

GLQ3402 - TRAITEMENT STATISTIQUE DES DONNÉES GÉOLOGIQUES (été 2006)

PLAN DE COURS

Crédits: 2 (2-2-2)  
Préalable: MTH2301

Professeur: Denis Marcotte  
Local: B-649  
tél : 4620  
courriel : denis.marcotte@polymtl.ca

DISPONIBILITÉ : en tout temps, particulièrement les journées 3 et 4 en après-midi.

BUT DU COURS : Ce cours vise à donner une connaissance d'usage des concepts et calculs relatifs aux principales méthodes multidimensionnelles d'analyse des données ainsi qu'à rendre l'étudiant capable d'interpréter en termes géologiques les résultats de telles analyses.

OBJECTIFS : À la fin du cours, l'étudiant saura:

- i. identifier les finalités et les particularités des différentes méthodes vues au cours, à savoir : corrélations, régression, analyse en composantes principales, analyse discriminante et classification automatique;
- ii. reconnaître les étapes de calculs propres à ces méthodes et effectuer lui-même ces calculs à l'aide du logiciel MATLAB;
- iii. distinguer les éléments communs aux différentes méthodes factorielles;
- iv. juger de la pertinence d'une analyse des données face à un problème géologique donné et évaluer quelle(s) méthode(s) est adéquate dans un contexte défini.

MÉTHODE D'ENSEIGNEMENT :

Cours magistraux avec parfois petits exercices en classe. Travaux pratiques plus élaborés permettant d'appliquer et de comprendre les méthodes vues en classe. On insistera sur les aspects géométriques des méthodes présentées. Des notes polycopiées du professeur constituent la référence principale. Elles sont disponibles sur le site internet du cours. Le logiciel MATLAB est utilisé lors des TP.

SITE INTERNET : <http://geo.polymtl.ca/~marcotte/glq3402.html>

Horaire : cours : lundi 13h - 15h et jeudi 8h30 – 10h30 au B620.  
labo : mardi 8h30 – 12h30 au L-6612

## ÉVALUATION:

L'évaluation de l'acquisition des connaissances pourra se faire selon la répartition suivante :

Travaux pratiques :	5 TP :	30 %
Examen mi-session :		35 %
Examen final (non-cumulatif) :		35 %

Les critères d'évaluation des travaux pratiques sont, par ordre d'importance :

- justesse des énoncés
- justifications soutenant les énoncés
- exactitude des calculs
- validité des interprétations
- présentation et orthographe et syntaxe

Les TP doivent être réalisés seuls ou en équipes de 2 (1 équipe de 3 sera acceptée dans le cas d'un nombre impair d'étudiants dans la classe). Les rapports doivent être remis préférablement en fin de séance, sinon au plus tard au cours suivant la séance de laboratoire. Il est fortement suggéré de répondre au brouillon à toutes les questions avant de quitter le laboratoire. **La participation aux TP est obligatoire.** La note 0 est attribuée pour une absence non motivée. Pour une absence motivée, la pondération du TP est reportée sur les examens.

## RÉFÉRENCE :

Notes de cours du professeur disponibles sur le site internet.

Autres références indiquées dans chaque chapitre des notes de cours.

## SITUATION DU COURS DANS LE PROGRAMME :

Le cours GLQ3402 constitue un complément au cours MTH2301 (statistiques en ingénierie). En plus de reprendre les notions de corrélation et de régression en les généralisant au contexte multivariable et en les appliquant à des données et problèmes géologiques, il introduit un ensemble de nouvelles méthodes multivariées d'analyse des données. Les notions vues au cours ING1006 Algèbre linéaire trouvent ici plusieurs applications. La matière vue au cours GLQ3401 (géostatistique et géologie minières) peut également être utile.

## UTILITÉ DU COURS :

Le cours fournit un ensemble d'outils utiles à l'ingénieur géologue dans sa vie professionnelle. Il est certain que tout ingénieur géologue a un jour à réaliser des régressions à deux ou plusieurs variables, à juger de la validité du modèle, au besoin, à l'améliorer, et à en interpréter les résultats. Bien que plus spécialisées, les autres méthodes vues au cours ont aussi beaucoup d'utilité pour plusieurs applications.

Bonne session à tous!

Denis Marcotte

## HORAIRE

Date (J. type)	Contenu	Remarques
23 mai (1)	Plan de cours, introduction , corrélations simples, Matlab, régression	
24 mai (2)	TP 1 : Matlab et Régression linéaire (2 variables)	
26 mai (4)	Régression linéaire multiple	
29 mai (1)	Régression linéaire multiple	
30 mai (2)	TP-2 Régression linéaire multiple	6%
1 juin (4)	Régression linéaire multiple	
5 juin (1)	Régression logistique et régression non-linéaire	
6 juin (2)	TP-3 Régression linéaire multiple et régression logistique	6%
8 juin (4)	Analyse en composantes principales	
12 juin (1)	Analyse en composantes principales	
13 juin (2)	Tp-4 Analyse en composantes principales	6%
15 juin (4)	Examen intra (Régression)	35%
19 juin (1)	Analyse discriminante	
20 juin (2)	TP-5 Analyse discriminante	6%
22 juin (4)	Analyse discriminante; Classification automatique	
26 juin (1)	Classification automatique; statistiques de données vectorielles	
27 juin (2)	TP-6 : Classification automatique; statistiques de données vectorielles	6%
29 juin (4)	Statistiques de données vectorielles	
5 juillet	Examen final (non-cumulatif; tout sauf régression)	35% de 9h30 à 12h

Journées types : (1) : 13h à 15h; (2) 8h30 à 12h30; (4) 8h30 à 10h30.

### Horaire détaillé – été 2006

X: journée sans cours

1 à 4: journée avec cours suivant le tableau 1

**Tableau 1 : les 4 journées type**

Journée	AM	PM	Total
1	GLQ3202 (3h cours) 9h-12h	GLQ3402 (2h cours) 13h-15h	5h
2	GLQ3402 (4h TP) 8h30-12h30	GLQ2201 (3h cours) 13h30-16h30	7h
3	GLQ3202 (3h cours) 9h-12h	GLQ2201 (3h cours) 13h-16h	6h
4	GLQ3402 (2h cours) 8h30-10h30	GLQ3202 (6h TP) 11h-17h	8h

Les journées 1 à 4 ne sont pas fixes, elles se déplacent au gré des aléas de l'horaire causés par les différents congés (fête de la Reine; Saint-Jean Baptiste; fête du Canada). Voir le tableau 2:

**Tableau 2: le calendrier**

Lundi	Mardi	Merc.	Jeudi	Vendredi
22 mai - X	23 - 1	24 - 2	25 - 3	26 - 4
29 - 1	30 - 2	31 - 3	1 <sup>er</sup> juin- 4	2 - X
5 - 1	6 - 2	7 - 3	8 - 4	9 - X
12 - 1	13 - 2	14 - 3	15 - 4	16 - X
19 - 1	20 - 2	21 - 3	22 - 4	23 - X
26 - 1	27 - 2	28 - 3	29 - 4	30 - X
3 - examen (GLQ2201)	4 - X	5- examen (GLQ3402)	6 - X	7- examen (GLQ3202)

Notes :

- le 1<sup>er</sup> chiffre indique la date, le 2<sup>e</sup> chiffre la journée type suivie ce jour calendrier;
- on a en tout 6 journées de chaque type + 3 journées réservées aux examens.