαίδευσης, η συμπεριφορά του σε σχέση με τα ζεύγη ελέγχου είναι χειρότερη. Το φαινόμενο αυτό ονομάζεται υπερβολική εκπαίδευση. Μια τεχνική για να αποφευχθεί τέτοιου είδους συμπεριφορά της εκπαίδευσης είναι οι ακόλουθη: Γίνεται εκκίνηση της διαδικασίας εκπαίδευσης με τυχαίες τιμές των βαρών. Μετά από κάποιο συγκεκριμένο αριθμό κύκλων εκπαίδευσης γίνεται έλεγχος της συμπεριφοράς του δικτύου μέσω των ζευγών ελέγχου, ενώ παράλληλα αποθηκεύονται

οι τιμές των βαρών οι οποίες αντιπροσωπεύουν την τρέχουσα κατάσταση του δικτύου. Η κατάσταση αυτή ονομάζεται προσωρινά βέλτιστη. Η διαδικασία εκπαίδευσης συνεχίζεται και μετά από κάποιους κύκλους, ο έλεγχος πρόβλεψης επαναλαμβάνεται. Αν η πρόβλεψη του δικτύου είναι καλύτερη της προσωρινά βέλτιστης κατάστασης, τότε η τρέχουσα κατάσταση του δικτύου αντικαθιστά την προσωρινά βέλτιστη και η διαδικασία εκπαίδευσης συνεχίζεται. Η διαδικασία τερματίζεται όταν για κάποιο αριθμό ελέγχων δεν υπάρχει βελτίωση της προσωρινά βέλτιστης κατάστασης.

2.3.9 Αρχιτεκτονική δικτύου

Το πλήθος των κόμβων στο(-α) κρυμμένο(-α) επίπεδο(-α) δεν είναι γνωστό εκ των προτέρων και συνήθως υπολογίζεται με δοκιμές. Ως βέλτιστη τοπολογία ενός δικτύου θεωρείται η μικρότερη δυνατή, όσον αφορά το πλήθος των κόμβων στα κρυμμένα επίπεδα, η οποία επιτρέπει την σωστή εκπαίδευση του δικτύου για την επίτευξη επιτρεπτών περιθωρίων σφάλματος στις προβλέψεις. Μεταξύ των διαφόρων τεχνικών που έχουν προταθεί για την εύρεση της βέλτιστης τοπολογίας, είναι η χρήση γενετικών αλγορίθμων, οι οποίοι όμως αποδεικνύονται ιδιαίτερα χρονοβόροι και αμφιβόλου αξιοπιστίας. Στα πλαίσια των εφαρμογών της παρούσης μεταπτυχιακής εργασίας έγινε επιλογή δικτύων με δύο κρυμμένα επίπεδα. Όσον αφορά το πλήθος των κόμβων στα κρυμμένα επίπεδα διαφέρει ανάλογα με το είδος της φόρτισης (μονοτονική ή ανακυκλιζόμενη).

Στην βιβλιογραφία έχουν προταθεί τεχνικές δυναμικής προσθήκης ή αφαίρεσης κόμβων (dynamic node pruning or node growing), ώστε να επιτυγχάνεται η βέλτιστη τοπολογία. Μία τέτοια τεχνική εντοπισμού της βέλτιστης τοπολογίας ξεκινάει με την επιλογή ενός δικτύου με υπερβολικό πλήθος κόμβων στα κρυμμένα επίπεδα. Κατά τη διαδικασία εκπαίδευσης αφαιρούνται οι πλεονάζοντες κόμβοι [44,45]. Για να εντοπιστούν οι κόμβοι που δεν συνεισφέρουν στην εκπαίδευση και συνεπώς είναι δυνατόν να διαγραφούν, ελέγχεται η έξοδος όλων των κόμβων για όλα τα ζεύγη εκπαίδευσης. Αν οι έξοδοι κάποιων κόμβων είναι σχεδόν σταθερού μεγέθους, ανεξαρτήτου του ζεύγους εκπαίδευσης, τότε οι κόμβοι αυτοί πρέπει να αφαιρεθούν. Αντίστοιχα αν δύο κρυμμένοι κόμβοι δίνουν ακριβώς την ίδια ή αντίθετη τιμή εξόδου τότε ένας από τους δύο κόμβους είναι δυνατό να αφαιρεθεί. Με την αφαίρεση των μη χρήσιμων κόμβων η τοπολογία του δικτύου έχει τροποποιηθεί και συνεπώς πρέπει να επαναληφθεί η διαδικασία εκπαίδευσης και να ελεγχθεί η νέα πρόβλεψη του δικτύου.

Μια δεύτερη τεχνική είναι εκείνη, η οποία βασίζεται στην προσθήκη κόμβων, ώστε να επιτευχθεί η βέλτιστη τοπολογία. Η αύξηση του πλήθους των κόμβων πραγματοποιείται μέχρι να επιτευχθεί το επιθυμητό επίπεδο σφάλματος. Στην συνέχεια

γίνεται προσπάθεια μείωσης του μεγέθους του δικτύου με μείωση των κόμβων που προστέθηκαν. Υπάρχουν δύο βασικοί λόγοι οι οποίοι καθιστούν την τεχνική αυτή περισσότερο αξιόπιστη: **(i)** Το πλήθος των κόμβων στο κρυμμένο επίπεδο είναι άγνωστο. **(ii)** Η διαδικασία εκπαίδευσης είναι πιθανό να παγιδευτεί σε ένα τοπικό ελάχιστο και έτσι με την προσθήκη κόμβων δίνεται η δυνατότητα στο δίκτυο να ξεφύγει από το τοπικό ελάχιστο, δεδομένου ότι με αυτό τον τρόπο αλλάζει μορφή η συνάρτηση σφάλματος στον χώρο των βαρών.

2.3.10 Επιλογή του συνόλου εκπαίδευσης

Η επιλογή του συνόλου εκπαίδευσης αποτελεί ένα από τα πιο κρίσιμα βήματα στη διαδικασία εκπαίδευσης ενός δικτύου. Η λάθος επιλογή του συνόλου εκπαίδευσης είναι δυνατό να οδηγήσει σε λανθασμένα εκπαιδευμένο δίκτυο, το οποίο να προβλέπει σωστά μόνο ένα περιορισμένο αριθμό προτύπων. Αυτό οφείλεται στη μη σωστή κάλυψη του πεδίου ενδιαφέροντος, με αποτέλεσμα το δίκτυο να μην μπορεί να προβλέψει τιμές που ανήκουν στο πεδίο που δεν έχει εκπαιδευτεί. Πολλοί ερευνητές, στην προσπάθειά τους να καλύψουν το σύνολο του χώρου αναζήτησης, διαμέρισαν τον χώρο αναζήτησης σε περιοχές προσπαθώντας να συνδυάσουν ζεύγη εκπαίδευσης από τις περιοχές αυτές. Η μέθοδος αυτή οδηγεί σε μεγάλο πλήθος ζευγών εκπαίδευσης. Στην παρούσα μεταπτυχιακή εργασία έγινε τυχαία επιλογή των ζευγών εισόδου – εξόδου (ζεύγη τάσεων – παραμορφώσεων) από όλο το εύρος αναζήτησης.

2.4 ΑΛΛΕΣ ΤΕΧΝΙΚΕΣ ΕΚΠΑΙΔΕΥΣΗΣ

Το πρόβλημα εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου με πολλά επίπεδα αποτελεί ένα πρόβλημα βελτιστοποίησης το οποίο μπορεί να αντιμετωπιστεί με διάφορες μεθόδους βελτιστοποίησης. Στη μέχρι τώρα περιγραφή έγινε αναφορά σε μαθηματικές μεθόδους βελτιστοποίησης, ενώ σε αυτό το εδάφιο θα γίνει περιγραφή ορισμένων στοχαστικών τεχνικών οι οποίες εμφανίζονται να είναι ιδιαίτερα αποτελεσματικές σε μια πληθώρα προβλημάτων βέλτιστου σχεδιασμού.

2.4.1 Στοχαστικές μέθοδοι

Οι στοχαστικές μέθοδοι παρουσιάζουν σημαντικά πλεονεκτήματα σε σχέση με τις ντετερμινιστικές μεθόδους, εκ των οποίων το σημαντικότερο είναι η δυνατότητα διαφυγής από περιοχές τοπικών ελαχίστων (hill climbing) κατά τη διαδικασία της εκπαίδευσης. Σε αυτό το εδάφιο θα γίνει περιγραφή του τρόπου εφαρμογής στοχαστικών μεθόδων στην εκπαίδευση πολυεπίπεδων τεχνητών νευρωνικών δικτύων.

Η βασική ιδέα της εκπαίδευσης, όπως παρουσιάστηκε σε προηγούμενες παραγράφους του κεφαλαίου, είναι η διόρθωση των βαρών του δικτύου ώστε η εφαρμογή ενός συνόλου εισόδων στο δίκτυο να παράγει το επιθυμητό σύνολο εξόδων. Οι ντετερμινιστικές μέθοδοι εκπαίδευσης ακολουθούν μία, βήμα προς βήμα, διαδικασία διόρθωσης των βαρών του δικτύου, η οποία βασίζεται στις τρέχουσες τιμές των βαρών, τις τιμές των εισόδων, καθώς και τις τιμές των πραγματικών και επιθυμητών εξόδων.

Οι στοχαστικές μέθοδοι εκπαίδευσης παράγουν ψευδοτυχαίες αλλαγές στις τιμές των βαρών, και διατηρούν εκείνες τις αλλαγές που έχουν ως αποτέλεσμα τη μείωση του σφάλματος. Η αρχιτεκτονική των δικτύων που χρησιμοποιείται παραμένει η ίδια όπως περιγράφηκε σε προηγούμενες παραγράφους. Η διαδικασία εκπαίδευσης με βάση μία στοχαστική μέθοδο μπορεί να περιγραφεί με τα παρακάτω βήματα:

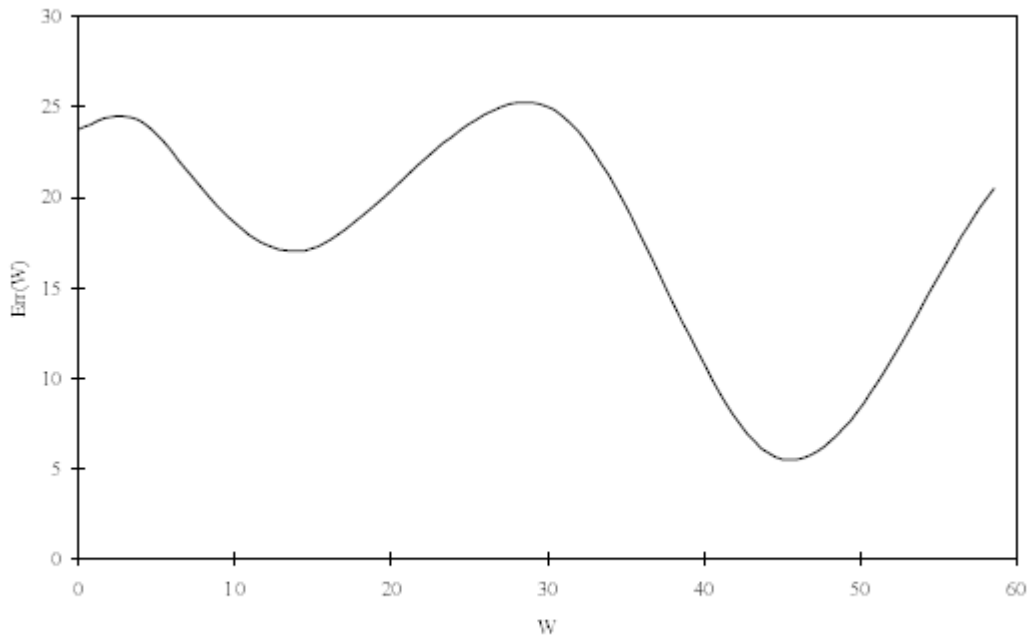
- ✓ Ένα σύνολο εισόδων επιδρά στο δίκτυο και υπολογίζονται οι έξοδοι του δικτύου.
- ✓ Συγκρίνονται οι έξοδοι του δικτύου out_i με τις επιθυμητές εξόδους tar_i και υπολογίζεται το σφάλμα πρόβλεψης. Αντικείμενο της εκπαίδευσης είναι η ελαχιστοποίηση της συνάρτησης σφάλματος:

$$\min \sum_{i=1}^p (tar_i - out_i)^2 \quad (2.40)$$

όπου p είναι το πλήθος των νευρώνων στο επίπεδο εξόδου.

- ✓ Επιλέγεται με τυχαίο τρόπο κάποιο συναπτικό βάρος του δικτύου και διορθώνεται κατά ένα μικρό τυχαίο ποσό. Αν η διόρθωση οδηγήσει στην ελάττωση της τιμής της συνάρτησης του σφάλματος, η διόρθωση διατηρείται, διαφορετικά η τιμή του βάρους παίρνει πάλι την προηγούμενη τιμή της.
- ✓ Επαναλαμβάνονται τα βήματα (i) έως (iii) μέχρι να εκπαιδευτεί το δίκτυο σε κάποιο επιθυμητό βαθμό.

Σκοπός της διαδικασίας που περιγράφηκε είναι να ελαχιστοποιηθεί η τιμή του σφάλματος της εξίσωσης (2.41), υπάρχει όμως περίπτωση η διαδικασία να οδηγηθεί σε περιοχή τοπικού ελαχίστου. Στο σχήμα (2.11) παρουσιάζεται η περίπτωση του τοπικού ελαχίστου.



Σχήμα 2.11: Το πρόβλημα του τοπικού ελαχίστου

Θα εξεταστεί η περίπτωση εγκλωβισμού σε περιοχή τοπικού ελαχίστου και η δυνατότητα διαφυγής από αυτή. Έστω ότι οι τιμές των βαρών του δικτύου ξεκινούν από κάποια τυχαία τιμή. Θεωρούμε το μοντέλο που περιγράψαμε προηγουμένως στο οποίο επιλέγεται τυχαία ένα συναπτικό βάρος το οποίο αρχικά έχει την τιμή w_A . Αν το τυχαίο βήμα διόρθωσης του βάρους είναι πολύ μικρό, η κίνηση από το σημείο w_A προς όλες τις διευθύνσεις οδηγεί σε αύξηση της τιμής της αντικειμενικής συνάρτησης και συνεπώς η διόρθωση θα απορριφθεί. Το καθολικό ελάχιστο, που βρίσκεται στην θέση w_B δεν θα εντοπιστεί ποτέ και το σύστημα θα παραμείνει παγιδευμένο στο *τοπικό ελάχιστο* (*local minimum*) της θέσης w_A . Αν υποθεθεί ότι, το τυχαίο βήμα διόρθωσης του βάρους είναι πολύ μεγάλο, τότε τα σημεία w_A και w_B όπως και όλα τα άλλα σημεία της επιφάνειας σφάλματος θα εντοπίζονται κατά τη διαδικασία υπολογισμού βέλτιστης τιμής του βάρους αλλά ο εντοπισμός δε θα είναι μόνιμος. Η τυχαία αλλαγή των βαρών θα είναι τέτοιου μεγέθους, ώστε ποτέ δεν θα εντοπίσει την επιθυμητή θέση w_B .

Μια στρατηγική για την αποφυγή τέτοιου είδους προβλημάτων είναι να γίνεται η εκκίνηση με μεγάλες τιμές βημάτων και βαθμιαίως να μειώνεται το μέγεθος της μέσης τιμής των τυχαίων αυτών βημάτων. Αυτό επιτρέπει στο δίκτυο να δραπέτεύει από τα τοπικά ελάχιστα, ενώ θα κάνει πιο βέβαιη την πορεία του δικτύου προς την θέση w_B .

Η παγίδευση του δικτύου σε τοπικά ελάχιστα αποτελεί κυρίαρχο πρόβλημα όλων των αλγορίθμων εκπαίδευσης, συμπεριλαμβανομένου και του κλασικού αλγορίθμου της πίσω διάδοσης. Οι στοχαστικές μέθοδοι δύνανται να ξεπεράσουν αυτό το πρόβλημα, με μια στρατηγική διόρθωσης των βαρών που μπορεί να οδηγήσει στον υπολογισμό

της τιμής του βάρους, ώστε η αντικειμενική συνάρτηση να φτάσει στην θέση w_B . Τα τεχνητά νευρωνικά δίκτυα μπορούν να εκπαιδευτούν μέσω μιας αντίστοιχης διαδικασίας τυχαίας διόρθωσης των βαρών. Στην αρχή λαμβάνουν χώρα διορθώσεις μεγάλου εύρους και διατηρούνται οι αλλαγές εκείνες που μειώνουν την τιμή της αντικειμενικής συνάρτησης. Ο μέσος όρος βήματος διόρθωσης των βαρών στη συνέχεια μειώνεται βαθμιαία και τελικά εντοπίζεται το καθολικό ελάχιστο.

Αυτή η διαδικασία έχει μεγάλη ομοιότητα με τη διαδικασία *ανόπτησης* των μετάλλων, και ονομάζεται “*προσομοίωση ανόπτησης*” (Simulated annealing). Όταν η θερμοκρασία ενός μετάλλου έχει υπερβεί το σημείο τήξης, τα άτομα του μετάλλου βρίσκονται σε τυχαία κίνηση. Όπως σε όλα τα φυσικά συστήματα τα άτομα τείνουν να κινηθούν προς την θέση με την ελάχιστη ενέργεια (την απλή κρυσταλλική δομή στην περίπτωση των μετάλλων). Καθώς το μέταλλο κρυώνει, βαθμιαία όλο και χαμηλότερες ενεργειακά καταστάσεις εντοπίζονται, μέχρι τελικά η χαμηλότερη δυνατή όλων των καταστάσεων, δηλαδή το καθολικό ελάχιστο, να εντοπιστεί.

2.4.1.1 Εκπαίδευση τύπου Boltzmann

Η εφαρμογή των στατιστικών μεθόδων στην εκπαίδευση των τεχνητών νευρωνικών δικτύων μπορεί να περιγραφεί με τα ακόλουθα βήματα:

- ✓ Ορίζεται μία μεταβλητή T_n , η οποία καλείται τεχνητή θερμοκρασία και αρχικά λαμβάνει κάποια υψηλή τιμή.
- ✓ Επιδρά ένα σύνολο εισόδων στο δίκτυο και υπολογίζεται η έξοδος του δικτύου και ακολούθως η τιμή του σφάλματος.
- ✓ Γίνεται τυχαία διόρθωση της τιμής ενός τυχαία επιλεγέντος βάρους, και υπολογίζεται, με βάση τη νέα τιμή του βάρους η έξοδος του δικτύου καθώς και η αλλαγή της τιμής της αντικειμενικής συνάρτησης, που οφείλεται στην τυχαία αυτή αλλαγή.
- ✓ Αν η τιμή της αντικειμενικής συνάρτησης δηλαδή του σφάλματος, έχει μειωθεί τότε διατηρείται η αλλαγή του βάρους διαφορετικά υπολογίζεται η πιθανότητα να γίνει αποδεκτή η διόρθωση.
- ✓ Η πιθανότητα να γίνει αποδεκτή η συγκεκριμένη διόρθωση ορίζεται με την κατανομή *Boltzmann* ως ακολούθως

$$P(\Delta w) = \exp(-\Delta w / kT) \quad (2.41)$$

όπου $\mathbf{P}(\Delta\mathbf{w})$ είναι η πιθανότητα του βάρους αλλαγής μεγέθους $\Delta\mathbf{w}$, \mathbf{k} είναι μια σταθερά, ενώ η παράμετρος \mathbf{T} αποτελεί την τεχνητή θερμοκρασία. Στην συνέχεια επιλέγεται ένας τυχαίος αριθμός \mathbf{r} ο οποίος ακολουθεί την ομοιόμορφη κατανομή με πεδίο τιμών το διάστημα $[0,1]$. Αν το $\mathbf{P}(\Delta\mathbf{w})$ είναι μεγαλύτερο του \mathbf{r} , τότε η διόρθωση διατηρείται, διαφορετικά το βάρος λαμβάνει την προηγούμενη τιμή του. Η διαδικασία αυτή επιτρέπει στο σύστημα να κάνει περιστασιακά βήματα σε διευθύνσεις που “χειροτερεύουν” την αντικειμενική συνάρτηση, και αυτό δίνει τη δυνατότητα στο δίκτυο να δραπτετεύει από τα τοπικά ελάχιστα.

Για να ολοκληρωθεί η στρατηγική εκπαίδευσης Boltzmann επαναλαμβάνονται τα βήματα (3) και (4) για όλα τα βάρη του δικτύου μειώνοντας βαθμιαίως την θερμοκρασία \mathbf{T} μέχρι να επιτευχθεί μια χαμηλή τιμή της αντικειμενικής συνάρτησης. Σε αυτό το σημείο ένα διαφορετικό διάνυσμα εισόδου επιδρά και η διαδικασία εκπαίδευσης επαναλαμβάνεται. Το δίκτυο εκπαιδεύεται με όλα τα διανύσματα του συνόλου εκπαίδευσης, μέχρι η αντικειμενική συνάρτηση να βρεθεί σε αποδεκτά χαμηλή τιμή για το σύνολο των ζευγών εκπαίδευσης.

Το μέγεθος της τυχαίας αλλαγής βάρους στο βήμα (3) μπορεί να ορισθεί με πολλούς τρόπους. Για παράδειγμα η αλλαγή βάρους μεγέθους $\Delta\mathbf{w}$ μπορεί να επιλεγεί με βάση την κατανομή *Gauss*

$$P(\Delta\mathbf{w}) = \exp(-\Delta\mathbf{w}^2 / \mathbf{T}^2) \quad (2.42)$$

όπου $\mathbf{P}(\Delta\mathbf{w})$ είναι η πιθανότητα της αλλαγής του βάρους μεγέθους $\Delta\mathbf{w}$, και \mathbf{T} είναι η τεχνητή θερμοκρασία. Αυτή η μέθοδος διόρθωσης των βαρών παράγει ένα σύστημα ανάλογο με τη μέθοδο που περιγράφεται από το Metropolis (1953) [46].

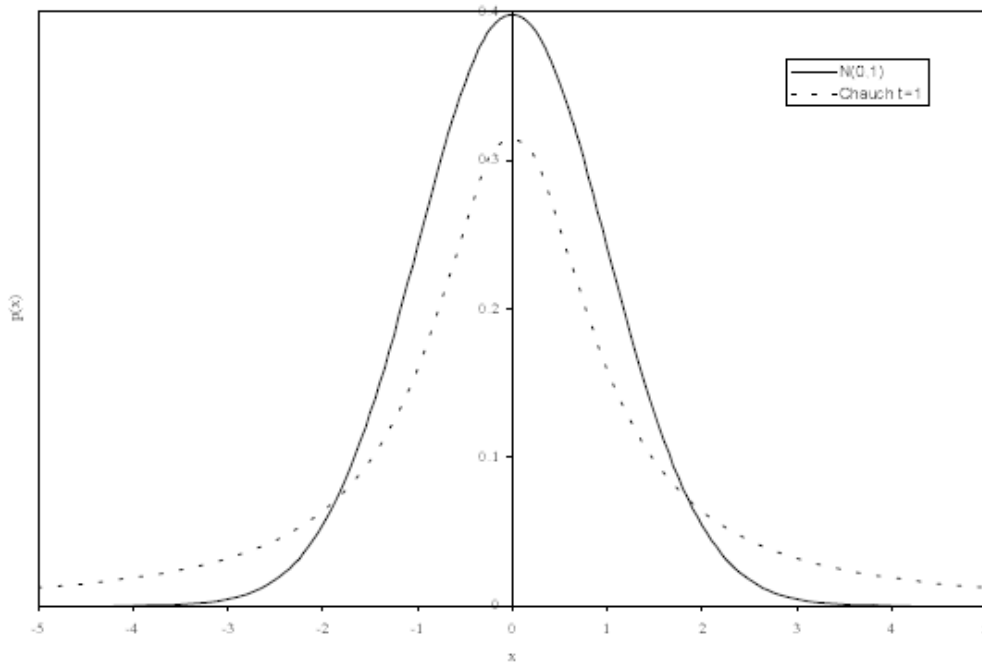
Τα χαρακτηριστικά της εκπαίδευσης κατά *Boltzmann* μελετήθηκαν εκτεταμένα από τους Geman και Geman [47] το 1984, οι οποίοι απέδειξαν ότι η μείωση της θερμοκρασίας πρέπει να είναι ανάλογη του αντιστρόφου του χρόνου. Ο βαθμός ψύξης του συστήματος εκφράζεται ως ακολούθως:

$$T(t) = T_0 / \log(1 + t) \quad (2.43)$$

όπου $\mathbf{T}(t)$ είναι η τεχνητή θερμοκρασία ως συνάρτηση του χρόνου, \mathbf{T}_0 είναι αρχική τεχνητή θερμοκρασία και t είναι ο τεχνητός χρόνος.

2.4.1.2 Εκπαίδευση Cauchy

Οι Szu και Hartley [48] το 1987 ανέπτυξαν μια μέθοδο για την πιο γρήγορη εκπαίδευση δικτύων με στατιστικές μεθόδους. Η βασική ιδέα της μεθόδου είναι η αντικατάσταση της κατανομής *Boltzmann* με την κατανομή *Cauchy* στον υπολογισμό της πιθανότητας αλλαγής κάποιου μεγέθους. Όπως φαίνεται στο σχήμα (2.12) η κατανομή *Cauchy* έχει μεγαλύτερη διασπορά από την κατανομή *Boltzmann*, και αυτό αυξάνει την πιθανότητα βημάτων μεγάλου μεγέθους. Η κατανομή *Cauchy* έχει απείρου μεγέθους διασπορά.



Σχήμα 2.12: Οι κατανομές Cauchy και Boltzmann

Μαζί με την αλλαγή κατανομής πιθανότητας, γίνεται αλλαγή και στον τρόπο μείωσης της θερμοκρασίας όπου χρησιμοποιείται μία αντίστροφη γραμμική συνάρτηση του χρόνου και όχι μία αντίστροφη λογαριθμική που ίσχυε στον αλγόριθμο εκπαίδευσης *Boltzmann*:

$$T(t) = T_0 / (1 + t) \quad (2.44)$$

Ενώ η κατανομή *Cauchy* δίνεται από την εξίσωση:

$$P(\Delta w) = T(t) / (T(t)^2 + \Delta w^2) \quad (2.45)$$

όπου $P(\Delta w)$ είναι η πιθανότητα διόρθωσης βάρους του μεγέθους Δw . Η πιθανότητα $P(\Delta w)$ μπορεί εύκολα να υπολογιστεί με απλές μεθόδους ολοκλήρωσης.

Παρά το γεγονός ότι η μέθοδος εκπαίδευσης *Cauchy* βελτίωσε την αξιοπιστία των μεθόδων εκπαίδευσης, ο χρόνος σύγκλισης παραμένει πολύ μεγάλος σε σχέση με τον κλασικό αλγόριθμο της πίσω διάδοσης. Το φαινόμενο της παράλυσης που είδαμε στο εδάφιο 2.3.7 εμφανίζεται και κατά την εκπαίδευση με τη μέθοδο *Cauchy* ιδιαίτερα σε δίκτυα με μη γραμμική συνάρτηση μεταφοράς. Η απείρου μήκους διασπορά της κατανομής *Cauchy* επιτρέπει στις αλλαγές των βαρών να παίρνουν απεριόριστου μεγέθους τιμές.

2.4.1.3 Ο αλγόριθμος της πίσω διάδοσης και εκπαίδευση *Cauchy*

Ο αλγόριθμος της πίσω διάδοσης έχει το πλεονέκτημα της “απευθείας έρευνας”, δηλαδή τα βάρη συνεχώς διορθώνονται κατά τη διεύθυνση, κατά την οποία η συνάρτηση σφάλματος ελαχιστοποιείται. Παρά το γεγονός, ότι ο χρόνος εκμάθησης είναι συχνά μεγάλος, είναι συγκριτικά πιο μικρός σε σχέση με την στατιστική μέθοδο εκπαίδευσης κατά *Cauchy*. Οι στοχαστικές μέθοδοι εκ της φύσεως τους πραγματοποιούν αρκετά βήματα προς τη λάθος κατεύθυνση και συνεπώς παρουσιάζουν αυξημένο υπολογιστικό κόστος. Η αξιοπιστία τους όμως να εντοπίζουν το καθολικό ελάχιστο, καθιστά τέτοιου είδους μεθόδους ιδιαίτερα ελκυστικές σε σχέση με την κλασική μέθοδο της πίσω διάδοσης.

Ο συνδυασμός των δύο μεθόδων έχει δοκιμαστεί από τον Wasserman [33] το 1988 με ικανοποιητικά αποτελέσματα. Ο συνδυασμός των δύο μεθόδων παράγει διόρθωση των βαρών ίση με το άθροισμα της διόρθωσης, που υπολογίζεται από τον αλγόριθμο *Backpropagation*, και του τυχαίου βήματος του αλγορίθμου *Cauchy*. Με αυτό τον τρόπο το σύστημα εκπαιδεύεται και εντοπίζει το ολικό ελάχιστο πολύ πιο γρήγορα από ένα σύστημα εκπαιδευμένο με τη μία ή την άλλη μέθοδο ξεχωριστά. Η διόρθωση των βαρών γίνεται με τις εξισώσεις:

$$\Delta w_{k,ji}^{t+1} = \eta \cdot [a\Delta w_{k,ji}^t + (1-a)\delta_{k,i} \text{out}_{p,j}] + (1-\eta)x_c \quad (2.46)$$

$$w_{k,ji}^{t+1} = w_{k,ji}^t + \Delta w_{k,ji}^{t+1} \quad (2.47)$$

όπου η είναι ο συντελεστής, ο οποίος ελέγχει το μέγεθος των όρων διόρθωσης από τον αλγόριθμο *Backpropagation* και από τον αλγόριθμο *Cauchy*. Στην περίπτωση που $\eta = 0$ τότε ο αλγόριθμος εκφυλλίζεται στον αλγόριθμο *Cauchy*, ενώ αν $\eta = 1$ τότε ταυτίζεται με τον αλγόριθμο *Backpropagation*.

2.4.2 Γενετικοί αλγόριθμοι και εκπαίδευση

Οι Γενετικοί Αλγόριθμοι (Genetic Algorithms GAs) είναι αποδεδειγμένα ένα πολύ ισχυρό εργαλείο βελτιστοποίησης. Η επιτυχής εφαρμογή τους, σε πολλά προβλήματα βελτιστοποίησης, οδήγησε πολλούς ερευνητές στη συνδυαστική χρήση τους με τα Τεχνητά Νευρωνικά Δίκτυα [49-51]. Ο συνδυασμός των δύο τεχνικών εμφανίζεται κυρίως σε δύο τομείς: **(i)** στην εύρεση της κατάλληλης τοπολογίας του νευρωνικού δικτύου και **(ii)** στην εκπαίδευση ενός νευρωνικού δικτύου συγκεκριμένης τοπολογίας. Στο πρόβλημα καθορισμού της βέλτιστης τοπολογίας οι GAs επιλέγουν τον αριθμό των κρυμμένων επιπέδων, τον αριθμό των κρυμμένων κόμβων αλλά και τον τρόπο σύνδεσης των επιπέδων μεταξύ τους. Αντίστοιχα κατά την εκπαίδευση ενός δικτύου με τους GAs το πρόβλημα της εκπαίδευσης καταλήγει σε ένα πρόβλημα βελτιστοποίησης, στο οποίο η αντικειμενική συνάρτηση είναι η συνάρτηση σφάλματος, ενώ οι μεταβλητές σχεδιασμού είναι τα συναπτικά βάρη του δικτύου.

Ένας τρόπος αυτόματης επιλογής της κατάλληλης αρχιτεκτονικής του δικτύου περιέχει δύο βήματα: γενετική αναζήτηση της κατάλληλης αρχιτεκτονικής και εκπαίδευση με τον αλγόριθμο της πίσω διάδοσης, ώστε να αξιολογηθεί η ποιότητα της κάθε αρχιτεκτονικής. Σύμφωνα με τα όσα αναφέρθηκαν για τους GAs η αναζήτηση θα ξεκινήσει με την παραγωγή ενός τυχαίου πληθυσμού νευρωνικών αρχιτεκτονικών η κάθε μία από τις οποίες αξιολογείται με βάση την πρόβλεψη που επιτυγχάνει. Η διαδικασία συνεχίζεται με τη δημιουργία νέων γενεών, πιθανών νευρωνικών αρχιτεκτονικών και τελικά επιλέγεται η κατάλληλη αρχιτεκτονική. Είναι φανερό ότι το υπολογιστικό κόστος που απαιτείται, κάνει την τεχνική μη εφαρμόσιμη.

Εκεί όμως που ο συνδυασμός των GAs με τα ΤΝΔ αποδεικνύεται πολύ αποτελεσματικός είναι στην εκπαίδευση ενός δικτύου, του οποίου η τοπολογία παραμένει σταθερή κατά την διαδικασία εκπαίδευσης. Η αναζήτηση ξεκινάει με την τυχαία δημιουργία ενός πληθυσμού από ομάδες τιμών των βαρών. Η κάθε μία από τις οποίες αξιολογείται με βάση την πρόβλεψη που επιτυγχάνει για τα πρότυπα εκπαίδευσης σε σχέση με τις επιθυμητές προβλέψεις. Κάθε κύκλος εκπαίδευσης ισοδυναμεί με μία γενιά, η επόμενη γενιά ομάδων βαρών δημιουργείται μέσω γενετικών τελεστών και είναι μερικώς εξαρτημένη από την προηγούμενη γενιά.

ΚΕΦΑΛΑΙΟ 3^ο

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Όπως αναφέρθηκε και στην εισαγωγή σκοπός της παρούσας εργασίας είναι η προσομοίωση καταστατικών νόμων των υλικών μέσω νευρωνικών δικτύων. Στο παρών κεφάλαιο παρουσιάζεται μια βιβλιογραφική ανασκόπηση σε άρθρα και εργασίες που αφορούν στην εφαρμογή των νευρωνικών δικτύων στην μοντελοποίηση καταστατικών νόμων και γενικά υλικών. Τα άρθρα παρουσιάζονται με χρονολογική σειρά από την παλαιότερη μέχρι και την πιο πρόσφατη ημερομηνία.

3.1 Material model for composites using neural networks (1993)

Συγγραφείς: Pidaparti R. M., Palakal M. J.

Στο συγκεκριμένο άρθρο αναπτύσσονται δύο διαφορετικά νευρωνικά δίκτυα [neural networks (NN)] πίσω μετάδοσης, για την αναπαράσταση της μη γραμμικής συμπεριφοράς (σχέση τάσεων - παραμορφώσεων) ελασμάτων γραφίτη κάτω από μονοτονική και κυκλική φόρτιση. Οι προβλέψεις των NN για την μονοτονική και την κυκλική φόρτιση συμφωνούν πολύ καλά με τα πειραματικά δεδομένα που προκύπτουν από τη βιβλιογραφία. Αρχικά αποτελέσματα υποστηρίζουν την χρήση NN για τη μοντελοποίηση συνθέτων υλικών. Το NN που αναπτύχθηκε στη συγκεκριμένη εργασία βοηθά στην εξακρίβωση σημαντικών θεμάτων που αφορούν την σχέση τάσεων – παραμορφώσεων, όπως είναι η τάση θραύσης και η τάση ρήξης. Ο τρόπος προσέγγισης για μοντελοποίηση με NN, που παρουσιάζεται στο παρών άρθρο, μπορεί επίσης να χρησιμοποιηθεί και στην πρόβλεψη της αστοχίας μηχανισμών.

3.2 New nested adaptive neural networks (NANN) for constitutive modeling (1997)

Συγγραφείς: J. Ghaboussi, D. E. Sidarta

Η ιδέα της χρήσης νευρωνικών δικτύων (NN) στην μοντελοποίηση καταστατικών νόμων έχει προταθεί από τον πρώτο συγγραφέα και τους συναδέλφους του. Σε αυτή τη μεθοδολογία, που παρουσιάζεται στο παρών άρθρο, τα NN εκπαιδεύονται άμεσα, με αποτελέσματα από δοκιμές υλικών και τα εκπαιδευμένα NN μπορούν να χρησιμοποιηθούν στην ανάλυση προβλημάτων συνοριακών τιμών, παρόμοια με οποιοδήποτε άλλο μοντέλο υλικών. Επίσης εισάγεται ένας καινούριος τύπος NN, τα νευρωνικά δίκτυα ένθετης προσαρμοστικής δομής (nested adaptive neural networks (NANN)), τα οποία αναπτύχθηκαν από τον Ghaboussi και τους συναδέλφους του και τα εφαρμόζουν στην μοντελοποίηση της καταστατικής συμπεριφοράς των γεωυλικών. Τα NN προσαρμοστικής δομής (adaptive neural networks) ονομάζονται έτσι διότι προσαρμόζουν αυτόματα την εσωτερική τους δομή (αριθμό κόμβων ή νευρώνων σε κάθε «κρυμμένο επίπεδο») κατά τη διάρκεια της εκπαίδευσης του NN, ενώ τα NN ένθετης προσαρμοστικής δομής (nested adaptive neural networks) εκμεταλλεύονται

την εσωτερική δομή των δεδομένων από δοκιμές σε υλικά και την αντικατοπτρίζουν στην εσωτερική δομή τους (δομή των νευρώνων). Τέλος ο καινούριος τύπος NN εφαρμόζεται στην μοντελοποίηση της αποστραγγισμένης και μη συμπεριφοράς της άμμου σε τριαξονικές δοκιμές.

3.3 Application of neural network to the modeling of some constitutive law (1998)

Συγγραφείς: S. Pernot, C., H. Lamarque

Αυτό το άρθρο ερευνά την μοντελοποίηση καταστατικών νόμων διαφόρων υλικών με νευρωνικά δίκτυα. Η συμπεριφορά του υλικού δεν εκπροσωπείται μαθηματικά πλέον, αλλά περιγράφεται με κάποιο μοντέλο από νευρώνες (νευρωνικό δίκτυο). Κύριος σκοπός της παρούσας εργασίας είναι να δημιουργηθεί ένα νευρωνικό δίκτυο, άμεσα, από πειραματικά αποτελέσματα (στάδιο εκπαίδευσης). Δίνονται αρκετά παραδείγματα μοντελοποίησης καταστατικών νόμων (νόμος Hooke, νόμος Sargin κ.α.) χρησιμοποιώντας έναν αλγόριθμο πίσω μετάδοσης. Τέλος αναπτύσσεται μία μέθοδος κατηγοριοποίησης των καταστατικών νόμων, που θα γίνεται και πάλι χρήση ενός NN, το οποίο ανάλογα με τα δεδομένα που του εισάγονται για κάποιο υλικό, θα προσπαθεί μέσα από μια «βιβλιοθήκη» καταστατικών νόμων να προσδιορίσει σε ποια κατηγορία ανήκει το συγκεκριμένο υλικό.

3.4 An Interface for implicit constitutive modeling (1998)

Συγγραφείς: Furukawa T., Yagawa G.

Στην εργασία αυτή προτείνεται ένα μοντέλο νευρωνικού δικτύου για την μοντελοποίηση της πεπλεγμένης ιξοδοπλαστικότητας των υλικών με σκοπό το προτεινόμενο μοντέλο να είναι όσον το δυνατό καλύτερα εφαρμόσιμο σε πραγματικά υλικά. Τα μοντέλα νευρωνικών δικτύων που είχαν αναπτυχθεί μέχρι εκείνη τη στιγμή περιέγραφαν την ανελαστική συμπεριφορά μεγάλου εύρους υλικών. Το μεγάλο πρόβλημα, όμως της χρήσης τέτοιων μοντέλων είναι η ύπαρξη λαθών αυτών των μοντέλων, το οποίο παραμένει αναπόφευκτα όσο το μοντέλο του υλικού γράφεται αναλυτικά. Σε αυτό το κείμενο προτείνεται μία τεχνική για την εξαγωγή δεδομένων εισόδου – εξόδου για την κατασκευή του προτεινόμενου μοντέλου νευρωνικού δικτύου από πειραματικά δεδομένα. Δοκιμασμένο με πραγματικά πειραματικά δεδομένα το προτεινόμενο μοντέλο κατέληξε σε ένα μηδαμινό πλήθος λαθών δείχνοντας έτσι την υπεροχή του σε σχέση με όλα τα ήδη υπάρχοντα αναλυτικά μοντέλα ως προς την ακρίβεια.

3.5 Implicit constitutive modeling for viscoplasticity using neural networks (1998)

Συγγραφείς: Furukawa T., Yagawa G.

Το παρών άρθρο έχει το ίδιο αντικείμενο με το προηγούμενο. Οι συγγραφείς προτείνουν ένα μοντέλο νευρωνικού δικτύου για την μοντελοποίηση της πεπλεγμένης ιξοδοπλαστικότητας των υλικών, με σκοπό να αντικαταστήσει τα ήδη υπάρχοντα, μέχρι εκείνη τη στιγμή, αναλυτικά μοντέλα ανελαστικής συμπεριφοράς υλικών. Περιγράφεται επίσης και η τεχνική με την οποία εξάγονται τα δεδομένα εισόδου – εξόδου για το νευρωνικό δίκτυο από πειραματικά δεδομένα. Τέλος δοκιμάζεται με πραγματικά δεδομένα και καταλήγει σε μικρό αριθμό λαθών, γεγονός που αναδεικνύει την υπεροχή του, σε σχέση με τα αναλυτικά μοντέλα.

3.6 Neural network based constitutive model for rubber material (2004)

Συγγραφείς: Shen V., Chandhashekhara K., Breig W. F., Oliver L. R.

Το συγκεκριμένο άρθρο προτείνει μια καινούρια μέθοδο προσδιορισμού της συνάρτησης ενέργειας παραμόρφωσης υλικών όπως το λάστιχο, μέσω ενός νευρωνικού δικτύου πίσω μετάδοσης. Η υπερελαστικότητα του λάστιχου χαρακτηρίζεται από μία συνάρτηση ενέργειας παραμόρφωσης. Οι συναρτήσεις ενέργειας παραμόρφωσης κατηγοριοποιούνται, κατά κύριο λόγο, σε δύο κατηγορίες: **(α)** η μία βασίζεται στη στατιστική θερμοδυναμική και **(β)** η δεύτερη βασίζεται σε μία φαινομενολογική προσέγγιση που θεωρεί το υλικό συνεχές. Αυτή η εργασία επικεντρώνεται στην δεύτερη κατηγορία, την φαινομενολογική προσέγγιση. Για τον προσδιορισμό των σταθερών της συνάρτησης της ενέργειας παραμόρφωσης με αυτή τη μέθοδο, απαιτείται προσαρμογή στην καμπύλη που προκύπτει από δεδομένα δοκιμών πάνω σε υλικά από λάστιχο. Μια ανασκόπηση σε διαθέσιμες συναρτήσεις ενέργειας παραμόρφωσης που βασίζονται στη φαινομενολογική προσέγγιση δείχνει ότι απαιτείται αρκετή προσπάθεια για να επιτευχθεί προσαρμογή στην καμπύλη με καλή ακρίβεια. Αυτό το πρόβλημα, προσπαθεί να λύσει η προτεινόμενη μέθοδος προσδιορισμού της ενέργειας παραμόρφωσης. Στο παρών άρθρο επεξηγείται λεπτομερώς ο υπολογισμός αυτής και των παραγώγων της μέσω νευρωνικού δικτύου, καθώς επίσης και η προετοιμασία για την εκπαίδευση του νευρωνικού δικτύου από δεδομένα δοκιμών πάνω σε λάστιχα. Τα αποτελέσματα από την προσαρμογή στην πειραματική καμπύλη με την καινούρια μέθοδο δείχνουν την αποτελεσματικότητα και την ακρίβεια της προσέγγισης μέσω νευρωνικών δικτύων. Ένα μοντέλο υλικού, που βασίζεται σε νευρωνικό δίκτυο, τέθηκε σε εφαρμογή και εφαρμόστηκε στην προσομοίωση ιμάντα τύπου V χρησιμοποιώντας κώδικα πεπερασμένων στοιχείων στο πρόγραμμα ABAQUS.

3.7 Self-learning elements for inverse estimation of thermal constitutive models (2005)

Συγγραφείς: Wilkins Aquino, John C. Brigham

Στο παρών κείμενο εισάγεται μια καινούρια μεθοδολογία για την αντίστροφη εκτίμηση των θερμικών καταστατικών μοντέλων. Η μέθοδος αυτή συνδυάζει την υπολογιστική ευφυΐα με την ανάλυση πεπερασμένων στοιχείων για την αντίστροφη επίλυση προβλημάτων μεταφοράς θερμότητας. Παρουσιάζεται η αναπαράσταση της θερμικής καταστατικής συμπεριφοράς από ένα νευρωνικό δίκτυο και η εφαρμογή του σε μη γραμμική ανάλυση πεπερασμένων στοιχείων. Η μεθοδολογία αυτό – εκπαίδευσης (self-learning simulations) χρησιμοποιεί μια καινούρια ιδέα για την ανάπτυξη μοντέλων υλικών κάνοντας χρήση πειραματικών δεδομένων και επαναληπτική ανάλυση πεπερασμένων στοιχείων. Η προτεινόμενη μεθοδολογία ψάχνει για ολοκληρωμένα θερμικά καταστατικά μοντέλα, σε αντίθεση με τον προσδιορισμό παραμέτρων σε προκαθορισμένες συναρτησιακές μορφές. Η εφαρμογή της καινούριας μεθοδολογίας επεξηγείται χρησιμοποιώντας ένα προσομοιωμένο πρόβλημα σταθερής μετάδοσης θερμότητας με επαφή. Βρέθηκε λοιπόν, ότι σε προσομοιωμένα πειράματα η μέθοδος αυτό – εκπαίδευσης (SelfSim) πεπερασμένων στοιχείων μπορεί να ανακτήσει αντίστροφα θερμικά καταστατικά μοντέλα NN χρησιμοποιώντας απλές μετρήσεις θερμοκρασιών. Επιπλέον η μέθοδος έδωσε ευστάθεια στην παρουσία ατελών δεδομένων ή δεδομένων θορύβου. Τέλος αποδείχθηκε ότι η χρήση NN στην αναπαράσταση καταστατικών μοντέλων βελτιώνει την ευστάθεια των λύσεων, φυσικά σε σχέση με το περιθώριο ανακρίβειας του NN.

3.8 Neural network material model enhancement: Optimization through visualization and selective data removal (2005)

Συγγραφείς: Butkovich J. N., Hashash Y. M. A.

Τα μοντέλα καταστατικών νόμων βασιζόμενα στα νευρωνικά δίκτυα (NN) έχουν χρησιμοποιηθεί πρόσφατα για να αναπαραστήσουν την καταστατική συμπεριφορά του εδάφους. Τα NN εκπαιδεύονται με μεγάλο αριθμό δεδομένων. Σε αυτό το άρθρο παρουσιάζεται και δοκιμάζεται μια καινούρια μέθοδος εμπλουτισμού των δεδομένων εκπαίδευσης, δείχνοντας ότι ένα περιορισμένο σύνολο δεδομένων μπορεί να επεκταθεί σε ένα μεγαλύτερο εύρος προβλημάτων. Επίσης παρουσιάζεται μία μέθοδος επιλεκτικής μείωσης του αριθμού των δεδομένων εκπαίδευσης, σύμφωνα με την οποία, η συμπεριφορά για την οποία εκπαιδεύταν ένα NN μέσω ενός μεγάλου αριθμού δεδομένων, η ίδια συμπεριφορά μπορεί να προκύψει τώρα, από ένα ελάχιστο αριθμό δεδομένων, επιλεγμένων κατάλληλα με την παραπάνω μέθοδο. Τέλος,

χρησιμοποιώντας νέες εικονικές τεχνικές για την αξιολόγηση της απόκρισης των εκπαιδευμένων μοντέλων NN, γίνεται κατανοητό ότι η μεγάλη μείωση του αριθμού των δεδομένων εκπαίδευσης δεν επηρεάζει σημαντικά την απόδοση του μοντέλου του NN.

3.9 Characterizing rate-dependent material behaviors in self-learning simulation (2006)

Συγγραφείς: Sungmoon Jung, Jamshid Ghaboussi

Στις δομικές δοκιμές, η μη ομοιογενής κατανομή των τάσεων που υπάρχει μέσα στα δείγματα, παρέχει περισσότερες πληροφορίες για την συμπεριφορά του υλικού σε σχέση με τις συμβατικές δοκιμές των υλικών όπου υπάρχει ομοιόμορφη κατανομή των τάσεων. Οι συγγραφείς παρουσιάζουν μία μεθοδολογία αναπαράστασης της συμπεριφοράς υλικών που εξαρτώνται από το ρυθμό των τάσεων και των παραμορφώσεων κάνοντας χρήση μετρήσεων φορτίου – μετατόπισης από δομικές δοκιμές. Η ικανότητα της μεθόδου αυτό – εκπαίδευσης (SelfSim) για τα μοντέλα υλικών από NN που εξαρτώνται από το ρυθμό τάσεων και παραμορφώσεων, που προτάθηκε νωρίτερα από τους συγγραφείς, χρησιμοποιήθηκε σε συνδυασμό με αυτή τη μεθοδολογία. Σε αντίθεση με άλλες μεθόδους παραμετρικής βελτιστοποίησης δεν χρειάζεται προγενέστερη γνώση της συμπεριφοράς του υλικού. Το μοντέλο είναι επίσης ικανό να βελτιώσει την απόδοσή του, όσο τα δεδομένα δοκιμών είναι διαθέσιμα. Τέλος σαν ένα επεξηγηματικό παράδειγμα όλων των παραπάνω, η μέθοδος εφαρμόζεται, ώστε να περιγράψει την μη γραμμική συμπεριφορά ερπυσμού του κράματος SRR99.

3.10 Neural network constitutive model for rate – dependent materials (2006)

Συγγραφείς: Sungmoon Jung, Jamshid Ghaboussi

Τα καταστατικά μοντέλα νευρωνικών δικτύων (NN) προσαρμόζονται έτσι ώστε, να μπορούν να περιγράψουν την σχέση τάσεων – παραμορφώσεων ενός υλικού. Είναι ικανά να αποδώσουν πολύπλοκες συμπεριφορές υλικών χρησιμοποιώντας ζεύγη τάσεων – παραμορφώσεων από πειράματα. Σε αυτό το άρθρο παρουσιάζεται ένα καταστατικό μοντέλο NN για υλικό, εξαρτώμενο από το ρυθμό τάσεων και παραμορφώσεων και επίσης, η εφαρμογή του στην ανάλυση πεπερασμένων στοιχείων. Το προτεινόμενο μοντέλο NN επαληθεύεται για ένα πρότυπο μοντέλο βισκοελαστικού στερεού και στη συνέχεια εφαρμόζεται στην ανάλυση της εξαρτώμενης από το ρυθμό των σ και ϵ συμπεριφοράς του σκυροδέματος. Το μοντέλο αυτό έχει τη δυνατότητα να περιγράψει οποιοδήποτε μοντέλο υλικού εξαρτώμενου από το ρυθμό των σ και ϵ , παρέχοντας αρκετά δοσμένα ζεύγη δεδομένων. Τέλος προτείνεται να

ερευνηθεί μελλοντικά, από τι πρέπει να αποτελείται ένα επαρκές σύνολο δεδομένων για την εκπαίδευση του NN.

3.11 Constitutive soil behavior representation via artificial neural networks: A shift from soil models to soil behavior data (2006)

Συγγραφείς: Hashash Y. M. A., Ghaboussi J., Fu Q., Maruranda C.

Την τελευταία δεκαετία τα νευρωνικά δίκτυα (NN) έχουν προταθεί ως εναλλακτικές λύσεις έναντι των συμβατικών ελαστοπλαστικών καταστατικών μοντέλων. Το μοντέλο NN για τα υλικά μπορεί να χρησιμοποιηθεί με παρόμοιο τρόπο όπως τα συμβατικά μοντέλα στην επίλυση προβλημάτων συνοριακών συνθηκών. Στην παρούσα εργασία οι συγγραφείς έχουν αναπτύξει τη μέθοδο αυτό – εκπαίδευσης (self-learning simulations) στην ανάλυση προσομοιωμένης μηχανικής, η οποία εξάγει σχετικές πληροφορίες για τη συμπεριφορά του εδάφους κάνοντας χρήση συνοριακών μετρήσεων φορτίου – μετατόπισης, διευκολυνόμενη από το μοντέλο NN. Σύμφωνα με αυτή τη μέθοδο το πιο σημαντικό κομμάτι της μοντελοποίησης είναι η συλλογή των δεδομένων τάσεων – παραμορφώσεων που θα πρέπει να χρησιμοποιηθούν για την εκπαίδευση και την ανάπτυξη του υλικού μοντέλου.

3.12 Novel approach to integration of numerical modeling and field observation for deep excavations (2006)

Συγγραφείς: Hashash Y. M. A., Maruranda C., Ghaboussi J., Jung S.

Η παρατήρηση της επίδοσης μιας κατασκευής είναι βασικό κομμάτι της σχεδίασης και της κατασκευής στην γεωτεχνική μηχανική. Στις βαθιές αστικές εκσκαφές οι σχεδιαστές βασίζονται σε εμπειρικά δεδομένα για την εκτίμηση δυνητικών μετατοπίσεων και επιπτώσεων σε παρακείμενες δομές. Οι αριθμητικές προσομοιώσεις χρησιμοποιούνται για την εκτίμηση των προκαλούμενων παραμορφώσεων του εδάφους. Σημαντικοί πόροι προορίζονται για την παρακολούθηση των δραστηριοτήτων κατασκευής και ελέγχου των προκαλούμενων παραμορφώσεων του εδάφους. Ενώ οι μηχανικοί έχουν τη δυνατότητα να μάθουν από τις παρατηρήσεις, διάφορες αριθμητικές προσομοιώσεις ήταν σε θέση να επωφεληθούν πλήρως από τις πληροφορίες που περιέρχονται σε μια δεδομένη τοποθεσία ή από το ιστορικό πριν τις ανασκαφές στην ίδια περιοχή. Μια νέα μέθοδος ανάλυσης, η μέθοδος της αυτό – εκπαίδευσης (SelfSim) εισάγεται στην προσομοιωμένη μηχανική για την ολοκλήρωση των αριθμητικών προσομοιώσεων. Η μέθοδος SelfSim είναι μια τεχνική αντίστροφης ανάλυσης, που συνδυάζει τη μέθοδο των πεπερασμένων στοιχείων, μοντέλα υλικών εμπνευσμένα από τη βιολογία (τεχνητά νευρωνικά δίκτυα) και τις μετρήσεις εδάφους. Μπορεί να εξάγει σχετικές πληροφορίες για την καταστατική σχέση του εδάφους από μετρήσεις του εδάφους στο χώρο της

ανασκαφής, όπως είναι οι παράπλευρες παραμορφώσεις στα τοιχώματα και η επιφανειακή καθίζηση. Το προκύπτων, από την παραπάνω μεθοδολογία, μοντέλο NN του εδάφους χρησιμοποιείται στην αριθμητική ανάλυση, παρέχοντας σωστές παραμορφώσεις του εδάφους και μπορεί επίσης να χρησιμοποιηθεί στην εκτίμηση παραμορφώσεων παρόμοιων εκσκαφών. Το πλεονέκτημα του είναι πως μπορεί συνεχώς να εξελίσσεται χρησιμοποιώντας επιπρόσθετες πληροφορίες από το έδαφος. Τέλος για την απόδειξη της αξιοπιστίας της μεθόδου SelfSim, γίνεται χρήση της σε δύο περιπτώσεις εκσκαφών στη Βοστώνη και το Σικάγο.

3.13 Constitutive modeling of Sn/Ag and Sn/Ag/Cu solder alloys (2006)

Συγγραφείς: Min P., Jianmin Q.

Οι συγκολλήσεις χρησιμοποιούνται εκτενέστερα σαν ηλεκτρικές συνδέσεις στις μικροηλεκτρονικές συσκευές. Τα κράματα βασισμένα στο μόλυβδο υπάρχουν εκ των προτέρων ως μέταλλα συγκόλλησης, λόγω της χαμηλής θερμοκρασίας τήξης τους, της υψηλής ολκιμότητας και της καλής διαβρεκτικής ικανότητας με άλλα μέταλλα. Λόγω των περιβαλλοντικών ζητημάτων τα κράματα μολύβδου πρόκειται να αντικατασταθούν από κράματα βασισμένα στα υλικά Sn/Ag και Sn/Ag/Cu. Δεδομένου της θερμομηχανικής αξιοπιστίας των σύγχρονων ηλεκτρονικών συσκευών που εξαρτάται, σε μεγάλο βαθμό, από την κόπωση και την συμπεριφορά ερπυσμού των συνδέσεων συγκόλλησης, είναι επιτακτική ανάγκη να κατανοηθεί η κόπωση και η συμπεριφορά ερπυσμού αυτών των νέων χωρίς μόλυβδο κραμάτων. Αρκετά, ήδη υπάρχοντα, ελαστοπλαστικά, βισκοπλαστικά μοντέλα και μοντέλα ερπυσμού έχουν χρησιμοποιηθεί ευρέως στο παρελθόν για να περιγράψουν την θερμομηχανική συμπεριφορά των κραμάτων μολύβδου. Αυτά τα καταστατικά μοντέλα μπορούν επίσης να χρησιμοποιηθούν και στην περίπτωση των κραμάτων από Sn/Ag και Sn/Ag/Cu. Παρόλ' αυτά η καταλληλότητα αυτών των καταστατικών μοντέλων για τα νέα κράματα δεν έχει ευρέως καθιερωθεί. Επιπροσθέτως και σε αντίθεση με τα κράματα μολύβδου για συγκολλήσεις, υπάρχει πολύ περιορισμένος αριθμός πειραματικών δεδομένων για τις θερμομηχανικές ιδιότητες και ιδιότητες ερπυσμού για τα κράματα Sn/Ag και Sn/Ag/Cu.

Στο παρών άρθρο λοιπόν, αναφέρονται κάποια αποτελέσματα από μία πρόσφατη μελέτη που έγινε για την ανάπτυξη καταστατικών νόμων κατάλληλων για τα νέα κράματα. Εκτεταμένες θερμομηχανικές δοκιμές έχουν διεξαχθεί σε αυτά τα κράματα κάτω από ποικίλες θερμοκρασίες (-55C ~ 150C) και ρυθμό τάσης (10^{-7} /s ~ 10^{-1} /s) υποβαλλόμενες και σε μονοτονική φόρτιση και κυκλική φόρτιση. Τα πειραματικά δεδομένα χρησιμοποιούνται στη συνέχεια για την ανάπτυξη ενός βισκοελαστικού καταστατικού μοντέλου για την περιγραφή της θερμομηχανικής συμπεριφοράς των

νέων κραμάτων. Επιπλέον ένα νευρωνικό δίκτυο (NN) σαν υπολογιστικό εργαλείο αναπτύσσεται και εκπαιδεύεται. Το εργαλείο αυτό χρησιμοποιείται για να εξάγει διάφορες παραμέτρους του καταστατικού μοντέλου που βασίζεται σε πειραματικά δεδομένα. Η εγκυρότητα και η ακρίβεια αυτού του εργαλείου έχει ήδη επαληθευτεί από τη χρήση του μοντέλου Anand σε ένα εφικτικό κράμα Sn/Pd. Όταν το NN εκπαιδευτεί είναι ικανό να εξάγει και τις 11 παραμέτρους του μοντέλου Anand χρησιμοποιώντας, πειραματικά λαμβανόμενες καμπύλες τάσης – παραμόρφωσης. Οι παραπάνω παράμετροι μπορούν άμεσα να χρησιμοποιηθούν σε ένα κοινό λογισμικό πεπερασμένων στοιχείων όπως το ANSYS ή το ABAQUS για την ανάλυση της αξιοπιστίας των συνδέσεων συγκόλλησης. Τέλος κάνοντας χρήση αυτού του NN μπορούν να αποκτηθούν όλες οι παράμετροι (και οι 11) στο μοντέλο Anand για το Sn/Ag και για το Sn/Ag/Cu.

3.14 A new neural network – based model for hysteretic behavior (2007)

Συγγραφείς: Gun Jin Yun, Jamshid Ghaboussi, Amr S. Elnashai

Η κυκλική συμπεριφορά των υλικών είναι πολύ περίπλοκη και δύσκολη στο να μοντελοποιηθεί. Ένας συνδυασμός από κανόνες σκλήρυνσης της κλασικής θεωρίας πλαστικότητας, είναι μία πιθανή λύση για τη μοντελοποίηση αυτής της πολύπλοκης συμπεριφοράς. Έχει δειχθεί στο παρελθόν ότι τα καταστατικά μοντέλα νευρωνικών δικτύων έχουν την ικανότητα να μοντελοποιήσουν πολύπλοκες υλικές συμπεριφορές, άμεσα, από τα αποτελέσματα των δοκιμών σε υλικά. Σε αυτό το άρθρο προτείνεται μία καινούρια προσέγγιση για μοντελοποίηση της κυκλικής συμπεριφοράς των υλικών, μέσω νευρωνικών δικτύων (NN). Το προτεινόμενο μοντέλο υλικού με NN χρησιμοποιεί καινούριες εσωτερικές μεταβλητές που διευκολύνουν την εκμάθηση (στάδιο εκπαίδευσης του NN) της υστερητικής συμπεριφοράς των υλικών. Η ίδια προσέγγιση μπορεί να χρησιμοποιηθεί για τη μοντελοποίηση της υστερητικής συμπεριφοράς των δομικών συστημάτων ή των δομικών στοιχείων κάτω από κυκλικές φορτίσεις. Αποδεικνύεται επίσης ότι το παρών μοντέλο NN είναι ανώτερο από τα προγενέστερα μοντέλα NN για τέτοιου είδους υλικά. Αν και τα προγενέστερα μοντέλα NN ήταν αποτελεσματικά στην απόδοση της πολυαξονικής συμπεριφοράς των υλικών, υφίσταντο δοκιμές μόνο κάτω από κυκλική μονοαξονική τάση. Το προτεινόμενο μοντέλο NN έχει τη δυνατότητα να εκπαιδευτεί για την υστερητική συμπεριφορά των υλικών ακόμα και σε κατάσταση μη – ομοιόμορφης τάσης για πολυδιάστατα προβλήματα. Η απόδοση του μοντέλου NN αποδεικνύεται μέσα από μία σειρά παραδειγμάτων κάνοντας χρήση πειραματικών δεδομένων και δεδομένων από προσομοιωμένες δοκιμές.

3.15 Nonlinear constitutive models for FRP composites using artificial neural networks (2007)

Συγγραφείς: Rami Haj – Ali, Hoan – Kee Kim

Αυτό το άρθρο παρουσιάζει μία καινούρια προσέγγιση για τη δημιουργία μη γραμμικών και πολυαξονικών καταστατικών μοντέλων για σύνθετα υλικά από ενισχυμένες πολυμερικές ίνες, με χρήση τεχνητών νευρωνικών δικτύων (ANNs). Τα νέα μη γραμμικά καταστατικά μοντέλα NN είναι πλήρη και έχουν ολοκληρωθεί, ακόμα περισσότερο, με ένα λογισμικό πεπερασμένων στοιχείων για την μη γραμμική ανάλυση δομών από σύνθετα υλικά. Τα προτεινόμενα καταστατικά μοντέλα νευρωνικών δικτύων (NN) εκπαιδεύονται με πειραματικά αποτελέσματα που έχουν προκύψει από δοκιμές εφελκυσμού / συμπίεσης εκτός άξονα και καθαρής διάτμησης (Arcan). Επίσης τα μοντέλα NN έχουν δημιουργηθεί για κατάσταση επίπεδης τάσης με υποθετική λειτουργική απόκριση σε ορισμένα τμήματα της πολυαξονικής τάσης χωρίς πειραματικά δεδομένα. Η ικανότητα των εκπαιδευμένων μοντέλων NN να προβλέπουν την απόκριση των υλικών, εξετάζεται άμεσα μέσα από την ανάλυση πεπερασμένων στοιχείων σε μία αυλακοφόρο πλάκα από σύνθετα υλικά. Το πειραματικό κομμάτι αυτής της μελέτης εμπεριέχει και μία δοκιμή διάβρωσης μιας λεπτής διατομής από ένα υλικό ενισχυμένης πολυμερικής ίνας από γυαλί και πολυεστέρα με βάση συγκεκριμένη διαδικασία παραγωγής για προηγμένες συνθέσεις. Η μη γραμμική αντίδραση ήταν έντονη στην κατεύθυνση των ινών λόγω του σχετικά χαμηλού συνολικού δείκτη όγκου των ινών. Η λαμαρίνες από σύνθετο υλικό με εγκοπή δοκιμάστηκαν επίσης για να επαληθεύσουν την ανάλυση πεπερασμένων στοιχείων, με μοντέλα ANN, για την πρόβλεψη γενικών μη ομογενοποιημένων αποκρίσεων σε δομικό επίπεδο.

3.16 Integration of laboratory testing and constitutive modeling of soils (2007)

Συγγραφείς: Qingwei Fu, Youssef M. A. Hashash, Sungmoon Jung, Jamshid Ghaboussi

Ένα καταστατικό μοντέλο του εδάφους, το οποίο αποτυπώνει σωστά τη συμπεριφορά του εδάφους κάτω από γενική φόρτιση είναι αναγκαίο για την επίλυση πολύπλοκων συνοριακών προβλημάτων που αφορούν στην γεωτεχνική μηχανική. Διαθέσιμες εργαστηριακές δοκιμές παρέχουν πληροφορίες για τη συμπεριφορά του υλικού μέσα σε ένα πολύ περιορισμένο εύρος διαδρομών τάσης – παραμόρφωσης και δεν καλύπτουν όλο το εύρος των φορτίσεων στις οποίες υποβάλλεται σε προβλήματα συνοριακών τιμών. Οι πληροφορίες για τη συμπεριφορά του εδάφους, που αναπτύχθηκαν από τις περισσότερες εργαστηριακές δοκιμές είναι συχνά περιορισμένες

και ανεπαρκής να επικυρώσουν την απόδοση του καταστατικού μοντέλου κάτω από συνθήκες γενικής φόρτισης.

Στο παρών άρθρο ερευνάται μία νέα προσέγγιση για τη σύνδεση των εργαστηριακών δοκιμών του εδάφους με την ανάπτυξη καταστατικών μοντέλων, μέσα από τη χρήση ενός καινούριου υπολογιστικού πλαισίου: προσομοιώσεις αυτό – εκπαίδευσης (SelfSim). Η μέθοδος SelfSim είναι μία αντίστροφη ανάλυση, που εξάγει την καταστατική σχέση που αντιπροσωπεύει το υλικό από μετρήσεις συνοριακών φορτίων και μετατοπίσεων. Η προτεινόμενη μέθοδος εφαρμόζεται σε δύο προσομοιωμένες εργαστηριακές δοκιμές, μία δοκιμή τριαξονικής συμπίεσης χωρίς τριβή ολίσθησης και μία δοκιμή τριαξονικής στρέψης χωρίς τριβή ολίσθησης. Η τριβή ολίσθησης είναι αποτέλεσμα της μη ομοιόμορφης τάσης και παραμόρφωσης μέσα σε όλο το δείγμα. Η μέθοδος SelfSim εξάγει με επιτυχία μεγάλη ποικιλία από ζεύγη τάσης – παραμόρφωσης που υφίστανται μέσα στο δείγμα. Ένα καταστατικό μοντέλο NN αναπτύσσεται χρησιμοποιώντας την εξαγόμενη συμπεριφορά του εδάφους και από τα δύο είδη δοκιμών και χρησιμοποιείται με επιτυχία στην πρόβλεψη της καθίζησης του εδάφους από την προσομοίωση ενός πέλματος σε σχήμα λωρίδας. Τα αποτελέσματα δείχνουν, ότι η μέθοδος SelfSim καθιερώνει μία άμεση σύνδεση ανάμεσα στις εργαστηριακές δοκιμές και τη καταστατική μοντελοποίηση του εδάφους, ώστε να εξάγει τη συμπεριφορά του, κάτω από πολύπλοκες φορτίσεις και να οδηγήσει στην άμεση ανάπτυξη αντίστοιχων καταστατικών μοντέλων.

3.17 Extracting inelastic metal behavior through inverse analysis: a shift in focus from material models to material behavior (2007)

Συγγραφείς: Y. M. A. Hashash, H. Song, S. Jung, J. Ghaboussi

Στην παρούσα εργασία τίθεται σε εφαρμογή μία καινούρια μέθοδος αντίστροφης ανάλυσης, η μέθοδος SelfSim (self learning simulations) από την οποία μπορεί να παραχθεί η συμπεριφορά διαφόρων μετάλλων. Η μέθοδος SelfSim χρησιμοποιεί μετρήσεις φορτίου – μετατόπισης από δοκιμές σε διάφορες δομές, μέσω των οποίων το υλικό υποβάλλεται σε μη ομοιόμορφες τάσεις και παραμορφώσεις, ώστε να προκύψει η καταστατική σχέση που αντιπροσωπεύει τη συμπεριφορά του υλικού, στη μορφή δεδομένων ζευγών τάσης – παραμόρφωσης ενός προκαθορισμένου μοντέλου του υλικού. Η προτεινόμενη μέθοδος επαληθεύεται με χρήση τόσο προσομοιωμένων όσο και φυσικών πειραμάτων σε μεταλλικές δομές. Έτσι από πολλαπλές δοκιμές που έγιναν στο αλουμίνιο αποδεικνύεται ότι η μέθοδος SelfSim παρέχει με επιτυχία την ανισοτροπική συμπεριφορά του. Τέλος γίνεται αντιληπτό ότι η προτεινόμενη μέθοδος, αυτή της εργασίας, απλουστεύει την επίπονη και παρατεταμένη διαδικασία δημιουργίας

ενός συμβατικού μοντέλου, το οποίο χαρακτηρίζει την καταστατική συμπεριφορά ενός νέου υλικού.

3.18 Neural network based constitutive model for elastomeric foams (2007)

Συγγραφείς: G. Liang, K. Chandrashekhara

Τα υλικά από ελαστομερή αφρό βρίσκουν, ευρείας κλίμακας, εφαρμογή λόγω των εξαιρετικών ιδιοτήτων τους στην απορρόφηση ενέργειας. Οι μηχανικές ιδιότητες των ελαστομερών αφρών είναι έντονα μη γραμμικές και είναι ουσιώδεις για την εφαρμογή μαθηματικών καταστατικών μοντέλων, ικανών να αναπαραστήσουν με ακρίβεια την συμπεριφορά των αφρών σε μορφή τάσεων – παραμορφώσεων. Σε αυτό το άρθρο παρουσιάζεται μία νέα μέθοδος καταστατικής μοντελοποίησης για τον προσδιορισμό της συνάρτησης της υπεραφρώδης ενέργειας παραμόρφωσης, μέσω νευρωνικών δικτύων (NN). Επίσης περιγράφεται η δομή των τεχνητών νευρωνικών δικτύων και η προετοιμασία του NN από δεδομένα δοκιμών σε αφρούς, ενώ επεξηγείται λεπτομερώς ο τρόπος υπολογισμού της ενέργειας παραμόρφωσης και των παραγώγων της, μέσω του NN. Τα αποτελέσματα από την προσαρμογή καμπυλών δίνονται για να παρουσιάσουν την απόδοση και την ακρίβεια του NN. Το καταστατικό μοντέλο NN για ελαστομερή αφρούς εφαρμόζεται στην προσομοίωση μιας διαδικασίας αυλακώσεως αφρώδους υλικού, επίπεδης παραμόρφωσης, για να καταδείξει την εφαρμογή και την απόδοση του NN στην ανάλυση πεπερασμένων στοιχείων. Τα αποτελέσματα δείχνουν ότι το NN παρέχει καλύτερη αναπαράσταση των δεδομένων από τις δοκιμές, σε σχέση με το ήδη χρησιμοποιούμενο μοντέλο Hyperfoam.

3.19 Application of artificial neural network for constitutive modeling in finite element analysis (2007)

Συγγραφείς: Javadi A. A., Tan T. P., Elkassas A. S. I.

Σε αυτό το άρθρο παρουσιάζεται ένα μοντέλο πεπερασμένων στοιχείων και νευρωνικών δικτύων για τη μοντελοποίηση προβλημάτων γεωτεχνικής μηχανικής. Η προτεινόμενη μεθοδολογία προϋποθέτει την ένταξη του καταστατικού μοντέλου NN σε ένα πρόγραμμα πεπερασμένων στοιχείων σαν αντικατάσταση των συμβατικών καταστατικών μοντέλων υλικών. Η ανάπτυξη του αλγορίθμου που θα παρουσιαστεί ακολουθείται από την εφαρμογή μιας εφυσής μεθόδου πεπερασμένων στοιχείων σε έναν αριθμό από προβλήματα μηχανικής. Οι προς μελέτη περιπτώσεις θεωρείται, ότι περιλαμβάνουν γραμμικά ελαστική, μη γραμμικά ελαστική και ελαστοπλαστική συμπεριφορά υλικών. Τα αποτελέσματα παρέχονται από την ανάλυση με χρήση της εφυσής μεθόδου πεπερασμένων στοιχείων, τα οποία συγκρίνονται με τα

αποτελέσματα που παρέχονται από την πρότυπη μέθοδο των πεπερασμένων στοιχείων, με διαφορετικά συμβατικά καταστατικά μοντέλα. Τα τελικά αποτελέσματα δείχνουν ότι τα τεχνητά NN έχουν καλύτερη δυνατότητα εκμετάλλευσης, σαν μια ενιαία προσέγγιση για την καταστατική μοντελοποίηση στην ανάλυση των πεπερασμένων στοιχείων. Τέλος γίνεται λεπτομερής αναφορά στα πλεονεκτήματα και τους περιορισμούς της μοντελοποίησης της καταστατικής συμπεριφοράς υλικών μέσω NN.

3.20 Development of neural network based hysteretic models for steel beam- column connections through self-learning simulation (2007)

Συγγραφείς: Yun G. J., Ghaboussi J., Elnashai A. S.

Οι συνδέσεις δοκαριού – υποστηλώματος είναι ζώνες με έντονα πολύπλοκη δράση και παραμορφώσεις αλληλεπίδρασης που συχνά οδηγούνται σε αστοχία υπό την επίδραση σεισμικών δονήσεων. Η μοντελοποίηση της σύνδεσης δοκαριού – υποστηλώματος είναι σημαντική τόσο για την κατανόηση της συμπεριφοράς της σύνδεσης, όσο και για τον σχεδιασμό. Στο παρών άρθρο, προτείνεται ένα πλαίσιο για την ανάπτυξη ενός μοντέλου νευρωνικού δικτύου για την σύνδεση δοκαριού-υποστηλώματος μέσω δοκιμών στη συγκεκριμένη δομή. Το νευρωνικό δίκτυο, που αντιπροσωπεύει την ανελαστική υστερητική συμπεριφορά αυτής της δομής, συνδυάζεται με ένα νέο μοντέλο πεπερασμένων στοιχείων αποτελούμενο από περισσότερα του ενός πεπερασμένα στοιχεία κάτω από το πλαίσιο της μεθόδου self – learning simulation (προσομοιώσεις αυτό - εκπαίδευσης). Η μέθοδος SelfSim έχει ένα μοναδικό πλεονέκτημα να μπορεί να χρησιμοποιήσει την απόκριση της δομής ώστε να εξάγει μοντέλα NN των υλικών. Επίσης βασίζεται στον αλγόριθμο auto – progressive, που εφαρμόζει τις αρχές της ισορροπίας και της συμβατότητας και την ικανότητα αυτό – οργάνωσης των νευρωνικών δικτύων. Το συστατικό μοντέλο είναι ένα σύνολο από στερεά στοιχεία και στοιχεία ελατηρίων, το οποίο αντιπροσωπεύει την καταστατική συμπεριφορά των στοιχείων που το αποτελούν, τόσο μη γραμμική ελαστική, όσο και μη γραμμική ανελαστική συμπεριφορά. Το μοντέλο πεπερασμένων στοιχείων με περισσότερα του ενός στοιχεία επαληθεύεται από την τριδιάστατη ανάλυση πεπερασμένων στοιχείων. Η προτεινόμενη μεθοδολογία παρουσιάζεται μέσα από τη μέθοδο SelfSim για μία συγκολλούμενη σύνδεση δοκαριού – υποστηλώματος. Τέλος ορίστηκε ένα πλαίσιο για την εφαρμογή της προτεινόμενης μεθοδολογίας σε άλλους τύπους συνδέσεων.

3.21 Neural network material model enhancement: optimization through selective data removal (2007)

Συγγραφείς: Butcovich J. N., Hashash Y. M. A.

Τα καταστατικά μοντέλα νευρωνικών δικτύων (NN) έχουν χρησιμοποιηθεί κατά κόρον για την περιγραφή της συμπεριφοράς του εδάφους. Όταν τα NN συνδυάστηκαν με τη μέθοδο SelfSim (μέθοδος αντίστροφης ανάλυσης), τα μοντέλα NN μπορούν να χρησιμοποιηθούν για την εξαγωγή της συμπεριφοράς του εδάφους με δοσμένες μετρήσεις από συνοριακές παραμορφώσεις και φορτία. Παρολ' αυτά τα ζεύγη δεδομένων που χρησιμοποιούνται για να εκπαιδεύσουν και να επανεκπαιδεύσουν τα μοντέλα NN είναι μεγάλα και οι χρόνοι εκπαίδευσης, ειδικά όταν χρησιμοποιούνται στην μέθοδο SelfSim, είναι μεγάλοι. Μία ποικιλία ζευγών από δεδομένα τάσης – παραμόρφωσης εξάγονται από ένα προσομοιωμένο πρόβλημα εκσκαφής για την εκπαίδευση του καταστατικού μοντέλου NN. Αρκετές μέθοδοι, για τη μείωση του μεγέθους των δεδομένων, προτείνονται και αξιολογούνται. Κάθε μία από αυτές τις μεθόδους διαγράφουν επιλεκτικά τα δεδομένα εκπαίδευσης, ώστε το μικρότερο πλήθος δεδομένων να χρησιμοποιείται για την εκπαίδευση του NN. Η μέθοδος των σημείων Gauss διαγράφει τα δεδομένα στηριζόμενη στη θέση του κάθε πεπερασμένου στοιχείου στο μοντέλο. Η μέθοδος πλέγματος (lattice) διαγράφει τα δεδομένα, έτσι ώστε τα σημεία που απομένουν διατάσσονται ομαλά στο χώρο των τάσεων. Τέλος, η μέθοδος τροχιάς φόρτισης (loading path method) συγκρίνει την ιστορία τάσης – παραμόρφωσης κάθε σημείου Gauss και διαγράφει εκείνα τα σημεία με παρόμοιες ιστορίες φόρτισης. Κάθε μια από τις μεθόδους αυτές δείχνει ότι ένα μεγάλο μέρος των δεδομένων εκπαίδευσης (πάνω από το 94%) μπορεί να διαγραφεί χωρίς να επηρεαστεί ουσιαστικά η απόδοση του μοντέλου του NN, ενώ με τη μέθοδο τροχιάς φόρτισης φαίνεται η καλύτερη και η πιο σταθερή απόδοση του NN. Οι χρόνοι εκπαίδευσης του μοντέλου μειώνονται κατά 20 φορές. Η απόδοση της μεθόδου της τροχιάς φόρτισης καταδεικνύεται, επίσης, χρησιμοποιώντας δεδομένα εξαγόμενα από την προσομοιωμένη εργαστηριακή δοκιμή τριαξονικής συμπίεσης με ολίσθηση τριβής.

3.22 Fire analysis of steel frames with the use of artificial neural networks

Συγγραφείς: T. Hozjan, G. Turka, S. Srpčič

Η εργασία παρουσιάζει μια εναλλακτική προσέγγιση για τη μοντελοποίηση της μηχανικής συμπεριφοράς των μεταλλικών σκελετών, όταν εκτίθενται στις υψηλές θερμοκρασίες που αναμένονται σε περιπτώσεις πυρκαγιές. Με βάση μια σειρά καμπυλών τάσεων – παραμορφώσεων, για διάφορα επίπεδα θερμοκρασίας, που προέρχονται από πειράματα, ένα τεχνητό νευρωνικό δίκτυο (ANN) εφαρμόζεται για τη μοντελοποίηση του χάλυβα. Γεωμετρικά και υλικά, μια μη γραμμική ανάλυση για

επίπεδες κατασκευές που υπόκειται σε φωτιά πραγματοποιείται από τη μέθοδο πεπερασμένων στοιχείων (FEM). Τα αριθμητικά αποτελέσματα από μια απλή στηριζόμενη δοκό συγκρίνονται με τα αποτελέσματα της παρούσας εργασίας, τα οποία δείχνουν να ταυτίζονται αρκετά καλά, αν και οι καμπύλες θερμοκρασίας – παραμόρφωσης παρουσιάζουν μάλλον ακανόνιστο σχήμα. Από όλα τα παραπάνω μπορεί να εξαχθεί το συμπέρασμα, ότι το NN είναι ένα αποτελεσματικό εργαλείο για την μοντελοποίηση των ιδιοτήτων του υλικού, στα χαλύβδινα πλαίσια, τα οποία χρησιμοποιούνται στις μελέτες του τεχνολογικού σχεδιασμού, σε περιπτώσεις πυρκαγιών.

3.23 Artificial Neural Networks in numerical modeling of composites (2008)

Συγγραφείς: M. Lefik, D.P. Boso, B.A. Schrefler

Σε αυτό το άρθρο δίνονται μερικές διαφορετικές ιδέες για τη χρήση των τεχνητών νευρωνικών δικτύων (ANN) στη μοντελοποίηση των συνθέτων υλικών και ιεραρχικών δομών. Από ένα σχετικά μικρό σύνολο κατάλληλων αριθμητικών πειραμάτων, που πραγματοποιούνται σε μία δομική μονάδα, δημιουργείται ένα κατάλληλο σύνολο αντίστοιχων δεδομένων εισόδου – εξόδου για την εκπαίδευση του νευρωνικού δικτύου και τον προσδιορισμό των πραγματικών ιδιοτήτων. Επιπλέον οι διαδικασίες που βασίζονται στα NN μπορούν να αξιοποιηθούν σε μια πολυδιάστατη ανάλυση, σαν εργαλείο για την ανάκτηση ζευγών τάσεων – παραμορφώσεων σε χαμηλότερα επίπεδα της ιεραρχικής δομής και / ή για την εκτίμηση του επιπέδου διαρροής της ευκαμψίας των υλικών. Αυτό το είδος της εφαρμογής έχει μεγάλη υπολογιστική σημασία, δεδομένου ότι η μη - γραμμικότητα του υλικού επιτρέπει σε αυτά τη, σημαντικά, βελτιωμένη υπολογιστική αποδοτικότητα τους.

Τέλος, τα NN μπορεί να χρησιμοποιηθούν για τον προσδιορισμό των ομογενοποιημένων ιδιοτήτων μιας κατηγορίας παραμετροποιημένων δομικών μονάδων ή όταν τα χαρακτηριστικά του υλικού εξαρτώνται πάνω από μία παράμετρο (π.χ. θερμοκρασία κ.α.). Το πρόβλημα του καλύτερου ANN για κάθε κατηγορία εφαρμογών συζητείται με τη βοήθεια των παραδειγμάτων που παρουσιάζονται στο κείμενο.

3.24 Self-learning FEM/NMM approach to identification of equivalent material models for plane stress problem (2008)

Συγγραφείς: Pabisek E.

Στο παρών άρθρο γίνεται αξιολόγηση και σύγκριση των αλγορίθμων αυτοπροοδευτικού και του (cumulative algorithm) αθροιστικού αλγορίθμου, που βασίζονται στον «on line» σχηματισμό των δομών και στην εκπαίδευση των νευρωνικών μοντέλων υλικών (NMM), σε προβλήματα μονοαξονικής τάσης. Αποδεικνύεται ότι για προβλήματα συνοριακών τιμών επίπεδης τάσης ο αλγόριθμος αυτοπροοδευτικός, κατά τον οποίο τα NMM εκπαιδεύονται σε κάθε βήμα φόρτισης, είναι ανώτερος από τον cumulative algorithm. Με σκοπό τη δημιουργία ενός μικρού NMM και την επιτάχυνση της σύγκλισης των επαναλήψεων των υπολογιστικών αποτελεσμάτων στα πειραματικά αποτελέσματα, ένας μικρός αριθμός από δεδομένα εισόδου του NMM χρησιμοποιείται και μία τροποποιημένη επιλογή των εκπαιδευμένων δομών εφαρμόζεται. Επίσης παρουσιάζεται μία καινούρια μέθοδος που συνδυάζει την «on line» και «off line» εκπαίδευση του NN. Τα προς συζήτηση προβλήματα επεξηγούνται με δύο περιπτώσεις προς μελέτη. Σχετίζονται με το σχηματισμό NN, για τον προσδιορισμό ισοδύναμων υλικών: (α) για επίπεδα μεταλλικά υποστηρίγματα από υλικό Ramberg – Osgood και (β) για προβλήματα συνοριακών τιμών ελαστοπλαστικής επίπεδης τάσης.

3.25 Comparison of two inverse analysis techniques for learning deep excavation response (2009)

Συγγραφείς: Youssef M.A. Hashash, Siverine Levasseur, Abdolreza Osouli, Richard Finno, Yann Malecot

Η παρατήρηση της απόδοσης μιας κατασκευής είναι ένα απαραίτητο κομμάτι για το σχεδιασμό και την διαδικασία κατασκευής στα γεωτεχνικά προβλήματα. Στις βαθιές αστικές εκσκαφές, οι εμπειρικές και οι αριθμητικές μέθοδοι χρησιμοποιούνται για την πρόβλεψη δυνατών παραμορφώσεων και των επιπτώσεων τους σε αντίστοιχες δομές. Δύο προσεγγίσεις αντίστροφης ανάλυσης συζητούνται και συγκρίνονται σε ένα αναπτυξιακό έργο εκσκαφών στην πόλη του Σικάγο. Η πρώτη προσέγγιση είναι μια μέθοδος παραμετρικής βελτιστοποίησης που βασίζεται σε γενετικό αλγόριθμο (GA). Η μέθοδος GA είναι γενική στοχαστική τεχνική για την βελτιστοποίηση μιας αντικειμενικής συνάρτησης με γραμμικούς ή μη γραμμικούς περιορισμούς. Η δεύτερη προσέγγιση Self – learning simulations (προσομοιώσεις αυτοεκπαίδευσης) (SelfSim) είναι μία μέθοδος αντίστροφης ανάλυσης, που συνδυάζει τη μέθοδο των πεπερασμένων στοιχείων, συνεχώς εξελισσόμενα μοντέλα υλικών και μετρήσεις του εδάφους. Η μέθοδος βελτιστοποίησης που βασίζεται σε γενετικό αλγόριθμο προσδιορίζει τις

ιδιότητες του υλικού από ένα ήδη υπάρχων μοντέλο εδάφους, ενώ η μέθοδος SelfSim περιγράφει τη «κρυμμένη» συμπεριφορά του εδάφους από μία συγκεκριμένη υπόθεση στην καταστατική συμπεριφορά του εδάφους. Οι δύο μέθοδοι αντίστροφης ανάλυσης προσεγγίζουν πολύ καλά τις παράπλευρες αποκλίσεις των τοιχωμάτων και την μέγιστη επιφανειακή καθίζηση. Η βελτιστοποίηση GA τείνει να υπερπροβλέπει την επιφανειακή καθίζηση σε μερική απόσταση από την εκσκαφή, καθώς περιορίζεται από μία συγκεκριμένη μορφή του υλικού καταστατικού μοντέλου (π.χ. προενταμένο μοντέλο εδάφους), ενώ οι επιφανειακές καθιζήσεις που υπολογίζονται από τη μέθοδο SelfSim ταιριάζουν με τις παρατηρούμενες λόγω της ικανότητας του NN να «μαθαίνει» μικρές μη γραμμικές παραμορφώσεις του εδάφους που συνεπάγονται τις μετρούμενες καθιζήσεις.

3.26 Intelligent finite element method and application to simulation of behavior of soils under cyclic loading (2009)

Συγγραφείς: A.A. Javadi, T.P. Tan, and A.S.I. Elkassas

Σε αυτό το άρθρο παρουσιάζεται μία μέθοδος που συνδυάζει νευρωνικά δίκτυα (NN) με τη μέθοδο των πεπερασμένων στοιχείων για τη μοντελοποίηση της συμπεριφοράς του εδάφους κάτω από κυκλική φόρτιση. Σε αυτή τη μέθοδο ένα NN εκπαιδεύεται χρησιμοποιώντας πειραματικά δεδομένα (ή in situ) που αναπαριστούν την μηχανική απόκριση του υλικού υπό την εφαρμογή ενός φορτίου. Το εκπαιδευμένο NN, στη συνέχεια, ενσωματώνεται στην ανάλυση των πεπερασμένων στοιχείων για την πρόβλεψη των καταστατικών σχέσεων του υλικού. Η ανάπτυξη και η επικύρωση της μεθόδου ακολουθείται από την εφαρμογή της, στη μελέτη της συμπεριφοράς του εδάφους κάτω από κυκλική φόρτιση. Στη συνέχεια, τα αποτελέσματα αυτής της «ευφυούς» μεθόδου πεπερασμένων στοιχείων συγκρίνονται με εκείνα που προκύπτουν από την πρότυπη ανάλυση των πεπερασμένων στοιχείων κάνοντας χρήση των συμβατικών καταστατικών μοντέλων. Επίσης τονίζονται τα πλεονεκτήματα και οι ικανότητες των καταστατικών μοντέλων NN και της «ευφυούς» μεθόδου πεπερασμένων στοιχείων. Τέλος δείχνεται ότι τα καταστατικά μοντέλα NN προσφέρουν μια αποτελεσματική και ενοποιημένη προσέγγιση στην καταστατική μοντελοποίηση των υλικών με πολύπλοκη συμπεριφορά στην ανάλυση πεπερασμένων στοιχείων για προβλήματα συνοριακών τιμών.

Γίνεται αντιληπτό από την παρουσίαση όλων των παραπάνω άρθρων, ότι εδώ και μία δεκαετία, τα τεχνητά νευρωνικά δίκτυα (ANN) έχουν χρησιμοποιηθεί ευρύτατα για τη μοντελοποίηση της συμπεριφοράς διαφόρων υλικών, μεμονωμένα, αλλά και σε συνδυασμό με τη μέθοδο των πεπερασμένων στοιχείων, σε αντίστοιχα λογισμικά,

όπως το ABAQUS. Συμπερασματικά λοιπόν, χρειάζεται να αναφερθεί ότι τα NN δεν αποτελούν απλά μια εναλλακτική λύση για τη δημιουργία καταστατικών μοντέλων υλικών, αλλά ένα χρήσιμο εργαλείο, το οποίο καθιερώνεται όλο και περισσότερο όσο οι έρευνες προχωρούν στο συγκεκριμένο πεδίο της μηχανικής.

ΚΕΦΑΛΑΙΟ 4^ο

ΕΦΑΡΜΟΓΕΣ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Στο παρών κεφάλαιο παρουσιάζονται παραδείγματα εφαρμογών του νευρωνικού δικτύου (NN) στην μοντελοποίηση καταστατικών νόμων διαφορών υλικών. Τα υλικά που επιλέχθηκαν είναι **(i)** ο χάλυβας, **(ii)** το σκυρόδεμα και **(iii)** το έδαφος, τα οποία εξετάζονται τόσο σε μονοτονική φόρτιση, όσο και σε ανακυκλιζόμενη. Τα δεδομένα εκπαίδευσης του νευρωνικού δικτύου για το κάθε υλικό προέκυψαν επί το πλείστον από πειραματικές καμπύλες τάσεων – παραμορφώσεων. Μόνο σε δύο περιπτώσεις υλικών, συγκεκριμένα του σκυροδέματος και του εδάφους, που αφορούν τη μονοτονική φόρτιση τα δεδομένα εκπαίδευσης προέκυψαν από αναλυτικούς τύπους. Στη συνέχεια γίνεται παρουσίαση των βασικών χαρακτηριστικών του NN, που χρησιμοποιήθηκε στις συγκεκριμένες εφαρμογές και ανάλυση της κάθε εφαρμογής χωριστά.

4.1 Γενικά χαρακτηριστικά του νευρωνικού δικτύου NN

Το νευρωνικό δίκτυο που χρησιμοποιήθηκε στην υλοποίηση των εφαρμογών, που θα παρουσιαστούν σε επόμενες παραγράφους, είναι πίσω διάδοσης (Backpropagation), ελεγχόμενο από την παραμόρφωση (strain – controlled), με δύο κρυμμένα επίπεδα και σταθερή δομή κατά τη διάρκεια της εκπαίδευσης. Η δομή του στη μονοτονική φόρτιση, διαφέρει από αυτή της ανακυκλιζόμενης και η διαφορά έγκειται τόσο στα δεδομένα εισόδου, όσο και στον αριθμό των κόμβων (νευρώνων) του κάθε κρυμμένου επιπέδου. Περισσότερες λεπτομέρειες για τη δομή του NN δίνονται σε κάθε περίπτωση χωριστά.

Η συνάρτηση μεταφοράς που χρησιμοποιείται για την «ενεργοποίηση» (διέγερση) του νευρωνικού δικτύου, ώστε να παράγει το διάνυσμα εξόδου είναι η σιγμοειδής συνάρτηση (§2.2.3, Κεφ. 2^ο), ενώ η μέθοδος βελτιστοποίησης της συνάρτησης σφάλματος είναι η μέθοδος RPROP, η περιγραφή της οποίας έγινε στην §2.3.2 του Κεφ. 2^ο.

Τέλος η μέθοδος εκπαίδευσης του NN γίνεται με τη μέθοδο της επίβλεψης. Τα δεδομένα εκπαίδευσης του NN στην περίπτωση της μονοτονικής φόρτισης αποτελούνται από ζεύγη τάσεων – παραμορφώσεων ($\sigma_n - \epsilon_n$) στην τρέχουσα κατάσταση. Στην περίπτωση της ανακυκλιζόμενης φόρτισης, κατά την οποία η συμπεριφορά όλων των υλικών, άρα και αυτών που μελετώνται σε αυτή την εργασία, εξαρτάται από την ιστορία της φόρτισης, τα δεδομένα εισόδου του NN εκτός από τα ζεύγη τάσεων – παραμορφώσεων ($\sigma_n - \epsilon_n$) στην τρέχουσα κατάσταση αποτελούνται και από τα ζεύγη τάσεων – παραμορφώσεων του προηγούμενου βήματος φόρτισης ($\sigma_{n-1} - \epsilon_{n-1}$) και από ένα σύνολο εσωτερικών μεταβλητών κατάστασης που σχετίζονται με την υστερητική συμπεριφορά των υλικών. Οι τιμές της τρέχουσας τάσης τόσο για την μονοτονική, όσο και για την ανακυκλιζόμενη φόρτιση αποτελούν το διάνυσμα στόχου του NN, όπου οι προβλεπόμενες τιμές της τάσης που θα προκύψουν από αυτό,

θα πρέπει να διαφέρουν ελάχιστα, σύμφωνα πάντα με την παράμετρο ελέγχου της σύγκλισης που έχει επιλεγεί. Η επιλογή των ζευγών τάσεων παραμορφώσεων για την διαδικασία της εκπαίδευσής του, έγινε τυχαία μέσα από το εύρος τιμών που προκύπτει από τις πειραματικές καμπύλες ή από τους αναλυτικούς τύπους.

4.2 Μονοτονική φόρτιση

Η δομή του νευρωνικού δικτύου (NN) που χρησιμοποιήθηκε για την μοντελοποίηση των καταστατικών νόμων των τριών υλικών (χάλυβας, σκυρόδεμα, έδαφος) στη μονοτονική φόρτιση είναι η ακόλουθη:

$$\sigma = \sigma_{NN}(\varepsilon/10/10/1) \quad (4.1)$$

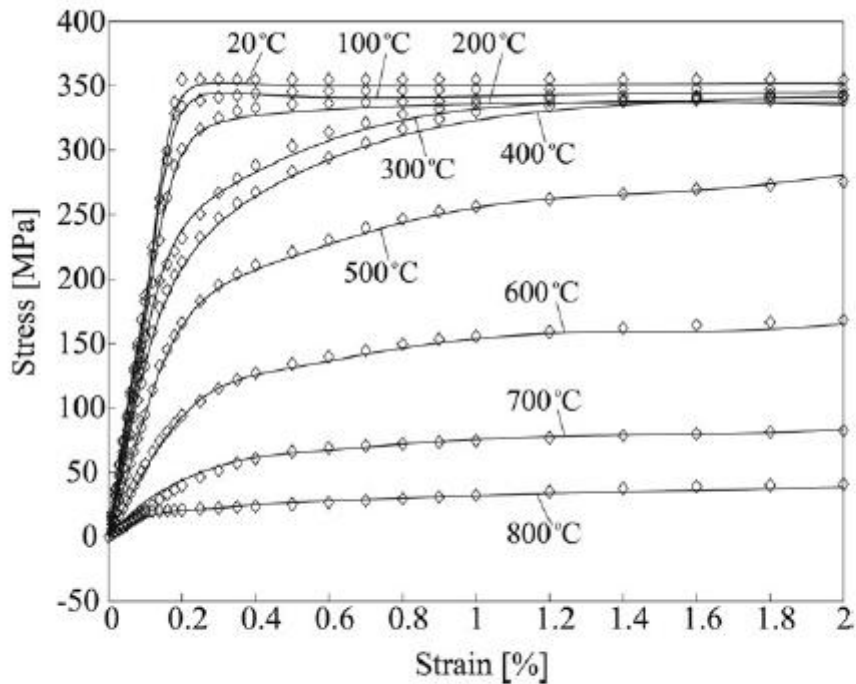
Επίσης ο αριθμός των επαναλήψεων (Epochs) που επιλέγονται για την εκπαίδευση του νευρωνικού δικτύου σε όλες της εφαρμογές αυτού του είδους φόρτισης είναι 5000.

4.2.1 Μονοαξονικός εφελκυσμός δομικού χάλυβα θερμής έλασης σε μεταβατικές συνθήκες θέρμανσης

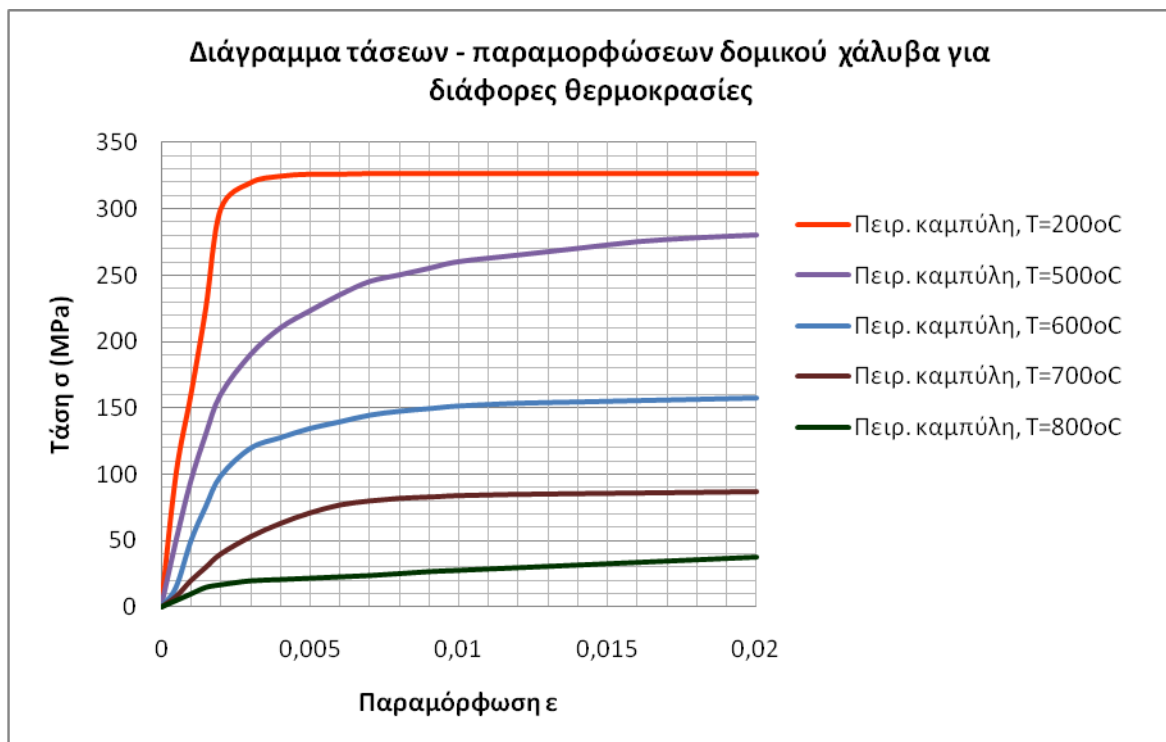
Οι Kirby και Preston [52] στην προσπάθειά τους να μελετήσουν τη συμπεριφορά του δομικού χάλυβα, υπό την επίδραση υψηλών θερμοκρασιών, σαν και αυτών που αναμένονται σε περιπτώσεις πυρκαγιών, πραγματοποίησαν δοκιμές εφελκυσμού σε δομικούς χάλυβες θερμής έλασης σε μεταβατικές συνθήκες θέρμανσης, με ρυθμό θέρμανσης 10 °C/min. Τα αποτελέσματα αυτών των δοκιμών για θερμοκρασίες από 20 °C έως 800 °C και για τιμές παραμόρφωσης έως 2%, παρουσιάζονται στο σχήμα 4.1. Όπως παρατηρείται από το διάγραμμα, η μορφή των καμπυλών, που αντιστοιχούν σε θερμοκρασίες από 400 °C – 800 °C, δεν παραπέμπουν στο γνωστό διγραμμικό μοντέλο του χάλυβα για μονοαξονικές καταπονήσεις (εφελκυσμός, θλίψη), όπως γίνεται με τις καμπύλες για θερμοκρασίες από 20 °C – 300 °C. Αυτό συμβαίνει, διότι ύστερα από έρευνες, διαπιστώθηκε ότι η συμπεριφορά του χάλυβα, πάνω από τους 400 °C εμφανίζει εξάρτηση, τόσο από το ρυθμό θέρμανσης, όσο και από τη χρονική διάρκεια των υψηλών θερμοκρασιών, με αποτέλεσμα να εμφανίζονται παραμορφώσεις ερπυσμού, οι οποίες γίνονται εντονότερες με την αύξηση της θερμοκρασίας και υπερισχύουν των ελαστικών και πλαστικών παραμορφώσεων.

Κάποιες, λοιπόν, από τις καμπύλες του παραπάνω σχήματος χρησιμοποιήθηκαν για την προσπάθεια μοντελοποίησης τους μέσω νευρωνικών δικτύων (NN). Οι καμπύλες που επιλέχθηκαν αντιστοιχούν σε θερμοκρασίες 200 °C, 500 °C, 600 °C, 700 °C και 800 °C, και παρουσιάζονται στο διάγραμμα 4.1. Από αυτές τις καμπύλες προκύπτουν τα δεδομένα εκπαίδευσης, ζεύγη τιμών τάσης - παραμόρφωσης. Όπως αναφέρθηκε

και προηγουμένως η επιλογή έγινε τυχαία για διάφορες τιμές παραμόρφωσης, οπότε και αντίστοιχες τιμές τάσης, μεταξύ των τιμών 0 % έως και 2%.



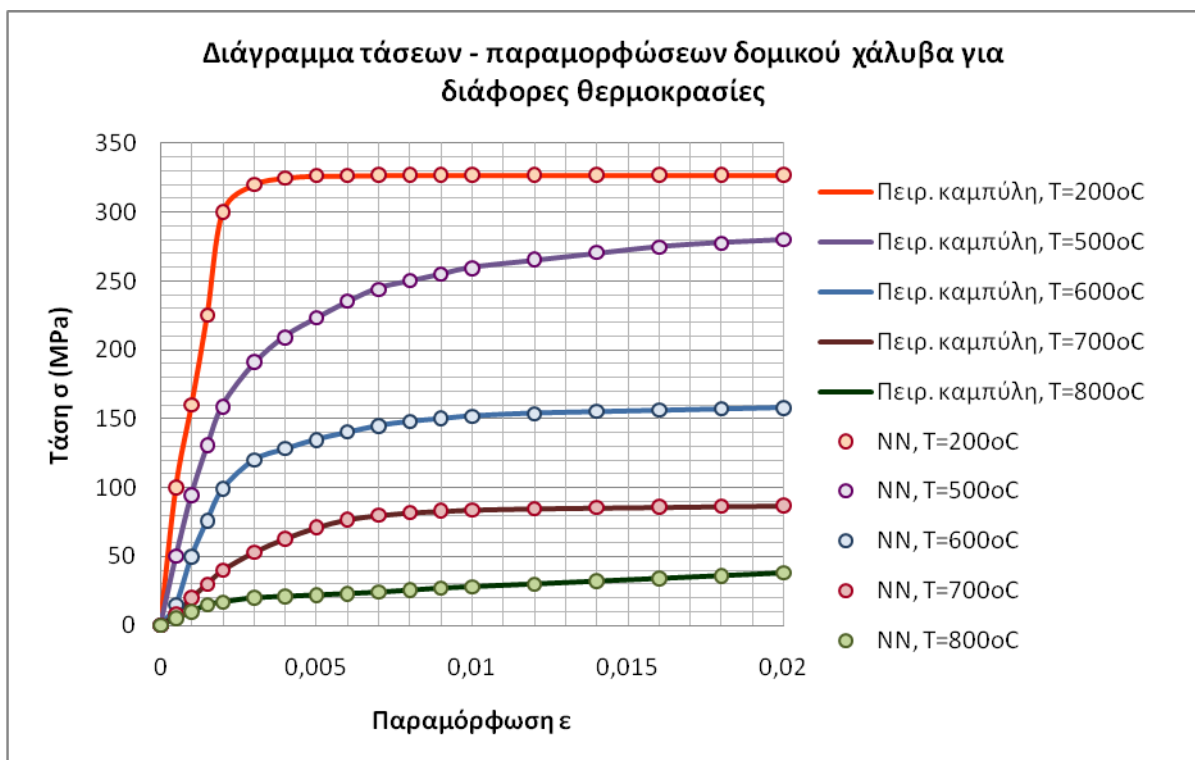
Σχήμα 4.1: Καμπύλες τάσης – παραμόρφωσης δομικού χάλυβα για διάφορες θερμοκρασίες



Διάγραμμα 4.1: Καμπύλες τάσης – παραμόρφωσης δομικού χάλυβα για θερμοκρασίες 200 °C, 500, °C 600 °C, 700 °C και 800 °C

Στο διάγραμμα 4.2 παρουσιάζονται τα αποτελέσματα που προέκυψαν από την εκπαίδευση του NN. Όπως φαίνεται από το διάγραμμα, αυτό, οι προβλεπόμενες, από το NN, τιμές της τάσης συμπίπτουν πολύ καλά με τις πειραματικές τιμές της τάσης (ή τιμές «στόχοι» για το NN). Αυτό οφείλεται στην προσεχτική επιλογή του συνόλου των δεδομένων εκπαίδευσης και στην σχετικά απλή περίπτωση, καταστατικού νόμου του δομικού χάλυβα που επιλέχθηκε για μοντελοποίηση μέσω του νευρωνικού δικτύου.

Χρειάζεται επίσης να σημειωθεί ότι ο αριθμός των επαναλήψεων (Epochs) για την διαδικασία εκπαίδευσης, αλλά και το πλήθος των κόμβων ή νευρώνων των κρυμμένων επιπέδων του NN, φαίνεται να μην έχουν επίδραση στα αποτελέσματα του. Από όλα τα παραπάνω, μπορεί να προκύψει ο ισχυρισμός, ότι το νευρωνικό δίκτυο που επιλέχθηκε (δομή:1/10/10/1) μπορεί να αποτελέσει το μοντέλο του καταστατικού νόμου που περιγράφει τη συμπεριφορά του χάλυβα σε δοκιμή εφελκυσμού και για διάφορες θερμοκρασίες.



Διάγραμμα 4.2: Πρόβλεψη NN για τις καμπύλες τάσης – παραμόρφωσης του δομικού χάλυβα για θερμοκρασίες 200 °C, 500, °C 600 °C, 700 °C και 800 °C

4.2.2 Αξονική συμπίεση οπλισμένου σκυροδέματος

Ο καταστατικός νόμος που περιγράφει τη αξονική συμπίεση σκυροδέματος προέρχεται από τον Sargin (Mougin, 1992) [53] και ο αναλυτικός του τύπος είναι ο εξής:

$$\sigma = \sigma_{ref} \cdot \left(\frac{k_1 \cdot \left(\frac{\varepsilon}{\varepsilon_{ref}} \right) + (k_2 - 1) \cdot \left(\frac{\varepsilon}{\varepsilon_{ref}} \right)^2}{1 + (k_1 - 2) \cdot \left(\frac{\varepsilon}{\varepsilon_{ref}} \right) + k_2 \cdot \left(\frac{\varepsilon}{\varepsilon_{ref}} \right)^2} \right) \quad (4.2)$$

όπου ισχύουν τα παρακάτω:

$$30 < \sigma_{ref} < 60 \text{ MPa}$$

$$\varepsilon_{ref} = 0.62 \cdot 10^{-3} \sqrt[3]{\sigma_{ref}}$$

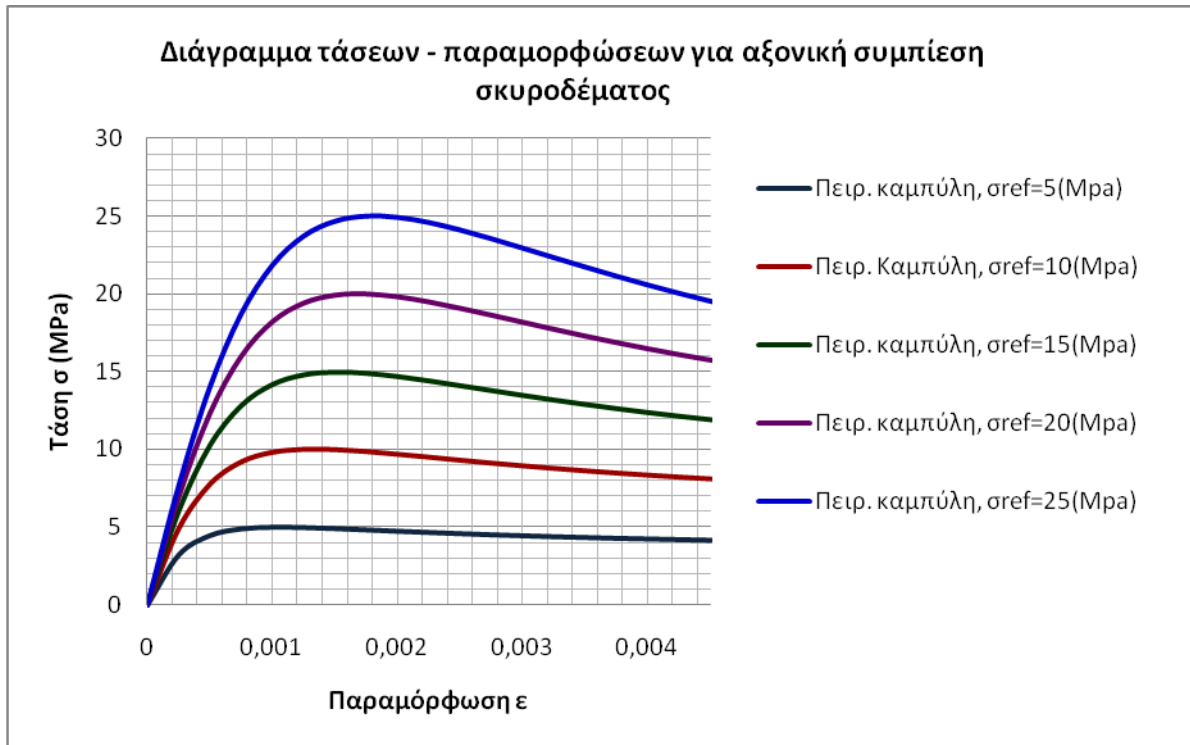
$$k_1 = \frac{E \cdot \varepsilon_{ref}}{\sigma_{ref}} \quad \mu\epsilon \quad E = 11000 \sqrt[3]{\sigma_{ref}}$$

$$k_2 = \begin{cases} k_1 - 1 & , \text{if } \sigma_{ref} \leq 30 \text{ MPa} \\ 0 & , \text{if } \sigma_{ref} \geq 55 \text{ MPa} \\ (k_1 - 1) \cdot \left(\frac{55 - \sigma_{ref}}{25} \right) & , \text{if } 30 < \sigma_{ref} < 55 \text{ MPa} \end{cases} \quad (4.3)$$

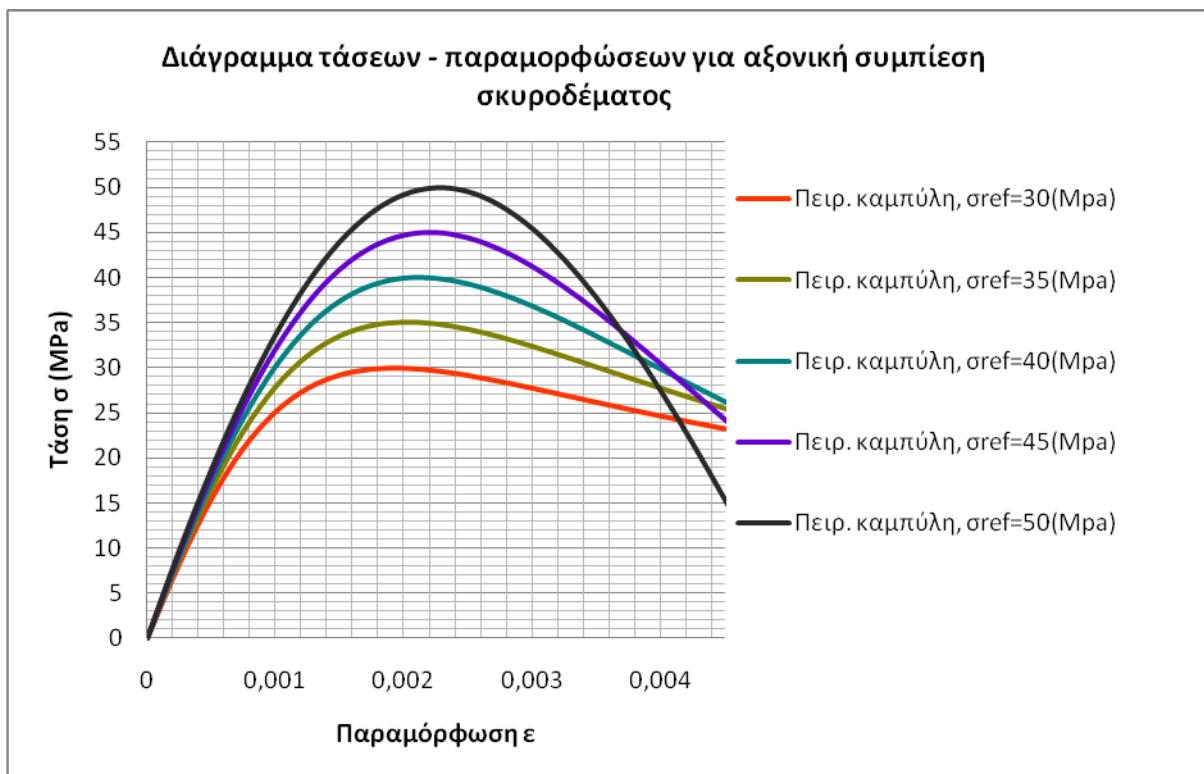
- ✓ σ : είναι η αξονική τάση,
- ✓ σ_{ref} : είναι η χαρακτηριστική τάση του σκυροδέματος,
- ✓ ε : είναι η αξονική παραμόρφωση,
- ✓ ε_{ref} : είναι η κρίσιμη παραμόρφωση που αντιστοιχεί την χαρακτηριστική τάση σ_{ref}

Σύμφωνα, λοιπόν με τον παραπάνω τύπο για εύρος τιμών της παραμόρφωσης ($0 - 5 \times 10^{-3}$) και για τιμές της χαρακτηριστικής τάσης σ_{ref} που ανήκουν στο πεδίο τιμών [5:5:50] προκύπτουν οι καμπύλες τάσης – παραμόρφωσης που φαίνονται στα διαγράμματα 4.3 και 4.4.

Από τα παρακάτω διαγράμματα προκύπτουν και τα δεδομένα εκπαίδευσης του νευρωνικού δικτύου NN. Η επιλογή, όπως και προηγουμένως, έγινε τυχαία μέσα από το εύρος τιμών της παραμόρφωσης και της τάσης, αντίστοιχα, που δόθηκε παραπάνω.



Διάγραμμα 4.3: Διάγραμμα τάσεων – παραμορφώσεων σκυροδέματος για $\sigma_{ref} \in [5:5:25]$

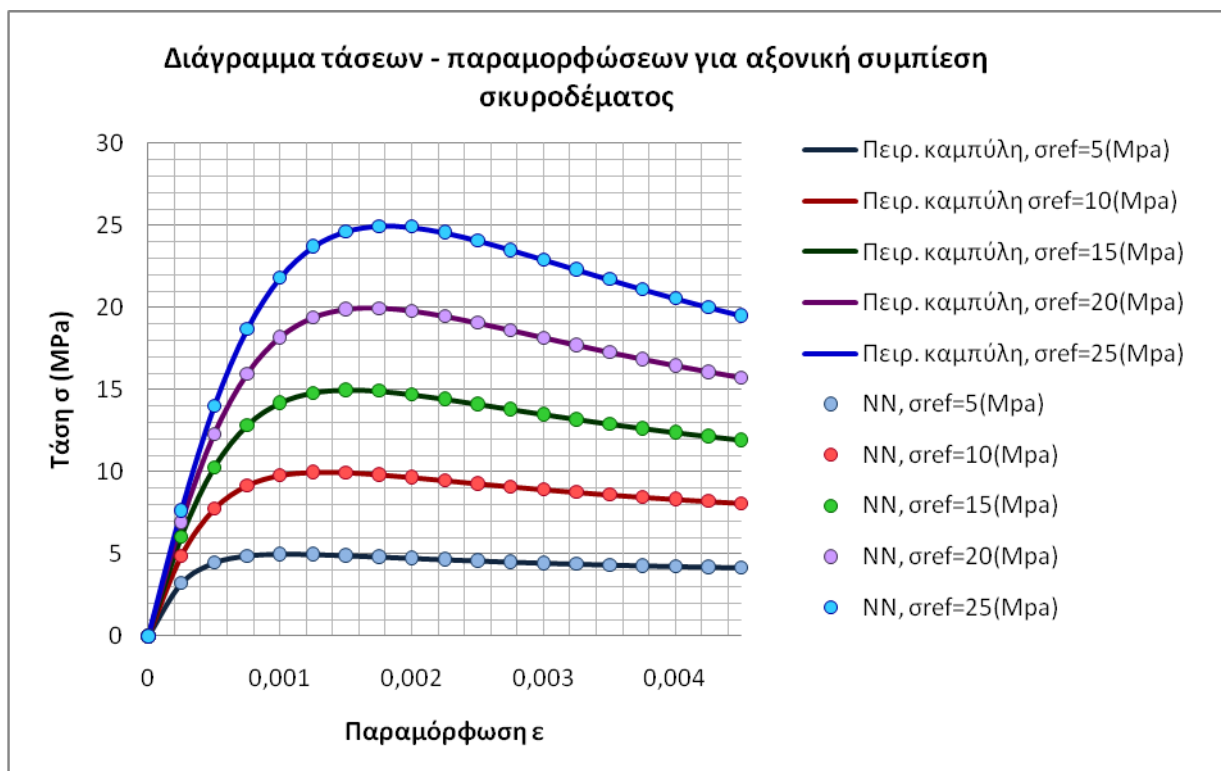


Διάγραμμα 4.4: Διάγραμμα τάσεων – παραμορφώσεων σκυροδέματος για $\sigma_{ref} \in [30:5:50]$

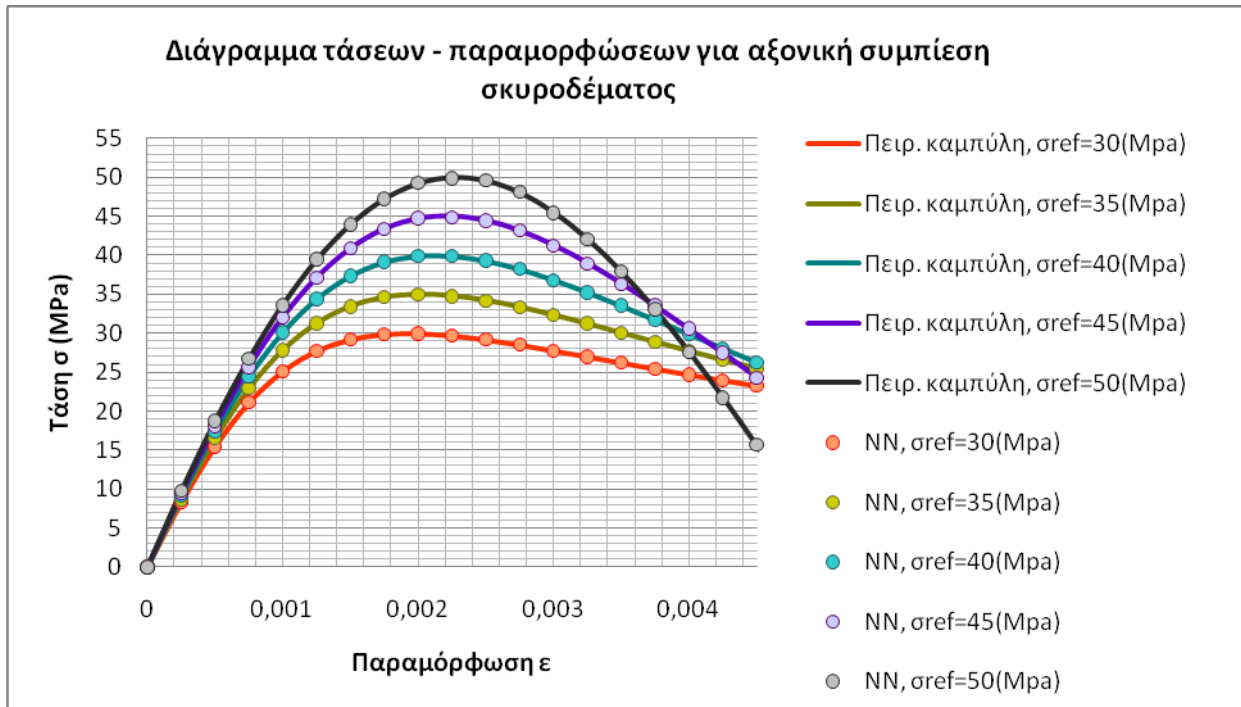
Στα διαγράμματα 4.5 και 4.6 που ακολουθούν εμφανίζονται τα αποτελέσματα από την εκπαίδευση – πρόβλεψη του νευρωνικού δικτύου NN. Οι προβλεπόμενες, από το NN, τιμές των τάσεων, που αντιστοιχούν στις τιμές των παραμορφώσεων που

χρησιμοποιήθηκαν ως διάνυσμα εισόδου του NN, συμπίπτουν πολύ καλά με τις πραγματικές τιμές των τάσεων, που αποτελούν το διάνυσμα στόχο του NN. Η πολύ καλή συμπεριφορά του νευρωνικού δικτύου και η ελάχιστη απόκλιση πειραματικών και προβλεπόμενων τιμών της τάσης είναι αποτέλεσμα της επιλογής του συνόλου των δεδομένων εκπαίδευσης, στο οποίο περιέχεται πληροφορία από ολόκληρο το εύρος τιμών των παραμορφώσεων και κατ' επέκταση και των τάσεων. Επίσης η «καλή» πρόβλεψη των τιμών της τάσης (διάνυσμα στόχος) από το NN οφείλεται και στο είδος της φόρτισης (μονοτονική, μονοαξονική), η οποία δεν έχει ιδιαίτερες απαιτήσεις όσον αφορά τη μοντελοποίηση της.

Πρέπει επίσης να σημειωθεί, όπως και πριν, ότι ο αριθμός των επαναλήψεων (Epochs) για την διαδικασία εκπαίδευσης, αλλά και το πλήθος των κόμβων ή νευρώνων των κρυμμένων επιπέδων του NN, φαίνεται να μην έχουν επίδραση στα αποτελέσματα του. Σύμφωνα με όλα τα προηγούμενα ο καταστατικός νόμος του σκυροδέματος για μονοαξονική συμπίεση είναι ένας ακόμα νόμος που μπορεί να μοντελοποιηθεί με τη χρήση των νευρωνικών δικτύων.



Διάγραμμα 4.5: Πρόβλεψη NN για τις καμπύλες τάσης – παραμόρφωσης αξονικής συμπίεσης σκυροδέματος για $\sigma_{ref} \in [5:5:25]$



Διάγραμμα 4.6: Πρόβλεψη NN για τις καμπύλες τάσης – παραμόρφωσης αξονικής συμπίεσης σκυροδέματος για $\sigma_{ref} \in [30:5:50]$

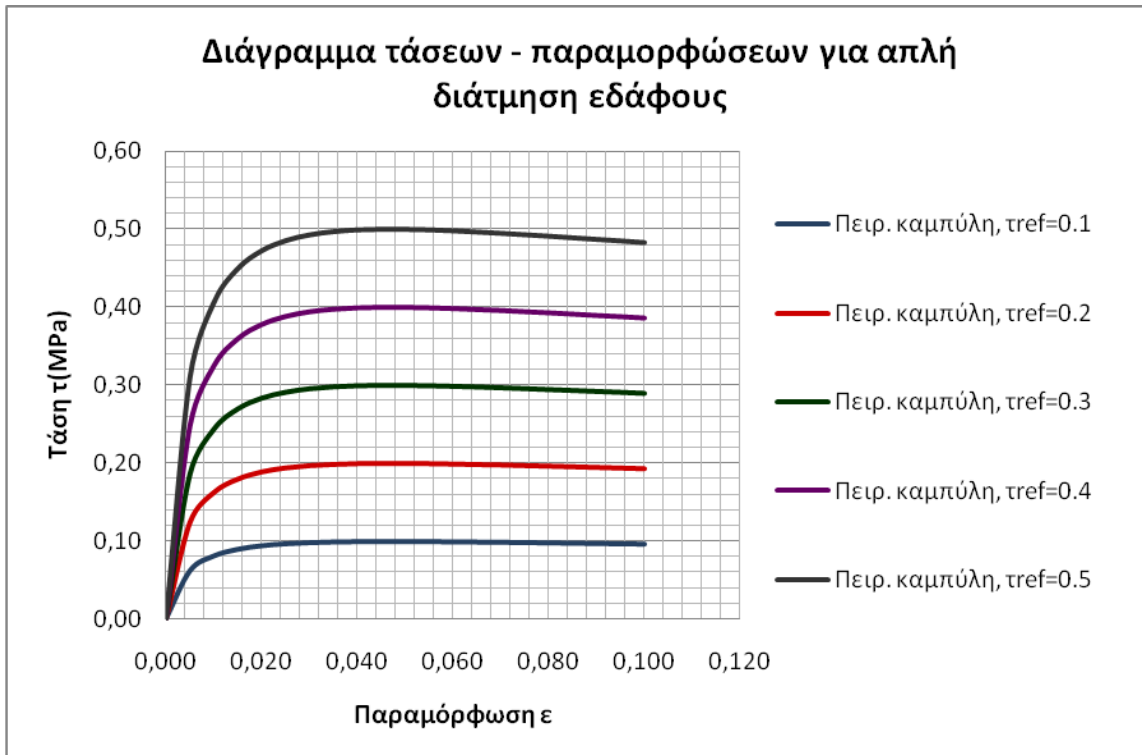
4.2.3 Καταστατικός νόμος διάτμησης εδάφους

Ο καταστατικός νόμος που περιγράφει την απλής διάτμηση του «χαλαρού» εδάφους (υψηλός δείκτης κενών) δίνεται από την εξής σχέση [54]:

$$\tau = \tau_{ref} \cdot \left(\frac{\frac{\gamma}{\gamma_{ref}}}{\left(1 - \frac{\gamma}{\gamma_{ref}}\right)^\delta} \right)^{\frac{k + \frac{\gamma}{\gamma_{ref}}}{\left(1 - \frac{\gamma}{\gamma_{ref}}\right)^\delta}}, \text{ με } k = \frac{G \cdot \gamma_{ref}}{\tau_{ref}} \quad (4.4)$$

όπου τ_{ref} είναι η μέγιστη διατμητική αντοχή του εδάφους, γ_{ref} είναι η κρίσιμη διατμητική παραμόρφωση που αντιστοιχεί στην διατμητική τάση, k είναι ο λόγος του μέτρου διάτμησης και ενός μέτρου αναφοράς και δ είναι η παράμετρος πόρων του εδάφους.

Από την εφαρμογή του παραπάνω καταστατικού νόμου για πεδίο τιμών της παραμόρφωσης γ (0 – 10%), για εύρος τιμών της χαρακτηριστικής διατμητικής αντοχής τ_{ref} (0.1 – 0.5), για $k = 0.1$ και για $\delta = 0.1$, προκύπτει το επόμενο διάγραμμα με τάσεων – παραμορφώσεων.

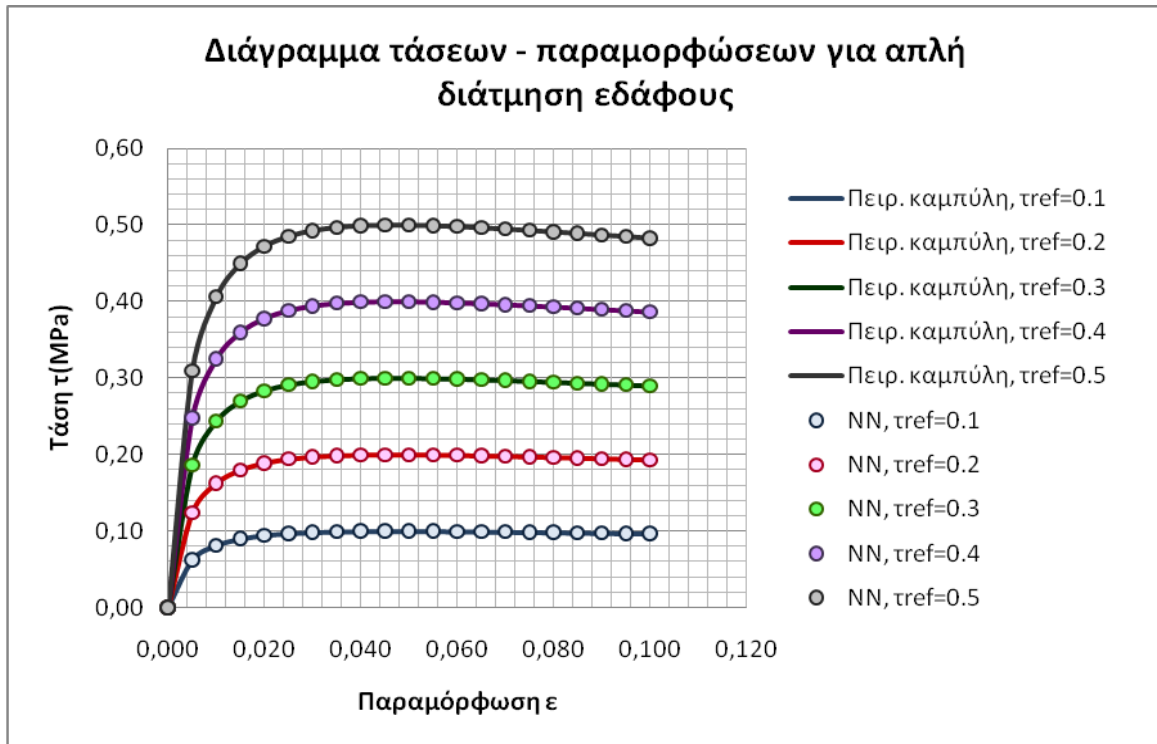


Διάγραμμα 4.7: διάγραμμα τάσεων – παραμορφώσεων για απλή διάτμηση εδάφους με $\tau_{ref} \in [0.1:0.1:0.5]$, $k = 0.1$ και $\delta = 0.1$

Από τις καμπύλες του διαγράμματος 4.7 δημιουργήθηκαν τα δεδομένα εκπαίδευσης του νευρωνικού δικτύου, τα οποία αποτελούνται από ζεύγη τάσεων - παραμορφώσεων. Οι τιμές της παραμόρφωσης αποτελούν το διάνυσμα εισόδου του NN και οι τιμές της τάσης σχηματίζουν το διάνυσμα στόχου, το οποίο πρέπει να προσεγγίσει όσο το δυνατόν καλύτερα το NN. Όπως και στις δύο προηγούμενες περιπτώσεις, επιλέχθηκαν τυχαία από το εύρος τιμών, τόσο της παραμόρφωσης γ , όσο και της διατμητικής τάσης τ .

Τα αποτελέσματα από την εκπαίδευση και την πρόβλεψη του νευρωνικού δικτύου φαίνονται στο διάγραμμα 4.8 που έπεται. Από την παρατήρηση του διαγράμματος φαίνεται ότι τα αποτελέσματα του νευρωνικού δικτύου προσεγγίζουν πολύ καλά τα πειραματικά δεδομένα, γεγονός που επιτρέπει την χρήση νευρωνικών δικτύων τόσο για τη μοντελοποίηση του συγκεκριμένου καταστατικού νόμου, όσο και άλλων νόμων γενικότερα. Τέλος ο αριθμός των επαναλήψεων (Epochs) για την διαδικασία εκπαίδευσης του νευρωνικού δικτύου, αλλά και το πλήθος των κόμβων ή νευρώνων

των κρυμμένων επιπέδων του, φαίνεται να μην έχουν επίδραση στα αποτελέσματα που προέκυψαν από τη διαδικασία εκπαίδευσης.



Διάγραμμα 4.8: Πρόβλεψη NN για τις καμπύλες ($\sigma - \epsilon$) απλής διάτμησης του εδάφους για $k = 0.1$, $\delta = 0.1$ και για $\tau_{ref} \in [0.1:0.1:0.5]$

4.3 Ανακυκλιζόμενη φόρτιση

4.3.1 Μοντελοποίηση της υστερητικής συμπεριφοράς των υλικών

Η κυκλική συμπεριφορά των υλικών είναι πολύ περίπλοκη και δύσκολη στο να μοντελοποιηθεί. Ένας συνδυασμός από κανόνες σκλήρυνσης της κλασικής θεωρίας πλαστικότητας, είναι μία πιθανή λύση για τη μοντελοποίηση αυτής της πολύπλοκης συμπεριφοράς. Τα καταστατικά μοντέλα νευρωνικών δικτύων, λόγω των ιδιαίτερων χαρακτηριστικών τους, έχουν την ικανότητα να μοντελοποιήσουν πολύπλοκες υλικές συμπεριφορές, άμεσα, από τα αποτελέσματα των δοκιμών σε υλικά. Γενικά υπάρχουν δύο τρόποι προσέγγισης της μοντελοποίησης της κυκλικής συμπεριφοράς των υλικών με εφαρμογή των νευρωνικών δικτύων (NN), οι οποίες είναι:

- ✓ Η τρέχουσα κατάσταση των υλικών εξαρτάται από την τρέχουσα και την (τις) προηγούμενη (-ες) κατάσταση (-ες) μόνο κάποιων παρατηρήσιμων μεταβλητών, όπως ολικές τάσεις – παραμορφώσεις, επαυξητικά βήματα τάσεων – παραμορφώσεων κ.α.)

- ✓ Η τρέχουσα κατάσταση των υλικών εξαρτάται από την τρέχουσα και/ή προηγούμενη (-ες) κατάσταση (-ες) κάποιων παρατηρήσιμων μεταβλητών και από ένα σύνολο εσωτερικών μεταβλητών που σχετίζονται με την κινηματική και την ιστροπική σκλήρυνση.

Πρόσφατα προτάθηκε μία καινούρια προσέγγιση (Yun, Ghaboussi, Elnashai, 2007) [55] για μοντελοποίηση της κυκλικής συμπεριφοράς των υλικών, μέσω νευρωνικών δικτύων (NN), η οποία βασίζεται στη δεύτερη, από τις δύο προσεγγίσεις που αναφέρθηκαν προηγουμένως. Σύμφωνα με την καινούρια μέθοδο το νευρωνικό δίκτυο χρησιμοποιεί καινούριες εσωτερικές μεταβλητές που διευκολύνουν την εκμάθηση (στάδιο εκπαίδευσης του NN) της υστερητικής συμπεριφοράς των υλικών και μπορεί να χρησιμοποιηθεί για διάφορα υλικά σε ανακυκλιζόμενη φόρτιση.

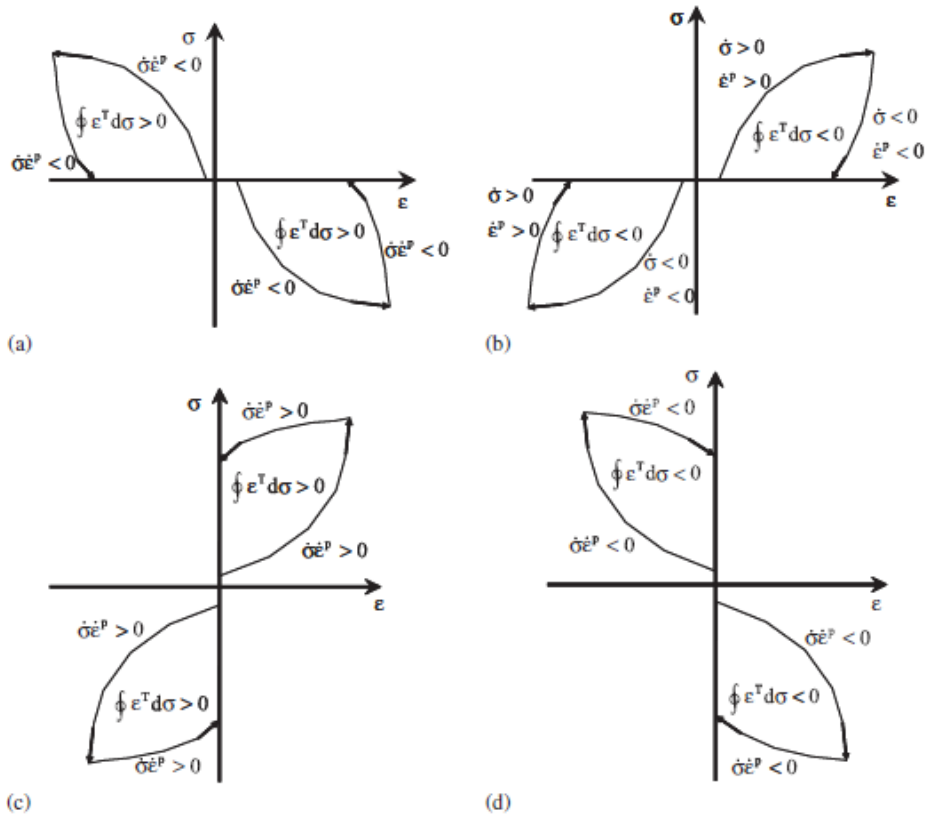
4.3.1.1. Περιγραφή της μεθόδου μοντελοποίησης της υστερητικής συμπεριφοράς των υλικών

Υπάρχουν τέσσερις τύποι υστέρησης για τη μονοδιάστατη σχέση τάσεων – παραμορφώσεων οι οποίες φαίνονται στο σχήμα (4.2).

Από τις τέσσερις παραπάνω περιπτώσεις μόνο η περίπτωση II είναι αποδεκτή για την εκπροσώπηση της υστερητικής συμπεριφοράς των υλικών σε μονοαξονικές φορτίσεις, διότι είναι η μόνη που δεν παραβιάζει τους νόμους διατήρησης της μηχανικής και τις συνθήκες ισορροπίας. Αυτό αποδεικνύεται από τον παρακάτω ορισμό:

Δίνεται η συνάρτηση $\hat{\sigma}: A \rightarrow B$ και $\sigma \in B$, Κατά μήκος μιας κλειστής καμπύλης της τάσης η συνάρτηση $\hat{\sigma}$ με μια οποιαδήποτε τιμή της τάσης σ γράφεται από τη σχέση $I_{\hat{\sigma}}(\sigma) = \{\varepsilon \in A : \hat{\sigma}(\varepsilon) = \sigma\}$. Εάν ισχύει ότι $n[I_{\hat{\sigma}}(\sigma)] \geq 2$, τότε η συνάρτηση $\hat{\sigma}$ περιγράφει ένα βρόχο υστέρησης. Η κλειστή καμπύλη της τάσης ικανοποιεί την ισορροπία και τους ενεργειακούς περιορισμούς, $d\sigma^T d\varepsilon^e \geq 0$ και $\oint \varepsilon^T d\sigma \leq 0$. Ο αντίθετος ορισμός ισχύει αν αντικαταστήσουμε την τάση σ με την παραμόρφωση ε .

Από με τον παραπάνω ορισμό παρατηρείται ότι σε μία τιμή της παραμόρφωσης αντιστοιχούν δύο τιμές της τάσης. Αυτό αποτελεί τη σημαντικότερη αιτία που εμποδίζει τη χρήση των NN από την εκμάθηση της υστερητικής συμπεριφοράς των υλικών.



Σχήμα 4.2: Περιπτώσεις της υστερητικής συμπεριφοράς των υλικών σε μονοαξονική κατάσταση: (a) περίπτωση I, (b) περίπτωση II, (c) περίπτωση III και (d) περίπτωση IV

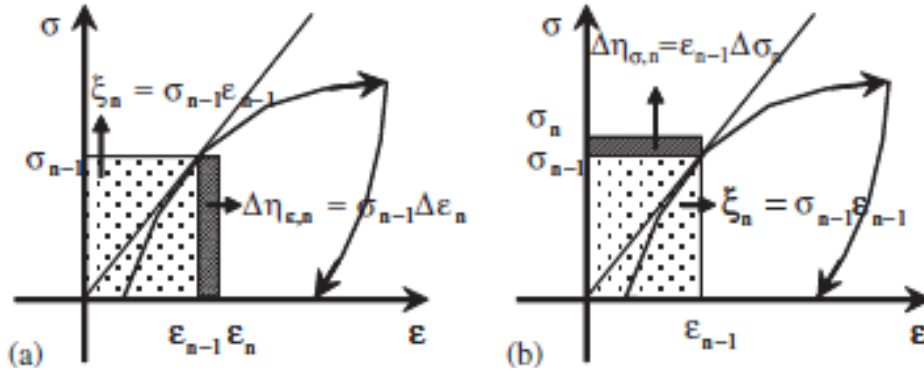
Για να επιλυθεί αυτό το πρόβλημα και να υπάρχει αντιστοίχιση μεταξύ των τιμών ϵ (διάνυσμα εισόδου του NN) και των τιμών σ (διάνυσμα εξόδου του NN) μία προς μία, εισάγονται δύο νέες εσωτερικές μεταβλητές. Οι μεταβλητές αυτές είναι οι ακόλουθες και παρουσιάζονται στο σχήμα (4.3).

$$\xi_n = \sigma_{n-1} \epsilon_{n-1} \quad (4.5)$$

$$\Delta \eta_{\epsilon,n} = \sigma_{n-1} \Delta \epsilon_n \quad (4.6)$$

$$\Delta \eta_{\sigma,n} = \epsilon_{n-1} \Delta \sigma_n \quad (4.7)$$

Ο δείκτης n στις σχέσεις (4.5) – (4.7) δείχνει το v – οστό επαυξητικό βήμα. Η πρώτη εσωτερική μεταβλητή (εξ. 4.5) υποδηλώνει την προηγούμενη κατάσταση κατά μήκος της καμπύλης τάσεων – παραμορφώσεων. Η δεύτερη και η τρίτη εσωτερική μεταβλητή υποδηλώνουν την κατεύθυνση για το επόμενο χρονικό βήμα ή βήμα φόρτισης, κατά μήκος της καμπύλης $\sigma - \epsilon$ και χρησιμοποιούνται για την ελεγχόμενη από την παραμόρφωση (strain control) και την ελεγχόμενη από την τάση (stress control) κατάσταση, αντίστοιχα.



Σχήμα 4.3: Εσωτερικές μεταβλητές για τη μοντελοποίηση της κυκλικής συμπεριφοράς των υλικών με NN : (a) ελεγχόμενη, από την παραμόρφωση, μορφή και (b) ελεγχόμενη, από την τάση, μορφή

Η εισαγωγή των παραπάνω εσωτερικών μεταβλητών στα δεδομένα εισόδου του νευρωνικού δικτύου διασφαλίζει την μία προς μία απεικόνιση της εισόδου ϵ ή σ με την έξοδο σ ή ϵ . Έτσι ισχύουν οι παρακάτω σχέσεις:

$$\sigma_n = \{\sigma(\xi_n, \Delta \eta_{\epsilon,n}, \epsilon_n) : A_1 B_1 \leq \xi_n \leq A_2 B_2, A_1 C_1 \leq \Delta \eta_{\epsilon,n} \leq A_2 C_2, B_1 \leq \epsilon_n \leq B_2 \quad (4.8)$$

$$\epsilon_n = \{\epsilon(\xi_n, \Delta \eta_{\sigma,n}, \sigma_n) : A_1 B_1 \leq \xi_n \leq A_2 B_2, B_1 D_1 \leq \Delta \eta_{\sigma,n} \leq B_2 D_2, A_1 \leq \sigma_n \leq A_2 \quad (4.9)$$

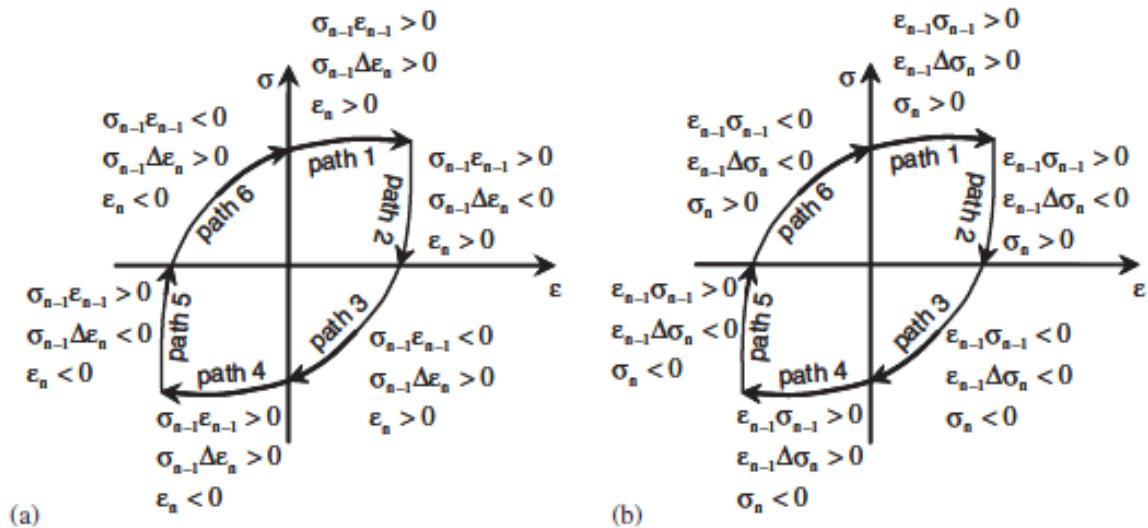
όπου $A_1 \leq \sigma_n \leq A_2$, $B_1 \leq \epsilon_n \leq B_2$, $C_1 \leq \Delta \epsilon_n \leq C_2$, $D_1 \leq \Delta \sigma_n \leq D_2$

Για να γίνει ακόμα περισσότερο κατανοητό ότι οι σχέσεις (4.8) και (4.9) υποδηλώνουν απεικόνιση μία προς μία, ο βρόχος υστέρησης χωρίζεται σε έξι κομμάτια και σε κάθε ένα από αυτά, παρουσιάζεται το πρόσημο κάθε μίας από τις τρεις εσωτερικές μεταβλητές εισόδου του NN. Το σχήμα (4.4) παρουσιάζει τα έξι κομμάτια της καμπύλης υστέρησης και τα αντίστοιχα πρόσημα των εσωτερικών μεταβλητών και για τα δύο είδη ελεγχόμενης καμπύλης (από την παραμόρφωση ή από την τάση).

Σύμφωνα με τη διαδικασία εκμάθησης της υστερητικής συμπεριφοράς κάποιου υλικού που περιγράφηκε παραπάνω, οι δύο εξισώσεις που περιγράφουν την ελεγχόμενη, από την παραμόρφωση, μορφή και την ελεγχόμενη από την τάση, μορφή είναι:

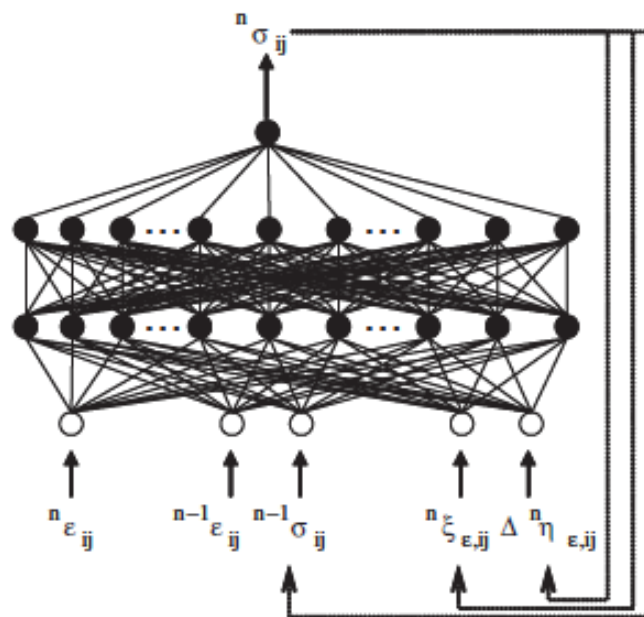
$$\sigma_n = \hat{\sigma}_{NN}(\{\epsilon_n, \epsilon_{n-1}, \sigma_{n-1}, \xi_{\epsilon,n}, \Delta \eta_{\epsilon,n}\} : \{\text{δομή του NN}\}) \quad (4.10\alpha)$$

$$\epsilon_n = \hat{\epsilon}_{NN}(\{\sigma_n, \sigma_{n-1}, \epsilon_{n-1}, \xi_{\sigma,n}, \Delta \eta_{\sigma,n}\} : \{\text{δομή του NN}\}) \quad (4.10\beta)$$



Σχήμα 4.4: Ταξινόμηση των επιμέρους μερών της καμπύλης υστέρησης και οι συνδυασμοί προσήμων των εσωτερικών μεταβλητών για (a) ελεγχόμενη από την ϵ , μορφή και (b) ελεγχόμενη από την σ , μορφή

Το επόμενο σχήμα αντιστοιχεί στη εξίσωση (4.10α), όπου το υλικό μοντέλο του NN ελεγχόμενο από την παραμόρφωση, βρίσκεται στο n – οστό βήμα της εκπαίδευσης του.

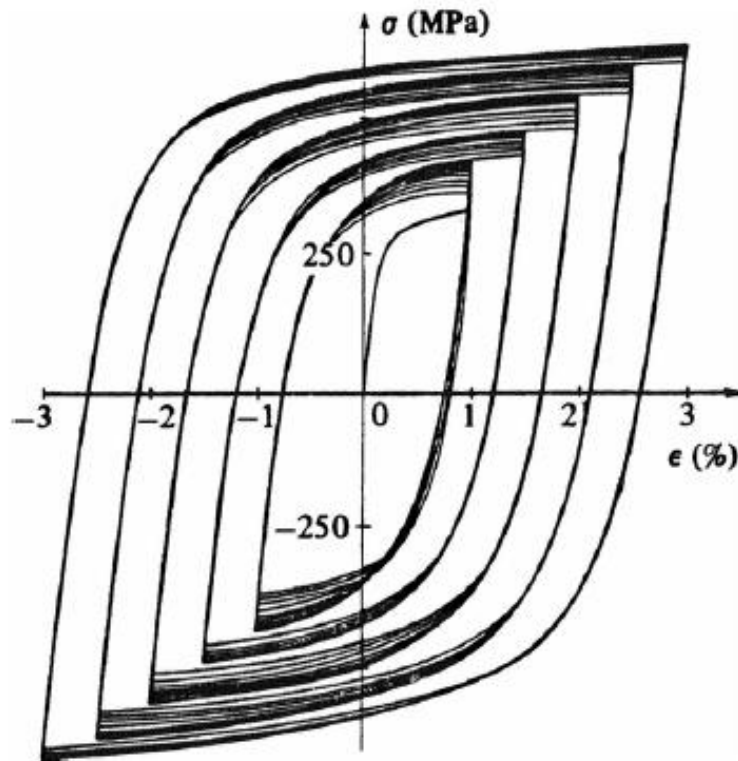


Σχήμα 4.5: Υλικό μοντέλο NN στο n – οστό βήμα εκπαίδευσης

Στα επόμενα παραδείγματα που ακολουθούν επιλέγεται η ελεγχόμενη, από την παραμόρφωση, μορφή του NN.

4.3.2 Μονοαξονική συμπίεση ανοξειδωτου χάλυβα 316

Ο Chaboche (1979) στην προσπάθειά του να ερμηνεύσει τα φαινόμενα μνήμης της προηγούμενης εντατικής κατάστασης (strain memory effects) των υλικών, όπως ο ανοξειδωτος χάλυβας, πραγματοποίησε δοκιμή μονοαξονικής συμπίεσης στον ανοξειδωτο χάλυβα 316 με ανακυκλιζόμενες συνθήκες φόρτισης. Τα αποτελέσματα που προέκυψαν από τη δοκιμή, αυτή, απεικονίζονται στο σχήμα (4.6) με τη μορφή καμπύλης τάσης – παραμόρφωσης.



Σχήμα 4.6: Πειραματική καμπύλη τάσης – παραμόρφωσης ανοξειδωτου χάλυβα σε δοκιμή συμπίεσης υπό ανακυκλιζόμενες συνθήκες φόρτισης

Ένα μέρος της καμπύλης του σχήματος (4.6) επιλέχθηκε για την διαδικασία εκπαίδευση και κατά συνέπεια τη μοντελοποίηση του νόμου του υλικού από το νευρωνικό δίκτυο (NN). Το τμήμα της καμπύλης που επιλέχθηκε παρουσιάζεται στο διάγραμμα 4.9. Τα δεδομένα εκπαίδευσης του NN επιλέχθηκαν από αυτό το διάγραμμα τυχαία, μέσα από το εύρος τιμών τόσο της παραμόρφωσης, όσο και της τάσης.

Σύμφωνα με την ανάλυση της μεθόδου για την μοντελοποίηση της υστερητικής συμπεριφοράς ενός υλικού με NN (§4.3.1.1) και ύστερα από την επιλογή της ελεγχόμενης, από την παραμόρφωση, μορφή του NN η δομή του είναι η ακόλουθη:

$$\sigma_n = \hat{\sigma}_{NN}(\{\varepsilon_n, \varepsilon_{n-1}, \sigma_{n-1}, \xi_{\varepsilon,n}, \Delta\eta_{\varepsilon,n}\}; \{5/30/30/1\})$$

Ο επιλογή του αριθμού των κόμβων ή των νευρώνων των κρυμμένων επιπέδων, έγινε τυχαία. Χρειάζεται βέβαια να αναφερθεί, ότι στη συγκεκριμένη περίπτωση, υπάρχει ένα κάτω όριο στο πλήθος του αριθμού των κόμβων ανά επίπεδο, που είναι οι 15 κόμβοι και κάτω από αυτό η μέθοδος βελτιστοποίησης της συνάρτησης σφάλματος δεν παρουσιάζει τα σωστά αποτελέσματα σύγκλισης που παρουσιάζονται πάνω από αυτό το όριο. Περισσότερες λεπτομέρειες όσον αφορά τον αριθμό των κόμβων των δύο κρυμμένων επιπέδων θα δοθούν, στη συνέχεια, στην παρουσίαση των αποτελεσμάτων.



Διάγραμμα 4.9: Τμήμα της καμπύλης ($\sigma - \epsilon$) του ανοξειδωτού χάλυβα που χρησιμεύει στην εκπαίδευση του NN

Στο διάγραμμα 4.10 παρουσιάζονται τα αποτελέσματα από την εκπαίδευση του νευρωνικού δικτύου. Παρατηρείται ότι τα προβλεπόμενα, από το NN, αποτελέσματα ταυτίζονται με τα πειραματικά. Η καλή συμπεριφορά του NN οφείλεται αρχικά στην εισαγωγή των εσωτερικών μεταβλητών, ως διανύσματα εισόδου του NN, σύμφωνα με την μέθοδο της §4.3.1.1, η οποία βοήθησε το NN στην εκμάθηση της ιστορίας της φόρτισης (ή της πορείας της καμπύλης $\sigma - \epsilon$). Επίσης η ταύτιση των αποτελεσμάτων οφείλεται στη προσεχτική επιλογή των ζευγών τάσεων – παραμορφώσεων στην τρέχουσα κατάσταση, όπου οι μεν παραμορφώσεις αποτελούν ένα από τα διανύσματα εισόδου του NN και οι τάσεις αποτελούν το διάνυσμα «στόχο», που πρέπει να προσεγγίσει, όσο το δυνατόν καλύτερα.



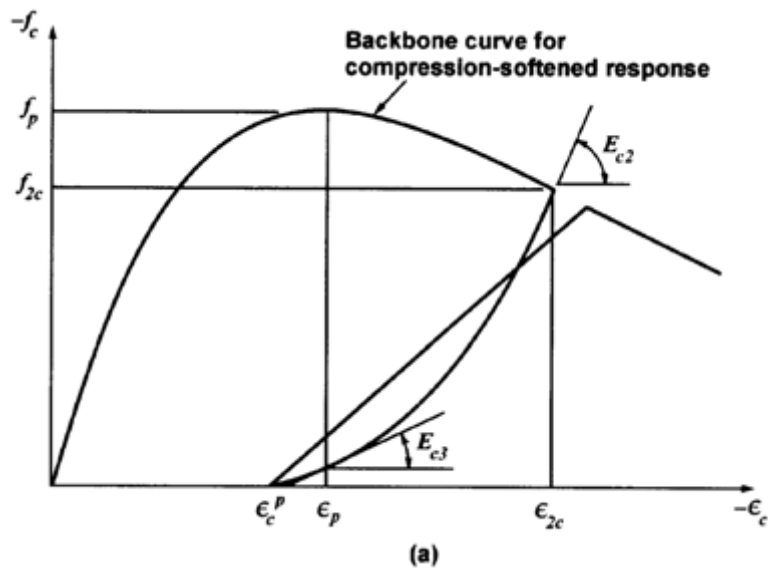
Διάγραμμα 4.10: Πρόβλεψη NN για την καμπύλη $\sigma - \epsilon$ ανοξειδωτού χάλυβα σε μονοαξονική συμπίεση

Επιπλέον χρειάζεται να αναφερθεί, ότι ύστερα από δοκιμές, τόσο το πλήθος του αριθμού των κόμβων των κρυμμένων επιπέδων, πάνω φυσικά από το όριο των 15 κόμβων ανά επίπεδο, όσο και ο αριθμός των επαναλήψεων (Epochs) για την διαδικασία εκπαίδευσης, φαίνεται να μην έχουν επίδραση στην ακρίβεια των αποτελεσμάτων. Ο αριθμός των Epochs επιλέχθηκε στις 5000 και ο αριθμός των κόμβων των κρυμμένων επιπέδων επιλέχθηκε, όπως φαίνεται και παραπάνω στους 30 ανά κρυμμένο επίπεδο.

4.3.3 Μονοαξονική συμπίεση οπλισμένου σκυροδέματος

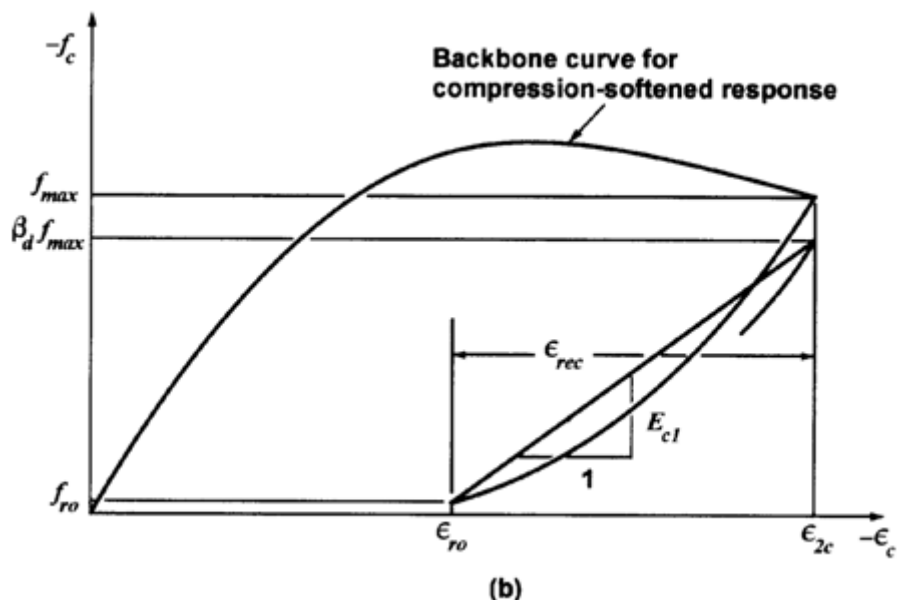
Η γενική μορφή της καμπύλης τάσεων – παραμορφώσεων που αντιπροσωπεύει την μονοαξονική συμπίεση σκυροδέματος σε κυκλικές συνθήκες φόρτισης απεικονίζεται στα σχήματα (4.7) και (4.8).

Όπως παρατηρείται από τα παρακάτω σχήματα το σχήμα και η κλίση της καμπύλης φόρτισης και αποφόρτισης εξαρτάται από την πλαστική παραμένουσα παραμόρφωση ϵ_c^p , η οποία προέρχεται από τη θραύση του σκυροδέματος, από εσωτερικές ρωγμές ή από την ύπαρξη εσωτερικών κενών.



Σχήμα 4.7: Υστερητικό μοντέλο σκυροδέματος σε συμπίεση – συμπεριφορά αποφόρτισης

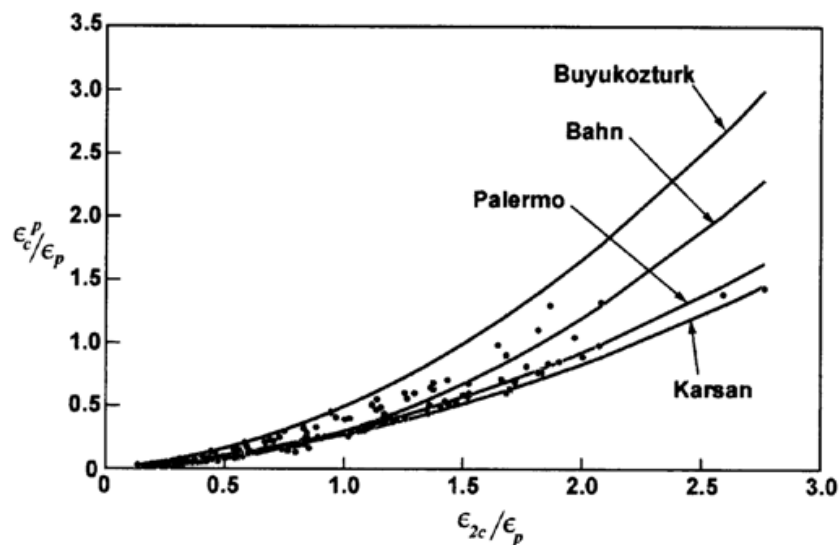
όπου: f_p είναι η μέγιστη κύρια τάση συμπίεσης του σκυροδέματος, ϵ_p είναι η παραμόρφωση που αντιστοιχεί στην μέγιστη τάση συμπίεσης, ϵ_c^p είναι η παραμένουσα πλαστική παραμόρφωση του σκυροδέματος, f_{2c} είναι η τάση αποφόρτισης του σκυροδέματος στην περιβάλλουσα καμπύλη συμπίεσης, ϵ_{2c} είναι η παραμόρφωση που αντιστοιχεί στην προηγούμενη τάση, E_{c2} είναι το αρχικό μέτρο ελαστικότητας και E_{c3} είναι το μέτρο ελαστικότητας στην φάση της αποφόρτισης για μηδενική τάση.



Σχήμα 4.8: Υστερητικό μοντέλο σκυροδέματος σε συμπίεση - συμπεριφορά επαναφόρτισης

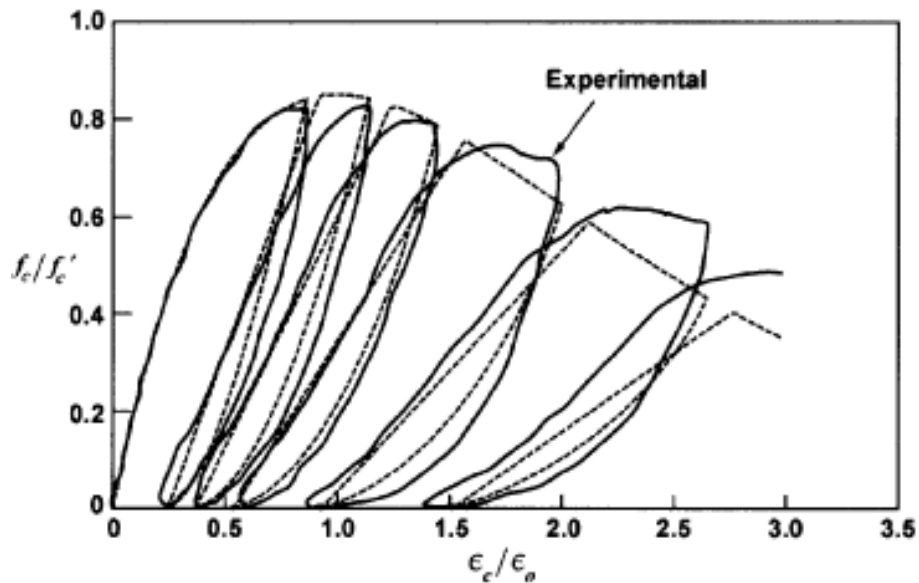
όπου: f_{max} είναι η μέγιστη τάση σκυροδέματος για την τρέχουσα καμπύλη αποφόρτισης, ϵ_{2c} είναι η παραμόρφωση που αντιστοιχεί στην προηγούμενη τάση, β_d είναι ένας συντελεστής μείωσης της αντοχής, f_{r0} είναι η τάση συμπίεσης στην έναρξη της καμπύλης επαναφόρτισης, ϵ_{r0} είναι η παραμόρφωση που αντιστοιχεί στην προηγούμενη τάση, ϵ_{rec} είναι η «ελαστική» παραμόρφωση κατά την καμπύλη αποφόρτισης και E_{c1} είναι το μέτρο ελαστικότητας του σκυροδέματος της καμπύλης επαναφόρτισης στη συμπίεση.

Υπάρχουν πολλά μοντέλα πλαστικής παραμένουσας παραμόρφωσης ϵ_c^p για την συμπίεση του σκυροδέματος, τα οποία έχουν καταγραφή στη βιβλιογραφία. Το πρώτο που αναφέρθηκε είναι αυτό, των Karsan και Jirsa [57], σύμφωνα με το οποίο παρουσιάζεται η εξάρτηση της πλαστικής παραμένουσας παραμόρφωσης ϵ_c^p από την παραμόρφωση στην έναρξη της καμπύλης αποφόρτισης ϵ_{2c} . Στο σχήμα που ακολουθεί παρουσιάζεται γραφικά το μοντέλο των Karsan και Jirsa, καθώς και άλλα αντίστοιχα μοντέλα. Επίσης στο σχήμα αυτό οι παραμορφώσεις ϵ_c^p και ϵ_{2c} είναι ανηγμένες ως προς την παραμόρφωση ϵ_p .



Σχήμα 4.9: Μοντέλα πλαστικής παραμόρφωσης για την συμπίεση του σκυροδέματος

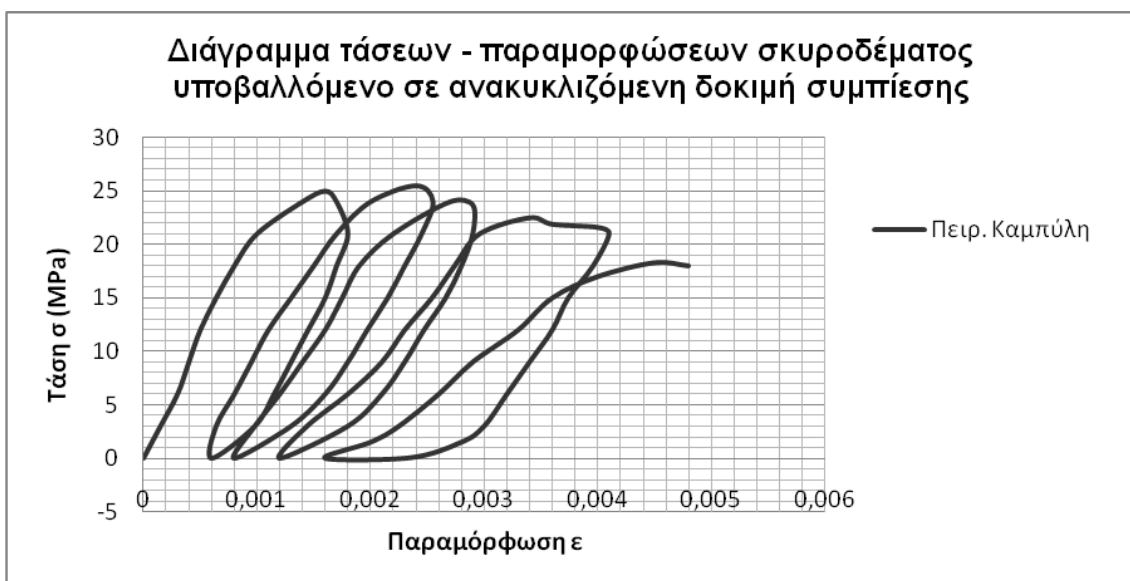
Στη συνέχεια παρουσιάζονται τα πειραματικά αποτελέσματα από δοκιμές συμπίεσης που πραγματοποίησαν οι Karsan και Jirsa σε οπλισμένο σκυρόδεμα με κυκλικές συνθήκες φόρτισης.



Σχήμα 4.10: Πειραματικά αποτελέσματα από δοκιμή συμπίεσης σε σκυροδέμα

Όπως παρατηρείται από το παραπάνω σχήμα, οι τιμές των τάσεων f_c και των παραμορφώσεων ϵ_c είναι ανηγμένες στην χαρακτηριστική μέγιστη τάση του σκυροδέματος f_c' και στην παραμένουσα πλαστική παραμόρφωση ϵ_0 , αντίστοιχα.

Για τις δεδομένες τιμές των $f_c' = 30\text{MPa}$ και $\epsilon_0 = 0.002$ προκύπτει το επόμενο διάγραμμα, το οποίο αποτελεί μέρος του σχήματος των πειραματικών αποτελεσμάτων. Το διάγραμμα αυτό, χρησιμοποιήθηκε για την εκπαίδευση του νευρωνικού δικτύου. Τα δεδομένα εκπαίδευσης του νευρωνικού δικτύου (NN) προέκυψαν με τυχαία επιλογή από το παρακάτω διάγραμμα, από όλο το εύρος τιμών των τάσεων και παραμορφώσεων, αντίστοιχα.



Διάγραμμα 4.11: Τμήμα της καμπύλης ($\sigma - \epsilon$) του σκυροδέματος υποβαλλόμενο σε δοκιμή ανακυκλιζόμενης συμπίεσης που χρησιμεύει στην εκπαίδευση του NN

Όπως και στην περίπτωση του χάλυβα σε ανακυκλιζόμενη φόρτιση, έτσι και εδώ η δομή του δικτύου, σύμφωνα πάντα και με την μέθοδο μοντελοποίησης της υστερητικής συμπεριφοράς των υλικών (§4.3.1.1), είναι η ακόλουθη:

$$\sigma_n = \hat{\sigma}_{NN}(\{\varepsilon_n, \varepsilon_{n-1}, \sigma_{n-1}, \xi_{\varepsilon,n}, \Delta\eta_{\varepsilon,n}\}; \{5/30/30/1\})$$

Το όριο των 15 κόμβων ανά κρυμμένο επίπεδο ισχύει και σε αυτή την περίπτωση. Περισσότερες λεπτομέρειες θα δοθούν στην συνέχεια.

Στο διάγραμμα 4.12 απεικονίζονται τα αποτελέσματα από την εκπαίδευση του NN. Όπως παρατηρείται τα προβλεπόμενα, από το NN, αποτελέσματα συμπίπτουν με τα πειραματικά του διαγράμματος 4.11. Η καλή προσέγγιση του NN στα πειραματικά αποτελέσματα οφείλεται στην εισαγωγή των εσωτερικών μεταβλητών $\xi_{\varepsilon,n}$ και $\Delta\eta_{\varepsilon,n}$, ως διανύσματα εισόδου του NN, οι οποίες βοήθησαν το NN στην εκμάθηση της ιστορίας της φόρτισης (ή της πορείας της καμπύλης $\sigma - \varepsilon$). Επίσης η ταύτιση των αποτελεσμάτων οφείλεται στη κατάλληλη επιλογή των ζευγών τάσεων (διάνυσμα στόχου) και παραμορφώσεων (ένα από τα διανύσματα εισόδου του NN) στην τρέχουσα κατάσταση.



Διάγραμμα 4.12: Πρόβλεψη NN για την καμπύλη $\sigma - \varepsilon$ σκυροδέματος σε μονοαξονική συμπίεση

Τέλος χρειάζεται να σημειωθεί, ότι ύστερα από δοκιμές, τόσο το πλήθος του αριθμού των κόμβων των κρυμμένων επιπέδων, πάνω φυσικά από το όριο των 15 κόμβων ανά επίπεδο, όσο και ο αριθμός των επαναλήψεων (Epochs) για την διαδικασία εκπαίδευσης, φαίνεται να μην έχουν επίδραση στην ακρίβεια των αποτελεσμάτων. Ο αριθμός των Epochs επιλέχθηκε στις 5000 και ο αριθμός των κόμβων των κρυμμένων επιπέδων επιλέχθηκε, όπως φαίνεται και παραπάνω στους 30 ανά κρυμμένο επίπεδο.

4.3.4 Τριαξονική δοκιμή του εδάφους κάτω από κυκλική αξονική φόρτιση

Οι Javadi, Tan και Elkassas [58] στην προσπάθειά τους να μελετήσουν και να μοντελοποιήσουν μέσω νευρωνικού δικτύου την συμπεριφορά του εδάφους κάτω από κυκλική φόρτιση πραγματοποίησαν τριαξονικές δοκιμές. Για τις δοκιμές αυτές χρησιμοποίησαν το τροποποιημένο μοντέλο Cam clay του εδάφους, με τις παρακάτω παραμέτρους:

$$\lambda = 14$$

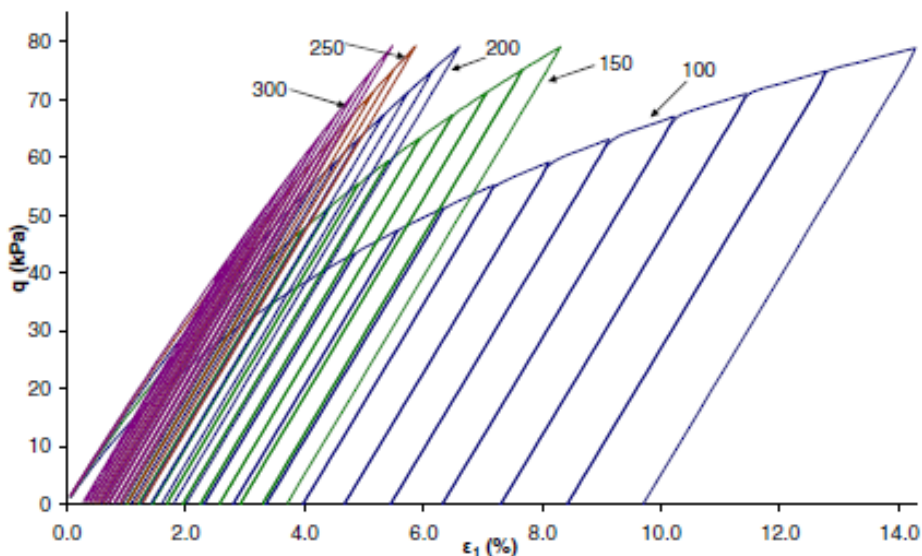
$$\kappa = 0.015$$

$$M = 0.8$$

$$P_0 = 100\text{kPa}$$

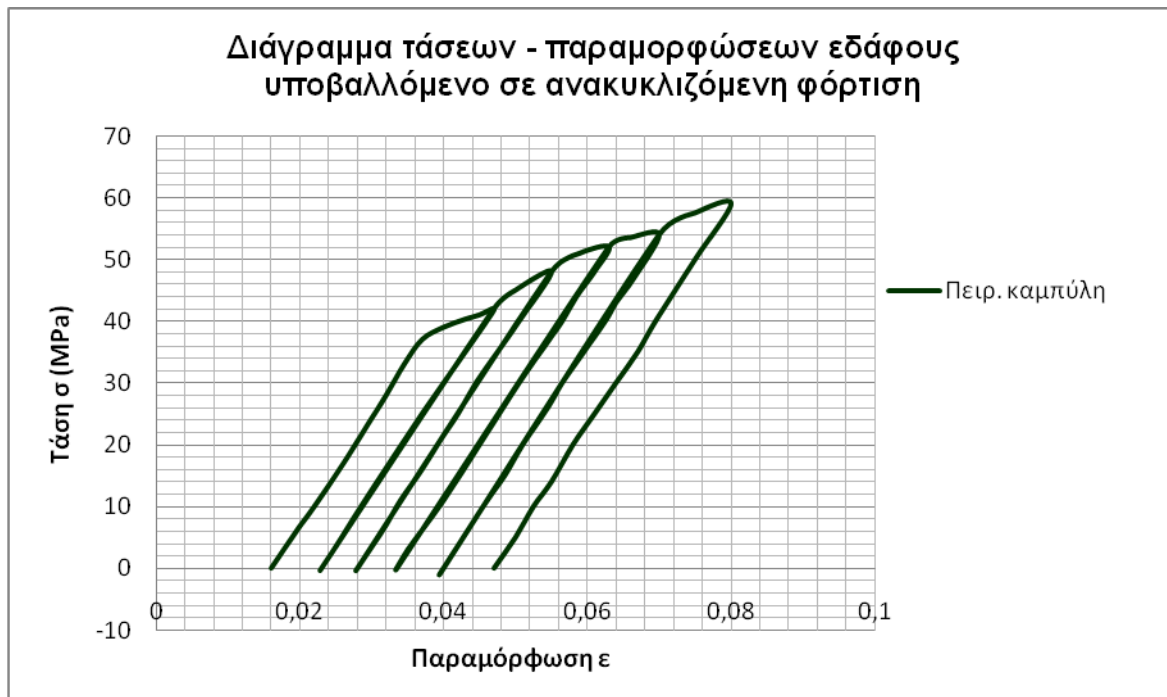
$$N = 2.68$$

Το επόμενο σχήμα δείχνει τα αποτελέσματα των πειραμάτων για διάφορες τιμές της πίεσης περιορισμού 100, 150, 200, 250 και 300 kPa.



Σχήμα 4.11: Πειραματικά αποτελέσματα από τριαξονικές δοκιμές στο μοντέλο εδάφους Cam clay για τιμές πιέσεων 100, 150, 200, 250 και 300 kPa

Από το παραπάνω σχήμα γίνεται επιλογή της καμπύλης, που αντιστοιχεί σε πίεση 100kPa, για την εκπαίδευση του νευρωνικού δικτύου (NN). Μέρος της καμπύλης αυτής παρουσιάζεται στο διάγραμμα 4.13.



Διάγραμμα 4.13: Τμήμα της καμπύλης ($\sigma - \varepsilon$) του εδάφους υποβαλλόμενο σε τριαξονική δοκιμή ανακυκλιζόμενης φόρτισης που χρησιμεύει στην εκπαίδευση του NN

Η επιλογή των δεδομένων για τη διαδικασία εκπαίδευσης του NN έγινε από το παραπάνω διάγραμμα, με τυχαίο τρόπο από το εύρος τιμών των παραμορφώσεων και τάσεων. Όπως και στις δύο προηγούμενες περιπτώσεις της ανακυκλιζόμενης φόρτισης για το χάλυβα και το σκυρόδεμα, έτσι και σε αυτή την περίπτωση η δομή του νευρωνικού δικτύου (σύμφωνα με την §4.3.1.1) είναι:

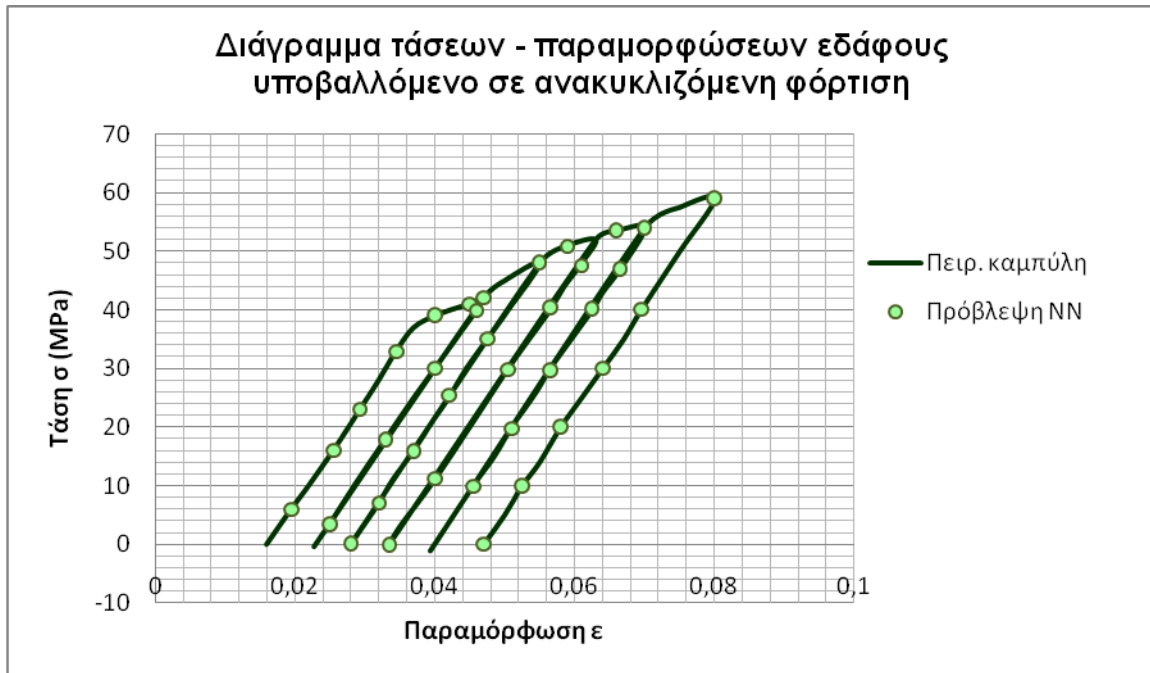
$$\sigma_n = \hat{\sigma}_{NN}(\{\varepsilon_n, \varepsilon_{n-1}, \sigma_{n-1}, \xi_{\varepsilon,n}, \Delta\eta_{\varepsilon,n}\}; \{5/30/30/1\})$$

Επίσης το όριο των 15 κόμβων ανά κρυμμένο επίπεδο ισχύει και σε αυτή την περίπτωση. Στο διάγραμμα 4.14 παρουσιάζονται τα αποτελέσματα από την εκπαίδευση και πρόβλεψη του νευρωνικού δικτύου.

Παρατηρώντας το παρακάτω διάγραμμα γίνεται φανερό ότι υπάρχει πολύ καλή προσέγγιση των αποτελεσμάτων του NN με τα πειραματικά αποτελέσματα. Αυτό οφείλεται όπως και στις δύο προηγούμενες περιπτώσεις στις εσωτερικές μεταβλητές $\xi_{\varepsilon,n}$ και $\Delta\eta_{\varepsilon,n}$, ως δύο από τα διανύσματα εισόδου και στην προσεχτική επιλογή των ζευγών τάσεων - παραμορφώσεων στην τρέχουσα κατάσταση.

Επιπλέον ο αριθμός των κόμβων ανά επίπεδο, πάνω από το όριο των 15/επίπεδο και ο αριθμός των επαναλήψεων της διαδικασίας εκμάθησης, δεν επηρεάζουν την ακρίβεια

των αποτελεσμάτων. Έτσι επιλέχθηκε και πάλι αριθμός Epochs ίσος με 5000 και πλήθος κόμβων ανά κρυμμένο επίπεδο ίσο με 30.



Διάγραμμα 4.14: Πρόβλεψη NN για την καμπύλη $\sigma - \epsilon$ του εδάφους σε τριαξονική δοκιμή για πίεση 100 kPa

Συμπερασματικά, μετά την ανάλυση όλων των παραπάνω, προκύπτει ότι τα νευρωνικά δίκτυα με κατάλληλα δεδομένα εισόδου, ανάλογα με την περίπτωση που μελετάται κάθε φορά (μονοτονική ή ανακυκλιζόμενη φόρτιση), μπορούν να αποτελέσουν ένα αξιόπιστο εργαλείο για τη μοντελοποίηση της καταστατικής συμπεριφοράς των υλικών. Επίσης η χρήση πειραματικών αποτελεσμάτων διευκολύνει ακόμα περισσότερο την διαδικασία εκπαίδευσης του, πρώτον διότι ελαχιστοποιεί το χρόνο εκμάθησης, ο οποίος είναι ήδη μικρός και δεύτερον γιατί αποκτά άμεσα την απαιτούμενη πληροφορία για τη συμπεριφορά ενός υλικού, υπό συγκεκριμένες συνθήκες φόρτισης, χωρίς να γίνεται απαραίτητη η μεσολάβηση περίπλοκων, σε πολλές περιπτώσεις, εξισώσεων που προσεγγίζουν τη δεδομένη συμπεριφορά. Για τους παραπάνω λόγους η χρήση των νευρωνικών δικτύων επιτρέπει την άμεση ανάπτυξη καταστατικών μοντέλων διαφόρων υλικών και για διάφορες καταστάσεις φόρτισης. Οι δύο απαιτήσεις που υπάρχουν για την αξιόπιστη «εκμάθηση» της συμπεριφοράς ενός υλικού σε δεδομένη φόρτιση από τα νευρωνικά δίκτυα είναι αρχικά η κατάλληλη επιλογή των δεδομένων εισόδου, όπως τάσεις ή παραμορφώσεις στην τρέχουσα κατάσταση ή στην τρέχουσα κατάσταση, σε προηγούμενες ή/και άλλες εσωτερικές μεταβλητές, που εκφράζουν την κάθε περίπτωση χωριστά. Δεύτερον είναι η επιλογή του κατάλληλου εύρους των τιμών των δεδομένων εισόδου.

Γενικά πάντως τα καταστατικά μοντέλα νευρωνικών δικτύων, ειδικά τα τελευταία χρόνια, πληθαίνουν όλο και περισσότερο, παρά τους δύο τελευταίους περιορισμούς που αναφέρθηκαν, οι οποίοι στις περισσότερες των περιπτώσεων, είναι εύκολα προσπελάσιμοι. Τέλος πρόκληση, αλλά και έρευνα για το μέλλον, αποτελεί και η μοντελοποίηση τριαξονικών καταστατικών μοντέλων υλικών.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Martin H. Sadd, "Elasticity: Theory, applications and numerics", Elsevier, Oxford 2005.
- [2] Menegotto M. and Pinto P., "Method of Analysis of Cyclically Loaded Reinforced Concrete Plane Frames Including Changes in Geometry and Nonelastic Behavior of Elements Under Combined Normal Geometry and Nonelastic Behavior of Elements Under Combined Normal Force and Bending." Proceedings of the IABSE Symposium on the Resistance and Ultimate Deformability of Structures Acted on by Well-Defined Repeated Loads. Lisbon, 1973.
- [3] Stanton J.F. and McNiven H.D., "The Development of a Mathematical Model to Predict the Flexural Response of Reinforced Concrete Beams to Cyclic Loads, Using System Identification", Report UCB/EERC-79/02. Berkeley: EERC, University of California, 1979.
- [4] Filippou F.C., Popov E. and Bertero V.V., "Effects of Bond Deterioration on Hysteretic Behavior of Reinforced Concrete Joints", Report No. UCB/EERC 83-19. Berkeley: EERC, University of California, 1983.
- [5] Chang, G.A. and Mander J.B., "Seismic Energy Based Fatigue Damage Analysis of Bridge Columns: Part I - Evaluation of Seismic Capacity", Technical Report NCEER-94-0006, NCEER, SUNY, Buffalo, New York, 1994.
- [6] Kent D.C. and Park R., "Flexural members with confined concrete." Journal of Structural Division, ASCE; Vol. 97: pp. 1964 – 1990, 1971.
- [7] Ulrich Smolczyk, Geotechnical engineering handbook, Publisher: Ernst & Sohn, Berlin 2003.
- [8]. Adachi, T and Oka, F., "Constitutive equations for normally consolidated clay based on elasto-viscoplasticity", Soils & Foundations, Vol. 22, No. 5, pp 57-70, 1982.
- [9]. Goldscheider, M., "True triaxial tests on dense sands". In: G. Gudehus, F.Darve, I. Vardoulakis; editors. Constitutive relations for soils. Balkema, Rotterdam, 1984.
- [10]. Burland, JB., "The yielding and dilation of clay correspondence", Geotechnique, Vol. 15, No. 2, pp 211-219, 1965.

- [11]. Duncan, JM and Chang, CY., "Nonlinear analysis of stress and strain in soils", ASCE J. of the Soil Mech. And Found. Div. Vol. 96, pp 1629-1653, 1970.
- [12] Brinkgreve, RBJ and Vermeer, PA., "Plaxis finite element code for soil and rock analysis-Version 7", Balkema, Rotterdam, 1997.
- [13] Papadrakakis, M., Papadopoulos, V., and Lagaros, N.D., "Structural reliability analysis of elastic-plastic structures using Neural Networks and Monte Carlo simulation", Comp. Meth. Appl. Mechanics & Engrg, vol. 136, pp. 145-163, 1996.
- [14] Papadrakakis, M., Lagaros, N.D., and Tsompanakis, Y., "Structural optimization using Evolution Strategies and Neural Networks", Comp. Meth. Appl. Mechanics & Engrg, vol. 156, pp. 309-333, 1998.
- [15] Papadrakakis M., Lagaros N.D., and Tsompanakis Y., Optimization of large-scale 3D trusses using Evolution Strategies and Neural Networks, Space Structures Special Issue, (to appear), 1997.
- [16] Berke, L., and Hajela, P., "Applications of Artificial Neural Nets in Structural Mechanics", NASA pp. 331-348 TM-102420, 1990.
- [17] Berke, L., Patnaik, S.N., and Murthy, P.L.N., "Optimum Design of Aerospace Structural Components using Neural Networks", Computers & Structures, Vol. 48, pp. 1001-1010, 1993.
- [18] Fu, B., and Hajela, P., "Minimizing Distortion in Truss Structures-A Hopfield Network Solution", Comp. Systems in Enging., Vol. 4, pp. 69-74, 1993.
- [19] Theocaris, P.S., and Panagiotopoulos, P.D., "Neural networks for computing in fracture mechanics. Methods and prospects of applications", Comp. Meth. in Appl. Mech. Engrg., Vol. 106, pp. 213-228, 1993.
- [20] Khan, A.I., Topping, B.H.V., and Bahreininejad, A., "Parallel training of neural networks for finite element mesh generation", Neural Networks and Combinatorial Optimisation in Civil and Structural Engineering, Topping, B.H.V., and Khan A.I., (eds), Civil-Comp Press, pp. 81-94, 1993.
- [21] McCulloch, W.S., and Pitts, W., "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133, 1943.

- [22] Rosenblatt, F., "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, Spartan Press, Washington, DC, 1962.
- [23] Windrow, B., "Generalization and information storage in networks of Adaline neurons", In Self-Organizing Systems, Yovitz, M., Jacobi, G., and Goldstein, G., (eds.), Spartan books, Washington, D.C., pp. 435-461, 1962.
- [24] Rummelhart, D.E., and McClelland, J.L., "Parallel Distributed Processing", Volume 1: Foundations, The MIT Press, Cambridge, 1986.
- [25] Lippmann, R.P., "An introduction to computing with neural nets", IEEE Acoustics, Speech and Signal Processing Magazine, pp. 4-22, April 1987.
- [26] Windrow, B., and Lehr, M.A., "30 years of adaptive neural networks: Perceptron, Madaline, and Backpropagation", Proceedings of the IEEE, vol. 78, pp. 1415-1442, 1990.
- [27] Hopfield, J.J., "Neural networks and physical systems with emergent collective computational abilities", Proceedings of the National Academy of Sciences, USA, vol. 79, pp. 2554-2558, 1982.
- [28] Kohonen, T., "Self-organization and associative memory", Series in Information Science, vol. 8, Berlin, Springer Verlag, 1984.
- [29] Hebb, D.O., "Organization of behavior", New York: Science Editions 1961.
- [30] Cauchy, A., "Method generale pour la resolution des systemes d'equations simultanees". Compt. Rend. Aca. Sci., vol. 25, pp. 536-538, 1847.
- [31] Fletcher, R., and Powell, M.J.D., "A rapidly convergent descent method for minimization". The Computer Journal, vol. 6, pp. 163-180, 1963.
- [32] Marquardt, D.W., "An algorithm for least squares estimation of nonlinear parameters", SIAM J., vol. 15, pp. 641-652, 1963.
- [33] Wasserman, P.D., "Combined backpropagation/Cauchy machine", Neural Networks: Abstracts of the First INNS meeting, vol. 1, pp. 556, Elmsford, NY, Boston, Pergamon Press, 1988.

[34] Wasserman, P.D., "Experiments in translating Chinese characters using back-propagation", Proceedings of the thirty-third IEEE Computer Society International Conference, Washington, D.C., Computer Society Press of the IEEE, 1988.

[35] Martin R. and Heinrich B., "A Direct Method for Faster Backpropagation Learning: The RPROP Algorithm", IEEE, 1993

[36] Battiti, R., "First- and second-order methods for learning: between steepest descent and Newton's method", Neural Computation, vol. 4, pp.141-166, 1992.

[37] Cichocki, A., and Unbehauen, R., "Neural networks for optimization and signal processing", John Wiley & Sons, pp. 142-152, 1993.

[38] Yu, Y-H, and Simmons, R.F., "Descending Epsilon in Back-Propagation: A Technique for better generalization", International Joint Conference on Neural Networks, 1990.

[39] Allen, R., "Several studies on natural language and Back-propagation", IEEE International Conference on Neural Networks, 1987.

[40] Hornik, K., "Multilayer Feedforwrd Networks are Universal Approximators", Neural Networks, Vol. 2, pp. 359-366, 1989.

[41] Rummelhart, D.E., and McClelland, J.L., "Parallel Distributed Processing", Volume 1: Foundations, The MIT Press, Cambridge, 1986.

[42] Sejnowski, T.J., and Rosenberg, C.R., "Parallel networks that learn to pronounce English text", Complex Systems, vol. 1, pp. 145-168, 1987.

[43] Wasserman, P.D., "Neural computing, Theory and Practice", ANZA Research Inc., Van Nostrand Reinhold, New York, 1989.

[44] Sietsma, J., Dow, R.J.F., "Creating artificial neural networks that generalize", Neural Networks, vol. 4, pp.67-79, 1991.

[45] Hirose, K. Yamashita and S. Hijiya, 'Back-propagation algorithm which varies the number of hidden units', Neural Networks, Vol. 4, pp. 61-66, 1991.

[46] Metropolis, N., Rosenbluth, A., Rosenbluth, M.N., Teller, A.H., and Teller, E., "Equations of state calculations by fast computing machines", Journal of Chemistry and Physics, Vol. 21, pp. 1087-1091, 1953.

- [47] Geman, S., and Geman, D., "Stochastic relaxation, Gibbs distributions and Bayesian restoration of images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 6, pp. 721-741, 1984.
- [48] Szu, H. and Hartley, R., "Fast simulated annealing", Physics Letters, Vol. 1222 (3,4), pp. 157-162, 1987.
- [49] Chu, C.H., and Chow, C.R. "A Genetic Algorithm Approach to Supervised Learning for Multilayered Networks", Proceedings of WCNN93, pp. IV744-IV747, 1993.
- [50] P. de la Brassinne, "Generic algorithms and learning of neural nets", Bulletin Scientifique de l'Association des Ingenieurs Electriciens sortis de l'Institut Electrotechnique Montefiore, Vol. 106(1), pp. 41-58, 1993.
- [51] Arena, P., and Caponetto, R., Fortuna, I., and Xibilia, M.G., "MLP Optimal Topology via Genetic Algorithms", Proceedings of ANNGA93, pp. 670-674, 1993.
- [52] Kirby BR, Preston RR. "High temperature properties of hot-rolled, structural steels for use in fire engineering design studies", Fire Safety Journal, pp 27–37, 1988.
- [53] Mougín, J. P., "Calcul des éléments simples et des structures de bâtiments", Eyrolles (Ed.), Paris, 1992.
- [54] S. Pernot, C.-H. Lamarque, "Application of neural networks to the modelling of some constitutive laws", Elsevier, 1998.
- [55] Gun Jin Yun, Jamshid Ghaboussi and Amr S. Elnashai, "A new neural network - based model for hysteretic behavior of materials", John Wiley & Sons, 2007.
- [56] J.L. Chaboche, "A review of some plasticity and viscoplasticity constitutive theories", Elsevier, 2008.
- [57] Karsan, I. K. and Jirsa, J. O., "Behaviour of Concrete Under Compressive Loadings", ASCE, pp. 2543 – 2563, 1969.
- [58] A.A. Javadi, T.P. Tan, and A.S.I. Elkassas, "Intelligent Finite Element Method and Application to Simulation of Behavior of Soils under Cyclic Loading", SCI 205, pp. 317–338, 2009.