VARIABILITÉ DES MESURES DE CARACTÉRISTIQUES HYDRODYNAMIQUES

par

A. LOBERT * et Y. CORMARY * *

SOMMAIRE

On se propose, sur quelques méthodes d'analyse de physique du sol, de montrer l'importance de la variabilité des résultats et comment en aborder l'étude par la statistique. Celle-ci permet, en particulier, de préciser dans quelle mesure chaque étape de l'obtention d'un résultat intervient dans la dispersion totale. On peut séparer, par exemple, la variabilité du sol de celle du laboratoire. Comme on le verra, la variabilité du sol, pour les méthodes exploitées, l'emporte sur celle du laboratoire. C'est donc d'abord en multipliant les prélèvements qu'on gagnera en précision, plutôt qu'en exigeant plus de soins au laboratoire.

On trouvera encore des exemples de calcul du nombre de répétitions exigées pour accéder à une précision donnée et on constatera combien cette valeur peut varier d'une méthode à l'autre.

* Société Centrale pour l'Equipement du Territoire, Pédologue au Centre de Génie rural de Tunisie.
** Chef de Mission de la Société Centrale pour l'Equipement du Territoire, Centre du Génie rural de Tunisie.

INTRODUCTION

1 - Position du problème

En matière de dispersion de mesures, aucune étude n'est définitive et extrapolable puisque cette dispersion dépend du type de mesure, des méthodes, du personnel.

Dans le cas présent il s'agit de séparer la variance systématique dûe au type de sol qui nous intéresse, de la variabilité aléatoire et nuisible dûe à tout autre facteur. Chaque fois qu'un technicien fournit des résultats, il devrait joindre une estimation de la précision de ses mesures, ce qui suppose qu'il les aura répétées et analysées. Si cette habitude était acquise, l'étude qui suit aurait peu d'intérêt.

En attendant, elle peut déjà permettre :

- à ceux qui emploient les résultats dans leurs projets, de voir combien la dispersion varie d'une mesure à l'autre, et jusqu'où peut mener l'utilisation sans précautions d'un résultat.

  Ces précautions iraient jusqu'à éviter chaque fois que possible certaines méthodes. Nous avons résumé ci-dessous les résultats essentiels, pour ne pas avoir à les chercher parmi de fastidieux calculs.

- à ceux qui font les mesures, d'y trouver des "recettes" parmi d'autres pour aborder le problème, et en particulier la nécessité de faire un certain nombre de répétitions pour chaque mesure.

Nous n'avons étudié ici que la variabilité de certaines méthodes appliquées à un certain sol, laissant de côté des problèmes plus vastes et théoriques, qui reviennent à savoir ce que représente réellement une mesure ; par exemple :

- une certaine caractéristique a-t-elle bien la signification physique qu'on lui attribue (ex. : le mouvement de l'eau vers le bas s'arrête quand le sol atteint la capacité de rétention) ;
- la méthode de mesure donne-t-elle en moyenne une valeur exacte de la caractéristique à mesurer (ex. : capacité de rétention au champ mesurée par l'humidité équivalente à la centrifugeuse).

Les prélèvements ont été faits à Tunis dans une zone d'alluvions limono-argileuses pédologiquement homogène. On dispose en outre, pour la perméabilité, de prélèvements à Hendi-Zitoun, également dans des alluvions limono-argileuses.

Les perméabilités sont mesurées par la méthode Vergières (prélèvement d'un cube de terre d'arête 10 cm, non remanié, auquel on applique, en laboratoire, une charge constante).

La densité apparente est également mesurée par la méthode Vergières sur le bloc de terre décrit précédemment. (1)

L'humidité équivalente est mesurée à la centrifugeuse et l'humidité en place par prélèvement, à la tarière, d'un échantillon d'une centaine de grammes qu'on sèche à l'étude pendant 24 heures.

Deux composantes interviennent dans la dispersion des résultats, la variabilité de la mesure au laboratoire et la variabilité du sol.

Pour connaître la première, il est nécessaire de faire sur un même prélèvement au moins deux mesures.

Pour la variabilité du sol, on a cherché à estimer non seulement l'importance de cette variabilité et ses conséquences (nombre d'échantillons à prélever, valeur de la méthode) mais encore, si la variabilité sur de petites surfaces ($\neq$ 4 m2) est inférieure à la variabilité sur des parcelles plus grandes.

Par la suite, on appellera d'ailleurs petites parcelles, des parcelles inférieures ou égales à 10 m2, et grandes parcelles, celles qui couvrent au moins 300 m2.

2 - Principaux résultats

Pour les méthodes testées, la variabilité du sol est toujours bien supérieure à celle de la variabilité de la mesure au laboratoire.

(1) Voir Cahiers C.R.E.G.R., n°2.

C'est donc cette variabilité du sol qui détermine essentiellement le nombre de prélèvements pour chaque type de mesure. En pratique, on peut donc admettre que ce nombre de prélèvements reste un minimum valable pour d'autres méthodes.

D'un autre point de vue, il faut prendre soin de répartir les prélèvements sur toute la surface de la zone à étudier.

2.1 - HUMIDITÉ ÉQUIVALENTE (Centrifugeuse)

La précision de la mesure de laboratoire est excellente et la variabilité du sol relativement faible. En pratique deux à trois prélèvements suffisent.

Dans ce cas, 95 fois sur cent l'erreur sera inférieure à 1,4 point.

2.2 - DENSITÉ APPARENTE (Méthode VERGIÈRES)

La précision au laboratoire est bonne mais la variabilité du sol impose de faire trois prélèvements pour lesquels l'intervalle de confiance au seuil 0,05 est encore de 0,06 point.

Par exemple, pour Da mesurée = 1,4, il arrivera cinq fois sur cent que la vraie valeur soit supérieure à 1,46 ou inférieure à 1,34.

2.3 - STOCK THÉORIQUE ET STOCK UTILE

Ces données sont calculées à partir de l'humidité équivalente et de la densité apparente. La densité apparente est la composante essentielle de l'erreur dans le calcul de ces stocks. Avec deux à trois prélèvements pour l'humidité équivalente, et trois pour la densité apparente, on arrive pour les stocks à une précision de 5 %, suffisante, puisque l'imprécision vient surtout ici de la difficulté de définir exactement la profondeur de sol exploitée par les racines.

2.4 - PERMÉABILITÉ

La dispersion des méthodes "ponctuelles" (Vergières, Porchet, Muntz, etc..) est telle qu'il faut distinguer deux cas :

- on ne cherche qu'un élément qui permette avec d'autres de classer les sols suivant leur aptitude à l'irrigation. Pour cela trois prélèvements par horizon assurent qu'on ne se trompera qu'une fois sur dix, affirmant que la perméabilité est comprise dans un intervalle de un à dix (ex. : entre 1 et 10 mm/heure ou entre 10 et 100, etc..).

  On prend la moyenne arithmétique des trois prélèvements en éliminant éventuellement un des résultats suivant une règle empirique décrite dans l'exposé;
  L'application de cette règle au périmètre de Sidi Ali Salem a donné des résultats satisfaisants, mais il ne faut pas oublier que la mesure n'amène qu'à une définition largement qualitative de la perméabilité.

- au contraire, dès qu'il faut inclure la perméabilité (ou tout autre paramètre analogue : taux d'infiltraction, etc..) dans le calcul de projets d'irrigation ou de drainage, il est hautement préférable de faire un essai direct, de mettre en oeuvre un modèle utilisant les éléments qu'on veut déterminer (ex. : essais d'avancement de l'eau dans les sillons, essais d'asperseurs, variation du niveau de la nappe).
  En dernier recours, s'il est impossible de procéder à de tels essais on ne peut que conseiller de faire des échantillons de 10 prélèvements au moins par horizon et d'ajuster graphiquement chaque échantillon à une distribution gausso-logarithmique.

Il serait intéressant d'utiliser un mode de prélèvement plus rapide que la méthode Vergières. L'intérêt de celle-ci se limite au premier cas où le nombre de prélèvements est de trois, sur lesquels on peut également mesurer la densité apparente. Signalons encore que la méthode de Muntz, qui permet de mesurer pour chaque sol la loi d'infiltration d'une lame d'eau dans un sol sec, est beaucoup plus proche de la réalité pour l'étude de l'irrigation, alors que la méthode Vergières, où l'on mesure la perméabilité en milieu saturé s'appliquerait mieux au drainage (ou la méthode de Hoogdhoudt, si on se trouve en présence d'une nappe).

2.5 - HUMIDITÉ DU SOL (Méthode gravimétrique)

La variabilité du sol est très forte : le nombre de sondages, la taille du prélèvement (en particulier l'épaisseur prélevée dans chaque tranche de sol), le nombre de prélèvements par sondage influent sur la précision de l'estimation d'un profil hydrique moyen.

Avec des échantillons de 100 gr et de 10 cm de hauteur par tranche de sol de 25 cm, l'écart type était égal à 5. Avec des échantillons de 300 à 400 gr, de 20 cm, par tranche de sol de 25 cm, l'écart-type d'un résultat s'abaisse à 2,5.

La variabilité augmentera en général avec la taille de la parcelle. Les humidités successives dans un sondage ne sont pas indépendantes. Donc, si on veut avoir les humidités moyennes d'une parcelle on a intérêt à multiplier les trous de sondage plutôt que les prélèvements à l'intérieur d'un sondage.

Actuellement on prélève des échantillons de 20 cm de hauteur par tranche de sol de 25 cm et on fait au moins neuf sondages pour 1 à 2 hectares. Cette méthode donne une bonne précision pour le calcul de la réserve en eau du sol à un certain moment, mais reste peut-être insuffisante pour le calcul d'un bilan.

I - HUMIDITÉ ÉQUIVALENTE

1 - Plan de prélèvement et analyse de variance

Les prélèvements ont été effectués à quatre profondeurs : 12, 37, 62 et 87 cm, en quatre "points" distants de 60 m et disposés en carré. En chaque "point" on fait deux "profils" distants de un mètre.

On a donc : $4 \times 2 \times 4 = 32$ prélèvements.

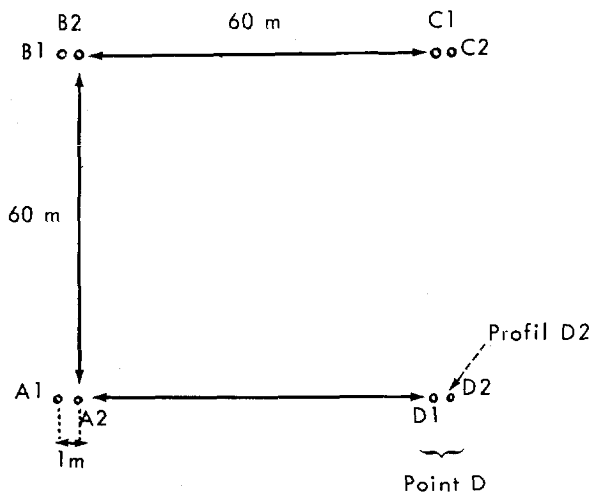
Sur chaque prélèvement tamisé à 2 mm, on fait deux mesures au laboratoire, soit 64 en tout.

Chaque analyse statistique ne porte évidemment que sur les échantillons et mesures issus d'une même profondeur, soit 16 résultats par analyse.

L'analyse de variance (voir en annexe) permet de :

- connaître la variabilité de la mesure au laboratoire ;
- savoir si la variabilité entre profils distants d'un mètre est importante ;
- savoir si la variabilité entre points distants de 60 m est supérieure à la précédente.

Plan de prélèvement



2 - Résultats de l'analyse

Profondeur	Moyenne observée	Ecart-type		
		S	Sj	Si
12	28	0,21	1,4	-
37	26,9	0,09	0,66	1,2
62	27,3	0,15	0,61	1,1
87	26,5	0,1	1,3	1,4

- On s'est assuré qu'il n'y avait pas de variation systématique (ex. C et D supérieurs à A et B) c'est-à-dire que la zone est bien homogène ;
- S est l'écart-type d'une mesure au laboratoire. Il est très toujours très faible. La moyenne des quatre estimations de S donne S moyen $\#$ 0,2, ce qui signifie que sur une série de 100 mesures effectuées sur un même prélèvement aussi bien homogénéisé que possible, 5 seulement s'écarteront d'environ 0,4 points de la moyenne des 100 mesures.
  Donc la dispersion de la méthode laboratoire est négligeable. On pourrait se contenter d'une mesure par prélèvement, mais la nécessité d'équilibrer la centrifugeuse impose qu'on en fasse deux.
- Sj est une estimation de l'écart-type moyen entre valeurs issues de deux profils distants de un mètre. On élimine pour calculer Sj, la variance précédente dûe à la mesure. Sj (comme Si ci-dessous) n'a pour origine que la variabilité dans le sol.
- Si est de même l'estimation de l'écart-type entre valeurs d'humidité dans des points écartés de 60 m (on élimine les deux variances précédentes S_2 et $S_{2}j$).

On constate :

- que la variabilité du sol existe toujours et l'emporte de beaucoup sur la variabilité de la mesure ;
- que trois fois sur quatre la variabilité entre les "points" éloignés est supérieure à la variabilité entre les "profils" proches l'un de l'autre.

C'est-à-dire qu'il existe de petits effets locaux et des effets plus importants sur de plus grandes surfaces, même à l'intérieur d'une zone qu'on admet être homogène.

Puisqu'en général on aura à déterminer l'humidité équivalente de grandes parcelles, on répartira donc les prélèvements sur toute la parcelle.

3 - Nombre optimum de prélèvements

Une estimation approchée de l'écart-type d'un seul résultat pris dans une "grande parcelle" et à condition de faire deux mesures de laboratoire (condition imposée par la centrifugeuse) est égale à 1,2.

Le tableau suivant donne l'écart-type $S_{\bar{x}}$ de moyennes de 1, 2, 3, 4, 5 prélèvements dans un horizon pédologiquement homogène et les intervalles de confiance I (x) de ces moyennes au seuil de probabilité 0,05.

Nombre de prélèvements	Ecart-type $S_{\bar{x}}$	I (x) P = 0,05
1	1,2	$\pm$ 2,4
2	0,9	1,8
3	0,7	1,4
4	0,6	1,2
5	0,5	1

Avec trois prélèvements on pourra calculer la moyenne, par exemple 25, en n'oubliant pas qu'il y a cinq chances sur cent pour que la valeur réelle de l'humidité équivalente de cet horizon soit supérieure à 25 + 1,4 = 26,4 ou inférieure à 25 - 2,4 = 23,6.

On peut, dans le tableau précédent, choisir un nombre de prélèvements correspondants à la précision requise. Il semble toutefois difficile d'en faire moins de deux. D'autre part, au delà de trois prélèvements le gain de précision est trop faible pour justifier des dépenses supplémentaires.

Enfin, on peut penser que, puisque la variabilité vient surtout du sol, toute méthode correcte de mesure autre que la centrifugeuse donnerait des dispersions à peu près analogues et ceci même si on ne fait qu'une mesure de laboratoire.

En résumé, sur une zone homogène il semble raisonnable de faire deux à trois trous répartis sur toute la zone et dans chaque trou, de prélever un échantillon par horizon ou couche de sol. On pourra dire en ne se trompant que cinq fois sur cent qu'on connaît la valeur de l'humidité équivalente à 1 ou 2 points près. Des écarts supérieurs à cinq points, sans être impossibles, sont fort peu probables (moins de cinq fois sur mille).

II - DENSITÉ APPARENTE

1 - Comparaison de la variabilité de la mesure au laboratoire à celle du sol

On a d'abord vérifié que la méthode de mesure ne changeait pas la densité apparente. Pour cela on a prélevé huit cubes de terre dans un carré de 1 x 1 m, à 40 cm de profondeur. On a fait une première série de mesures sur ces huit cubes. On a ensuite enlevé la paraffine de chaque cube, et on a recommencé une nouvelle série de huit mesures sur les mêmes cubes.

L'analyse statistique montre :

- qu'il n'y a pas de différence systématique entre les résultats des deux séries, donc que la manipulation au laboratoire n'altère pas la valeur de la densité apparente du sol (bien que le prélèvement puisse le faire) ;
- que l'écart-type d'une mesure de laboratoire est d'environ 0,02 ;
- que la différence entre les huit cubes, donc la variabilité du sol, même sur une petite surface, est plus importante que la variabilité de la mesure au laboratoire. On a pu estimer que l'écart-type d'un résultat (compte-tenu de la variabilité au laboratoire, de celle du sol, et de celle induite éventuellement par le prélèvement) est déjà de 0,05 pour des valeurs qui vont en général de 1,3 à 1,7.

2 - Autre estimation de la variabilité du sol

A l'intérieur des quatre carrés de un mètre de côté, et distants l'un de l'autre de 60 m, on a fait chaque fois quatre prélèvements, soit 16 au total, entre 10 et 20 cm de profondeur.

La variabilité entre carrés n'est pas supérieure à la variabilité des échantillons à l'intérieur d'un carré ce qui signifie qu'à la différence de l'humidité équivalente, la variabilité de la densité apparente dans une zone pédologique homogène est dûe surtout à des effets très localisés (fissures). On pourrait donc, par mesure d'économie, faire les prélèvements nécessaires à une estimation correcte de la densité apparente d'une zone homogène dans une seule tranchée. Il faut toutefois reconnaître que des prélèvements en différents points peuvent mettre en évidence une erreur dans la limitation de cette zone homogène.

3 - Nombre optimum de prélèvements

A partir de l'ensemble de tous les résultats, on peut estimer que l'écart-type d'un résultat reste égal à 0,05.

Tableau des intervalles de confiance I (x) de la moyenne, en fonction

du nombre de prélèvements

Nombre de prélèvements	Ecart-type $S_{\bar{x}}$	I (x) P = 0,05
1	0,05	$\pm$ 0,1
2	0,035	0,070
3	0,029	0,06
4	0,025	0,05
5	0,022	0,044
6	0,0204	0,040

On remarquera :

- qu'avec un seul prélèvement, en affirmant par exemple que la densité apparente est de 1,4, il arrivera une fois sur vingt qu'elle séra inférieure à 1,3 ou supérieure à 1,5.
  Comme la densité apparente des sols varie de 1,25 à 1,75, qu'elle est souvent voisine de 1,3 dans les labours, de 1,5 à 1,6 dans les horizons profonds limono-argileux, il revient à peu près au même d'admettre que ces valeurs sont exactes dans tous les cas, que de faire un seul prélèvement.
- l'optimum semble être de trois prélèvements compte-tenu du temps nécessaire à réussir un prélèvement Vergières. Pour une densité moyenne de 1,4 par exemple, calculée d'après trois résultats, il arrivera une fois sur vingt que la valeur réelle soit supérieure à 1,46 ou inférieure à 1,34. Mais des écarts de 0,1 point (comme pour le cas d'un seul prélèvement) n'arriveront que cinq fois sur mille.

Enfin, comme pour l'humidité équivalente, la variabilité du sol est la composante essentielle de la variabilité totale d'un résultat, laquelle ne dépendra pas beaucoup de la méthode de mesure.

4 - Dispersion des résultats de stock théorique et de stock utile

On calcule le stock théorique S que peut contenir le sol d'après :

$$S \text{ en mm} = \sum_i \frac{Da_i}{10} \cdot HE_i \cdot E_i = \sum S_i \quad (1)$$

D_{a_i} = Densité apparente de la tranche de sol i

HE_i = Humidité équivalente de la tranche i

E_i = Epaisseur de la tranche i.

4.1 - ÉCART-TYPE DE $Z = Da \times HE$

On a :

$$\frac{S_Z}{Z} = \sqrt{\left(\frac{S_{Da}}{Da}\right)^2 + \left(\frac{S_{HE}}{HE}\right)^2} \qquad (2)$$

S_Z, S_{Da}, S_{HE} = Ecarts-types

Exemple : pour Da = 1,5, HE = 25 et une seule mesure, on trouve d'après les estimations précédentes de S_{Da} et S_{HE} : $\frac{S_Z}{Z} \# 0,06$.

4.2 - ÉCART-TYPE DU STOCK S_i D'UN HORIZON i

Soit y, variable aléatoire, combinaison linéaire de variables aléatoires x_i .

$$y = \sum a_i \quad x_i$$

On a $S^2 y = \sum a_i^2 \quad S^2 x_i$ (3)

d'où : $S_{s_i} = \frac{Ei}{10} \quad S_Z$

Et si on admet que S_Z reste très peu différent de 0,06 Z quel que soit Z, on a :

$$\frac{S_{s_i}}{s_i} = \frac{S_Z}{Z} = 0,06$$

4.3 - ÉCART-TYPE DU STOCK THÉORIQUE $S = \sum S_i$

D'après (3) on a :

$$S_s^2 = \sum S^2_{s_i} = \sum (0,06 \; s_i)^2$$

$S_s = 0,06 \sqrt{\sum s_i^2}$, pour une seule mesure de Da et de HE dans chaque horizon

Pour N mesures : $S_s = \frac{\sqrt{\sum s_i^2}}{17 \sqrt{N}}$

et l'intervalle de confiance I (S) au seuil de probabilité 0,05 sera environ :

$$I\ (S) = \pm\ 0,12 \sqrt{\frac{\sum s_i^2}{N}}$$

Exemple :

Profondeur	Da	HE	E	S_i	S_i^2
0 - 30	1,3	25	30	97,5	9.506
30 - 60	1,5	28	30	126	15.876
60 - 100	1,6	20	20	128	16.384
			Totaux	351,5	41.766

Pour $I(S) = 0,12 \sqrt{\frac{41.766}{N}}$ on obtient :

N	I (S)
1	24
2	17
3	14

N = nombre de mesures de Da et de HE. L'intervalle de confiance de 14 mm peut être admis (précision de 4 à 5 %) compte-tenu de l'emploi de ces stocks théoriques. L'erreur vient beaucoup plus de la difficulté de définir exactement la tranche de sol exploitée par les racines (tranche de sol à amener à la capacité de rétention) que de la variabilité des mesures. Le **Stock utile**, ou la dose unitaire (dose à amener à chaque irrigation) est définie comme une fraction du stock théorique R.

$D = k R \qquad k = 0,5$ à $0,75$ généralement

Intervalle de confiance $I(D) = k\, I(R)$.

C'est-à-dire que la précision relative ne change pas et reste, avec trois mesures, d'environ 4 à 5 %.

5 - Application

En déterminant la densité apparente d'une petite zone irriguée, on a obtenu huit séries de trois résultats, chacune relative à un horizon différent.

L'analyse statistique a montré que l'écart-type d'un résultat est égal à 0,06, donc très peu différent du résultat précédent.

III - PERMÉABILITÉ

1 - Résultats

On dispose :

- de trois séries de résultats pris à Tunis en une même tranchée à des profondeurs de 15, 55 et 85 cm ;
- des cubes ayant servi à l'analyse de la densité apparente, cubes prélevés en surface (15 cm) au même endroit que les trois séries ci-dessus et suivant le plan donné au chapitre précédent ;
- de prélèvements à Hendi-Zitoun : on n'a gardé que ceux entre 25 et 50 qui forment une série et ceux entre 50 et 75 qui en forment une autre.

Les prélèvements de surface à Hendi-Zitoun étalés sur plusieurs mois ont été éliminés car des travaux de mise en culture ont pu briser l'homogénéité de la série. On a donc cinq échantillons, pris chacun dans un horizon pédologique homogène.

Résultats en mm/heure (coefficient K de Darcy).

Tunis (15 cm)

7 - 11 - 12 - 14 - 16 - 20 - 22 - 29 - 31 - 32 - 47 - 54 - 59 - 68 - 76 - 85 - 92 - 115 - 122 - 130 - 142 - 151 - 180 - 576 - 684.

Tunis (55 cm)

1,8 - 5,6 - 9,4 - 9,5 - 10,4 - 11,7 - 14,5 - 16,5 - 18,2 - 20.

Tunis (85 cm)

1,4 - 2 - 2 - 3 - 4,4 - 6 - 8,8 - 14,3.

Hendi-Zitoun (25-50 cm)

0,12 - 0,4 - 0,54 - 0,63 - 0,9 - 1 - 1 - 1,1 - 1,7 - 1,8 - 2,9 - 3,8 - 4,9 - 6 - 6,3 - 7,2 - 17 - 35 - 41.

Hendi-Zitoun (50-75 cm)

0,03 - 0,05 - 0,05 - 0,07 - 0,12 - 0,2 - 0,2 - 0,29 - 0,3 - 0,3 - 0,32 - 0,33 - 0,4 - 0,45 - 0,5 - 4,2 - 9,9.

2 - Analyse des résultats

On peut vérifier graphiquement que la distribution n'est pas gaussienne, qu'elle semble, pour les échantillons recueillis, plus "aplatie" qu'une distribution gaussienne (c'est-à-dire que des écarts très importants apparaissent plus souvent que dans la distribution gaussienne) et qu'elle est dissymétrique (Il peut apparaître assez souvent des écarts positifs importants par rapport à la moyenne ou à la médiane de la population).

Il est probable que pour une couche homogène de sol, les surfaces des pores sont distribuées normalement, mais les débits, donc les cœfficients K, sont proportionnels au carré des surfaces, selon la loi de Poiseuille. Ils ne peuvent donc être distribués à peu près normalement que dans le cas exceptionnel où les surfaces des pores varient très peu autour de leur moyenne, soit en des sols de structure très homogène. Dans les sols lourds, hétérogènes, susceptibles de se fissurer par dessiccation on aura souvent des distributions dissymétriques (cf. distribution de χ^2, somme de carrés de variables normales réduites).

Devant une telle dispersion, chercher la forme de la distribution ne semble pas de première nécessité, d'autant plus que cette distribution risque de changer avec le type de sol.

Par contre, le fait que la distribution ne soit pas normale pose le problème du choix des paramètres à retenir pour représenter un échantillon de plus de dix prélèvements : moyenne arithmétique, géométrique, médiane, et de la dispersion de ces paramètres : écart-type de la moyenne, écart-médian, etc..

- si on garde la moyenne arithmétique pour caractériser un échantillon, donc une couche d'une zone homogène, on risque d'avoir des résultats systématiquement supérieurs à la réalité. Un cube de terre prélevé au champ peut être traversé de part en part, par une fissure importante, et donner au laboratoire un débit exceptionnel. Sur le terrain, quand la fissure sera remplie d'eau, ce seront les pores spécifiques de la micro-structure du sol qui deviendront facteur limitant de l'infiltration.

  De plus le calcul des intervalles de confiance de cette moyenne à partir de la variable t conduit à des résultats aberrants comme on a pu le vérifier puisque la distribution n'est pas normale.

- il serait peut-être préférable de prendre la moyenne géométrique qui donne moins d'importance aux fortes valeurs, ou la médiane qui est moins influencée par les valeurs extrêmes de la distribution. Pour un échantillon de 2p + 1 (= nombre impair de prélèvements) la médiane est la (p + l)ème valeur. Ex.: Tunis 15 cm : (25 résultats). La médiane est la treizième valeur : 59 mm/heure. Il y a autant de résultats avant la médiane qu'après (12).

Pour un échantillon de 2p prélèvements on fait la moyenne du pème et du (p+ 1)ème résultat : Ex.: Tunis 85 cm (huit résultats) médiane $= \frac{3 + 4,4}{2}$.

Cependant, de telles questions ne viennent qu'au second plan : dans la réalité on ne peut se permettre des séries de 10 à 25 échantillons par horizon. Comment faire le nombre minimum de prélèvements quand on se trouve en face d'une telle dispersion, tout en limitant les risques d'erreur grossière.

3 - Nombre de prélèvements

Les calculs suivants ne sont que très approchés et exigent un certain nombre de conventions.

On admettra que des méthodes aussi "ponctuelles" que la méthode Vergières, ne peuvent donner qu'un ordre de grandeur de la perméabilité.

En considérant les cinq séries de résultats on admet que toutes les valeurs comprises :
entre 14 et 142 pour la première série,
0,63 et 6,3 pour la deuxième série,
0,05 et 0,5 pour la troisième série,
1,8 et 18,2 pour la quatrième série,
2 et 14,3 pour la cinquième série,
sont des valeurs satisfaisantes, en ce sens qu'elles sont comprises entre deux bornes définies ci-après : dans chaque série, les deux bornes ont été choisies de manière à ce que la borne supérieure soit dix fois plus forte que l'inférieure et qu'il y ait le plus grand nombre de valeurs possible à l'intérieur des bornes.

Les valeurs en dehors des bornes sont donc considérées comme des erreurs.

La probabilité moyenne, pour l'ensemble des séries, qu'un cube de terre donne une "erreur" telle qu'on l'a définie, est de 0,23, soit environ une chance sur quatre.

En faisant usage de la loi binomiale on peut trouver pour des séries comprenant 1, 2, 3 . . n prélèvements dans une zone homogène, les probabilités P (0), P (1), P (2). . . P (n) d'avoir 0, 1, 2 . . n "erreurs" dans une série.

Probabilités	Nombre de prélèvements			
	1	2	3	4
P (0)	0,77	0,593	0,456	0,351
P (1)	0, 23	0,354	0,409	0,420
P (2)		0,053	0,122	0,188
P (3)			0,012	0,037
P (4)				0,003
Totaux	1	1	1	1

Si on prend pour définir une couche homogène de sol :

- 1 seul prélèvement : on a une chance sur quatre qu'il soit en dehors de l'intervalle défini ci-dessus, donc une chance sur quatre de commettre une "erreur" importante ;
- 2 prélèvements : on a 60 chances sur cent (P (0) : 0,593) d'avoir deux bons prélèvements, cinq chances sur cent d'en avoir deux mauvais, mais 35 chances sur cent d'en avoir un bon et un mauvais. Comme on ne sait évidemment pas lequel est le bon, en faisant la moyenne on commettra souvent une erreur importante ;
- 3 prélèvements : on a quarante-cinq chances sur cent qu'ils soient tous les trois bons et quarante chances sur cent qu'il y en ait deux bons et un mauvais. Ce cas présente un intérêt particulier puisqu'il permet d'éliminer éventuellement une des trois valeurs s'écartant beaucoup des deux autres.
  Cependant on se trompera encore douze fois sur cent en ayant deux mauvaises valeurs pour une bonne et douze fois sur mille avec trois mauvaises valeurs. Mais ces deux mauvaises valeurs peuvent être réparties de chaque côté de l'intervalle contenant la valeur correcte et leur moyenne risque de se trouver dans cet intervalle. Finalement on risque à peu près une "erreur" sur dix en faisant trois prélèvements par couche homogène ;
- 4 prélèvements : ce cas ne présente pas plus d'intérêt que le précédent car on a presque vingt chances sur cent d'avoir deux bonnes valeurs et deux mauvaises et de ne pouvoir lever le doute.

Il faut donc faire un nombre impair de prélèvements, mais dépasser trois prélèvements par horizon sera rarement possible.

4 - Exemple d'interprétation des résultats

Dans les séries Tunis 15 cm et Hendi-Zitoun 25-50, on a tiré au sort 20 fois trois valeurs. A l'intérieur d'un tirage de trois valeurs, on ne remet pas la valeur tirée dans la population, mais on la remet quand le tirage des trois valeurs est complet.

Chaque échantillon de trois valeurs aurait pu être l'échantillon de trois prélèvements qu'on aurait effectué pour déterminer la perméabilité de la couche de sol, en appliquant la règle précédente.

On a comparé cinq méthodes pour le calcul de la moyenne arithmétique.

- on garde les trois valeurs ;
- on élimine une des valeurs si la valeur la plus forte est plus de 10 fois supérieure à la plus faible.
  Ex.: 7, 76, 180. 180 est supérieur à 7 x 10. On élimine une valeur en faisant le rapport 180/76 et 76/7. 180/76 < 76/7, donc on garde 76 et 180.
- même règle que ci-dessus, mais on fait la différence 180 - 76 et 76 - 7 . 76 - 7 < 180 - 76. On garde 76 et 7.
- on élimine systématiquement une des trois valeurs en faisant les rapports ;
- on élimine systématiquement une des trois valeurs en faisant les différences.

Les résultats sont résumés dans les tableaux suivants (les colonnes 1, 2, 3, 4, 5, correspondant à ces cinq méthodes).

Tableau 1 : Série Tunis 15 cm - (en mm /heure)

Echantillons			Nombre d'erreurs valeurs extérieures à (14 - 142)				Moyennes arithmétiques				
			0	1	2	3	1°	2°	3°	4°	5°
7	11	130			+		49	9	9	9	9
7	47	68		+			41	41	41	58	58
32	130	180		+			114	114	114	155	155
12	16	31		+			20	20	20	14	14
12	47	576			+		212	30	30	30	30
14	22	180		+			72	18	18	18	18
29	76	142	+				82	82	82	109	53
20	31	68	+				40	40	40	25	25
16	47	122	+				62	62	62	85	32
59	92	130	+				94	94	94	111	76
7	76	180			+		88	128	42	128	42
11	20	31		+			21	21	21	26	16
14	31	68	+				38	38	38	50	23
16	68	130	+				71	71	71	99	42
7	61	76		+			48	69	69	69	69
31	54	151		+			79	79	79	43	43
11	14	54		+			26	26	26	13	13
14	122	130	+				89	89	89	126	126
29	59	684		+			257	44	44	44	44
12	20	76	+				36	36	36	16	16
Total			8	9	3	0					
Moyennes							76,95	55,55	51,25	61,40	45,20
Ecart - type							58	34	30	44	37

Tableau 2 : Série Hendi-Zitoun 25-50 cm en mm/heure

Echantillons			Nombre d'erreurs valeurs extérieures à (0,63-6,3)				Moyennes arithmétiques				
			0	1	2	3	1°	2°	3°	4°	5°
1	17	41			+		19,7	29,00	9,00	29,00	9,00
0,12	6,3	17			+		7,81	11,65	3,21	11,65	3,21
1,1	1,8	6,0	+				2,97	2,97	2,97	1,45	1,45
0,12	2,9	4,9		+			2,64	3,90	3,90	3,90	3,90
1,8	4,9	6,3	+				4,33	4,33	4,33	5,60	5,60
1	1	6	+				2,67	2,67	2,67	1,00	1,00
0,63	1	2,9	+				1,51	1,51	1,51	0,81	0,81
2,9	6,3	7,2		+			5,47	5,47	5,47	6,75	6,75
1	2,9	4,9	+				2,93	2,93	2,93	3,90	1,95
6,3	17	41			+		21,43	21,43	21,43	29,00	11,65
0,54	1,7	3,8		+			2,01	2,01	2,01	2,25	1,12
0,9	17	41			+		19,63	29,00	8,95	29,00	8,95
0,54	1	3,8		+			1,78	1,78	0,77	0,77	0,77
0,4	0,9	35			+		12,10	0,65	0,65	0,65	0,65
0,54	3,8	35			+		13,11	2,17	2,17	2,17	2,17
0,12	3,80	41			+		14,97	22,40	1,96	22,40	1,96
0,12	1,10	3,8		+			16,73	2,45	0,61	2,45	0,61
0,63	6	7,2		+			4,61	6,60	6,60	6,60	6,60
1	6	7,2		+			4,73	6,60	4,10	6,60	6,60
1	2,9	41		+			14,97	1,95	1,95	1,95	1,95
Total			5	8	7	0	150,12	161,47	87,19	167,90	76,70
Moyennes							7,5	8,07	4,35	8,39	3,83
Ecart-type							7,50	9,40	4,70	10,18	3,36

On constate que :

- dans la série Tunis les fréquences de 0, 1, 2, 3 erreurs sont respectivement de 8/20, 9/20, 3/20 et 0/20, donc peu différentes des probabilités calculées d'après la loi binomiale. Dans la série Hendi-Zitoun plus dispersée, la fréquence de deux erreurs est plus élevée que la normale, mais dans ce cas, cinq fois sur sept, elles sont de sens opposées donc s'annulent plus ou moins.
  Finalement, pour Tunis, on a fait une "erreur grossière" 15 fois sur cent et pour Hendi-Zitoun dix fois sur cent.
- les méthodes 3 et 5 donnent des moyennes et des écarts-types plus faibles que les autres. Dans la série Tunis normalement dispersée, c'est la méthode 3 qui donne le plus faible écart-type. On peut d'ailleurs penser qu'il ne faut éliminer une valeur qu'avec prudence.

En résumé, on peut fixer une règle, d'ailleurs très empirique :

- prendre trois prélèvements qui donnent trois résultats : $a < b < c$.
- éliminer éventuellement une valeur si $c > 10\,a$, en faisant les différences $c - b$ et $b - a$ et en gardant les deux valeurs les plus proches.
- faire la moyenne arithmétique des valeurs restantes. On remarquera que pour la densité apparente et la perméabilité, qui se déterminent sur le même bloc de terre, on a trouvé chaque fois un optimum "économique" de trois prélèvements.

5 - Variabilité de la mesure sur de petites et de grandes parcelles

On pourrait penser que la variabilité est beaucoup plus faible à l'intérieur de petites parcelles et qu'on peut y suivre de manière assez précise, l'effet de certains traitements (fumures, façons culturales, etc..) sur l'évolution de la perméabilité.

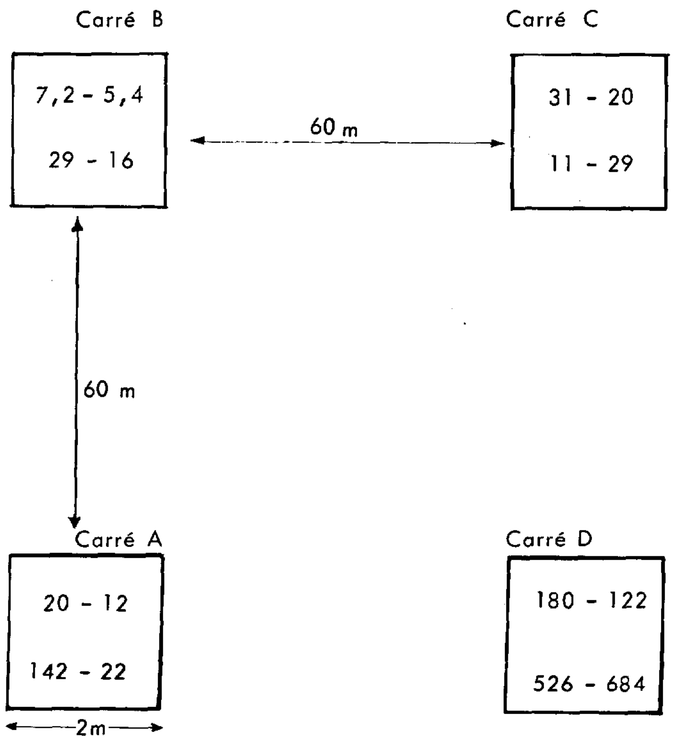
Voici trois séries de résultats pris à trois profondeurs dans une même tranchée pédologique, et qui montrent combien la variabilité est forte même sur de très petites surfaces.

K en mm /heure

Série I 10 - 20 cm	59 - 68 - 76 - 85 - 92 - 115 - 130 - 142 - 151
Série 2 50 - 60	0,18 - 0,55 - 0,94 - 1 - 1,2 - 1,4 - 1,6 - 1,8 - 2
Série 3 80 - 90	0,08 - 0,14 - 0,19 - 0,20 - 0,31 - 0,44 - 0,6 - 0,9 - 1,4

On a encore fait quatre prélèvements en **surface**, à l'intérieur de quatre carrés distants de 60 m environ (cf Densité apparente).

Les résultats sont reportés ci-dessous (en mm /h).



Carrés	A	B	C	D
Moyennes arithmétiques	49	26,5	22,75	378

Donc les moyennes sont assez stables dans trois carrés, mais le quatrième s'en écarte nettement. Il s'agit de prélèvements de surface et il est probable que ces variations systématiques, puisqu'on n'a observé aucune différence pédologique, sont dues à des façons culturales (En D par exemple, deux prélèvements sont très différents des deux autres). Ces variations peuvent ne pas exister aussi nettement en profondeur, mais, dans le doute, on ne peut travailler au mieux qu'en répartissant plusieurs prélèvements sur toute la zone à étudier et en se méfiant des valeurs anormalement élevées dès qu'on est certain qu'elles ne sont pas issues d'une zone pédologique différente.

Si on veut donc connaître la perméabilité avec plus de précision que celle donnée dans ces paragraphes, il est nécessaire de procéder différemment.

6 - Cas de nombreux prélèvements par horizon

On peut dire que le procédé décrit précédemment permettra seulement de classer les sols suivant leur perméabilité et, en recoupant par d'autres éléments pédologiques, de les classer selon leur aptitude à l'irrigation. Les surfaces irrigables retenues, il semble toujours préférable d'établir le projet sur des mesures moins "ponctuelles" que celles données par les méthodes Vergieres, Porchet, Muntz, etc..

- **E n i r r i g a t i o n** : suivant qu'on a choisi l'irrigation par sillon, calants, aspersion, faire une expérimentation directe en mesurant l'avancement de l'eau dans les sillons avec différents débits, ou essayer quelques asperseurs (camion-citerne et moto-pompe à différentes pressions, etc..

- **D r a i n a g e** : installer un système rudimentaire de drains et mesurer les débits, la variation de la nappe entre les drains.

Si on ne peut faire autrement qu'utiliser les méthodes "ponctuelles", il est nécessaire et économique de multiplier les prélèvements. En effet, on peut imaginer facilement les pertes économiques qu'entraîne une erreur du simple au double sur la mesure de la perméabilité en surface dans un périmètre : perte d'eau en colature, si les sillons sont trop courts, mauvaise répartition de l'eau sous les sillons, stagnation de l'eau en surface, etc..

En effet, l'horizon de surface (25 cm) peut stocker les doses d'irrigation courantes. Après que l'eau est entrée dans cet horizon, qu'elle mette 10 heures ou 50 heures pour se répartir sur l'ensemble du profil a généralement peu d'importance. C'est-à-dire que les prélèvements "ponctuels" peuvent définir ce caractère, parmi d'autres, d'aptitude à l'irrigation.

Au contraire, il est indispensable de savoir si l'eau mettra une heure ou trois heures pour pénétrer dans le sol et l'examen des résultats de la méthode Vergières montre son incapacité à définir ce temps d'une manière aussi précise, si on ne fait que peu de prélèvements.

Des études analogues à celle-ci (en Belgique et Hollande, notamment) et des observations qu'on a pu faire, montrent que la méthode du trou (Porchet ou Hoogdhoudt) ou la méthode de Muntz donnent des dispersions analogues.

Il faut se méfier en particulier de la précision illusoire de certaines pratiques, telles que celles qui consistent pour chaque mesure à tracer approximativement une droite passant par un certain nombre de points à calculer à peu près la pente de cette droite, et à en tirer le cœfficient K (on peut le faire dans la méthode Porchet).

Dès qu'on aura tracé la droite de la première mesure, on aura tendance à tracer des droites parallèles pour les mesures suivantes, issues de la même zone. De plus, le fait de n'utiliser qu'une échelle pour toute la gamme des perméabilités, entraîne que, pour les perméabilités faibles, les pentes sont très faibles, si bien qu'on en arrive à trouver et à mesurer graphiquement la même pente pour ces faibles valeurs.

Finalement, on peut formuler les étapes du problème de la manière suivante :

Toutes les méthodes ponctuelles donnent des résultats très dispersés, et les distributions ne sont pas, en général, gaussiennes. La forme de ces distributions est à priori inconnue et probablement variable pour différents sols.

- la seule méthode rigoureuse qui permette de calculer la moyenne consiste à faire un ajustement à une distribution et à le tester. Si on calcule la moyenne et l'écart-type sur l'échantillon, on ne peut utiliser les tables habituelles de Student (qui ne sont valables que si la distribution est normale) pour estimer l'intervalle de confiance de la moyenne. Au contraire, si on connait la forme de la distribution, on peut toujours connaître l'intervalle dans lequel sont comprises par exemple 90 % des valeurs.

- mais pour tester un ajustement, il faut au moins une cinquantaine de points de prélèvements par horizon homogène.

- de plus l'ajustement à certaines distributions et le test d'ajustement exigent un travail considérable.

- après cela on pourra très bien ne pas être satisfait de la précision de la moyenne.

- il existe d'autres méthodes moins rigoureuses mais qui exigent beaucoup moins de travail et de prélèvements. Dans cet ordre d'idées, on peut suggérer l'ajustement graphique à une distribution gausso-logarithmique. Sur les séries de perméabilités dont nous disposons cet ajustement s'est révélé **g r a p h i q u e m e n t** satisfaisant.

Donc, en dernier recours, quand on ne pourrait vraiment pas mettre en œuvre des méthodes directes de mesures des paramètres d'un projet d'irrigation, c'est cette méthode d'ajustement qu'il faudrait utiliser, bien qu'elle ne comporte pas, rappelons-le, le test de l'ajustement. Elle exige au minimum une dizaine de prélèvements par horizon.

IV - HUMIDITÉ DU SOL EN PLACE

1 - Plan de prélèvement et analyse de variance

Ils sont identiques à ceux utilisés pour l'analyse de l'humidité équivalente.

Les prélèvements ont été faits après des pluies répétées, en hiver, donc en période d'évapotranspiration faible. La dispersion qu'on va calculer risque donc d'être différente de celle qui existe dans des périmètres d'irrigation. Par contre, l'humidité moyenne du sol est élevée (20 %) et il est probable que la dispersion est proportionnelle au taux d'humidité et diminuerait pour des sols plus secs.

Mais il est intéressant de montrer combien des calculs effectués à partir de combinaisons de résultats aléatoires (tel que le calcul du stock d'eau dans la couche de sol exploitée par les racines) sont susceptibles d'erreurs importantes, même si les résultats de base apparaissent initialement peu dispersés.

Remarque : On ne peut pas faire à proprement parler deux mesures sur un prélèvement. Donc chaque prélèvement est émietté et soigneusement mélangé dans un bac puis réparti dans deux boîtes en aluminium qui passeront à l'étuve. A partir de ces différences entre deux boîtes contenant le même prélèvement, on calcule l'écart-type de la mesure de laboratoire.

2 - Résultats de l'analyse

Profondeur	Humidité moyenne	S	S_j	S_i
12	26,3	0,29	0,92	-
37	19	0,44	1,97	2,9
62	17,9	0,15	0,84	1,94
85	16,6	0,21	0,74	1,95

Pour la signification de S, S_j, S_i, voir au chapitre humidité équivalente.

- S : écart-type d'une mesure au laboratoire, est faible.
- S_j : écart-type entre prélèvements distants d'un mètre, est déjà important et significativement distinct du prélèvement. Il est en outre particulièrement élevé pour la tranche 25-50 qui reçoit l'eau de la couche supérieure labourée, où les perméabilités sont assez bonnes, mais très variables d'un point à l'autre.
- S_i : écart-type entre prélèvements distants de 60 m, est significativement supérieur à S_j, sauf dans la tranche 0-25.

On retrouve ici, comme pour l'humidité équivalente, des variabilités très locales et des variabilités plus importantes sur de grandes surfaces.

La tranche de surface ameublie, en contact direct avec la pluie, ne présente au contraire que des variabilités locales.

3 - Écart-type et nombre optimum de prélèvements dans le cas de grandes parcelles

On n'a pas retenu la tranche supérieure puisqu'on admet que l'homogénéité est exceptionnelle par rapport à ce qui se présentera le plus souvent en période plus sèche.

L'écart-type d'un résultat estimé à partir des résultats des trois tranches 25-50, 50-75 et 75-100, est égal à 5.

Le tableau suivant donne l'intervalle de confiance de l'humidité moyenne d'une couche de sol déterminée, à partir de 1, 2,..8 prélèvements répartis sur la parcelle.

Nombre de prélèvements	s_H	I (H) P = 0,05
1	5	$\pm$ 10
2	3,53	$\pm$ 7
3	2,89	$\pm$ 5,8
4	2,5	$\pm$ 5
5	2,23	$\pm$ 4,5
6	2,01	$\pm$ 4
7	1,89	$\pm$ 3,6
8	1,77	$\pm$ 3,5

On voit qu'il est inutile d'espérer connaître l'humidité d'une parcelle de plus de 300 à 400 m2 en faisant moins de quatre prélèvements. Encore arrivera-t-il une fois sur vingt que l'erreur relative sera de 5/20 environ, soit 25 %.

4 - Dispersion des résultats de réserve d'eau

Réserve en eau : c'est la quantité d'eau R (exprimée par une hauteur d'eau en mm) contenue dans le sol sur une profondeur donnée au moment de la mesure.

Bilan : c'est la différence B, entre la réserve initiale (au début d'une campagne) et la réserve finale, ou, plus généralement, entre deux réserves aux temps T1 et T2.

On sait que ces données interviennent dans de nombreux problèmes de recherche appliquée ou d'exploitation d'un périmètre.

On admettra que, pour connaître l'humidité d'une tranche de sol d'épaisseur E, l'opérateur prélève une carotte de sol d'épaisseur E ; sinon, il introduit en plus de la variabilité d'échantillonnage horizontal qu'on a calculé, une variabilité verticale. Or, s'il prélève une épaisseur de 10 cm et qu'il l'extrapole à 25 cm, il commet une erreur qui n'est pas purement aléatoire, mais qui dépend de la forme du profil hydrique dans la dite tranche. L'étude de cette erreur n'est donc pas du domaine statistique.

Il est donc préférable, et particulièrement dans les tranches superficielles du sol où le profil hydrique varie fortement, de prélever toute l'épaisseur à déterminer. La formule donnant R est identique à la formule (1) du chapitre 2, paragraphe 4, en remplaçant HE par H, humidité en place.

$$R = \sum_i \frac{Da_i}{10} \cdot H_i \cdot E_i$$

Des calculs identiques à ceux du paragraphe 4 donnent, pour un seul sondage (4) :

$$S_R = 0,25 \sqrt{\sum R^2_i}$$

Plus on fait de prélèvements dans un même sondage, plus l'épaisseur des tranches est faible et plus les valeurs élémentaires R_i sont petites. La formule ci-dessus montre que S_R diminue en conséquence.

Malheureusement, la formule 4 permet seulement de confirmer ce fait. En effet, l'équation dont on est parti pour l'établir :

$$S_y^2 = a_i^2 \cdot S^2_{x_i} \qquad \text{(voir chapitre 2, paragraphe 4)} \quad (3)$$

n'est valable que si les variables aléatoires x_i sont indépendantes.

Dans le cas présent, x_i = humidité d'une tranche de sol. On peut supposer qu'entre les humidités d'un même sondage, il existe une corrélation, certainement positive (certains sondages ont toutes leurs humidités inférieures ou supérieures aux moyennes respectives de chaque tranche). Il en résulte une sous-estimation de S_R si on utilise l'équation 4.

On a d'ailleurs pu vérifier que les humidités d'un sondage n'étaient pas indépendantes pour nos résultats.

Si on part des formules 3 et 4 pour calculer S_R, on trouve S_R = 33.

Si on calcule les réserves réelles dans chaque sondage, puis l'écart-type de l'échantillon de réserves, on trouve S_R = 55. Soit un intervalle de confiance de ± 100 mm pour R $\#$ 300 mm, si on fait un seul sondage de quatre prélèvements.

Puisqu'il faut augmenter les prélèvements, il résulte de ceci que :

- en général, on n'a pas intérêt à augmenter le nombre de prélèvements par sondage au détriment du nombre de sondages. Puisque les humidités ne sont pas indépendantes, on risquerait avec un nombre faible de sondages de trouver des réserves systématiquement différentes de la moyenne de la parcelle ;
- puisque la variabilité est plus faible entre sondages très proches (c'est-à-dire que ces sondages ne sont pas indépendants non plus) on a intérêt à répartir les prélèvements sur toute la parcelle. Si toutefois on ne s'intéresse qu'à certains phénomènes théoriques on peut, pour bénéficier de cet effet, travailler sur de petites parcelles ;
- il est difficile, en raison du nombre de phénomènes, de fixer un nombre optimum de prélèvements. Signalons encore que la précision dépend beaucoup de l'épaisseur de terre prélevée et du volume de l'échantillon. Les résultats présentés ci-dessus venaient de déterminations sur les échantillons de 100 g, d'épaisseur 10 cm pour une tranche de 25 cm. On travaille actuellement avec des bocaux de 500 cc qui permettent de prélever 20 cm dans une tranche de 25 cm, à la tarière de 6 cm de diamètre. Pour des humidités de 20 %, l'écart-type d'un résultat s'est abaissé à 2,5, soit une précision double de celle trouvée avec les échantillons de 10 cm.

L'écart-type d'une réserve de 300 mm (sur un seul sondage avec un prélèvement tous les 25 cm) descend à 30 mm. On fait au minimum neuf sondages par champ d'essai (de 1 à 2 ha), soit un intervalle de confiance I (R) = $\frac{2 S_R}{\sqrt{9}}$ = 20 mm, et une précision de 7 % environ.

5 - Dispersion des résultats de bilan

$B = R_1 - R_2$

R_1 = réserve au temps T_1

R_2 = réserve au temps T_2

D'où $S_B = \sqrt{2}\, S_R$ d'après l'équation (3)

Donc, pour les bilans dont les valeurs sont inférieures à celles des stocks, l'écart-type est plus élevé, alors qu'ils sont les éléments les plus intéressants. Et le nombre de neuf sondages apparaît vraiment comme un minimum.

Ex. : $S_B = 1,4 \cdot 20 = 28$

I (B) $\neq$ 60 au seuil de probabilité 0,05.

CONCLUSION

Les exemples énumérés ont l'avantage de montrer certains des risques encourus quand on procède à des mesures ou lorsqu'on en utilise les résultats :

- la précision change beaucoup suivant les méthodes. Il faut à chaque fois s'en faire une opinion et employer les résultats avec les précautions correspondantes ;
- pour améliorer la précision il faut analyser la variabilité de chaque cause, de chaque opération à l'intérieur d'une méthode et porter l'effort (répétitions, travail plus soigné) sur celles qui introduisent le plus de variabilité. On notera à ce propos qu'on s'efforce souvent d'améliorer les techniques de laboratoire sans répéter les mesures et surtout les prélèvements dans une zone homogène alors qu'en général la variabilité du sol est supérieure à celle du laboratoire (nous l'avons constaté également sur des analyses chimiques ; en particulier conductivité chlore, etc..) ;
- les chiffres qui intéressent le praticien sont souvent des fonctions des données fournies par le laboratoire. Ces données peuvent être relativement précises mais, quand elles se combinent entre elles, conduisent à des résultats douteux par cumul des erreurs (bilan d'eau) ;
- d'autre part, il est inutile d'obtenir une grande précision ou même d'utiliser certaines méthodes, si on introduit les résultats dans des combinaisons à cœfficients empiriques douteux (dose d'irrigation).

En conséquence, et puisque les laboratoires sont souvent surchargés, il serait peut-être préférable, en prospection, de ne pas étaler des demandes d'analyse sur tous les profils ou toutes les caractéristiques pédologiques rencontrés. En général, les appréciations sur le terrain (texture, structure, porosité, présence et nature de la matière organique, du calcaire) feront aussi bien qu'une analyse unique. On pourrait au contraire porter plus d'attention à certaines zones ou problèmes d'intérêt théorique ou pratique, et y faire une campagne de prélèvements et d'analyses planifiée, de sorte qu'elle permette d'estimer la précision des mesures et de répondre avec suffisamment de garanties à ces problèmes limités, mais qui prendront alors valeur d'exemple, de référence. Le pédologue pourra juger de manière valable d'après ces résultats quantitatifs la qualité de ses appréciations de terrain et, inversement, en se donnant un thème de recherche dans chaque prospection, garder l'habitude de se plier aux chiffres.

ANNEXE

Exemples d'analyses statistiques

1 - DISPERSION DE MESURES D'HUMIDITÉ ÉQUIVALENTE

Exemple : *résultats entre 25 et 50 cm*

Point	Profil	Résultats au laboratoire		Total par profil X_{ij}	Total par point X_i
		I	II		
A	A_1	27,55	27,48	55,03	110,29
	A_2	27,66	27,60	55,26	
B	B_1	26,52	26,57	53,09	105,23
	B_2	26,11	26,03	52,14	
C	C_1	28,89	28,69	57,58	112,64
	C_2	27,53	27,53	55,06	
D	D_1	25,07	25,26	50,33	101,61
	D_2	25,55	25,73	51,28	
Totaux		T^1 = 214,88,	T^2 = 214,89	429,77	429,77
		429,77			

X_{ijs} = un résultat élémentaire

$$C = \frac{T^2}{M} = \frac{184.702,25}{16} = 11.543,89$$

$$QT = \sum X^2_{ijs} - C = 11.564,51 = 11.543,89 = 20,62$$

$$Qi = \frac{\sum X^2_i}{K_j K_s} - C = \frac{46.249,60}{4} - 11.543,89 = 18,51$$

$$Qj = \frac{\sum X^2_{ij}}{K_s} - \frac{\sum X^2_i}{K_j K_s} = \frac{23.128,90}{2} - \frac{46.249,60}{4} = 2,05$$

$$Qs = \sum X^2_{ijs} - \frac{\sum X^2_{ij}}{K_s} = 11.564,51 - 11.564,45 = 0,06$$

Tableau d'analyse de variances

(1) Origine de la variance	(2) Somme **des carrés**	**(3) Degrés de liberté**	(4) Carré moyen (variat. moyenne)	(5) Estimation
Entre points	Qi = 18,51	3	6,17	$\sigma^2 + 2\sigma_j^2 + 4\sigma_i^2$
Entre profils	Qj = 2,05	4	0,51	$\sigma^2 + 2\sigma_j^2$
Entre mesures (résiduel)	QS = 0,06	8	0,0075	σ^2
Total	QT = 20,62			

$F_b = \frac{0,51}{0,0075}$ hautement significatif (F.3 ; 4 (0,01)= 16)

$F_a = \frac{6,17}{0,51}$ hautement significatif (F.4 / 8 (0,01)= 7)

$S^2 = \hat{E}\ (\sigma^2) = 0,0075$; $S \# 0,09$

$S^2_j = \hat{E}\ (\sigma_j^2) = \frac{0,51 - 0,0075}{2} = 0,435$; $S_j = 0,66$

$S^2_i = \hat{E}\ (\sigma_i^2) = \frac{6,17 - 0,51}{4} = 1,41$; $S_i = 1,2$

$\hat{E}\ (\sigma^2)$ = estimation de σ^2.

On peut trouver une explication détaillée de cette méthode d'analyse de variance dans : Bennett et Franklin "Statistical analysis in chemistry". Les anglo-saxons la nomment "nested classification" ou "classification hiérarchique" par opposition aux "classifications croisées" d'usage plus courant en expérimentation agricole.

Dans une classification croisée (méthode des blocs par exemple) on dispose pour chaque bloc et chaque traitement d'un résultat (somme de répétitions). Dans une classification "hiérarchique", on n'atteint au résultat d'un "point", dans cet exemple, qu'à travers ceux du profil et eux-mêmes à travers ceux du résultat de laboratoire. On remonte en quelque sorte les ramifications d'une chaîne. Cette méthode s'applique spécialement dès qu'on veut diviser une grande population en sous-populations dans lesquelles on prend un échantillon. Dans cet échantillon, on peut d'ailleurs tirer un échantillon plus petit, etc..

Il est particulièrement important, dans cette méthode, de reconnaître ce qu'estiment exactement les carrés moyens de la colonne 4 du tableau d'analyse. D'où la dernière colonne, où est portée en face de chaque carré moyen, la combinaison linéaire de variances théoriques estimées.

A défaut de cela, on risquerait :

- de ne pas faire de test F correct.

Exemple : pour savoir si σ_i^2, variance entre points, est significativement différente de zéro, il faut faire :

$$F = \frac{6,17}{0,51} = \frac{\hat{E}\ (\sigma^2 + 2\sigma_j^2 + 4\sigma_i^2)}{\hat{E}\ (\sigma^2 + 2\sigma_j^2)} = \frac{A}{B} \quad (1)$$

Si $\sigma_i^2 = 0$, $\frac{A}{B}$ devrait être inférieur ou égal à la variable F de Fisher au seuil de probabilité 0,05, avec trois et quatre degrés de liberté, ce qui n'est pas le cas, d'où $\sigma_i^2 \neq 0$.

En bref, les tests ne se font pas automatiquement par rapport à la variance "résiduelle".

- de ne pas calculer exactement les estimations de σ_i^2, σ_j^2, etc.., dès qu'on sait qu'elles sont différentes de zéro.
  Exemple : pour σ_i^2 son estimation est ici égale d'après (1) à : $S^2_i = \frac{A - B}{4}$

Les autres analyses pour l'humidité équivalente, la densité apparente, l'humidité, les réserves, sont identiques.

2 - ANALYSE STATISTIQUE DE RÉSULTATS DE DENSITÉ APPARENTE

Il s'agit d'abord de prouver que la manipulation au laboratoire ne change pas le résultat. On a pris une série de huit cubes dans un carré de 1 x 1 m. Après avoir mesuré la densité apparente, on a enlevé la paraffine des cubes et on a recommencé la mesure. Chaque cube subit donc deux mesures.

On peut isoler par l'analyse statistique :

- un carré moyen dû aux différences entre cubes, donc à la variabilité du sol (Q_B) ;
- un carré moyen dû au fait que la seconde série de mesures pourrait être significativement différente de la première (Q_T) ;
- un carré moyen dû à l'erreur de laboratoire (Q_E).

Tableau d'analyse de variance (les résultats ont été multipliés par 100)

Origine de la variation	Somme des carrés	Degrés de liberté	Carré moyen	Estimation
Q_B	176	7	25,1	$\sigma^2 + 2\,\sigma^2_i$
Q_T	2	1	2	$\sigma^2 + 8\,\sigma_t^2$
Q_E	35	7	5	σ^2
Q total	213	15		

$F_T = \frac{2}{5}$ non significatif : il n'y a pas de différence systématique entre les deux séries de mesures d'où $\sigma_t^2 = 0$.

$F_B = \frac{25,1}{5}$ significatif : la variabilité du sol est supérieure à la variabilité de la mesure au laboratoire d'où $\sigma_i \neq 0$.

L'écart-type S de la mesure au laboratoire est correctement éstimé par :

$$S = \sqrt{\frac{2 + 35}{1 + 7}} \;/\; 100 \neq 0,01.$$

L'écart-type S_i, dû uniquement à la variabilité du sol dans un carré de 1 x 1 m, est estimé par :

$$\sqrt{\frac{25 - 5}{2}} \;/\; 100 \neq 0,032 \quad > 0,02.$$

Enfin, l'écart-type d'un résultat est estimé par :

$$\sqrt{\frac{213}{15}} \; / \; 100 \neq 0{,}04$$

D'autre part, l'analyse hiérarchique montre que les résultats ne sont pas plus variables entre carrés qu'à l'intérieur d'un carré.

On peut donc ajouter les deux sommes de carrés de la première et de la seconde analyse pour avoir une meilleure estimation de la variance et de l'écart-type d'un résultat, soit :

$$S = \sqrt{\frac{213 + 382}{15 + 15}} \; / \; 100 = 0{,}045 \text{ qu'on a arrondi à } 0{,}05.$$

Enfin, sur un périmètre irrigué, on a obtenu les échantillons suivants de trois unités et issus chacun de populations qu'on suppose différentes.

Tableau de résultats (multipliés par 100)

Echantillons	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
Répétitions	143	137	139	150	153	178	139	147
	126	141	124	149	154	179	155	147
	134	137	128	152	145	166	158	154

Analyse de la variance

Origine de la variation	Somme des carrés	DDl	Carré moyen
Entre échantillon	3910	7	559
Erreur	675	16	42
Total	4585	23	

$F = \frac{559}{42}$ significatif : certains échantillons sont bien issus de populations différentes (zone ou horizons pédologiques différents).

Ecart-type = $\sqrt{42} \; / \; 100 \neq 0{,}065$

On avait une première fois obtenu un écart-type de 0,045 et maintenant de 0,065.

Une nouvelle estimation de l'écart-type à partir de l'ensemble des résultats montre que le choix de 0,05 est justifié.