

ANALÝZA PROSTOROVÉ ZÁVISLOSTI A VZÁJEMNÝCH VZTAHŮ PŮDNÍCH VLASTNOSTÍ

JIŘÍ ŠIMŮNEK

Článek obsahuje analýzu půdních vlastností získaných klasickým pedologickým průzkumem. Byly použity metody geostatistické analýzy a metody analýzy mnoho-rozměrného pole. Byly studovány proměnné popisující geometrii půdního profilu, texturální charakteristiky, kořenovou zónu, obsah organických látek atd. Pro každou proměnnou byly zjištěny její izotropní a směrové semivariogramy. Pro odhalení vzájemných vztahů mezi půdními vlastnostmi a jednotlivými půdními profily a pro zkoumání struktury celého datového souboru byly použity klastrová analýza, analýza hlavních komponentů a faktorová analýza.

Jiří Šimůnek: Analysis of Spatial Dependence and Soil Properties Interrelations. Vodohosp. Čas., 38, 1990, 5; 8 figs, 8 tabs, 20 refs.

The artikel contains the analysis of the soil survey data of Hupselse Beck. Geostatistical analysis of the spatial variability and the methods of multivariate analysis are used. The variables describing the geometry of the soil profile, the textural characteristics, the root zone, the organic matter content etc. are investigated. For each variable its statistical moments and semivariograms are calculated. Cluster analysis, principal component analysis and factor analysis are used to investigate interrelation between variables and observations and to reveal the structure in multivariate data.

KEY WORDS: Semivariogram, Multivariate Analysis, Variability of Soil Properties.

1. Úvod

Měření půdních hydraulických charakteristik je relativně časově a finančně náročné. Naskytá se otázka, zda není možné částečně překonat tento problém odhadem hydraulických charakteristik ze snázeji zjistitelných parametrů, tj. např. z informací zjišťovaných při klasickém půdním průzkumu (dále jen půdní vlastnosti) (Wösten a van Genuchten [20], Haverkamp a Parlange [7], Bloemen [2]). Různé techniky mohou být použity pro využití informací získaných studiem půdních vlastností, popř. vzájemných vztahů mezi půdními vlastnostmi a hydraulickými charakteristikami; např. „cokriging“ (Vauclin a kol. [16]) nebo různé varianty krigingu s dalšími pod-

Ing. Jiří Šimůnek, Výzkumný ústav zúrodňování zemědělských půd, Žabovřeská 250,
150 00 Praha 5-Zbraslav.

Tabulka 1. Seznam proměnných

Číslo	Jméno proměnné
1	Nadmořská výška
2	Hloubka jílové vrstvy
3	Roční nejvyšší hladina podzemní vody
4	Roční nejnižší hladina podzemní vody
5	Hloubka zakořenění
6	Pozorovaná hloubka kořenové zóny
7	Tloušťka horizontu A
8	Obsah organických látek v horizontu A
9	Obsah jílových látek v horizontu A
10	Mediánová hodnota písčité frakce v horizontu A
11	Tloušťka horizontu BC
12	Obsah organických látek v horizontu BC
13	Obsah jílových látek v horizontu BC
14	Mediánová hodnota písčité frakce v horizontu BC
15	Stupeň vrstvení horizontu D
16	Odpor proti vrtání horizontu D

půdními proměnnými (Ahmed a De Marsily [1]). Než je však možné přistoupit k predikci hydraulických charakteristik z půdních vlastností, je nutné studovat prostorovou závislost a variabilitu uvnitř a mezi oběma těmito třídami dat. Technikám použitelným k tomuto účelu je věnován mezinárodní projekt „Aplikace a vyhodnocení postupů sloužících k popisu prostorové variability půdně-fyzikálních a hydrologických veličin“, jehož koordinačním pracovištěm je Katedra hydrauliky a hydrologie Zemědělské university ve Wageningen.

Na experimentálním povodí Hupselse Beek ve východonizozemské provincii Gelderland o ploše 650 ha bylo shromážděno velké množství informací týkajících se jak půdních vlastností (1128 profilů) (Burrough a kol. [5], Wösten a kol. [18]), tak půdně fyzikálních charakteristik (Wösten a kol. [18], Brom [4], Booltink [3]). Detailní popis geologické struktury tohoto povodí je možné nalézt ve zprávě Burrough a kol. [5]. Na půdně fyzikální charakteristiky byly aplikovány různé statistické techniky Hopmansem [8], [9] a Hopmansem a Strickerem [10], [11]. Studium půdních vlastností získaných klasickým pedologickým průzkumem se zabýval Šimůnek [15]. Této problematice je věnován také tento článek.

Na základě předchozích studií (Wösten a kol. [19]) byl půdní profil charakterizován třemi základními horizonty — orníčním horizontem A, písčítým horizontem BC a písčito nebo hlinito-jílovitým horizontem D. Šestnáct proměnných (tab. 1) bylo vybráno pro následující analýzu.

2. Analýza prostorové závislosti

Proměnné jsou charakterizovány svojí pozicí v prostoru. Rozdíl mezi velikostí proměnné měřené na dvou různých stanovištích je závislý na vzdálenosti h mezi těmito body; proměnné měřené na méně vzdálených stanovištích jsou více korelované než proměnné z odlehlejších stanovišť. Tato závislost je popisována semivariogramem. Semivariogramy, jako kvantifikovaný souhrn všech dostupných strukturních informací, jsou používány k identifikaci struktury prostorového rozdělení.

Obyčejně jsou později použity pro další druhy analýz, např. kriging, mapování atd. Semivariogram proměnné Z je definován:

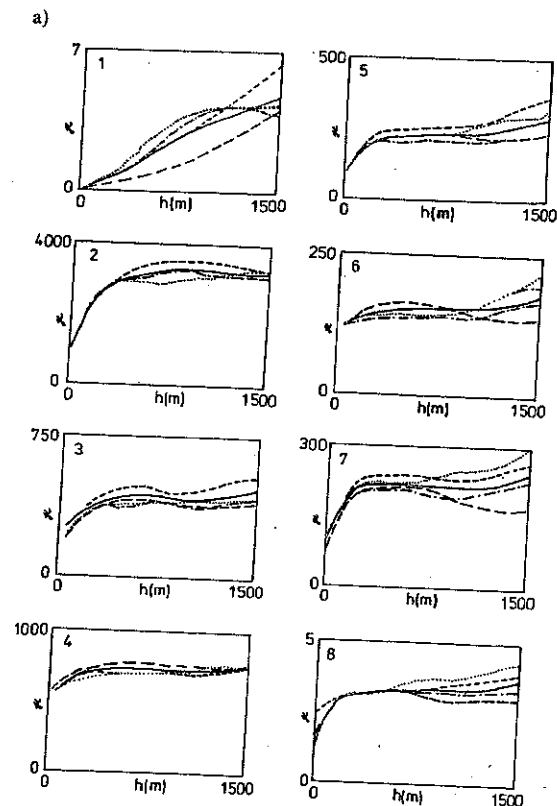
$$2\gamma(h) = \text{Var}[Z(x+h) - Z(x)], \quad (1)$$

kde h je vzájemná vzdálenost. Pokud očekávaná střední hodnota náhodné funkce $E(Z)$ je konstantní v prostoru, potom rovnice (1) může být přepsána:

$$2\gamma(h) = E\{[Z(x+h) - Z(x)]^2\}. \quad (2)$$

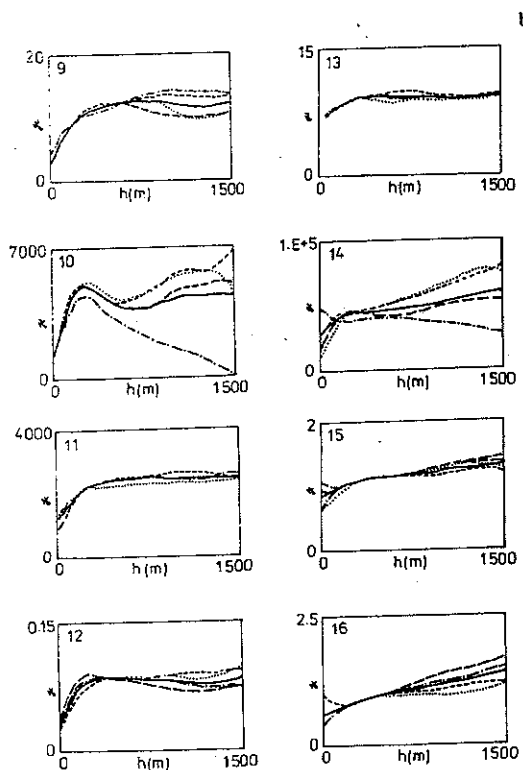
Plný popis teorie lze nalézt v Journal a Huijbregts [12].

Na počátku semivariogram obecně vzrůstá z malé hodnoty v anglické literatuře označované jako „nugget effect“. Od určité hodnoty vzdálenosti h (tato vzdálenost bývá označována jako rozsah — v anglické literatuře „range“) dosahuje konstantní hodnoty (práh — „sill“). V heterogenních podmínkách není důvod předpokládat, že rozsah a práh jsou nezávislé na směru, tj., že semivariogram je izotropní. Studium $\gamma(h)$ v různých směrech lze určit možnou anizotropii.



Pro všech šestnáct proměnných byly počítány čtyři směrové semivariogramy a jeden semivariogram nezávislý na směru. Směrové semivariogramy byly konstruovány takovým způsobem, že každá hodnota proměnné byla slučována s každou další hodnotou nacházející se zároveň uvnitř daného délkového i směrového intervalu. Směrový interval byl určen uvažovaným směrem a úhlovou tolerancí (22.5°) a délkový interval danou vzdáleností s délkovou tolerancí (25 m). Empirické semivariogramy byly vyhlazeny použitím Fourierovy transformace (Press et al. [13]). Výsledné semivariogramy jsou na obr. 1. Plná čára reprezentuje izotropní semivariogram. Ostatní čáry představují směrové semivariogramy: dlouze čárkovaná čára semivariogram pro úhel 0° (směr I), tečkovaná čára 90° (směr II), čerchovaná čára 45° (směr III) a čárkovaná čára 135° (směr IV). Úhly definující směr jsou vztaženy k ose západ—východ.

Praktické pokyny navržené Journalem a Huijbregtsem [12] udávají, aby experimentální semivariogram byl uvažován pouze pro dostatečně velký počet párů (30 až 50) a pro malé vzdálenosti ve vztahu k rozměrům experimentální plochy, pro kterou je semivariogram počítán. První pravidlo je snadné splnit, protože pouze několik bodů na počátku je počítáno z méně než tisíce



Obr. 1. Směrové a izotropní variogramy.

Tabulka 2. Parametry analytických semivariogramů

I	Model	C_0	C_1	a
1	Exponenciální	0.0404	0.000124	1.5286
2	Sférický	543.86	2351.74	239.00
3	Sférický	0.00	388.43	160.96
4	Sférický	0.00	668.05	106.65
5	Sférický	5.06	205.69	75.88
7	Sférický	16.70	191.67	102.56
8	Sférický	0.00	3.14	59.14
9	Sférický	1.30	9.47	257.42
10	Gaussův	—	4997.12	74.02
11	Sférický	706.08	1469.17	174.40
12	Sférický	0.00	0.08	80.00
13	Sférický	2.10	6.94	175.30
14	Sférický	18746.59	24112.68	136.76
15	Lineární	0.8472	0.0007	—
16	Lineární	0.6232	0.0008	—

datových párů. Druhé pravidlo znamená, že interval spolehlivosti je okolo 1500 m. Skutečnost, že některé body na počátku byly odhadovány z mnohem menšího počtu datových párů než zbytek semivariogramu, vysvětluje některé fluktuace v této části.

Většina proměnných vykazuje více či méně izotropní chování na počátku se vzrůstající anizotropií se zvětšující se vzdáleností. Protože byly zjištěny pouze malé rozdíly na počátku směrových semivariogramů, analytické modely byly proloženy pouze semivariogramy izotropními. Byly použity pouze proměnné vykazující prostorovou závislost v intervalu 0—500 m.

Použité modely:

— Gaussův model pro mediánovou hodnotu písčité frakce horizontu A.

$$\gamma(h) = C_1 [1 - \exp(-h^2/a^2)], \quad (3)$$

— lineární model pro proměnné popisující horizont D

$$\gamma(h) = C_0 + C_1 h, \quad (4)$$

— exponenciální model pro topografickou výšku

$$\gamma(h) = C_0 + C_1 h^a, \quad (5)$$

— sférický model pro ostatní proměnné

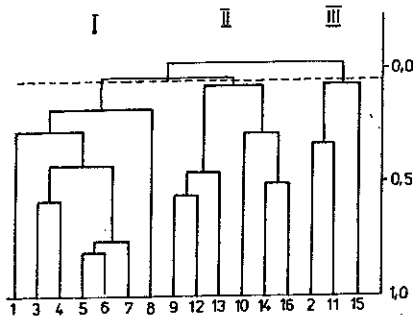
$$\gamma(h) = C_0 + C_1 \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right] \quad h \in [0, a], \quad (6)$$

$$\gamma(h) = C_0 + C_1, \quad h \geq a. \quad (7)$$

Sférický model dosahuje práh $(C_0 + C_1)$ ve vzdálenosti a . Gaussův model dosahuje práh asymptoticky a může být uvažován jako $a' = a\sqrt{3}$. Lineární a exponenciální modely práh nemají. Pro určení neznámých koeficientů byla použita obyčejná regrese, popř. Levenbergova—Marquardtova nelineární metoda nejmenších čtverců. Výsledky jsou uvedeny v tab. 2.

Semivariogram pro nadmořskou výšku vykazuje parabolické chování bez výrazného nuggetového efektu na počátku s nejvyšší heterogenitou směrových semivariogramů ze všech proměnných.

Tato významná heterogenita má jasné fyzikální vysvětlení. Povodí Hupselse Beck je údolí s nejnižší částí na západě a s nejvyšší na jihu. To znamená, že směr IV musí vykazovat významný trend. Stejný trend vykazuje i směr I, ale s menšími hodnotami rozptylu. Semivariogramy pro druhé směry vykazují Gaussovo chování s praktickým rozsahem okolo 1000 m. Semivariogram pro hloubku jílové vrstvy a tloušťku horizontu BC ukazují podobně nejpravidelnější chování bez výrazné anizotropie. Semivariogram pro roční nejvyšší hladinu podzemní vody a pro kořenovou zónu vykazují přechodový jev mezi počátkem a vzdáleností asi 1000 m. Za touto vzdáleností dochází k dalšímu nárůstu hodnot semivariogramu indikující přítomnost trendu (quasistacionarita). Pro obě proměnné popisující pohyb hladiny podzemní vody byl použit sférický model s předepsanou hodnotou



Obr. 2. Dendrogram korelační matice (tab. 3).

$C_0 = 0$. Proměnné 5, 7, 8 a 9 vykazují velmi podobné prostorové chování: téměř izotropní chování na počátku se vzrůstající anizotropií se vzrůstající vzdáleností. Nárůst mediánové hodnoty písčité frakce horizontu A není monotónní, ale vykazuje značné kolísání („hole effect“). Semivariogramy proměnných popisujících horizont BC s výjimkou mediánové hodnoty písčité frakce vykazují velmi pravidelné a izotropní chování. Proměnné popisující horizont D vykazují lineární trend v celém intervalu spolehlivosti.

3. Analýza mnohorozměrného statistického souboru

Po prohlídce měřených dat a výběru proměnných, které vykazují dostatečnou variabilitu, byla získána matice o N řádcích a M sloupcích. Každý řádek charakterizuje jeden profil (pozorování) a každý sloupec jednu půdní vlastnost nebo geometrickou veličnost (proměnná). Tato matice představuje mnohorozměrný datový soubor. Pokud si přejeme pracovat s celou touto maticí a klasifikovat půdu na základě všech proměnných, je možné použít metod analýzy více proměnných. Tyto metody umožňují manipulovat s několika proměnnými současně a uvažovat jejich změny, ukázat vztahy mezi vlastnostmi a odhalit seskupování pozorování nebo proměnných. Proměnné jsou sloučeny do různých množin a vztahy mezi množinami a uvnitř těchto množin jsou studovány. Klastrová analýza, analýza hlavních komponentů a faktorová analýza budou diskutovány v dalším textu a jejich použití bude demonstrováno na datovém souboru získaném na povodí Hupselse Beck.

Klastrová analýza

Cílem klastrové analýzy je studovat vztahy mezi pozorováními nebo mezi proměnnými a odhalit strukturu vícerozměrného datového souboru. Úkolem je uspořádat pozorování (proměnné) do

Tabulka 3. Korelační matice

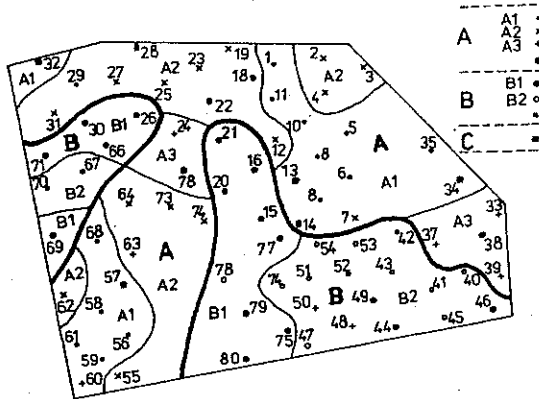
	1	2	3	4	5	6	7	8
1	1.0000	0.0840	0.3529	0.1780	0.3013	0.3350	0.2187	0.1479
2	0.0840	1.0000	0.2239	0.1664	0.1419	0.1457	0.1222	-0.0271
3	0.3529	0.2239	1.0000	0.5713	0.6061	0.6679	0.4565	0.1233
4	0.1780	0.1664	0.5713	1.0000	0.2896	0.3330	0.2673	-0.0005
5	0.3013	0.1419	0.6061	0.2896	1.0000	0.8076	0.7914	0.1947
6	0.3350	0.1457	0.6679	0.3330	0.8076	1.0000	0.7229	0.2093
7	0.2187	0.1222	0.4565	0.2673	0.7914	0.7229	1.0000	0.4159
8	0.1479	-0.0271	0.1233	-0.0005	0.1947	0.2093	0.4159	1.0000
9	-0.1438	0.0182	-0.0920	-0.0257	0.0018	-0.0360	-0.0201	0.2647
10	0.2607	-0.0218	0.1436	0.0785	-0.0708	-0.1770	0.2922	-0.0609
11	-0.1835	0.4314	0.2943	0.3761	-0.1567	-0.3131	-0.4068	-0.1491
12	0.0897	0.0113	-0.1758	-0.1210	0.0590	0.4434	0.1147	0.1828
13	-0.0715	-0.0281	-0.0210	0.1040	-0.0635	-0.0414	-0.0677	0.0595
14	-0.2985	-0.0484	0.3760	-0.2039	-0.2206	-0.5096	-0.3394	0.1100
15	0.0164	0.0994	0.1255	0.1179	-0.1416	-0.1560	-0.1251	-0.1040
16	-0.0356	-0.2176	-0.1050	0.0640	-0.0161	0.1181	0.0926	0.0240
	9	10	11	12	13	14	15	16
1	-0.1438	0.2607	-0.1835	0.0897	-0.0715	-0.2985	0.0164	-0.0356
2	0.0182	-0.0218	0.4314	0.0113	-0.0281	-0.0484	0.0994	-0.2176
3	-0.0920	0.1436	0.2943	-0.1758	-0.0210	0.3760	0.1255	-0.1050
4	-0.0257	0.0785	0.3761	-0.1210	0.1040	-0.2039	0.1179	0.0640
5	0.0018	-0.0708	-0.1567	0.0590	-0.0635	-0.2206	-0.1416	-0.0161
6	-0.0360	-0.1770	-0.3131	0.4434	-0.0414	-0.5096	-0.1560	0.1181
7	-0.0201	0.2922	-0.4068	0.1147	-0.0677	-0.3394	-0.1251	0.0926
8	0.2647	-0.0609	-0.1491	0.1828	0.0595	0.1100	-0.1040	0.0240
9	1.0000	0.1779	-0.0441	0.5526	0.3791	-0.1262	-0.0447	0.0711
10	0.1779	1.0000	0.0128	-0.1000	0.2093	0.4052	-0.2522	0.1552
11	-0.0441	0.0128	1.0000	-0.1845	0.0170	-0.0957	-0.0471	-0.1270
12	0.5526	-0.1000	-0.1845	1.0000	0.5313	-0.3371	0.2078	0.2921
13	0.3791	0.2093	0.0170	0.5313	1.0000	0.0616	-0.0104	0.0627
14	-0.1262	0.4052	-0.0957	-0.3371	0.0616	1.0000	0.3072	0.5044
15	-0.0447	-0.2522	-0.0471	0.2078	-0.0104	0.3072	1.0000	0.1464
16	0.0711	0.1552	-0.1270	0.2921	0.0627	0.5044	0.1464	1.0000

smysluplného pořadí tak, aby mohly být získány vztahy mezi jednotlivými pozorováními (proměnnými). Potom pozorování mohou být zařazena do několika více či méně homogenních skupin, s kterými může být jednotně zacházeno pro účely plánování nebo řízení.

Prvním krokem je získat nějaké měřítko podobnosti. V minulosti bylo použito mnoho různých měřítek (Se y h a n [14]), nejčastěji korelační koeficient nebo standardizovaná m -rozměrná Euklidova vzdálenost. Euklidova vzdálenost, resp. korelační koeficient, se zdá být vhodnější jako měřítko podobnosti mezi pozorováními, resp. proměnnými. Pro korelační koeficient je maximální podobnost reprezentována hodnotou 1, a maximální nepodobnost hodnotou -1 ; pro Euklidovu vzdálenost maximální podobnost má hodnotu 0. Čím větší vzdálenost, tím větší je nepodobnost. Koeficienty podobnosti jsou poté uspořádány do matice podobnosti. Pokud je počítána podobnost mezi pozorováními, resp. proměnnými, matice podobnosti má dimenzi $N \times N$, resp. $M \times M$. Dalším

krakem je výběr prvků z matice podobnosti s největší vzájemnou podobností a jejich umístění do skupin. Tyto skupiny objektů jsou poté spojovány s dalšími skupinami s nejvyšší podobností atd. Je možné použít několik klastrovacích postupů. V této analýze byla použita metoda „vážených párových skupin“ (Davis [6]). Klastrová analýza byla použita jak pro proměnné, tak pro pozorování.

Korelační matice (tab. 3) a dendrogram konstruovaný na základě této matice odhaluje vztahy mezi proměnnými a jejich hierarchickou strukturou. Je možné nalézt tři hlavní klastry. První hlavní klastr obsahuje nadmořskou výšku, obsah organických látek v horizontu A a dva subklastry; první popisuje pohyb hladiny podzemní vody (roční nejvyšší a nejnižší hladina podzemní vody) a druhý geometrii horizontu A (hloubka zakořenění, pozorovaná hloubka kořenové zóny a tloušťka horizontu A). Tento klastr obsahuje proměnné, které jsou nejdůležitější pro půdňí produktivitu. Druhý hlavní klastr obsahuje půdňí vlastnosti popisující půdňí texturu horizontů A a BC. Je opět tvořen dvěma subklastry. V prvním podklastru jsou obsahy jílových látek v obou horizontech a obsah organických látek v horizontu BC, druhý klastr obsahuje mediánovou hodnotu písčité frakce obou horizontů a odpor proti vrtání horizontu D. Třetí hlavní klastr je tvořen podklastrem obsahujícím hloubku k jílové vrstvě a tloušťku horizontu BC a proměnnou, která vykazuje nejvyšší nepodobnost ke všem ostatním proměnným — stupeň zvrstvení horizontu D.



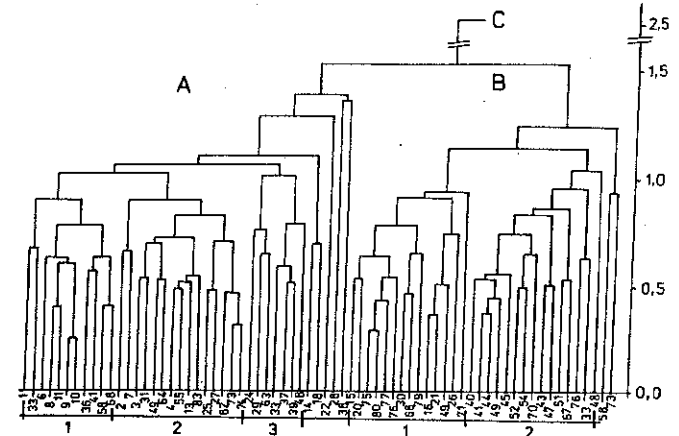
Obr. 3. Experimentální plocha rozdělená do dvou a pěti tříd.

Použití klastrové analýzy pro pozorování je demonstrováno pouze pro část povodí o ploše 67 ha (obr. 3). Na této ploše je 77 pozorování. Jako měřítko podobnosti byl použit koeficient Euklidovské vzdálenosti:

$$d_{ij} = \frac{\sum_k (X_{ik} - X_{jk})^2}{m} \quad (8)$$

kde m je počet proměnných, X_{ik} je k -tá proměnná v profilu i a X_{jk} je k -tá proměnná měřená v profilu j . Výsledný dendrogram může být rozdělen do tří hlavních klastřů. Na obr. 4 není vykreslen klastr C, který obsahuje osm profilů bez písčitého horizontu BC. Vzájemný vztah mezi zbývajícími dvěmi hlavními klastry A a B a mezi klastry nižší úrovně lze odvodit z tab. 4. Hlavní příčinou nepodobnosti mezi klastry je geometrie půdňího profilu, tj. hloubka k jílové vrstvě,

tloušťka horizontu BC, hloubka zakořenění a fluktuace hladiny podzemní vody. Některá pozorování není možné zahrnout do klastřů nižší úrovně. Toto je způsobeno extrémní v půdňích vlastnostech. V případě pozorování 14 a 18 to je vysoký obsah organických látek v horizontu BC, v případě profilů 22, 58 a 78 vysoká mediánová hodnota písčité frakce. Vysoký obsah jílových látek v horizontu A u profilu 28 a velká hloubka zakořenění v profilu 38 způsobují velkou vzdálenost od ostatních klastřů. Výsledky klastrové analýzy mohou být použity pro rozdělení plochy na různé parcely.



Obr. 4. Dendrogram půdňích profilů z plochy na obr. 3.

Tabulka 4. Aritmetické průměry některých proměnných v jednotlivých klastrech

Klas.	Hloubka jílové vrstvy	Roční nejvyšší HPV	Roční nejnižší HPV	Tloušťka horizontu BC	Hloubka zakořenění					
A1	44.5	10.0	120.1	25.5	39.1					
A2	76.1	99.0	20.4	24.7	123.2	121.3	45.3	61.0	45.7	47.3
A3	104.3	31.6	125.7			57.1				52.9
B1	167.3	144.3	42.1	40.0	153.2	144.3	118.1	105.7		46.4
B2	180.7	180.7	45.3	45.3	153.2	162.6		133.0		51.7

Předpokladem bylo, že klastr C s chybějícím horizontem BC může být považován za část klastru A1 s velmi malou tloušťkou horizontu BC. Pokud jsou uvažovány pouze dva hlavní klastry A a B, lze rozdělit celou plochu na menší části, přičemž nelezání hranic je velmi snadné. Pro další dělení již nelze uvažovat každý bod, pokud individuální části mají být rozumně velké a kompaktní s relativně hladkými hranicemi. Zanedbání některých bodů může být odůvodnitelné, neboť mezi klastry nižší úrovně není velká nepodobnost. Výsledná mapa je na obr. 3.

Analyza hlavních komponentů

Úkolem analýzy hlavních komponentů je interpretovat strukturu variační-kovariační matice mnohorozměrného datového souboru. M originálních proměnných je lineárně transformováno ve stejný počet nových proměnných — hlavních komponentů, kde každý nový hlavní komponent je lineární kombinací originálních proměnných. Hlavní komponenty jsou uspořádány v takovém pořadí, že každý nový komponent zaujímá co možná největší část z celkového rozptylu. Tudiž část celkového rozptylu reprezentovaná nejdlejší hlavní osou je značně větší než část reprezentovaná kteroukoli originální osou. Předpokládá se, že je možné studovat pouze několik hlavních komponentů, které představují velkou část celkového rozptylu a zanedbat ostatní bez velké ztráty informací v datovém souboru a tak snížit dimenzi originálních dat.

Prvním krokem je standardizovat proměnné. V opačném případě by orientace hlavních os byla kontrolována převážně proměnnými s největším rozptylem. Protože kovariační matice standardizovaných proměnných je matice korelační, je možné vzít matici podobnosti z klastrové analýzy (tab. 3) jako výchozí bod analýzy hlavních komponentů. Najít hlavní komponenty znamená najít vlastní čísla a vlastní vektory kovariační matice. Osm největších vlastních čísel, procento celkového rozptylu, které reprezentuje každé vlastní číslo a kumulativní procento celkového rozptylu jsou v tab. 5.

Tabulka 5. Vlastní čísla korelační matice

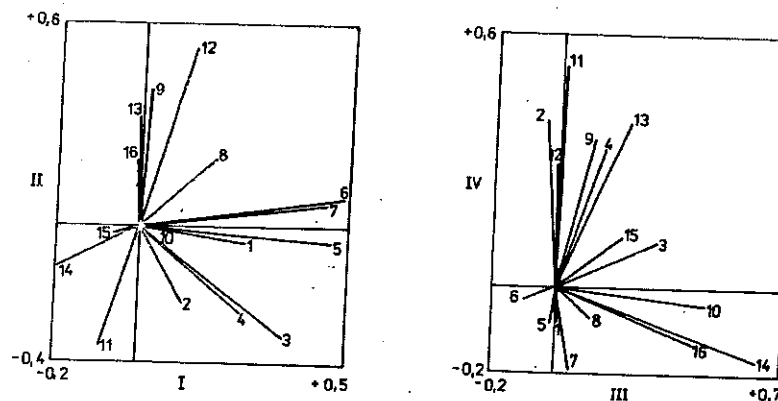
Pořadí	Vlastní číslo	Procento rozptylu	Kumulativní procento
1	3.7127	23.2041	23.2041
2	2.3470	14.6688	37.8729
3	1.9547	12.2167	50.0895
4	1.7904	11.1902	61.2798
5	1.4250	8.9061	70.1859
6	1.0827	6.7669	76.9528
7	0.9761	6.1007	83.0535
8	0.8159	5.0993	88.1528

Procento celkového rozptylu, které zaujmají jednotlivá vlastní čísla, je možné počítat, protože součet všech vlastních čísel se rovná součtu M originálních rozptylů, v případě standardizovaných proměnných přímo číslu M . První vlastní číslo je zdaleka největší, první tři vlastní čísla reprezentují více než polovinu rozptylu datového souboru a osm vlastních čísel reprezentuje téměř devadesát procent rozptylu.

Část matice hlavních komponentů, obsahující pouze první čtyři vlastní vektory, je v tab. 6. Vlastní vektory jsou ve sloupcích, relativní příspěvky každé proměnné do daného komponentu zvané zatížení jsou v řádcích. Tato matice obsahuje důležité informace pro interpretaci hlavních os. Absolutní hodnota relativního příspěvku blízko 1 znamená, že osa reprezentující originální proměnnou má směr blízko směru hlavní osy. Na druhé straně, pokud zatížení je blízko 0, obě osy jsou navzájem téměř kolmé a příspěvek této proměnné do hlavního komponentu je malý. Z tab. 6, nebo lépe z projekce vektorů do roviny (obr. 5), lze získat informace pro určení významu hlavních os. Grafy vektorů znázorňujících příspěvky proměnných do prvních, resp. druhých dvou hlavních komponentů jsou na obr. 5a), resp. 5b). Proměnné popisující geometrii horizontu A a fluktuaci hladiny podzemní vody přispívají výrazně ke komponentu 1. Tyto proměnné jsou shodné s proměnnými, které tvoří první hlavní klastr v klastrové analýze. Obsah organických látek horizontu BC společně s obsahem jílových látek v obou horizontech a tloušťkou horizontu BC v opačném směru přispívají významně ke komponentu 2. Opět lze najít podobnost s klastrovou analýzou. Hlavní

Tabulka 6. Matice hlavních komponentů: sloupce — vlastní vektory, řádky — proměnné

Proměnná	Vektory			
	1	2	3	4
1	0.2538	-0.0456	0.0050	-0.0806
2	0.1085	-0.2212	-0.0408	0.3836
3	0.3519	-0.3248	0.3099	0.1065
4	0.2444	-0.2507	0.1323	0.3174
5	0.4533	-0.0457	-0.0139	-0.0903
6	0.4896	0.0838	-0.1002	-0.0331
7	0.4460	0.0667	0.0521	-0.1957
8	0.1761	0.1972	0.1125	-0.0749
9	-0.0181	0.4052	0.1002	0.3400
10	0.0218	-0.0063	0.4382	-0.0384
11	-0.0876	-0.3533	0.0013	0.5212
12	0.1230	0.5301	-0.0096	-0.2885
13	-0.0031	0.3194	0.2138	0.3852
14	-0.1946	-0.1216	0.6243	-0.1735
15	-0.0604	-0.0218	0.1940	0.1152
16	-0.0049	0.2022	0.4289	-0.1316

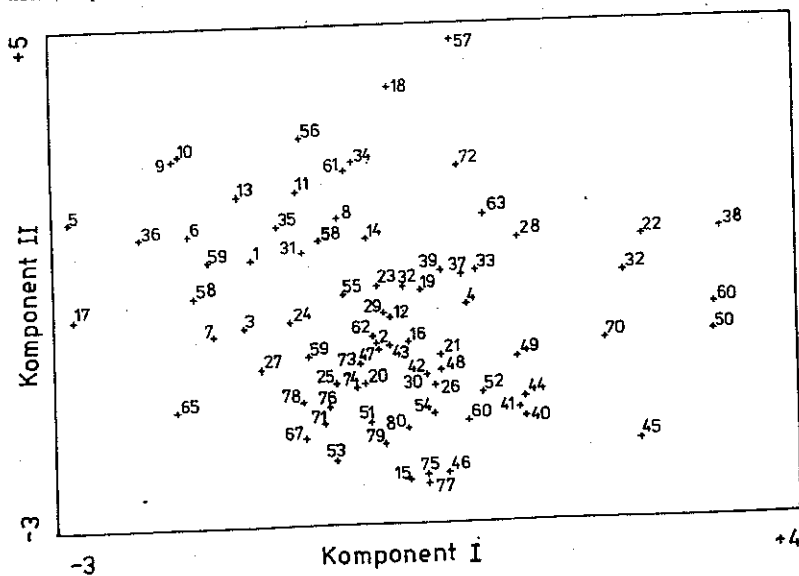


Obr. 5. Zatížení na první čtyři hlavní komponenty.

komponent 3 se skládá z mediánové hodnoty písčité frakce obou horizontů a odporu proti vrtání horizontu D. Hlavní komponent 4 obsahuje tloušťku horizontu BC a hloubku jílové vrstvy. Také v případě těchto dvou komponentů je možné odhalit těsnou shodu s výsledky klastrové analýzy. To není překvapující závěr, neboť výchozí bod pro obě analýzy byl shodný — korelační matice.

Vynásobením matice hlavních komponentů maticí obsahující originální data získáme skóre hlavních komponentů. Pro demonstrování výsledků byla vybrána opět pouze malá část povodí

(stejná jako v případě klastrové analýzy). Projekce skóre 77 profilů na první dva hlavní komponenty je na obr. 6. Protože tyto dva komponenty reprezentují téměř 40 % celkového rozptylu, tato projekce dává nejinformativnější pohled na vztahy v celém prostoru. Pro porozumění rozmístění jednotlivých bodů je nutné vzít v úvahu význam přisouzený jednotlivým osám. Profily se silným horizontem A a s hlubokou kořenovou zónou jsou umístěny doprava, zatímco profily s tenkým horizontem A a s kořenovou zónou jsou v levé části. Podél druhé hlavní osy jsou pozorování uspořádána podle obsahu jílových látek v obou horizontech, s nejvyšším, resp. nejnižším obsahem v horní, resp. dolní části.



Obr. 6. Měřené profily v rovině prvních dvou hlavních komponentů.

V kontrastu s výsledky klastrové analýzy, kde hlavní příčina nepodobnosti byla nalezena v hloubce jílové vrstvy, tloušťce horizontu BC a pohybu hladiny podzemní vody, se zdá být fakt, že první dvě vlastnosti přispívají významně pouze k čtvrtému hlavnímu komponentu. Avšak při detailnější prohlídce matice hlavních komponentů a vlastních čísel lze zjistit, že mezi velikostí druhého, třetího a čtvrtého vlastního čísla není velký rozdíl, a že tyto vlastnosti přispívají výrazně také ke druhému komponentu. Takže jejich význam je výrazný, přestože nemají rozhodující vliv na žádný z prvních tří komponentů. Další rotaci hlavních os lze získat jejich exaktnější interpretaci.

Faktorová analýza

Použitím analýzy hlavních komponentů jsme získali M hlavních komponentů, které reprezentují veškerý rozptyl v originálním datovém souboru. Protože prvních několik komponentů reprezentuje velkou část celkového rozptylu originálního datového souboru, není nutné pro vysvětlení struktury

Tabulka 7. Faktorová matice: sloupce — faktory, řádky — proměnné

Proměnná	Vektory			
	1	2	3	4
1	0.4891	-0.0698	0.0070	-0.1079
2	0.2090	-0.3388	-0.0571	0.5133
3	0.6780	-0.4976	0.4332	0.1425
4	0.4710	-0.3841	0.1850	0.4246
5	0.8735	-0.0700	-0.0194	-0.1208
6	0.9434	0.1284	-0.1401	-0.0443
7	0.8593	0.1021	0.0729	-0.2619
8	0.3393	0.3022	0.1573	-0.1002
9	0.0349	0.6207	0.1401	0.4550
10	0.0420	-0.0096	0.6126	-0.0514
11	-0.1687	-0.5413	0.0018	0.6975
12	0.2370	0.8121	-0.0134	0.3860
13	-0.0059	0.4893	0.2990	0.5155
14	-0.3749	-0.1863	0.8728	-0.2321
15	-0.1164	-0.0334	0.2713	0.1542
16	-0.0095	0.3098	0.5996	-0.1761

toto souboru použít všech M komponentů. Ale pozice těchto několika prvních komponentů v prostoru je významně ovlivněna existencí všech ostatních os. Pokud jsou tyto osy zanedbány, je možné rotovat vybrané komponenty a nelézt pro ně novou pozici, ve které jsou mnohem snáze interpretovatelné.

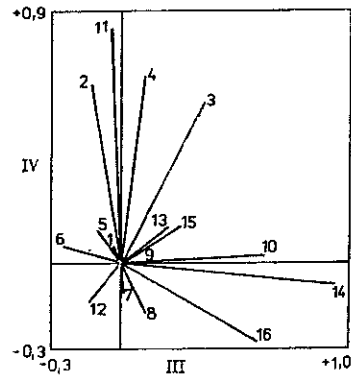
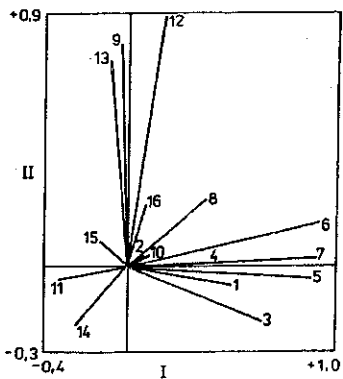
Protože výchozím bodem analýzy hlavních komponentů byla korelační matice (tab. 3) a vlastní čísla (tab. 5) byla počítána v normalizované formě, je možné transformovat hlavní komponenty ve faktory vynásobením každého prvku v normalizovaném vlastním vektoru odmocninou odpovídajícího vlastního čísla. Faktory jsou po této transformaci úměrné odmocnině části celkového rozptylu, který reprezentují a každé faktorové zatížení je úměrné odmocnině rozptylu, kterým přispívá daná proměnná k faktoru. Část faktorové matice obsahující pouze první čtyři faktory je v tab. 7.

Pro rotaci prvních čtyř faktorů byla použita technika zvaná „Kaiser's varimax“ (Davis [6]). Tato metoda se snaží, pokud je to možné, rotovat faktory takovým způsobem, aby každá originální proměnná byla blíže ukloněna k jednomu z nových faktorů a v pravém úhlu ke všem ostatním. Potom pro každý faktor existuje pouze několik významně vysokých příspěvků od originálních proměnných a mnoho nevýznamných příspěvků. Faktorové osy je nyní snadnější interpretovat pomocí originálních proměnných.

Výsledky po rotaci jsou ukázány v tab. 8, na obr. 7 a 8. Vzájemné vztahy mezi jednotlivými proměnnými jsou ve čtyřrozměrném prostoru přesně zachovány, ale jejich pozice vzhledem k faktorovým osám je změněna. Například roční nejvyšší a nejnižší hladina podzemní vody, které přispívaly ke všem prvním čtyřem hlavním komponentům, nyní přispívají významně pouze prvnímu a čtvrtému faktoru. První faktor je tvořen stejnými proměnnými, jako první hlavní komponent, tj. tloušťkou horizontu A, hloubkou zakořenění, pozorovanou hloubkou kořenové zóny, roční nejvyšší a nejnižší hladinou podzemní vody a obsahem organických látek v horizontu A. V interpretaci druhého hlavního komponentu byla velká nejistota, neboť kromě hlavního příspěvku od obsahu organických látek v horizontu BC existovalo dalších sedm proměnných s významnými příspěvky k tomuto komponentu. Tato nejistota byla odstraněna po rotaci faktorů. Nyní jsou zde pouze tři

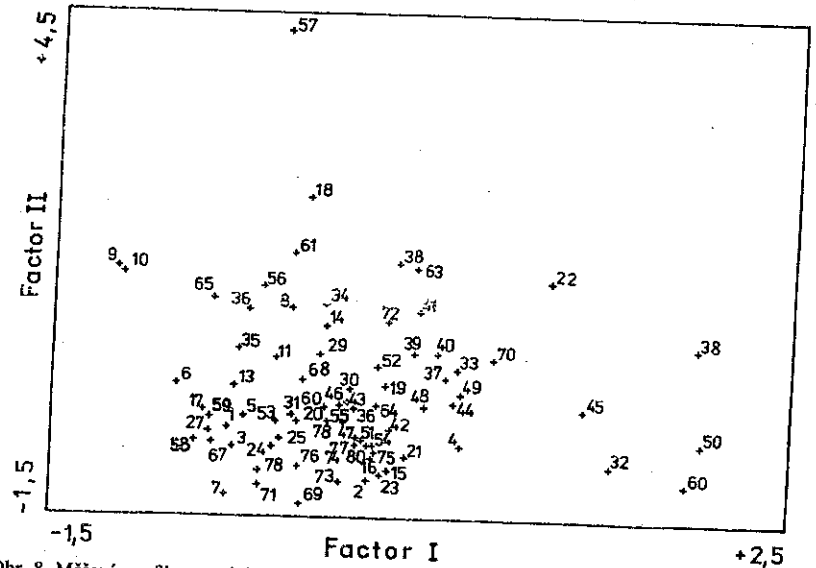
Tabulka 8. Matice rotovaných faktorů: sloupce — faktory, řádky — proměnné

Proměnná	Vektory			
	1	2	3	4
1	0.4975	-0.0688	-0.0262	0.0534
2	0.0717	0.0390	-0.1216	0.6354
3	0.6423	-0.1941	0.3677	0.5743
4	0.3626	0.0080	0.1036	0.6645
5	0.8716	-0.0416	-0.0899	0.1159
6	0.9194	0.1531	-0.2361	0.0603
7	0.9020	0.0249	0.0103	-0.0912
8	0.3732	0.2373	0.1161	-0.1791
9	-0.0392	0.7807	0.0453	0.0006
10	0.0977	0.0439	0.6064	0.0247
11	-0.3387	-0.0466	-0.0363	0.8305
12	-0.1672	0.8947	-0.1288	-0.1404
13	-0.0848	0.7265	0.2059	0.1310
14	-0.2523	-0.2105	0.9371	-0.0692
15	-0.1288	0.0850	0.2624	0.1392
16	0.0850	0.2207	0.5966	-0.2734



Obr. 7. Zatížení na první čtyři rotované faktory.

významné příspěvky k druhému faktoru — obsah jílových látek v obou horizontech společně s obsahem organických látek v horizontu BC. To znamená, že tento faktor popisuje texturu půdního profilu. Význam třetího faktoru je stejný jako význam třetího hlavního komponentu, tj. popisuje složení písčité frakce. Největší přínos přinesla rotace v případě čtvrtého faktoru, který je nyní definován roční nejvyšší a nejnižší hladinou podzemní vody, hloubkou jílové vrstvy a tloušťkou horizontu BC. To znamená, že tento faktor reprezentuje geometrické proměnné, které jsou ovlivněny hloubkou jílové vrstvy. Projekce faktorových skóre 77 profilů do roviny prvních dvou faktorů



Obr. 8. Měřené profily v rovině prvních dvou rotovaných faktorů.

je na obr. 8. Interpretace horizontální osy je stejná jako v případě analýzy hlavních komponentů, tj. půdní profily s hlubokým orničním horizontem a hlubokou kořenovou zónou jsou umístěny doprava, zatímco profily s tenkým horizontem A a tenkou kořenovou zónou se nacházejí v levé části. Podél druhého faktoru jsou pozorování tříděna podle textury obou horizontů, těžší texturní profily v horní a lehčí v dolní části. Je zajímavé si všimnout, jak jsou půdní profily rozděleny do jednotlivých kvadrantů. Více než jedna třetina pozorování se nachází napravo od středu s negativní hodnotou druhého faktoru, zbytek je rozmístěn téměř rovnoměrně ve zbývajících kvadrantech. Podle výše uvedeného interpretace faktorů to znamená, že v této části povodí je tendence pro texturně lehčí profily k hlubším orničním horizontům a hlubší kořenové zóně, zatímco texturně těžší profily nevykazují žádnou takovou závislost.

Závěr

Hlavním cílem této studie bylo zkoumat prostorovou variabilitu půdních vlastností a jejich vzájemné vztahy. Geostatistické metody a metody analýzy mnohorozměrného souboru (klastrová analýza, analýza hlavních komponentů a faktorová analýza) byly použity k charakterizaci variability šestnácti proměnných popisujících geometrii půdního profilu, texturní charakteristiky, obsah organických látek atd.

Pomocí prostorové analýzy bylo zjištěno, že kromě kvalitativních proměnných charakterizujících horizont D a pozorované velikosti kořenové zóny, všechny půdní vlastnosti vykazují prostorovou závislost nejméně na vzdálenost 60 metrů. Aby bylo možné tento závěr zvýraznit, bylo by nutné získat více informací pro menší vzdálenosti.

Metody vícerozměrné analýzy ukázaly, že v daném případě vybraných šestnáct půdních vlastností může být rozděleno do čtyř hlavních skupin. První skupina zahrnuje proměnné, které jsou v těsném vztahu s půdní produktivitou půdního profilu. Druhá skupina odráží texturu půdního profilu a třetí složení písčité frakce. Druhá a třetí skupina obsahují půdní vlastnosti, u kterých lze očekávat nejtěsnější vztah k hydraulickým charakteristikám. Ve čtvrté skupině se nacházejí proměnné, které jsou ovlivněny geologickým podložím.

Dalším směrem výzkumu by mělo být použití metod vícerozměrné analýzy na datový soubor obsahující nejen půdní vlastnosti získávané klasickým pedologickým průzkumem, ale též parametry popisující hydraulické charakteristiky.

LITERATURA

- [1] AHMED, S.—De MARSILY, G.: Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Wat. Resour. Res.*, 23, 1987, 9, s. 1717—1737.
- [2] BLOEMEN, G. W.: Calculation of hydraulic conductivities of soils from texture and organic matter content. *Z. Pflanzennahr. und Bodenkd.*, 143, 1980, 5, s. 581—615.
- [3] BOOLTINK, P. I.: Bodemfysische eigenschappen op zeer korte afstand Hupsel. Concept rapport. Vakgroep Hydraulica en afvoerhydrologie, Wageningen. Landbouwniversiteit 1985.
- [4] BROM, A.: Bodemfysisch-, hydrologisch veldonderzoek op een 7-tal lokaties in Hupsel in 1982. Rapport no. 2. Studie-groep Hupselse Beek. Vakgroep Hydraulica en afvoerhydrologie, Wageningen, Landbouwniversiteit 1983.
- [5] BURROUGH, P. A.—OERLEMANS, G.—STOFFELSEN, G.—WITTER, J. V.: Reconnaissance soil studies to determine optimum survey scales and mapping legend for soil moisture research in the Hupselse Beek hydrological catchment; Oost Gelderland, Report no. 1 of study group „Soil Variability Hupselse Beek“. Vakgroep Hydraulica en afvoerhydrologie, Wageningen. Landbouwniversiteit 1983.
- [6] DAVIS, J.: Statistical analysis in geology. New York, J. Wiley & Sons 1973.
- [7] HAVERKAMP, R.—PARLANGE, J. V.: Predicting the water-retention curve from particle-size distribution: 1. Sandy soils without organic matter. *Soil Sci.*, 142, 1986, 6, s. 325—339.
- [8] HOPMANS, J. W.: Application and evaluation of techniques which describe the spatial variability of soil physical and hydrological variables. Publication 75. Dept. of Hydraulics and Catchment hydrology, Wageningen, Agricultural University 1986.
- [9] HOPMANS, J. W.: A comparison of various methods to scale soil hydraulic properties. *J. Hydrol.*, 93, 1987, s. 241—256.
- [10] HOPMANS, J. W.—STRICKER, J. N. M.: Soil hydraulics properties in the study area Hupselse Beek as obtained from three different scales of observation: An overview. Report Dept. of Hydraulics and Catchment hydrology, Wageningen, Agricultural University 1987.
- [11] HOPMANS, J. W.—STRICKER, J. N. M.: Stochastic analysis of soil-water regime in a watershed. *J. Hydrol.*, 1988, (v tisku).
- [12] JOURNAL, A. G.—HUIBREGTS, Ch. J.: Mining geostatistics. London. Academic Press 1978.
- [13] PRESS, W. H.—FLANNERY, B. P.—TEUKOLSKY, S. A.—VETTERLING, W. T.: Numerical recipes — The art of scientific computing. Cambridge, University Press 1987.
- [14] SEYHAN, E.: Introduction to multivariate statistical analysis in hydrology. Free University, Amsterdam, Institute of earth science de Boelelaan 1981.
- [15] ŠIMŮNEK, J.: Analysis of soil survey data of Hupselse Beek. Report. Dept. of Hydraulics and Catchment hydrology, Wageningen, Agricultural University 1988.
- [16] VAUCLIN, M.—VIEIRA, S. R.—VACHAUD, G.—NIELSEN, D. R.: The use of cokriging with limited field soil observations. *Soil Sci. Soc. Am. J.*, 47, 1983, s. 175—184.
- [17] WEBSTER, R.: Quantitative and numerical methods in soil classification and survey. Oxford, Clarendon Press 1975.
- [18] WÖSTEN, J. H. M.—STOFFELSEN, G. H.—TEUNISSEN, J. W. M.—HOLST, A. F.—BOUMA, J.: Proefgebied Hupselse Beek; regionaal bodemkundig en bodemfysisch onderzoek. Rapport. no. 1706. Wageningen, Stichting voor Bodemkartering 1983.
- [19] WÖSTEN, J. H. M.—BOUMA, J.—STOFFELSEN, G. H.: Use of soil survey data for regional soil water simulation models. *Soil Sci. Soc. Am. J.*, 49, 1985, s. 1238—1244.
- [20] WÖSTEN, J. H. M.—van GENUCHTEN, M. Th.: Using texture and other soil properties to predict the unsaturated soil hydraulics functions. *Soil Sci. Soc. Am. J.*, 52, 1988.

Do redakcie došlo 24. 8. 1989
Diskusia otvorená do 30. 4. 1991.

АНАЛИЗ ПРОСТРАНСТВЕННОЙ ЗАВИСИМОСТИ И ВЗАИМНЫХ ОТНОШЕНИЙ ПОЧВЕННЫХ ХАРАКТЕРИСТИК

Иржи Шимunek

Статья решает вопрос анализа почвенных свойств в бассейне Гупселсе Бек. Были применены геостатистический анализ пространственного варьирования и методы многомерного анализа. Были исследованы переменные величины описывающие геометрию почвенного профиля, текстурные характеристики, коренную зону, содержания органических веществ и т.д. Вычислены статистические моменты и семи-вариограммы для каждой переменной величины. Для исследования отношения между величинами и для открытия структуры многомерного сочетания данных были применены кластерный анализ, анализ главных компонентов и факторный анализ.

- Рис. 1. Указательные и изотропические семи-вариограммы.
Рис. 2. Дендрограм корреляционной гайки из табл. 3.
Рис. 3. Наблюдательная площадка разделенная в два и пять классов.
Рис. 4. Дендрограм почвенных профилей из рис. 3.
Рис. 5. Нагрузка на первые четыре главные компоненты.
Рис. 6. Почвенные профили в плане первых двух главных компонентов.
Рис. 7. Нагрузка на первые четыре фактора после ротации.
Рис. 8. Почвенные профили в плане первых двух факторов после ротации.

Табл. 1. Список обозначений. Номер — переменная величина; 1 — топографическая высота, 2 — глубина глинистого слоя, 3 — годовой самый высокий уровень грунтовой воды, 4 — годовой самый низкий уровень грунтовой воды, 5 — глубина укоренения, 6 — размер коренного слоя, 7 — размер горизонта А, 8 — содержание органических веществ — горизонт А, 9 — содержание ила — горизонт А, 10 — средний размер песчаной фракции — горизонт А, 11 — размер горизонта В, 12 — содержание органических веществ — горизонт ВС, 13 — содержание ила — горизонт ВС, 14 — средний размер песчаной фракции — горизонт ВС, 15 — степень слоения — горизонт D, 16 — сопротивление бурению — горизонт D.

Табл. 2. Параметры аналитических семи-вариограмм.

Табл. 3. Корреляционная гайка.

Табл. 4. Средние величины некоторых переменных в отдельных кластерах.

Табл. 5. Собственные числа корреляционной гайки.

Табл. 6. Гайка главных компонентов.

Табл. 7. Гайка факторов.

Табл. 8. Гайка факторов после ротации.

ANALYSIS OF SPATIAL DEPENDENCE AND SOIL PROPERTIES INTERRELATIONS

Jiří Šimůnek

The article contains the analysis of the soil survey data of Hupselse Beck. Geostatistical analysis of the spatial variability and the methods of multivariate analysis are used. The variables describing the geometry of the soil profile, the textural characteristics, the root zone, the organic matter content etc. are investigated. For each variable its statistical moments and semivariograms are calculated. Cluster analysis, principal component analysis and factor analysis are used to investigate interrelation between variables and observations and to reveal the structure in multivariate data.

Fig. 1. Directional and isotropic semi-variograms.

Fig. 2. Dendrogram of correlation matrix in Tab. 3.

Fig. 3. Experimental area divided into two and five classes.

Fig. 4. Dendrogram for soil profiles from area on Fig. 3.

Fig. 5. Plot of loadings on the first four principal components.

Fig. 6. Scatter of 77 sampling sites plotted on the plane of the first two principal components.

Fig. 7. Plot of loadings on the first four rotated factors.

Fig. 8. Scatter of 77 sampling sites plotted on the plane of the first two rotated factors.

Table 1. List of variables. Number — name of variable; 1 — topographic height, 2 — depth to a clay layer, 3 — annual highest groundwater level, 4 — annual lowest ground-water level, 5 — rootable depth, 6 — root zone observed, 7 — thickness — A horizon, 8 — organic matter content — A horizon, 9 — clay content — A horizon, 10 — median sand size fraction — A horizon, 11 — thickness — BC horizon, 12 — organic matter content — BC horizon, 13 — clay content — BC horizon, 14 — median sand size fraction — BC horizon, 15 — degree of layering — D horizon, 16 — resistance to sampling — D horizon.

Table 2. Parameters of analytical semivariograms.

Table 3. Correlation matrix.

Table 4. Means of some variables in different clusters.

Table 5. Eigenvalues of the correlation matrix: (order, eigenvalue, percentage of known variance, cumulative percentage).

Table 6. The principal component matrix: columns — eigenvectors, rows — variables.

Table 7. The factor matrix: columns — factors, rows — variables.

Table 8. The rotated factor matrix: columns — factors, rows — variables.