

Evaluation de différentes méthodes d'interpolation spatiale pour la production d'un MNT à partir de données topographiques dans un SIG

■ Noamen REBAI, Tarek SLAMA et Mohamed Moncef TURKI

La vraie puissance du système d'information géographique (SIG) réside dans l'analyse de l'information spatiale. Cette dernière nécessite le paramétrage de la continuité spatiale ainsi qu'un suivi de régularité surfacique qui ne peut être assuré que par interpolation. Cet article présente une approche méthodologique quantitative pour l'évaluation de six méthodes d'interpolation utilisées dans la production d'un modèle numérique de terrain (MNT) à partir de données topographiques vectorielles dans un SIG. Divers paramètres statistiques sont utilisés et analysés pour évaluer ces méthodes et déterminer le ou les meilleurs interpolateurs.

■ mots-clés

MNT, SIG, interpolation, analyse statistique, données topographiques

En sciences de la terre et particulièrement dans les disciplines de la géologie, la géographie et la géomorphologie, la modélisation du terrain par la production du modèle numérique de terrain (MNT) est très importante. C'est à partir de ce modèle spatial continu et régulier que se sont effectués plusieurs analyses et traitements. Le MNT contient une véritable mine d'informations sur la configuration du relief sous forme de paramètres morphométriques et topographiques (Grohmann, 2004). De nombreuses cartes comme celle de la distribution ci-dessus des glissements de terrain sont produites en se basant principalement sur l'étude du MNT (Gaspar *et al.*, 2004).

Actuellement, les sources de données pour la production d'un MNT sont très variées, mais la carte topographique (échelle 1: 50 000 ou 1: 25 000) reste la plus facile d'accès et la plus répandue. Le SIG constitue l'unique environnement permettant une production par la procédure d'interpolation spatiale. Cette puissante technique permet la génération d'une surface continue ou grille régulière selon une résolution contrôlée. Mais il existe plusieurs méthodes

d'interpolation dans le SIG, et il est bien démontré qu'il n'existe pas une méthode universelle parfaite à tous types de données spatiales ((Rhind, 1975 ; Lam, 1983 ; Sarközy, 1999 ; Hamrah *et al.*, 2006).

Le présent article, qui entre dans le cadre des travaux de recherches visant à exploiter le MNT d'une région d'un contexte géologique et géomorphologique assez compliqué, présente une étude comparative de six méthodes d'interpolation les plus utilisées en sciences de la terre: (1) l'Inverse Distance (IDW), (2) la Courbure Minimale, (3) le Voisin Naturel "Natural Neighbor", (4) le krigeage ordinaire par point, (5) le krigeage ordinaire par bloc et (6) le Polynôme Local. Pour l'évaluation de leurs performances, une méthodologie d'analyse statistique quantitative est mise en place. Elle se base sur: (i) l'examen de l'ensemble des données topographiques interpolées par la méthode de validation croisée (Cross Validation), (ii) l'analyse statistique et visuelle (sous la forme de cartes des erreurs d'interpolation) des résiduels et (iii) l'étude détaillée des échantillons de points altimétriques $z f(x, y)$ de l'ensemble des méthodes.

Un inventaire statistique détaillé du secteur d'étude est également nécessaire pour examiner la base de données et mieux comprendre sa distribution géographique et spatiale. Pour ce faire une approche d'analyse visuelle et statistique est aussi mise en place.

Le but principal de cette étude est d'estimer statistiquement, suivant une méthodologie quantitative, la performance des interpolateurs testés pour générer, dans un SIG, un MNT à partir des données topographiques d'une zone présentant une géomorphologie plutôt compliquée et ainsi de mettre en évidence le rapport entre fonction d'interpolation et nature des données spatiales et donc d'examiner les facteurs agissant sur cette fonction qui peuvent être une source potentielle d'erreur.

Présentation de la zone d'étude

Le secteur de Jebel Kechtilou-Jebel Jebes est une importante zone située en Tunisie septentrionale qui fait partie de l'orogène atlasique nord africain édifié au cours des serrages de l'Eocène et du Miocène (Figure 1, A). Géologiquement,

■ ■ ■ il présente une complication structurale majeure à cause des extrusions triasiques qui ont affecté, par les mouvements halocinétiques, les séries post-triasiques et ont engendré une morphologie spécifique.

La figure 1(B) présente la configuration topographique des deux massifs majeurs de la zone étudiée. Le premier est celui du Jebel Kechtilou (encadré bleu d'en bas) et le second est celui du Jebel Jebbs (encadré bleu d'en haut dans la même figure).

Les techniques d'interpolation spatiale

■ Vue d'ensemble

Généralement, plusieurs informations géographiques de différentes natures peuvent être mesurées et largement utilisées dans des procédures variées d'analyse spatiale. A l'origine, ces informations sont de nature continue dans l'espace géographique tridimensionnel et il est évidemment impossible de mesurer leurs valeurs en tous points de l'espace qu'elles occupent, mais uniquement au niveau des endroits types d'une manière irrégulière et selon la technique d'échantillonnage utilisée (the sampling).

Ainsi, l'interpolation est la technique qui permet de donner par estimation des valeurs aux points non échantillonnés en se basant primordialement sur ceux dont les valeurs et les positions sont connues dans le même espace. Le résultat immédiat de cette opération est la génération d'une surface continue et régulière (dépendant de la résolution) qui est une représentation bidimensionnelle de la surface tridimensionnelle appelée grille ou MNT dans le cas des données topographiques.

Dans le SIG il existe plusieurs méthodes d'interpolation qui donnent généralement des résultats variés et différents. Il existe plusieurs approches pour classer ces méthodes (Hartkamp *et al.*, 1991 ; Sarközy, 1999) mais sommairement on distingue trois critères de classification selon que l'interpolateur soit :

- local ou global

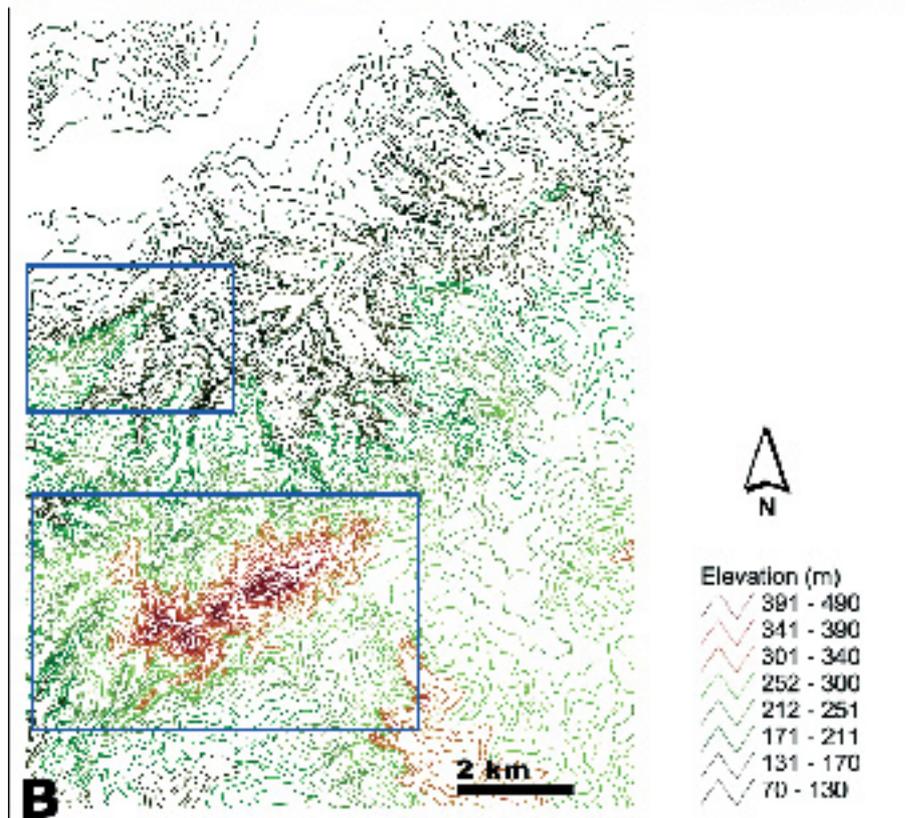
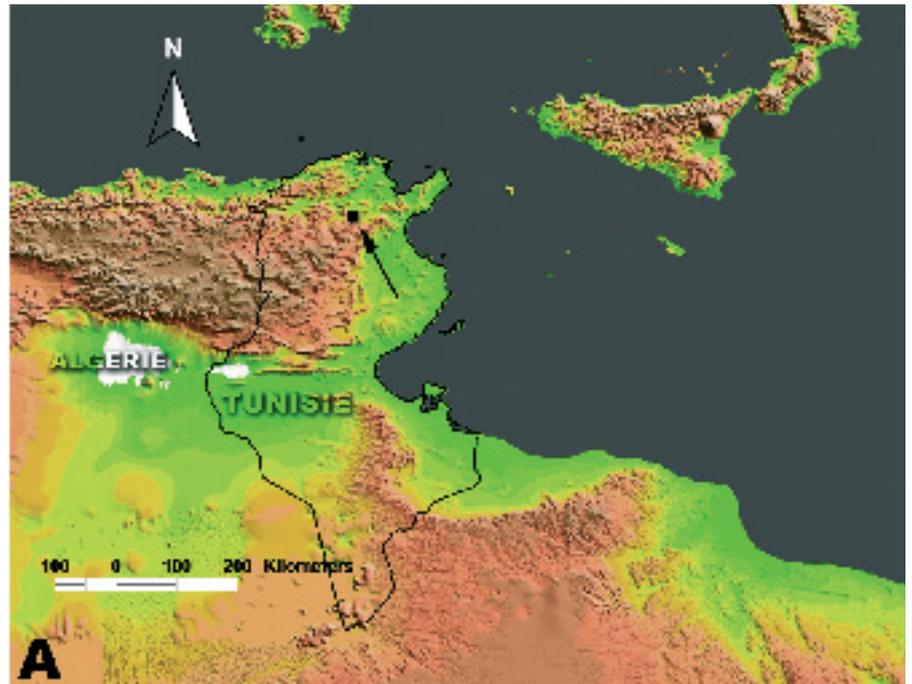


Figure 1 : Localisation de la zone d'étude. En A sa situation géographique et en B sa carte topographique représentée par des courbes de niveau avec une équidistance de 10 m.

- stochastique ou déterminatif (empirique)
- exact ou approximatif

Les méthodes globales comme "la surface de tendance ou the trend surface "

et les équations polynomiales sont largement utilisées pour modéliser des variations de grandes étendues et, comme son nom l'indique, estimer la spécificité régionale du phénomène étu-

dié. La plupart des techniques d'interpolation sont locales et déterminatives à l'exception du krigeage qui est de nature stochastique. Ce dernier est généralement considéré comme un interpolateur exact.

De nombreux travaux de recherche sur l'analyse des données représentant l'espace géographique ont recours aux méthodes d'interpolation pour l'estimation et la prédiction des modèles de distribution. En climatologie l'interpolation est utilisée pour l'estimation des températures journalières (Collins et Bolstad, 1996) et des précipitations au niveau des zones montagneuses (Creutin et Obled, 1982 ; Tabios et Salas, 1985 ; Phillips *et al.*, 1992) avec la recommandation d'utiliser les méthodes de l'Inverse Distance et la Régression Polynomiale comme interpolateurs. En hydrologie, Nash et Sutcliffe (1970) présentent une approche statistique pour prédire le flux fluvial. Pour les données du sol, les techniques les plus fiables sont le krigeage et la surface de tendance (Stein *et al.*, 1988, 1989).

D'autres travaux ont modélisé par interpolation la distribution du pH dans le sol (Laslett *et al.*, 1987). Dans le domaine de la géophysique, la méthode de la courbure minimale est la plus fréquemment utilisée avec les données gravimétriques et magnétiques pour l'estimation des pièges potentiels des hydrocarbures (Mendoza et Silva, 1995 ; Cooper, 2000), alors que dans l'industrie minière c'est l'inverse distance (IDW) qui est l'interpolateur le plus recommandé (Collins et Bolstad, 1996).

Le champ d'application des méthodes d'interpolation dans un SIG est très vaste et intéresse presque toutes les disciplines des sciences de la vie (médecine, biologie, botanique, etc.) et des sciences de la terre. De ce fait, il est vraiment indispensable d'avoir mis en place une méthodologie d'analyse pour tester et quantifier les erreurs des estimateurs mathématiques.

■ Méthodes utilisées

Six méthodes d'interpolation spatiale sont comparées dans l'étude afin d'évaluer la ou les plus performantes d'entre-

elles pour la production d'un MNT à partir des données topographiques d'un SIG. Ces méthodes sont couramment utilisées dans les sciences de la terre pour modéliser différents paramètres de la réalité géographique du territoire (cf. supra).

Les techniques évaluées sont exposées ici brièvement sans entrer dans le détail mathématique des équations qui sont bien énoncées et précisées dans de nombreuses publications dont celle de Sarkozy (1999).

- L'inverse distance (couramment désigné par les lettres IDW: Inverse Distance Weighting) est une méthode d'estimation locale et déterminative qualifiée d'interpolateur exact. Elle suppose que les valeurs tout près de l'emplacement non échantillonnées sont plus représentatives de la valeur à être estimée que celles des échantillons plus éloignés (Franke, 1980). Ses résultats sont bons lorsque les valeurs à interpoler sont régulièrement échantillonnées. Il s'agit d'une méthode simple à utiliser.

- La courbure minimale (the minimum curvature method) de Briggs (1974) est une technique déterminative locale présentant une interpolation approximative. Elle produit la surface la plus lisse possible à partir des données disponibles (pour plus de détail voir Smith et Wessel, 1990). Cette méthode est largement adoptée par les géophysiciens.

- Le voisinage naturel (the natural neighbor) est également une méthode déterminative locale qui, dans son algorithme d'interpolation, se base sur la moyenne pondérée des valeurs voisines pour effectuer une estimation de la valeur non échantillonnée (Sibson, 1981). Cette méthode est commune à certains domaines de la géoscience (géostatistique et estimation des tenseurs en particulier).

- Le krigeage est une méthode locale stochastique (ou géostatistique) développée par Matheron (1963, 1970). Il s'agit d'un interpolateur exact qui tient compte de la taille du champ à estimer et de la position des points entre eux

(c'est-à-dire l'interdépendance des variables), et prend en considération la continuité spatiale de la zone étudiée. C'est une méthode d'estimation linéaire minimisant la variance d'estimation telle que calculée à l'aide du variogramme. Il existe généralement deux types de krigeage: universel et ordinaire, avec un avantage du dernier qui est le plus fréquemment utilisé et considéré comme un bon estimateur des données interpolées. Dans l'étude comparative deux méthodes de krigeage ordinaire par point et par bloc seront testées et évaluées.

- Le polynôme local est une méthode globale simple. Elle se base sur le "moindre carré pondéré" lors de l'interpolation pour la génération de la grille régulière. La tendance locale de cette technique peut être intéressante dans le domaine de l'interpolation des variables topographiques d'un terrain géologique.

Méthodologie et approche d'analyse

L'évaluation rigoureuse de ou des meilleurs interpolateurs est soumise à une méthode d'analyse statistique et visuelle. Celle-ci intègre de nombreux paramètres et coefficients statistiques et se développe comme suit.

■ Analyse des données topographiques

Les données topographiques sous format vectoriel du secteur d'étude (Figure 1, B) sont issues d'une numérisation de la carte topographique de la région de Mejez el Bab à l'échelle 1:50 000 avec une équidistance des courbes de niveaux surfacique de 10 m. Ces dernières sont transformées en une base de données "ponctuelles" composée de 38712 valeurs d'altitude au format (x, y, z) stockées dans un fichier d'extension .dbf.

L'analyse montre que l'ensemble des données présente une répartition légèrement asymétrique reflétant une faible valeur du coefficient d'obliquité (coefficient of skewness) comme le montre le tableau 1. Ce coefficient est sensible à la

■■■ distribution des données dans l'espace géographique et permet de mesurer son asymétrie. Pour une distribution parfaitement symétrique sa valeur est égale à zéro. Ce coefficient est calculé par :

$$Coef_Skewness = \left(\frac{1}{N} \sum_{i=1}^N (Z_i - Z_m) \left(\frac{1}{SD} \right) \right) [1]$$

Avec N est le nombre total des données d'altitude, Z_i et Z_m sont respectivement l'altitude d'un point en un emplacement, $i = f(x, y)$ et la moyenne des altitudes. SD représente la déviation standard.

La distribution des données présente une forme dite "Leptokurtique" comme le montre l'histogramme de la Figure 2, ainsi qu'une valeur de Kurtosis égale à 2.74 indiquant également une légère asymétrie dans la distribution des données. Ce paramètre est calculé par le coefficient de Kurtosis (équation 2). Une valeur égale à 3 montre qu'il existe une symétrie dans la distribution.

$$Coef_Kurtosis = \left(\frac{1}{N * SD} \sum_{i=1}^N (Z_i - Z_m)^3 \right) - 3 [2]$$

Il est important d'avoir un aperçu visuel sur la distribution moyenne des données topographiques ainsi que sur la répartition de leurs densités dans l'espace géographique. Ainsi, deux types de cartes ont été créés pour examiner ces deux aspects. La carte des altitudes (Figure 3) montre clairement l'existence de trois zones distinctes qui sont du

Nombre de points	38712
Minimum	70
Maximum	490
Moyenne	214.92
Erreur Standard	0.4079
Variance	6462
Déviatoin moyenne	65.28
Déviatoin standard	80.25
Coefficient de variation	0.37362
Coefficient d'obliquité (Skewness)	0.443
Coefficient de Kurtosis	2.743
Médiane	210

Tableau 1 : Différents paramètres statistiques extraits de la base de données.

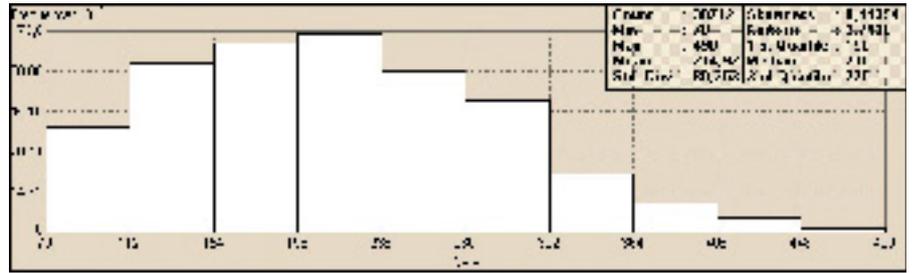


Figure 2 : Histogramme et statistique sommaire des valeurs topographiques de la base de données.

Sud au Nord: une première où se concentrent des valeurs d'altitude élevées, une deuxième des valeurs moyennes et une troisième de faibles valeurs (ne dépassant pas 150 m). La densité des données topographiques est très élevée essentiellement au niveau de la première zone et généralement faible au niveau des autres (Figure 4).

■ **Analyse des données interpolées par "cross validation"**

Les données topographiques ainsi préparées (38712 points ou valeurs altimé-

triques) sont interpolées par les différentes techniques pour produire les modèles numériques de terrain correspondants. Une analyse visuelle sur ces modèles ne donne absolument aucune information sur les différences entre les interpolateurs testés et même montre que les MNT sont identiques (Figure 5).

La méthode de la validation croisée ou "cross validation" est un ensemble des paramètres statistiques permettant d'évaluer l'erreur d'estimation pour chaque méthode d'interpolation testée et cela pour la totalité des données des grilles ou MNT générés (Tableau 2).

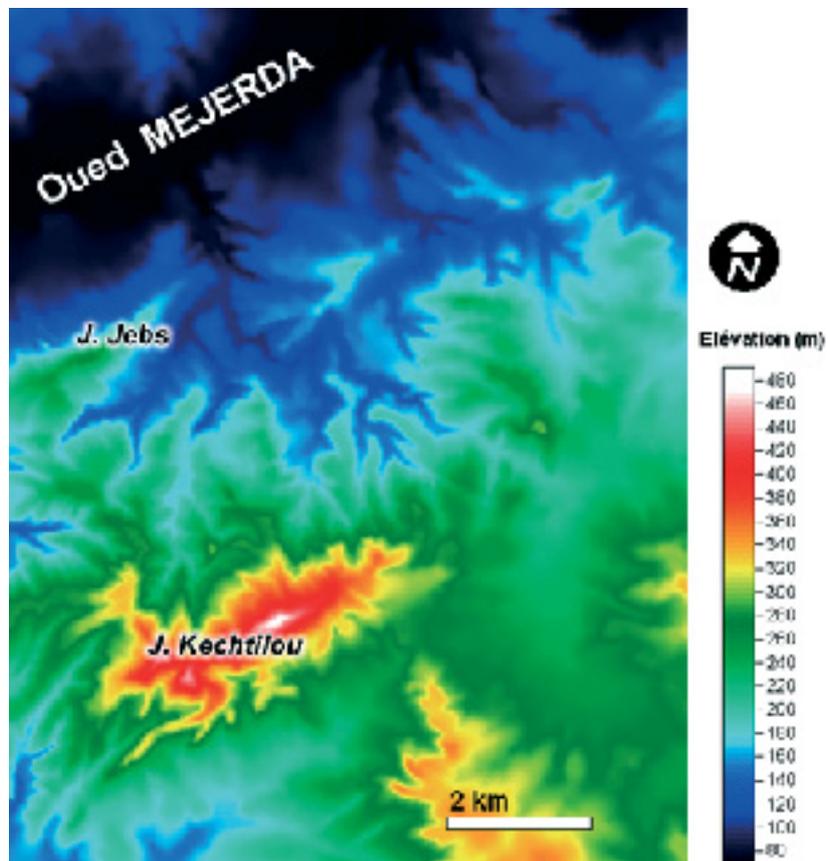


Figure 3 : Carte de distribution des altitudes.

Le tableau 4 montre que l'erreur standard (Es) présente des valeurs faibles pour les méthodes de: krigeage, l'inverse distance et le voisin naturel, alors qu'elles sont plus élevées pour la courbure minimale et le polynôme local. Ce qui est confirmé par les valeurs de l'erreur moyenne (EM) et l'erreur moyenne quadratique (RMSE). Cette dernière est considérée comme une méthode d'estimation absolue (équation 3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad [3]$$

Avec n est le nombre des données interpolées, O_i et P_i sont respectivement les valeurs observées et interpolées dans un emplacement i de coordonnées (x, y) dans l'espace géographique bidimensionnel. Plus la valeur du RMSE est faible plus l'estimation de l'interpolateur est bonne.

Le coefficient d'obliquité (Sk) indique que la méthode de la courbure minimale permet d'avoir une faible distribution asymétrique des valeurs interpolées. Pour le coefficient de régression (R) plus la valeur tend vers 1 plus l'interpolation est exacte. Or, pour l'ensemble des techniques testées R est proche de 1 sauf pour celle du polynôme local (0.781) et avec un léger avantage du krigeage par point (0.933).

■ Analyse des résiduels

Le résiduel est la différence entre la valeur observée (O_i) et la valeur interpolée ou prédite (P_i) dans un emplacement i, ce qui correspond à l'erreur de l'interpolation. Ainsi, l'étude comparative du résiduel de l'ensemble des techniques d'interpolation spatiale testées permet d'avoir d'importantes informations sur chaque méthode et d'évaluer plus précisément la marge d'erreur qu'elle produite lors de l'interpolation.

Pour le calcul du résiduel il a été tenu compte de la différence entre les valeurs d'altitude de la base de données initiales, produite à partir des données topographiques avant l'interpolation, et celles d'une base de données extraites de la grille interpolée aux mêmes emplacements. Cette opération est effectuée pour chaque méthode d'interpolation.

Méthode d'interpolation	T en s	Es	Valeur mal estimée (en m)	(EM)	RMSE	R	Sk
Inverse Distance (Inverse Distance Weighting)	17.7	0.64	260	0.009	1.42	0.848	0.66
Courbure minimale (Minimum Curvature)	4.1	0.74	310	0.049	1.52	0.814	0.38
Voisin naturel (Natural Neighbor)	4.8	0.52	150	0.004	1.13	0.870	0.68
Polynôme local (local polynomial)	8.9	1.08	320	0.032	2.12	0.781	0.61
Krigeage par point (point kriging)	38.4	0.22	140	0.003	0.63	0.933	0.52
Krigeage par bloc (block kriging)	31.5	0.23	140	0.003	0.97	0.884	0.61

Tableau 2: Résultat de l'analyse par "Cross Validation".

	Minimum (m)	Maximum (m)	Erreur standard	SD	K	RMSE
Inverse Distance	-37.5	32.6	0.3	6.0	1.18	4.82
Courbure minimale	-31.9	32.2	0.27	5.48	1.63	5.72
Voisin Naturel	-40.3	30.1	0.26	5.11	3.25	2.13
Polynôme local	-37.9	30.1	0.28	5.57	1.62	8.12
Krigeage par point	-20.5	24.9	0.24	4.83	4.07	1.73
Krigeage par bloc	-41.4	32.1	0.21	5.86	2.67	1.97

Tableau 3: Analyse statistique des résiduels de chaque méthode d'interpolation. SD c'est la déviation standard, K est le coefficient de Kurtosis et le RMSE est l'erreur moyenne quadratique.

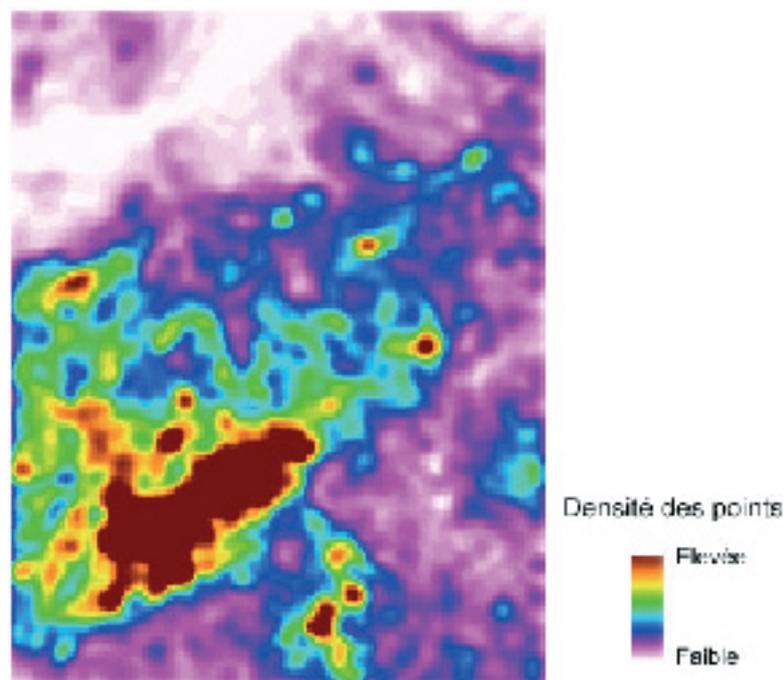


Figure 4: Carte de répartition de la densité des valeurs altimétriques $z = f(x, y)$.

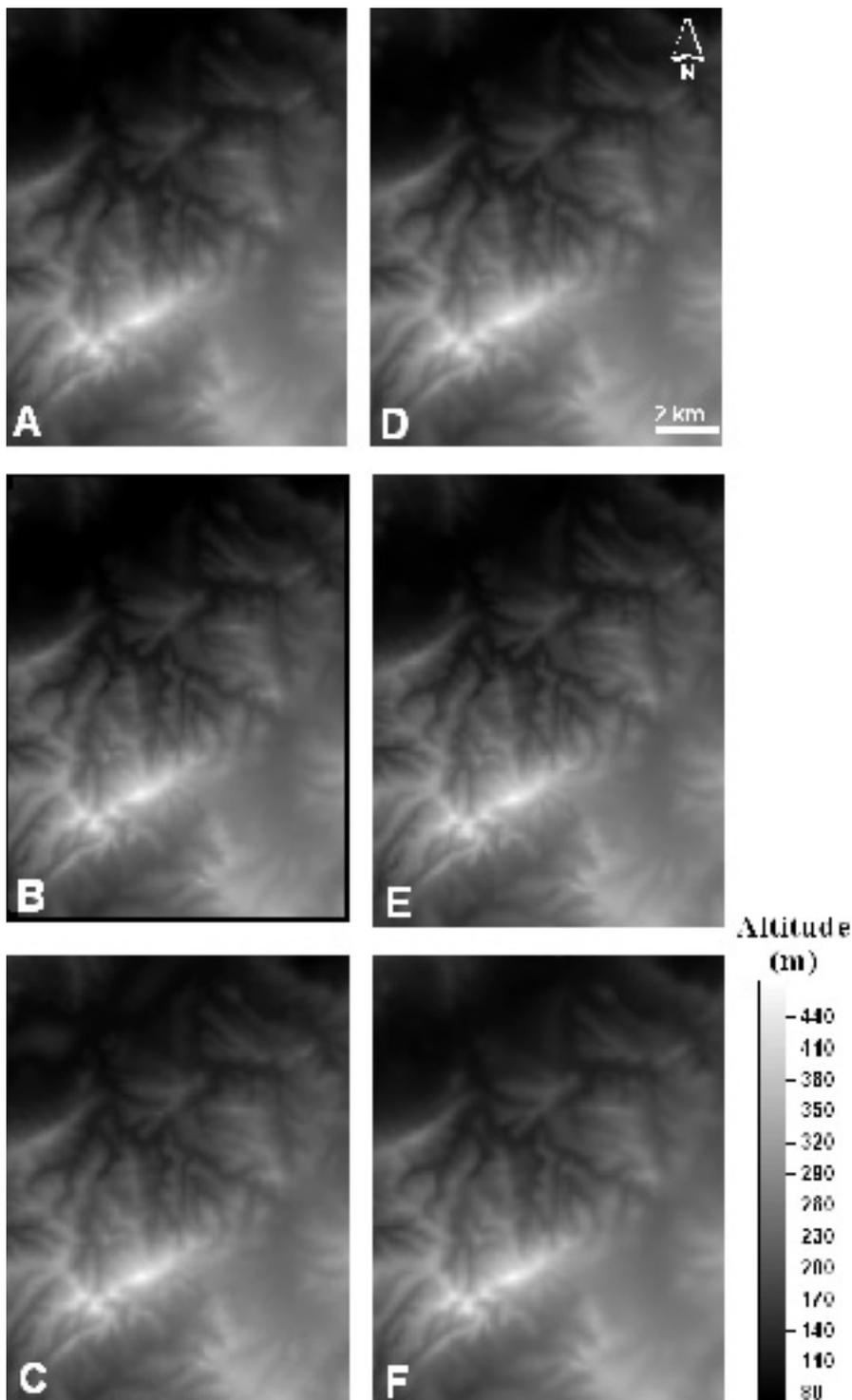


Figure 5 : Les MNT produits par les méthodes testées. A. Inverse distance, B. Voisin naturel, C. Courbure minimale, D. Krigeage par point, E. Krigeage par bloc, F. Polynôme local. La résolution des grilles est de 100 x 78 (donc 7800 cellules).

■■■ Le tableau 5 indique que la distribution de l'erreur est presque symétrique pour le krigeage par point et le voisin naturel puisqu'elles présentent les valeurs les plus élevées du coefficient de Kurtosis

(K). Ce qui veut dire que ces deux méthodes tendent à homogénéiser l'erreur au niveau de la grille produite par interpolation. Pour le krigeage par point, la différence entre le minimum et le

maximum du résiduel est la plus faible ainsi que pour la valeur du RMSE qui est très élevée dans le cas du polynôme local (8.12).

Les bases de données des valeurs résiduelles établies précédemment sont exploitées visuellement pour produire des cartes de la distribution des erreurs d'interpolation (ou résiduels) pour chaque méthode étudiée (Figure 6). Cela constitue une approche originale permettant à la fois de connaître la diffusion des erreurs dans le secteur étudié et de mettre en évidence la relation entre l'importance des résiduels et les caractéristiques topographiques du terrain (voir paragraphe *Analyse des données topographiques*).

Ces cartes constituent une importante source d'information et permettent l'exploitation statistique de plusieurs types de relations spatiales que pouvaient avoir les résiduels avec les autres paramètres topographiques du terrain.

L'analyse de la figure 6 montre que le krigeage par point présente sommairement une distribution homogène d'une valeur d'erreur généralement faible (entre -2 m et 2 m) avec localement la présence de valeurs élevées (représentées par de petites taches de couleur blanche). Pour l'inverse distance et le polynôme local la distribution des valeurs élevées de résiduel est importante sur la totalité du secteur étudié, alors que pour le reste des techniques cette distribution se localise au niveau des zones à forte topographie (partie sud). Cependant, cette dernière présente les densités les plus élevées des données d'altitude (Figure 4), mais du point de vue géomorphologique cette zone se caractérise par des changements brusques et brutaux d'altitudes ainsi que la présence de nombreuses discontinuités structurales.

Cette observation peut suggérer l'utilisation des résiduels comme une méthode de localisation de certains paramètres purement de nature géomorphologique présentant une sensibilité particulière à une ou plusieurs techniques d'interpolation spatiale. Cette perspective de recherche à envisager reste à détailler et à examiner.

■ Approche d'une analyse statistique détaillée

Pour mieux encore comparer les différentes méthodes d'interpolation étudiées, une approche statistique détaillée est mise en œuvre. Elle s'appuie sur l'analyse des échantillons extraits à partir des bases de données définissant pour chacune des méthodes les valeurs interpolées à des emplacements identiques. Six échantillons renferme donc six valeurs. (cf. tableau 4).

Deux paramètres statistiques sont utilisés pour analyser l'ensemble des échantillons par rapport aux valeurs observées à savoir absolu et relatif.

Paramètres absolus

Beaucoup de paramètres d'évaluation statistique ont été passés en revue par Legates et McCabe (1999). Ces auteurs recommandent fortement l'utilisation d'au moins un paramètre de mesure d'erreur absolue (RMSE ou EAM) pour une évaluation complète et efficace de la performance du modèle produit. Généralement, les plus utilisées sont :

- l'erreur moyenne quadratique ou RMSE (Equation 3) qui place une importance particulière aux valeurs périphériques dans le jeu de données.
- l'erreur moyenne absolue ou EMA :

$$EMA = n^{-1} \sum_{i=1}^n |O_i - P_i| \quad [4]$$

Avec n est le nombre des données utilisées (qui sont de six valeurs dans notre cas).

Paramètres relatifs

Ils sont considérés comme des indices non dimensionnels qui fournissent une comparaison relative d'un modèle par rapport à un autre (Kneale et Smith, 2001). Ces paramètres sont très recommandés pour l'évaluation des données interpolées par rapport à celles non interpolées. Il existe trois paramètres communément employés :

- le coefficient de détermination (Kneale et Smith, 2001)

$$R^2 = \left[\frac{\sum_{i=1}^n (O_i - O_m)(P_i - P_m)}{\sqrt{\sum_{i=1}^n (O_i - O_m)^2} \sqrt{\sum_{i=1}^n (P_i - P_m)^2}} \right]^2 \quad [5]$$

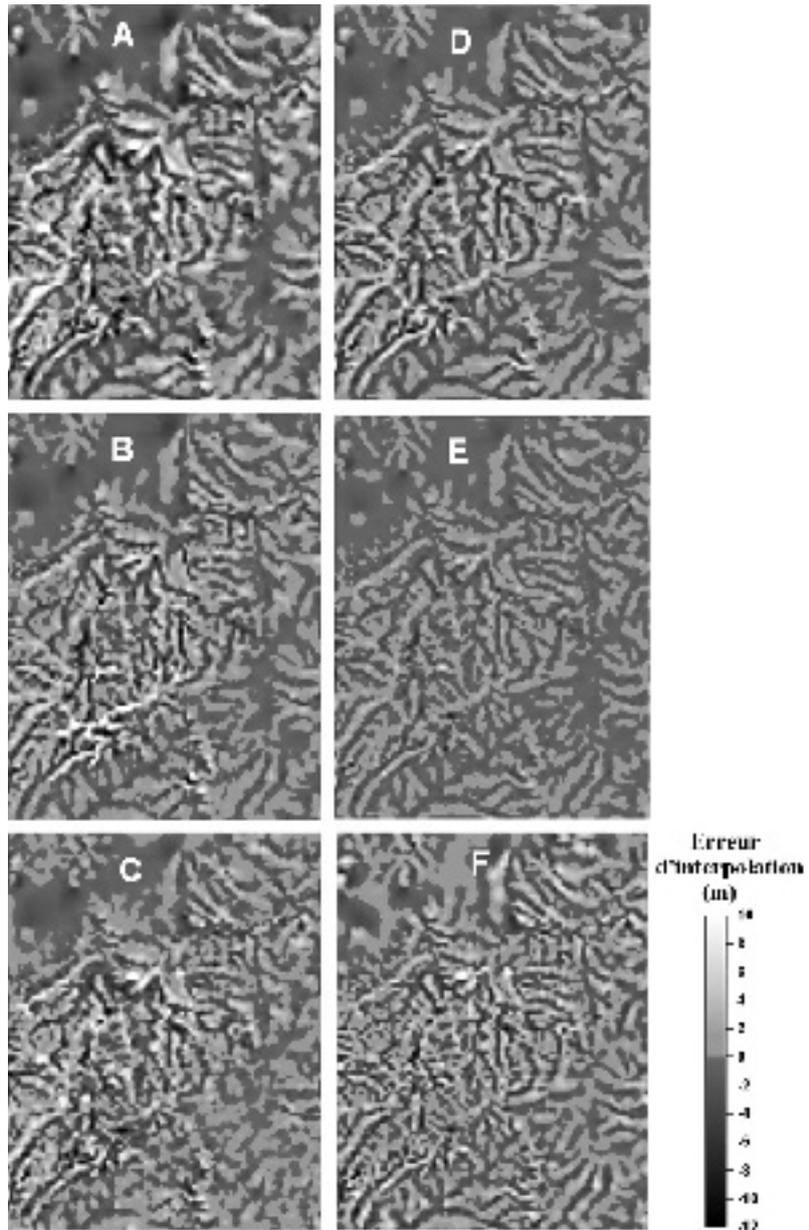


Figure 6 : Cartes de distribution des erreurs d'interpolation (ou résiduels) pour chaque méthode. (Même légende que la figure 5)

O _i	P _i					
	Krigeage par point	Krigeage par bloc	Inverse Distance	Polynôme local	Courbure minimale	Voisin Naturel
430,00	433,34	444,33	451,88	449,43	442,71	438,07
350,00	363,33	365,71	376,18	375,15	371,72	366,01
260,00	257,87	261,32	248,52	269,36	265,01	258,01
200,00	203,93	205,04	206,50	209,13	203,79	209,20
110,00	111,67	102,00	102,75	119,07	110,24	111,95
70,00	69,67	68,51	60,18	69,57	69,26	69,27

Tableau 4 : Les différents échantillons utilisés. O_i est la valeur observée dans un emplacement i et P_i est celle interpolée (prédite) au niveau du même emplacement.

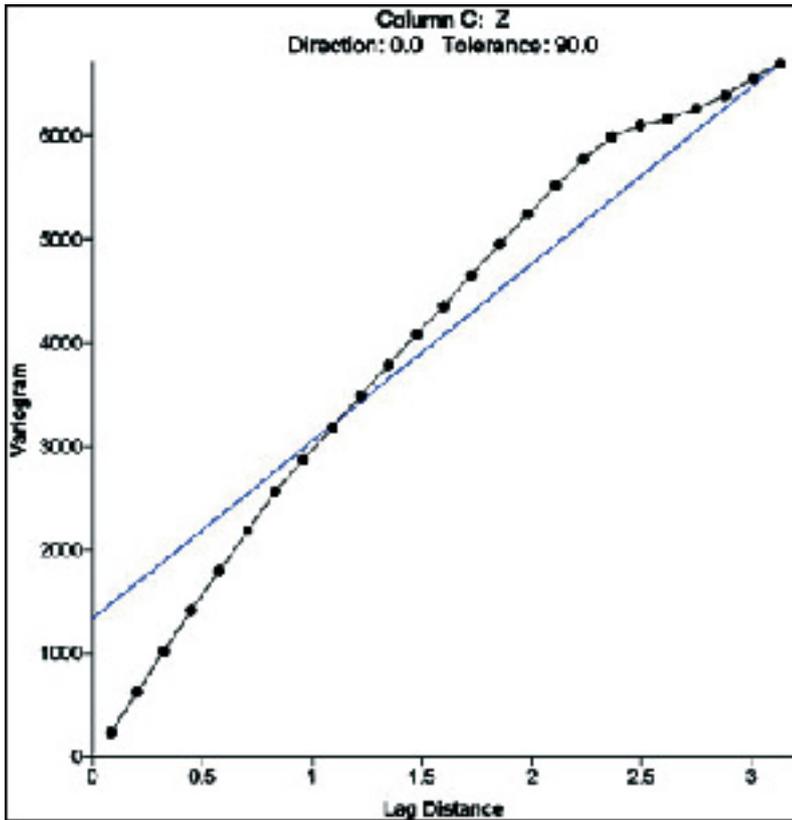


Figure 7: Variogramme expérimental obtenu à partir des données topographiques testées.

également constaté que le voisin naturel présente des résultats proches de ceux du krigeage par point. Ces deux dernières méthodes s'adaptent bien aux données topographiques testées.

L'important avantage du krigeage réside dans sa capacité à fournir la meilleure estimation linéaire possible. Il fournit également une variance d'estimation qui est fonction de la continuité spatiale, telle qu'exprimée par le "variogramme" et de la configuration (c'est-à-dire la quantité et la répartition) de l'information disponible, ce qui peut être déduit de l'analyse de la carte des densités (Figure 4).

Il est également démontré que chaque phénomène géologique possède un variogramme qui lui est propre et que généralement la topographie présente un variogramme très continu, ce qui est constaté dans l'étude (Figure 7).

De ce fait lors de l'interpolation par krigeage, on applique un variogramme de type linéaire lorsqu'il s'agit des données topographiques. Ceci a été également démontré par les travaux de Hamrah *et al.* (2006).

Enfin, la figure 8 qui est issue de l'exploitation graphique du tableau 4 résume bien le comportement de chaque méthode utilisée avec les données topographiques. Principalement l'inverse distance présente une sous-estimation au niveau des valeurs inférieures à 100 m et à 150 m et une sur-estimation à partir de 300 m sauf pour le krigeage par point où cette surestimation est la plus faible. Pour des valeurs inférieures à 200 m, la courbure minimale, le voisin naturel et le krigeage par point sont presque confondus avec la courbe de l'interpolation exacte.

Généralement, la plupart des interpolateurs présente des déficiences d'estimation des valeurs prédites principalement au niveau des zones où les altitudes changent brusquement sous l'effet d'un relief accidenté présentant des valeurs élevées de pendage. Le krigeage par point, combiné à un variogramme de type linéaire, semble être la méthode la plus adaptée.

■ ■ ■ Avec O_m est la moyenne des valeurs observées et P_m est la moyenne des valeurs interpolées (ou prédites). n est le nombre des observations utilisé. Il s'agit du "coefficient de corrélation" élevé au carré. Cette mesure permet de quantifier la force du lien linéaire entre deux variables. R^2 est compris entre 0.0 et 1.0 avec une valeur élevée indiquant un haut degré de "colinéarité".

• le coefficient d'efficacité (Nash et Sutcliffe, 1970)

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - O_m)^2} \quad [6]$$

Une valeur de 1.0 présente une interpolation parfaite. Ce paramètre est largement utilisé en hydrologie.

• l'indice d'accord (Willmott, 1981)

$$d = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n [|P_i - O_m| + |O_i - O_m|]} \quad [7]$$

d varie entre 0.0 et 1.0, si il est élevé cela indique un bon accord entre les valeurs observées (O_i) et celles interpolées (P_i). Une valeur de 1.0 indique un accord parfait alors qu'une valeur de 0.0 indique un désaccord complet. L'ensemble des résultats de cette approche est indiqué dans le tableau 5.

L'analyse du tableau 5 montre que toutes les méthodes présentent un intéressant niveau de colinéarité reflétant directement une relation linéaire entre les valeurs observées, c'est-à-dire non interpolées, et les valeurs prédites (R^2 est très proche de 1.0). Cela montre que ces méthodes fonctionnent correctement lors de l'interpolation des données topographiques et ne génèrent pas de très grandes erreurs, mais ce qui les différencie c'est leur niveau de précision altimétrique (ce qui est confirmé par le coefficient d'efficacité E). Les valeurs des paramètres absolus (RMSE et EMA) sont clairement les plus faibles pour le krigeage par points, tandis qu'elles sont les plus élevées pour l'inverse distance et le polynôme local. Il est

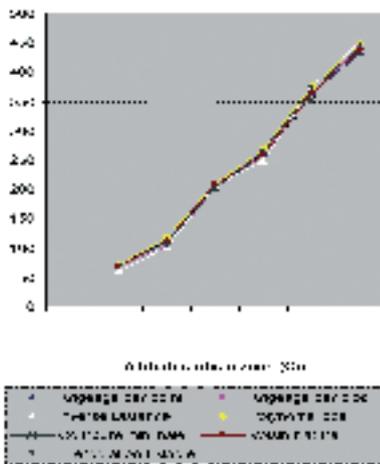


Figure 8 : Comparaison graphique entre les valeurs interpolées issues de différentes techniques utilisées avec celles des données observées.

Conclusions

Plusieurs conclusions, remarques et paramètres découlent de cette étude dont les plus importantes sont :

- la topographie du relief présente des paramètres qui peuvent influencer le résultat de l'interpolation surtout au niveau de la distribution de la densité des valeurs altimétriques ;
- une densité élevée n'implique pas forcément un bon résultat d'interpolation puisque certaines méthodes comme la courbure minimale et le polynôme local fonctionnent mal dans les zones où la densité est élevée ;
- les particularités géomorphologiques influencent considérablement la marge d'erreur de l'interpolation ;
- l'ensemble des interpolateurs montre qu'il existe une relation linéaire entre les valeurs observées et celles interpolées ;
- il est clair que les méthodes testées ne produisent pas les mêmes résultats et qu'une interprétation visuelle comparative des MNT ne le permet pas ;
- le polynôme local, la courbure minimale et l'inverse distance ne sont pas de bons interpolateurs pour les données topographiques ;
- grâce à ses propriétés variographiques le krigeage par point produit le meilleur modèle de la réalité géomorphologique du terrain avec une

	n	O _m	P _m	RMSE	EMA	R ²	E	d
Inverse Distance	6	236,67	241,00	15.74	13.85	0.9991	0.96	0.92
Courbure minimale	6		243,79	10.59	7.36	0.9989	0.98	0.96
Voisin Naturel	6		242,09	8.31	6.32	0.9992	0.99	0.97
Polynôme local	6		248,62	14.51	12.09	0.9984	0.99	0.93
Krigeage par point	6		239,97	5.94	4.12	0.9998	0.99	0.98
Krigeage par bloc	6		241,15	9.53	7.64	0.9995	0.98	0.97

Tableau 5 : Etude statistique des échantillons de différentes méthodes d'interpolation. O_m et P_m sont respectivement les moyennes des valeurs observées et interpolées.

faible marge d'erreur surtout au niveau des zones à relief accentué (de 1 à 2 m) ;

- la marge d'erreur diminue considérablement lorsqu'on augmente la résolution de la grille ;
- le krigeage par point nécessite une densité élevée de données pour avoir un meilleur résultat ainsi qu'une faible marge d'erreur. Le temps de calcul est plus élevé avec cette méthode ;
- enfin, une méthode d'interpolation est bonne pour une situation spécifique et suivant le contexte logiciel (software) dans lequel elle est exécutée et appliquée.

Les résultats montrent que le krigeage ordinaire est la technique la plus fiable pour modéliser la réalité du terrain géologique. ●

Contacts

Noamen REBAI, Tarek SLAMA et Mohamed MONCEF TURKI

FST, 02/UR/10-02: Dynamique des Bassins Sédimentaires, Paléoenvironnements et Structures Géologiques, Département de Géologie, Faculté des Sciences de Tunis, El Manar II, P.O. 2092. Campus Universitaire, Tunisie
noamenrebai@yahoo.com

Bibliographie

- Briggs I.C.** "Machine contouring using Minimum Curvature", Geophysics, vol. 39, 1974, 39-48.
- Collins F.C. - Bolstad P.V.** "A comparison of spatial interpolation techniques in temperature estimation", Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling, New Mexico, 21-25 janvier, 1996, Santa Barbara, California, CD-ROM.
- Cooper G.R.J.** "Gridding gravity data using an equivalent layer", Computers & Geosciences, vol. 26, 2000, 227-233.
- Creutin J.D. - Obled C.** "Objective analysis and mapping techniques for rainfall fields: An objective comparison", Water Resources Research, vol. 18, 1982, 413-431.
- Franke R.** "Scattered data interpolation: tests of some methods", Mathematics and Computation, vol. 38, 1980, 181-200.
- Gaspar J.L. - Goulart C. - Queiroz G. - Silveira D., Gomes A.** "Dynamic structure and data sets of a GIS database for geological risk in the Azores volcanic islands", Natural Hazards and Earth System Sciences, vol. 4, 2004, 233-242.
- Grohmann C.H.** "Morphometric analysis in geographic information systems: applications of free software GRASS and R", Computer and Geosciences, vol. 30, 2004, 1055-1067.

- ■ ■ **Hamrah M. - Shajae D. - Mosavi, A.** "Evaluation of DTM generation in Surfer 8.0", MapIndia, 2006, www.gisdevelopment.net
- Kneale P. - See L. - Smith A.** "Towards defining measures for neural network forecasting models", Proceeding of the 6th International Conference on Geocomputation, University of Queensland, 24-26 September 2001, Brisbane, Australia, 61-72.
- Lam N.S.** "Spatial interpolation methods: A review", American Cartography, vol. 10, 1983, 129-149.
- Laslett G.M. - McBratney A.B. - Pahl P.J. - Hutchinson M.F.** "Comparison of several spatial prediction methods for soil pH", Journal of Soil Science, vol. 38, 1987, 325-341.
- Legates D.R. - McCabe G.J.** "Evaluating the use of the "goodness-of-fit" measure in hydrologic and hydroclimatic model validation", Water Resources Research, vol. 35, 1999, 233-241.
- Matheron G.** "Principales of geostatistics", Economic Geology, vol. 58, 1963, 1246-1266.
- Matheron G.** *La théorie des variables régionalisées et ses applications*. Les cahiers du centre de morphologie mathématique de Fontainebleau, Ecole Nationale Supérieure des Mines, Paris, Fascicule 5, 1970, p. 212.
- Mendonça C.A. - Silva J.B.C.** "Interpolation of potential field data by equivalent layer and minimum curvature: a comparative analysis", Geophysics, vol. 60, 1995, 399-407.
- Nash J.E. - Sutcliffe J.V.** "River flow forecasting through conceptual models, I, A discussion of principals", Journal of Hydrology, vol. 10, 1970, 282-290.
- Phillips D.L. - Dolph J. - Marks D.** "A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain", Agriculture and Forest Meteorology, vol. 58, 1992, 119-141.
- Rhind D.** "A skeletal overview of spatial interpolation techniques", Computer Applications, vol. 2, n° 3/4, 1975, 293-309.
- Sarközy F.** "GIS functions - Interpolation", Periodica Polytechnica

Civil Engineerings, vol. 43, n° 1, 1999, 63-86.

Sibson R. "A Brief Description of Natural Neighbor Interpolation, Interpreting Multivariate Data", John Wiley and Sons, Barnett editor, New York, 1981, 21-36.

Smith W. H. F. - Wessel P. "Gridding with Continuous Curvature Splines in Tension", Geophysics, vol. 55, n° 3, 1990, 293-305.

Stein A. - Hoogerwerf M. - Bouma J. "Use of soil map delineations to improve kriging of point data on moisture deficit", Geoderma, vol. 43, 1988, 163-177.

Stein A. - Bouma J. - Kroonenberg S.B. - Cobben S. "Sequential sampling to measure the infiltration rate within relatively homogeneous soil units", Catena, vol. 16, 1989, 91-100.

Tabios G.Q. - Salas J.D. "A comparative analysis of techniques for spatial interpolation of precipitation", Water Resources Bulletin, vol. 21, 1985, 365-380.

Willmott C.J. "On the validation of models", Physical Geography, vol. 2, 1981, 148-194.

ABSTRACT

Keywords: DEM, GIS, interpolation, statistical analysis, topographic data.

The main Strength of geographic information system (GIS) is found in the spatial information analysis. These latter inquires the spatial continuity parameter and a regular gridded surface that can be performed only by interpolation procedure.

This paper presents a quantitative approach to evaluate six interpolation methods used in DEM production from vector topographic data into GIS environment.

Different statistical parameters are used and analyzed to assess these methods and to determine the best one(s). Our results show that ordinary kriging performing better overall in the field of geological terrain modeling.

Journées de la Topographie de Strasbourg 2006 Remise des prix AFT 2005 - Erratum



Dans le précédent numéro d'XYZ (n°109 p22) la légende faisait défaut. On peut reconnaître sur cette photo : De gauche à droite Alain Gaudet (OGE), André Bailly (AFT), Sébastien Vera (2^e du prix AFT), Pauline Lucas (2^e du prix AFT), Alexander Hoffmann-Blodau (Topcon) et Pierre Grussenmeyer (AFT).