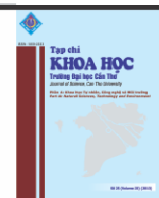




Tạp chí Khoa học Trường Đại học Cần Thơ
website: sj.ctu.edu.vn



PHÂN TÍCH ĐỊNH LƯỢNG TÍNH KHÔNG CHẮC CHẮN CỦA KẾT QUẢ MÔ HÌNH DỰ ĐOÁN LƯỢNG PHÁT THẢI KHÍ DINITƠ MONOXID CỦA KHU VỰC TỰ NHIÊN Ở PHẦN LAN

Trương Ngọc Phương¹

¹ Khoa Môi trường & Tài nguyên Thiên nhiên, Trường Đại học Cần Thơ

Thông tin chung:

Ngày nhận: 07/11/2014

Ngày chấp nhận: 24/04/2015

Title:

Uncertainty analysis of N₂O
emission prediction from
natural areas in Finland

Từ khóa:

Phân tích tính không chắc
chắn, mô hình môi trường,
dinitơ monoxide, khu vực tự
nhiên, Phần Lan

Keywords:

Uncertainty analysis,
environmental model, nitrous
oxide, natural areas, Finland

ABSTRACT

One of the main purposes of environmental modelling is to do spatial and temporal predictions for the values of environmental variables over a study site. The environmental models are often developed based on measured data, (inter) relations between environmental variables, and the nature of underlying environmental processes. In this study, the model for predicting nitrous oxide (N₂O) emission was applied with a case study from natural areas of Finland. The prediction outcomes could be affected by errors in calibration data and model's inputs. Uncertainty quantification in the form of a probability distribution function was used and the Monte Carlo approach was applied for uncertainty propagation analysis. The results show that the prediction of N₂O emission in average ranged from 0.4 to 3.1 kgN/ha/year over the natural areas of Finland. The standard deviation of the prediction values over the study area varied from 0.1 to 5.5 kgN/ha/year. This interval quantified the uncertainty about the prediction outcomes of N₂O from the model used for the natural areas in Finland.

TÓM TẮT

Mục đích của các mô hình môi trường là nhằm dự đoán giá trị của các biến môi trường và xu hướng thay đổi của các giá trị này trong tương lai. Các mô hình môi trường thường được phát triển sử dụng các dữ liệu đo đạc, dựa trên các mối tương quan giữa các biến môi trường và các đặc tính của các quá trình môi trường. Trong nghiên cứu này, mô hình dự đoán lượng phát thải khí dinitơ monoxid (N₂O) được áp dụng trên khu vực đất tự nhiên của Phần Lan. Kết quả dự đoán của mô hình có khả năng không chính xác do sai số từ các dữ liệu đo đạc và dữ liệu nội suy được sử dụng để điều chỉnh mô hình và làm dữ liệu đầu vào cho mô hình. Hàm phân bố xác suất được áp dụng trong nghiên cứu này để định lượng tính không chắc chắn. Phương pháp Monte Carlo được áp dụng để phân tích sự lan truyền của tính không chắc chắn đến kết quả đầu ra của mô hình. Kết quả nghiên cứu cho thấy giá trị dự đoán lượng phát thải N₂O nằm trong khoảng 0,4 – 3,1 kgN/hecta/năm. Độ lệch chuẩn của giá trị dự đoán nằm trong khoảng 0,1 – 5,5 kgN/hecta/năm. Đây là kết quả định lượng tính không chắc chắn của kết quả từ mô hình dự đoán lượng phát thải N₂O cho khu vực tự nhiên của Phần Lan.

1 GIỚI THIỆU

Các nghiên cứu môi trường đã và đang tập trung vào thực hiện việc giám sát, đánh giá và dự báo tác động của các tác nhân môi trường, ví dụ như các tác nhân gây ô nhiễm có ảnh hưởng đến con người và sinh vật. Nghiên cứu về sự phân bố trên bề mặt không gian của các tác nhân môi trường đòi hỏi việc sử dụng các phân tích không gian và mô hình địa lý (Longley *et al.*, 2006). Các kết quả từ những phân tích không gian và các mô hình địa lý thường không phải là các kết quả chính xác do sai số trong các dữ liệu đầu vào của mô hình và các thông số của mô hình. Các sai số trong đo đạc, lấy mẫu, nội suy và thay đổi tỷ lệ xích của dữ liệu là các nguyên nhân chính gây ra sai số cho các dữ liệu đầu vào và các thông số của mô hình (Brown & Heuvelink, 2007). Các sai số này dẫn đến hậu quả là sự không chắc chắn về các kết quả dự đoán; và do đó, là sự không chắc chắn về các tác động tiêu cực của các tác nhân môi trường lên con người và sinh vật. Sự không chắc chắn này cần được định lượng và thông tin cho các nhà ra quyết định để họ có thể thực hiện một cách có hiệu quả việc lựa chọn hoặc ra các quyết định có liên quan đến các chiến lược giảm nhẹ hoặc xử lý mà có thể có tác động lớn đến con người và sinh vật.

Một số phương pháp đã được nghiên cứu để sử dụng cho việc định lượng tính không chắc chắn trong các phân tích không gian và mô hình địa lý tương ứng với mục đích của từng nghiên cứu, từng loại dữ liệu không gian và các đặc trưng thuộc tính của chúng (Attah-Okine và Ayyub, 2005; Longley *et al.*, 2006). Ví dụ, Linkov và Burmistrov (2003) nghiên cứu tính không chắc chắn của mô hình, bao gồm việc xây dựng vấn đề, việc thực hiện mô hình và lựa chọn các thông số, bằng cách so sánh các kết quả từ các mô hình khác nhau được phát triển cho cùng một tác nhân môi trường, đó là nồng độ phóng xạ hạt nhân. Heuvelink *et al.* (2007) đề nghị phương pháp xác suất sử dụng hàm phân bố xác suất (probability density function - pdf) để biểu

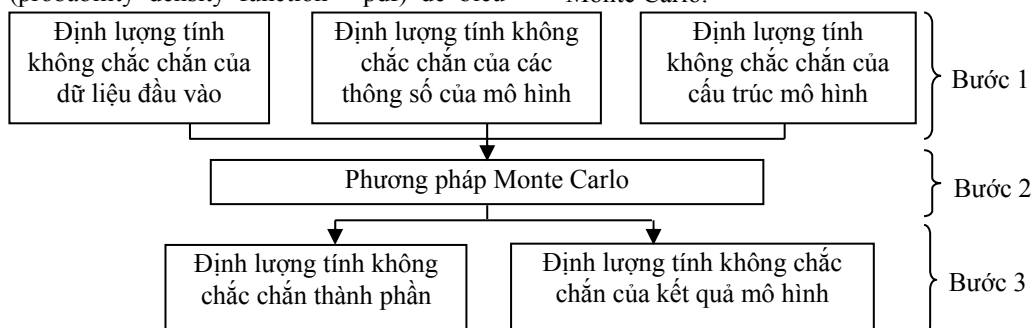
diễn tính không chắc chắn về giá trị các biến môi trường. Gần đây, các ý kiến của các chuyên gia đã được sử dụng để định lượng tính không chắc chắn của sự biến đổi trên bề mặt không gian của các tính chất của đất (Trương Ngọc Phương và Heuvelink, 2013).

Trong nghiên cứu này, phương pháp xác suất được áp dụng, trong đó tính không chắc chắn được định lượng bằng một khoảng giá trị xác suất, mà bất kỳ sự lặp lại nào của việc ước tính sẽ cho ra một kết quả mới mà chủ yếu là nằm trong khoảng giá trị đó. Các nguồn khác nhau gây nên sự không chắc chắn có thể được phân loại thành bốn nguồn chính: (1) Sự không chắc chắn về dữ liệu đầu vào; (2) Sự không chắc chắn về tham số mô hình; (3) Sự không chắc chắn về cấu trúc mô hình; và (4) Sự không chắc chắn về giải pháp kỹ thuật (Heuvelink, 1998). Nghiên cứu này cũng trả lời câu hỏi làm thế nào định lượng sự lan truyền của tính không chắc chắn từ dữ liệu đầu vào và các tham số của mô hình đến kết quả đầu ra của mô hình áp dụng phương pháp Monte Carlo.

2 PHƯƠNG PHÁP MONTE CARLO

Nguyên tắc của phương pháp Monte Carlo là liên tục lặp lại việc tính toán kết quả của mô hình với các giá trị đầu vào được lấy mẫu ngẫu nhiên từ các phân bố xác suất của chúng. Các yếu tố đầu vào có thể là các dữ liệu đầu vào, các thông số của mô hình, các sai số trong cấu trúc của mô hình. Các kết quả lặp lại của mô hình sẽ tạo thành một tập mẫu ngẫu nhiên của phân bố xác suất của kết quả đầu ra. Phân tích thống kê mô tả tập mẫu này bằng cách tính giá trị trung bình và độ lệch chuẩn (std) cung cấp giá trị đại diện cho mức độ không chắc chắn về kết quả đầu ra của mô hình với điều kiện là kích thước mẫu đủ lớn.

Hình 1 trình bày ba bước cơ bản được thực hiện trong việc phân tích tính không chắc chắn và sự lan truyền tính không chắc chắn áp dụng phương pháp Monte Carlo.



Hình 1: Các bước trong phương pháp phân tích tính không chắc chắn

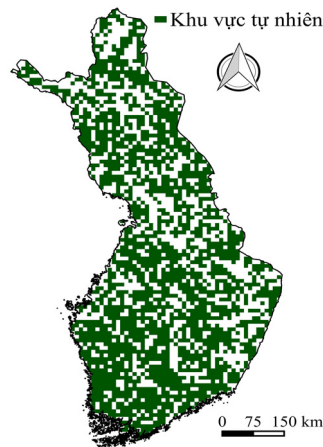
3 ỨNG DỤNG TRONG DỰ ĐOÁN LƯỢNG PHÁT THẢI KHÍ N₂O Ở CÁC KHU VỰC TỰ NHIÊN CỦA PHÂN LAN

Để minh họa cho việc áp dụng phương pháp Monte Carlo, trường hợp nghiên cứu sử dụng mô hình hồi quy tuyến tính để dự đoán lượng phát xạ của chất gây ô nhiễm không khí dinitơ monooxit từ đất (N₂O - kgN/ha/năm) ở khu vực tự nhiên của Phần Lan (Hình 2) đã được chọn. Đây là một tiêu mô hình của mô hình INTEGRATOR ở Châu Âu (Reis *et al.*, 2007). Mô hình hồi quy tuyến tính (Phương trình 1) tính toán lượng phát thải dinitơ monooxit (N₂O) từ đất như là một hàm của nhiều yếu tố môi trường, bao gồm cả các biến khí hậu (lượng mưa hàng tháng, nhiệt độ tối thiểu), pH, hàm lượng carbon hữu cơ, hàm lượng nitơ lắng đọng và các loại thực vật ôn đới.

$$\text{Log(N}_2\text{O)} = \beta_0 + \beta_1 * P + \beta_2 * (\text{Tháng có } T < 0^\circ\text{C}) + \beta_3 * \text{Log(N}_{\text{ld}}) + \beta_4 * \text{pH} + \beta_5 * \text{C}_{\text{hữu cơ}} + \beta_6 * (\text{thực vật: cây ôn đới}) \quad (1)$$

Trong đó: β_i là các hệ số hồi quy; các biến khí hậu gồm có: P là lượng mưa trung bình hàng tháng trong khoảng thời gian đo (mm) và số tháng có nhiệt độ tối thiểu $T < 0^\circ\text{C}$ trong khoảng thời gian đo; dữ liệu về đất được sử dụng từ cơ sở dữ liệu WISE & SPADE, lấy giá trị trung bình trên lớp đất ở độ sâu 0-20 cm: độ pH-H₂O (tất cả các giá trị pH khác nhau được chuyển đổi sang pH-H₂O sử dụng hàm hồi quy) và hàm lượng carbon hữu cơ (g/kg đất); giá trị nitơ lắng đọng được lấy từ nguồn dữ liệu của Chương trình giám sát và đánh giá ở Châu Âu (EMEP): N_{ld} là lượng nitơ lắng đọng (kgN/ha/năm); thực vật ôn đới bao gồm: rừng rụng lá, rừng lá kim, thảm thực vật thấp (bao gồm cây thạch nam và cỏ) và các loại đất rừng. Nội suy kriging hồi quy (Hengl *et al.*, 2007) được sử dụng để lập bản đồ và định lượng tính không chắc chắn cho các biến đầu vào của mô hình, đó là pH và OC.

Nội suy Kriging đơn giản được sử dụng cho giá trị hồi quy thẳng dư. Sự không chắc chắn về các hệ số hồi quy được định lượng bằng hàm phân bố xác suất. Phân tích sự lan truyền tính không chắc chắn được thực hiện thông qua phương pháp Monte Carlo, trong đó sự đóng góp của từng thông số đầu vào cũng được tính toán.

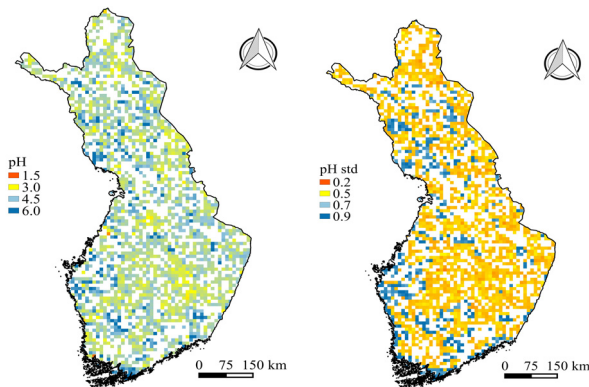


Hình 2: Khu vực tự nhiên ở Phần Lan

4 KẾT QUẢ VÀ THẢO LUẬN

4.1 Tính không chắc chắn về pH

Hình 3 trình bày bản đồ mô phỏng pH (Hình 3A) và giá trị độ lệch chuẩn của kết quả mô phỏng giá trị pH (Hình 3B) trên khu vực nghiên cứu. Giá trị mô phỏng pH trên toàn bộ khu vực tự nhiên ở Phần Lan khá thấp ở mức axit, nằm trong khoảng từ 1.5 đến 6. Ở đây không có xu hướng thay đổi giá trị pH theo các chiều hướng không gian xác định. Độ lệch chuẩn của giá trị pH cung cấp thông tin định lượng của mức độ không chắc chắn về các giá trị mô phỏng của pH khi sử dụng mô hình nội suy Kriging hồi quy và các giá trị pH tại các vị trí được đo đạc. Giá trị định lượng tính không chắc chắn này nằm trong khoảng từ 0.2 đến 0.9.



Hình 3: Kết quả mô phỏng giá trị pH (A) và độ lệch chuẩn (B)

4.2 Tính không chắc chắn của OC

Hình 4 trình bày bản đồ mô phỏng OC (Hình 4A) và giá trị độ lệch chuẩn của kết quả mô phỏng giá trị OC (Hình 4B). Giá trị mô phỏng OC thay đổi trong khoảng giá trị rộng 0 - 100 g OC/kg đất trên toàn bộ khu vực tự nhiên ở Phần Lan với mức độ phân tán 0 - 31 g OC/kg đất. Hình 4 cũng cho thấy ở những khu vực có giá trị OC cao thì độ lệch chuẩn cũng cao; điều này có nghĩa là mức độ không chắc chắn về các giá trị mô phỏng này cũng cao.

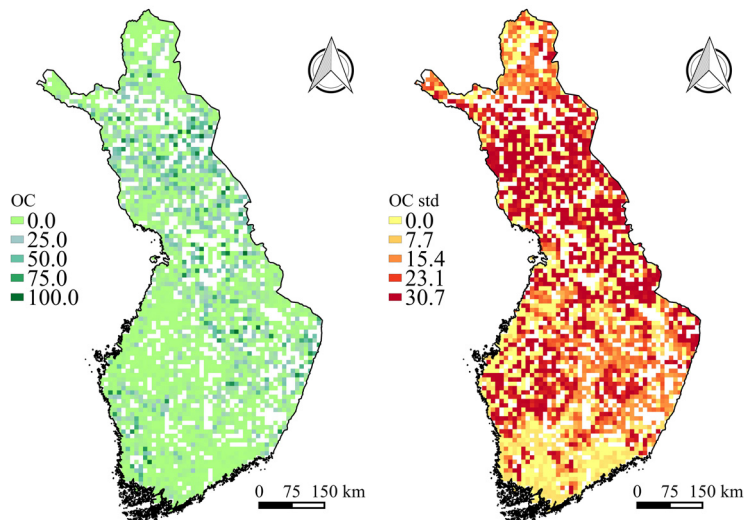
4.3 Tính không chắc chắn về các thông số của mô hình

Phương trình 2 trình bày kết quả tính toán các hệ số cho mô hình hồi quy tuyến tính mà mô tả những mối quan hệ giữa giá trị LogN_2O và các

biến độc lập (như được trình bày ở Phần 3). Mô hình này đã được hiệu chỉnh bởi bộ dữ liệu của toàn bộ các nước thành viên của Cộng đồng chung Châu Âu.

$$\text{Log}(\text{N}_2\text{O}) = 0.133 - 0.007 \cdot P + 1.537 \cdot (\text{Tháng có } T < 0^\circ\text{C}) + 0.497 \cdot \text{Log}(\text{N}_{\text{ld}}) - 0.137 \cdot \text{pH} - 0.003 \cdot C_{\text{hữu cơ}} + 0.056 \cdot (\text{thực vật: cây ôn đới}), R^2 \text{ điều chỉnh} = 0.2 \quad (2)$$

Bảng 1 cho thấy các giá trị ước tính của tất cả các hệ số hồi quy và độ lệch chuẩn tương ứng với từng hệ số. Độ lệch chuẩn này là giá trị định lượng tính không chắc chắn về các giá trị ước tính của các hệ số hồi quy. Trong số tất cả các hệ số thì tham số tự do và biến thực vật: cây ôn đới có độ lệch chuẩn cao so với giá trị ước tính của hai hệ số này.



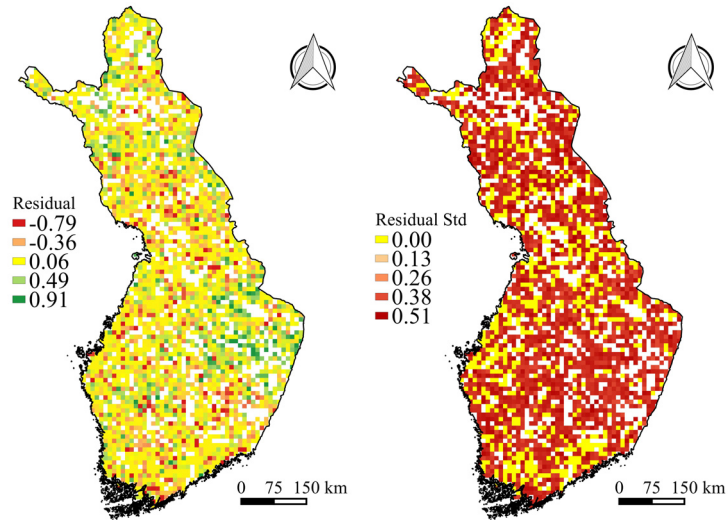
Hình 4: Kết quả mô phỏng OC (A) và độ lệch chuẩn (B)

Bảng 1: Giá trị ước lượng của các hệ số hồi quy và mức độ phân tán của các giá trị này

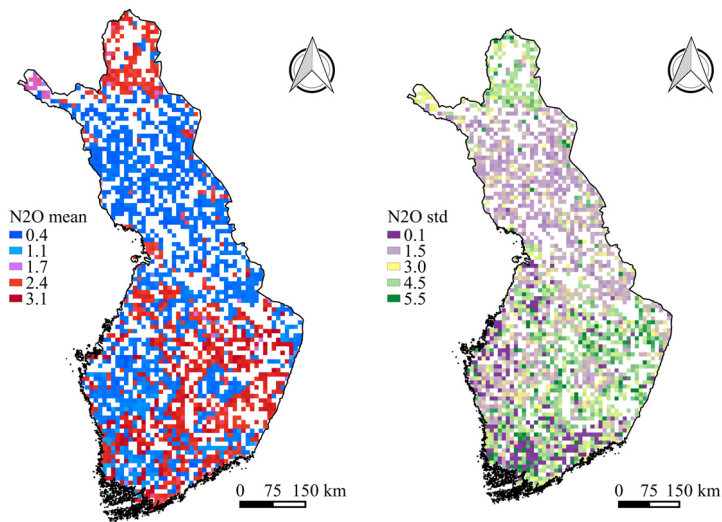
Hệ số	Giá trị ước tính	Độ lệch chuẩn	Hệ số biến phân
Tham số tự do	0.133	0.563	4.215
Lượng mưa	-0.007	0.003	0.429
Tháng có $T < 0^\circ\text{C}$	1.537	0.397	0.258
LogN_{ld}	0.497	0.290	0.584
pH	-0.137	0.062	0.448
OC	-0.003	0.001	0.292
Cây ôn đới	0.056	0.107	1.908

4.4 Tính không chắc chắn về kết quả dự đoán của mô hình

Bản đồ giá trị độ lệch chuẩn của kết quả mô phỏng N_2O (Hình 6B) cho thấy mức độ của sự phân tán của giá trị dự đoán trung bình của N_2O (Hình 6A) là rất cao. Như vậy, sự không chắc chắn về kết quả dự đoán là rất lớn ở một số khu vực. Trên khu vực nghiên cứu, giá trị dự đoán trung bình của N_2O biến đổi trong khoảng 0,5 đến 3 kgN/hecta/năm, trong khi đó giá trị độ lệch chuẩn biến đổi trong khoảng 0,1 – 5,5 kgN/hecta/năm.



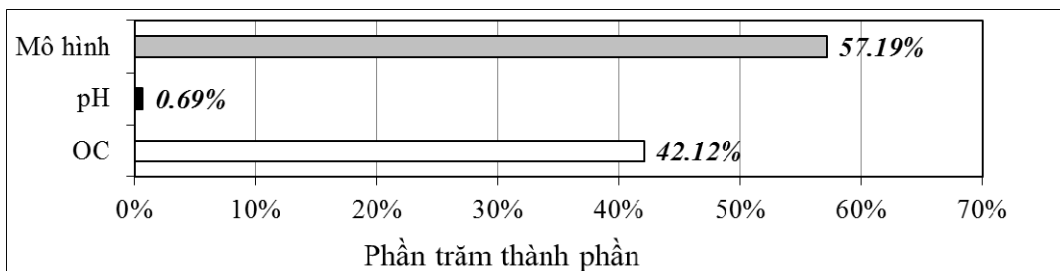
Hình 5: Kết quả mô phỏng giá trị hồi quy thẳng dư (A) và độ lệch chuẩn (B)



Hình 6: Kết quả trung bình lượng phát thải khí N₂O (A) và độ lệch chuẩn (B)

Trong Hình 7, ta có thể thấy tính không chắc chắn của mô hình hồi quy có tỷ lệ đóng góp lớn nhất vào kết quả định lượng tính không chắc chắn của kết quả đầu ra của mô hình. Đóng góp của mô hình hồi quy là khoảng 57,19%. Tỷ lệ đóng góp

của OC cũng cao tương đương. pH có sự đóng góp nhỏ nhất, chưa đến 1% tổng giá trị định lượng tính không chắc chắn về kết quả dự đoán N₂O sử dụng mô hình như phương trình 2.



Hình 7: Tính không chắc chắn thành phần

5 KẾT LUẬN

Thông qua nghiên cứu này, phương pháp định lượng tính không chắc chắn về các dữ liệu đầu vào và các thông số của mô hình đã được minh họa; trong đó, tính không chắc chắn được định lượng bằng độ lệch chuẩn của các kết quả mô phỏng. Phương pháp Monte Carlo đã được sử dụng để phân tích sự lan truyền tính không chắc chắn thông qua mô hình từ các yếu tố đầu vào của mô hình. Đây là phương pháp rất đơn giản về mặt nguyên tắc nhưng lại rất hiệu quả.

Các kết quả chính của nghiên cứu này cho thấy: (1) Các mô hình thống kê đã được sử dụng là phù hợp cho việc định lượng tính không chắc chắn về các dữ liệu đầu vào và các thông số của mô hình; (2) Sự không chắc chắn về ước lượng giá trị phát thải N_2O được biểu diễn bởi độ lệch chuẩn của kết quả dự đoán; giá trị định lượng này biến đổi trên khu vực nghiên cứu; (3) Sự không chắc chắn về mô hình hồi quy tuyến tính (tức là sự không chắc chắn về tham số mô hình) là nguồn lan truyền quan trọng nhất đến kết quả định lượng sự không chắc chắn về kết quả dự đoán lượng phát thải khí N_2O từ đất ở Phần Lan.

TÀI LIỆU THAM KHẢO

1. Attoh-Okine, N.O., Ayyub, B.M., 2005. Applied Research in Uncertainty Modeling and Analysis. Springer US, Boston, MA, 545 pp.
2. Brown, J.D., Heuvelink, G.B.M., 2007. The Data Uncertainty Engine (DUE): A software tool for assessing and simulating uncertain environmental variables. Computers & Geosciences 33, 172–190, doi:10.1016/j.cageo.2006.06.015.
3. Hengl, T., Heuvelink, G.B.M., David, G.R., 2007. About regression-kriging: From equations to case studies. Computers & Geosciences 33, 1301–1315.
4. Heuvelink, G.B.M., 1998. Error propagation in environmental modelling with GIS. London, Taylor & Francis, 150 pp.
5. Heuvelink, G.B.M., Brown, J.D., van Loon, E.E., 2007. A probabilistic framework for representing and simulating uncertain environmental variables. International Journal of Geographical Information Science 21, 497–513.
6. Linkov, I., Burmistrov, D., 2003. Model Uncertainty and Choices Made by Modelers: Lessons Learned from the International Atomic Energy Agency Model Intercomparisons. Risk Analysis 23, 1297–1308, doi:10.1111/j.0272-4332.2003.00402.x.
7. Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., 2005. Geographic Information systems and science, 2nd edition. New York: Wiley, 536 pp, Chapter 6, 9.
8. Reis, S., Sutton, M. A., de Vries, W., Kros, H., 2007. Modelling multidimensional problems – The case of integrated nitrogen management. In: Oxley, L. and Kulasiri, D. (eds). MODSIM 2007, International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2007, 2083–2089.
9. Truong Ngoc Phuong., Heuvelink, G.B.M., 2013. Uncertainty quantification of soil property maps with statistical expert elicitation. Geoderma 202–203, 142–152.