

ANALYSE ET COMPARAISON DE SERIES CHRONOLOGIQUES RELATIVES A LA MESURE INDIRECTE DE LA TENEUR EN EAU DU SOL

C. DEBOUCHE

Résumé

Ce travail présente une méthodologie de comparaison de séries chronologiques relatives à la mesure indirecte de la teneur en eau d'un sol par des cellules à résistance.

Le modèle proposé est une somme de deux composantes qui explique les variations du logarithme de la résistance. La première composante est une relation linéaire avec la teneur en eau du sol estimée au moyen d'une sonde à neutrons. La deuxième composante caractérise l'inertie de la cellule par une relation également linéaire avec la mesure précédente.

Les paramètres de ce modèle, ajustés aux différentes situations permettent de comparer ces dernières entre elles. Il en ressort une différence significative entre le comportement des deux types de fabrication de cellule qui étaient présentés à l'analyse.

Mots-clés

Modèles, teneur en eau du sol.

1. INTRODUCTION

Trois cellules à résistance d'un premier type de fabrication (type I) et trois cellules à résistance d'un autre type de fabrication (type II) sont placées dans un même sol à 10 cm de profondeur. Les aspects techniques de ce dispositif de mesures ont été présentés par ailleurs (Ben Harrath et al., 1981). Les résistances mesurées

C. Debouche - Dept. Hydraulique et Topographie, Faculté des Sciences Agronomiques de l'Etat, Passage des Déportés, 2, 5800 Gembloux, Belgique.

varient entre 10 et $5 \cdot 10^5$ ohms, les valeurs faibles correspondant à des teneurs en eau élevées et réciproquement.

La période de mesure s'étend du 26 avril au 26 octobre 1976. En raison de l'absence de relevé pendant les week-end, l'intervalle entre deux mesures est de deux, trois ou quelques fois quatre jours.

Le problème posé est la comparaison des mesures indirectes d'humidité effectuées par les deux types de cellules. Cette comparaison ne porte pas sur la moyenne des humidités mesurées. En effet, les cellules peuvent avoir des évolutions parallèles qui donnent toutes une image exacte de la teneur en eau à une constante près. La comparaison doit plutôt porter sur la qualité de la mesure de l'évolution de la teneur en eau.

2. MODELE : PRINCIPE ET NOTATIONS

Etant donné les valeurs des résistances, celles-ci sont transformées en logarithmes. Dans la suite de ce texte, le terme "mesure" signifie donc : logarithme de la résistance mesurée.

Dans la mesure effectuée par une cellule on considère l'existence de trois composantes additives :

$$y_{ijk} = u_{ij}(\theta_k) + v_{ij}(y_{ijk-1}, y_{ijk-2}, \dots) + \varepsilon_{ijk}$$

où y_{ijk} est la mesure effectuée à l'instant k , par la $j^{\text{ème}}$ cellule du type i (logarithme népérien d'une résistance),

i est l'indice du type de cellule (1 ou 2),

j est l'indice de la cellule (1, 2 ou 3),

k est l'indice du jour (de 1 à 73),

u_{ij} est la première composante qui est liée à la teneur en eau dans le sol θ_k par une fonction à définir dont les paramètres peuvent être différents d'une cellule à l'autre,

v_{ij} est la deuxième composante qui est fonction de la ou des mesures précédentes et qui caractérise l'inertie de la cellule,

ε_{ijk} est la dernière composante aléatoire, de distributions normales (α, σ_{ij}) indépendantes entre elles (d'un jour à l'autre et d'une cellule à l'autre).

La comparaison des deux types de cellule revient donc à identifier les composantes u_{ij} et v_{ij} , à en calculer les paramètres et à tester l'égalité de ces derniers d'une cellule à l'autre à l'intérieur de chaque type de fabrication et d'un type à l'autre.

3. RELATION ENTRE LE LOGARITHME DE LA RESISTANCE ET LA TENEUR EN EAU

3.1. Calcul

La composante u_{ij} est la partie de la mesure qui est liée à la teneur en eau du sol :

$$u_{ijk} = f(\theta_k, \alpha_{ij}, \beta_{ij}, \dots)$$

où α_{ij} et β_{ij} sont des paramètres.

La teneur en eau du sol est évidemment inconnue. Cependant, on dispose d'une autre mesure de cette teneur en eau effectuée dans les mêmes conditions par une sonde à neutrons (Ben Harrath et al., 1981). La composante u_{ij} va donc être identifiée en mettant en relation la mesure effectuée par chaque cellule avec la mesure faite par la sonde. La composante u_{ij} n'est donc pas exactement liée à la teneur en eau du sol θ mais à son estimation $\hat{\theta}$ au moyen d'une sonde à neutrons (exprimée en % volume). Une relation linéaire suffit à rendre compte du lien existant entre le logarithme de la résistance d'une cellule et la mesure par la sonde. Cette relation peut s'écrire :

$$y_{ijk} = \alpha_{ij} + \beta_{ij} \hat{\theta}_k + d_{ijk},$$

où d_{ijk} est un résidu de la régression.

Les valeurs caractéristiques de cette relation pour les différentes cellules sont présentées au tableau 1. Elles sont obtenues à partir de la considération de 61 mesures en raison de quelques données manquantes dans les observations de la sonde à neutrons.

Tableau 1.

Régression simple entre cellules et sondes à neutrons.

	Type I (i=1)			Type II (i=2)		
	j=1	j=2	j=3	j=1	j=2	j=3
$\hat{\alpha}_{ij}$	14,29	13,41	13,27	14,40	14,66	13,95
$\hat{\beta}_{ij}$	-0,031	-0,023	-0,021	-0,033	-0,046	-0,039
$t_{\hat{\beta}_{ij}}$	12,21	18,15	10,91	11,51	17,48	12,25
r^2	0,72	0,85	0,67	0,69	0,84	0,72
$\hat{\sigma}_{y.x}$	0,746	0,369	0,559	0,854	0,778	0,933
$\hat{\sigma}_y$	1,388	0,940	0,966	1,525	1,917	1,742

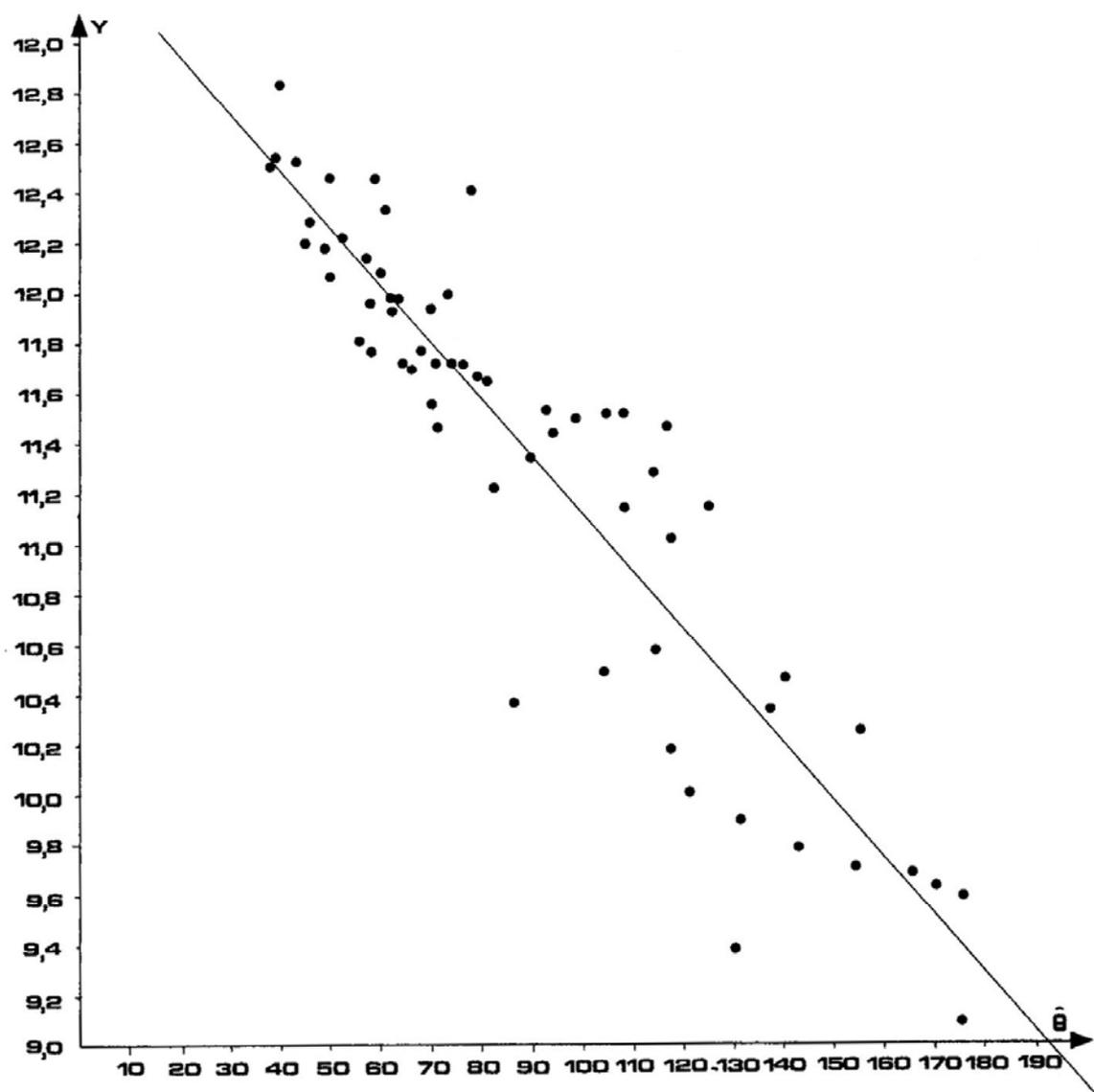


Fig. 1.

Teneurs en eau mesurées à la sonde à neutrons ($\hat{\theta}$ en %) et logarithmes des résistances électriques mesurées à la cellule n° 2 du premier type de fabrication (Y en logarithme de la résistance).

D'après le modèle proposé, les résidus de ces regressions peuvent s'écrire :

$$d_{ijk} = v_{ij} (y_{ijk-1}, y_{ijk-2}, \dots) + \varepsilon_{ijk}$$

Ils sont donc liés entre eux par la composante d'inertie v_{ij} et ne sont pas indépendants comme le supposent généralement les règles de l'inférence statistique sur la régression linéaire. L'examen de ces résidus (fig. 1 et 2) confirme cependant le caractère linéaire de la composante $u_{ij}(\theta_k)$.

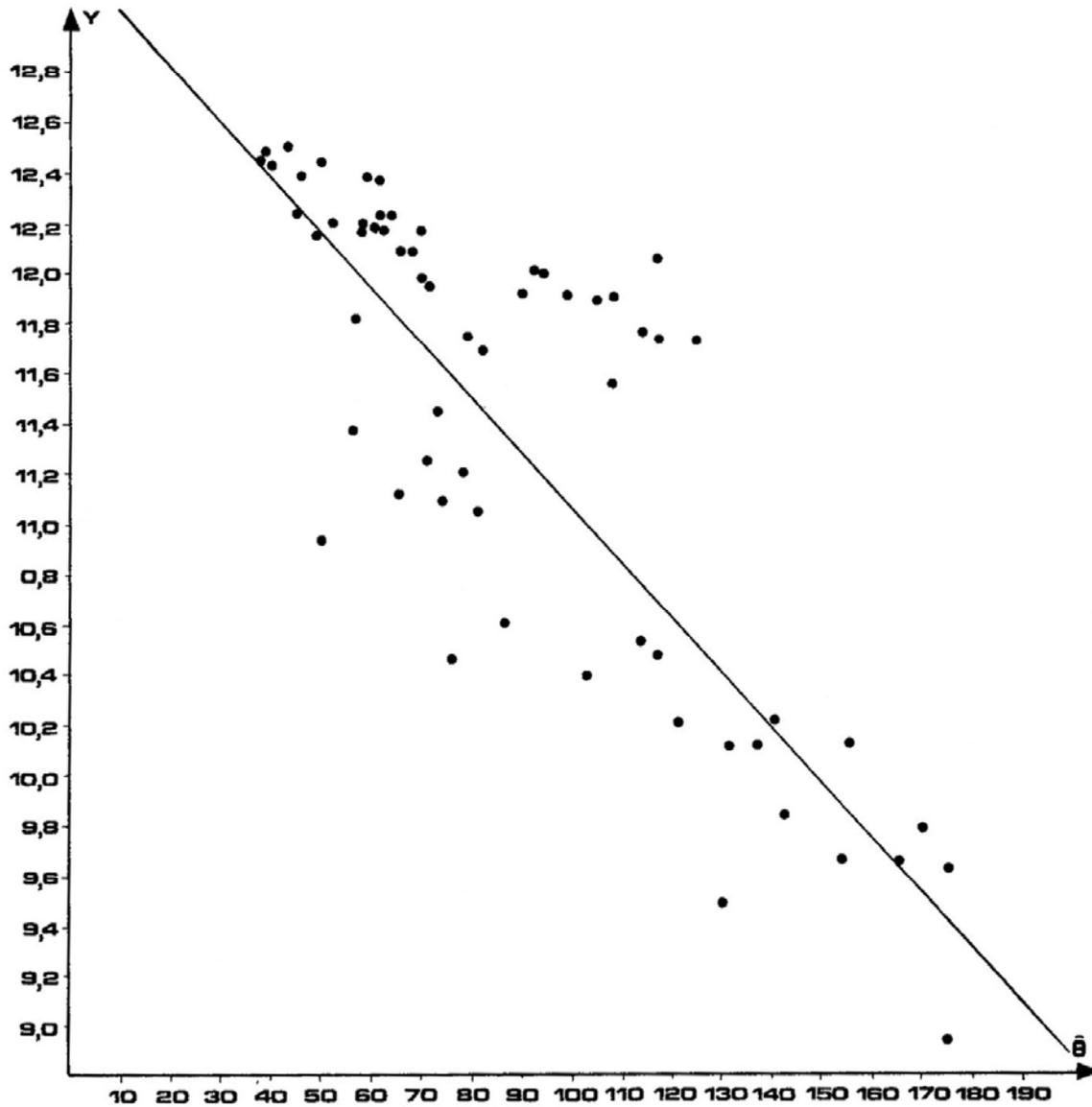


Fig. 2.

Teneurs en eau mesurées à la sonde à neutrons ($\hat{\theta}$ en %) et logarithmes des résistances électriques mesurées à la cellule n° 3 du premier type de fabrication (Y en logarithme de la résistance).

3.2. Critique des résultats

Une relation linéaire entre la teneur en eau estimée par la sonde et le logarithme de la résistance d'une cellule explique 70 à 80 % de la variance de ce dernier.

Les écarts-types résiduels sont plus élevés pour le type II que pour le type I. Cela signifie que la plus grande variabilité des mesures des cellules du second type n'est pas expliquée par une plus grande fidélité à la teneur en eau ou du moins à celle qui est estimée par la sonde à neutrons.

Comme on cherche moins à identifier des différences constantes entre cellules que des différences dans les relations cellule-teneur en eau, les ordonnées à l'origine ne doivent pas faire l'objet de test particulier. La comparaison essentielle doit porter sur les coefficients de régression β_{ij} .

Trois hypothèses sont à vérifier :

$$H_0 \quad \beta_{11} = \beta_{12} = \beta_{13},$$

$$H'_0 \quad \beta_{21} = \beta_{22} = \beta_{23},$$

$$H''_0 \quad \beta_{1.} = \beta_{2.}$$

où $\beta_{1.}$ et $\beta_{2.}$ sont les coefficients de régression moyens relatifs respectivement à l'ensemble des cellules du premier et du second type.

Le première hypothèse consiste à dire que les trois cellules du premier type ont la même réponse à une même variation de la teneur en eau. La seconde hypothèse est identique mais concerne les cellules du second type. La troisième hypothèse consiste à dire que, en moyenne, les cellules du premier type ont la même réponse à une variation de la teneur en eau que les cellules du deuxième type.

Ces trois hypothèses peuvent être vérifiées par un test de parallélisme (Rao, 1965). Les conditions d'application de ce test supposent, entre autres, l'indépendance des résidus successifs, ce qui n'est pas le cas. Il convient donc d'éliminer cette autocorrélation avant de procéder au test de parallélisme.

4. ETUDE DE L'INERTIE DES CELLULES

4.1. Autocorrélation

Nous avons étudié ci-dessus la relation existante entre la mesure effectuée par une cellule et la teneur en eau mesurée par la sonde. Cela permet de soustraire du logarithme de la résistance de la cellule, la part qui est liée à la teneur en eau. Le solde se compose d'une partie liée aux situations précédentes, c'est-à-dire de l'inertie de la cellule et d'une partie aléatoire appelée résiduelle.

Le tableau 2 présente les coefficients d'autocorrélation des résidus des régressions linéaires simples présentées au tableau 1. Ces valeurs sont calculées sur des effectifs variant de 47 à 56 en raison des données manquantes. La signification de ces coefficients d'autocorrélation peut être testée grossièrement en les comparant à $2/\sqrt{n-1}$, c'est-à-dire à 0,28. Les valeurs soulignées sont considérées comme significativement différentes de 0.

L'allure générale de l'évolution de ces coefficients d'autocorré-

Tableau 2.

Coefficients d'autocorrélation des résidus des régressions simples.

Ordre	Type I			Type II		
	j=1	j=2	j=3	j=1	j=2	j=3
1	<u>0,70</u>	<u>0,38</u>	<u>0,76</u>	<u>0,70</u>	<u>0,35</u>	<u>0,49</u>
2	<u>0,50</u>	<u>0,46</u>	<u>0,65</u>	<u>0,68</u>	<u>0,40</u>	<u>0,34</u>
3	<u>0,43</u>	0,06	<u>0,53</u>	<u>0,55</u>	0,19	<u>0,32</u>
4	<u>0,43</u>	0,11	<u>0,45</u>	<u>0,49</u>	0,05	0,07
5	<u>0,33</u>	-0,11	<u>0,38</u>	<u>0,41</u>	0,00	0,12

lation peut faire penser à une décroissance exponentielle, ce qui conduit à envisager un modèle autorégressif. L'ordre de ce dernier peut être indiqué par le calcul des coefficients d'autocorrélation partielle dont les valeurs sont reprises au tableau 3. Les valeurs soulignées correspondent à des coefficients d'autocorrélation considérés comme significativement différents de zéro. L'ordre maximum du coefficient d'autocorrélation partielle significatif indique l'ordre du processus autorégressif (Box et Jenkins, 1976). On peut donc en conclure qu'il existe une composante autorégressive dans les résidus des régressions simples présentées au tableau 1, dont l'ordre est de 1 ou 2.

Tableau 3.

Coefficients d'autocorrélation partielle des résidus des régressions simples.

Ordre	Type I			Type II		
	j=1	j=2	j=3	j=1	j=2	j=3
1	<u>0,70</u>	<u>0,38</u>	<u>0,76</u>	<u>0,70</u>	<u>0,35</u>	<u>0,49</u>
2	0,01	<u>0,38</u>	0,17	<u>0,36</u>	<u>0,32</u>	0,14
3	0,15	-0,26	-0,04	-0,01	-0,02	0,15
4	0,16	-0,03	0,04	0,00	-0,15	-0,22
5	-0,12	-0,06	0,00	-0,02	-0,05	0,13

Cette conclusion ne doit cependant être considérée qu'à titre indicatif, car elle n'est pas valable que si la série est stationnaire, ce qui n'est pas évident a priori. De plus, il ne faut pas perdre de vue que les mesures ne sont pas équidistantes, ce qui complique de toute manière l'interprétation de l'inertie des cellules.

4.2. Modèle autorégressif

L'analyse des coefficients d'autocorrélation simple et partielle permet de supposer l'existence d'un processus autorégressif d'ordre 1 ou 2 dans les résidus des régressions simples.

Avant d'estimer le modèle autorégressif, il convient de préciser si les caractéristiques de l'inertie d'une cellule sont indépendantes ou non de la teneur en eau mesurée par cette même cellule. Si l'hypothèse d'indépendance peut être retenue, l'estimation du modèle autorégressif peut se faire sur les résidus des régressions simples. Si l'hypothèse d'indépendance ne peut être retenue, l'estimation de la partie autorégressive du modèle général doit se faire en même temps que l'estimation de la partie liée à la teneur en eau. Dans ce cas également les comparaisons des deux types de cellules devront se faire simultanément sur tous les paramètres du modèle.

Même si cela conduit à plus de difficultés, il semble néanmoins réaliste d'abandonner l'hypothèse d'indépendance entre les deux parties du modèle.

Le tableau 4 présente les résultats des régressions doubles qui supposent que la mesure effectuée par une cellule au jour k dépend linéairement de la teneur en eau à ce jour k et de la mesure effectuée par la même cellule deux ou trois jours auparavant. Le modèle utilisé s'écrit :

$$y_{ijk} = \alpha_{ij} + \beta_{ij} \theta_k + \gamma_{ij} y_{ijk-1} + \varepsilon_{ijk}$$

La lecture de ce tableau indique que cette composante d'inertie intervient de manière significative dans l'explication de la mesure effectuée par la cellule au même titre si pas davantage que la teneur en eau du jour de la mesure.

L'écart-type résiduel reste plus élevé pour les mesures du type II que pour les mesures du type I. La variabilité supérieure du type II n'est donc pas imputable à une inertie d'ordre 1, c'est-à-dire d'environ trois jours.

L'étude des autocorrélations a indiqué l'éventuelle présence d'une composante autorégressive d'ordre 2. Le tableau 5 présente les résultats des régressions triples dans lesquelles figure le terme explicatif supplémentaire $\delta_{ij} y_{ijk-2}$.

Tableau 4.
Régressions doubles.

	Type I (i=1)			Type II (i=2)		
	j=1	j=2	j=3	j=1	j=2	j=3
$\hat{\alpha}_{ij}$	5,20	8,17	4,51	4,89	7,14	5,61
$\hat{\beta}_{ij}$	-0,013	-0,015	-0,008	-0,013	-0,024	-0,019
$t_{\hat{\beta}_{ij}}$	5,36	9,30	4,18	4,40	7,38	6,64
$\hat{\gamma}_{ij}$	0,65	0,40	0,67	0,67	0,53	0,62
$t_{\hat{\gamma}_{ij}}$	9,43	6,01	8,47	9,08	8,10	10,10
R^2	0,88	0,90	0,85	0,87	0,92	0,89
$\hat{\sigma}_{y.x}$	0,473	0,292	0,377	0,553	0,538	0,567

Tableau 5.
Régressions triples.

	Type I (i=1)			Type II (i=2)		
	j=1	j=2	j=3	j=1	j=2	j=3
$\hat{\alpha}_{ij}$	5,19	8,17	4,73	5,01	6,88	5,64
$\hat{\beta}_{ij}$	-0,013	-0,015	-0,008	-0,012	-0,023	-0,018
$t_{\hat{\beta}_{ij}}$	5,17	9,08	4,22	4,27	6,96	6,49
$\hat{\gamma}_{ij}$	0,69	0,40	0,81	0,92	0,74	0,75
$t_{\hat{\gamma}_{ij}}$	5,25	4,09	6,68	7,69	6,67	6,87
δ_{ij}	-0,04	-0,003	-0,16	-0,26	-0,20	-0,13
$t_{\hat{\delta}_{ij}}$	0,34	0,04	1,53	2,56	2,30	1,39
R^2	0,88	0,90	0,85	0,88	0,93	0,90
$\hat{\sigma}_{y.x}$	476	294	373	529	519	562

Il en ressort très clairement que ce terme supplémentaire n'apporte rien de significatif dans l'explication des mesures du type I et que son apport dans l'explication des mesures du type II est très réduit, même s'il est significatif. Il faut remarquer le signe négatif de ces coefficients autorégressifs d'ordre 2. La variation résiduelle pour le type II est très légèrement réduite par ce terme supplémentaire. Elle reste cependant supérieure à celle du type I. Il convient donc d'abandonner ces régressions triples.

Revenant aux régressions doubles, le tableau 6 présente les coefficients d'autocorrélation simple des résidus. Les valeurs soulignées correspondent à des coefficients considérés comme significativement différents de 0. Comparées aux coefficients d'autocorrélation des résidus des régressions simples (tableaux 2 et 3), ces valeurs permettent de conclure à la disparition d'autocorrélations importantes et systématiques dans les résidus de la régression double. Les autocorrélations partielles correspondantes sont reprises au tableau 7. Elles n'ajoutent pas grand chose à ces conclusions.

Tableau 6.

Coefficients d'autocorrélation des résidus des régressions doubles.

Ordre	Type I			Type II		
	j=1	j=2	j=3	j=1	j=2	j=3
1	0,22	0,18	<u>0,45</u>	<u>0,37</u>	0,21	0,05
2	0,16	<u>0,33</u>	0,17	0,23	0,27	0,15
3	0,12	-0,01	0,04	0,22	0,21	0,26
4	-0,01	0,21	0,14	0,21	0,13	0,05
5	<u>0,49</u>	0,04	0,19	<u>0,29</u>	0,24	<u>0,31</u>

Tableau 7.

Coefficients d'autocorrélation partielle des résidus des régressions doubles.

Ordre	Type I			Type II		
	j=1	j=2	j=3	j=1	j=2	j=3
1	0,22	0,18	<u>0,45</u>	<u>0,37</u>	0,21	0,05
2	0,11	<u>0,31</u>	-0,05	0,11	0,24	0,15
3	0,07	-0,12	-0,02	0,12	0,13	0,25
4	-0,06	0,15	0,17	0,09	0,02	0,02
5	<u>0,52</u>	0,03	0,08	0,19	0,16	0,26

4.3. Comparaisons

Ce qui a été dit au paragraphe précédent nous permet de nous en tenir aux régressions doubles qu'il nous reste à comparer. Ces comparaisons porteront sur les paramètres β et γ . Elles seront réalisées au moyen du test de parallélisme portant sur les trois hypothèses suivantes :

$$H_0 \quad \beta_{11} = \beta_{12} = \beta_{13} \quad \text{et} \quad \gamma_{11} = \gamma_{12} = \gamma_{13} ,$$

$$H'_0 \quad \beta_{21} = \beta_{22} = \beta_{23} \quad \text{et} \quad \gamma_{21} = \gamma_{22} = \gamma_{23} ,$$

$$H''_0 \quad \beta_{1.} = \beta_{2.} \quad \text{et} \quad \gamma_{1.} = \gamma_{2.} ,$$

où $\beta_{1.}$ et $\gamma_{1.}$ sont les coefficients de régression partielle moyens relatifs à l'ensemble des cellules du premier type et de même $\beta_{2.}$ et $\gamma_{2.}$ pour le deuxième type. Ces trois hypothèses portant simultanément sur les coefficients β et γ ont le même sens que les hypothèses énoncées au paragraphe 3.2.

Le tableau d'analyse de la variance (tableau 8) présente les résultats de ces tests de parallélisme.

Tableau 8.

Test de parallélisme des régressions doubles : analyse de la variance.

Sources de variation	d.l.	SCE	CM	F _{obs}
Régression linéaire moyenne	1	678.185.152	678.185.152	2.976
Non parallélisme H ₀	4	2.126.940	531.735	2,33
H' ₀	4	2.256.440	564.110	2,48 x
H'' ₀	2	5.015.463	2.507.732	11,01 xxx
Résidus	348	79.294.061	227.856	
TOTAL	360	766.878.056		

On peut en outre procéder aux estimations globales suivantes :

$$\hat{\beta}_{1.} = - 11,32 \quad \text{et} \quad \hat{\gamma}_{1.} = 0,62 ,$$

$$\hat{\beta}_{2.} = - 17,94 \quad \text{et} \quad \hat{\gamma}_{2.} = 0,62.$$

En supposant vérifiées les conditions d'application du test on peut en tirer les conclusions suivantes :

1° dans le cadre du modèle choisi, les trois cellules du type I sont caractérisées par des paramètres qui ne diffèrent pas significativement entre eux;

2° il existe peut-être une différence significative entre les paramètres caractérisant les trois cellules du type II;

3° il existe une différence très hautement significative entre les paramètres caractérisant les cellules du premier type et les paramètres caractérisant les cellules du deuxième type. Les valeurs globales calculées pour ces paramètres indiquent que la différence porte uniquement sur la liaison entre valeur mesurée et teneur en eau, la caractéristique d'inertie étant identique d'un type à l'autre.

L'inégalité des variances résiduelles des deux types de cellule justifie la prudence de la conclusion 2. Les hypothèses H_0 et H'_0 peuvent d'ailleurs être testées respectivement par rapport au carré moyen résiduel calculé pour le premier type et pour le second type. Cette procédure donne une valeur de F_{obs} égale à 3,53 pour H_0 et 1,85 pour H'_0 , ce qui inverse les conclusions 1 et 2.

Il faut également remarquer que les 348 degrés de liberté du carré moyen résiduel global, supposent l'indépendance des résidus des différentes cellules, ce qui est peu plausible. La matrice des coefficients de corrélation simple entre les résidus des six cellules est très hautement significative ($\chi^2_{obs} = 214^{xxx}$ à 15 degrés de liberté, Dagnelie, 1975, p. 150). Il convient donc de réduire le nombre de degrés de liberté, ce qui réduira les différents F_{obs} . Toutefois, étant donné la valeur élevée du F_{obs} relatif à H''_0 , la conclusion 3 reste valable.

5. CONCLUSIONS

Il existe une relation linéaire satisfaisante entre le logarithme de la résistance électrique mesurée sur une cellule et la teneur en eau mesurée par une sonde à neutrons. Les paramètres caractérisant cette relation varient d'un type de fabrication à l'autre et, peut-être même d'une cellule à l'autre pour un même type.

La mesure effectuée par ces cellules est aussi fortement liée à la mesure effectuée deux ou trois jours auparavant. Cette inertie peut également être représentée par une composante linéaire, qui s'ajoute à la composante liée à la teneur en eau. Les paramètres de cette composante d'inertie sont identiques pour les deux types de cellule.

La plus grande variabilité des mesures des cellules du type II n'est expliquée ni par une plus grande fidélité aux variations de la teneur en eau, ni par une inertie différente de celle des cellules du type I.

Il faut signaler que les mesures analysées dans ce travail ont été réalisées à faible profondeur et en période relativement peu humide. D'autres circonstances moins favorables nécessiteront peut-être quelques raffinements du modèle proposé.

BIBLIOGRAPHIE

Ben Harrath, A., Ledieu, J. & Cogels, O. (1981)
Comparaison de deux types de cellules à résistance pour la mesure de l'humidité du sol in situ. Communication personnelle.

Box, G.E.P. & Jenkins, G.M. (1976)
Time series analysis forecasting and control.
San Francisco, Holden Day, 575 p.

Dagnelie, P. (1975)
Analyse statistique à plusieurs variables.
Presses agronomiques, Gembloux, 362 p.

Rao, C.R. (1965)
Linear statistical inference and its applications.
John Wiley, New York, 522 p.

Analyse en vergelijking van chronologische series met betrekking tot de indirecte meting van het water in de bodem

Samenvatting

Een chronologische serie analyse werd gebruikt om indirecte metingen van de bodemvochtigheid met elektrische celweerstand te vergelijken.

Het model bestaat uit een som van twee componenten voor de beschrijving van de variatie van de weerstandslogaritme. De eerste component vormt een lineaire relatie met de bodemvochtigheid gemeten met een neutronsonde. De tweede component is een maat voor de cel-inertie.

De vergelijking van de parameters in het model laat toe twee types van cellen tegen elkaar af te wegen.

Analysis and comparison of chronological series for the indirect measurement of soil moisture contents

Summary

Time series analysis is used to compare indirect measurements of soil moisture by electric resistance of cell.

The model is a sum of two components for the description of the variation of a resistance logarithm. The first component is a linear relation with the soil moisture estimated by a neutron source. The second component is a measure of the cell's inertia.

The comparison of the parameters in the model permits to evaluate the two types of cells.
