**ΔΙΑΧΕΙΡΙΣΗ- ΑΝΑΛΥΣΗ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΓΕΩΧΗΜΙΚΩΝ ΔΕΔΟΜΕΝΩΝ**

**ΜΕ ΧΡΗΣΗ** **ΓΣΠ**

1. **Εισαγωγή**

Τα ΓΣΠ είναι ένα πολύ χρήσιμο εργαλείο για τον σύγχρονο γεωχημικό, παρέχοντας δυνατότητες επεξεργασίας, ανάλυσης και παρουσίασης δεδομένων που έχουν συλλεχθεί στο ύπαιθρο και το εργαστήριο. Πριν την εμφάνιση των ΓΣΠ, ο κλασικός τρόπος ανάλυσης γεωχημικών δεδομένων περιλάμβανε τη χρήση στατιστικών μεθόδων. Σήμερα, αυτές οι μέθοδοι μπορούν να εμπλουτιστούν και να συνδυαστούν με χωρικά δεδομένα μέσω των ΓΣΠ, συμβάλλοντας στην καλύτερη κατανόηση των διεργασιών που ελέγχουν την κατανομή των στοιχείων στον πλανήτη μας. Έτσι ο συνήθης τρόπος εργασίας για την ανάλυση γεωχημικών δεδομένων είναι πλέον η συνδυασμένη χρήση λογισμικών ΓΣΠ και στατιστικών πακέτων.

Το εύρος των δυνατοτήτων των ΓΣΠ εξαρτάται από το συγκεκριμένο πρόβλημα που εξετάζεται, το λογισμικό που χρησιμοποιείται και την ευχέρεια του χρήστη στο συγκεκριμένο λογισμικό. Η παρούσα άσκηση αποτελεί ένα απλό παράδειγμα χρήσης του ArcView© GIS λογισμικού για τον **προσδιορισμό της επίδρασης της συχνότητας του πλημμυρικού φαινομένου ενός ποταμού στη διασπορά της ρύπανσης από Zn στις ποτάμιες αποθέσεις**. Για την επίλυση του προβλήματος εκτός από το ArcView© θα χρησιμοποιηθεί και το πρόγραμμα Microsoft Excel για τη βασική στατιστική ανάλυση των δεδομένων.

Τα δεδομένα της άσκησης (Πίνακας 1) περιλαμβάνουν μετρήσεις διαφόρων παραμέτρων για 98 δείγματα επιφανειακού, αλλούβιου εδάφους (βάθους 0 – 20 cm) σε μία περιοχή 5x2 km που ανήκει στην πλημμυρική επιφάνεια του ποταμού Maas στη Νότια Ολλανδία (Burrough and McDonnell, 1998). Κάθε ένα από τα **σημειακά** αυτά **δεδομένα** αντιπροσωπεύει τιμές παραμέτρων εντός μιας επιφάνειας εμβαδού 10x10 m, από την οποία συλλέχθηκαν σύνθετα δείγματα βάσει κανάβου (stratified random sampling). Τα δείγματα αναλύθηκαν χημικά για τον προσδιορισμό της συγκέντρωσης βαρέων μετάλλων (Zn, Cd, Cu, Pb). Η ανάλυσή μας θα εστιάσει στα δεδομένα των χημικών αναλύσεων Zn και τις κατηγορίες συχνότητας του πλημμυρικού φαινομένου (Fldf). Το **κριτήριο για την κατηγοριοποίηση** των δεδομένων Fldf είναι η συχνότητα εκδήλωσης πλημμύρας ως εξής:

**Κατηγορία Συχνότητα πλημμύρας**

1. Κάθε χρόνο
2. Κάθε 2-5 χρόνια
3. Σπάνια (> 5 χρόνια)

Οι τρεις κατηγορίες συχνότητας του πλημμυρικού φαινομένου καθώς και τα σημεία δειγματοληψίας απεικονίζονται χωρικά στο Χάρτη 1.

1. **Θεωρητική βάση για λύση του προβλήματος**

Σκοπός μας είναι να προσδιορίσουμε την μεταβολή της συγκέντρωσης Zn (όπως έχει προσδιοριστεί στα σημεία δειγματοληψίας) στα πλημμυρικά ιζήματα του ποταμού σε συνάρτηση με τη συχνότητα του πλημμυρικού φαινομένου (όπως περιγράφεται χωρικά από τις κατηγορίες 1-3 του χάρτη). Υιοθετώντας μια τέτοιου είδους θεώρηση μπορούμε να υποθέσουμε ότι εφ’ όσων ο Zn στην περιοχή προέρχεται από τα ιζήματα απόθεσης του ποταμού, οι περιοχές με μεγαλύτερη συχνότητα πλημμύρας θα εμφανίζουν τις υψηλότερες περιεκτικότητες Zn. Θα εξετάσουμε την ισχύ αυτής της υπόθεσης χρησιμοποιώντας τη στατιστική τεχνική της απλής ανάλυσης της μεταβλητότητας (ANOVA). Θυμηθείτε το απλό μοντέλο ANOVA:

*z(xo) = μ + αk + ε* , όπου

*z* η τιμή μιας μεταβλητής στο σημείο *xo*,

*μ* η μέση τιμή της μεταβλητής *z*,

*αk* η απόκλιση μεταξύ της *μ* και της μέσης τιμής στη μονάδα *k* και

*ε* η μεταβλητότητα των τιμών εντός κάθε κατηγορίας (θόρυβος).

Τα αποτελέσματα της ANOVA μπορούν να αποτελέσουν κριτήριο για την επιτυχία της κατηγοριοποίησης μιας παραμέτρου. Μας ενδιαφέρει η σχετική μεταβλητότητα (s2w/ s2t), όπου s2w είναι η μεταβλητότητα των τιμών εντός κάθε κλάσης και s2t η συνολική μεταβλητότητα του πληθυσμού. Όσο μικρότερο είναι το κλάσμα (s2w/ s2t )τόσο καλύτερη είναι η κατηγοριοποίηση.

Η εφαρμογή του μοντέλου ANOVA σε χωρικά δεδομένα προϋποθέτει τα εξής:

* Η μεταβλητότητα εντός κάθε χωρικής μονάδας του χάρτη ακολουθεί τυχαία κατανομή και δεν εξαρτάται από τη θέση.
* Η μεταβλητότητα εντός κάθε χωρικής μονάδας του χάρτη δεν μεταβάλλεται
* Όλες οι παράμετροι ακολουθούν την κανονική κατανομή
* Όλες οι χωρικές μεταβολές λαμβάνουν χώρα ακριβώς πάνω στα όρια των χωρικών κλάσεων.

Στην πραγματικότητα πολλές από τις προϋποθέσεις είναι δύσκολο να ισχύουν. Ιδιαίτερα η κανονικότητα της κατανομής είναι σπάνια για φυσικές παραμέτρους και ειδικά για γεωχημικά δεδομένα. Έτσι πολλές φορές προκειμένου να χρησιμοποιηθεί η μέθοδος, τα δεδομένα μετατρέπονται πριν την ανάλυση σε κατάλληλη κλίμακα, π.χ. φυσικούς λογάριθμους.

1. **Μεθοδολογία επίλυσης**

* Αρχίστε εξάγοντας τα στατιστικά περιγραφικά στοιχεία για τον πληθυσμό του Ζn.

Στο **Excel** ανοίξτε το αρχείο c:\mydocuments\PG\_GISexercise\maasdata.xls και υπολογίστε τη μέση τιμή και την τυπική απόκλιση των δεδομένων. Επιλέξτε κλάσεις εύρους 100 ppm και δημιουργήστε το ιστόγραμμα κατανομής των τιμών Zn στο έδαφος. Είναι η κατανομή των δεδομένων κανονική; Αν όχι μετατρέψτε τα δεδομένα σε φυσικούς λογαρίθμους και περιγράψτε στατιστικά τον τροποποιημένο πληθυσμό (μέση τιμή, τυπική απόκλιση, ιστόγραμμα).

* Ξεκινήστε τη λειτουργία του προγράμματος **ArcGis 🡪 Arc Map**.
* Προσθέστε στον χάρτη (Map document) τα θεματικά επίπεδα (Layers):

1. c:\mydocuments\PG\_GISexercise\maasmap.shp

2. c:\mydocuments\PG\_GISexercise\floodclass.shp

3. c:\mydocuments\PG\_GISexercise\maasdata.txt

* Δημιουργήστε κατάλληλα υπομνήματα (symbology) ώστε να απεικονίζονται σωστά τα στοιχεία του χάρτη:
* Κάντε διπλο κλικ στην περιοχή του τίτλου των θεματικών επιπέδων και επιλέξτε κατάλληλα χρώματα και σχήματα στο παράθυρο που εμφανίζεται.
* Δημιουργήστε ένα νέο σημειακό θεματικό επίπεδο χρησιμοποιώντας τα δεδομένα x, y του αρχείου maasdata.txt:
* Add x,y data
* Δώστε όνομα στο νέο αρχείο – θεματικό επίπεδο που θα δημιουργηθεί π.χ. maasdata.shp
* Υπολογίστε τη μέση τιμή και την τυπική απόκλιση των τιμών Zn για κάθε κατηγορία πλημμυρικής συχνότητας:
* Κάντε δεξί κλικ στο επίπεδο maasdata.shp και επιλέξτε attribute table
* Κάντε δεξί κλικ στον πίνακα και επιλέξτε Summarize
* Στο παράθυρο που εμφανίζεται επιλέξτε το πεδίο με τις κατηγορίες πλημμυρικής συχνότητας και τις περιγραφικές παραμέτρους average, standard deviation για το πεδίο τιμών του Zn. Τέλος, προσδιορίστε το όνομα του αρχείου όπου θα αποθηκευτούν οι πληροφορίες.
* Εξετάστε τις τιμές που προκύπτουν σε σχέση με το ιστόγραμμα κατανομής των τιμών Zn. Τι συμπεράσματα βγάζετε σχετικά με τη διάκριση των τιμών Zn σε κάθε κατηγορία; Αν στην ανάλυση χρησιμοποιηθούν οι φυσικοί λογάριθμοι αντί των πραγματικών τιμών συγκέντρωσης Zn \\ υπάρχει βελτίωση στην κατηγοριοποίηση;
* Δημιουργήστε τρία νέα επίπεδα, ένα για κάθε κατηγορία πλημμυρικής συχνότητας:
* Στο menu selection επιλέξτε select by attribute και στο παράθυρο που εμφανίζεται διαμορφώστε τη σχέση αναζήτησης [Fldf = 1]. Επιλέξτε new set.
* Κάντε δεξί κλικ στον πίνακα του επιπέδου maasdata.shp και επιλέξτε data 🡪 export data και ονομάστε το νέο αρχείο που θα δημιουργηθεί από τα επιλεγμένα στοιχεία π.χ. flood1.shp
* Επαναλάβετε τη διαδικασία για να δημιουργήσετε νέα επίπεδα για τις άλλες δύο κατηγορίες (flood2.shp, flood3.shp).
* Επιστρέψτε στο Excel και πραγματοποιήστε ANOVA στα δεδομένα συγκέντρωσης Zn:
* Σε ένα νέο φύλλο εργασίας εισάγετε τα δεδομένα από τους πίνακες flood1.dbf, flood2.dbf, flood3.dbf
* Επιλέξτε Tools🡪 Data Analysis🡪 ANOVA και αποθηκεύστε τα αποτελέσματα σε ένα νέο φύλλο εργασίας.
* Τι συμπεράσματα εξάγετε από την ανάλυσή σας για την επίδραση της συχνότητας του πλημμυρικού φαινομένου στη διασπορά της ρύπανσης Zn στο έδαφος της περιοχής;

# Βιβλιογραφία

Burrough, A. and McDonnell, A. (1998) *Principles of Geographical Information Systems*, Oxford University Press, New York.