

ESTIMAREA PRECIZIEI MODELULUI NIRS DE PREDICȚIE ÎN FUNCȚIE DE TIPUL DE SOL

CZU: 631.42

DOI: <https://doi.org/10.52673/18570461.22.2-65.08>

Doctorandă Marina ILUȘCA

E-mail: ilushca.marina@gmail.comORCID: <https://orcid.org/0000-0003-3866-1157>

Institutul de Cercetări pentru Culturile de Câmp „Selecția”

Universitatea Agrară de Stat din Moldova

ESTIMATING THE ACCURACY OF THE NIRS PREDICTION MODEL BASED ON SOIL TYPES

Summary. Near infrared reflectance spectroscopy (NIRS) proved to be a rapid, reliable and cost-effective method for soil analysis, but its prediction accuracy relies on the proper calibration model. In this paper, NIR spectroscopy in combination with partial least-squares regression (PLSR) was used to predict soil chemical properties on regional and local scales. The results obtained show a higher potential of NIRS to assess soil fertility on local scale, compared to a regional scale, with excellent prediction for organic C, carbonates and total N content ($R^2 \geq 0.98$ and $RPD > 6$), and good prediction for soil pH ($R^2=0.88$ and $RPD > 3$).

Keywords: soil analysis, soil organic carbon, near infrared reflectance spectroscopy, local calibration, soil type.

Rezumat. Spectroscopia de reflexie în regiunea infraroșu apropiat (NIRS) s-a dovedit a fi o metodă rapidă, fiabilă și rentabilă pentru analiza solului, însă precizia predicției se bazează pe modelul de calibrare corespunzător. În această lucrare, spectroscopia NIR în combinație cu regresia parțială minimă pătrată (PLSR), a fost utilizată pentru estimarea proprietăților agrochimice ale solului la scară regională și locală. Rezultatele obținute arată un potențial mai mare al NIRS de a evalua fertilitatea solului la nivel local, comparativ cu o scară regională, cu predicție excelentă pentru conținutul de C organic, carbonați și azot total ($R^2 \geq 0,98$ și $RPD > 6$), și predicție bună pentru pH-ul solului ($R^2 = 0,88$ și $RPD > 3$).

Cuvinte-cheie: analiza solului, carbon organic al solului, spectroscopie de reflexie în regiunea infraroșu apropiat, calibrare locală, tip de sol.

INTRODUCERE

Rezultatele cercetărilor au demonstrat că analiza spectroscopică de absorbție în domeniul infraroșu apropiat (NIRS) este o tehnică analitică rapidă și nedistructivă de analiză a proprietăților solului [1]. Analiza probelor de sol prin metoda NIRS poate reduce considerabil costurile analizelor de laborator convenționale, cu o precizie rezonabilă. Datorită caracteristicilor de absorbție a parametrilor esențiali ai solului, regiunea NIR conține informații utile privind conținutul compușilor organici și anorganici din sol [2].

Predicția proprietăților solului prin intermediul reflexiilor spectrale depinde de modelele statistice de calibrare, care explică relația dintre ele, iar succesul procedurii de calibrare, precum și al procedurii de validare se bazează pe selectarea unui set de probe corespunzător fiecărei etape.

În acest context, M. Chodak [3] a subliniat că setul de calibrare trebuie să reprezinte variația spectrală a întregii populații de probe pentru care se efectuează calibrarea. De asemenea, setul de calibrare urmează să

include toate variațiile posibile [4] și relevante în funcție de destinația de utilizare. Astfel, J. Wetterlind [2] a atenționat că dacă modelul va fi utilizat pe probele de sol din întreaga țară, setul de calibrare trebuie să includă probe cu toate tipurile de sol existente, iar în cazul utilizării modelului la nivel local, setul de calibrare va include variația locală, pentru a îmbunătăți rezultatele predicției. Totodată, pentru descrierea variației spațiale a proprietăților solului la nivel de câmp sau gospodărie este necesară colectarea unui număr suficient de probe pentru modelarea spectroscopică [5].

Calibrările locale se bazează pe setul de date al probelor de sol prelevate la nivel de câmp, iar calibrările regionale sau zonale se bazează pe datele obținute dintr-un areal geografic mai mare decât calibrările locale, care se distinge printr-o varietate de condiții pedoclimatice.

Cercetările efectuate de K.A.Sudduth și J.W.Hummel (1996), citați de K.S. Lee și al. [6], au demonstrat că arealul geografic din care se prelevează setul de probe de sol afectează precizia estimării NIR a carbonului organic din sol (COS). Predicția COS s-a redus atunci

când în crearea modelului de calibrare/predicție au fost folosite probe de sol dintr-un areal geografic mai larg.

C. Nduwamungu și coautorii [7] au subliniat necesitatea studierii diverselor modele de predicție bazate pe tipul de sol în scopul elaborării unor modele de calibrare mai stabile și mai robuste decât modelele de calibrare universale ce includ toate tipurile de sol și care s-ar putea dovedi a fi mai puțin stabile.

Modelul universal de calibrare la nivel zonal s-a dovedit a fi robust, atunci când s-a utilizat diferite tipuri și subtipuri de sol, precum și straturi de sol prelevate la adâncimi diferite [1].

Scopul acestor cercetări constă în estimarea preciziei de predicție a modelelor de calibrare la nivel local, pe un set de probe mai mic, care reprezintă un singur subtip de sol, și anume cernoziomul tipic.

MATERIALE ȘI METODE

Cercetările științifice au fost efectuate în experiența polifactorială de lungă durată a Laboratorului de Tehnologii și sisteme agrotehnice din cadrul IP Institutul de Cercetări pentru Culturile de Câmp „Selectia”, mun. Bălți. Scopul experienței rezidă în optimizarea sistemelor de rotație a culturilor, de lucrare și fertilizare a solului în vederea reducerii cheltuielilor de producere, restabilirii efective a fertilității solului și adaptării la schimbările climatice. Solul câmpului experimental este cernoziom tipic luto-argilos. Condițiile pedologice și cele ale reliefului solul din cadrul terenului folosit pentru experiențe sunt tipice zonei de nord a Republicii Moldova.

Au fost prelevate 18 profile de sol volumetrice cu ajutorul unui burghiu ce permite extragerea acestora în tuburi de plastic până la adâncimea solului de 0-50 cm și 50-100 cm. În rezultatul divizării profilelor de sol pe orizonturi genetice s-au obținut 211 probe, care au fost uscate, cernute până la 2 mm și depozitate în recipiente de sticlă, dintre care 85 de probe din numărul total au fost supuse analizelor chimice de laborator prin metode clasice/standard.

Conținutul de carbon (C) și azot (N) total a fost determinat prin metoda ANCA-MS (Automated Nitrogen Carbon Analysis – Mass Spectrometry), pe spectrometrul de masă cu raportul izotopului în flux continuu Europa 20-20x după arderea la 1 000 °C în analizorul CN Europa ANCA-GSL. Conținutul de C organic a fost determinat prin aceeași metodă ANCA-MS ca și în cazul C total, dar după îndepărtarea carbonaților. Carbonații au fost îndepărtați prin metoda fumeșării acide cu HCl, ulterior conținutul fiind calculat prin diferența C total și C organic. Anali-

zele de referință ale probelor de sol au fost efectuate în laboratorul Catedrei de Nutriție a Plantelor (*Lehrstuhl für Pflanzenernährung*) din cadrul Universității Tehnice din München (Freising, Germania).

Achiziția datelor spectrale a fost efectuată prin scanarea probelor de sol cu spectrofotometrul 6 500 NIR Systems (Foss, NIR System, Silver Spring, Md), echipat cu un modul de transport vertical (Foss), în intervalele lungimii de undă de 700-2 500 nm (regiunea infraroșu apropiat) cu rezoluția spectrală de 2 nm între punctele datelor colectate. Analiza NIRS a fost efectuată în cadrul Centrului de Științe Agricole Hans Eisenmann (HEZ – Hans Eisenmann-Zentrum für Agrarwissenschaften Weihenstephan) din Freising, Germania.

Pentru crearea modelelor de calibrare/predicție la nivel local, probele au fost modelate după cum urmează: setul de date din 85 de probe, ce aparțin unui singur subtip de sol – cernoziom tipic, a fost împărțit în 62 de probe pentru setul de calibrare și 23 de probe pentru setul de validare. Setul de calibrare a fost utilizat pentru a elabora ecuația de predicție, iar setul de validare – pentru a valida ecuația predictivă.

Datele spectrale transformate și rezultatele analizelor de laborator au fost utilizate în crearea modelelor de calibrare (predicție) folosind algoritmul de Regresie parțială minimă pătrată (*Partial Least Squares Regression* sau PLSR), care face legătură între datele spectrale cu datele solului de referință obținute în laborator, extrăgând astfel informațiile despre parametrii solului din spectrele de reflectanță ale solurilor din regiunea NIR. Pentru optimizarea calibrării a fost utilizată o procedură de validare încrucișată (cross-validare). Modelele de calibrare au fost testate și validate cu datele setului de validare. După aceasta, modelele de calibrare au fost utilizate pentru a prezice proprietățile de sol ale probelor vizate. Procesarea datelor spectrale, calibrarea și crearea modelului de predicție a fost efectuată după același principiu ca și pentru calibrarea universală la nivel zonal, descrisă de M. Ilușca [1].

Modelele au fost evaluate pentru calitatea predictivă pe baza a trei indici principali: (1) *coeficientul de determinare* (R^2), care este utilizat pentru a măsura reușita potrivirii; (2) *eroarea medie pătrată de predicție* (RMSE), care este utilizată pentru a măsura diferența dintre valorile măsurate și estimate din model; (3) *raportul dintre performanță și abatere* (RPD), care este utilizat pentru a măsura precizia modelului [8].

Analiza componentelor principale (PCA), analiza datelor spectrale, corelarea și toate analizele statistice au fost efectuate folosind software-ul Unscrambler[®]X 10.4 (CAMO Software AS) și Microsoft Excel 2010.

REZULTATE ȘI DISCUȚII

Rezultatele obținute prin aplicarea algoritmului PLSR sunt prezentate în figurile 1 și 2. În graficele de suprapunere este prezentată relația liniară a valorilor de referință ale C organic, carbonați, N total și pH obținute prin analize clasice de laborator față de valorile precise prin metoda NIRS, la etapele de calibrare (punctele albastre) și validare (punctele roșii) a modelului de predicție. Cu alte cuvinte, valorile Y precise selectate din modelele de regresie sunt reprezentate grafic în raport cu valorile Y măsurate, pentru verificarea calității modelului de regresie/calibrare. Dacă modelul oferă o potrivire bună, graficul va afișa punctele aproape într-o linie dreaptă prin origine și cu panta (*slope*) apropiată de valoarea 1.

Valorile R^2 pentru majoritatea proprietăților solului, în special N total, C organic și carbonați, nu diferă semnificativ de valoarea 1. De asemenea, după parametrul pantei liniei de regresie (*slope*) s-a constatat că acești indici ai solului se potrivesc semnificativ cu linia 1:1. În ambele cazuri, excepție face pH-ul solului a cărui parametri R^2 și panta liniei de regresie (*slope*) au înregistrat valori mai mici.

Analiza statistică a cross-validării (validării-încruciate) la etapa calibrării și rezultatele validării independente a modelelor NIR-PLSR pentru predicția

proprietăților solului investigate sunt prezentate în tabelul 1. Spectrele măsurate au reflectat variabilitatea găsită în datele de referință analizate prin metode chimice clasice.

Pentru setul de calibrare, valorile coeficientului de determinare R^2 au variat între 0,83-0,97 și 0,91-0,99 pentru calibrarea zonală și locală, respectiv. Valorile R^2 din setul de validare au fost mai mici și valorile RMSEP mai mari decât valorile corespunzătoare din setul de calibrare, dar performanța statistică nu s-a diferențiat mult. Predicția rezultată în urma validării modelelor de calibrare la nivel zonal a avut un $R^2 \geq 0,94$ pentru estimarea conținutului de C organic, carbonați și N total, în timp ce cel al pH-lui solului a fost de 0,77.

Rezultatele obținute au arătat că a existat o corelație puternică dintre spectrele NIR și indicatorii măsurați ai solului. Cu o precizie excelentă s-a prezis conținutul de carbonați, C organic și N total, obținându-se un coeficient de determinare $R^2 \geq 0,98$ și RPD > 6. O predicție bună s-a obținut la estimarea pH-lui solului ($R^2 = 0,88$ și RPD > 3), în urma validării pe un set independent.

Se consideră că estimarea cu precizie a C organic, N total și a carbonaților este asociată cu energia absorbantei și reflectanței legăturilor moleculare din regiunea NIR datorită combinației grupelor C-H, N-H, C-O, C-N și O-H. Conform acesteia, predicția

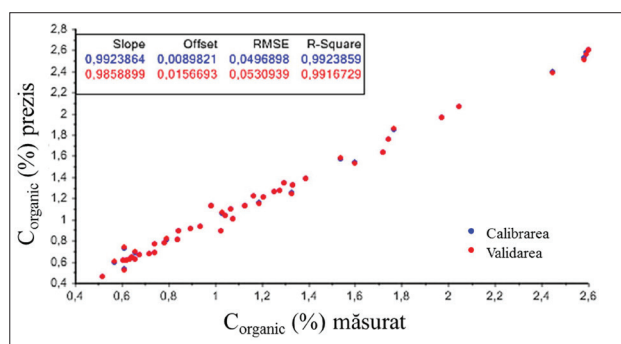


Figura 1. Valorile măsurate ale carbonului organic și ale carbonaților din sol reprezentate grafic în raport cu valorile estimate/precizie la etapele de calibrare și validare a modelului.

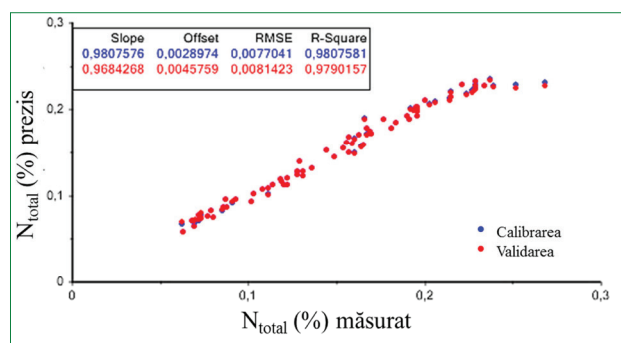
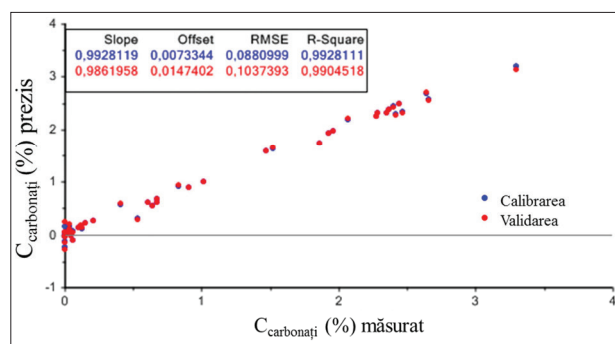
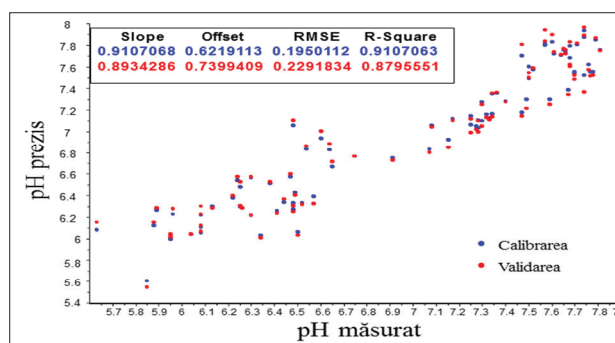


Figura 2. Valorile măsurate ale azotului total și pH-lui solului reprezentate grafic în raport cu valorile estimate/precizie la etapele de calibrare și validare a modelului.



Parametrii statistici ai modelului NIR-PLSR de calibrare și performanța (validarea) predicției acestuia

Indicii solului	Factori PLS	Etapa de calibrare				Etapa de validare			
		SD	RMSE	R ²	RPD	SD	RMSEP	R ²	RPD
		Calibrarea universală (zonală)							
pH (CaCl ₂)	9	0,77	0,32	0,83	2,38	0,80	0,38	0,77	2,11
N total, %	5	0,07	0,02	0,94	4,32	0,07	0,02	0,94	3,88
C organic, %	5	0,87	0,15	0,96	5,72	0,84	0,18	0,94	4,58
Carbonați, %	6	0,85	0,16	0,97	5,45	0,74	0,18	0,96	4,10
Calibrarea locală									
pH (CaCl ₂)	7	0,62	0,20	0,91	3,20	0,75	0,23	0,88	3,28
N total, %	6	0,05	0,01	0,98	6,95	0,06	0,01	0,98	7,62
C organic, %	3	0,67	0,05	0,99	13,46	0,79	0,05	0,99	14,85
Carbonați, %	4	0,91	0,09	0,99	10,29	1,00	0,10	0,99	9,60

cu succes a C organic, N total și carbonaților se datorează faptului că indicatorii respectivi au o influență directă asupra spectrelor ca urmare a legăturilor moleculare, a compoziției fizice sau a transmisiei și reflexiei luminii [9, 4, 10]. Predicția slabă a pH-ului solului este cauzată de faptul că această proprietate nu are un răspuns primar sau spectral direct în regiunea NIR [11]. Astfel, capacitatea NIR-PLS de a-l prezice ar depinde de relația cu materia organică și conținutul de argilă [4].

Calibrarea universală zonală pe un set de date cu o gamă largă de variație (diverse tipuri și subtipuri de sol) a generat o precizie mai mică atunci când a fost testată pe întregul set de probe (n = 554) format din probe ce reprezintă o varietate de tipuri și subtipuri de sol, decât calibrarea locală pe un set de date

cu o gamă îngustă (un singur subtip de sol). RMSE pentru calibrările zonale și locale au oscilat între 0,02-0,32 % și 0,01-0,20 %, respectiv, ceea ce arată că ambele calibrări au capacitate de predicție pentru intervalele respective. În urma validării modelelor de calibrare, valorile RMSE au oscilat între 0,02-0,38 % pentru calibrările zonale și 0,01-0,23 % pentru cele zonale. Cu toate acestea, rezultatele au fost mai bune atunci când au fost elaborate calibrări numai pentru un singur subtip de sol, și anume cernoziomul tipic.

Pe baza statisticilor descriptive ale valorilor prezise, este evident că proprietățile solului (Nt, Corg, carbonați, pH) prezise din datele spectrale nu au fost semnificativ diferite de cele determinate prin metode standard de laborator (figura 3).

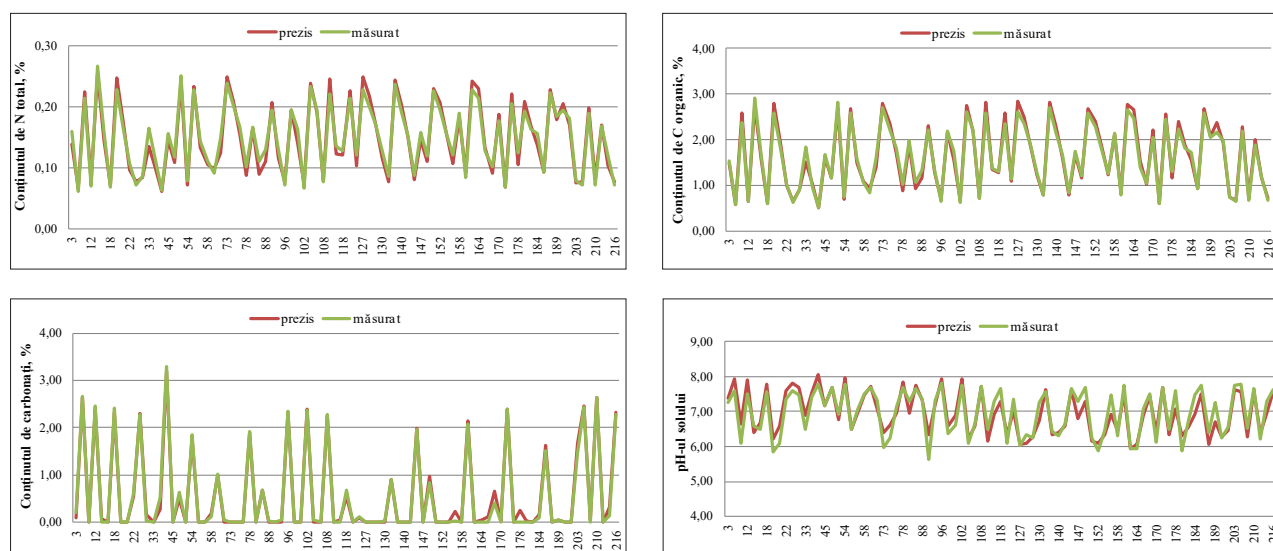


Figura 3. Graficele de suprapunere a valorilor prezise și măsurate ale conținutului de N total, C organic, carbonați și pH-ul solului în experiența polifactorială a ICC „Selecția”.

Potrivit graficelor de suprapunere prezentate în figura 3, s-a evidențiat că majoritatea valorilor N total, C organic, carbonaților și a pH-ului prezise din datele spectrale au fost aproape identice cu valorile măsurate, cu o mică deviere.

M. Todorova și S. Atanassova [12] au subliniat că pentru analiza cantitativă a proprietăților solului, precizia predicțiilor prin metoda spectroscopiei NIR este îmbunătățită semnificativ prin gruparea probelor pe baza claselor taxonomice ale solului și a similitudinii spectrale, așa cum fiecare tip de sol conține caracteristici spectrale specifice în regiunea NIR.

B. Kuang și A.M.Mouazen [11] au comparat precizia modelelor generale de calibrare în baza probelor colectate din trei gospodării din Europa cu cea a modelelor de calibrare la nivel de gospodărie. Cercetările efectuate de acești autori au demonstrat o predicție cu succes atât pentru modelele de calibrare generale, cât și pentru cele individuale la nivel de câmp. Din cele trei modele de calibrare individuale la nivel de câmp, calibrarea și validarea conținutului de C total, de N total și de C organic a solului pentru două modele s-au dovedit a fi mai eficiente decât în cazul unui model general. Validarea modelelor a înregistrat un R^2 în limitele 0,74-0,96 și un RPD între 2,00-4,95. O predicție mai slabă a fost obținută pentru pH-ul solului. Ei considerau că eroarea cauzată de variabilitatea mai mare a culorii, texturii și conținutului de umiditate a solului în modelele de calibrare generale explică rezultatul mai puțin reușit în comparație cu modelele individuale la nivel de gospodărie [11].

De asemenea, autorii au ajuns la concluzia că la nivel de gospodărie sau câmp calibrările nu întotdeauna oferă cea mai bună performanță, din cauza devierii standard (SD) și a intervalului îngust al concentrațiilor probelor. O deviere standard mai mare și o gamă mai largă a setului de date în calibrare au dus la valori R^2 și RPD mai mari, dar și la valori RMSEP mai mari [11].

În acest studiu, pentru crearea modelelor de calibrare universale zonale, probele de sol au fost prelevate de pe o suprafață geografică vastă a zonei de Nord a Republicii Moldova, care se caracterizează printr-o gamă largă de proprietăți fizice și chimice datorită variațiilor în modul de gestionare a terenului, covorul vegetal și condițiile climatice specifice. Acest fapt a generat obținerea de spectre NIR distincte, oferind oportunitatea de a testa eficacitatea modelelor spectroscopice NIR în combinație cu algoritmul PLSR pentru estimarea sau predicția concentrațiilor de C (C total, C organic și carbonați), N total și a valorilor pH-ului în sol la diferite scări (regionale, zonale sau locale), cuprinzând zonele agricole cheie din nordul țării.

Separarea întregii populații în subgrupuri mai omogene, cu tip de sol similar, a contribuit la o îmbunătățire a preciziei predicției pentru carbonul organic, carbonați, azot total și pH-ul solului. O calibrare universală zonală este mai puțin precisă în comparație cu calibrarea locală sau specifică bazată pe tipul de sol și practicile agricole. Acest lucru arată că variabilitatea solurilor și a tipurilor de sol folosite în crearea modelului de predicție influențează direct performanța acestuia.

CONCLUZII

Rezultatele obținute în studiul dat demonstrează că diapazonul calibrării trebuie să fie selectat corespunzător relevanței scopului practic pentru a obține beneficiul maxim al metodei analitice NIRS.

Rezultatele obținute au arătat că precizia de predicție a modelului de calibrare la nivel local ($R^2 \geq 0,98$ și RPD > 6 pentru conținutul de carbonați, C organic și N total; $R^2 = 0,88$ și RPD > 3 pentru pH-ul solului) a fost mai bună decât cea a modelului de calibrare zonal ($R^2 \geq 0,94$ și RPD > 3 pentru conținutul de carbonați, C organic și N total; $R^2 = 0,77$ și RPD > 2 pentru pH-ul solului).

Totodată, rezultatele studiului confirmă utilitatea spectroscopiei NIR pentru predicția diferitor proprietăți chimice ale cernoziomurilor Republicii Moldova. Prin urmare, spectroscopia NIR ar putea fi utilizată drept un instrument analitic rapid pentru monitorizarea managementului solului și evaluarea fertilității solului.

BIBLIOGRAFIE

1. Ilușca M. Potențialul metodei NIRS în determinarea texturii solurilor din zona de Nord a Republicii Moldova, in: *Tradiție și inovare în cercetarea științifică*. Ediția 10, Vol. I, 8 octombrie 2021, Bălți. Universitatea de Stat „Alecu Russo” din Bălți, 2021, pp. 183-189.
2. Wetterlind J., Stenberg B., Viscarra Rossel R. Soil analysis using visible and near infrared spectroscopy, in: *Methods in molecular biology* (Clifton, N.J.). 2013, 953, 95-107.
3. Chodak M. Application of near infrared spectroscopy for analysis of soils, litter and plant materials, in: *Polish Journal of Environmental Studies*. 2008, 17(5), 631-642.
4. Zornoza R., Guerrero C., Mataix-Solera J., Scow K.M., Arcenegui V., Mataix-Beneyto, J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils, in: *Soil Biology and Biochemistry*. 2008, 40(7), 1923-1930, <https://doi.org/10.1016/j.soilbio.2008.04.003>
5. Shi Z., Ji W., Viscarra Rossel R.A., Chen S., Zhou Y. Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chi-

nese vis-NIR spectral library, in: *European Journal of Soil Science*, 2015, 66(4), 679-687.

6. Lee K.S., Sudduth K.A., Drummond S.T., Lee D.H., Kitchen N.R., Chung S.O. Calibration Methods for Soil Property Estimation Using Reflectance Spectroscopy, in: *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*. 2010, 53(3), 675-684, doi: 10.13031/2013.30059.

7. Nduwamungu C., Ziadi N., Parent L-É., Tremblay G.F., Thuriès L. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review, in: *Canadian Journal of Soil Science*, 2009, 89(5), 531-541, <https://doi.org/10.4141/CJSS08076>

8. Heil K., Schmidhalter U. An evaluation of different NIR-spectral pre-treatments to derive the soil parameters C and N of a humus-clay-rich soil, in: *Sensors*, 2021, 21(4), 1423, <https://doi.org/10.3390/s21041423>

9. Ben-Dor E., Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties, in: *Soil Science Society of American Journal*, 1995, 59(2), 364-372.

10. Askari M., O'Rourke S., Holden N. Evaluation of soil quality for agricultural production using visible-near-infrared spectroscopy, in: *Geoderma*, 2015, 243, 80-91.

11. Kuang B., Mouazen A.M. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms, in: *European Journal of Soil Science*, 2011, 62(4), 629-636.

12. Todorova M., Atanassova S. Near infrared spectra and soft independent modelling of class analogy for discrimination of Chernozems, Luvisols and Vertisols, in: *Journal of Near Infrared Spectroscopy*, 2016, 24(3), 271-280.



Lică Sainciuc. *Eterna țigară*, 1974, tuș, pânză, 90 × 60 cm.