

Original Research

Soil Quality Analysis Using Modern Statistics and NIR Spectroscopy Procedure

Ibrahim M. Almanjahie¹, Ishfaq Ahmad^{1,2*}, Zouaoui Chikr Elmezouar¹,
Ali Laksaci¹

¹Department of Mathematics, College of Science, King Khalid University, 61413, Abha, Kingdom of Saudi Arabia

²Department of Mathematics and Statistics, International Islamic University, Islamabad, Pakistan

Received: 1 August 2018

Accepted: 21 October 2018

Abstract

In this paper, we combine the recent development on mathematical statistics with modern chemistry in order to provide a new approach for soil quality analysis. Precisely, from modern chemistry we use near-infrared reflectance (NIR) spectroscopy procedure as a fast, accurate and inexpensive tool to evaluate chemical properties. Based on the collected data, the relationship between soil quality variables is modeled by using the functional statistics, which allows for analyzing a data as a curve or an image. The used predictor models are functional classical regression (FCR), functional local linear regression (FLLR), functional relative error regression (FRER) and functional robust regression (FRR). We prove that the performance of these models is closely linked to the homogeneity of the data. Considering the Abisko soil data, we show that FNR and FLLR are suitable for soil organic matter data, while for the Ergosterol concentration data the use of FRER and FRR are adequate. Furthermore, the proposed functional approach, permits us to avoid many drawbacks of the classical approach as principal component regression (P.C.R.).

Keywords: soil quality, near infrared spectroscopy, functional relative error regression

Introduction

Assessment of soil quality plays a vital role in the management of environmental and agricultural resources. In order to get optimum results from the inputs in agriculture, cost-effective soil analysis is very important [1-5]. We have to make efficient use of different inputs, such as pesticides and fertilizers not only for crop production but also to mitigate and abolish pollution risks resulting from the over-application

of different agrochemicals [6-10]. NIR spectroscopy is a rapid, inexpensive, and precise technology for predicting a range of soil properties in ecological studies [11-12].

In soil sciences, NIR spectroscopy is used to measure different factors such as clay content, cation-exchange capacity, specific surface area, organic matter, moisture content, and concentrations of organic and inorganic fractions of nitrogen and carbon [13-14]. In literature, some studies have also tried to predict microbial, fungal biomass, respiration in boreal [15-18] and temperate forest soils [19-21]. As silica, being the major part of the inorganic substance in the

*e-mail: Ishfaq.ahmad@iiu.edu.pk

soil, is transparent to infrared radiation, on the other side, NIR spectroscopy, when applied on high organic matter soils, produces different results as compared to high mineral content soils. However, measurements on forest humus advocate for the use of NIR spectroscopy on organic soil [15-16]. Arctic soils, which account for 14% of the total world organic carbon in the soil [22], experience rapid progressive climatic warming. Greater attention is being paid to the science of climate change, and therefore the applicability of NIR spectroscopy techniques to this highly organic soil.

NIR spectroscopy usually measures the bond structure of molecules through their vibrational transitions. A wide range of molecules belonging to organic compounds absorbs in the NIR spectroscopy range, which makes it a reasonable technique not only for the qualitative analysis but also for quantitative analysis of organic matters. NIR spectroscopy is fast and invasive because, in NIR spectroscopy mostly, samples are analyzed without any pretreatment.

The goal of this study was to search the prediction models of NIR spectroscopy to determine the essential quality indices of the soil. A total of 108 samples of soil data from Abisko, northern Sweden were collected. Our purpose in this study was to evaluate the performance of NIR spectroscopy in measuring both the Ergosterol concentration and the soil organic matter. Some alternative statistical models to the PCR or the linear partial least squares regression are employed in this prediction setting. The proposed models come from the recent development of the modern Statistics that allow for analyzing the big data without reducing the dimension. These new models are taken using the notion of functional statistics, which allows us to analyze the data as curves instead of some numerical numbers. Notice that the functional statistics have encountered a strong infatuation in these last years, as evidenced by the several special issues dedicated to this topic by various papers [e.g., 23-24]. Such statistical analysis is motivated by the recent technological development of the measuring instruments and the informatics tools that allow for the recovery of increasingly bulky data being recorded densely over a thinner discretization grid, which makes them intrinsically a continuous curve. One of the advantages of this approach is that it takes into account the whole curves in its continuous path instead of taking some discrete points. Of course with this consideration we keep all the information in the sampled data. In this contribution, we will predict the chemical and the microbiological properties of soil quality by using a various nonparametric functional regression approach. More precisely, we use four regression models: FNR, FLLR, FRER and FRR. A comparison between these models is carried out by evaluating their rapidness, accuracy and robustness in the prediction of the ergosterol concentration and the peregrinate soil organic matter.

Methods and Materials

Samples Collection and their Pretreatment

Soil samples were obtained in the result of a long field experiment in Abisko, northern Sweden (681210 N, 181490 E). The soil samples taken from the site have a depth of organic horizon of 12–15 cm, while the vegetation is controlled by *Cassiope tetragona* (L.) *D. Don*. The samples for analysis were obtained from 6 different replications, including: 1) high nutrient availability, provision of NPK fertilization; 2) high temperature, provision of passive open top greenhouse; 3) cloudiness, provision of shading treatment; 4) provision of mixture of fertilization as warming conditions; 5) mixture of fertilization and shading environment; and 6) use of control plots without any manipulation. In the result of this, we get 36 plots (120 *120 cm). The details of these plots are available in [25-26]. In short, only four soil cores of 10 cm depth were obtained from each of 36 plots. Furthermore, each core was divided into two equal parts of 0-5 cm and 5-10 cm. Two subsamples were formulated per plot for surface layer while only one sample for the deeper layer. In this way, we obtained 108 samples (72 for the surface layer and 37 for the deeper layer). The soil samples were also made homogenous after the removal of roots and stones, and these samples were further dried in the oven for 15 minutes at 70°C. Some other related studies are available [e.g., 27-29].

Recently Developed Regression Models

The chemical and microbiological properties of this data set have been described in detail by [30]. Nevertheless, from a statistical point of view, we have observed that there is wide variability in the Ergosterol concentration. Indeed, it ranges 53.7 to 434.15 µg/g soil and its standard deviation, whose measure of the dispersion of the data is 85.49 µg/g soil. On the other hand, the soil organic matter (SOM) index is more

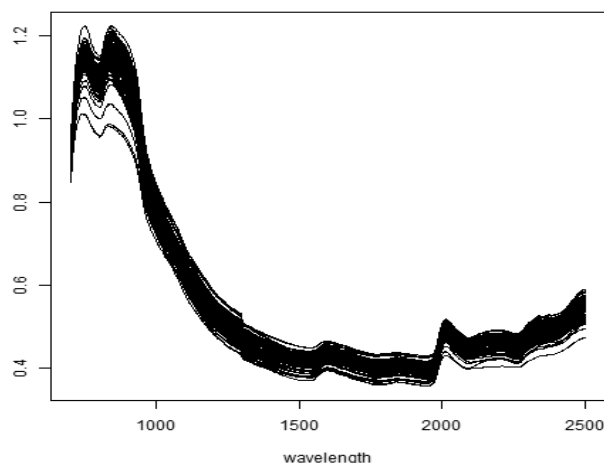


Fig. 1. NIR spectroscopy curves.

homogenous than the Ergosterol concentration. In particular, the SOM-percentage of 108 samples ranges from 43% to 95% and has a standard deviation equal to 10.82. We return to [30] for the NIR spectroscopy analysis of these data.

From Fig. 1, it appears clear that the shape of the spectra contains some peaks that hide information. Therefore, it is challenging to evaluate the quantity of the chemical and microbiological properties through the NIR spectroscopy analysis. Thus, some statistical models might be necessary to quantify these properties. In this context, multiple linear regression (MLR), PCR and partial least squares regression (PLSR) are the most popular models of this kind of data. However, the implementation of these models is based on the dimension reduction of the data. Alternatively, in this paper we use some new statistical models in order to investigate the soil quality without dimension reduction. These new models come from the recent development of modern statistics, allowing us to analyze the data as curves. The main advantage of this approach is the fact that it takes into account the whole curves in its continuous path and not only on some discrete points. Of course, with this consideration we keep all the information in the sampled data. As pointed out in the introduction, we will predict the chemical and the microbiological properties of the soil quality by using a varied nonparametric functional regression approach. To fix the ideas, let's present the mathematical formulation of the prediction problem, the regression formula:

$$Y = r(X) + error \tag{1}$$

...where Y is a real variable means – either the SOM-percentage or the Ergosterol concentration. The functional variable is the curve of the NIR spectroscopy (700-2500 wavelength). So, the prediction of Y is based on the determination of function $r(\cdot)$. In modern statistics there exists a various nonparametric approach that can be used to model this prediction issue. In order to provide the results to the users of NIR spectroscopy, we analyzed a varied collection of statistical predictors. We present in the rest of this section some regression models recently developed in functional statistics.

Functional Classical Regression (FCR)

Classical regression was introduced in functional statistics by [31-38], where they studied the functional version of the Nadaraya Watson estimator, which is obtained by

$$r(x) = \operatorname{argmin}_f E[(Y - f)^2 | X = x] \tag{2}$$

By a simple derivation we get explicitly the estimator

$$r(x) = E[Y | X = x] \tag{3}$$

Then for all given spectra curves we predict its SOM-percentage or the Ergosterol quantity by

$$\tilde{Y} = \tilde{r}(x) \tag{4}$$

...where $\tilde{r}(x)$ is the kernel estimator of $r(x)$ as defined by:

$$\tilde{r}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{\|x - X_i\|}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\|x - X_i\|}{h_n}\right)} \tag{5}$$

Functional Local Linear Regression (FLLR)

This regression model is obtained by combining the linear approach with the nonparametric model. The local linear regression inherits the advantages of the two regression models. In particular, from the linear property of this model we reduce the convergence rate and from the nonparametric one we explore the local structure of the curves. This statistical model is obtained by assuming that the function has a linear form at the neighborhood of a given spectra curve x . Mathematically, this assumption is expressed by:

$$\forall x \text{ in neighborhood of } x: r(x) \approx a + b\beta(x, x) \tag{6}$$

...where the quantities a and b are obtained by minimizing

$$(a, b) = \operatorname{argmin}_{f, g} \sum_{i=1}^n (Y_i - f - g\beta(X_i, x))^2 K\left(\frac{\|x - X_i\|}{h_n}\right) \tag{7}$$

...with $\beta(x, x)$ being the known function used to measure the proximity between the curves. The local linear estimator (LLE) $r(x)$ can be explicitly expressed as follows

$$\tilde{r}(x) = \frac{\sum_{i,j=1}^n W_{i,j} Y_i}{\sum_{i,j=1}^n W_{i,j}} \tag{8}$$

...where:

$$W_{i,j} = \beta(X_i, x) \left(\beta(X_i, x) - \beta(X_j, x) \right) K\left(\frac{\|x - X_i\|}{h_n}\right) K\left(\frac{\|x - X_j\|}{h_n}\right) \tag{9}$$

We point out that this estimator is a functional version of the LLE proposed by Fan et al. [37]. Such a version has been introduced by [40-44].

Functional Relative Error Regression (FRER)

This regression model was introduced in functional statistics by Demongeot et al. [43] and further used by

[46-47]. It constitutes an alternative of the classical regression model. The advantage of this model is the fact that it takes into account the largest value of the response variable. Indeed, the least square of this regression is defined relative to the values of the response variable as follows. For a given curve the predictor of the response variable is a solution of:

$$\min_f E \left[\frac{(Y-f)^2}{Y^2} | X = x \right]. \tag{10}$$

Once again this regression is very easy to implement in practice, because its expression can be explicitly defined by:

$$\tilde{r}(x) = \frac{\sum_{i=1}^n Y_i^{-1} K\left(\frac{\|x-X_i\|}{h_n}\right)}{\sum_{i=1}^n Y_i^{-2} K\left(\frac{\|x-X_i\|}{h_n}\right)} \tag{11}$$

Functional Robust Regression (FRR)

The M-regression or the robust regression is an old alternative model introduced by Huber (1964). It allows us to correct many drawbacks of the classical regression such as heteroscedasticity and the presence of the outliers. The functional version of this model has been introduced by [48] and also implemented in [49]. This model is defined as the solution of the following optimization problem:

$$\min_f E[\rho(Y, f) | x = x] \tag{12}$$

...where ρ is a real-valued Borel function chosen by the user according to the studied data. It is clear that this mode covers and includes many usual nonparametric models, for example for $\rho(Y, f) = (Y - f)^2$ we obtain the classical regression and $\rho(Y, f) = |Y - f|$ leads to the conditional median. More examples can be found in [50]. With this regression, the predictor is obtained by resolving the empirical version of the optimization problem that is:

$$\bar{r}(x) = \operatorname{argmin}_f \frac{\sum_{i=1}^n \rho(Y_i, f) K\left(\frac{\|x-X_i\|}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\|x-X_i\|}{h_n}\right)} \tag{13}$$

Results and Discussion

We point out that the performance of all these models is closely linked to the different parameters involved. In fact, the most influenced parameters in this prediction issue are smoothing parameter h_n , kernel K and the parameters related to the distance between the curves. Concerning this last point, most of the metrics are based on some usual interpolation such as trigonometrics, spline or by wavelet basis functions [32]. In all these cases, we approximate curve x by using a basis function, and we write:

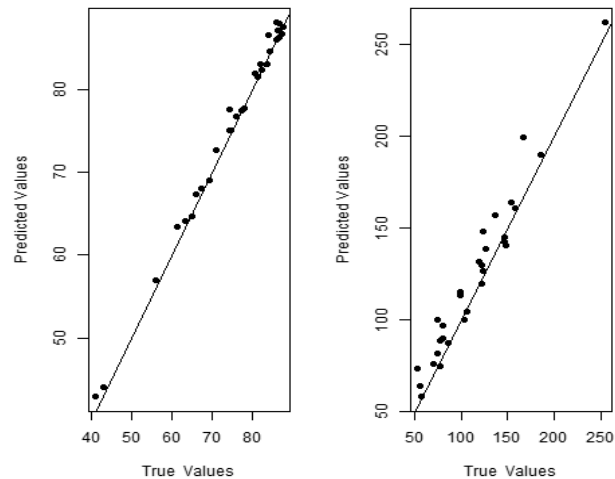


Fig. 2. Prediction results of functional classical regression (SOM on the left and Ergosterol at right).

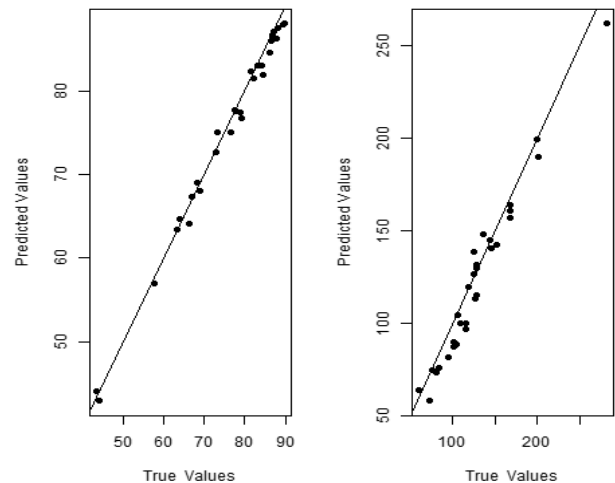


Fig. 3. Prediction results of functional local linear regression (SOM on the left and Ergosterol to the right).

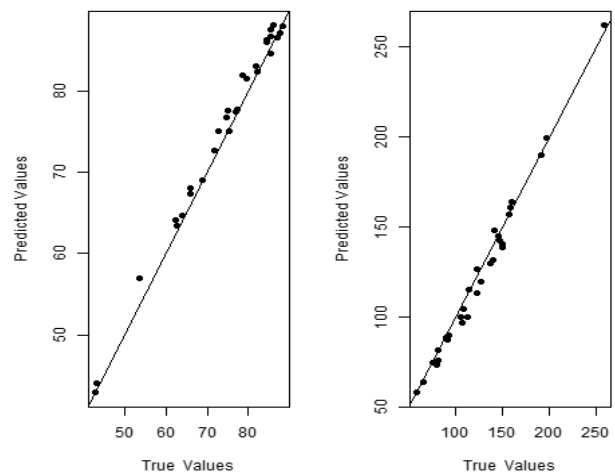


Fig. 4. Prediction results of functional relative error regression (SOM to the left and Ergosterol at right).

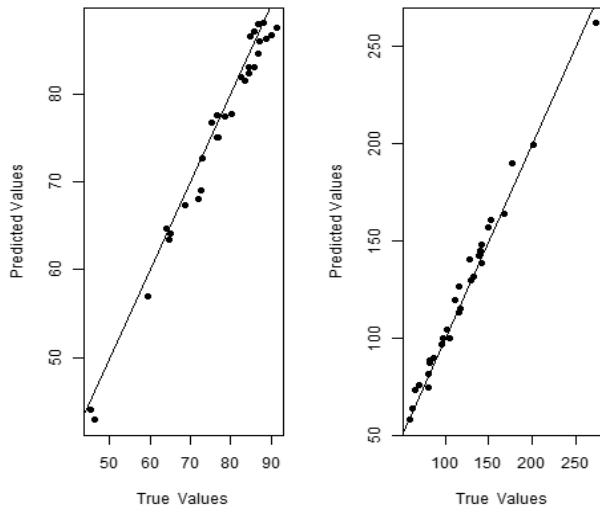


Fig. 5. Prediction results of functional robust regression (SOM to the left and Ergosterol at right).

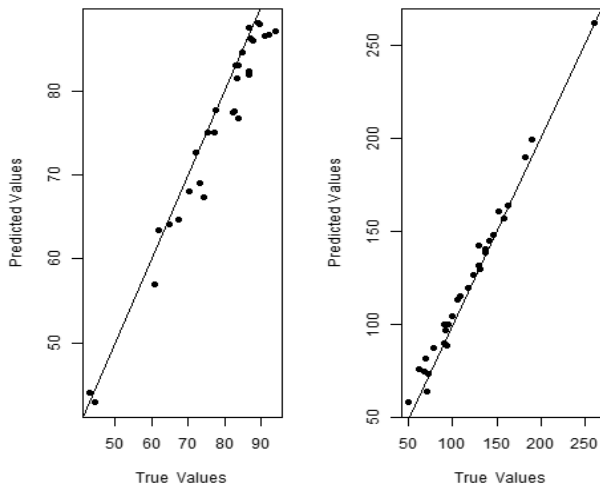


Fig. 6. Prediction results of principal component regression (SOM at left and Ergosterol to the right).

$$x(t) \approx \sum_{i=1}^{\infty} c_i B_i(t) \tag{14}$$

...where is $(B_i(\cdot))_i$ a given orthonormal basis function. In practice, we use a penalized version of this approximation, and we put:

$$x(t) \approx \sum_{i=1}^k c_i B_i(t) \tag{15}$$

Thus, the distance between the curves is computed between their interpolation functions. Readers can find in [32] more details on the mathematical formulation for this notion of a functional metric. For this prediction issue, after dividing the 108-sample into parts (learning sample and testing sample), we use the leave-one-out cross-validation technique to select parameters and the bandwidth parameter. We point out that the cross-validation procedure is implemented with respect to the least square error defined by

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \theta(X_i))^2 \tag{16}$$

...where $\theta(\cdot)$ means one of the previous regression models. Concerning the locative function, we choose the same function used by Rachedi et al. [51]. Using the same kernel and score function as in [52], we obtained the following results (Figs 2-5).

Finally, we compare these results to the PCR. The latter is obtained by using the routine `pcr` in the R-package `PLS`.

It is clear that these new regression models have satisfactory prediction results and they improve substantially the prediction results of the PCR studied by [30] (see Fig. 6). However, the performance of the studied models is varied with respect to the homogeneity of the data. For instance, the FCR and the FLLM are adequate for the SOM variable, which is more homogeneous than the Ergosterol concentration. On the other hand, the FREL and the FRR are more adapted to the Ergosterol concentration. Such a conclusion is justified by the cross-validated prediction error (CVPE) as defined in [16] and presented in Table 1.

Table 1. Cross-validated prediction error (CVPE) of the different regression models.

Model	CVPE-FCR	CVPE -FLLR	CVPE -FRER	CVPE -FRR	CVPE -PCR
SOM	0.87	0.74	0.98	0.94	2.91
Ergosterol	1.27	1.12	0.91	0.89	3.19

Table 2. Computational time (CT) of the different regression models.

Model	FCR-CT	FLLR-CT	FRER-CT	FRR-CT	PCR-CT
SOM	0.45	0.56	0.49	0.53	0.44
Ergosterol	0.43	0.51	0.42	0.57	0.50

Finally, we point out that these proposed models are very fast in predicting the chemical properties of the soil data. In Table 2 we give the computational time (CT) of each model using proc-time. It seems clear that, for the SOM case, the FCR and PCR have small CT as compared to the FLLR, FRER and FRR models. While for the Ergosterol concentration data the fast models are FCR and FRER.

Conclusion

In this work we have developed a new approach to predicting the soil quality of Alisko in terms of its chemical and microbiological properties. Typically, we have focused on both the prediction of the soil organic matter (SOM) and the Ergosterol concentration. This new approach combines recent developments in chemistry and modern statistics. Specifically, from chemistry we use NIR spectroscopy to analyze the collection of some regression models recently developed in nonparametric statistics. The technology of NIR spectroscopy is the fastest and most accurate method. Moreover, it reduces the need for conventional wet chemistry procedures. Next, the functional statistics allows us to explore all the information of the spectroscopy analysis where the spectra are viewed as curves. These models are easily implementable, and their efficiency is related to the homogeneity of the studied data. In this sense for each case, we can choose the adapted model. Moreover, the proposed models permit us to avoid the core drawback of the PCR model that is the loss of information after the predictors transformation. Indeed, this transformation is carried out through some directions chosen independently to the prediction problem. In particular, the response variable does not intervene in the choice of the optimal direction, which can influence negatively the prediction issue. Thus, the originality of the functional approach comes from the fact that the prediction problem is performed without any transformation of the data. We emphasize that such a combination of modern chemistry and modern statistics is very beneficial ecologically, and is more flexible, robust and fast. On the other hand, it is not necessary to add another chemical product, and with respect to the nature of data, we can choose an adequate model among the various functional models that exist. Last but not the least, in this study we have concentrated only on four models, but there exist a wide range of statistical models that can be used to describe this kind of data.

Acknowledgements

Authors are very grateful to the Deanship of Scientific Research at King Khalid University, Kingdom of Saudi Arabia for their administrative and technical

support and for funding this work through a research group program under project No. RGP1/50/39.

References

- MÖLLER K., OBERSON A., BÜNEMANN E.K., COOPER J., FRIEDEL J.K., GLAESNER N., HÖRTENHUBER S., LØES A.K., MÄDER P., MEYER G., MÜLLER T. Improved phosphorus recycling in organic farming: navigating between constraints. In: *Advances in Agronomy*. **147**, 159, **2018**.
- WOLLMANN I., GAURO A., MÜLLER T., MÖLLER K. Phosphorus bioavailability of sewage sludge-based recycled fertilizers. *Journal of Plant Nutrition and Soil Science*. **181**, 158, **2018**.
- TEN HOEVE M., BRUUN S., NAROZNOVA I., LEMMING C., MAGID J., JENSEN L.S., SCHEUTZ C. Life cycle inventory modeling of phosphorus substitution, losses and crop uptake after land application of organic waste products. *The International Journal of Life Cycle Assessment*, **45**, 1, **2017**.
- RAYMOND N.S., STÖVER D.M., JENSEN L.S., HÅKANSSON S. Survival and phosphate solubilisation activity of desiccated formulations of *Penicillium bilaiae* and *Aspergillus niger* influenced by water activity. *Journal of microbiological methods*, **1**, 150, **2018**.
- WEISSENGRUBER L., MÖLLER K., PUSCHENREITER M., FRIEDEL J.K. Long-term soil accumulation of potentially toxic elements and selected organic pollutants through application of recycled phosphorus fertilizers for organic farming conditions. *Nutrient Cycling in Agroecosystems*, **110**, 427, **2018**.
- SMITH L.E., SICILIANO G. A comprehensive review of constraints to improved management of fertilizers in China and mitigation of diffuse water pollution from agriculture. *Agriculture, Ecosystems & Environment*, **209**, 15, **2015**.
- HUANG J., XU C.C., RIDOUTT B.G., WANG X.C., REN P.A. Nitrogen and phosphorus losses and eutrophication potential associated with fertilizer application to cropland in China. *Journal of Cleaner Production*. **159**, 171, **2017**.
- PANG J., CHEN X., ZHANG Z., LI H. Measuring eco-efficiency of agriculture in China. *Sustainability*. **8**, 398, **2016**.
- LI J., RODRIGUEZ D., TANG X. Effects of land lease policy on changes in land use, mechanization and agricultural pollution. *Land Use Policy*, **64**, 405, **2017**.
- PAN D., KONG F., ZHANG N., YING R. Knowledge training and the change of fertilizer use intensity: Evidence from wheat farmers in China. *Journal of environmental management*. **197**, 130, **2017**.
- FOLEY W.J., MCILWEE A., LAWLER I., ARAGONES L., WOOLNOUGH A.P., BERDING N., Ecological applications of near infrared reflectance spectroscopy a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia* **116**, 293, **1998**.
- SUMMERS D., LEWIS M., OSTENDORF B., CHITTLEBOROUGH D. Visible NIRspectroscopy as a predictive indicator of soil properties. *Ecological Indicators*, **11**, 123, **2011**.
- ISLAM K., SINGH B., MCBRATNEY A. Simultaneous estimation of several soil properties by ultra-violet, visible, and NIRspectroscopy. *Australian Journal of Soil Research*, **41**, 1101, **2003**.

14. VELASQUEZ E., LAVELLE P., BARRIOS E., JOFFRE R., REVERSAT F., Evaluating soil quality in tropical agroecosystems of Colombia using NIRS. *Soil Biology & Biochemistry* **37**, 889, **2005**.
15. FRITZE H., JÄRVINEN P., HIUKKA R., Near-infrared characteristics of forest humus are correlated with soil respiration and microbial biomass in burnt soil. *Biology and Fertility of Soils* **18**, 80, **1994**.
16. PIETIKAINEN J., FRITZE H. Clear-cutting and prescribed burning in coniferous forest-comparison of effects on soil fungal and total microbial biomass, respiration activity and nitrification. *Soil Biology & Biochemistry* **27**, 101, **1995**.
17. PLUCHON N., VINCENT A.G., GUNDALE M.J., NILSSON M.C., KARDOL P., WARDLE D.A. The impact of charcoal and soil mixtures on decomposition and soil microbial communities in boreal forest. *Applied soil ecology*, **99**, 40, **2016**.
18. BARDGETT R.D., DE VRIES F.T., VAN DER PUTTEN W.H. Soil Biodiversity and Ecosystem Functioning. In *Microbial Biomass: A Paradigm Shift in Terrestrial Biogeochemistry*, **45**, 119, **2017**.
19. CHODAK M., KHANNA P., BEESE F. Hot water extractable C and N in relation to microbiological properties of soils under beech forests. *Biology and Fertility of Soils* **39**, 123, **2003**.
20. LUDWIG B., KHANNA P.K., BAUHUS J., HOPMANS P. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecology and Management* **171**, 121, **2002**.
21. CHODAK M., LUDWIG B., KHANNA P., BEESE F. Use of near infrared spectroscopy to determine biological and chemical characteristics of organic layers under spruce and beech stands. *Journal of Plant Nutrition and Soil Science*, **165**, 27, **2002**.
22. POST W.M., EMANUEL W.R., ZINKE P.J., STANGENBERGER A.G. Soil carbon pools and world life zones. *Nature*, **298**, 156, **1982**.
23. FERRATY F. High-dimensional data: a fascinating statistical challenge. *Journal of Multivariate Analysis*. **101**, 305, **2010**.
24. ANEIROS G., CAO R., FRAIMAN R., VIEU P. Special issue on functional data analysis and related topics. *J. Multivariate Anal.* (**2018** to appear).
25. HAVSTRÖM M., CALLAGHAN T.V., JONASSON S. Differential growth responses of *Cassiope tetragona*, an arctic dwarf-shrub, to environmental perturbations among three contrasting high-and subarctic sites. *Oikos*. **389**, **1993**.
26. RUMPF S.B., SEMENCHUK P.R., DULLINGER S., COOPER E.J. Idiosyncratic responses of high arctic plants to changing snow regimes. *PLoS One*. **11**, 862, **2014**.
27. XU Z., HU T., ZHANG Y. Effects of experimental warming on phenology, growth and gas exchange of treeline birch (*Betula utilis*) saplings, Eastern Tibetan Plateau, China. *European Journal of Forest Research.*, **131**, 811, **2012**.
28. GORNALL J.L., WOODIN S.J., JÓNSDÓTTIR I.S., VAN DER WAL R. Balancing positive and negative plant interactions: how mosses structure vascular plant communities. *Oecologia*. **166**, 769, **2011**.
29. VAN DER WAL R., STIEN A. High- arctic plants like it hot: a long- term investigation of between- year variability in plant biomass. *Ecology*. **95**, 341, **2014**.
30. RINNAN R., MICHELSEN A., BAÅTH E., JONASSON S. Fifteen years of climate change manipulations alter soil microbial communities in a subarctic heath ecosystem, *Global Change Biology* **13**, 28, **2007**.
31. FERRATY F., VIEU P. Nonparametric functional data analysis. Theory and Practice., Springer Series in Statistics. New York. **2006**.
32. FERRATY F., VIEU P. Nonparametric functional data analysis: theory and practice. Springer Science & Business Media; **2006**.
33. HORVÁTH L., KOKOSZKA P. Inference for functional data with applications. Springer Science & Business Media; **2012**.
34. PESARIN F., SALMASO L. Permutation tests for complex data: theory, applications and software. John Wiley & Sons; **2010**.
35. SUN Y., GENTON M.G. Functional boxplots. *Journal of Computational and Graphical Statistics*. **20**, 316, **2011**.
36. FEBRERO-BANDE M., DE LA FUENTE M.O. Statistical computing in functional data analysis: The R package *fd*. *usc. Journal of Statistical Software*. **51**, 1, **2012**.
37. VILAR J.M., CAO R., ANEIROS G. Forecasting next-day electricity demand and price using nonparametric functional methods. *International Journal of Electrical Power & Energy Systems*. **39**, 48, **2012**.
38. WANG J.L., CHIOU J.M., MÜLLER H.G. Functional data analysis. *Annual Review of Statistics and Its Application*. **3**, 257, **2016**.
39. FAN J., HECKMAN N.E., WAND M.P. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association*. **90**, 141, **1995**.
40. BARRIENTOS-MARIN J., FERRATY F., VIEU P. Locally modelled regression and functional data. *Journal of Nonparametric Statistics*. **22**, 617, **2010**.
41. GOIA A., VIEU P. An introduction to recent advances in high/infinite dimensional Statistics. **2016**.
42. FERRATY F., LAKSACI A., TADJ A., VIEU P. Kernel regression with functional response. *Electronic Journal of Statistics*. **5**, 159, **2011**.
43. BERLINET A., ELAMINE A., MAS A. Local linear regression for functional data. *Annals of the Institute of Statistical Mathematics*. **63**, 1047, **2011**.
44. DEMONGEOT J., LAKSACI A., RACHDI M., RAHMANI S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A*. **76**, 328, **2014**.
45. DEMONGEOT J., HAMIE A., LAKSACI A., RACHDI M. Relative-error prediction in nonparametric functional Statistics: Theory and practice. *Journal of Multivariate Analysis*. **146**, 261, **2016**.
46. ALTENDJI B., DEMONGEOT J., LAKSACI A., RACHDI M. Functional data analysis: estimation of the relative error in functional regression under random left-truncation model. *Journal of Nonparametric Statistics*. **30**, 472, **2018**.
47. HAO M., LIN Y., ZHAO X. A relative error-based approach for variable selection. *Computational Statistics & Data Analysis*. **103**, 250, **2016**.
48. AZZEDINE N., LAKSACI A., OULD-SAÏD E. On robust nonparametric regression estimation for a functional regressor. *Statistics & Probability Letters*. **78**, 3216, **2008**.
49. ATTOUCH M.K., LAKSACI A., SAÏD E.O. Robust regression for functional time series data. *Journal of the Japan Statistical Society*. **42**, 125, **2013**.
50. STONE C.J., Nonparametric M-regression with free knot splines. *J. Statist. Planning and Inference*, **130**, 183, **2005**.

-
51. RACHDI M., LAKSACI A., DEMONGEOT J., ABDALI A., MADANI F. Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics & Data Analysis*. **73**, 53, **2014**.
52. ATTOUCH M., LAKSACI A., OULD-SAID E. Asymptotic distribution of robust estimator for functional nonparametric models. *Communications in Statistics – Theory and Methods*. **38**, 1317, **2009**.