

ΚΕΦΑΛΑΙΟ 04

ΕΝΤΑΣΗ ΧΩΡΙΚΟΥ ΣΗΜΕΙΑΚΟΥ ΠΡΟΤΥΠΟΥ

Στόχος του κεφαλαίου είναι η επισκόπηση μεθόδων της περιγραφικής στατιστικής για την ποσοτικοποίηση της χωρικής κατανομής της έντασης σημειακών συμβάντων (πλήθος ανά επιφάνεια) σε μια περιοχή μελέτης.

4.1. Ανάλυση Έντασης Συμβάντων

Η έννοια της έντασης έχει να κάνει με το πλήθος των παρατηρήσεων ανά κάποια μονάδα μέτρησης. Είναι δηλαδή η ποσοτικοποίηση της συχνότητας (πλήθος, αριθμός) συμβάντων ανά επιφάνεια περιοχής. Αν για παράδειγμα μελετάμε θέσεις δέντρων σε ένα οικόπεδο, τότε η ένταση του φαινομένου είναι το πλήθος των δέντρων δια το εμβαδό του οικοπέδου. Συγκρίνοντας δυο οικόπεδα ίδιου εμβαδού (παρανομαστής), η μεγαλύτερη ένταση θα εμφανιστεί στο οικόπεδο που έχει μεγαλύτερο πλήθος δέντρων. Αντιθέτως, συγκρίνοντας δυο οικόπεδα με ίδιο πλήθος δέντρων, η μεγαλύτερη ένταση θα εμφανιστεί στο οικόπεδο με το μικρότερο εμβαδό.

Η ανάλυση της χωρικής έντασης ενός φαινομένου, έχει να κάνει με τη μελέτη της χωρικής κατανομής της έντασης συμβάντων και ταυτοποίηση τυχών θέσεων (ή υποπεριοχών) όπου πιθανώς η ένταση εμφανίζεται αυξημένη. Επίσης, η ανάλυση μπορεί να εστιάζει στη διερεύνηση πιθανών συσχετίσεων μεταξύ έντασης συμβάντων και άλλων χωρικών μεταβλητών, όπως π.χ. απόσταση από κάποια πηγή μόλυνσης ή απόσταση από κάποιο γνωστό σημείο.

Σε περιπτώσεις όμως όπου μελετάμε χωρικά συνεχείς μεταβλητές (π.χ. θερμοκρασία, υγρασία) η ανάλυση θα εστιάσει στον χαρακτηρισμό της χωρικής μεταβολής των χαρακτηριστικών που μετρήθηκαν σε συγκεκριμένες θέσεις.

Γενικά, η ανάλυση της χωρικής έντασης δεν υπεισέρχεται στη μελέτη της σχετικής θέσης των συμβάντων, δηλαδή στην ανάλυση της χωρικής ελκτικότητας ή απώθησης μεταξύ συμβάντων. Η ανάλυση της χωρικής έντασης ποσοτικοποιεί και αναδεικνύει τις αλλαγές της έντασης και δεν εστιάζει στις αιτίες που μπορεί να προκαλούν αυξημένες/μειωμένες τιμές χωρικής έντασης.

4.2. Βασικές έννοιες έντασης

Εστω ένα σύνολο από σημειακά συμβάντα, το οποίο περιλαμβάνει N θέσεις συμβάντων $\{u_i, i = 1, \dots, N\}, u_i \in D$ που καταγράφηκαν σε μια περιοχή μελέτης όπου u_i είναι ένα διάνυσμα συντεταγμένων της θέσης του συμβάντος i . Αν π.χ. έχουμε έναν δισδιάστατο χώρο, τότε κάθε θέση περιλαμβάνει δυο συντεταγμένες $u_i = (x_i, y_i)$ και η θέση ανήκει στην περιοχή μελέτης D . Ο υπολογισμός της χωρικής έντασης συμβάντων μπορεί να γίνει σε δυο διαφορετικά επίπεδα. Μπορεί δηλαδή να γίνει είτε σε γενικό είτε σε τοπικό επίπεδο. Η **γενικευμένη ένταση** λ εκτιμάται ως η διαίρεση του πλήθους θέσεων δια το εμβαδό της περιοχής μελέτης:

$$\hat{\lambda} = \frac{N}{|D|}$$

Εξίσωση 4.1: Γενικευμένη ένταση θέσεων

Σε δεύτερο επίπεδο, η τοπική ένταση θέσεων $\lambda(s)$ υπολογίζεται ως η διαίρεση του πλήθους θέσεων σε κάποια υποδιαίρεση του χώρου s (γειτονιά, ζώνη, φατνίο) προς το εμβαδό $|s|$ αυτής της υπό-περιοχής.

$$\lambda(s) = \frac{n(s)}{|d(s)|}$$

Εξίσωση 4.2: Τοπική ένταση θέσεων σε ένα υποστήριγμα (s =support) της περιοχής μελέτης

4.3. Συχνότητα - Ένταση - Πυκνότητα

Μελετώντας τις θέσεις σημείων σε μια περιοχή μελέτης, πολλές φορές χωρίζουμε σε μικρότερες υποπεριοχές για να κάνουμε υπολογισμούς και τελικά να εξάγουμε συμπεράσματα σχετικά με τη χωρική κατανομή των συμβάντων. Πριν προχωρήσουμε, είναι απαραίτητο να αποσαφηνιστούν κάποιοι όροι οι οποίοι σχετίζονται με την καταμέτρηση θέσεων σε μια περιοχή/υπό-περιοχή μελέτης. **Συχνότητα** (counts ή absolute frequency) είναι το πλήθος των παρατηρήσεων σε μια περιοχή. **Σχετική Συχνότητα** (relative frequency) είναι το πλήθος των παρατηρήσεων, σε μια υποπεριοχή, προς το συνολικό πλήθος παρατηρήσεων σε όλη τη περιοχή μελέτης. Είναι δηλαδή το ποσοστό παρατηρήσεων σε μια υποπεριοχή.

$$\text{Σχετική Συχνότητα} = \frac{\text{Πλήθος στο κελί}}{\text{Συνολικό Πλήθος στη περιοχή μελέτης}}$$

Εξίσωση 4.3: Υπολογισμός Σχετικής Συχνότητας

Ένταση, είναι το πλήθος θέσεων προς το εμβαδό, όπως είδαμε και στην προηγούμενη ενότητα 4.2 αυτού του κεφαλαίου.

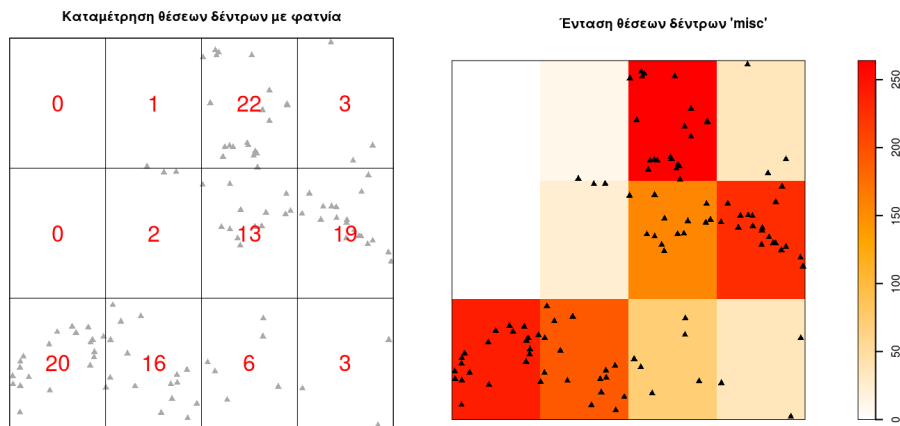
$$\text{Ένταση} = \frac{\text{Πλήθος ανα κελί}}{\text{Εμβαδό κελιού}}$$

Εξίσωση 4.4: Υπολογισμός χωρικής Έντασης θέσεων

Πυκνότητα για μια υπό-περιοχή, είναι η σχετική συχνότητα προς το εμβαδό της υπό-περιοχής

$$\text{Πυκνότητα} = \frac{\text{Σχετική Συχνότητα στο κελί}}{\text{Εμβαδό κελιού}}$$

Εξίσωση 4.5: Υπολογισμός χωρικής Πυκνότητας θέσεων



Εικόνα 4.1: Πλήθος θέσεων (αριστερά), Ένταση θέσεων (δεξιά)

Η παραπάνω Εικόνα 4.1, παρουσιάζει τις δυο έννοιες της συχνότητας συμβάντων (αριστερά) και της έντασης συμβάντων (δεξιά).

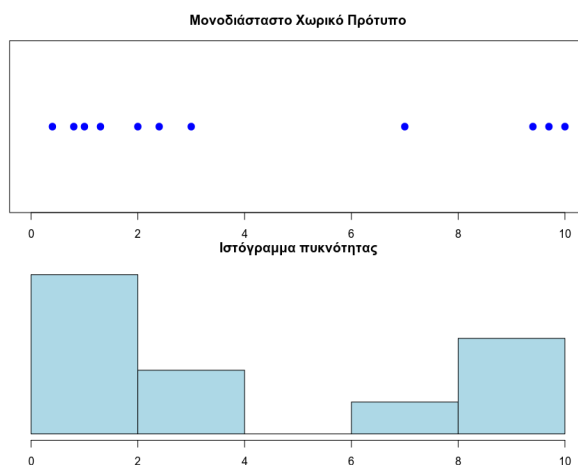
Πίνακας 4.1: Σχέση μεταξύ: Έντασης, Σχετικής Συχνότητας και Πυκνότητας συμβάντων σε υπό-περιοχές

Υπό-περιοχή	Εμβαδό	Συμβάντα	Ένταση	Σχετική Συχνότητα	Πυκνότητα
A	800	45	0.05625	0.3982	0.0004
B	200	45	0.225	0.3982	0.0019
Γ	800	23	0.02875	0.2035	0.0002
	1800	113		1	

Συγκρίνοντας χωρικές πληροφορίες μεταξύ ετερόκλητων περιοχών, είναι απαραίτητη η τυποποίηση των τιμών ώστε να νοείται η σύγκριση και η ιεράρχηση. Όπως φαίνεται και στον παραπάνω πίνακα 4.1 η ένταση είναι ένα μέτρο τυποποίησης με βάση το εμβαδό. Οι περιοχές Α και Β, αν και έχουν το ίδιο πλήθος συμβάντων (45) παρόλα αυτά δεν έχουν την ίδια ένταση φαινομένου καθώς το εμβαδό της περιοχής Α είναι πολύ μεγαλύτερο και έτσι η ένταση είναι μικρότερη. Οι περιοχές Α και Γ αν και έχουν το ίδιο εμβαδό, παρόλα αυτά η ένταση είναι μικρότερη στην περιοχή Γ καθώς το πλήθος συμβάντων είναι μικρότερο.

4.4. Ένταση σε μια και δύο διαστάσεις

Εστω ένα υποθετικό μονοδιάστατο χωρικό πρότυπο, το οποίο αποτελείται από $N=10$ συμβάντα. Για την εκτίμηση της τοπικής πυκνότητας τους μπορεί να χρησιμοποιηθεί ένα ιστόγραμμα πυκνοτήτων (Εικόνα 4.2). Αυτό το ιστόγραμμα αντιστοιχεί σε μια μονοδιάστατη κατανομή της συχνότητας συμβάντων σε χώρο μιας διάστασης. Ο στόχος εδώ είναι η περιγραφή της έντασης των x συνεταγμένων και μια εκτίμηση της κατανομής της έντασης η οποία μπορεί να δοθεί από το ιστόγραμμα πυκνότητας. Αυτή η προσέγγιση μπορεί να εφαρμοστεί στην ανάλυση θέσεων επιχειρήσεων σε έναν εμπορικό δρόμο καταδεικνύοντας τις υπό-περιοχές με μεγαλύτερη πυκνότητα επιχειρήσεων. Σε ένα τέτοιο παράδειγμα, ο οριζόντιος άξονας x θα περιλαμβάνει τις θέσεις των επιχειρήσεων π.χ. Ερμού 7, Ερμού 8, Ερμού 9 κ.ο.κ.



Εικόνα 4.2: Μονοδιάστατο χωρικό πρότυπο θέσεων συμβάντων και αντίστοιχο ιστόγραμμα πυκνότητας

ΧΩΡΙΚΗ ΠΑΡΕΜΒΟΛΗ

ΚΕΦΑΛΑΙΟ 07

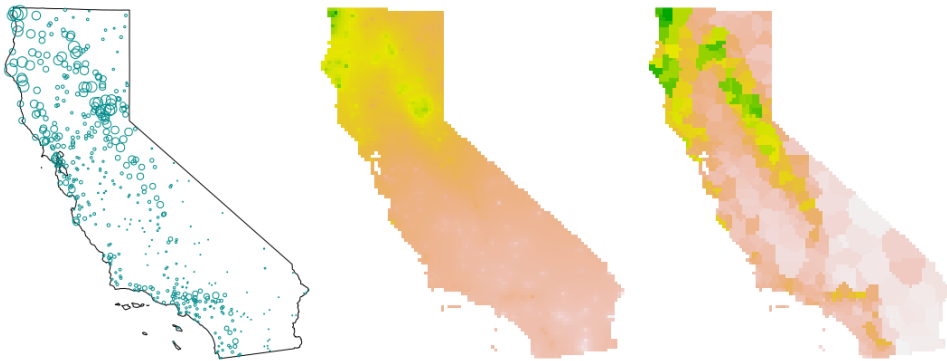
Μέχρι τώρα έχουμε εστιάσει στην ανάλυση δεδομένων που περιλαμβάνουν μετρήσεις μιας χωρικής μεταβλητής (spatial attribute) που έχει καταμετρηθεί σε διάφορες μονάδες παρατήρησης (σημεία δειγματοληψίας). Συνήθως, τα σημεία παρατήρησης είναι λίγα και χωρικά διεσπαρμένα (σε μη-κανονική διάταξη). Στο κεφάλαιο αυτό θα εστιάσουμε στην προσέγγιση της Χωρικής Παρεμβολής. Η οικογένεια μεθόδων της χωρικής παρεμβολής είναι ουσιαστικά η εκτίμηση των άγνωστων τιμών μιας μεταβλητής σε ένα οποιοδήποτε σημείο, με βάση τις γνωστές μετρήσεις. Συχνά, χωρική παρεμβολή μετρήσεων γίνεται στους κόμβους ενός καννάβου ώστε να μπορέσουμε να κατασκευάσουμε μια εκτιμώμενη επιφάνεια τιμών μιας μεταβλητής. Ο σκοπός του κεφαλαίου αυτού είναι η επισκόπηση των βασικών μεθοδολογικών προσεγγίσεων της **Χωρικής Παρεμβολής**. Στο παρόν κεφάλαιο θα εξετάσουμε ντετερμινιστικές μεθόδους Χωρικής Παρεμβολής οι οποίες παράγουν ξεχωριστά αποτελέσματα για ένα-ένα τα άγνωστα σημεία εκτίμησης. Οι μη-ντετερμινιστικές μέθοδοι Χωρικής Παρεμβολής (γεωστατιστικές) θα εξεταστούν στο Κεφάλαιο 10.

Εστω πως έχουμε μια μεταβλητή η οποία είναι **συνεχούς χωρικής κατανομής**, όπως η θερμοκρασία, υγρασία κ.ο.κ. Δεν γνωρίζουμε όμως την ακριβή της χωρική κατανομή σε μια περιοχή μελέτης. Δεν γνωρίζουμε δηλαδή τις τιμές της μεταβλητής μας σε όλη την περιοχή μελέτης, παρά μόνο σε κάποια σημεία δειγματοληψίας, όπως για παράδειγμα τις τιμές της θερμοκρασίας σε κάποιους μετεωρολογικούς σταθμούς.

Η γενικός σκοπός της Χωρικής Παρεμβολής είναι η ανακατασκευή (reconstruction) ή **εκτίμηση** (estimation) ή πρόβλεψη (prediction) της χωρικής κατανομής της μεταβλητής σε σημεία που δεν υπάρχουν γνωστές παρατηρήσεις στην περιοχή μελέτης. Συνήθως οι γνωστές τιμές είναι **λίγες** και προσπαθούμε να ανακατασκευάσουμε τις τιμές της μεταβλητής σε πολλά σημεία της περιοχής μελέτης με σκοπό να κατασκευάσει μια επιφάνεια της μεταβλητής για την περιοχής μελέτης.

7.1. Ορολογία και Συμβολισμός

Αρχικά είναι αναγκαίο να τεθούν οι ορολογίες και οι συμβολισμοί. Τα δεδομένα που χρησιμοποιούνται είναι ένα συνήθως ένα σύνολο N μετρήσεων μιας μεταβλητής $\{z(s_n), n=1, \dots, N\}$, π.χ., s_n είναι το σημείο παρατήρησης $\{n=1, \dots, N\}$ σε N μονάδες παρατήρησης. Ο γενικός σκοπός της Χωρικής Παρεμβολής είναι η εκτίμηση M άγνωστων τιμών $\{z(t_m), m=1, \dots, M\}$ σε ένα σύνολο M σημείων $\{t_m, m=1, \dots, M\}$, με βάση τις N γνωστές αρχικές μετρήσεις. Αυτά τα M σημεία είθισται να είναι διατεταγμένα στους κόμβους ενός καννάβου ώστε αφού παραχθούν τα αποτελέσματα, να είναι δυνατή η οπτικοποίηση υπό μορφή επιφάνειας για μια μεταβλητή.



Εικόνα 7.1: Επισκόπηση αρχικών δεδομένων σε σημεία και αποτελεσμάτων με δυο διαφορετικές μεθόδους.

Οι αρχικές μετρήσεις της μεταβλητής ονομάζονται **δεδομένα πηγής** (source data), ενώ οι άγνωστες τελικές τιμές καλούνται **τιμές στόχου** (target values). Στα πλαίσια αυτό του βιβλίου, κάνουμε την παραδοχή πως τα δεδομένα πηγής και οι τιμές στόχου αφορούν μια μεταβλητή Z και οι αντίστοιχες χωρικές μονάδες παρατήρησης είναι σημειακές. Οι μέθοδοι Χωρικής Παρεμβολής, χωρίζονται με βάση το πλήθος των δεδομένων πηγής, στις ακόλουθες ομάδες:

1. **Καθολική Χωρική Παρεμβολή (global)**: Όταν όλες οι N αρχικές μετρήσεις πηγής χρησιμοποιούνται για εκτίμηση σε κάθε σημείο t_m .
2. **Τοπική Χωρική Παρεμβολή (local)**: Όταν χρησιμοποιείται μόνο ένα υποσύνολο $N(t_m) \ll N$ μετρήσεων, π.χ. των 5 πλησιέστερων για κάθε σημείο t_m .

Οι μέθοδοι Χωρικής Παρεμβολής, χωρίζονται με βάση τη χρήση των τιμών των αρχικών δεδομένων πηγής:

1. **Ακριβής μέθοδος Χωρικής Παρεμβολής** (exact) είναι αυτή η προσέγγιση στην οποία όταν η θέση πρόβλεψης t_m ταυτίζεται με τη θέση παρατήρησης s_n , η εκτιμώμενη τιμή αναπαράγει την αντίστοιχη αρχική μέτρηση: $\hat{z}(s_n) = z(s_n)$
2. **Κυρτή μέθοδος Χωρικής Παρεμβολή** (convex) είναι η προσέγγιση στην οποία οι εκτιμώμενες τιμές δεν είναι μικρότερες/μεγαλύτερες από τη μικρότερη/μεγαλύτερη μέτρηση στο αρχικό χωρικό δείγμα. Είναι δηλαδή μια προσέγγιση η οποία θέτει ανώτερα ή/και κατώτερα όρια τιμών με βάση τα αρχικά δεδομένα πηγής.

Οι μέθοδοι Χωρικής Παρεμβολής, χωρίζονται με βάση τη διαδικασία εκτίμησης των χωρικών δεδομένων στόχου σε:

1. **Ντετερμινιστικές** οι οποίες εφαρμόζουν στη μεταβλητή, μαθηματικές σχέσεις για τη δημιουργία μιας συνεχούς εκτίμησης.
2. Μη-ντετερμινιστικές (ή **γεωστατιστικές**) μέθοδοι Χωρικής Παρεμβολής (Κεφάλαιο 10) οι οποίες βασίζονται στην έννοια της χωρικής αυτοσυσχέτισης παράγοντας μια συνεχή επιφάνεια εκτίμησης αλλά και ενός μέτρου ακριβείας.

7.2. Γραμμική Χωρική Παρεμβολή

Η πιο γενική μορφή Χωρικής Παρεμβολής, είναι η **Γραμμική Χωρική Παρεμβολή** στην οποία η εκτιμώμενη τιμή $\hat{z}(t_m)$ στο σημείο στόχος t_m , εκφράζεται ως ένα **σταθμισμένο άθροισμα** των N μετρήσεων $\{z(s_n), n=1, \dots, N\}$ όπως φαίνεται στην παρακάτω εξίσωση 7.1. Ο όρος w_{mn} είναι το βάρος που αποδίδεται στη μέτρηση $z(s_n)$ κατά την εκτίμηση της άγνωστης τιμής $z(t_m)$ στο σημείο στόχος t_m . Όταν τα σημεία στόχου αντιστοιχούν στους κόμβους ενός κανονικού καννάβου, τότε τα αποτελέσματα μπορούν να οπτικοποιηθούν ως μια “επιφάνεια” εκτιμώμενων τιμών.

$$\hat{z}(t_m) = \sum_{n=1}^N w_{mn} z(s_n)$$

Εξίσωση 7.1: Γενική μορφή της Γραμμικής Χωρικής Παρεμβολής

7.2.1. Προσδιορισμός Βαρών

Τα N βάρη $\{w_{mn}, n=1, \dots, N\}$ που χρησιμοποιούνται στη Χωρική Παρεμβολή, είναι συνήθως συνάρτηση της χωρικής κατανομής (χωρικής διάταξης) των σημείων πηγής και των σημείων στόχου. Ειδικότερα, τα βάρη αυτά είναι συνήθως συνάρτηση των N αποστάσεων στόχου-πηγής $\{d_{mn}, n=1, \dots, N\}$ μεταξύ των N αρχικών

σημείων παρατήρησης και των σημείων εκτίμησης t_m . Με άλλα λόγια, τα βάρη υπολογίζονται με βάση τη χωρική διάταξη των σημείων παρατήρησης και εκτίμησης και στη συνέχεια εφαρμόζονται ως μαθηματικά βάρη στις αρχικές παρατηρούμενες τιμές της μεταβλητής. Στα παρακάτω υποκεφάλαια, παρουσιάζονται 4 μέθοδοι για τον προσδιορισμό των βαρών της χωρικής παρεμβολής:

- Μέθοδος Βαρών 1) Μέθοδος Πλησιέστερου Γείτονα
- Μέθοδος Βαρών 2) Τριγωνισμός Delauney
- Μέθοδος Βαρών 3) Μέθοδος Τοπικού Μέσου Ορου
- Μέθοδος Βαρών 4) Μέθοδος Αντίστροφων Αποστάσεων (IDW)

7.3. Μέθοδος Βαρών 1: Μέθοδος Πλησιέστερου Γείτονα

Σύμφωνα με τη μέθοδο πλησιέστερου γείτονα, τα βάρη που χρησιμοποιούνται έχουν άθροισμα 1 και λαμβάνουν τιμές στο διάστημα $[0, 1]$. Τα βάρη είναι συνάρτηση των αποστάσεων μεταξύ σημείων εκτίμησης και σημείων αρχικών παρατηρήσεων. Κάποιο σημείο παίρνει $w=1$ για το αρχικό σημείο παρατήρησης που είναι πλησιέστερος γείτονας του σημείου στόχος t_m . Όλα τα υπόλοιπα σημεία παρατήρησης παίρνουν βάρος $w=0$ όπως φαίνεται και στην παρακάτω εξίσωση 7.2.

$$w_{mn} = \begin{cases} 1 & \text{αν } d_{mn} = \min\{d_{mn}, n = 1, \dots, N\} \\ 0 & \text{διαφορετικά} \end{cases}$$

Εξίσωση 7.2: Προσδιορισμός βαρών με τη προσέγγιση του πλησιέστερου γείτονα

Η εκτιμώμενη τιμή που αποδίδεται σε κάποιο σημείο στόχος $\hat{z}(t_m)$ ταυτίζεται με τη μέτρηση $z(s_n)$ του αρχικού σημείου παρατήρησης s_n που βρίσκεται πλησιέστερα στο σημείο εκτίμησης t_m όπως αναλύεται και στην παρακάτω εξίσωση 7.3.

$$\begin{aligned} \hat{z}(t_m) &= \sum_{n=1}^{n=N} w_{mn} z(s_n) \\ &= 1 \cdot z(s_n) + \sum_{\substack{n=1 \\ n \neq n}}^{n=N} 0 \cdot z(s_n) \end{aligned}$$

Εξίσωση 7.3: Υπολογισμός της εκτιμώμενης τιμής μόνο από το πλησιέστερο αρχικό σημείο παρατήρησης