



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΜΗΜΑ ΤΟΠΟΓΡΑΦΙΑΣ – ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

Διπλωματική εργασία

**Διερεύνηση μεθόδων βαθιάς μάθησης για φασματικό διαχωρισμό
υπερφασματικών απεικονίσεων**

Γιώργος Γιαννόπουλος

Επιβλέπουσα Καθηγήτρια:

Δρ Βασιλεία Καραθανάση

Αθήνα, Φεβρουάριος 2023

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την Καθηγήτρια Βασιλεία Καραθανάση που μου πρότεινε ένα τόσο ενδιαφέρον θέμα καθώς και για την παρουσία της, το ενδιαφέρον της και την άμεση ανταπόκρισή της όποτε χρειαζόμουν την καθοδήγησή της.

Επίσης θα ήθελα να ευχαριστήσω τον κ.Βασίλη Ανδρώνη, μέλος του εργαστηρίου Τηλεπισκόπησης, για την πολύτιμη βοήθεια του στην κατανόηση και τον προγραμματισμό ορισμένων μεθόδων επεξεργασίας των δεδομένων.

Πίνακας Περιεχομένων

Περίληψη.....	5
Abstract.....	6
1. Εισαγωγή.....	7
2. Φασματικός διαχωρισμός.....	8
2.1. Το γραμμικό μοντέλο.....	8
2.2. Φασματική γωνία και φασματική απόκλιση.....	9
3. Νευρωνικά δίκτυα.....	10
3.1. Gradient descent.....	10
3.2. Ο νευρώνας.....	10
3.3. Συναρτήσεις ενεργοποίησης.....	11
3.4. Πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης.....	12
3.5. Οπισθοδιάδοση.....	13
3.6. Batch Normalization.....	13
3.7. Συνελικτικά δίκτυα.....	14
3.8. Αυτοκωδικοποιητές.....	15
4. Βιβλιογραφική ανασκόπηση.....	16
5. Μέθοδοι.....	17
5.1. DAEU.....	17
5.2. SIDAEU.....	19
5.3. MTAEU.....	19
5.4. CNNAEU.....	21
5.5. EndNet.....	22
5.6. OSPAEU.....	24
6. Αποτελέσματα πειραμάτων.....	27
5.1. Καθαροί στόχοι.....	28
5.1.1 Samson.....	28
5.1.2 Urban με τέσσερις καθαρούς στόχους.....	34
5.1.3 Urban με πέντε καθαρούς στόχους.....	40
5.1.4. Αξιολόγηση ως προς την εύρεση των καθαρών στόχων.....	46
5.2. Αφθονίες.....	50
5.2.1. Samson.....	50
5.2.2. Urban με τέσσερις καθαρούς στόχους.....	52
5.2.3. Urban με πέντε καθαρούς στόχους.....	55
5.3. Συμπεράσματα.....	58
7. Πρόγραμμα εκτέλεσης φασματικού διαχωρισμού.....	59
8. Βιβλιογραφία.....	60

Περίληψη

Ο γραμμικός φασματικός διαχωρισμός είναι η διαδικασία έκφρασης των εικονοστοιχείων μια υπερφασματικής εικόνας σε καθαρούς στόχους και αφθονίες με ταυτόχρονη εξαγωγή τους. Σε αυτή την εργασία εκτελείται φασματικός διαχωρισμός με έξι διαφορετικούς αλγορίθμους βαθιάς μάθησης με σκοπό την συγκριτική αξιολόγησή τους. Αποδεικνύεται πως η αρχιτεκτονική ενός αποκωδικοποιητή, εφόσον σε αυτή τεθούν μια σειρά περιορισμών, ανταποκρίνεται πολύ καλά στο γραμμικό μοντέλο και μπορεί να απεικονίσει την φυσική και μαθηματική υπόσταση του. Ένας αποκωδικοποιητής είναι ένα πλήρως διασυνδεδεμένο νευρωνικό δίκτυο μη επιβλεπόμενης μάθησης που κωδικοποιεί τα δεδομένα εισόδου σε μια κρυφή αναπαράσταση χαμηλότερων διαστάσεων και έπειτα προσπαθεί να ανακατασκευάσει την εικόνα αποκωδικοποιώντας την κρυφή αναπαράσταση αυτή. Σε αυτό το μοντέλο η κρυφή αναπαράσταση αποτελεί τους χάρτες αφθονίας και τα βάρη του αποκωδικοποιητή τους καθαρούς στόχους.

Από την βιβλιογραφία επιλέχθηκαν έξι μέθοδοι βαθιάς μάθησης για την συγκριτική αξιολόγησή τους [6] - [11]. Οι μέθοδοι συγκρίθηκαν με χρήση δύο διαφορετικών ευρέως χρησιμοποιούμενων εικόνων που συνοδεύονται από εξαγομένους χάρτες αφθονιών και καθαρούς στόχους. Πραγματοποιήθηκαν δέκα πειράματα για κάθε μέθοδο σε κάθε σύνολο δεδομένων κατά τα οποία εξάχθηκαν οι καθαροί στόχοι και οι αντίστοιχες αφθονίες. Από τις μεθόδους πιο ακριβείς βρέθηκαν η ΜΤΑΕU[8] που ενσωματώνει την χωρική πληροφορία της υπερφασματικής εικόνας στο μοντέλο και η CNNΑΕU[9] που είναι συνελκτικός αποκωδικοποιητής και άρα εκμεταλλεύεται και τα χωρικά χαρακτηριστικά της λήψης. Οι υπόλοιπες μέθοδοι παρουσίασαν λιγότερο ακριβή και κυρίως λιγότερο σταθερά αποτελέσματα.

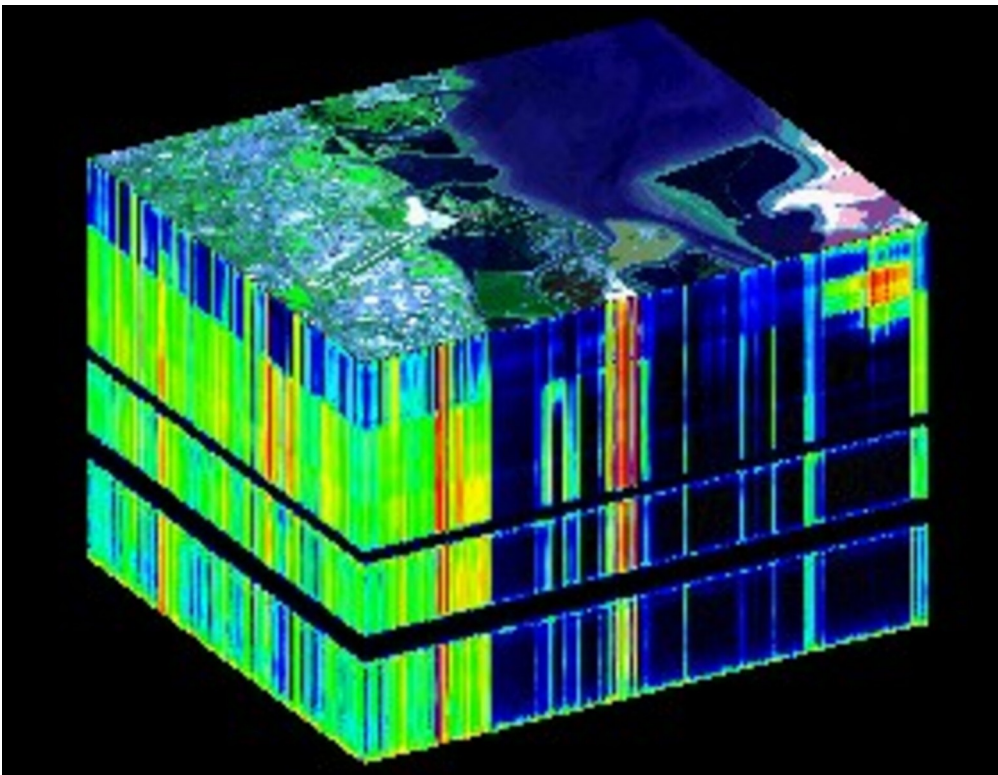
Abstract

Linear hyperspectral unmixing is the process of expressing the pixels of a hyperspectral image into endmembers and abundances while simultaneously extracting them. In this paper, spectral separation is performed with six different deep learning algorithms in order to benchmark them. It is shown that the architecture of a decoder, given a set of constraints, corresponds very well to the linear model and can reflect its physical and mathematical substance. A decoder is a fully connected neural network that uses unsupervised learning and encodes the input data into a lower dimensional hidden representation. Then it attempts to reconstruct the image by decoding this hidden representation. In this model the hidden representation is the abundance maps and the decoder weights are the endmembers.

Six deep learning methods were selected from the literature for their comparative evaluation [6]-[11]. The methods were compared using two different widely used images accompanied by extracted abundance maps and their corresponding endmembers. Ten experiments were performed for each method on each dataset. Of the methods, the most accurate were found to be the MTAEU[8] which incorporates the spatial information of the hyperspectral image into the model and CNNAEU[9] which is a convolutional decoder and therefore takes advantage of the spatial characteristics of the image. The other methods showed less accurate and especially less stable results.

1. Εισαγωγή

Η υπερφασματική τηλεπισκόπηση είναι ένα γρήγορα εξελισσόμενο πεδίο με πλήθος εφαρμογών. Κατά αυτή, ένας υπερφασματικός αισθητήρας είναι ευαίσθητος σε συνήθως εκατοντάδες πολύ κοντινά τμήματα του ηλεκτρομαγνητικού φάσματος με αποτέλεσμα την παραγωγή μιας πολύ φασματικά πυκνής λήψης που ονομάζεται υπερφασματικός κύβος. Αυτή η πολύ υψηλή φασματική ανάλυση είναι που επιτρέπει την λεπτομερή αναγνώριση στοιχείων του εδάφους σε εφαρμογές ταξινόμησης και κατάτμησης εικόνας με συχνά πεδία εφαρμογής την γεωλογία και την καλλιέργεια (smart agriculture). Όμως, η υψηλή φασματική ανάλυση περιορίζει την χωρική ανάλυση του δέκτη με αποτέλεσμα τα εικονοστοιχεία των λήψεων αυτών να αποτελούνται συνήθως από μια μίξη πολλών συστατικών. Έτσι απαιτείται ο διαχωρισμός των εικονοστοιχείων αυτών στα επιμέρους στοιχεία τους, μια διαδικασία υπολογιστικά βαριά και χρονοβόρα.



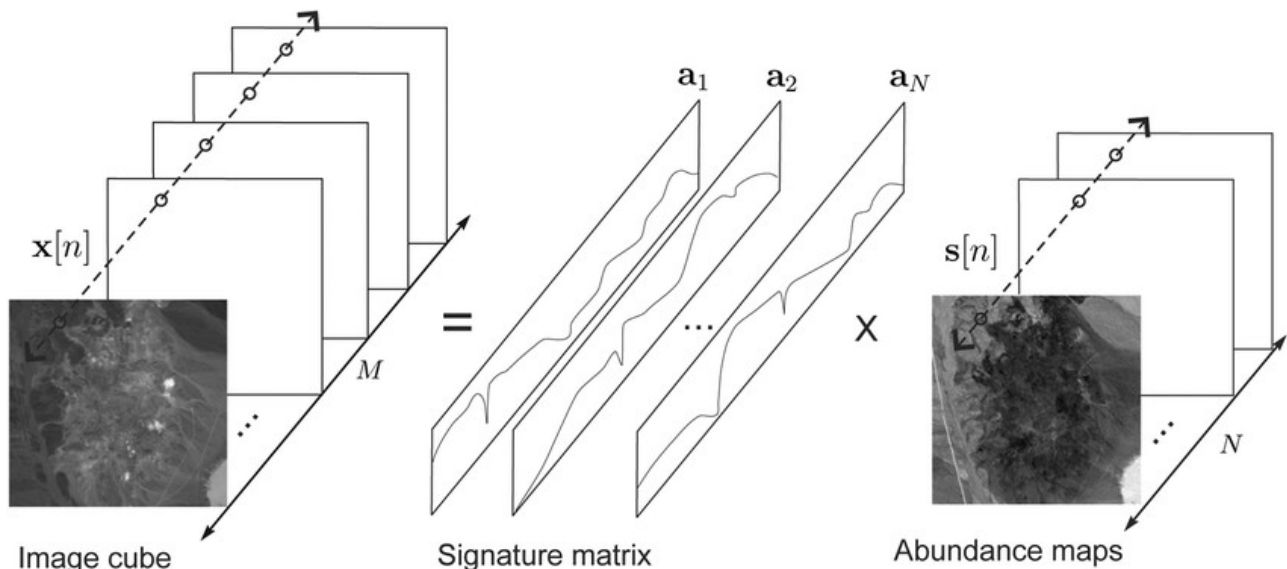
Σχήμα 1. Ο υπερφασματικός κύβος Πηγή: ESA

Οι μέθοδοι που χρησιμοποιούνται για την εκτέλεση της διαδικασίας αυτής μπορούν να χωριστούν σε γεωμετρικές που θεωρούν τους καθαρούς στόχους ως κορυφές ενός γεωμετρικού σχήματος που καλύπτει τα δεδομένα της εικόνας (N-FINDR, MVC-NMF, VCA), σε γραμμικές μεθόδους που αναπτύσσουν μια μαθηματική γραμμική σχέση και χρησιμοποιούν θεωρίες γραμμικής άλγεβρας (AMEE) και λοιπές μεθόδους που βασίζονται στην στατιστική ή που δεν μπορούν να καταταγούν σε κάποια από τις προηγούμενες κατηγορίες (IPP, ACA-AQA). Με την ανάπτυξη και την μεγάλη επιτυχία των μεθόδων βαθιάς μάθησης στην τηλεπισκόπηση τις τελευταίες δεκαετίες, έχουν αναπτυχθεί και αλγόριθμοι φασματικού διαχωρισμού βαθιάς μάθησης, ικανοί να εξαγάγουν τις αφθονίες και τους καθαρούς στόχους με ικανοποιητική ακρίβεια. Αυτές οι μέθοδοι είναι στην πλειοψηφία των περιπτώσεων υπολογιστικά ελαφριές με αποτέλεσμα να είναι γρηγορότερες των συμβατικών μεθόδων που αναφέρθηκαν ειδικά με την χρήση μονάδων επεξεργασίας GPU. Έτσι, θεωρήθηκε χρήσιμη η συγκριτική αξιολόγηση τους.

2.Φασματικός διαχωρισμός

2.1. Το γραμμικό μοντέλο

Όπως αναφέρθηκε, στις τηλεπισκοπικές παθητικές λήψεις, πολλά εικονοστοιχεία της σκηνής δεν αντιπροσωπεύουν ένα υλικό αλλά μία μίξη όλων των επιμέρους στοιχείων που βρίσκονται στο οπτικό πεδίο του δέκτη. Η μίξη αυτή μπορεί να συμβαίνει καθαρά λόγω του χωρικού περιορισμού του δέκτη ή ακόμα λόγω της φύσης της σκηνής όπου ο στόχος μπορεί να είναι μια ομοιογενής μίξη ξεχωριστών υλικών. Η φασματική υπογραφή των μεικτών αυτών εικονοστοιχείων είναι μια σύνθεση των υλικών που τα αποτελούν. Τα επιμέρους αυτά μακροσκοπικά υλικά ονομάζονται καθαροί στόχοι και τα ποσοστά κατά τα οποία συμμετέχουν στην σύνθεση της φασματικής υπογραφής αφθονίες. Έτσι, φασματικός διαχωρισμός είναι η διαδικασία ανάλυσης των μεικτών εικονοστοιχείων στους επιμέρους καθαρούς στόχους και τις αντίστοιχες αφθονίες που τα περιγράφουν. Η υψηλή φασματική ανάλυση των υπερφασματικών δεκτών επιτρέπει την ανάλυση αυτή. Γενικά, η διαδικασία του φασματικού διαχωρισμού αποτελείται από τρία βήματα: μείωση των διαστάσεων της εικόνας, εκτίμηση των καθαρών στόχων και υπολογισμός των αφθονιών.



Σχήμα 2. Γραμμικό μοντέλο φασματικού διαχωρισμού. Πηγή: [1]

Το κύριο μοντέλο που χρησιμοποιείται για την περιγραφή της ανάμειξης των υλικών σε μια σκηνή είναι το γραμμικό μοντέλο. Κατά αυτό η ένταση της ακτινοβολίας που καταγράφεται οφείλεται σε ένα μόνο υλικό και δεν επηρεάζεται από φαινόμενα πολλαπλής διάχυσης του φωτός μεταξύ στοιχείων. Έτσι, θεωρώντας πως η συνολική έκταση του οπτικού πεδίου έχει αναλογική σχέση με τις αφθονίες των καθαρών στόχων προκύπτει πως η φασματική υπογραφή ενός εικονοστοιχείου και οι καθαροί στόχοι συνδέονται με την γραμμική σχέση:

$$\mathbf{x} = \sum_{r=1}^R a_r s_r + \mathbf{n} \Rightarrow \mathbf{X} = \mathbf{S}\mathbf{A} + \mathbf{N}$$

όπου R ο αριθμός των καθαρών στόχων, \mathbf{X} πίνακας με στήλες τις φασματικές υπογραφές μεικτών εικονοστοιχείων μεγέθους $C \times P$ όπου C ο αριθμός των καναλιών και P ο αριθμός των εικονοστοιχείων, \mathbf{A} ο πίνακας μεγέθους $R \times P$ με στήλες τα διανύσματα με τις αφθονίες a , \mathbf{S} ο πίνακας μεγέθους $C \times R$ με στήλες τις φασματικές υπογραφές των καθαρών στόχων s και \mathbf{N} πρόσθετος θόρυβος.

Οι αφθονίες υπάγονται σε δύο περιορισμούς: πρέπει να είναι μη αρνητικές και το άθροισμά τους να ισούται με ένα καθώς αποτελούν ποσοστά συμμετοχής των καθαρών στόχων:

$$a_i \geq 0 \text{ και } s_i \geq 0, i=1, \dots, R$$

$$\sum_{i=1}^R a_i = 1$$

Φυσικά το γραμμικό μοντέλο δεν λειτουργεί εξίσου καλά με μη γραμμικές προσεγγίσεις σε περιπτώσεις ομοιογενούς μίξης υλικών. Σε αυτές η χωρική συσχέτιση των υλικών είναι σε κλίμακα μικρότερη του μήκους των φωτονίων με αποτέλεσμα το φως να επιδρά σε περισσότερα από ένα υλικά και το γραμμικό μοντέλο να χάνει την φυσική σημασία του. Παρόλα αυτά χρησιμοποιείται ευρέως λόγω της απλότητας του.

2.2. Φασματική γωνία και φασματική απόκλιση

Η φασματική γωνία (spectral angle distance - SAD) και η φασματική απόκλιση (spectral information divergence - SID) είναι δύο μέτρα ομοιότητας μεταξύ δύο φασματικών υπογραφών. Θεωρώντας $\mathbf{x}^{(i)}, \mathbf{x}^{(j)}$ δύο φασματικές υπογραφές, τα μέτρα ορίζονται ως:

$$SAD(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \cos^{-1}\left(\frac{\mathbf{x}^{(i)} \cdot \mathbf{x}^{(j)}}{\|\mathbf{x}^{(i)}\|_2 \|\mathbf{x}^{(j)}\|_2}\right), SAD \in [0, \frac{\pi}{2}]$$

$$SID(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \sum_{i=1}^C p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=1}^C q_i \log\left(\frac{q_i}{p_i}\right)$$

$$q_i = \frac{x_i^{(i)}}{\sum_{i=2}^C x_i^{(i)}}, p_i = \frac{x_i^{(j)}}{\sum_{i=2}^C x_i^{(j)}}$$

Η φασματική γωνία εκφράζει την γωνιακή απόσταση ανάμεσα στο δύο διανύσματα των φασματικών υπογραφών με μικρότερες τιμές να δηλώνουν μεγαλύτερη ομοιότητα και η φασματική απόκλιση την διαφορά ανάμεσα στις κατανομές πιθανότητας των δύο διανυσμάτων.

3. Νευρωνικά δίκτυα

Τα νευρωνικά δίκτυα αποτελούν ένα σύνολο μεθόδων μηχανικής μάθησης που προσπαθούν να μιμηθούν τον τρόπο με τον οποίο λειτουργεί ο ανθρώπινος εγκέφαλος. Υπό αυτή την φιλοσοφία, ένα νευρωνικό δίκτυο απαρτίζεται από ένα πλήθος βασικών υπολογιστικών στοιχείων που αλληλεπιδρούν μεταξύ τους και ονομάζονται νευρώνες. Οι νευρώνες συνδέονται μεταξύ τους μέσω βεβαρημένων, γραμμικών και μη, σχέσεων. Σύμφωνα με τους Aleksander και Morton:

Ένα νευρωνικό δίκτυο είναι ένας τεράστιος παράλληλος επεξεργαστής με κατανομημένη αρχιτεκτονική, ο οποίος αποτελείται από απλές μονάδες επεξεργασίας και έχει από την φύση του την δυνατότητα να αποθηκεύει εμπειρική γνώση και να την καθιστά διαθέσιμη για χρήση. Μοιάζει με τον ανθρώπινο εγκέφαλο σε δύο σημεία:

- Το δίκτυο προσλαμβάνει γνώση από το περιβάλλον του μέσω μιας διαδικασίας μάθησης.
- Η ισχύς των συνδέσεων μεταξύ των νευρώνων, που αποκαλείται συναπτικό βάρος, χρησιμοποιείται για την αποθήκευση της γνώσης που αποκτιέται.

3.1. Gradient descent

Οι περισσότεροι αλγόριθμοι μηχανικής μάθησης περιλαμβάνουν μια διαδικασία μάθησης. Αυτή η διαδικασία στα νευρωνικά δίκτυα είναι υπό την μορφή βελτιστοποίησης ενός συνόλου παραμέτρων που τα χαρακτηρίζουν. Η βελτιστοποίηση πραγματοποιείται ελαχιστοποιώντας μια συνάρτηση που ονομάζεται συνάρτηση κόστους. Αυτό που ενδιαφέρει είναι η εύρεση του ολικού ελάχιστου της συνάρτησης αυτής. Μια πολύ διαδεδομένη μέθοδος για την εύρεση του ελαχίστου είναι η μέθοδος gradient descent (κάθοδος με βάση την κλίση). Κατά αυτήν υπολογίζονται τα διαφορικά της συνάρτησης κόστους και οι μεταβλητές τροποποιούνται στην κατεύθυνση του χώρου που είναι αντίθετη στην μέγιστη κλίση. Μαθηματικά, η νέα τιμή μιας παραμέτρου προς βελτιστοποίηση εκφράζεται ως:

$$\dot{x} = x - \varepsilon \nabla_x f(x)$$

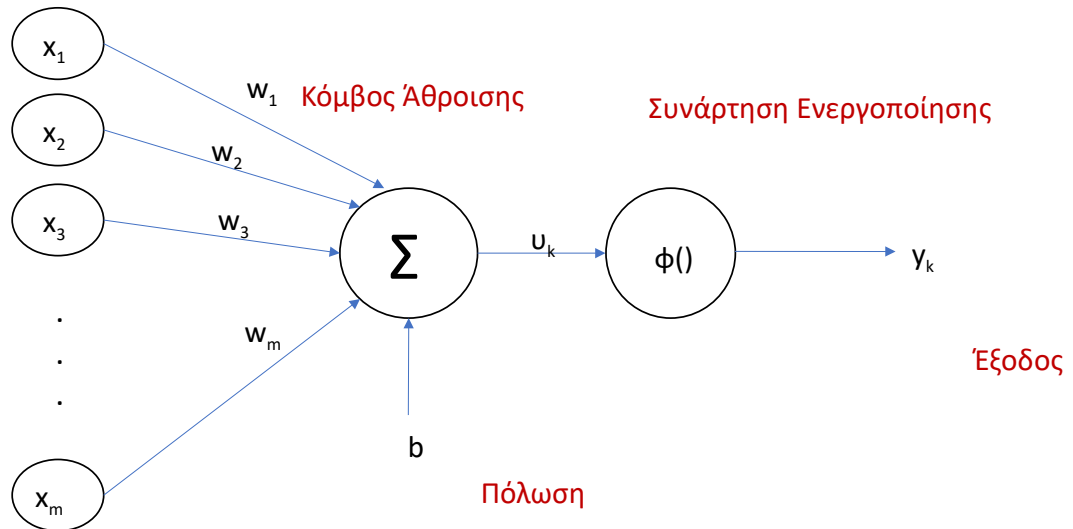
όπου ε ένας αριθμός που ονομάζεται ρυθμός μάθησης. Ο ρυθμός μάθησης ορίζει την ταχύτητα κατά την οποία θα συγκλίνει ο αλγόριθμος. Μεγάλες τιμές οδηγούν σε μεγάλες τροποποιήσεις στις παραμέτρους και σε γενικές γραμμές σε δυσχερέστερες λύσεις καθώς είναι πιθανότερο ο αλγόριθμος να παρουσιάσει παρεκκλίνουσα συμπεριφορά. Αντιθέτως μικρές τιμές οδηγούν σε αργή σύγκλιση και υπάρχει κίνδυνος υπερπροσαρμογής στα δεδομένα. Μια σωστή τιμή πρέπει να εξισορροπεί την ταχύτητα σύγκλισης με την ποιότητα του αποτελέσματος.

3.2. Ο νευρώνας

Ένας τυπικός νευρώνας δέχεται ως είσοδο ένα σήμα εισόδου M διαστάσεων, το οποίο μπορεί και να αποτελεί έξοδο προηγούμενων νευρώνων. Κάθε στοιχείο του σήματος εισόδου αθροίζεται σταθμισμένα με βάση τα συναπτικά βάρη w στον κόμβο άθροισης. Στο αποτέλεσμα του αθροίσματος προστίθεται η πόλωση b (bias), η οποία μεταθέτει το αποτέλεσμα κατά ένα μέγεθος για καλύτερη προσαρμογή στα δεδομένα. Το αποτέλεσμα του κόμβου άθροισης ονομάζεται δυναμικό ενεργοποίησης του νευρώνα το οποίο και περνάει μέσα από μια συνάρτηση ενεργοποίησης. Η διαδικασία αυτή περιγράφεται στο σχήμα 2 και στην παρακάτω εξίσωση.

$$y_k = \varphi(u_k) = \varphi\left(\sum_{i=1}^M (w_i x_i + b)\right) = \varphi(\mathbf{w} \cdot \mathbf{x} + b)$$

Σήμα εισόδου



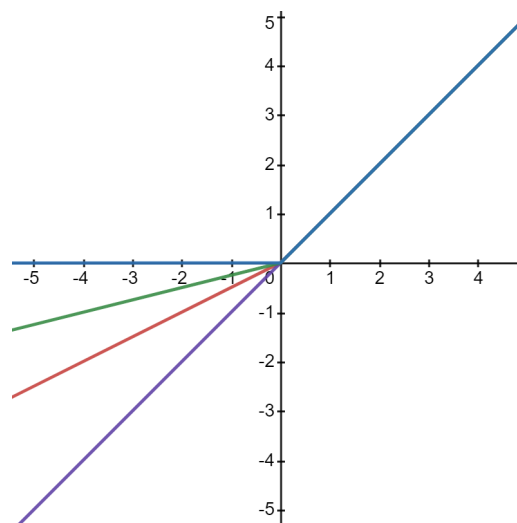
Σχήμα 3. Είσοδος και έξοδος ενός νευρώνα.

3.3. Συναρτήσεις ενεργοποίησης

Η συνάρτηση ενεργοποίησης έχει ως στόχο τον περιορισμό του εύρους του δυναμικού ενεργοποίησης, κυρίως στα εύρη $[-1,1]$, $[0,1]$ ή σε υποσύνολα αυτών. Μέσω αυτής εισάγεται και η μη γραμμικότητα στο δίκτυο και μάλιστα με ομοιόμορφο τρόπο. Όπως αναφέρθηκε, η μέθοδος gradient descent βασίζεται στον υπολογισμό των διαφορικών και συνεπώς η συνάρτηση ενεργοποίησης πρέπει να είναι διαφοροποιήσιμη σε όλο το πεδίο ορισμού της που συνήθως είναι το \mathbb{R} . Αυτό δεν συμβαίνει πάντα. Δύο πολύ χρήσιμες συναρτήσεις ενεργοποίησης για τον φασματικό διαχωρισμό είναι η ReLu και η softmax.

Η συνάρτηση ReLu δέχεται ως τιμές όλο το \mathbb{R} και μηδενίζει τις μη θετικές τιμές. Μάλιστα ο θετικός άξονας της συνάρτησης έχει κλίση 1. Μια γενικευμένη περίπτωση της ReLu είναι η LeakyReLu κατά την οποία δεν μηδενίζονται πλέον οι αρνητικές τιμές αλλά ορίζονται και αυτές από άξονα με κλίση α που είναι εκπαιδευσιμη παράμετρος. Το αρνητικό με τις εξισώσεις αυτές είναι πως δεν είναι διαφοροποιήσιμες στο 0. Παρόλα αυτά η βιβλιογραφία δείχνει πως παρουσιάζουν τα καλύτερα αποτελέσματα.

$$LeakyReLU(x) = \max(ax, x), a \in [0,1]$$



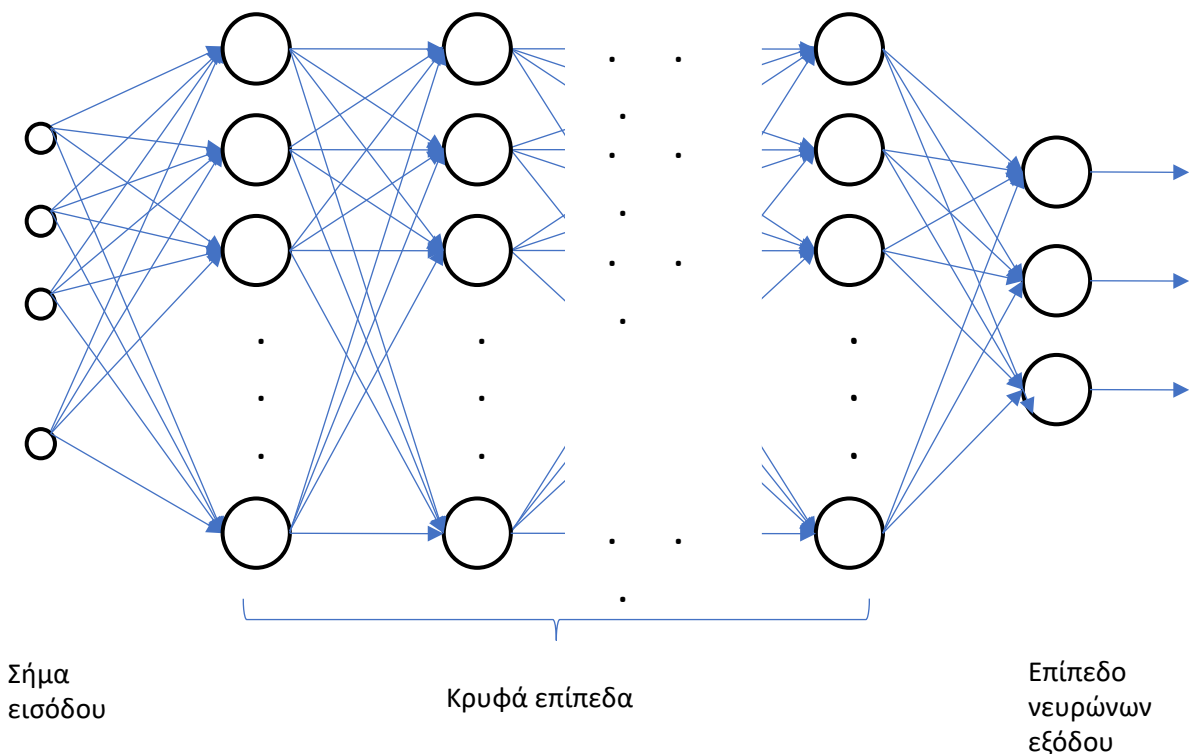
Σχήμα 4. Η συνάρτηση LeakyReLU για διάφορες τιμές του α . Για $\alpha=0$ προκύπτει η απλή ReLu (μπλε γραμμή)

Η συνάρτηση softmax έχει πεδίο ορισμού το \mathbb{R} και πεδίο τιμών το διάστημα $[0,1]$. Η συνάρτηση αυτή είναι πολυπαραμετρική και δέχεται όλες τις τιμές του προηγούμενου επιπέδου του δικτύου. Έχει την πολύ χρήσιμη ιδιότητα να κανονικοποιεί τα δυναμικά ενεργοποίησης ενός επιπέδου με τέτοιο τρόπο ώστε το άθροισμά τους να ισούται με 1. Επομένως μπορεί να χρησιμοποιηθεί για την εφαρμογή των περιορισμών του γραμμικού μοντέλου του φασματικού διαχωρισμού.

$$\text{Softmax}(x) = \frac{e^{x_i}}{\sum_{j=1}^k e^{x_j}}$$

3.4. Πολυεπίπεδα δίκτυα πρόσθιας τροφοδότησης

Όπως αναφέρθηκε, ένα νευρωνικό δίκτυο αποτελείται από ένα πλήθος νευρώνων οργανωμένων σε επίπεδα. Ένα απλό ρηχό δίκτυο έχει τρία επίπεδα: το επίπεδο εισόδου που είναι ένα διάνυσμα με τα δεδομένα εισόδου, ένα κρυφό επίπεδο που περιέχει μια σειρά από νευρώνες και ένα επίπεδο εξόδου. Κάθε νευρώνας του κρυφού επιπέδου είναι συνδεδεμένος με συναπτικά βάρη με όλες τις τιμές του προηγούμενου επιπέδου. Τα βαθιά δίκτυα έχουν πολλά κρυφά επίπεδα. Η μορφή ενός τέτοιου δικτύου φαίνεται στο σχήμα 5.



Σχήμα 5. Ένα πολυεπίπεδο δίκτυο πρόσθιας τροφοδότησης.

3.5. Οπισθοδιάδοση

Κατά την πρόσθια τροφοδότηση το δίκτυο δεν τροποποιεί τις τιμές των βαρών του. Αυτή η ενημέρωση γίνεται κατά το πέρασμα προς τα πίσω όπου με βάση τις τιμές τις εξόδου υπολογίζονται σταδιακά οι τροποποιήσεις στις τιμές των εκπαιδευσιμων παραμέτρων του τελευταίου επιπέδου, έπειτα του προτελευταίου και ούτω καθεξής. Με βάση την εξίσωση που περιγράφει έναν νευρώνα η και τον ορισμό της μεθόδου gradient descent προκύπτει μετά από συνεχείς διαφοροποιήσεις πως:

$$\begin{aligned}\Delta w_{ji} &= \epsilon \delta_j(n) y_i(n) \\ \delta_j(n) &= e_j(n) \phi_j'(v_j(n))\end{aligned}$$

Δηλαδή η αλλαγή στην τιμή που εφαρμόζεται στο συναπτικό βάρος είναι το γινόμενο του ρυθμού μάθησης ϵ , τις τιμές εξόδου του προηγούμενου νευρώνα $y_i(n)$ και του όρου $\delta_j(n)$ που ονομάζεται τοπική κλίση. Η τοπική κλίση είναι το γινόμενο της παραγώγου της συνάρτησης ενεργοποίησης του νευρώνα και του σήματος σφάλματος του νευρώνα $e_j(n)$ δηλαδή ενός μέτρου διαφοράς ανάμεσα στις προβλεπόμενες και τις πραγματικές τιμές του εξεταζόμενου φαινομένου. Η προηγούμενη εξίσωση τοπικής κλίσης ισχύει μόνο για το επίπεδο εξόδου. Για τα κρυφά επίπεδα αποδεικνύεται πως:

$$\delta_j(n) = \phi_j'(v_j(n)) \sum_k \delta_k(n) w_{kj}(n)$$

Δηλαδή, για έναν νευρώνα κρυφού επιπέδου η τοπική κλίση ισούται με την παράγωγο της συνάρτησης ενεργοποίησης του νευρώνα και το σταθμισμένο άθροισμα των ήδη υπολογισμένων τοπικών κλίσεων των νευρώνων του επόμενου επιπέδου. Τα βάρη του σταθμισμένου αθροίσματος είναι τα αντίστοιχα συναπτικά βάρη.

3.6. Batch Normalization

Ένα επίπεδο Batch Normalization εφαρμόζει μια κανονικοποίηση στα δυναμικά ενεργοποίησης του προηγούμενου επιπέδου, με αποτέλεσμα την πιο γρήγορη εκπαίδευση του δικτύου. Η κανονικοποίηση αυτή γίνεται με βάση την μέση τιμή και την τυπική απόκλιση των δυναμικών και τα αναγκάζει να έχουν μέση τιμή μηδέν. Στην συνέχεια, εφαρμόζεται αφινικός μετασχηματισμός με εκπαιδευσιμες παραμέτρους γ και β . Δηλαδή:

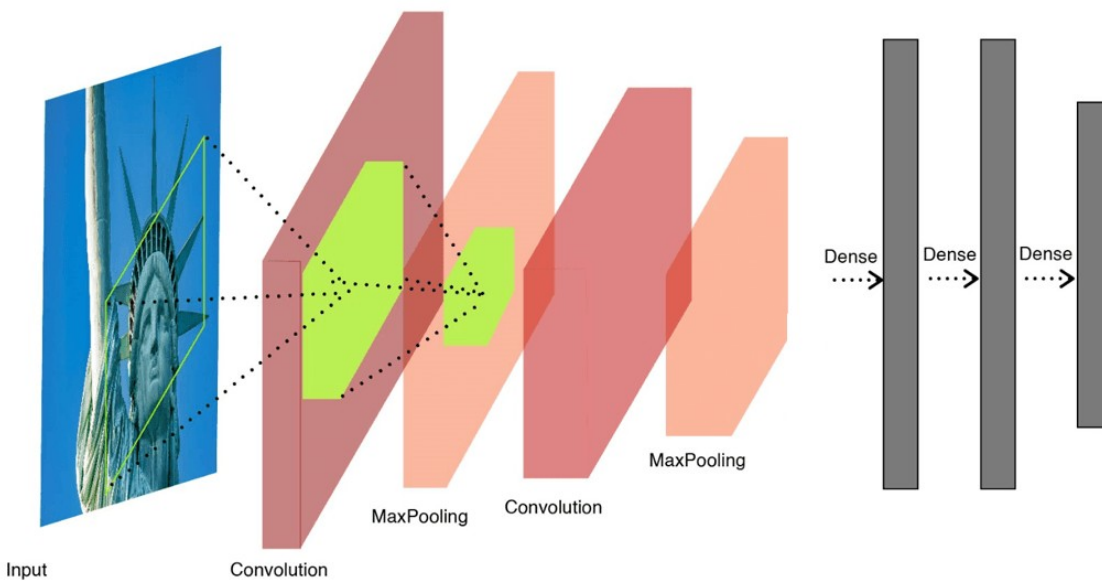
$$\begin{aligned}v^{(i)} &= \gamma v_{norm}^{(i-1)} + \beta \\ v_{norm}^{(i-1)} &= \frac{v^{(i-1)} - \mu}{\sqrt{\sigma^2 + \epsilon}}\end{aligned}$$

όπου ϵ ένας πολύ μικρός αριθμός.

3.7. Συνελικτικά δίκτυα

Τα συνελικτικά δίκτυα είναι δίκτυα που χρησιμοποιούν την γραμμική πράξη της συνέλιξης στο εσωτερικό του νευρώνα και χρησιμοποιούνται κυρίως για την επεξεργασία εικόνων ή χρονοσειρών. Έτσι, πλέον ο νευρώνας ενός επιπέδου δεν εξαρτάται από όλους του προηγούμενου άλλα από f^2 νευρώνων όπου f το μέγεθος του παραθύρου συνέλιξης. Αυτό σε συνδυασμό με το γεγονός πως οι τιμές του παραθύρου συνέλιξης είναι ίδιες για όλα δεδομένα, σημαίνει πως ένα συνελικτικό δίκτυο έχει πολύ λιγότερες παραμέτρους προς εκπαίδευση. Με λίγα λόγια ένα τέτοιο δίκτυο έχει αραιές συνδέσεις και διαμοιράζει τις παραμέτρους του. Λόγω της φύσης της πράξης της συνέλιξης το δίκτυο είναι πλέον αναλλοίωτο στους χωρικούς μετασχηματισμούς. Τα συνελικτικά επίπεδα αποτελούνται από τρία είδη επιπέδων: επίπεδα συνέλιξης, επίπεδα pooling και από πλήρως διασυνδεδεμένα επίπεδα.

Τα επίπεδα συνέλιξης είναι επίπεδα που εφαρμόζουν την πράξη της συνέλιξης στην εικόνα εισόδου, προσθέτουν την πόλωση και έπειτα υπολογίζουν το δυναμικό ενεργοποίησης μέσω μιας συνάρτησης ενεργοποίησης. Με τον όρο συνέλιξη εδώ εννοείται η μαθηματική πράξη που ονομάζεται cross-correlation. Κατά την εφαρμογή της πράξης αυτής η εικόνα χάνει διαστάσεις στα άκρα της και έτσι πολλές φορές προστίθενται στα όρια μηδενικά εικονοστοιχεία με τιμή 0 (padding). Τα επίπεδα pooling είναι επίπεδα που έχουν ως στόχο την μείωση του όγκου της πληροφορίας εφαρμόζοντας ένα φίλτρο στην εικόνα εισόδου. Μερικά συνηθισμένα φίλτρα είναι το μέγιστο, το ελάχιστο και το μέσο και αντίστοιχα τα επίπεδα ονομάζονται max pooling, min pooling και average pooling. Τα επίπεδα αυτά δεν έχουν παραμέτρους προς εκπαίδευση. Τέλος τα πλήρως διασυνδεδεμένα επίπεδα είναι επίπεδα σαν αυτά του σχήματος 5. Συνήθως τοποθετούνται στο τέλος του δικτύου για την εξαγωγή της πρόβλεψης αφού η έξοδος του προηγούμενου επιπέδου συνέλιξης ή pooling μετατραπεί σε μονοδιάστατο διάνυσμα. Στο σχήμα 6 φαίνεται ένα τυπικό συνελικτικό δίκτυο.



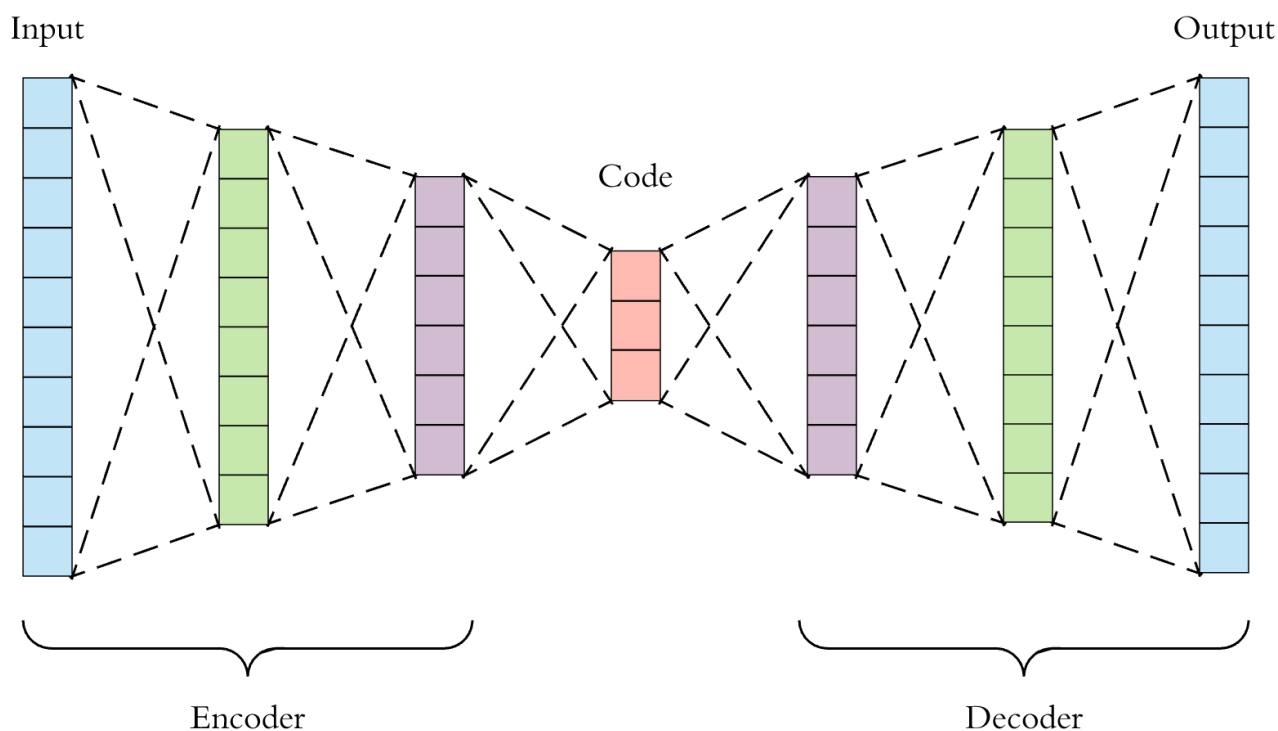
Σχήμα 6. Ένα τυπικό συνελικτικό δίκτυο.

3.8. Αυτοκωδικοποιητές

Ο αυτοκωδικοποιητής είναι ένα νευρωνικό δίκτυο, αποτελούμενο από δύο μέρη: τον κωδικοποιητή (encoder) και τον αποκωδικοποιητή (decoder). Ο κωδικοποιητής δέχεται τα αρχικά δεδομένα και τα μετατρέπει στην κρυφή αναπαράσταση (code) $h = E(x)$ και ο αποκωδικοποιητής δέχεται την κρυφή αναπαράσταση και προσπαθεί να ανακατασκευάσει τα αρχικά δεδομένα $\hat{x} = D(h)$. Δηλαδή:

$$\hat{x} = D(E(x))$$

Μέσω αυτής της διαδικασίας το δίκτυο μαθαίνει σημαντικές και χρήσιμες ιδιότητες των δεδομένων. Έτσι, αυτό που ενδιαφέρει στην πραγματικότητα είναι η κρυφή αναπαράσταση h και όχι η τελική τέλεια αντιγραφή των αρχικών δεδομένων. Για αυτό τον λόγο οι αυτοκωδικοποιητές εκπαιδεύονται να κάνουν αυτή την αντιγραφή με περιορισμούς. Ένας πολύ συνηθισμένος περιορισμός είναι η κρυφή αναπαράσταση να είναι ένας χώρος μικρότερων διαστάσεων από τον χώρο των αρχικών δεδομένων (undercomplete autoencoders). Με αυτό τον τρόπο το δίκτυο δυσκολεύεται να ανακατασκευάσει ακριβώς τα αρχικά δεδομένα με αποτέλεσμα να αναγκάζεται να μάθει σημαντικές ιδιότητες που τα χαρακτηρίζουν. Το γεγονός αυτό καθιστά τα μοντέλα αυτά πολύ ικανά στην μείωση των διαστάσεων των αρχικών δεδομένων (dimensionality reduction). Έτσι, αφού ο φασματικός διαχωρισμός μπορεί να θεωρηθεί ως μια διαδικασία εύρεσης μιας αναπαράστασης χαμηλότερων διαστάσεων, τα δίκτυα αυτά φαίνονται ιδανικά για την πραγματοποίησή του. Στο παρακάτω σχήμα (Σχήμα 7) φαίνεται ένας τυπικός αυτοκωδικοποιητής.



Σχήμα 7. Ένας τυπικός αυτοκωδικοποιητής (undercomplete autoencoder).

4. Βιβλιογραφική ανασκόπηση

Την τελευταία δεκαετία, με την εξέλιξη της υπολογιστικής ισχύς, ο τομέας της βαθειάς μάθησης έχει υποστεί ραγδαία ανάπτυξη. Έχει αποδειχθεί πως τα νευρωνικά δίκτυα είναι ικανά να εκτελέσουν ένα μεγάλο εύρος εργασιών, στις οποίες περιλαμβάνεται και ο φασματικός διαχωρισμός υπερφασματικών εικόνων. Βέβαια η ικανότητα των νευρωνικών δικτύων να εξαγάουν χάρτες αφθονιών και καθαρούς στόχους από υπερφασματικές λήψεις δεν έχει εξετασθεί στον ίδιο βαθμό, όπως για παράδειγμα έχει εξετασθεί η ικανότητα για ταξινόμηση κάλυψης γης. Παρόλα αυτά, η βιβλιογραφία δείχνει πως έχουν πραγματοποιηθεί αρκετές και επιτυχημένες προσπάθειες. Σε αυτές χρησιμοποιούνται αποκωδικοποιητές στους οποίους τα βάρη του αποκωδικοποιητή αποτελούν τους καθαρούς στόχους και η κρυφή αναπαράσταση τις αφθονίες.

Στο άρθρο [6] οι συγγραφείς προτείνουν έναν πλήρως διασυνδεδεμένο αποκωδικοποιητή για την εκτέλεση του φασματικού διαχωρισμού. Καταλήγουν στην φασματική γωνία ως πιο αποδοτική συνάρτηση κόστους και δείχνουν πως επειδή ο αποκωδικοποιητής πρέπει να είναι ρηχός, ένας βαθύς κωδικοποιητής δεν προσφέρει απαραίτητα καλύτερα αποτελέσματα. Το δίκτυο αυτό φαίνεται να δίνει καλύτερα αποτελέσματα από μεθόδους όπως η VCA και η NMF.

Στο άρθρο [7] ελέγχεται η φασματική απόκλιση ως συνάρτηση κόστους σε πλήρως διασυνδεδεμένο αποκωδικοποιητή αντί για το μέσο τετραγωνικό σφάλμα που χρησιμοποιείται ευρέως σε τέτοια μοντέλα. Αποδεικνύεται πως η φασματική απόκλιση αποτελεί καλύτερο μέτρο προς ελαχιστοποίηση.

Στο άρθρο [8] γίνεται μια απόπειρα να αξιοποιηθεί και η χωρική διάσταση της εικόνας με βάση το μοντέλο [6]. Σε αυτό εκτελείται φασματικός διαχωρισμός ταυτόχρονα σε όλα τα εικονοστοιχεία μια γειτονιάς, με τα παράλληλα δίκτυα να μοιράζονται τις παραμέτρους ορισμένων επιπέδων. Η μέθοδος αυτή δίνει καλύτερα αποτελέσματα από κλασσικές μεθόδους όπως η VCA και η NMF άλλα και από τις μεθόδους βαθειάς μάθησης [6] και [14]. Βέλτιστα αποτελέσματα έδωσε η χρήση γειτονιάς 3x3 εικονοστοιχείων.

Στο άρθρο [9] χρησιμοποιείται ένας συνελκτικός αυτοκωδικοποιητής αξιοποιώντας με αυτό τον τρόπο καλύτερα την χωρική πληροφορία της εικόνας ταυτόχρονα με την φασματική. Το δίκτυο αυτό χρησιμοποιεί μια τροποποίηση του γραμμικού μοντέλου που συμπεριλαμβάνει την συμβολή των γειτονικών εικονοστοιχείων ενός μεικτού στην φασματική υπογραφή του. Το δίκτυο συγκρίνεται με άλλα βαθειάς μάθησης ([6], [8] και [12]) και φαίνεται πως δίνει τα καλύτερα αποτελέσματα.

Στο άρθρο [10] προτείνεται ένας πλήρως διασυνδεδεμένος αυτοκωδικοποιητής που χρησιμοποιεί την φασματική γωνία αντί για το εσωτερικό γινόμενο στο εσωτερικό του νευρώνα άλλα και μια καινούρια συνάρτηση κόστους που λαμβάνει υπόψιν της το sparsity των χαρτών αφθονίας. Επίσης προτείνεται η μέθοδος VCA για την αρχικοποίηση των τιμών του βαρών του αποκωδικοποιητή. Η μέθοδος αυτή συγκρίνεται με μεθόδους όπως η VCA, NMF και DmaxD και δίνει καλύτερα αποτελέσματα και από τις τρεις.

Στο άρθρο [11] προτείνεται ένας πλήρως διασυνδεδεμένος αυτοκωδικοποιητής με τρεις καινοτομίες. Αρχικά προτείνεται η έννοια της ορθογωνικότητας των χαρτών αφθονίας που εκφράζει την ιδιότητα sparsity και η λαπλασιανή κατανομή ως την κατανομή που ακολουθεί το σφάλμα ανακατασκευής της εικόνας. Επιπλέον πραγματοποιείται αρχικοποίηση των βαρών του αυτοκωδικοποιητή εφαρμόζοντας ένα κατώφλι σε έναν χάρτη ομοιότητας που παράγεται από τους χάρτες αφθονίας. Τιμές μικρότερες του κατωφλίου απορρίπτονται και έπειτα χρησιμοποιείται η μέθοδος DmaxD για μια εκτίμηση των καθαρών στόχων. Η μέθοδος συγκρίνεται με μεθόδους όπως η VCA, DmaxD και ορισμένων αποκωδικοποιητών και παρουσιάζει καλύτερα αποτελέσματα.

5. Μέθοδοι

Θεωρώντας έναν αυτοκωδικοποιητή με ρηχό κωδικοποιητή χωρίς πολώσεις, ρηχό γραμμικό αποκωδικοποιητή χωρίς πολώσεις και είσοδο στο δίκτυο έναν πίνακα $\mathbf{X} \in \mathbb{R}^{C \times P}$ προκύπτει πως:

$$\begin{aligned}\mathbf{H} &= E(\mathbf{X}) = \varphi(\mathbf{W}_1 \mathbf{X}) \\ \hat{\mathbf{X}} &= D(\mathbf{H}) = \mathbf{W}_2 \mathbf{H}\end{aligned}$$

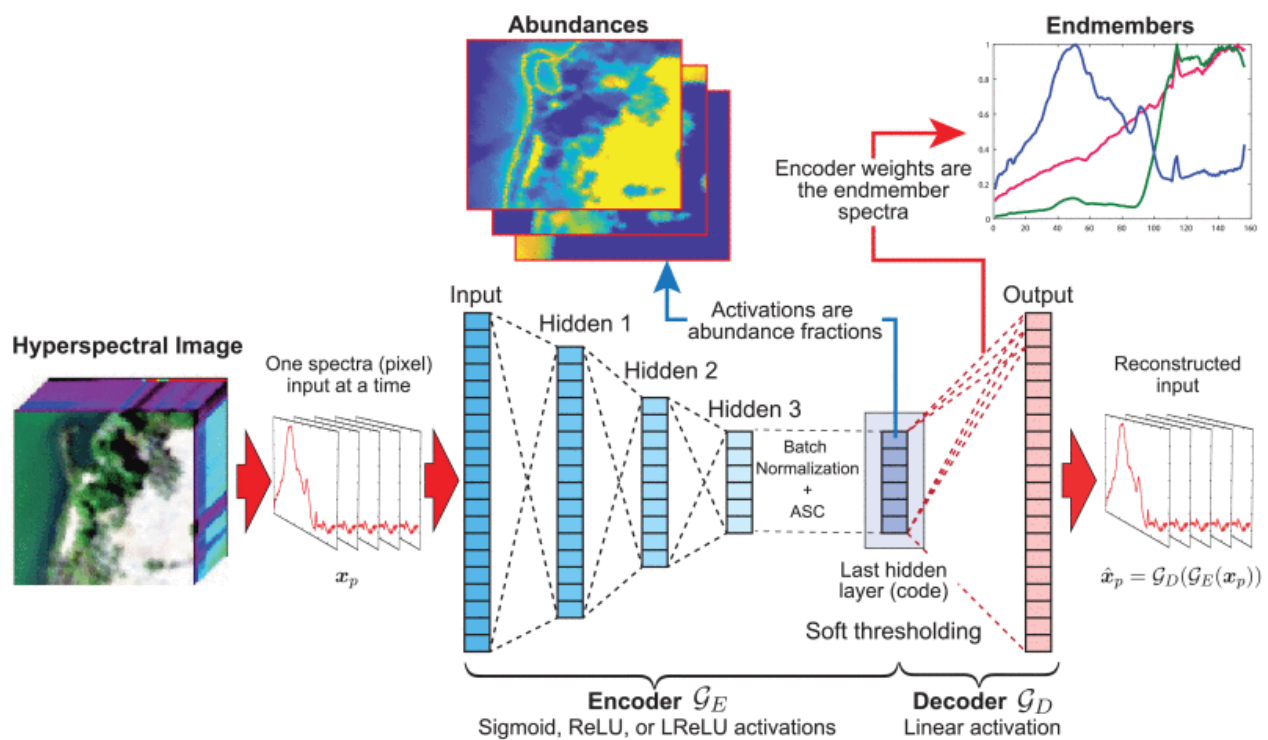
όπου φ η συνάρτηση ενεργοποίησης, \mathbf{W} τα βάρη των επιπέδων του δικτύου και \mathbf{X} η είσοδος του δικτύου που εδώ αποτελεί πίνακα με στήλες φασματικές υπογραφές εικονοστοιχείων. Επιπλέον, αν ορίσουμε την κρυφή αναπαράσταση με αριθμό διαστάσεων (δηλαδή αριθμό νευρώνων) ίσο με τον αριθμό των καθαρών στόχων προκύπτει πως

$$\mathbf{H} \in \mathbb{R}^{R \times P} \text{ και } \mathbf{W}_2 \in \mathbb{R}^{C \times R}$$

Από την εξίσωση του γραμμικού μοντέλου παρατηρείται πως η ανακατασκευή της εικόνας από την αναπαράσταση ταυτίζεται με το γραμμικό μοντέλο με $\mathbf{W}_2 = \mathbf{S}$ και $\mathbf{H} = \mathbf{A}$. Δηλαδή η κρυφή αναπαράσταση αποτελεί τις αφθονίες και τα βάρη του αποκωδικοποιητή αποτελούν τους καθαρούς στόχους. Αυτό σημαίνει πως πλέον το δίκτυο δεν είναι γενικεύσιμο σε ένα πληθυσμό καθώς η αρχιτεκτονική του και μέρος των βαρών του εξαρτάται άμεσα από την φύση της σκηνής ενδιαφέροντος και επομένως πρέπει να εκπαιδεύεται ξεχωριστά για κάθε σκηνή. Τέλος, το μοντέλο αυτό, έτσι όπως είναι διατυπωμένο, για να είναι απολύτως ικανό να εκτελέσει φασματικό διαχωρισμό, πρέπει πρώτα να τροποποιηθεί ώστε να υπάγεται στους περιορισμούς του γραμμικού μοντέλου, κάτι που γίνεται με διάφορες μεθόδους στην βιβλιογραφία όπως για παράδειγμα με μια απλή κανονικοποίηση ή με χρήση της συνάρτησης ενεργοποίησης softmax αμέσως πριν την κρυφή αναπαράσταση.

5.1 DAEU

Η μέθοδος αυτή υιοθετεί έναν αποκωδικοποιητή με βαθύ κωδικοποιητή και ρηχό αποκωδικοποιητή. Είναι ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο όπου ως είσοδο δέχεται την φασματική υπογραφή ενός εικονοστοιχείου υπερφασματικής εικόνας, δηλαδή το επίπεδο εισόδου έχει αριθμό νευρώνων όσα και τα κανάλια της εικόνας. Ο κωδικοποιητής αποτελείται από τέσσερα κρυφά πλήρως διασυνδεδεμένα επίπεδα με συνάρτηση ενεργοποίησης ReLU τα οποία καταλήγουν, με σταδιακή μείωση του αριθμού των νευρώνων, σε αριθμό νευρώνων ίσο με τον αριθμό των καθαρών στόχων. Στη συνέχεια ακολουθεί ένα επίπεδο Batch Normalization τα αποτελέσματα του οποίου περνάνε από μια δυναμική συνάρτηση ενεργοποίησης ReLU, εφαρμόζονται οι περιορισμοί του γραμμικού μοντέλου και τέλος ένα επίπεδο Gaussian Dropout. Τα αποτελέσματα του τελευταίου αποτελούν και την κρυφή αναπαράσταση h , δηλαδή τις αφθονίες. Ο αποκωδικοποιητής είναι ρηχός και γραμμικός. Τα βάρη του όπως αποδείχθηκε είναι οι καθαροί στόχοι. Τα επίπεδα φαίνονται αναλυτικά στον παρακάτω πίνακα και σχήμα.



Σχήμα 8. Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου DAEU. Source: [6]

Πίνακας 1. Τα επίπεδα του δικτύου DAEU.

A/α	Επίπεδο	Αριθμός νευρώνων
1	Είσοδος	C
2	Πλήρως διασυνδεδεμένο	9R
3	Πλήρως διασυνδεδεμένο	6R
4	Πλήρως διασυνδεδεμένο	3R
5	Πλήρως διασυνδεδεμένο	R
6	Batch Normalization	R
7	Δυναμική συνάρτηση ReLu	R
8	Κανονικοποίηση	R
9	Gaussian Dropout	R
10	Πλήρως διασυνδεδεμένο	C
11	Έξοδος	C

Σε μια υπερφασματική εικόνα παρατηρείται πως οι αφθονίες ενός εικονοστοιχείου έχουν υψηλή τιμή για συνήθως ένα ή δύο καθαρούς στόχους και σχεδόν μηδενική για όλους τους υπόλοιπους, μια ιδιότητα που ονομάζεται sparsity. Αυτήν ακριβώς την ιδιότητα προσπαθεί να εφαρμόσει και η δυναμική συνάρτηση ενεργοποίησης ReLu στην κρυφή αναπαράσταση h. Αυτό το επιτυγχάνει εφαρμόζοντας ένα κατώφλι σε μια κλασική συνάρτηση ενεργοποίησης ReLu:

$$v^{(i)} = \max(0, v^{(i-1)} - a)$$

όπου α εκπαιδεύσιμη παράμετρος. Με λίγα λόγια η συνάρτηση αυτή μηδενίζει τις αρνητικές τιμές (όπως και μια κλασική συνάρτηση ReLu) καθώς και τις τιμές που είναι μικρότερες από ένα κατώφλι. Οι περιορισμοί που προϋποθέτει το γραμμικό μοντέλο εφαρμόζονται με την παρακάτω κανονικοποίηση:

$$v^{(i)} = \frac{v^{(i-1)}}{\sum_{j=1}^R v_j^{(i-1)}}$$

Τέλος το επίπεδο Gaussian Dropout εισάγει λευκό γκαουσιανό θόρυβο στα δυναμικά ενεργοποίησης του προηγούμενου επιπέδου. Αυτό έχει αποδειχθεί πως εμποδίζει το overfitting. Ως συνάρτηση κόστους χρησιμοποιείται η φασματική γωνιακή απόσταση (Spectral Angle Distance).

5.2. SIDAEU

Το δίκτυο αυτό είναι παρόμοιος λογικής με DAEU μόνο που έχει ρηχό κωδικοποιητή και υιοθετεί την Spectral Information Divergence (SID) ως συνάρτηση κόστους. Συγκεκριμένα το δίκτυο αποτελείται από έξι επίπεδα: το επίπεδο εισόδου, ένα γραμμικό επίπεδο μεγέθους R, ένα επίπεδο batch normalization, την δυναμική συνάρτηση ενεργοποίησης ReLu της προηγούμενης μεθόδου, την κανονικοποίηση της προηγούμενης μεθόδου για την εφαρμογή των περιορισμών του γραμμικού μοντέλου και τέλος ένα πλήρως διασυνδεδεμένο γραμμικό επίπεδο μεγέθους M που ανακατασκευάζει την εικόνα με γραμμική συνάρτηση ενεργοποίησης. Τα επίπεδα φαίνονται στον πίνακα 2.

Πίνακας 2. Τα επίπεδα του δικτύου SIDAEU.

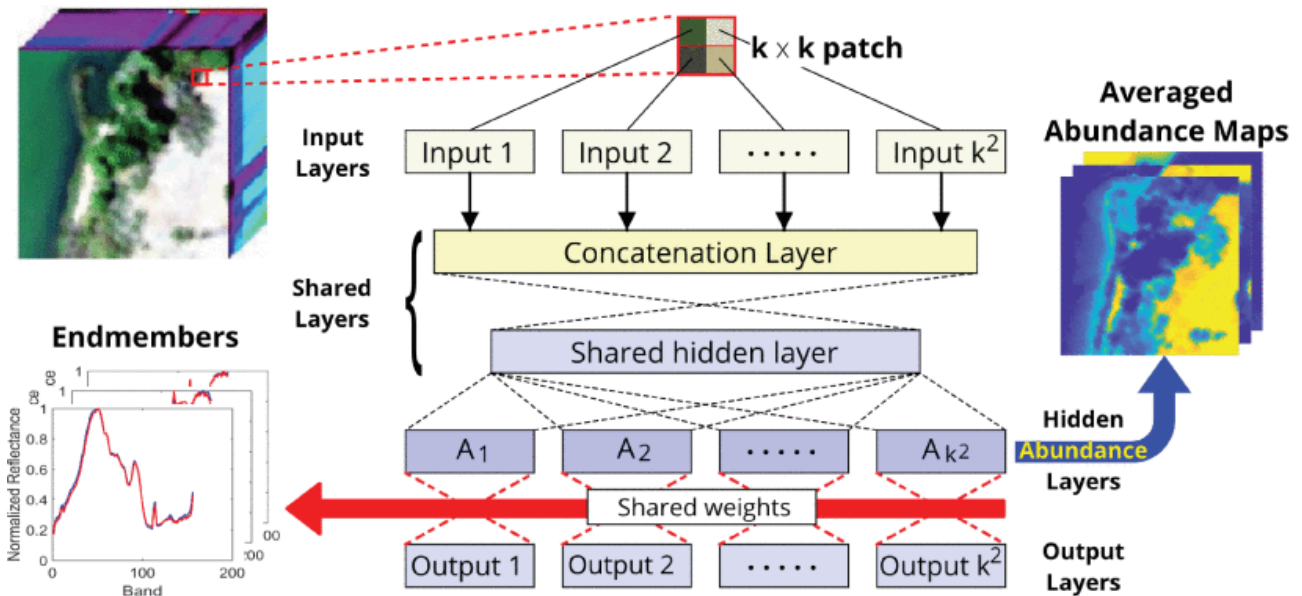
A/α	Επίπεδο	Αριθμός νευρώνων
1	Είσοδος	C
2	Πλήρως διασυνδεδεμένο	R
3	Batch normalization	R
4	Δυναμική συνάρτηση ReLu	R
5	Κανονικοποίηση	R
6	Πλήρως διασυνδεδεμένο	R
7	Έξοδος	C

5.3. MTAEU

Οι προηγούμενες μέθοδοι εκμεταλλεύονται μόνο την φασματική πληροφορία της εικόνας και αγνοούν την χωρική συσχέτιση των εικονοστοιχείων. Η μέθοδος αυτή αποπειράται να εισάγει και την χωρική διάσταση σε ένα πλήρως διασυνδεδεμένο δίκτυο που βασίζεται στο δίκτυο DAEU. Το δίκτυο αυτό εμπνέεται από την φιλοσοφία μεθόδων multitask learning.

Θεωρώντας μια γειτονιά k^2 εικονοστοιχείων, εκτελείται φασματικός διαχωρισμός με το ίδιο δίκτυο σε όλα τα εικονοστοιχεία της γειτονιάς. Η χωρική συσχέτιση εισάγεται στο δίκτυο υπό την μορφή μοιρασμένων βαρών σε ένα επίπεδο το οποίο μοιράζονται όλα τα παράλληλα δίκτυα. Αρχικά ενώνονται όλες οι εισοδοί των δικτύων και έπειτα εισέρχονται σε ένα κοινό πλήρως διασυνδεδεμένο επίπεδο με μικρότερη διάσταση και κοινά βάρη. Κατόπιν τα δίκτυα διαχωρίζονται με το καθένα να περνάει από ένα

επίπεδο Dropout με τιμή 0.4, Batch Normalization, ένα πλήρως διασυνδεδεμένο δίκτυο με αριθμό νευρώνων R, ένα ακόμα επίπεδο Batch normalization και τέλος από την συνάρτηση Softmax που εφαρμόζει τους περιορισμούς του γραμμικού μοντέλου. Το αποτέλεσμα είναι και το διάνυσμα αφθονίας, ξεχωριστό για κάθε εικονοστοιχείο της γειτονιάς, αλλά παραγόμενο με την χωρική πληροφορία της επιλεγόμενης γειτονιάς. Ακολουθεί ο αποκωδικοποιητής ο οποίος είναι ρηχός και γραμμικός. Αφού τα βάρη του αποκωδικοποιητή είναι οι καθαροί στόχοι πρέπει να είναι τα ίδια για όλα τα δίκτυα. Η αρχιτεκτονική του δικτύου φαίνεται στο σχήμα 9. και τον πίνακα 3. Ως συνάρτηση ενεργοποίησης για τα πλήρως συνδεδεμένα επίπεδα χρησιμοποιείται η LeakyRelu και ως συνάρτηση κόστους χρησιμοποιείται η μέση τιμή της φασματικής απόστασης για κάθε παράλληλο δίκτυο.



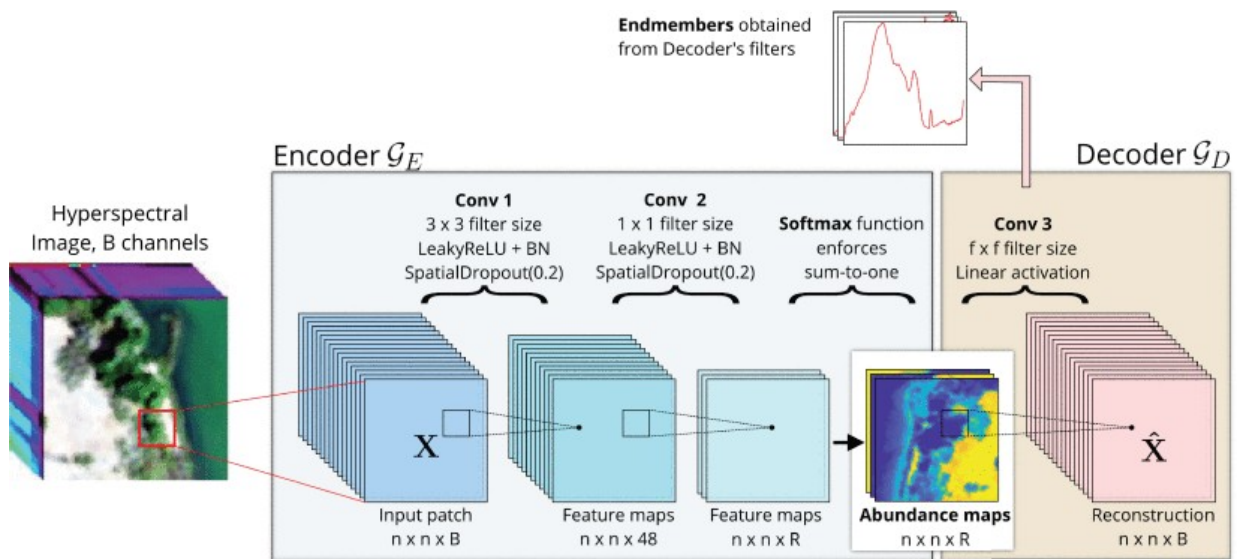
Σχήμα 9. Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου MTAEU. Source:[8]

Πίνακας 3. Τα επίπεδα του δικτύου MTAEU.

A/α	Επίπεδο	Αριθμός νευρώνων
1	Είσοδος	C
2	Ένωση	k^2C
3	Κοινό πλήρως διασυνδεδεμένο	3R
4	Batch normalization	3R
5	Πλήρως διασυνδεδεμένο	R
6	Batch normalization	R
7	Κανονικοποίηση	R
8	Πλήρως διασυνδεδεμένο	C
9	Έξοδος	C

5.4. CNNAEU

Το δίκτυο αυτό προσπαθεί επίσης να χρησιμοποιήσει την χωρική πληροφορία πραγματοποιώντας φασματικό διαχωρισμό με ένα συνελκτικό αυτοκωδικοποιητή χωρίς επίπεδα pooling και upsampling, άλλα μόνο επίπεδα συνέλιξης που διατηρούν την διάσταση της εικόνας ίδια. Οποιαδήποτε αλλαγή στις διαστάσεις της εικόνας μέσα στο δίκτυο θα οδηγούσε σε αφθονίες και καθαρούς στόχους με λάθος διαστάσεις. Τα δεδομένα εισόδου περνάνε από δύο συνελκτικά μπλοκ κάθε ένα από το οποίο αποτελείται από μια δισδιάστατη συνέλιξη με συνάρτηση ενεργοποίησης LeakyRelu, και ένα επίπεδο Dropout με παράμετρο 0,2. Η πρώτη συνέλιξη έχει φίλτρο μεγέθους 3x3 και η δεύτερη 1x1. Κατόπιν ακολουθεί η συνάρτηση ενεργοποίησης Softmax που εφαρμόζει τους περιορισμούς του γραμμικού μοντέλου και καταλήγει στην κρυφή αναπαράσταση η μεγέθους 3x3. Ο αποκωδικοποιητής αποτελείται από μια δισδιάστατη συνέλιξη μεγέθους 11x11 χωρίς κάποια συνάρτηση ενεργοποίησης. Ως συνάρτηση κόστους χρησιμοποιείται η μέση τιμή της φασματικής απόστασης για κάθε εικονοστοιχείου του παραθύρου συνέλιξης.



Σχήμα 10 .Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου CNNAEU. Source:[9]

Πίνακας 4. Τα επίπεδα του δικτύου CNNAEU.

A/α	Επίπεδο
1	Είσοδος
2	Δισδιάστατη συνέλιξη
3	LeakyReLU
4	Batch Normalization
5	Spatial Dropout
6	Δισδιάστατη συνέλιξη
7	LeakyReLU
8	Batch Normalization
9	Spatial Dropout
10	Softmax
11	Δισδιάστατη συνέλιξη
12	Έξοδος

Το μοντέλο αυτό θεωρεί μια τροποποίηση του γραμμικού μοντέλου κατά την οποία εισάγεται η χωρική διάσταση. Η φασματική υπογραφή ενός μεικτού εικονοστοιχείου που αποτελεί κέντρο μιας γειτονιάς μπορεί να γραφτεί ως:

$$x_p = M s_p + \sum_{i \in N_p: p} M_i \tilde{s}_i + \epsilon_p, p=1, \dots, P$$

Επιπλέον η ανακατασκευή της εικόνας στον αποκωδικοποιητή του μοντέλου εκφράζεται ως:

$$\hat{x}_p = \sum_{m=1}^{f^2} W_m h_m$$

όπου f^2 η γειτονιά γύρω από το εικονοστοιχείο και h τα δυναμικά ενεργοποίησης του προηγούμενου επιπέδου, δηλαδή οι αφθονίες. Με έναν μετασχηματισμό η προηγούμενη εξίσωση παίρνει την μορφή:

$$\hat{x}_p = \left(\sum_{m=1}^{f^2} W_m \right) h_c + \sum_{\substack{m=1 \\ m \neq c}}^{f^2} W_m (h_m - h_c)$$

Συγκρίνοντας το τροποποιημένο γραμμικό μοντέλο με το τελευταίο προκύπτει πως:

$$M = \sum_{m=1}^{f^2} W_m$$

5.5. EndNet

Το δίκτυο αυτό είναι ένα πλήρως διασυνδεδεμένο δίκτυο με ρηχό κωδικοποιητή και αποκωδικοποιητή. Τα δεδομένα εισόδου περνάνε από ένα πλήρως διασυνδεδεμένο επίπεδο με αριθμό νευρώνων ίσο με τον αριθμό των καθαρών στόχων, ένα ειδικά διαμορφωμένο επίπεδο Batch Normalization και από μια συνάρτηση ενεργοποίησης ReLu. Επιπλέον έχει γίνει και χρήση ενός επιπέδου Dropout με τιμή παραμέτρου 0,5 και ενός επιπέδου κανονικοποίησης l_1 αμέσως μετά την συνάρτηση ενεργοποίησης. Τα επίπεδα batch normalization και dropout βοηθάνε στην εφαρμογή της sparsity ιδιότητας των αφθονιών και η κανονικοποίηση στην εφαρμογή των περιορισμών του γραμμικού μοντέλου. Σημειώνεται πως στα δεδομένα εισόδου εισάχθηκε ισοτροπικός γκαουσιανός θόρυβος στο 40% των δεδομένων εισόδου για καλύτερη σύγκλιση στην σωστή λύση. Ο θόρυβος δεν εισάγεται σε όλα τα δεδομένα καθώς η φασματική γωνία είναι ευαίσθητη σε μεγάλες αλλαγές.

Πίνακας 5. Τα επίπεδα του δικτύου Endnet.

A/α	Επίπεδο	Αριθμός νευρώνων
1	Είσοδος	C
2	Πλήρως διασυνδεδεμένο	R
3	Batch Normalization	R
4	ReLu	R
5	Dropout	R
6	Κανονικοποίηση l_1	R
7	Πλήρως διασυνδεδεμένο	C
8	Έξοδος	C

Συγκεκριμένα από το επίπεδο batch normalization έχει αφαιρεθεί ο παράγοντας κλίμακας και επιπλέον όλες οι αρνητικές τιμές απορρίπτονται. Είναι ορισμένο ως:

$$v^{(i)} = v^{(i-1)} + \beta$$

$$v^{(i-1)} = \frac{v^{(i-1)} - \mu}{\sqrt{\sigma^2 + \epsilon}}$$

Μια σημαντική καινοτομία που εισάγει αυτή η μέθοδος είναι η αντικατάσταση του εσωτερικού γινομένου στο εσωτερικό του νευρώνα με την γωνιακή απόσταση. Δηλαδή πλέον το δυναμικό ενεργοποίησης ορίζεται ως μια κανονικοποιημένη εκδοχή της φασματικής απόστασης έτσι ώστε $C \in [0,1]$:

$$v = \varphi(C(x, w))$$

όπου:

$$C = 1.0 - \text{SAD} \frac{(x^{(i)}, x^{(j)})}{\pi}$$

Τέλος οι συγγραφείς ορίζουν και μια καινούρια συνάρτηση κόστους που βασίζεται στο μέσο τετραγωνικό σφάλμα και στην γωνιακή απόσταση και εκφράζεται ως:

$$L = \frac{\lambda_0}{2} \|x - \hat{x}\|_2^2 - \lambda_1 D_{KL}(1.0 \| C(\text{bold}x, \hat{x})) + \lambda_2 \|z\|_1 + \lambda_3 \|W^{(e)}\|_2 + \lambda_4 \|W^{(d)}\|_2 + \lambda_5 \|p\|_2$$

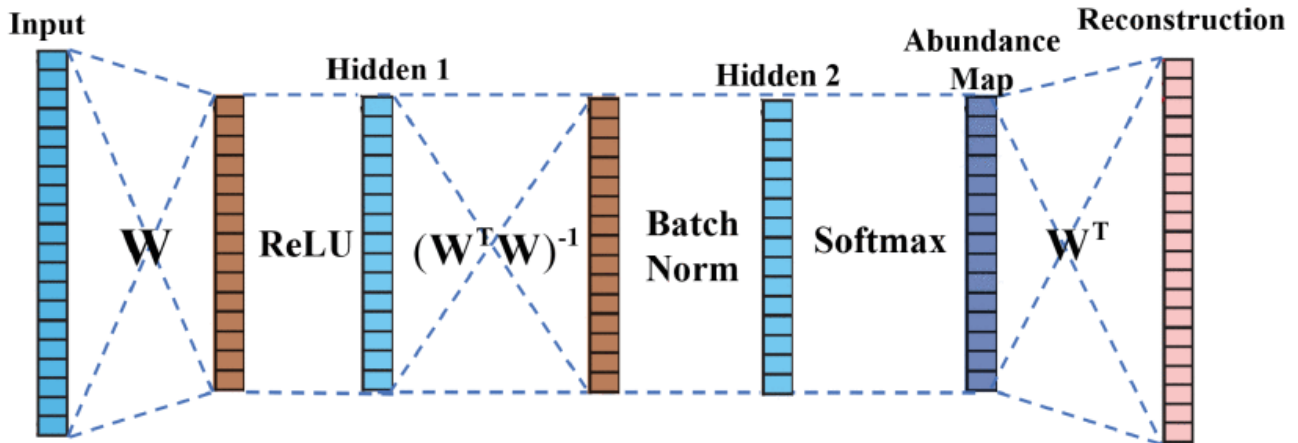
Ο πρώτος όρος εκφράζει το μέσο τετραγωνικό σφάλμα. Ο δεύτερος την απόσταση Kullback-Leibler ανάμεσα στην φασματική απόσταση των δεδομένων εισόδου και εξόδου και στην φασματική απόσταση των δεδομένων εισόδου με τον εαυτό τους. Ο τρίτος όρος αφορά την έξοδο του επιπέδου dropout. Ο τέταρτος, ο πέμπτος και ο έκτος είναι η δεύτερη νόρμα των βαρών του κωδικοποιητή, των βαρών του αποκωδικοποιητή και της παραμέτρου μετατόπισης του επιπέδου batch normalization αντίστοιχα. Οι τιμές των παραμέτρων των όρων φαίνονται στον παρακάτω πίνακα.

Πίνακας 6. Τιμές των παραμέτρων της συνάρτησης κόστους.

Παράμετρος	Τιμή
λ_0	0,01
λ_1	10
λ_2	0,1
λ_3	10e-5
λ_4	10e-5
λ_5	10e-3

5.6. OSPAEU

Το δίκτυο OSPAEU αποτελείται από έναν αποκωδικοποιητή με δύο πλήρως διασυνδεδεμένα επίπεδα. Το πρώτο επίπεδο ακολουθείται από μια συνάρτηση ReLU ενώ το δεύτερο από ένα επίπεδο Batch Normalization και μια συνάρτηση softmax η οποία και εφαρμόζει τους περιορισμούς του γραμμικού μοντέλου. Ο αποκωδικοποιητής είναι ρηχός και γραμμικός. Η αρχιτεκτονική του δικτύου φαίνεται στο παρακάτω σχήμα και πίνακα.



Σχήμα 11 .Σχηματική αναπαράσταση της αρχιτεκτονικής του δικτύου OSPAEU. Source:[11]

Πίνακας 7. Τα επίπεδα του δικτύου OSPAEU.

A/α	Επίπεδο	Αριθμός νευρώνων
1	Είσοδος	C
2	Πλήρως διασυνδεδεμένο	3R
3	ReLU	3R
4	Πλήρως διασυνδεδεμένο	R
5	Batch normalization	R
6	Softmax	R
7	Πλήρως διασυνδεδεμένο	C
8	Έξοδος	C

Η καινοτομία αυτού του δικτύου βρίσκεται στην πρόταση μιας καινούριας συνάρτησης κόστους βασισμένης σε μια ελαχιστοποίηση που ονομάζεται OSP (Orthogonal Sparse Prior). Η ελαχιστοποίηση αυτή βασίζεται στην ιδιότητα sparsity των αφθονιών κατά την οποία συνήθως το πολύ δύο καθαροί στόχοι επηρεάζουν το μεγαλύτερο μέρος της φασματικής υπογραφής του μεικτού εικονοστοιχείου. Συνεπώς δύο διαφορετικοί χάρτες αφθονίας θα πρέπει να είναι ορθογωνικοί μεταξύ τους. Το OSP εκφράζεται ως:

$$L_{OSP}(\mathbf{B}) = \sum_{i < j} \frac{\mathbf{B}_i \mathbf{B}_j}{\|\mathbf{B}_i\|_2 \|\mathbf{B}_j\|_2}$$

όπου \mathbf{B} οι παραγόμενοι χάρτες αφθονίας και i, j οι γραμμές και οι στήλες των χαρτών. Στην ουσία με την ελαχιστοποίηση αυτού του μεγέθους οι χάρτες αφθονίας γίνονται όλο και πιο ορθογωνικοί μεταξύ τους. Επίσης οι συγγραφείς υιοθετούν την λαπλασιανή κατανομή και όχι την γκαουσιανή για να περιγράψουν το σφάλμα που εμπεριέχεται στο γραμμικό μοντέλο. Έτσι, η συνάρτηση κόστους που προτείνεται είναι η εξής:

$$E = \|\mathbf{X} - \mathbf{R}\|_p^p - \sum_i \log\left(\frac{\mathbf{R}_i \cdot \mathbf{X}_i}{\|(\mathbf{WH})_i\|_2 \|\mathbf{X}_i\|_2}\right) + \rho_1 L_{OSP} + \rho_2 \sum_i L_1(\mathbf{h}_i) + \rho_3 S(\mathbf{M}_{w < 0} \odot \mathbf{W}) + \rho_4 S(\mathbf{M}_{w > 0} \odot \mathbf{W})$$

Ο πρώτος όρος αναφέρεται στην ελαχιστοποίηση της λαπλασιανής κατανομής του θορύβου που μεταφράζεται σε μεγιστοποίηση αυτού του όρου. Ο δεύτερος αφορά την γωνιακή απόσταση ανάμεσα στους καθαρούς στόχους. Ο τρίτος εκφράζει την ορθογωνική σχέση των αφθονιών. Οι τρεις τελευταίοι όροι περιγράφονται από τις παρακάτω εξισώσεις:

$$L_1(\mathbf{h}_i) = \sum_j |h_{ij}|^p$$

$$(\mathbf{M}_{w < 0})_{ij} = \begin{cases} 1 & \text{if } W_{ij} < 0 \\ 0, & \text{else} \end{cases}$$

$$(\mathbf{M}_{w > 1})_{ij} = \begin{cases} 1 & \text{if } W_{ij} > 1 \\ 0, & \text{else} \end{cases}$$

$$S(\mathbf{M}) = \sum_{ij} M_{ij}$$

Η ελαχιστοποίηση του τρίτου εφαρμόζει την ιδιότητα sparsity των αφθονιών και οι δύο τελευταίοι προσπαθούν να περιορίσουν τις αφθονίες στο διάστημα $[0,1]$ εξισώνοντας στη μονάδα το άθροισμα των αφθονιών αν αυτό είναι μικρότερο ή μεγαλύτερο του 0 και μηδενίζοντας το σε όλες τις άλλες περιπτώσεις. Οι τιμές των συντελεστών φαίνονται στον πίνακα 8.

Πίνακας 8. Τιμές των παραμέτρων της συνάρτησης κόστους.

Παράμετρος	Τιμή
ρ	0,7
ρ_1	0,5
ρ_2	1
ρ_3	10
ρ_4	10

Τέλος, οι συγγραφείς προτείνουν έναν νέο τρόπο αρχικοποίησης των τιμών των βαρών που πραγματοποιείται με τον εντοπισμό ακραίων τιμών (outliers) στην υπερφασματική εικόνα. Ως ακραίες τιμές θεωρούνται τα εικονοστοιχεία που βρίσκονται σε περιοχές μετάβασης των κατηγοριών χρήσης γης. Έτσι αρχικά υπολογίζεται ένας χάρτης ομοιότητας της υπερφασματικής εικόνας που σκοπό έχει να εντοπίσει τις περιοχές μετάβασης που αναμένεται να έχουν υψηλή ανομοιομορφία. Ο χάρτης υπολογίζεται σύμφωνα με την ομοιότητας heat kernel με γειτονιά εφαρμογής $N = 8$ εικονοστοιχεία. Μαθηματικά οι τιμές του χάρτη ομοιότητας ορίζονται ως:

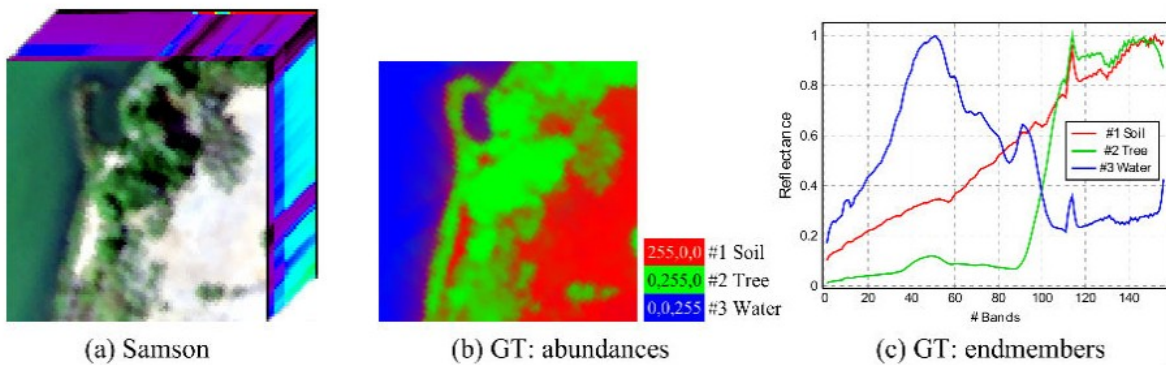
$$d_i = \frac{1}{|N_i|} \sum_{j \in N_i} s_{ij}$$

$$s_{ij} = e^{-\frac{\|x_i - x_j\|_2^2}{\sigma}}$$

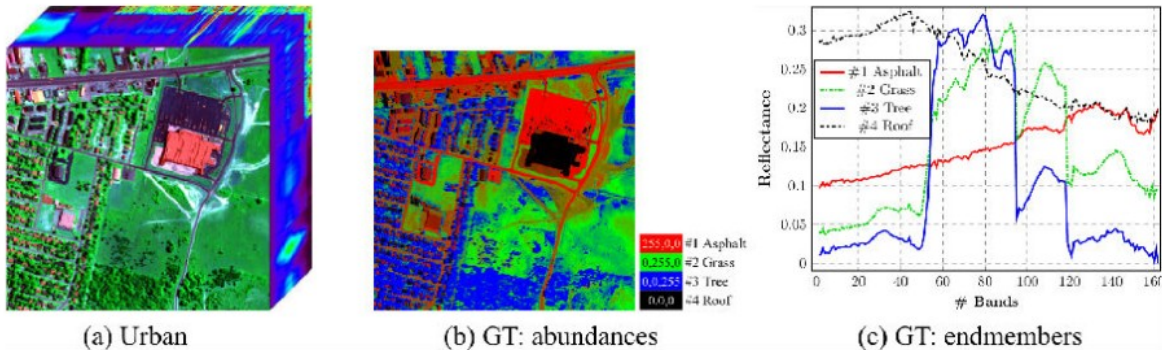
με $\sigma = 0,005$ μια τιμή ρύθμισης της αντίθεσης του χάρτη. Έτσι όλα τα εικονοστοιχεία πάνω από ένα κατώφλι $\tau = 0,5$ θεωρούνται κανονικά και χρησιμοποιούνται για την αρχικοποίηση των βαρών του δικτύου με βάση την μέθοδο απόστασης DMaxD.

6. Αποτελέσματα πειραμάτων

Για κάθε μέθοδο που προαναφέρθηκε εκτελέστηκαν δέκα πειράματα με σκοπό την συγκριτική αξιολόγηση τους. Τα πειράματα αυτά γίνανε πάνω σε δύο υπερφασματικές εικόνες: Samson (Σχήμα 12) και Urban (Σχήμα 13). Η Samson είναι σχετικά χαμηλή χωρική ανάλυση, μεγέθους 95x95 εικονοστοιχείων. Έχει 156 κανάλια. Την χαρακτηρίζουν τρεις καθαροί στόχοι: δέντρα, νερό και έδαφος. Η Urban έχει 307x307 εικονοστοιχεία και 162 κανάλια. Την χαρακτηρίζουν τέσσερις καθαροί στόχοι: άσφαλτος, γρασίδι, δέντρα και κεραμυδοσκεπές. Υπάρχει και σε εκδοχή με πέντε καθαρούς στόχους με τον πέμπτο να είναι το έδαφος. Οι εικόνες λήφθηκαν από [αυτήν](#) την ιστοσελίδα και τα δεδομένα επαλήθευσης προέκυψαν από κλασσικές μεθόδους φασματικού διαχωρισμού. Τα πειράματα έτρεξαν με τις τιμές παραμέτρων που προαναφέρθηκαν στην παρουσίαση των μεθόδων, με Adam optimizer και ρυθμό μάθησης ίσο με $10e-3$. Επίσης ο αριθμός των εποχών και του batch size φαίνονται στον πίνακα 9. Στην συνέχεια παρουσιάζονται οι εξαγόμενοι καθαροί στόχοι και χάρτες αφθονιών για κάθε μέθοδο και πείραμα.



Σχήμα 12 .Η εικόνα Samson με τις αφθονίες και τους καθαρούς στόχους που την συνοδεύουν.



Σχήμα 13 .Το σύνολο δεδομένων Urban με τις αφθονίες και τους καθαρούς στόχους που το συνοδεύουν.

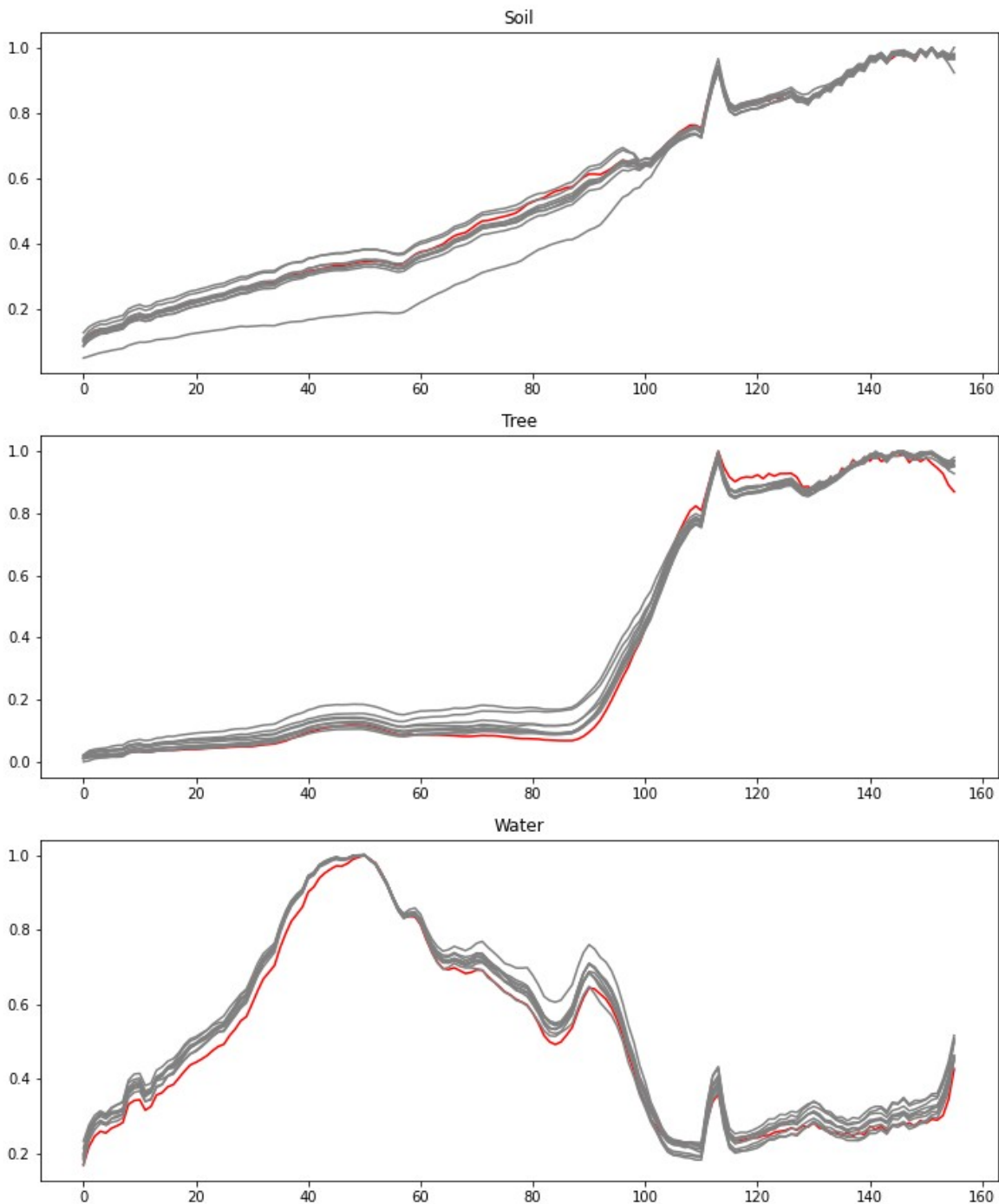
Πίνακας 9. Εποχές και batch size για την κάθε μέθοδο.

	DAEU	SIDAEU	MTAEU	CNNAEU	ENDNET	OSPAEU
Εποχές	40	40	60	40	40	13
Batch size	6	5	15	250	64	256

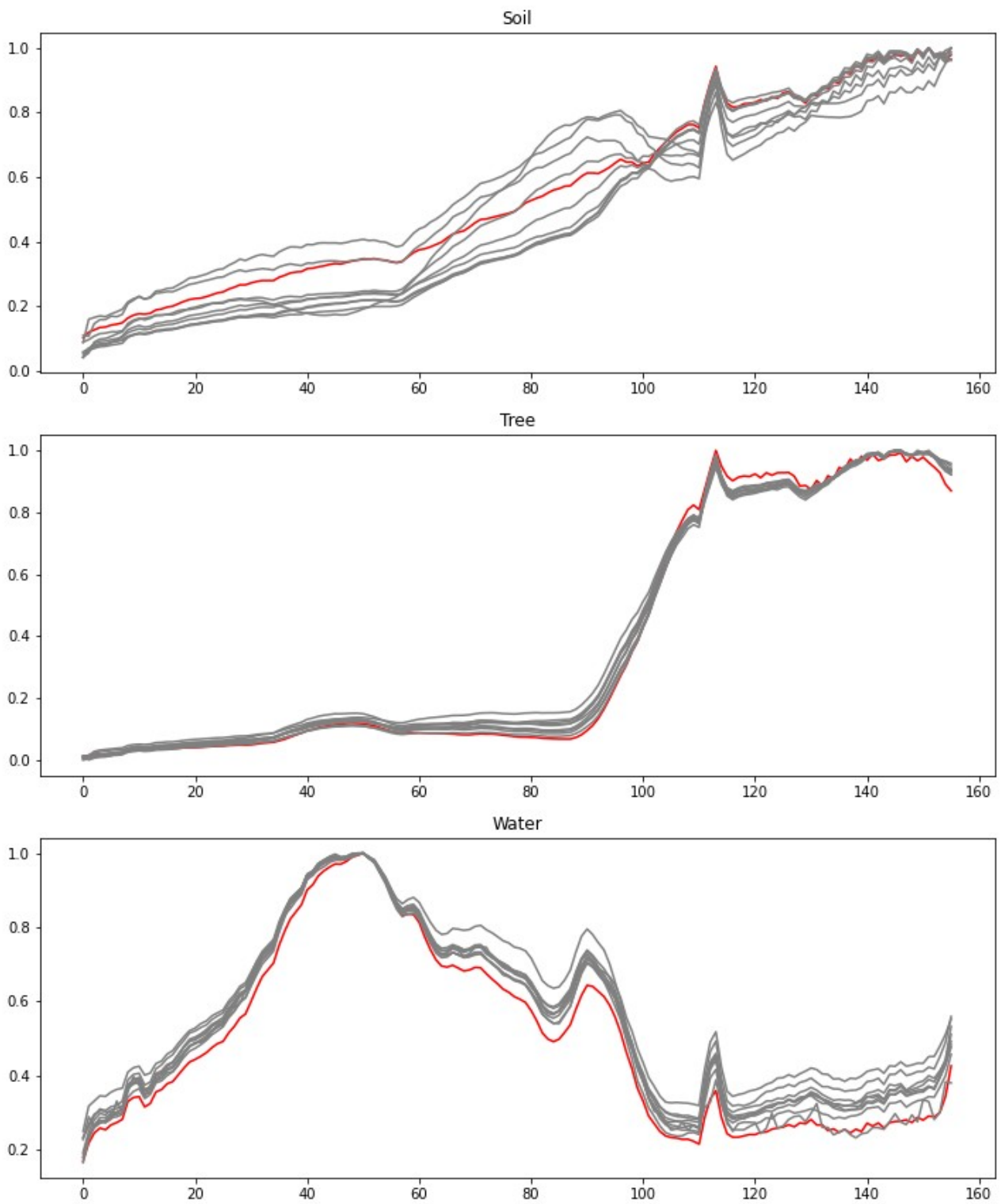
5.1. Καθαροί στόχοι

Παρακάτω, στα διαγράμματα 14-31 παρουσιάζονται οι εξαγόμενοι καθαροί στόχοι. Με γκρι χρώμα απεικονίζονται οι εξαγόμενοι καθαροί στόχοι των δέκα πειραμάτων και με κόκκινο τα δεδομένα επαλήθευσης. Με αυτή την οπτικοποίηση φαίνεται ποιοτικά η σταθερότητα άλλα και η ακρίβεια των μεθόδων όσον αφορά τους καθαρούς στόχους.

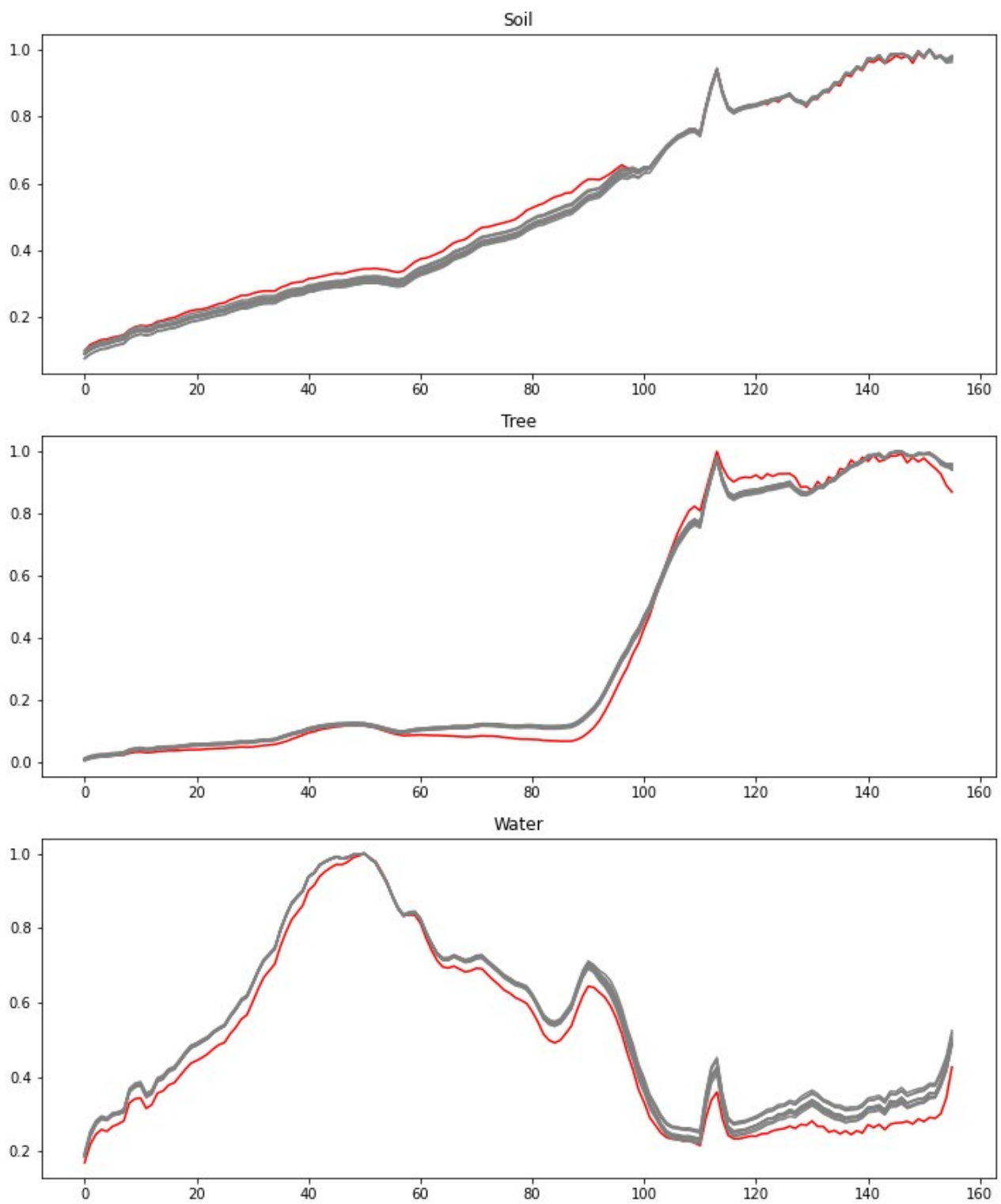
5.1.1 Samson



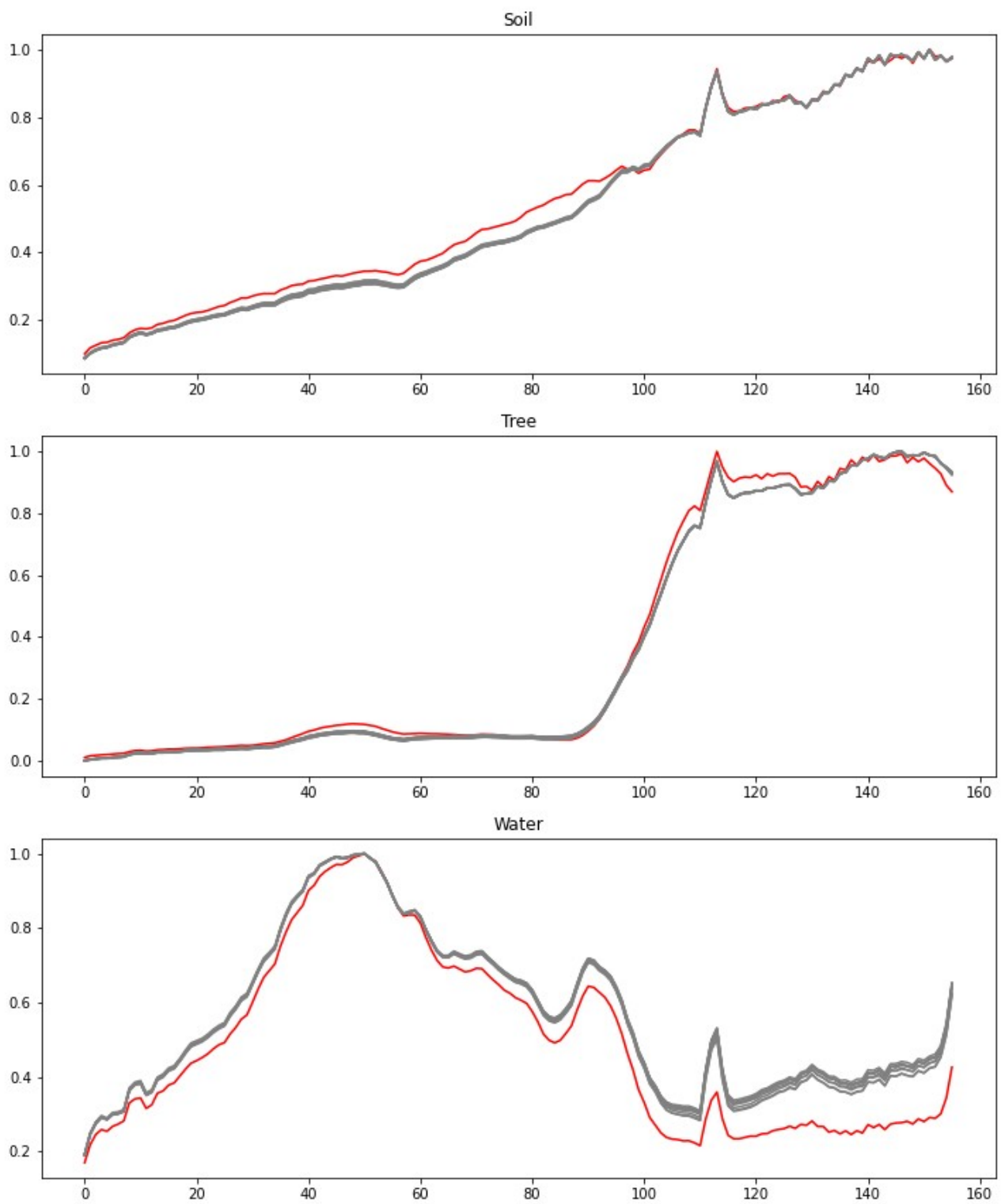
Σχήμα 14 .Παραγόμενοι καθαροί στόχοι με την μέθοδο DAEU στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



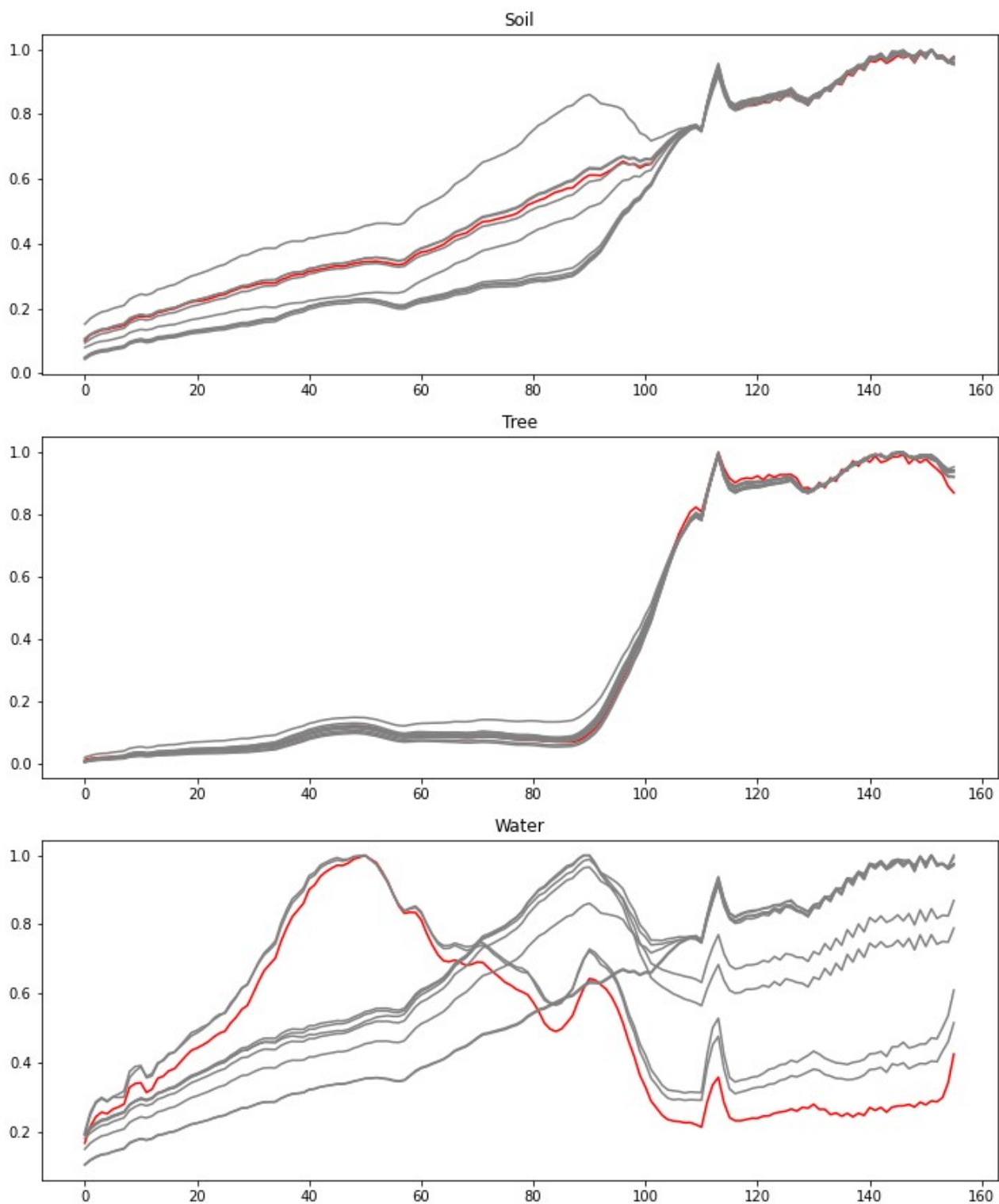
Σχήμα 15 .Παραγόμενοι καθαροί στόχοι με την μέθοδο SIDAEU στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



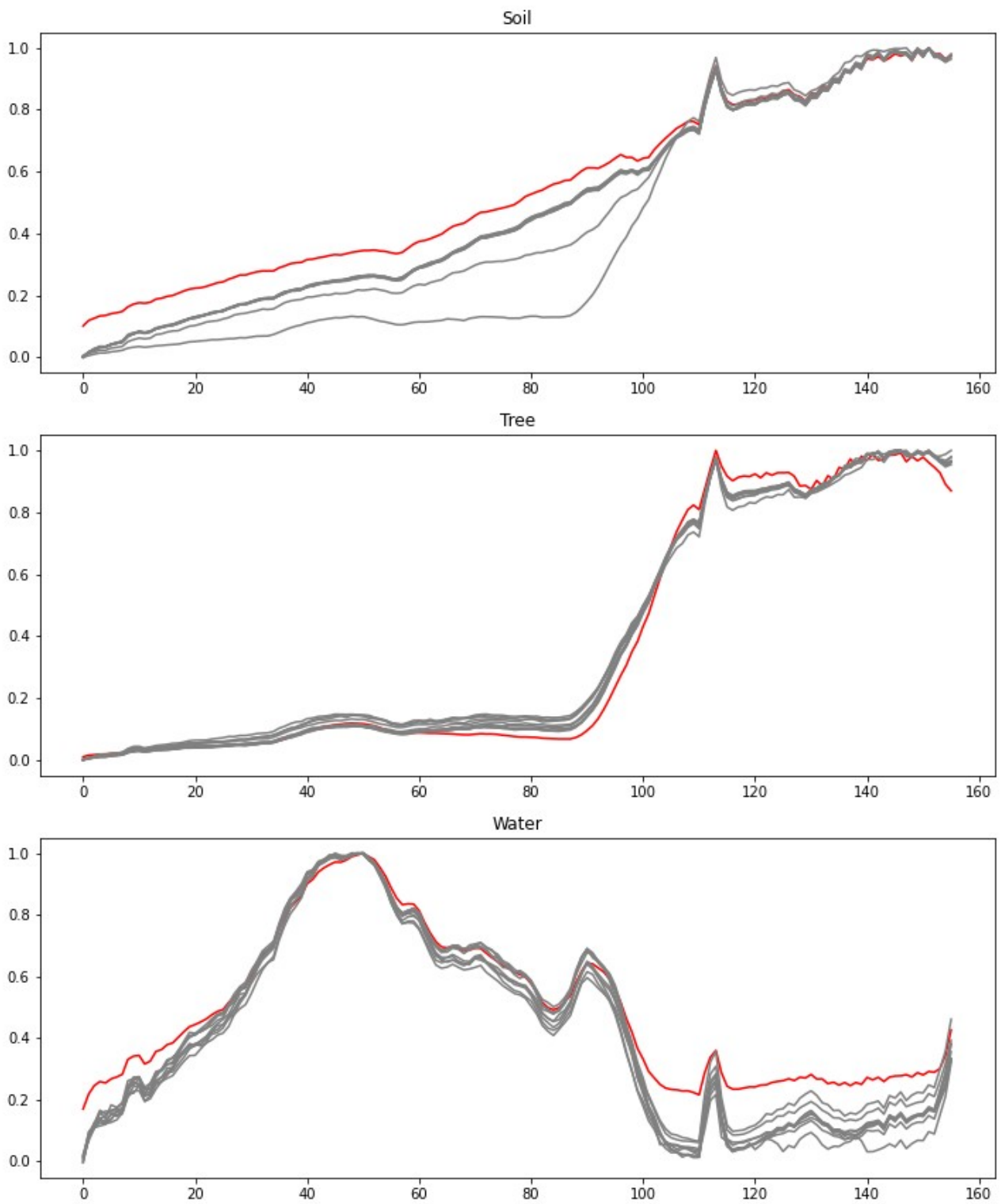
Σχήμα 16 .Παραγόμενοι καθαροί στόχοι με την μέθοδο ΜΤΑΕU στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



Σχήμα 17 .Παραγόμενοι καθαροί στόχοι με την μέθοδο CNNAEU στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

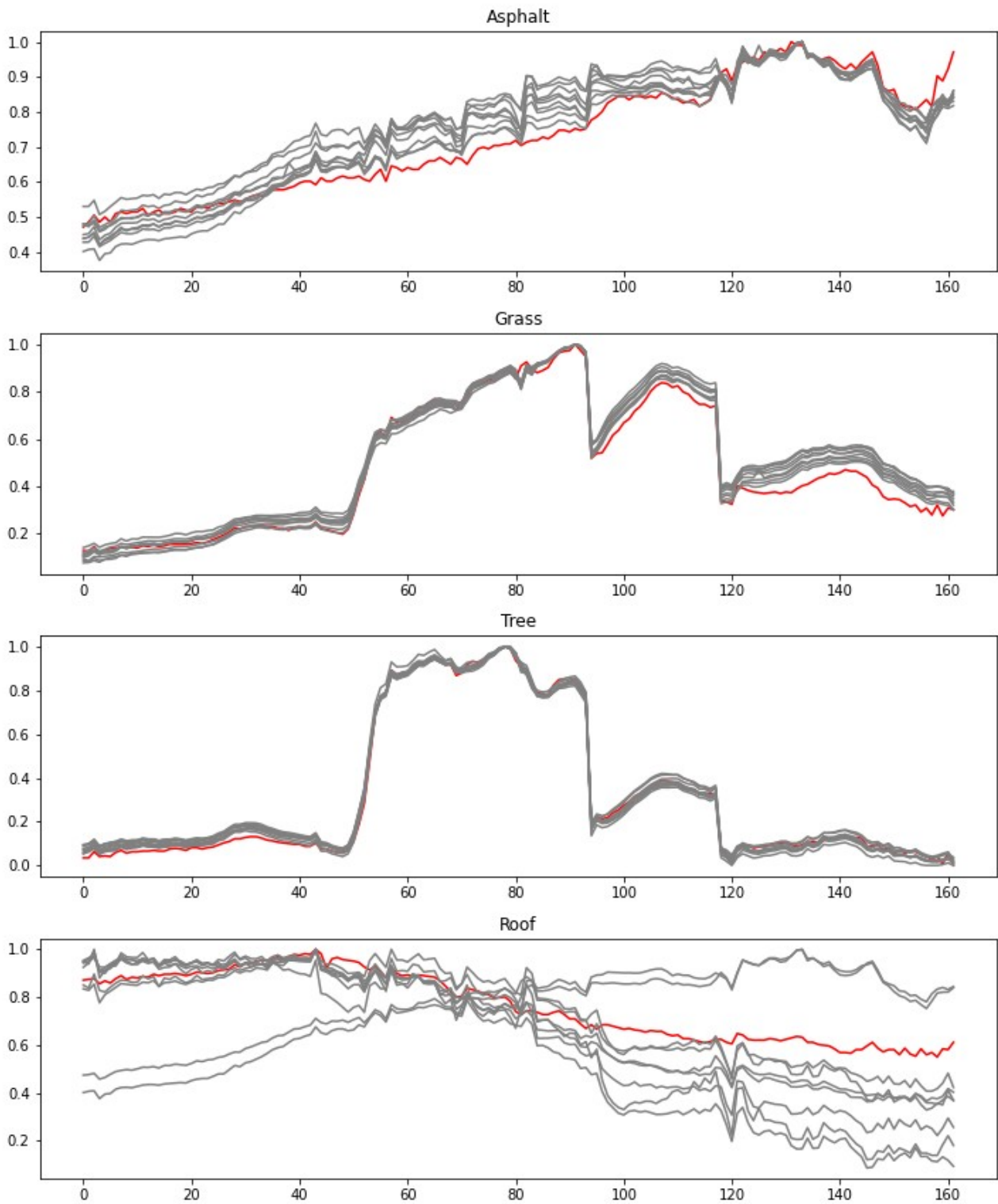


Σχήμα 18 .Παραγόμενοι καθαροί στόχοι με την μέθοδο Endnet στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

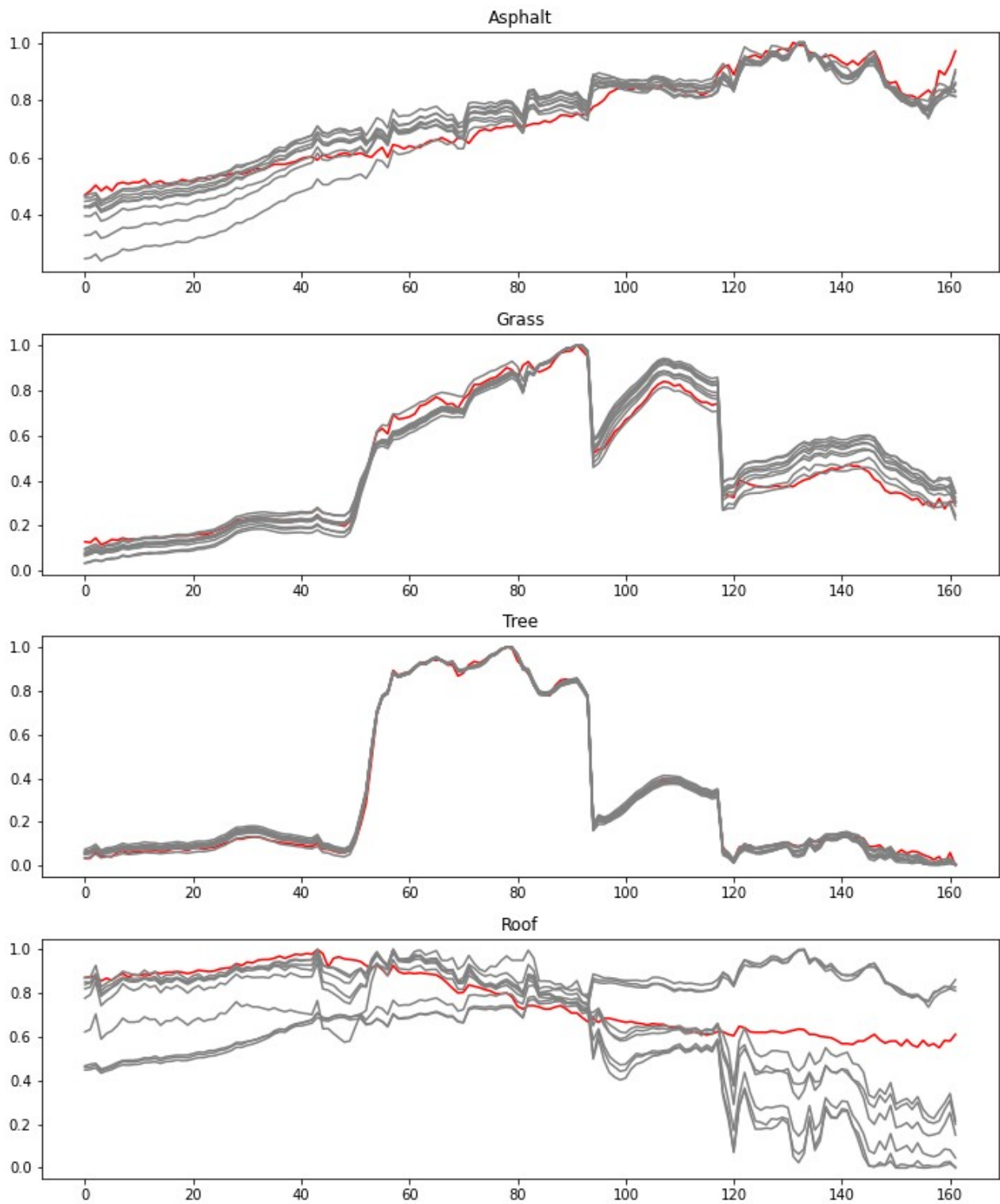


Σχήμα 19 .Παραγόμενοι καθαροί στόχοι με την μέθοδο OSPAEU στο σύνολο δεδομένων Samson. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

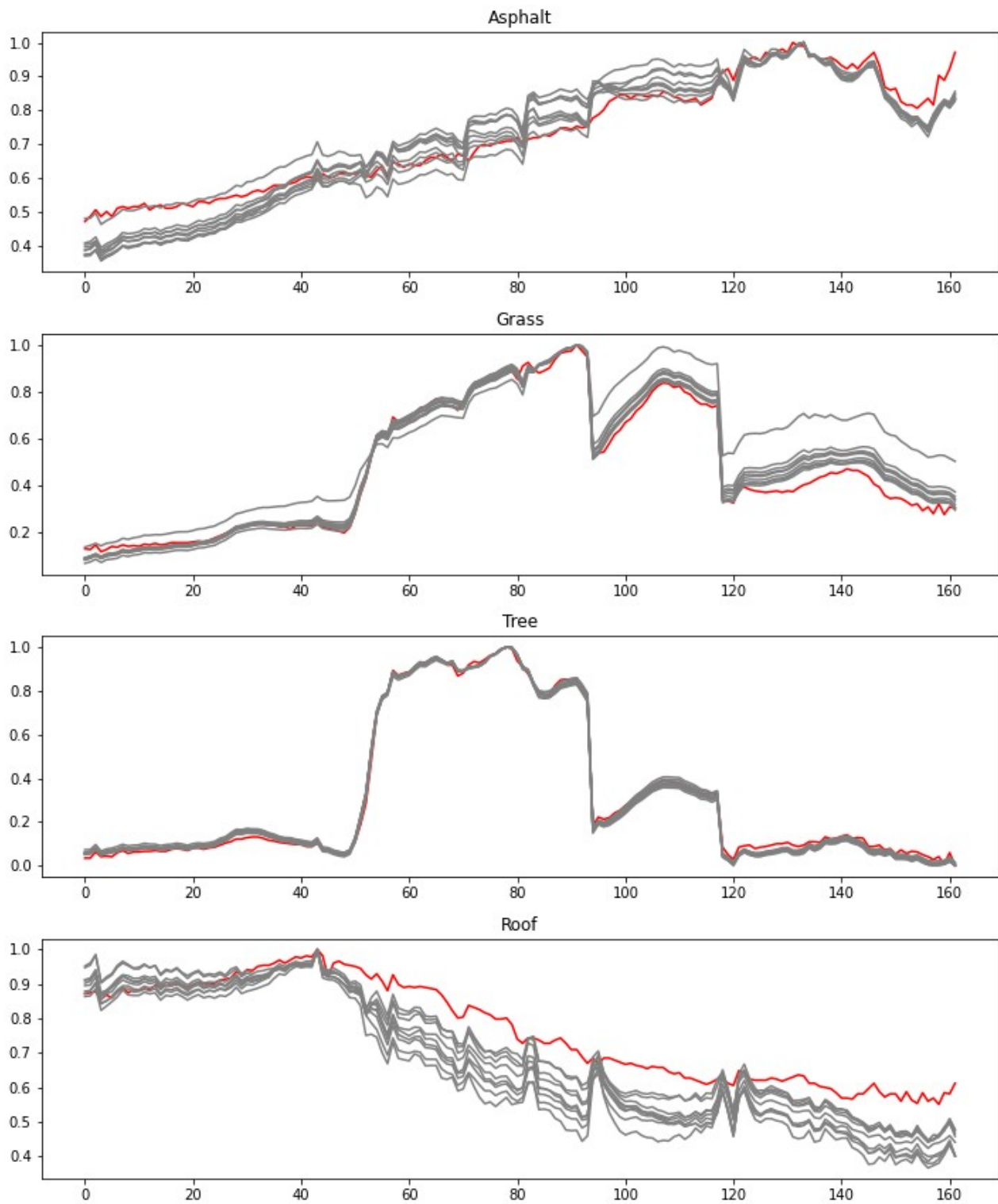
5.1.2 Urban με τέσσερις καθαρούς στόχους



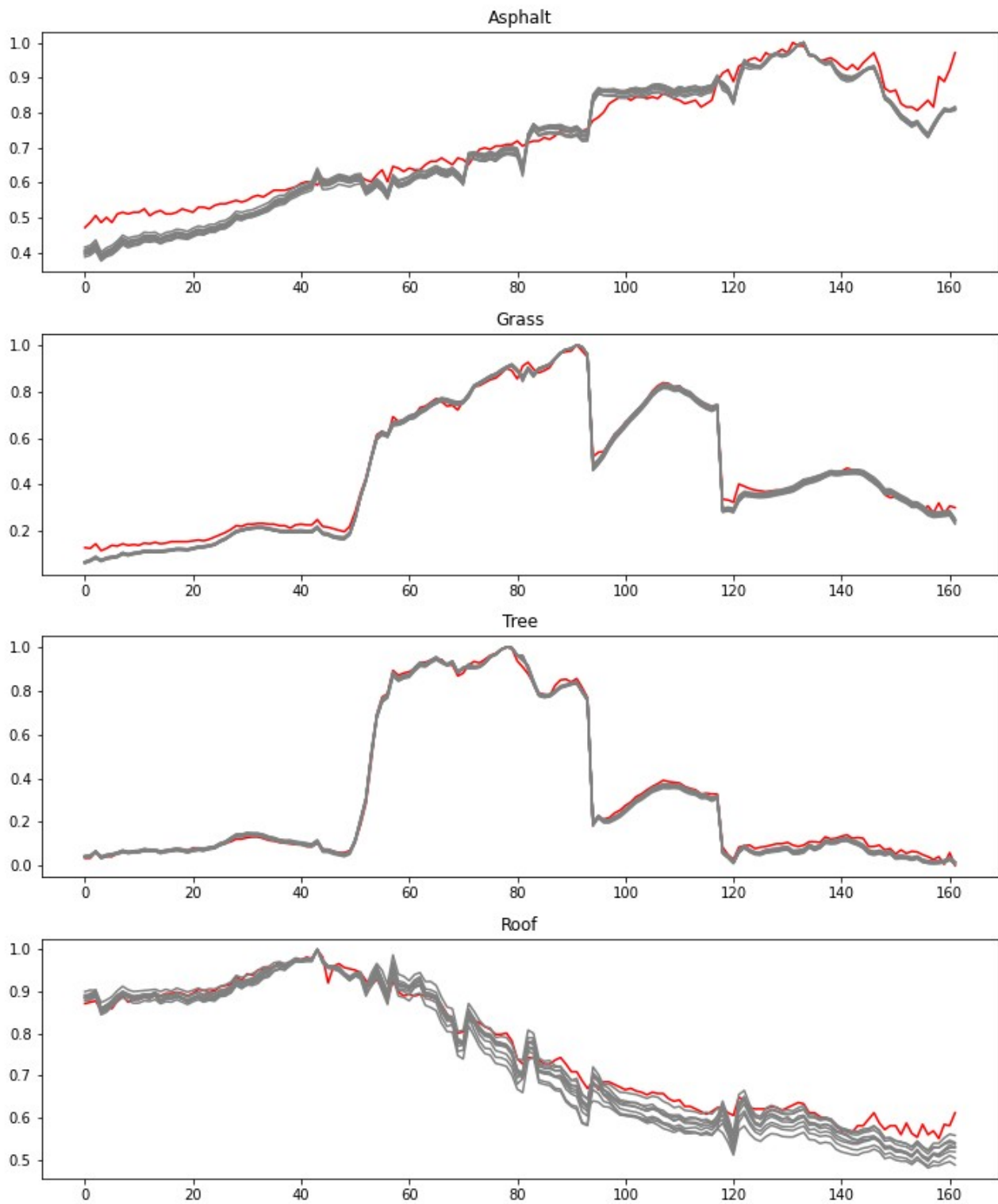
Σχήμα 20 .Παραγόμενοι καθαροί στόχοι με την μέθοδο DAEU στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



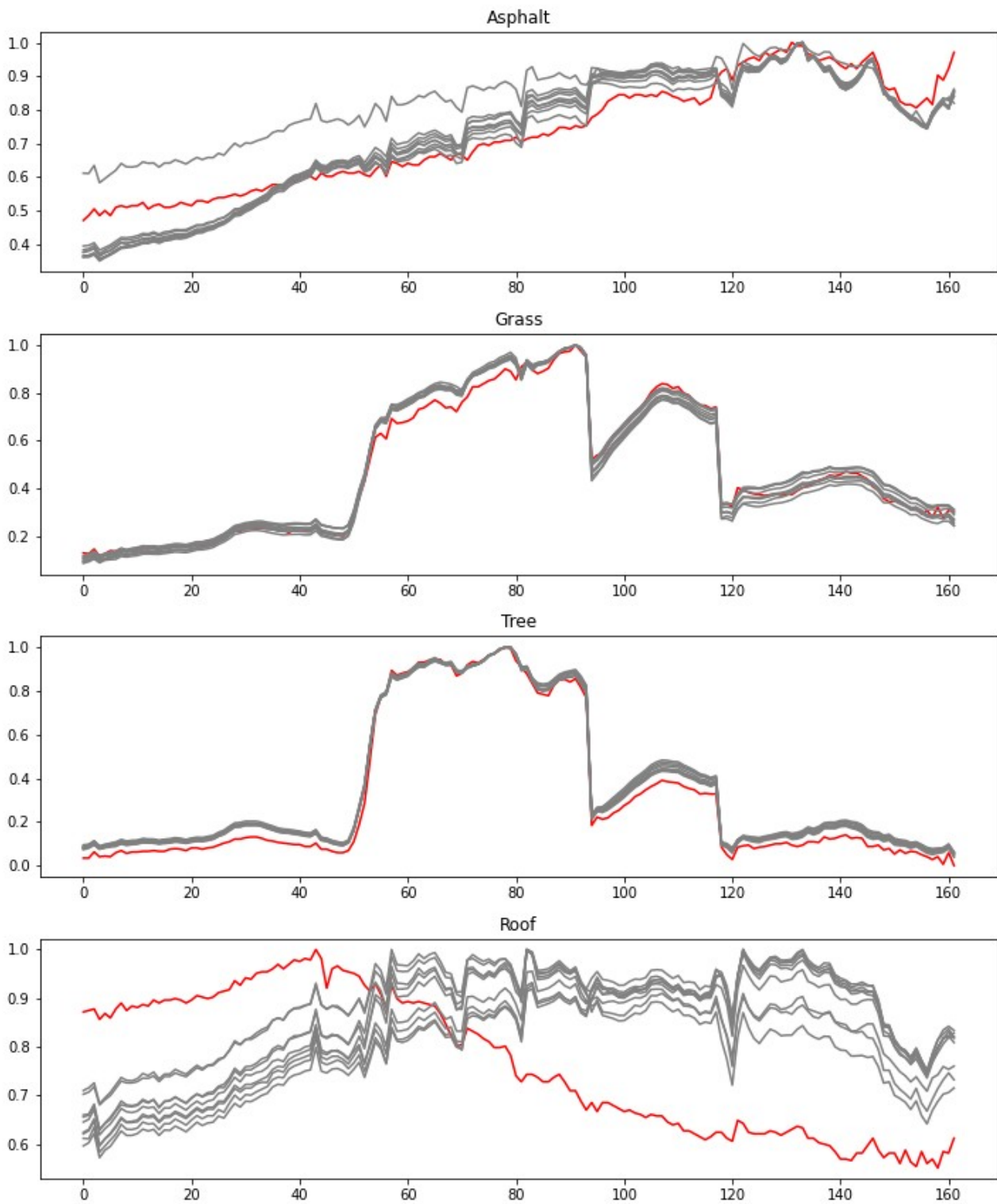
Σχήμα 21 .Παραγόμενοι καθαροί στόχοι με την μέθοδο SIDAEU στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



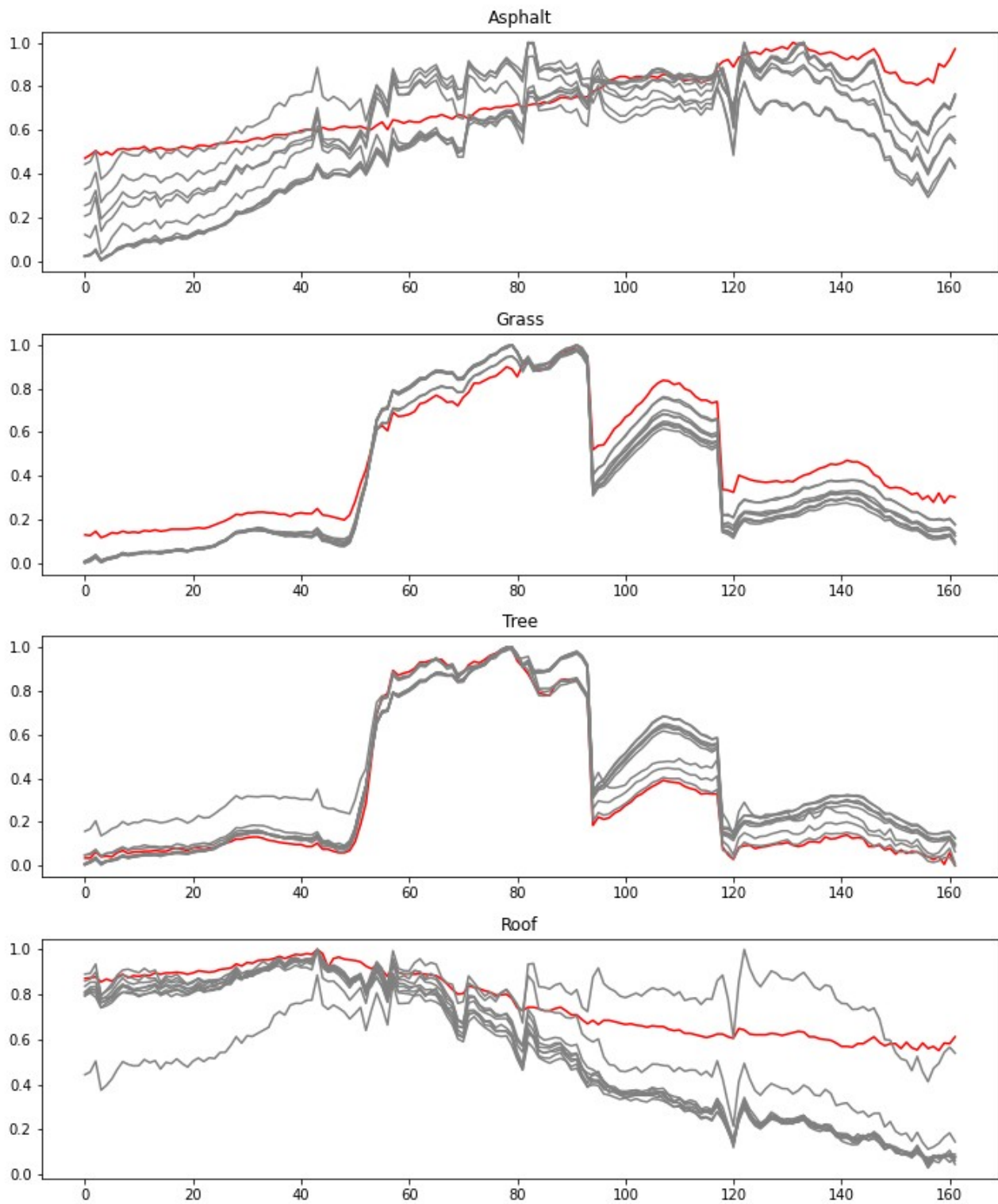
Σχήμα 22 .Παραγόμενοι καθαροί στόχοι με την μέθοδο ΜΤΑΕΥ στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



Σχήμα 23 .Παραγόμενοι καθαροί στόχοι με την μέθοδο CNNAEU στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

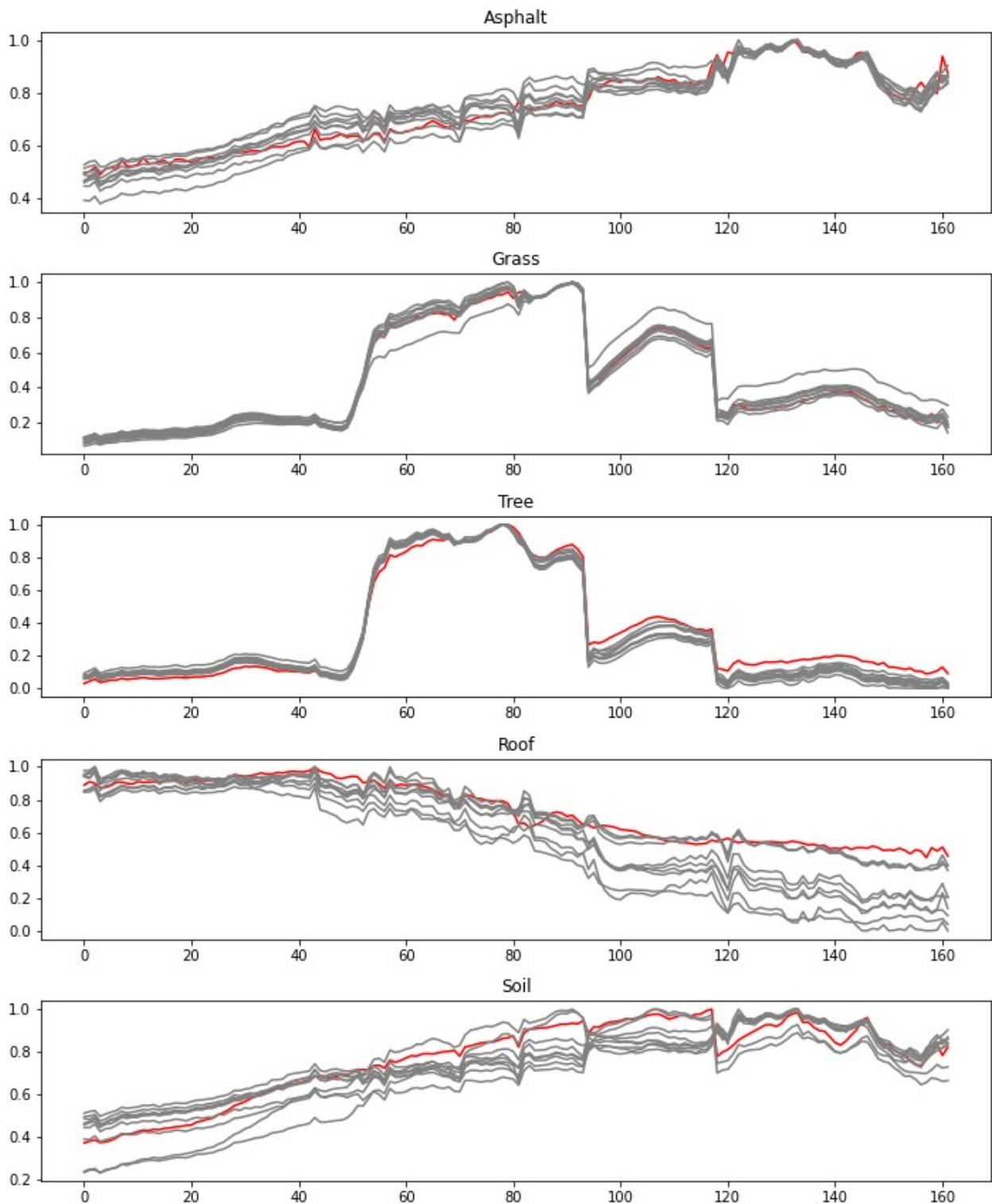


Σχήμα 24 .Παραγόμενοι καθαροί στόχοι με την μέθοδο Emdnet στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

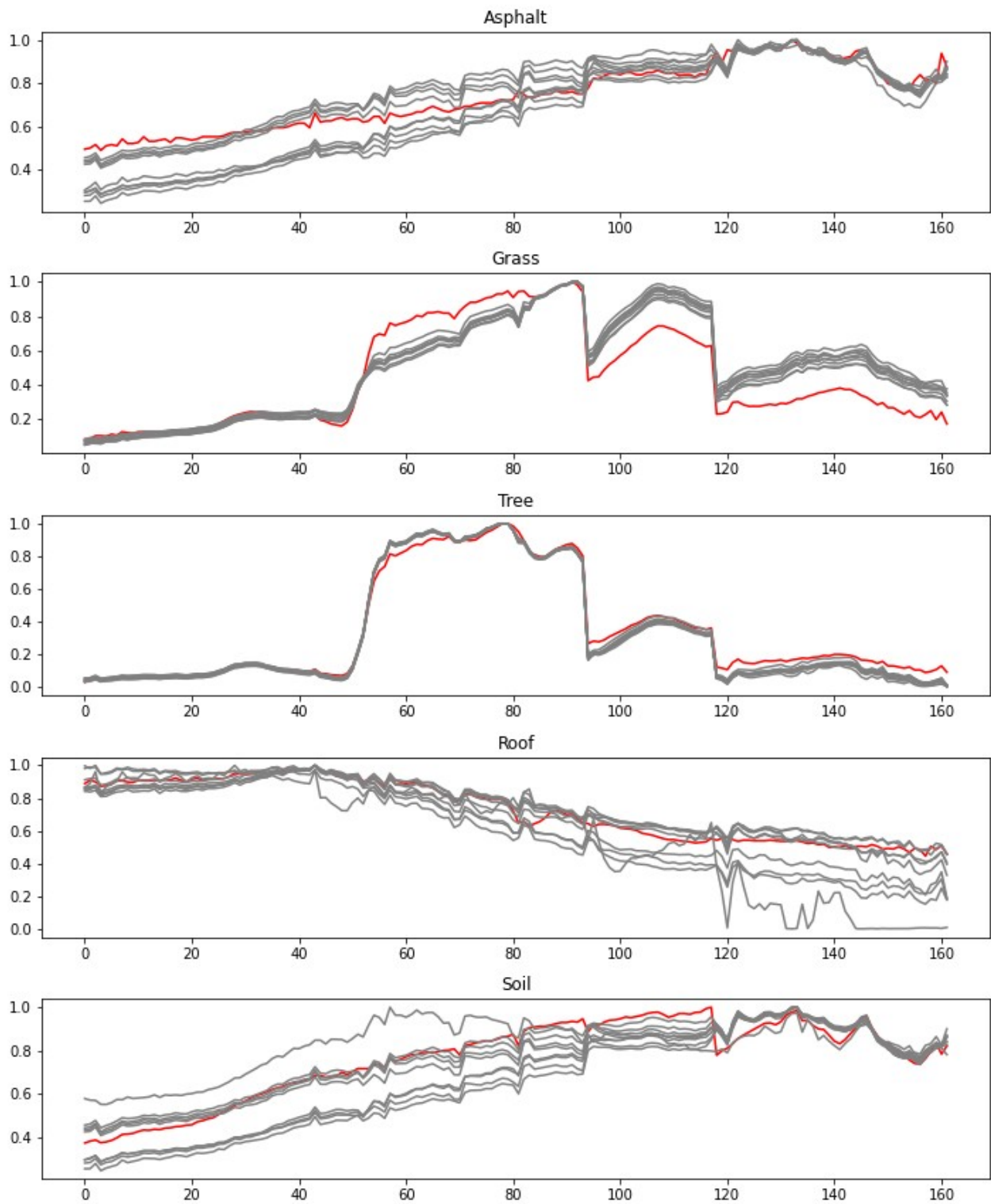


Σχήμα 25 .Παραγόμενοι καθαροί στόχοι με την μέθοδο OSPAEU στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

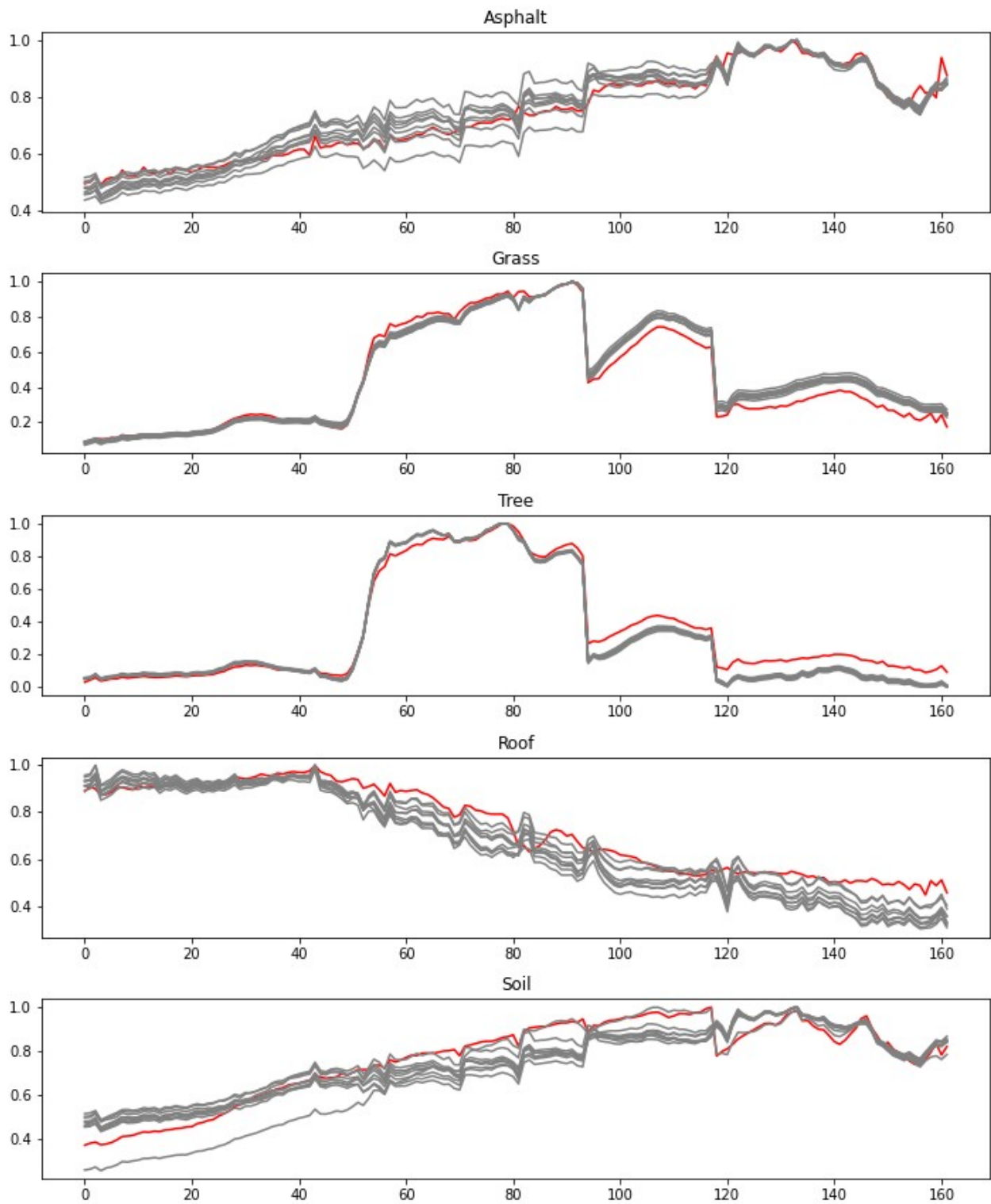
5.1.3 Urban με πέντε καθαρούς στόχους



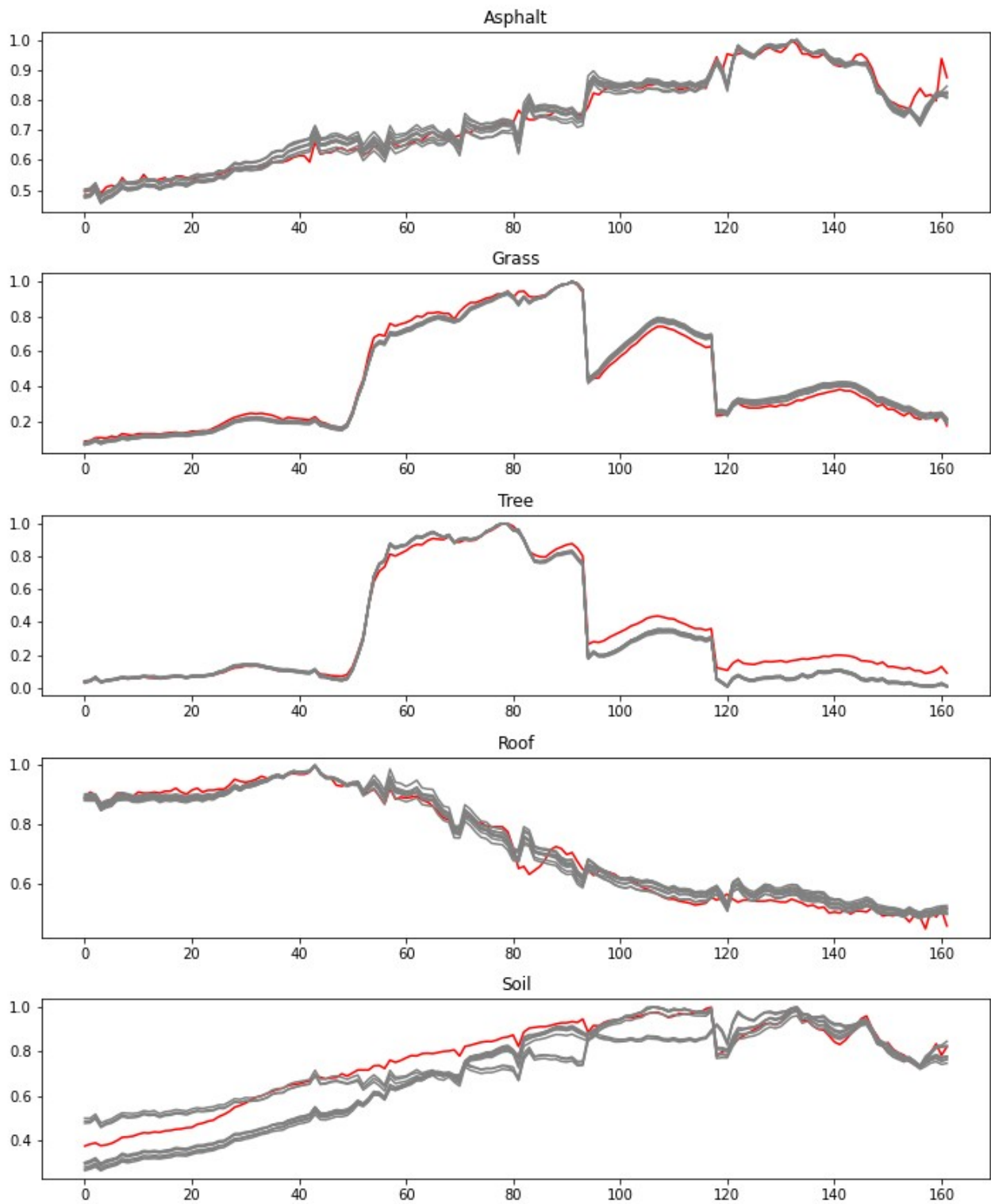
Σχήμα 26 .Παραγόμενοι καθαροί στόχοι με την μέθοδο DAEU στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



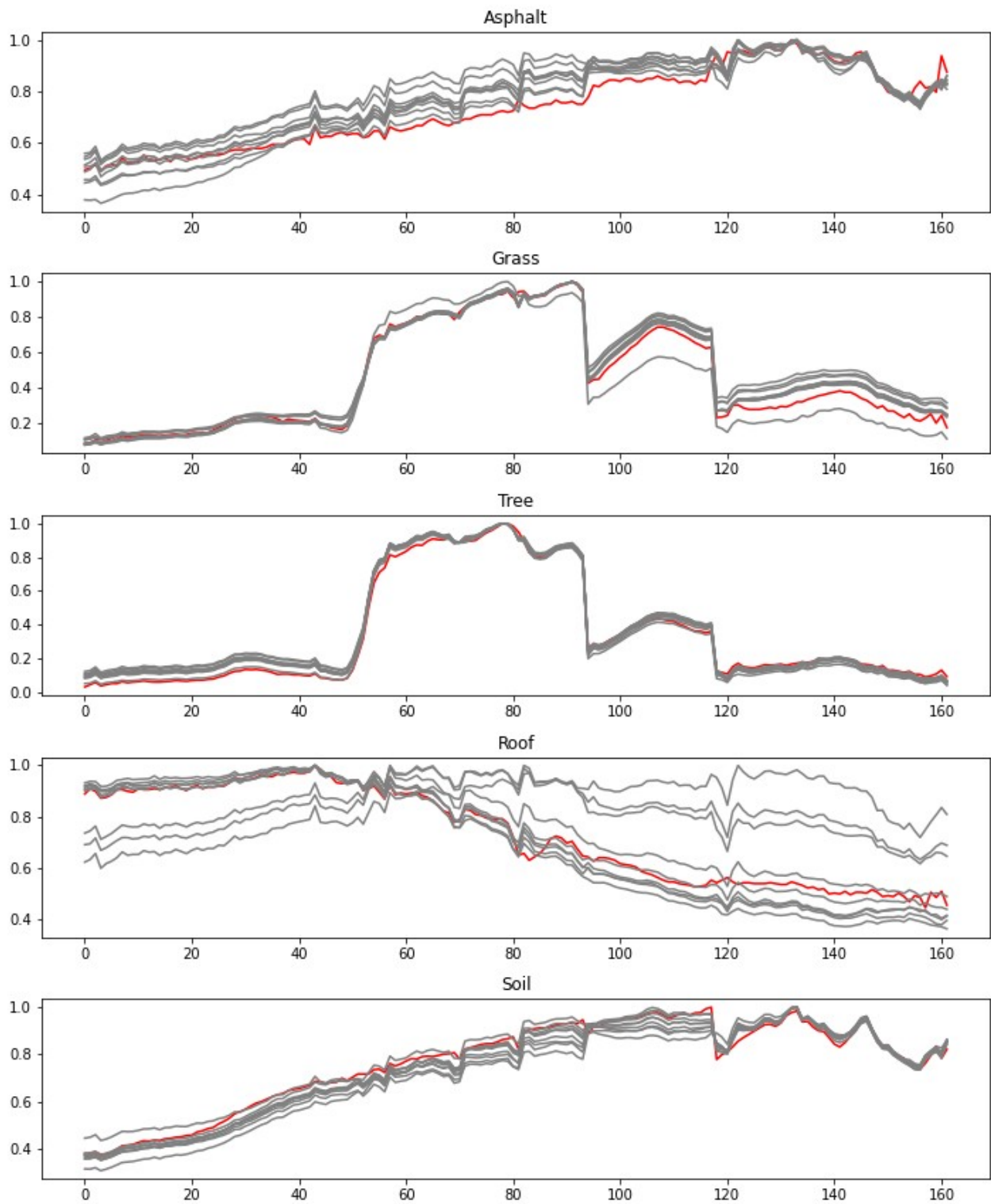
Σχήμα 27 .Παραγόμενοι καθαροί στόχοι με την μέθοδο SIDAEU στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



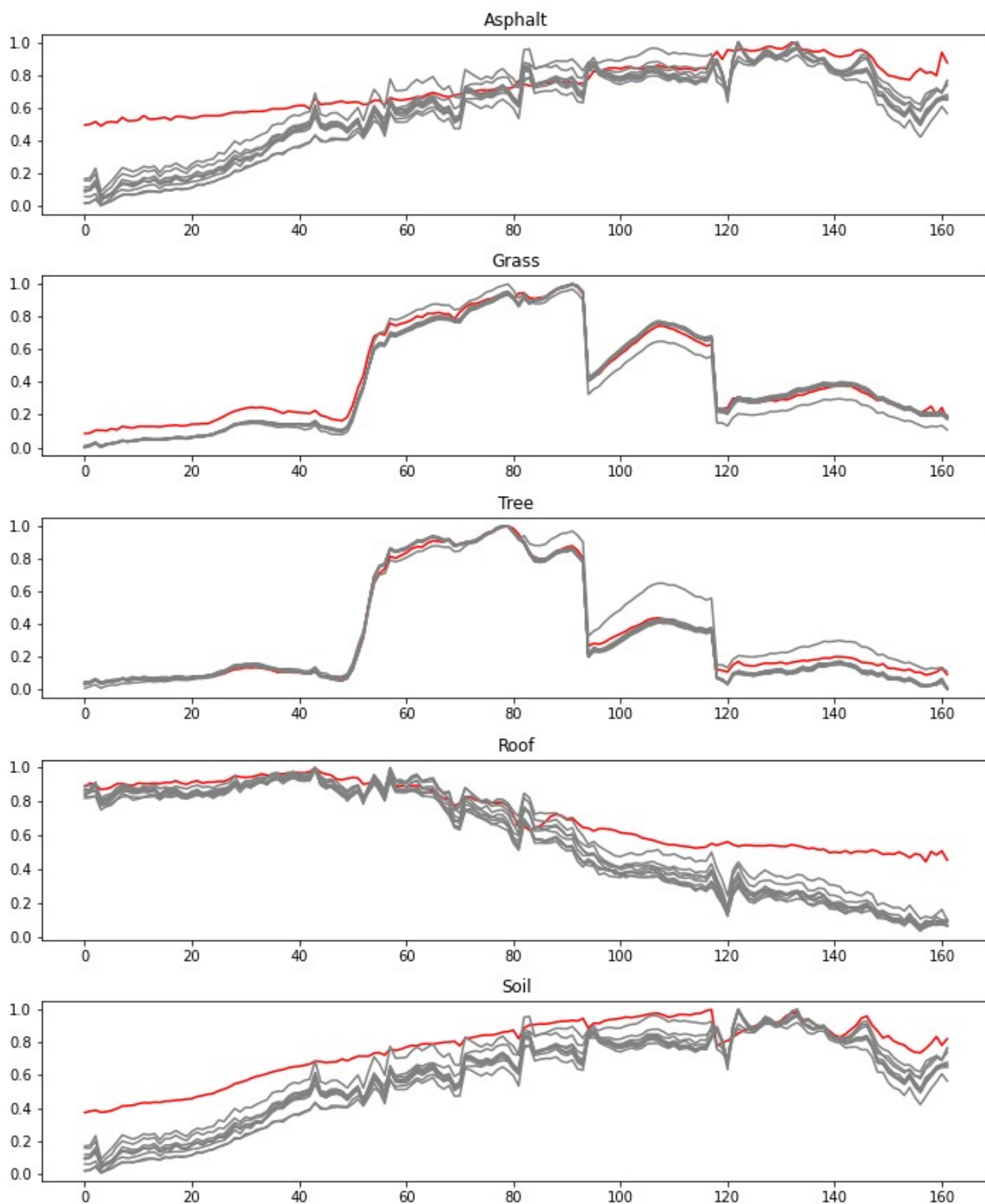
Σχήμα 28 .Παραγόμενοι καθαροί στόχοι με την μέθοδο MTAEU στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



Σχήμα 29 .Παραγόμενοι καθαροί στόχοι με την μέθοδο CNNAEU στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



Σχήμα 30. Παραγόμενοι καθαροί στόχοι με την μέθοδο Endnet στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.



Σχήμα 31 .Παραγόμενοι καθαροί στόχοι με την μέθοδο OSPAEU στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Με κόκκινο φαίνονται τα δεδομένα επαλήθευσης.

5.1.4. Αξιολόγηση ως προς την εύρεση των καθαρών στόχων.

Παρακάτω στους πίνακες 10-15 παρουσιάζονται οι υπολογισμένες φασματικές γωνίες ανάμεσα στους εξαγόμενους καθαρούς στόχους και τα δεδομένα επαλήθευσης για κάθε μέθοδο, σύνολο δεδομένων και πείραμα. Επίσης αναγράφεται και η μέση τιμή της φασματικής γωνίας όλων των καθαρών στόχων για κάθε πείραμα.

Πίνακας 10. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος DAEU.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.050	0.058	0.062	0.066	0.070	0.058	0.090	0.096	0.075	0.062
Grass	0.172	0.060	0.044	0.083	0.055	0.121	0.099	0.090	0.115	0.081
Tree	0.077	0.055	0.043	0.051	0.063	0.043	0.041	0.024	0.041	0.029
Roof	0.352	0.239	0.111	0.251	0.452	0.229	0.135	0.212	0.263	0.267
Average	0.163	0.103	0.065	0.113	0.160	0.113	0.091	0.105	0.124	0.110
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.036	0.309	0.046	0.038	0.047	0.057	0.040	0.050	0.047	0.061
Grass	0.044	0.175	0.035	0.035	0.069	0.030	0.045	0.051	0.031	0.031
Tree	0.150	0.134	0.148	0.101	0.147	0.136	0.185	0.106	0.080	0.128
Roof	0.179	0.363	0.200	0.064	0.060	0.362	0.056	0.237	0.247	0.331
Dirt	0.097	0.062	0.156	0.125	0.155	0.122	0.167	0.131	0.172	0.133
Average	0.101	0.175	0.117	0.073	0.096	0.142	0.099	0.115	0.115	0.137
Samson	1	2	3	4	5	6	7	8	9	10
Soil	0.015	0.029	0.166	0.018	0.024	0.042	0.021	0.037	0.022	0.011
Water	0.039	0.098	0.057	0.067	0.043	0.058	0.123	0.044	0.033	0.077
Tree	0.039	0.035	0.065	0.053	0.028	0.053	0.039	0.037	0.043	0.038
Average	0.031	0.054	0.096	0.046	0.032	0.051	0.061	0.039	0.033	0.042

Πίνακας 11. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος SIDAEU.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.067	0.074	0.069	0.052	0.102	0.075	0.087	0.069	0.092	0.087
Grass	0.130	0.105	0.110	0.134	0.174	0.078	0.073	0.161	0.102	0.204
Tree	0.043	0.038	0.030	0.034	0.108	0.053	0.047	0.029	0.053	0.052
Roof	0.376	0.276	0.444	0.333	0.098	0.215	0.134	0.373	0.262	0.408
Average	0.154	0.123	0.163	0.138	0.120	0.105	0.085	0.158	0.127	0.188
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.129	0.136	0.719	0.123	0.062	0.700	0.719	0.717	0.065	0.046
Grass	0.254	0.184	0.206	0.243	0.263	0.232	0.249	0.300	0.273	0.315
Tree	0.093	0.116	0.095	0.065	0.086	0.106	0.095	0.093	0.084	0.083
Roof	0.070	0.321	0.058	0.059	0.135	0.066	0.162	0.210	0.045	0.148
Dirt	0.698	0.139	0.142	0.673	0.665	0.133	0.056	0.047	0.646	0.292
Average	0.249	0.179	0.244	0.233	0.242	0.247	0.256	0.273	0.223	0.177
Samson	1	2	3	4	5	6	7	8	9	10
Soil	0.074	0.150	0.134	0.110	0.140	0.143	0.128	0.086	0.143	0.117

Water	0.080	0.097	0.067	0.122	0.050	0.067	0.053	0.086	0.060	0.065
Tree	0.063	0.066	0.029	0.039	0.061	0.051	0.093	0.049	0.039	0.052
Average	0.072	0.104	0.077	0.090	0.084	0.087	0.091	0.074	0.081	0.076

Πίνακας 12. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος ΜΤΑΕΥ.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.063	0.068	0.073	0.059	0.073	0.096	0.068	0.053	0.080	0.063
Grass	0.050	0.051	0.069	0.083	0.062	0.068	0.067	0.093	0.049	0.053
Tree	0.039	0.048	0.049	0.049	0.047	0.043	0.054	0.051	0.046	0.062
Roof	0.140	0.156	0.112	0.109	0.139	0.141	0.085	0.129	0.095	0.109
Average	0.073	0.081	0.076	0.075	0.080	0.087	0.068	0.081	0.068	0.072
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.034	0.042	0.045	0.054	0.047	0.056	0.067	0.046	0.041	0.068
Grass	0.102	0.098	0.098	0.122	0.101	0.082	0.076	0.099	0.075	0.091
Tree	0.133	0.126	0.115	0.139	0.139	0.135	0.136	0.150	0.130	0.123
Roof	0.072	0.057	0.090	0.061	0.103	0.120	0.110	0.092	0.107	0.051
Dirt	0.132	0.156	0.107	0.103	0.109	0.129	0.147	0.112	0.135	0.116
Average	0.095	0.096	0.091	0.096	0.100	0.104	0.107	0.100	0.098	0.089
Samson	1	2	3	4	5	6	7	8	9	10
Soil	0.031	0.054	0.043	0.038	0.050	0.029	0.040	0.045	0.026	0.034
Water	0.060	0.053	0.055	0.052	0.057	0.056	0.057	0.054	0.057	0.058
Tree	0.044	0.048	0.044	0.062	0.048	0.039	0.062	0.068	0.049	0.048
Average	0.045	0.051	0.047	0.051	0.051	0.041	0.053	0.056	0.044	0.047

Πίνακας 13. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος CNNAEU.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.0597	0.0538	0.0561	0.0600	0.0558	0.0636	0.0556	0.0511	0.0542	0.0553
Grass	0.0412	0.0438	0.0452	0.0476	0.0511	0.0382	0.0371	0.0365	0.0479	0.0447
Tree	0.0411	0.0349	0.0511	0.0323	0.0432	0.0361	0.0359	0.0350	0.0406	0.0369
Roof	0.0508	0.0447	0.0453	0.0366	0.0253	0.0525	0.0411	0.0424	0.0325	0.0380
Average	0.048	0.044	0.049	0.044	0.044	0.048	0.042	0.041	0.044	0.044
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.0409	0.0383	0.0412	0.0348	0.0373	0.0366	0.0471	0.0358	0.0359	0.0357
Grass	0.0648	0.0675	0.0685	0.0618	0.0536	0.0713	0.0640	0.0671	0.0662	0.0590
Tree	0.1246	0.1309	0.1397	0.1394	0.1415	0.1288	0.1428	0.01266	0.1382	0.1468
Roof	0.0457	0.0341	0.0470	0.0346	0.0322	0.0615	0.0367	0.0387	0.0660	0.0450
Dirt	0.0863	0.1063	0.1110	0.1063	0.0947	0.0855	0.1061	0.1032	0.1213	0.1060
Average	0.072	0.075	0.081	0.075	0.072	0.077	0.079	0.074	0.086	0.078
Samson	1	2	3	4	5	6	7	8	9	10
Soil	0.0361	0.0350	0.0421	0.0359	0.0396	0.0390	0.0341	0.0436	0.0425	0.0419

Tree	0.0418	0.0402	0.0416	0.0429	0.0413	0.0428	0.0411	0.0405	0.0425	0.0414
Water	0.0977	0.1105	0.1214	0.1109	0.1157	0.1129	0.0914	0.0940	0.1033	0.1104
Average	0.059	0.062	0.068	0.063	0.066	0.065	0.056	0.059	0.063	0.065

Πίνακας 14. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος Endnet.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.279	0.105	0.083	0.080	0.058	0.102	0.060	0.077	0.082	0.064
Grass	0.044	0.056	0.044	0.045	0.049	0.049	0.050	0.044	0.063	0.067
Tree	0.109	0.119	0.098	0.099	0.092	0.115	0.092	0.109	0.096	0.119
Roof	0.105	0.400	0.294	0.386	0.224	0.406	0.225	0.234	0.273	0.223
Average	0.134	0.170	0.130	0.153	0.106	0.168	0.107	0.116	0.129	0.118
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.057	0.049	0.074	0.052	0.087	0.094	0.088	0.047	0.093	0.076
Grass	0.084	0.062	0.064	0.095	0.058	0.057	0.078	0.086	0.057	0.067
Tree	0.102	0.089	0.086	0.085	0.072	0.080	0.100	0.108	0.088	0.084
Roof	0.138	0.101	0.267	0.082	0.181	0.223	0.087	0.087	0.032	0.040
Dirt	0.034	0.030	0.028	0.031	0.038	0.026	0.048	0.029	0.039	0.033
Average	0.083	0.066	0.104	0.069	0.087	0.096	0.080	0.071	0.062	0.060
Samson	1	2	3	4	5	6	7	8	9	10
Soil	0.192	0.094	0.203	0.012	0.195	0.018	0.013	0.148	0.185	0.193
Water	0.540	0.032	0.634	0.999	0.641	0.128	0.025	0.021	0.559	0.033
Tree	0.028	0.090	0.024	0.029	0.027	0.071	1.020	0.974	0.034	0.588
Average	0.254	0.072	0.284	0.347	0.288	0.072	0.352	0.381	0.259	0.271

Πίνακας 15. Φασματική απόσταση των παραγόμενων στόχων στα δεδομένα επαλήθευσης, μέθοδος OSPAEU.

Urban4	1	2	3	4	5	6	7	8	9	10
Asphalt	0.299	0.308	0.314	0.257	0.286	0.264	0.276	0.254	0.280	0.271
Grass	0.311	0.290	0.286	0.207	0.141	0.133	0.145	0.327	0.130	0.326
Tree	0.230	0.210	0.215	0.415	0.037	0.040	0.031	0.220	0.029	0.227
Roof	0.328	0.310	0.323	0.585	0.263	0.251	0.285	0.328	0.277	0.302
Average	0.292	0.279	0.285	0.366	0.102	0.172	0.104	0.282	0.179	0.281
Urban5	1	2	3	4	5	6	7	8	9	10
Asphalt	0.315	0.306	0.288	0.398	0.636	0.290	0.368	0.313	0.295	0.310
Grass	0.101	0.093	0.096	0.096	0.097	0.096	0.094	0.093	0.099	0.097
Tree	0.088	0.091	0.083	0.073	0.083	0.085	0.086	0.081	0.069	0.084
Roof	0.236	0.247	0.262	0.210	0.229	0.264	0.252	0.184	0.253	0.263
Dirt	0.128	0.133	0.218	0.241	0.238	0.115	0.217	0.207	0.153	0.141
Average	0.173	0.174	0.190	0.204	0.257	0.170	0.203	0.176	0.174	0.179
Samson	1	2	3	4	5	6	7	8	9	10

Soil	0.103	0.105	0.106	0.179	0.662	0.101	0.106	0.101	0.102	0.104
Water	0.179	0.202	0.138	0.181	0.121	0.223	0.179	0.183	0.180	0.099
Tree	0.050	0.087	0.050	0.093	0.074	0.064	0.083	0.056	0.054	0.084
Average	0.110	0.131	0.098	0.151	0.285	0.129	0.122	0.113	0.112	0.096

Συνοψίζοντας τους προηγούμενους πίνακες και διαγράμματα φαίνεται πως κατά μέσο όρο οι μέθοδοι MTAEU και CNNAEU δώσανε καθαρούς στόχους πιο κοντά στους πραγματικούς. Η μέθοδος DAEU δίνει πολύ καλά αποτελέσματα στην εικόνα Samson που έχει μόνο τρεις καθαρούς στόχους και δυσκολεύεται να εξάγει πιο σωστά συμπεράσματα στην Urban που περιγράφεται από τέσσερις και πέντε. Στους πίνακες 16-17 η μέθοδος Endnet φαίνεται πως παρουσιάζει καλά αποτελέσματα και ως προς την καλύτερη αλλά και την μέση τιμή όσον αφορά την εικόνα Urban με πέντε καθαρούς στόχους. Όμως από τα προηγούμενα διαγράμματα φαίνεται πως κάποιες φορές αποτυγχάνει να εξάγει τον καθαρό στόχο της σκεπής. Χειρότερα αποδίδουν οι μέθοδοι Endnet και OSPAEU.

Πίνακας 16. Καλύτερη φασματική γωνία των παραγόμενων στόχων και των δεδομένων επαλήθευσης για κάθε σύνολο δεδομένων και μέθοδο.

Best	DAEU	SIDAEU	MTAEU	CNNAEU	Endnet	OSPAEU
Samson	0.031	0.072	0.041	0.046	0.072	0.096
Urban4	0.091	0.085	0.068	0.041	0.116	0.102
Urban5	0.099	0.177	0.091	0.072	0.060	0.170

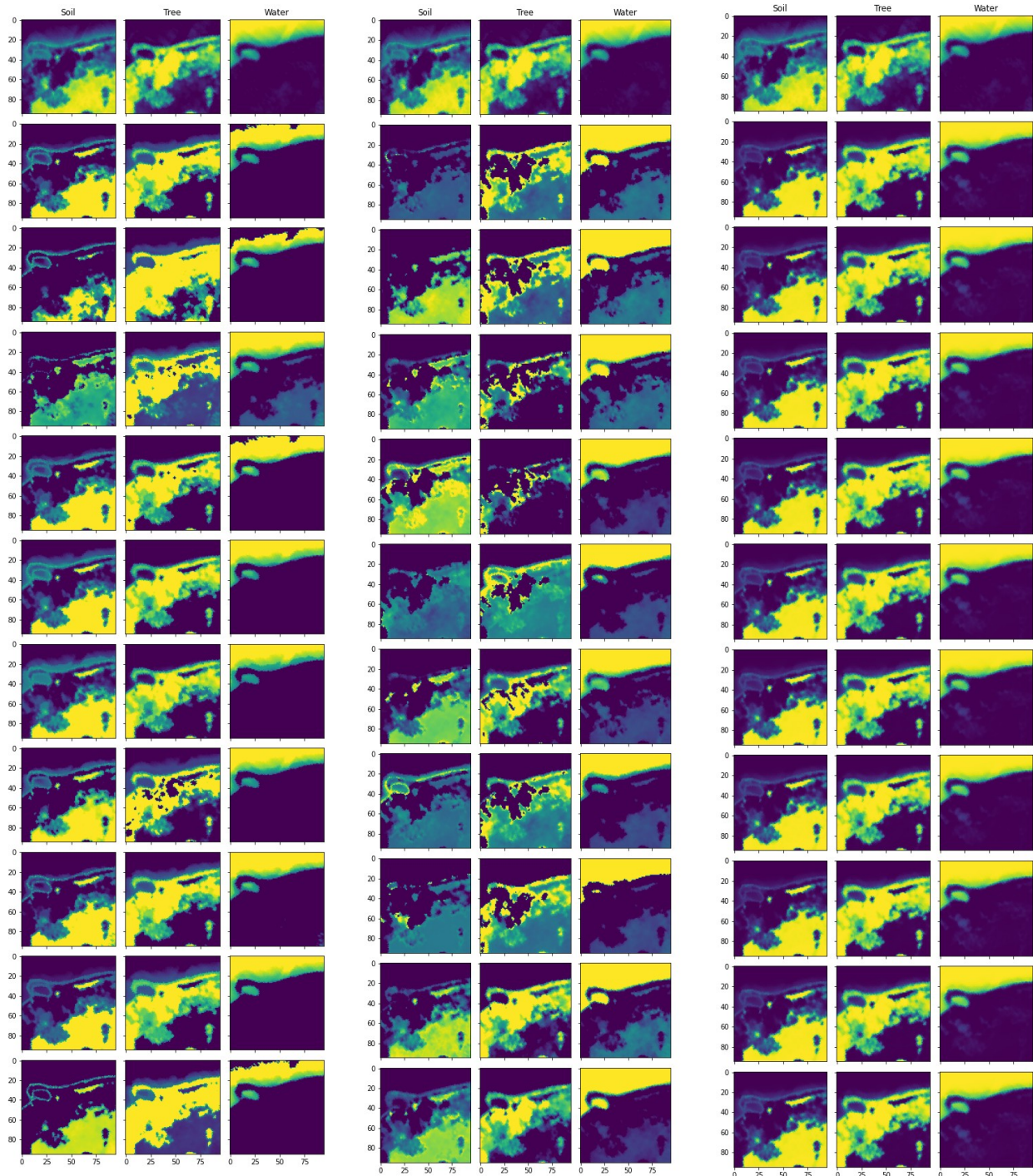
Πίνακας 17. Μέση φασματική γωνία των παραγόμενων στόχων και των δεδομένων επαλήθευσης για κάθε σύνολο δεδομένων και μέθοδο.

Average	DAEU	SIDAEU	MTAEU	CNNAEU	Endnet	OSPAEU
Samson	0,0485	0,0836	0,0486	0,0626	0,258	0,1347
Urban4	0,1147	0,1361	0,0761	0,0448	0,1331	0,2342
Urban5	0,117	0,2323	0,0976	0,0769	0,0778	0,19

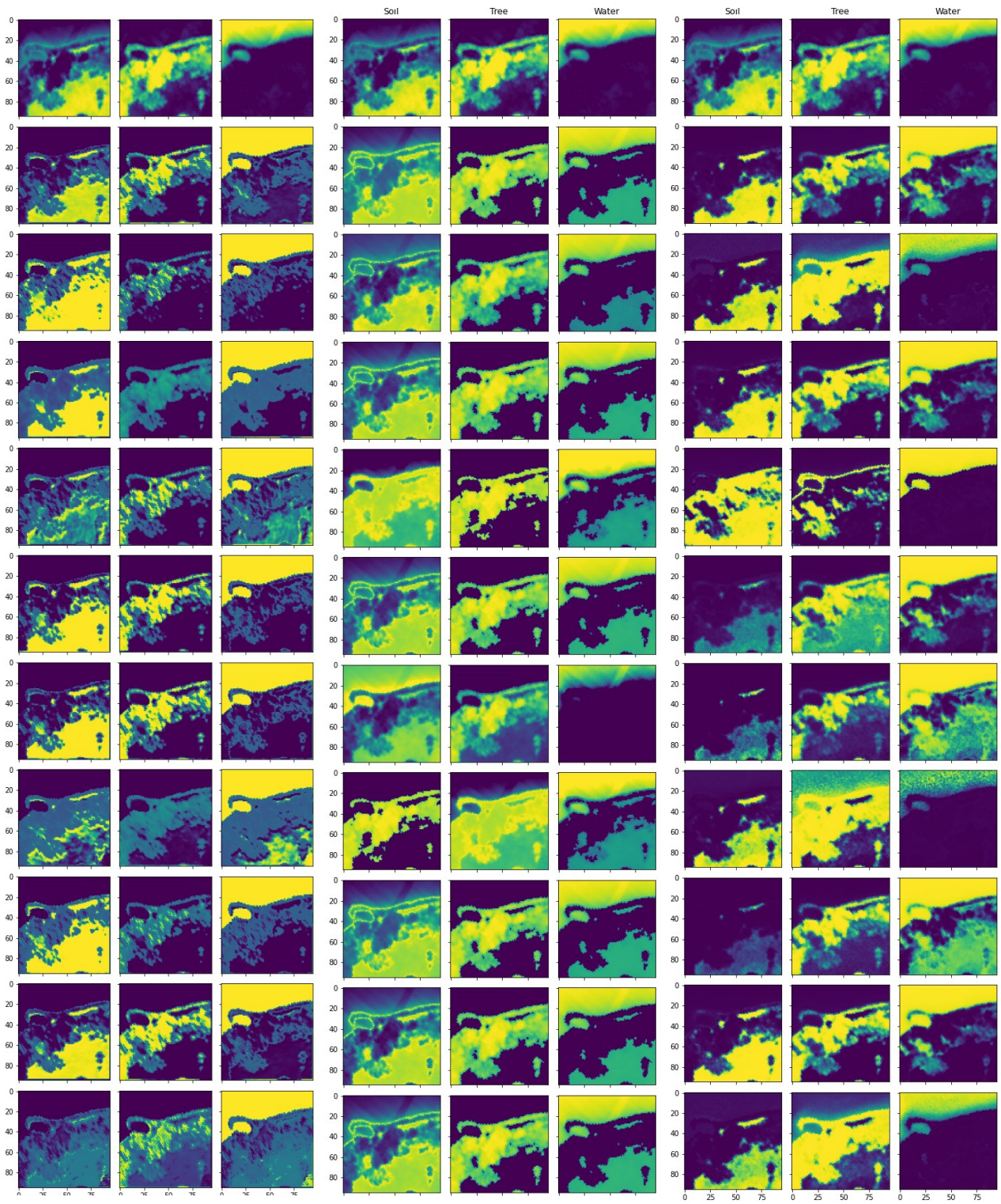
5.2. Αφθονίες

Για τις αφθονίες έγινε μόνο ποιοτική αξιολόγηση. Παρακάτω παρουσιάζονται οι εξαγόμενοι χάρτες αφθονίας για κάθε σύνολο δεδομένων, μέθοδο και πείραμα σε κλίμακα από μωβ έως κίτρινο με κίτρινο να απεικονίζονται οι υψηλές τιμές. Εδώ, κύριο κριτήριο αποτελεί η ομοιότητα των χαρτών με τα δεδομένα επαλήθευσης και συγκεκριμένα υφή τους.

5.2.1. Samson

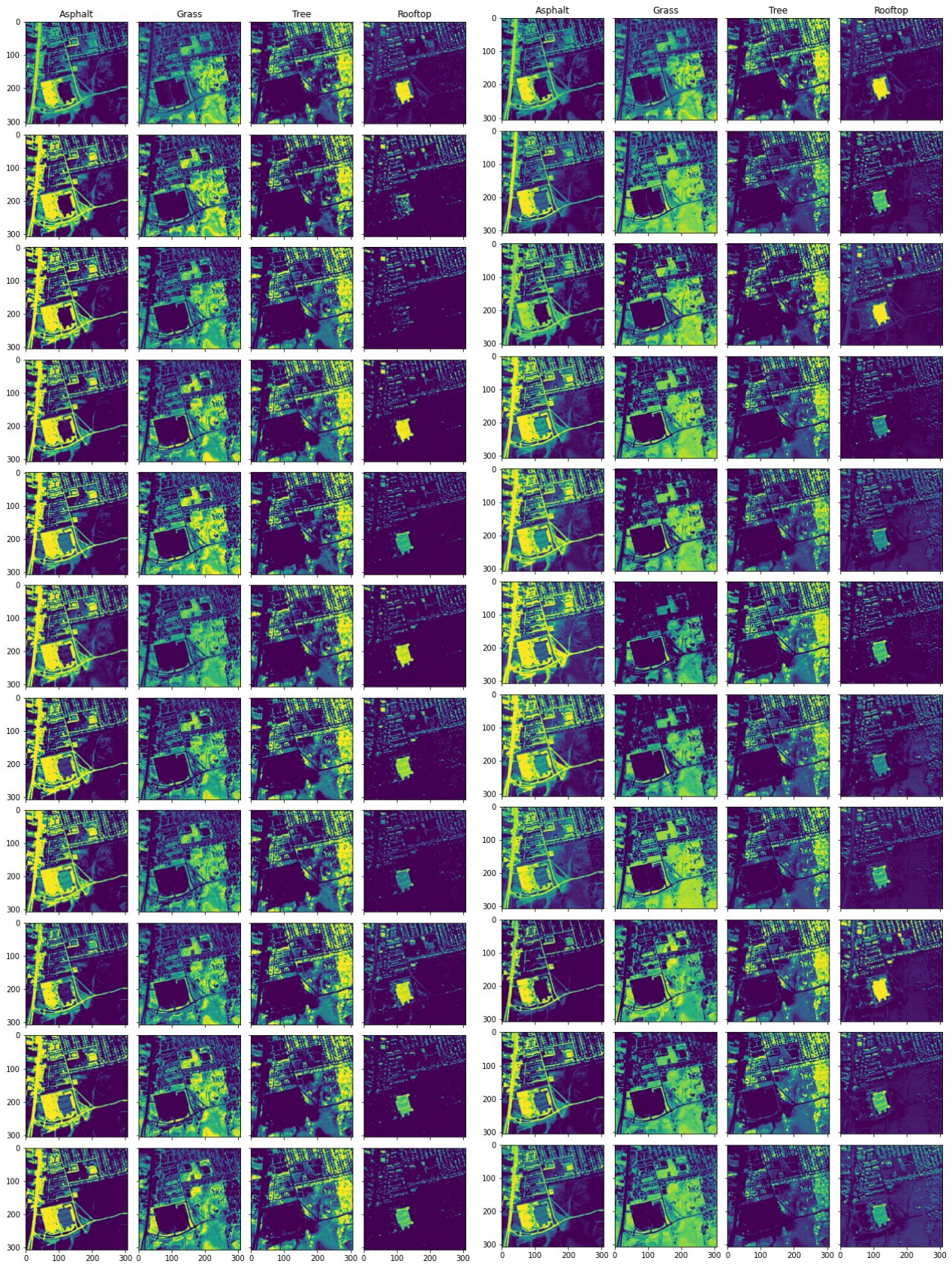


Σχήμα 32 .Παραγόμενες αφθονίες με την μέθοδο DAEU (αριστερά), SIDAEU (κέντρο) και MTAEU (δεξιά) στο σύνολο δεδομένων Samson. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης. Στήλες με σειρά εμφάνισης: Έδαφος, δέντρο και νερό.

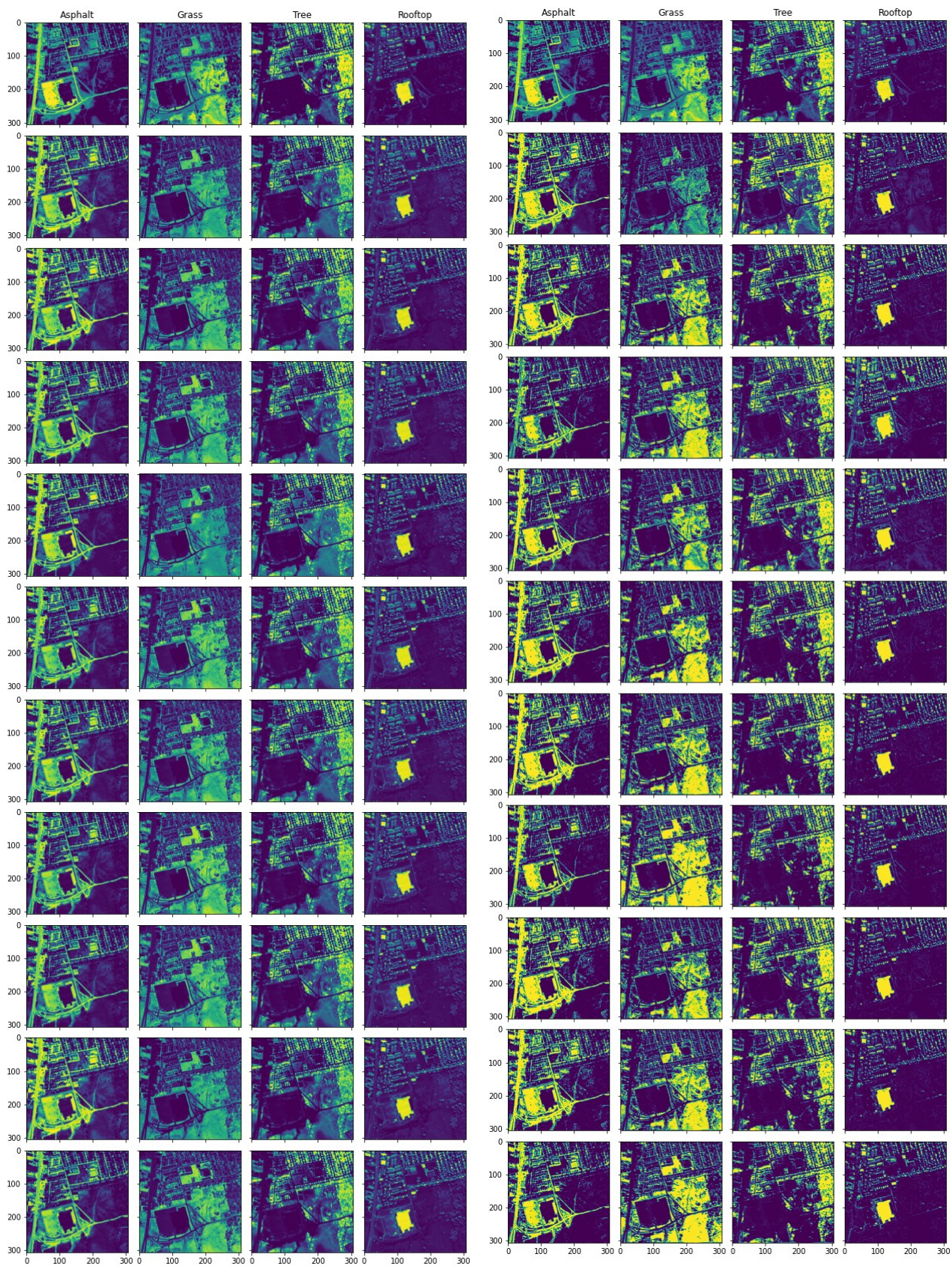


Σχήμα 33 .Παραγόμενες αφθονίες με την μέθοδο CNNAEU (αριστερά), Endnet (κέντρο) και OSPAEU (δεξιά) στο σύνολο δεδομένων Satson. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης. Στήλες με σειρά εμφάνισης: Έδαφος, δέντρο και νερό.

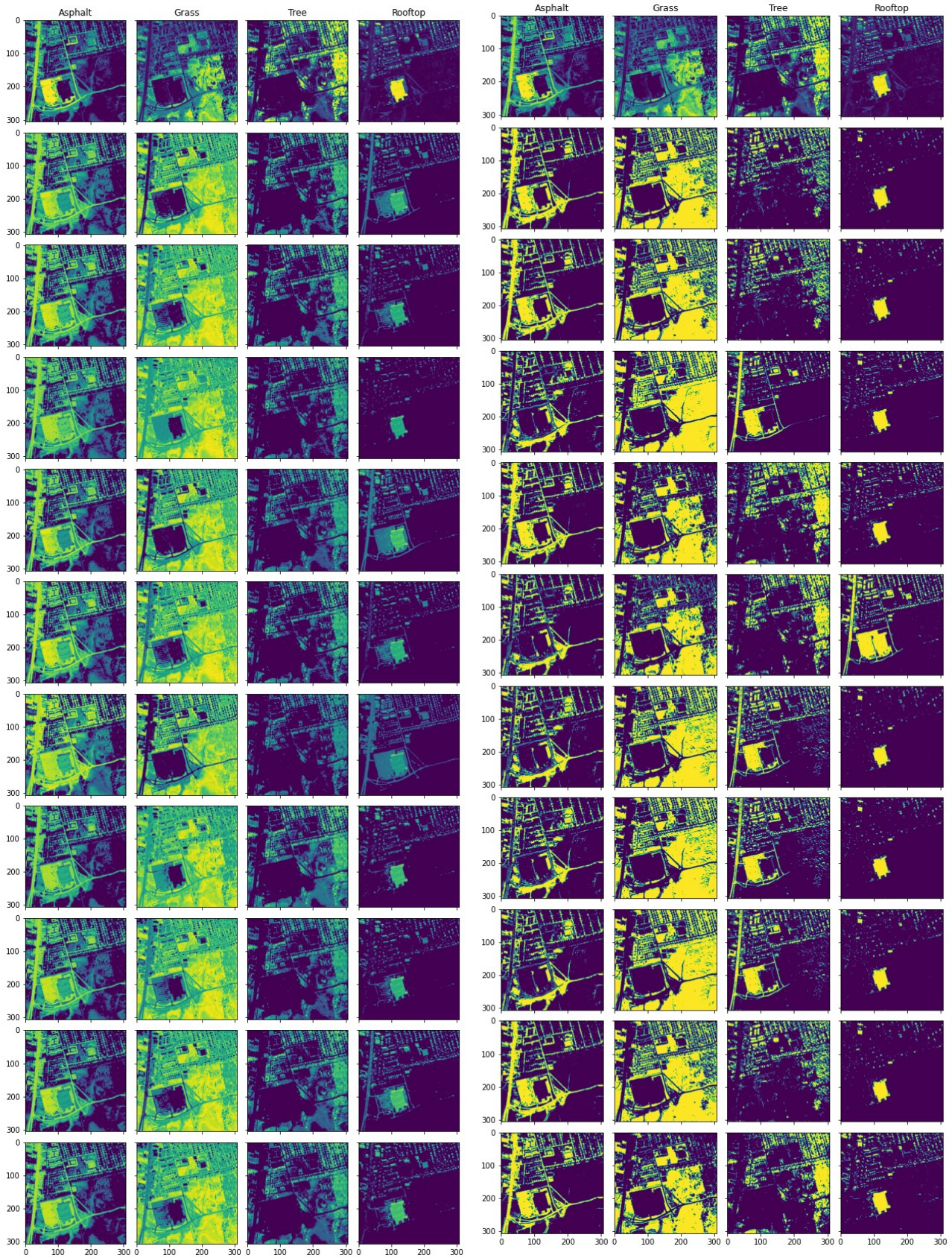
5.2.2. Urban με τέσσερις καθαρούς στόχους



Σχήμα 34 .Παραγόμενες αφηθινίες με την μέθοδο DAEU (αριστερά) και SIDAEU (δεξιά) στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.

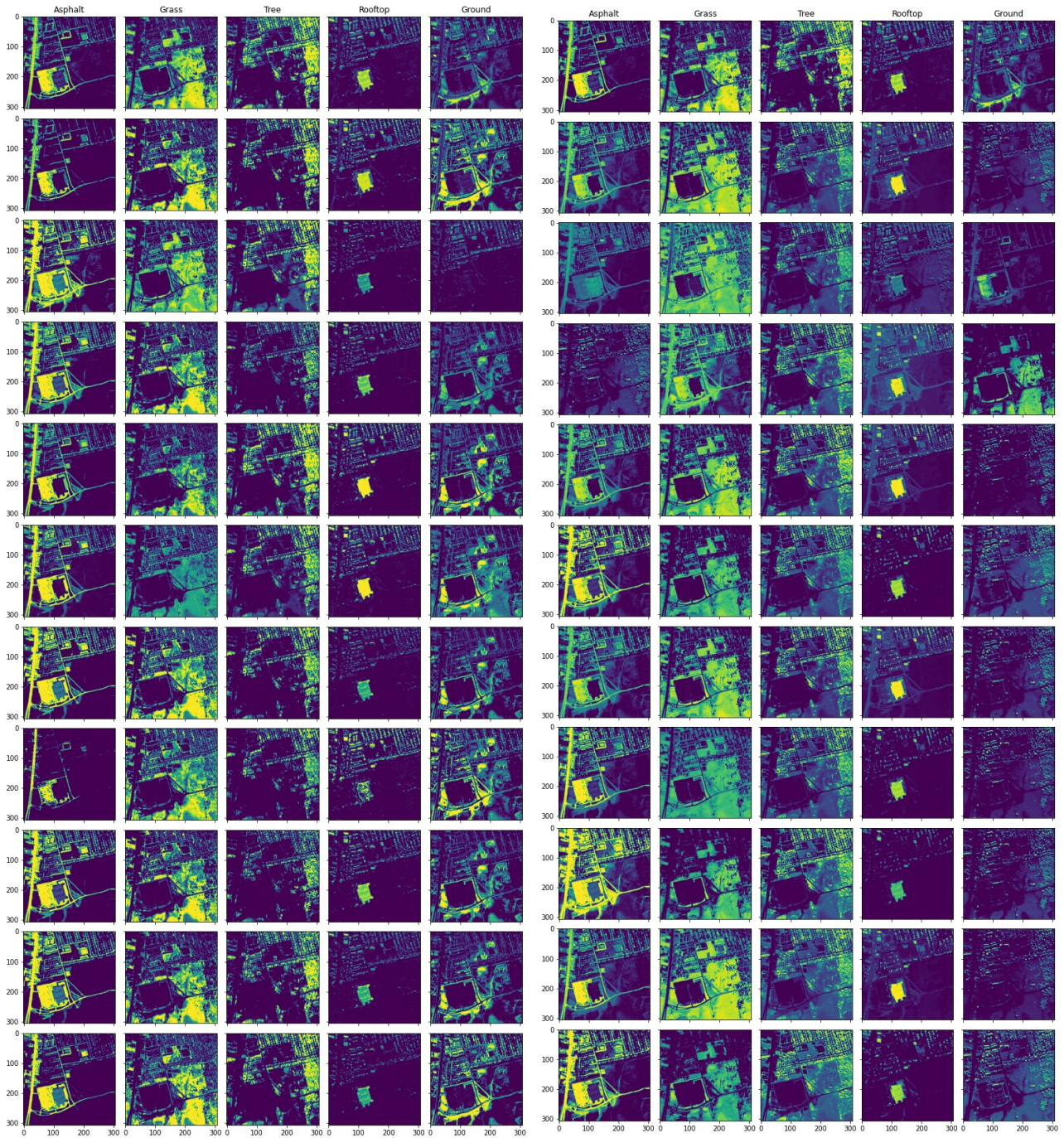


Σχήμα 35 .Παράγόμενες αφθονίες με την μέθοδο MTAU (αριστερά) και CNNAEU (δεξιά) στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.

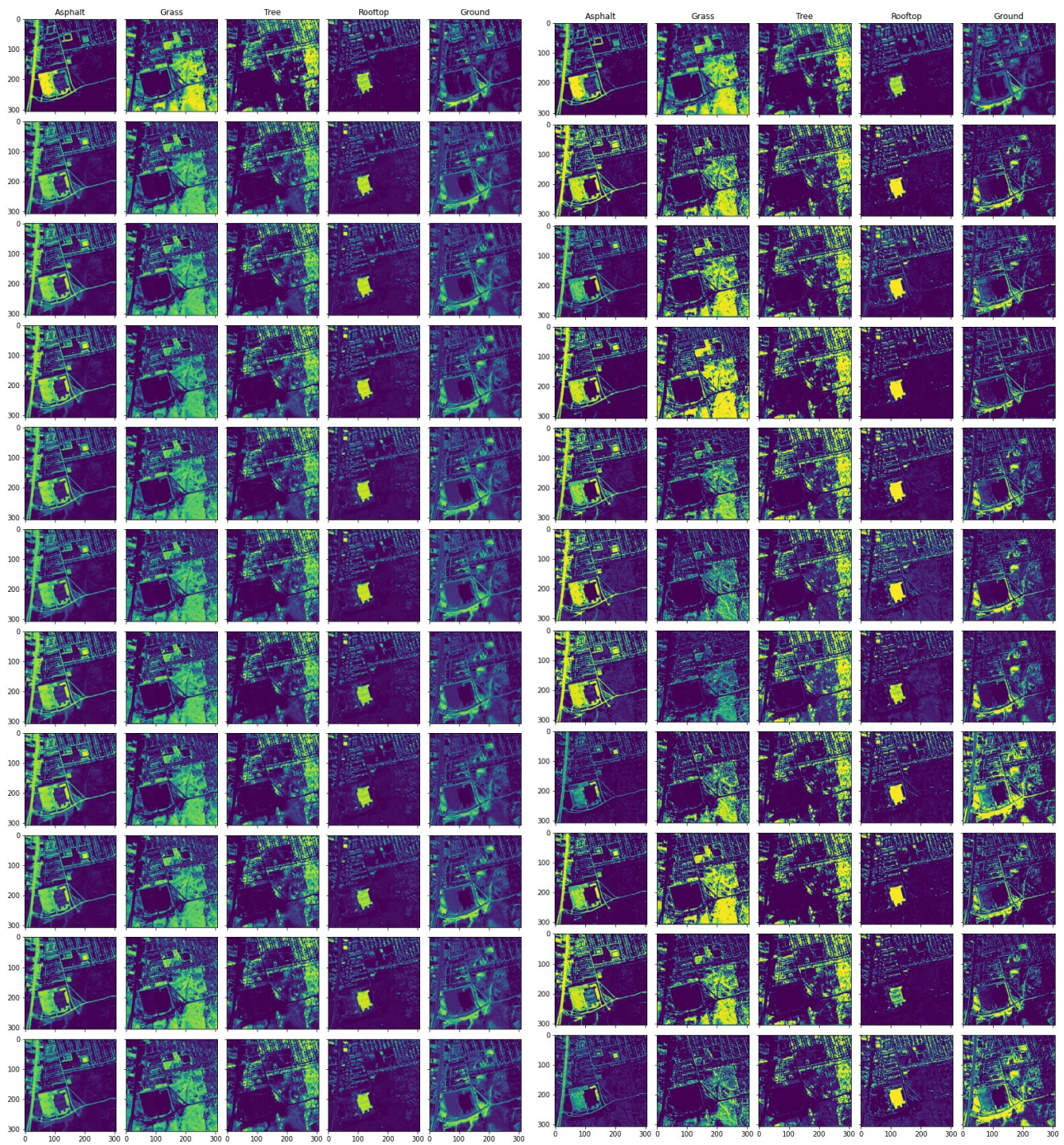


Σχήμα 36 .Παραγόμενες αφθονίες με την μέθοδο Endnet (αριστερά) και OSPAEU (δεξιά) στο σύνολο δεδομένων Urban με τέσσερις καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.

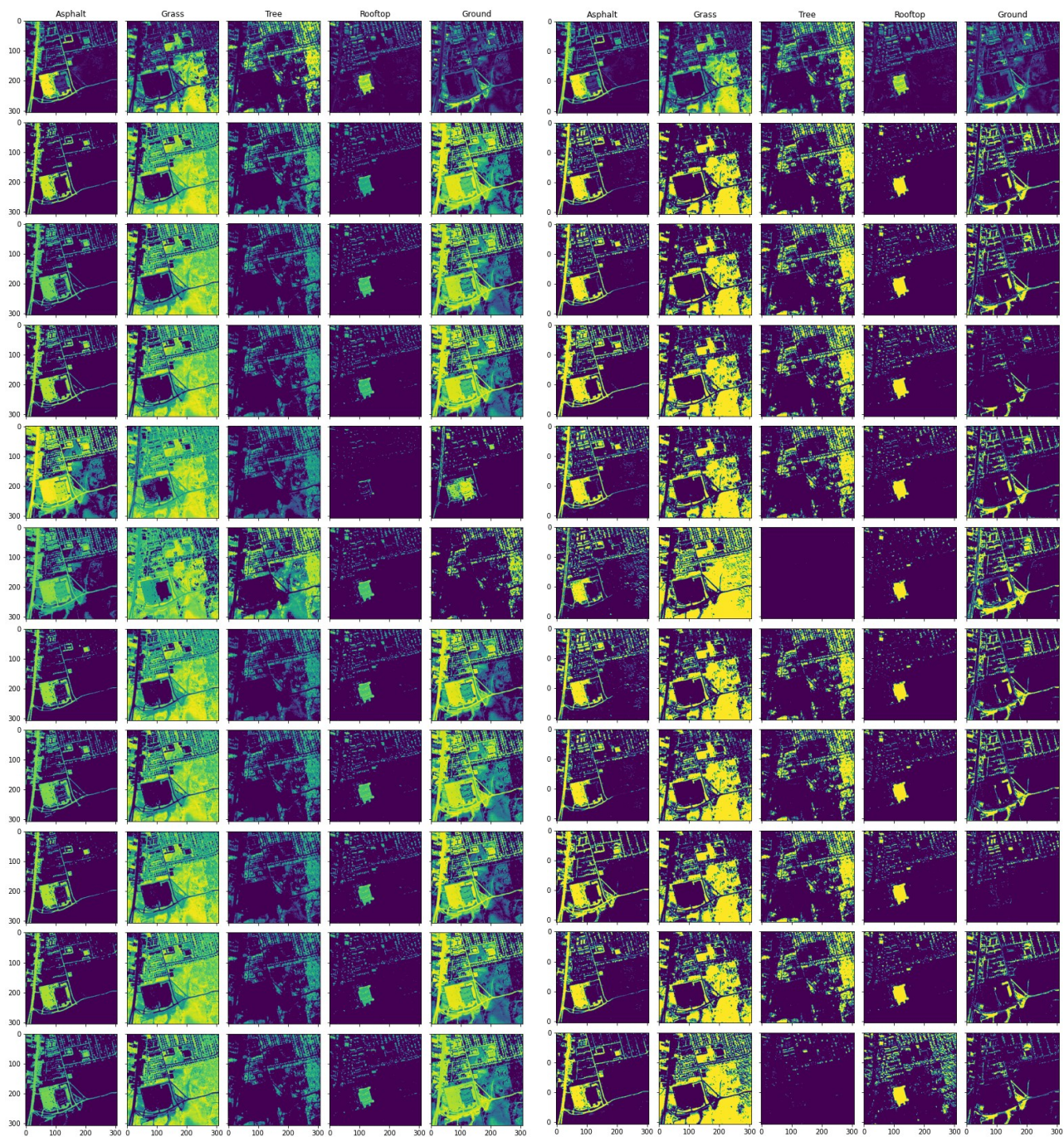
5.2.3. Urban με πέντε καθαρούς στόχους



Σχήμα 37 .Παραγόμενες αφθονίες με την μέθοδο DAEU (αριστερά) και SIDAEU (δεξιά) στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.



Σχήμα 38 .Παραγόμενες αφθονίες με την μέθοδο MTAEU (αριστερά) και CNNAEU (δεξιά) στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.



Σχήμα 39 .Παραγόμενες αφθονίες με την μέθοδο Endnet (αριστερά) και OSPAEU (δεξιά) στο σύνολο δεδομένων Urban με πέντε καθαρούς στόχους. Η πρώτη γραμμή είναι τα δεδομένα επαλήθευσης.

Από τους προηγούμενους χάρτες φαίνεται μια γενική τάση των μεθόδων αυτών να προτιμούν ακραίες τιμές. Αυτό το φαινόμενο είναι ιδιαίτερα εμφανές στις μεθόδους OSPAEU και Endnet μάλλον γιατί εισάγουν το sparsity με πολλούς τρόπους και πιο έντονα από τις άλλες. Μάλιστα η μέθοδος OSPAEU φαίνεται πως σε μερικά πείραμα απέτυχε εντελώς να εξάγει τους χάρτες αφθονιών του δέντρου στο σύνολο δεδομένων Urban με τέσσερις και πέντε καθαρούς στόχους και τον χάρτη αφθονίας του νερού στο σύνολο δεδομένων Samson. Η μέθοδος MTAEU παρουσιάζει ιδιαίτερα καλούς και σταθερούς χάρτες αφθονιών, πιο καλούς ακόμα και από την μέθοδο CNNAEU που σε γενικές γραμμές φαίνεται πως εξάγει ελαφρώς πιο ακριβείς καθαρούς στόχους. Το δίκτυο CNNAEU φαίνεται πως εμφανίζει μεγαλύτερη διακύμανση και ακραίες τιμές από το CNNAEU, ειδικά στο σύνολο δεδομένων Samson.

5.3. Συμπεράσματα

Για την σύγκριση των μεθόδων ελέγχθηκε η ποιότητα των καθαρών στόχων και των αφθονιών ξεχωριστά. Για τους καθαρούς στόχους ως κριτήριο θεωρήθηκε η ομοιότητα με τα δεδομένα επαλήθευσης εκφρασμένη με την φασματική γωνία ενώ οι αφθονίες εξετάστηκαν μόνο ποιοτικά. Να σημειωθεί πως η ποιότητα των χαρτών αφθονίας είναι άμεσα συνδεδεμένη και αντικατοπτρίζει την ποιότητα των εξαγόμενων καθαρών στόχων. Σημαντική είναι επίσης και η σταθερότητα των μεθόδων, η οποία ελέγχθηκε ποσοτικά με την σειρά των δέκα πειραμάτων και οπτικοποιείται στα αντίστοιχα διαγράμματα.

Σε γενικές γραμμές φαίνεται πως οι μέθοδοι CNNAEU και MTAEU έδωσαν τα καλύτερα και πιο συνεπή αποτελέσματα. Και οι δύο αυτές μέθοδοι χρησιμοποιούν με κάποιο τρόπο την χωρική πληροφορία της εικόνας και δεν θεωρούν κάθε εικονοστοιχείο χωρικά ανεξάρτητο. Όλες οι μέθοδοι φαίνεται πως δίνουν λιγότερο ακριβή αποτελέσματα με την αύξηση του αριθμού των καθαρών στόχων (εκτός από την μέθοδο CNNAEU στην εικόνα Urban με 4 καθαρούς στόχους). Επίσης κάποιοι καθαροί στόχοι όπως το έδαφος (και μερικές φορές το νερό) στο σύνολο δεδομένων Samson και η σκεπή στο Urban εξάγονται με μεγαλύτερη δυσκολία ενώ οι καθαροί στόχοι που σχετίζονται με την βλάστηση πολύ πιο εύκολα. Μάλιστα στην πλειοψηφία των περιπτώσεων οι μέθοδοι OSPAEU και Endnet σχεδόν αποτυγχάνουν να εξάγουν τον καθαρό στόχο που αντιστοιχεί στις σκεπές. Όσον αφορά την σταθερότητα, οι μέθοδοι MTAEU, CNNAEU και DAEU παρουσιάζουν τα πιο σταθερά αποτελέσματα. Οι υπόλοιπες μέθοδοι, ειδικά στους πιο δύσκολους καθαρούς στόχους εμφανίζουν μεγάλη διακύμανση.

Όπως αναφέρθηκε οι παραγόμενες αφθονίες είναι άμεσα εξαρτώμενες από την ποιότητα των παραγόμενων καθαρών στόχων. Δηλαδή όσο πιο ακριβείς είναι οι καθαροί στόχοι, τόσο πιο ακριβείς είναι και οι αντίστοιχες αφθονίες. Ένα γενικό πρόβλημα που παρατηρείται σε αυτές είναι πως εμφανίζουν μια προτίμηση στις ακραίες τιμές, δηλαδή προκύπτουν πιο sparse από αυτές των δεδομένων επαλήθευσης. Αυτό είναι ιδιαίτερα έντονο στις μεθόδους OSPAEU και Endnet, ίσως λόγω της μεγάλης έμφασης των μεθόδων αυτών στην εφαρμογή του sparsity. Η μέθοδος με τις πιο ποιοτικές και σταθερές αφθονίες φαίνεται πως είναι η MTAEU.

Όλα αυτά τα αποτελέσματα φαίνεται πως συνάδουν επί το πλείστον με αυτά της βιβλιογραφίας, τουλάχιστον όσον αφορά τις μεθόδους [6]-[9]. Οι μέθοδοι OSPAEU και Endnet δεν φαίνεται να δίνουν καλά αποτελέσματα ή τουλάχιστον αποτελέσματα που να ανταγωνίζονται τις υπόλοιπες μεθόδους βαθιάς μάθησης.

Μελλοντικά, φαίνεται χρήσιμη η αξιολόγηση των καινοτομιών που θέτει η κάθε μέθοδος ως ξεχωριστά στοιχεία. Δηλαδή:

- η διερεύνηση διαφορετικών συναρτήσεων κόστους (Endnet και OSPAEU),
- η εισαγωγή της χωρικής πληροφορίας με συνελκτικά δίκτυα (CNNAEU) ή multisk learning (MTAEU),
- η αρχικοποίηση των τιμών των βαρών με βάση της ακραίες τιμές (OSPAEU), την μέθοδο VCA (Endnet) ή τυχαία (υπόλοιπες μέθοδοι)
- και η εφαρμογή των περιορισμών του γραμμικού μοντέλου με softmax (OSPAEU, CNNAEU), κανονικοποίηση L_1 (Endnet) ή κανονικοποίηση με βάση το άθροισμα (υπόλοιπες μέθοδοι).

7. Πρόγραμμα εκτέλεσης φασματικού διαχωρισμού

Για την χρήση όλων των αλγορίθμων που διερευνήθηκαν συντάχτηκε πρόγραμμα σε γλώσσα προγραμματισμού python. Το πρόγραμμα αυτό τρέχει σε command line και δέχεται μια εικόνα σε μορφή tiff ή σε μορφή .mat και παράγει γεωαναφερμένους (αν η εικόνα είναι σε μορφή tiff) χάρτες αφθονιών και ένα αρχείο .csv που περιέχει έναν πίνακα με στήλες τους καθαρούς στόχους. Οι παράμετροι με την σειρά που τις δέχεται το πρόγραμμα είναι (με παύλα δηλώνονται οι προαιρετικές παράμετροι) :

```
hsipath method -epochs -learning_rate -batch -ODM -n_endm -runs save
```

όπου

- *hsipath*: το path της υπερφασματικής εικόνας
- *method*: η μέθοδος φασματικού διαχωρισμού (MTAEU, SIDAEU, ENDNET, CNNAEU, DAEU, OSPAEU)
- *epochs*: αριθμός εποχών εκπαίδευσης
- *--learning rate*: ο ρυθμός μάθησης για τον αλγόριθμο οπισθοδιάδοσης
- *--batch*: ο αριθμός των εικονοστοιχείων που δέχεται ο αλγόριθμος πριν την ανανέωση των βαρών
- *- ODM*: επιλογή εκτέλεσης του αλγορίθμου για εκτίμηση του αριθμού των καθαρών στόχων (0,1)
- *--n_endm*: ο αριθμός των καθαρών στόχων αν δεν επιλεγεί η χρήση του ODM
- *--runs*: πόσες φορές να τρέξει ο αλγόριθμος
- *save*: το path αποθήκευσης των αποτελεσμάτων

8. Βιβλιογραφία

- [1]. Chan, Tsung-Han & Chi, Chong-Yung & Huang, Yu-Min & Ma, Wing-Kin. (2009). A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing. *IEEE Transactions on Signal Processing*. 57. 4418-4432. 10.1109/TSP.2009.2025802.
- [2]. Carmen Quintano, Alfonso Fernández-Manso, Yosio E. Shimabukuro & Gabriel Pereira (2012) Spectral unmixing, *International Journal of Remote Sensing*, 33:17, 5307-5340, DOI:10.1080/01431161.2012.661095
- [3]. N. Keshava and J. F. Mustard, "Spectral unmixing," in *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44-57, Jan. 2002, doi: 10.1109/79.974727.
- [4]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- [5]. Haykin, S. S. (2009). *Neural networks and learning machines*. Upper Saddle River, NJ: Pearson Education.
- [6]. B. Palsson, J. Sigurdsson, J. R. Sveinsson and M. O. Ulfarsson, "Hyperspectral Unmixing Using a Neural Network Autoencoder," in *IEEE Access*, vol. 6, pp. 25646-25656, 2018, doi: 10.1109/ACCESS.2018.2818280.
- [7]. F. Palsson, J. Sigurdsson, J. R. Sveinsson and M. O. Ulfarsson, "Neural network hyperspectral unmixing with spectral information divergence objective," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 755-758, doi: 10.1109/IGARSS.2017.8127062.
- [8]. B. Palsson, J. R. Sveinsson and M. O. Ulfarsson, "Spectral-Spatial Hyperspectral Unmixing Using Multitask Learning," in *IEEE Access*, vol. 7, pp. 148861-148872, 2019, doi: 10.1109/ACCESS.2019.2944072.
- [9]. B. Palsson, M. O. Ulfarsson and J. R. Sveinsson, "Convolutional Autoencoder for Spectral–Spatial Hyperspectral Unmixing," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 535-549, Jan. 2021, doi: 10.1109/TGRS.2020.2992743.
- [10]. S. Ozkan, B. Kaya and G. B. Akar, "EndNet: Sparse AutoEncoder Network for Endmember Extraction and Hyperspectral Unmixing," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 482-496, Jan. 2019, doi: 10.1109/TGRS.2018.2856929.
- [11]. Z. Dou, K. Gao, X. Zhang, H. Wang and J. Wang, "Hyperspectral Unmixing Using Orthogonal Sparse Prior-Based Autoencoder With Hyper-Laplacian Loss and Data-Driven Outlier Detection," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6550-6564, Sept. 2020, doi: 10.1109/TGRS.2020.2977819.
- [12]. Y. Qu and H. Qi, "uDAS: An Untied Denoising Autoencoder With Sparsity for Spectral Unmixing," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1698-1712, March 2019, doi: 10.1109/TGRS.2018.2868690.
- [13]. Veganzones, Miguel & Graña, Manuel. (2008). Endmember Extraction Methods: A Short Review. *Knowledge-Based Intelligent Information and Engineering Systems*. 400-407. 10.1007/978-3-540-85567-5_50.