Review comments on the manuscript "**Subsoil Reconstruction in Geostatistics beyond Kriging: A Case Study in Veneto (NE Italy)**" submitted by Paolo Fabbri, Carlo Gaetan, Luca Sartore, and Nico Dalla Libera

Reviewer: **Weidong Li**
Date reviewer invited by *Journal of Hydrology*: 13 March 2019.
Date review submitted: 11 April 2019.

I carefully read the manuscript and find that in the whole manuscript the authors made lots of improper statements and they completely followed the jokes and misinterpretations of Allard's to both the MCRF model proposed by Li (2007a) and the BME model proposed by Bogaert (2002) (given the fact that Allard was a trained geostatistician in kriging and had worked for a long time in geostatistical application before 2011, I would rather believe Allard was making jokes in his research and articles than believe he misunderstood). The study has no value, except for messing the geostatistics field and causing more confusions. So I cannot support it. However, I wrote a lot of comments (much more than a normal manuscript review) to the text of the manuscript, and I sincerely hope the authors could realize their biases to the MCRF model and their misunderstandings to geostatistics. Below are my reasons and comments to the manuscript.

GENERAL COMMENTS:

The authors completely followed Allard's biases and misinterpretations (Allard et al. 2009, 2011) to the MCRF model proposed by Li (2007a, submitted in 2004). Sartore's articles (2013, 2016) and his spMC package with wrong interpretations about the MCRF model, Bogaert's BME model (Bogaert 2002) and Allard's MCP method have caused strong confusions and troubles to others in recent five years. Allard might misunderstand, but more possibly, he was making jokes, to both the MCRF approach and the BME approach. Li and Zhang (2012) commented Allard's article and pointed out some of his misunderstandings. It is not difficult to see that Allard made the following improper interpretations and claims in his article (Allard et al. 2011) and in his reply letter (Allard 2012) to the comment letter of Li and Zhang (2012):

(1) Allard (Allard et al. 2011) distorted the MCRF model (Li 2007a) by describing it as the CMC model of Elfeki and Dekking (2001), that is, a combination of two Markov chains, while the MCRF model as a solution solved the small class underestimation problem of the CMC model and clearly proved it.
(2) Allard distorted the spatial conditional independence assumption for nearest neighbors suggested by Li into an independence assumption of two Markov chains used in Elfeki and Dekking (2001).
(3) Allard distorted the idea of sequential Bayesian updating on nearest data within a neighborhood (that was used in deriving the MCRF model by Li 2007a) to be a sequential simulation algorithm like the sequential indicator simulation or sequential Gaussian simulation algorithms. However, sequential simulation is a strategy of performing stochastic spatial simulation of a whole random field using any spatial probability model.
(4) Allard thought that the conditional independence assumption used by Li (2007a), which may be simply expressed as P(B|A,C)=P(B|A) given A, was wrong, and that what he used, which may be simply expressed as P(B,C|A)=P(B|A)P(C|A) given A, was correct. However, these two expressions are actually equivalent. In fact, conditional independence assumption was thought to be wrong and invalid in geostatistics previously (Journel 2002). It was Li (2007a) who found its rationale from a property of Pickard random fields and then generalized and used it to nearest spatial data. It was not simply common knowledge in geostatistics before Li presented it.
(5) Allard used Li's ideas (spatial sequential Bayesian updating and spatial conditional independence assumption on nearest data) for deriving the MCRF model to derive a model that is actually the same

as the MCRF model, but claimed it was a simplified BME model and his MCP method. However, entropy maximization (used by Bogaert for deriving his BME model) cannot lead to the MCRF model, and Li's ideas also cannot lead to Bogaert's BME model.

(6) Allard ignored the CMC model and also ignored the fact that the MCRF model was obtained as a solution to the theoretical defect of the CMC model. Li first used the CMC model in 1999 and then explored the reasons of its defects, and then corrected its defects - The CMC model consists of two 1-D Markov chains and has the small class underestimation problem; Li used a single Markov chain (locally-conditioned on nearest data) to overcome the problem and then generalized the new idea to a geostatistical approach. Li was neither familiar with Christakos' BME approach previously nor followed it. Bogaert's BME model was even not known in mathematically geology, geography and soil science when Li derived his MCRF model, and it also had nothing to do with Li's ideas in deriving the MCRF model.

(7) Allard claimed he derived a simplified BME model using the conditional independence assumption. However, both Christakos's BME framework and Bogaert's BME model contain no multi-point probability terms (let alone multi-point likelihood terms) and thus have no place to apply the conditional independence assumption. Except for a simplified model formula, the MCRF model has its full model formula, which is also not Bogaert's BME model.

(8) Allard claimed that his MCP method (i.e., the MCRF model) was known as the "Naive Bayes classifier" used in machine learning and classification literature (Didn't that mean he plagiarized the naïve Bayes classifier?). However, Naïve Bayes classifier is a nonspatial model which classifies a dependent variable based on some predictor variables.

(9) Allard claimed that the spatial conditional independence assumption for nearest neighbors used by Li was known as the "Naive Bayes assumption". However, the latter is nonspatial and assumes that the predictor variables are conditionally independent given the dependent variable (this is similar to the independence assumption of predictor variables in multiple linear regression).

(10) The MCRF model was generalized in Li (2007a) with no limit on the number of nearest data, but Allard claimed that Li and Zhang used four nearest data in their simulation algorithms and it was his MCP method that generalized it and could use more nearest neighbors. This is a wrong excuse. Based on such a logic, a kriging model with different numbers of nearest neighbors could be claimed as different models, and cokriging could never use more than one ancillary variables because it was traditionally implemented with only one ancillary variable, isn't it?

(11) Allard accused that Li and Zhang claimed they were able to build explicitly a random field but he did not make that claim, as if Li and Zhang made a big mistake (Allard was accusing the model name used by Li - "Markov chain random field"). However, the truth is that every theoretically-sound spatial probability model actually expresses a random field.

(12) Allard claimed that the maximum entropy approach does not need stationarity assumptions at any stage. However, without the stationarity assumption, how could one estimate the (global) parameters from sample data? Stationarity assumption is the basic assumption for any spatial model. Nonstationary spatial models still need stationary assumption in local areas.

(13) Allard claimed that Markov chain Monte-Carlo was the real spatial Markov chain but the MCRF model was not a spatial Markov chain. However, 1-D Markov chain models have been widely used to simulate lithofacies sequences since 1960s. Although not practical in most situations, the CMC model has also been used to simulate spatial patterns. Weren't they all spatial? On the contrary, Markov chain Monte-Carlo is just an iterative algorithm that uses some properties of Markov chains, not regarded as a spatial Markov chain in common sense, especially in mathematical geology.

(14) Allard claimed that Li defined the MCRF theory on the random path. However, random path is just one way to implement the MCRF model in stochastic simulation. Kriging was also used in random path simulations, was kriging also defined on the random path? Since Markov chain is a stochastic process and MCRF is a typical spatial stochastic process, there seemed nothing wrong to intuitively describe a MCRF from the perspective of random motion.

(15) D'Or and Allard (2008) even called the MCP method (i.e., the MCRF model, which can be written in different form with the detailed balance principle) a multinomial regression in a conference presentation, while it has nothing to do with regression.

……

Allard misinterpreted the MCRF model almost on every point and from every perspective that he could imagine to cause confusions. If the MCRF model was indeed simplified from Bogaert's BME model and known as the Naïve Bayes classifier widely used in machine learning and classification, then what Allard should do if he had problems? What he should do was to tell readers his personal understanding or judgement with evidence rather than to claim a MCP method of his. In geostatistics, it is known that the kriging approach and the BME approach have disagreements. As a trained geostatistician in kriging, Allard suddenly meddled in the MCRF model and dragged it to the BME approach while completely ignoring the existence of the CMC model, whose defects were exactly the problems that the MCRF model aimed to solve. Was it reasonable? Given the irrational and confusing interpretations Allard made to the MCRF model, I believe the authors of this manuscript were completely confused.

In fact, the conditional independence assumption was mistakenly regarded as a wrong and invalid assumption in geostatistics previously by Journel (2002). It was Li who found its rationale from a property of Pickard random fields, and used it to simplify the full formula (with multi-point statistics) of the MCRF model. This point was clearly presented in Li (2007a) and Li and Zhang (2008). So the spatial conditional independence assumption for nearest spatial data was not common knowledge previously in geostatistics.

Sartore (2013, 2016) followed the wrong interpretations of Allard's to the MCRF model and aimed to solidify Allard's claims and mistakes. The authors of this manuscript continuously did the same. This is very improper, misleading to readers, and also disrespective to the efforts of others. In the whole manuscript, the authors did not mention the CMC model of Elfeki and Dekking (2001) and also did not mention the MCRF model. But they did cite Li (2007a) and Li and Zhang (2006) with misleading statements. However, they messed Li (2007a) that proposed the MCRF model and Li and Zhang (2006) that combined the CMC model with transiograms, and give them a wrong interpretation.

It should be noticed that, due to the quantitative relationship $p_{ij}(\mathbf{h}) = p_{ji}(\mathbf{h})p_j/p_i$, which is called "detailed balance" in Markov chain theory, the simplified formula of the MCRF model based on the spatial conditional independence assumption can be written in different forms:

$$\frac{p_{l_1k}(\mathbf{h}_1)\prod_{i=2}^{m}p_{kl_i}(\mathbf{h}_i)}{\sum_{f=1}^{n}[p_{l_1f}(\mathbf{h}_1)\prod_{i=2}^{m}p_{fl_i}(\mathbf{h}_i)]} = \frac{p_k\prod_{i=1}^{m}p_{kl_i}(\mathbf{h}_i)}{\sum_{f=1}^{n}[p_f\prod_{i=1}^{m}p_{fl_i}(\mathbf{h}_i)]} = \frac{p_k^{1-m}\prod_{i=1}^{m}p_{l_ik}(\mathbf{h}_i)}{\sum_{f=1}^{n}[p_f^{1-m}\prod_{i=1}^{m}p_{l_if}(\mathbf{h}_i)]},$$

which are mathematically equivalent. No matter which form Allard used, it is still the MCRF model.

The following website provided the information about the MCRF approach and its development: http://gis.geog.uconn.edu/weidong/Markov_chain_spatial_statistics.htm. (Li did his research in a very hard situation caused by the troublemaking of some people since the beginning of 2004).

SPECIFIC COMMENTS:

Line 1: Title: "Beyond the kriging"? Do the authors know that multiple-point geostatistics proposed and developed by Stanford University is also beyond the kriging? I am afraid the inverse distance weighting (IDW) is also beyond kriging.

Line 40-41: "Carle and Fogg (1996) introduced the concept of transition probability in geostatistics and their graphical representations called transiograms". First, Carle and Fogg (1996) is not the first article that introduced the concept of transition probability and their graphical representations in mathematical geology. Transition probability was used in mathematical geology as early as 1930s. Transition probability-lag diagrams appeared in earlier articles, such as Schwarzacher (1969) and Luo (1993). Second, Carle and Fogg did not call that as transiograms. It was Li (2007b) who called it "transiogram" and provided a concept system for it by following partially the variogram concept for the convenience of its use and description as a spatial measure in Markov chain geostatistics. Carle and Fogg's major theoretical contributions in spatial-lag transition probability mainly lie with (1) the quantitative relationship between transition probability and indicator covariance/variogram for reformulating indicator kriging equations and (2) the transition rate method (i.e., continuous-lag Markov chain model) for corregionalization of multiple classes (so that indicator cokriging can be used to simulate multiple lithofacies with interclass correlations).

Line 43-45: No such thing. Schwarzacher calculated transition probability-lag diagrams (curves) using a one-step transition probability matrix, and that kind of transiograms were called "idealized transiograms" in Li (2007b).

Line 47-49: This (incorporating mean thickness and proportion knowledge) was done by Carle and Fogg (1997) through transition rates for lateral direction transiograms, not for vertical direction transiograms. This is an advantage for subsurface modeling where sample data are normally insufficient for estimating meaningful lateral transiograms, but it becomes a disadvantage in other situations where sample data can provide meaningful experimental transiograms.

Line 52-53: Transiogram is just a spatial measure for providing transition probability parameters. It does not make spatial prediction by itself. Spatial prediction or simulation needs a spatial statistical model to estimate the local conditional probability distribution at any unsampled location. Carle and Fogg used reformulated indicator cokriging model.

Line 73-75: Misleading citations and interpretations! None of these two cited articles (Li 2007a, Li and Zhang 2006) used a random-path Markov chain algorithm, and the simulation cases were also not lithologies.

Li and Zhang (2006) implemented the CMC model of Elfeki and Dekking (2001) in a modified form (the TMC model used two CMCs to advance alternately to overcome the layer/patch inclination problem) using transiograms (with another multi-class corregionalization method) so that it can be conditioned on regular point sample data. Both the CMC model and the CMC-based TMC model used fixed paths. This was not surprising because Markov mesh models also used asymmetric neighborhoods and fixed paths for simulation.

However, Li (2007a) proposed the MCRF model, and it was the first published article about the MCRF model. The MCRF model is a new geostatistical model based on the ideas of sequential Bayesian updating on nearest data and spatial conditional independence assumption of nearest data in a neighborhood. It has its full model formula and simplified model formula for local conditional probability distribution. The MCRF model initially aimed to solve the small class underestimation problem of the CMC model by using a locally-conditioned Markov chain, but it was generalized and was then called "Markov chain random field".

The MCRF model manuscripts (Li and Zhang 2008, Li 2007a) were stuck there for a long time due to the misunderstandings about whether the CMC model was defective (somebody guessed that the CMC model was perfect, although even Elfeki and Dekking (2001) clearly showed the small class underestimation problem in their simulation cases). In order to publish both the MCRF model and the transiogram methods (Li proposed two corregionalization methods for MCRF simulation of multiple classes using transiograms), Li and Zhang had to use transiogram methods to the CMC model and first publish them (see Li and Zhang 2005, 2006). The MCRF model manuscripts (Li and Zhang 2008, Li 2007b) were finally accepted for publication in 2006 after Li and Zhang developed the random-path MCRF sequential simulation algorithm (see Li and Zhang 2007). Li and Zhang totally presented three simulation algorithms – two fixed paths (alternate advancing and middle insertion) and one random path for the MCRF model.

The simulation paths are not necessarily attached with a spatial model, and a generalized spatial model is also not necessarily applicable to only one path. Different spatial models can use the same path for stochastic simulation. Pease do not make irrational interpretation based on your own imagination and bias, and please also do not simply follow the bias from others.

Line 79-82: The authors made wrong and misleading citations here when talking about Bogaert's BME model (Bogaert 2002). The articles of Bogaert and D'Or (2002) and D'Or et al. (2001) were studies of continuous spatial variables (e.g., soil textural fractions) using Christakos' BME method, and none of them mentioned Bogaert's BME model. Bogaert's BME model was neither known in geostatistics nor in soil science when Li developed the MCRF model. It seemed that the authors never read or browsed these articles when they cited them to support their view points and claims.

Based on my understanding, Christakos's BME approach (Christakos 1990) is a generalized framework for spatial prediction with incorporation of soft information and it normally used simple kriging as its fundamental model for neighborhood modeling and used the Bayesian principle for incorporating soft information. According to Christakos (1990), he derived a linear geostatistical model using entropy maximization and it coincided with simple kriging. Entropy maximization is the basic requirement of Jaynes' maximum entropy principle and Christakos' BME approach. A statistical principle is not equal to a geostatistical model, especially not equal to a fundamental model for modeling on a neighborhood. As presented in Bogaert (2002), Bogaert derived a nonlinear maximum entropy (ME) model using entropy maximization for categorical variables and fit it into the BME framework of Christakos'.

Line 83-86: This is misunderstanding and Allard's joke. Allard (Allard et al 2011) used the ideas of Li's (conditional independence assumption of nearest neighbors and sequential Bayesian updating on nearest neighbors, both were presented in Li 2007a) for deriving the MCRF model, and what Allard derived was still the MCRF model but was written with a little difference. Allard misunderstood the MCRF model and messed both the MCRF model and the BME approach together. What was "Markovian type categorical prediction"? That cannot be a proper model name. Did the authors read the article of Allard et al. (2011) and compared it with Li (2007a) and Bogaert (2002)? If you did, you should be able to find it has many misunderstandings.

Line 87-89: Who proposed the spatial Bayesian updating formulation on nearest data of a neighborhood? It was Li in the MCRF model (Li 2007a, Li et al. 2015). It was Li (2007a) that presented the simplified MCRF model based on the spatial conditional independence assumption of nearest neighbors. Bayesian updating was previously only used for incorporating soft or ancillary information or for parameter estimation in spatial statistics. Huang (Huang et al. 2016) imagined a nonspatial model (for integrating expert opinion) into a transition probability-based indicator geostatistics. Such an imagination was not only irrational, but also had no value, because transition probability-based indicator geostatistics (i.e., TPROGS) had already existed (see Carle and Fogg 1996).

Line 90-91: This manuscript and other earlier articles of Sartore's (Sartore 2013, Sartore et al. 2016) and those using Sartore's spMC package are misleading due to the fact that Sartore followed the mistaken article of Allard et al. (2011) and its wrong interpretations about the MCRF model and geostatistics to explain his spMC package.

Line 109-121: The authors defined the transition probability to be a joint probability in equation (1). This is wrong. A transition probability and a joint probability are different things.

Schwarzacher (1969) and Luo (1993) presented transiograms (transition probability-lag curves) more early. Carle and Fogg (1997) found a multi-class corregionalization method (transiogram joint modeling) using the transition rate to implement the transition probability-reformulated indicator geostatistics, because geostatistical models that consider spatial cross correlations of multiple classes need a corregionalization method to obtain proper parameters (either indicator variogram models or transiogram models). Although the transition rate method is well theorized, it is over-idealized and not adjustable through setting parameters (e.g., sill and range), and thus cannot effectively fit the shapes of complex experimental transiograms. Unfortunately, Carle stopped his research after he developed the TPROGS software for 3D simulation of lithofacies and did not go further, probably because indicator kriging is not so flexible to other styles of transiogram models.

Spatial transition probability is the indispensable measure of Markov chain spatial modeling. While working on Markov chain spatial modeling, Li followed pioneer studies to fully establish the transiogram as a spatial correlation measure for categorical spatial data. Li proposed the transiogram concept system, developed two transiogram joint modeling (or corregionalization) methods for practical use, and provided a more accurate definition of transiograms. For example, Carle's definition of continuous-lag Markov chain (actually transition probability) (i.e., transiogram) was related to joint probability and indicator data, which made the transiogram lose the legitimacy of being unidirectional and unnecessarily attached to indicator data. Li recognized Carle's work (see Li 2007b).

The name of "transiogram" is apparently a good choice to use for transition probability-lag diagrams, and the previous names such as continuous-lag Markov chains (used by Carle) or Markov diagrams (used by Luo) were not convenient. That is probably why the authors also used the name "transiogram".

Line 125-136: It seems that the authors lack knowledge in geostatistics. That is not surprising given the facts that they did not work on geostatistics or Markov chains previously and Sartore started his research by following the mistakes of Allard's. Allard's articles are full of misunderstandings and misinterpretations to both the MCRF model and the whole spatial statistics. Following those can only confuse the authors themselves and also confuse readers.

The estimation of transition rate matrix and the choice of the conditional probability approximation are not two problems at the same level. The estimation of transition rate matrix is just an issue for one of the corregionalization methods of multiple classes – the transition rate method (or continuous-lag Markov chain method) proposed by Carle and Fogg (1997) for implementing the TPROGS. The choice of the conditional probability approximation is the whole geostatistics, the whole spatial statistics and even the whole statistical modeling. Any statistical models, including geostatistics, aim to approximate the conditional probability distribution for a variable to be estimated. Indicator kriging, MCRF, and MRF models are all spatial models for approximating the local conditional probability distribution of a categorical random variable. They all have their different formulas and different ways.

Line 137-138: "some authors more recently proposed approximations based on path methods (Li, 2007a)". I am afraid such a statement is an insult to the MCRF model and Li and his coauthors? Who told you the MCRF model was "path-based methods". What is a path method? Li's MCRF model is a new generalized geostatistical model. You followed the wrong attitude of Allard. Fixed paths were widely used before random path was adopted in stochastic simulation after 1980s. The MCRF model has its systematic theory. Fixed path was even used in extending 1-D autoregressive model into multiple dimensions (see Turner and Sharp 1994; Sharp and Turner 1999).

Line 139-140: "Bogaert (2002) introduced a Bayesian procedure that exploits the maximum entropy principle". Misunderstanding and misleading. Bogaert (2002) used the entropy maximization method to derive a ME model, as clearly presented in Bogaert (2002, p. 431-432, equation 27). As to how he derived it and whether his model has some relation with other existing models, he did not explain clearly. Bayesian principle was used conceptually for fitting the ME model into the BME framework of Christakos. In real applications, Bayesian principle was often used to incorporate soft/ancillary data.

Line 140-147: "Allard et al. (2011) avoided the entropy optimization by aggregating the transition probabilities (Markovian type Categorical Prediction; MCP) to approximate the optimal solution of the maximum entropy approach". Avoid the entropy optimization? How could one derive a BME or ME model by avoiding entropy optimization? Entropy maximization is the basic requirement of the BME approach and the maximum entropy principle. As to "aggregating the transition probabilities", is that a way to develop a geostatistical model? Is kriging just something aggregating weights? How could Allard suddenly have an idea to aggregate transition probabilities? The simplified MCRF model is indeed composed of spatial transition probability terms, but the MCRF model was strictly derived mathematically step by step, not a model obtained by aggregating transition probabilities arbitrarily. According to what theory did Allard aggregate transition probabilities? According to the MCRF model! If the authors dared not tell the truth, please do not write misleading articles!

Allard used Li's ideas to derive his MCP method. Equation (4) is exactly the MCRF model, which can be written in different forms by applying the detailed balance equation of Markov chain theory, as pointed out by Li et al. (2015). This model can be derived only by the spatial Bayesian updating and spatial conditional independence assumption (Li's ideas for locally-conditioned Markov chain to solve the small class underestimation problem of the CMC model), but cannot be derived by entropy maximization.

Line 181-182: 13 borehole logs are definitely insufficient to obtain any meaningful experimental transiograms in horizontal dimensions. Because the boreholes are unevenly distributed, the vertical experimental transiograms cannot be representative. These points have been proved by the experimental transiograms provided in Figure 4. So your theoretical transiograms (i.e., transiogram models) do not reflect the characteristics of the data.

Line 239-243: While you used the method proposed by Carle and Fogg (1997) for transiogram joint modeling, you did not use its advantage for subsurface modeling.

Line 253-254: What is the MCP algorithm? What is its difference from other algorithms? Any geostatistical simulation algorithm has to work on a geostatistical model. You said that you used a ME approach to approximate the posterior conditional probabilities. However, Allard's MCP method is the same as the MCRF model and it is not a ME model.

Line 277: MCP (Markovian type Categorical Prediction) is neither a proper name for a geostatistical model nor a proper name for a geostatistical simulation algorithm. It is actually the MCRF model. You were implementing the MCRF model but stated as if you were using another model.

Line 292-293: The authors said "The simulation performed by MCP takes into account only sums and products of univariate and bivariate probabilities (maximum entropy approach)". If using sums and products of univariate and bivariate probabilities is maximum entropy approach, is the Tau model proposed by Journel also a maximum entropy approach? Christakos' BME approach regarded kriging as a special case of BME and usually used simple kriging as its fundamental model for dealing with neighborhoods. Is simple kriging also sums and products of univariate and bivariate probabilities? The MCRF model is even not a maximum entropy model! According to Jaynes, Gibbs distribution is a maximum entropy model! Did the authors read Chrsitakos (1990), Bogaert (2002) and Li (2007a)? The simulated realizations generated by the MCRF model and Bogaert's BME model are different.

References

Allard, D., D. D'Or, and R. Froidevaux. 2009. Estimating and simulating spatial categorical data using an efficient maximum entropy approach. Avignon: Unité Biostatistique et Processus Spatiaux Institut National de la Recherche Agronomique. Technical Report No. 37.

Allard, D., D. D'Or, and R. Froidevaux. 2011. An efficient maximum entropy approach for categorical variable prediction. *European Journal of Soil Science*, 62(3): 381-393.

Allard, D. 2012. Response to W. Li and C. Zhang by D. Allard, D. D'Or & R. Froideveaux. *European Journal of Soil Science,* 63(1): 125-128.

Bogaert, P. 2002. Spatial prediction of categorical variables: the Bayesian maximum entropy approach. *Stochastic Environmental Research and Risk Assessment* (SERRA), 16 (6): 425–448.
Bogaert, P., and D. D'Or. 2002. Estimating soil properties from thematic soil maps. *Soil Science Society of America Journal*, 66: 1492-1500.
Carle, S.F., and G.E. Fogg. 1997. Modeling spatial variability with one and multidimensional continuous-lag Markov chains. *Mathematical Geology*, 29: 891–918.
Carle, S.F., and G.E. Fogg. 1996. Transition probability-based indicator geostatistics. *Mathematical Geology*, 28: 453–476.
D'Or, D., D. Allard, P. Biver, R. Froidevaux, and A. Walgenwitz. 2008. Simulating categorical random fields using the multinomial regression approach. In: Ortiz JM, Emery X (eds) *Proceedings of the eight international geostatistics congress*, Santiago de Chile. GECAMIN Ltd.
D'Or, D., P. Bogaert, and G. Christakos. 2001. Application of the BME approach to soil texture mapping. *Stochastic Environmental Research and Risk Assessment*, 15: 87–100.
Christakos, G. 1990. A Bayesian/maximum-entropy view to the spatial estimation problem. *Mathematical Geology*, 22 (7): 764–777.
Elfeki, A. 1996. Stochastic characterization of geological heterogeneity and its impact on groundwater contaminant transport. Ph.D. Dissertation, Delft University of Technology, The Netherlands.
Elfeki, A.M., and F.M. Dekking. 2001. A Markov chain model for subsurface characterization: Theory and applications. *Mathematical Geology*, 33: 569-589.
Journel, A.G. 2002. Combining knowledge from diverse sources: an alternative to traditional data independence hypothesis. *Mathematical Geology*, 34: 573–596.
Li, W., and C. Zhang. 2005. Application of transiograms to Markov chain modeling and spatial uncertainty assessment of land cover classes. *GIScience & Remote Sensing*, 42(4): 297-319.
Li, W., and C. Zhang. 2006. A generalized Markov chain approach for conditional simulation of categorical variables from grid samples. *Transactions in GIS*, 10: 651–669.
Li, W. 2007a. Markov chain random fields for estimation of categorical variables. *Mathematical Geology*, 39 (3): 321–335.
Li, W. 2007b. Transiograms for characterizing spatial variability of soil classes. *Soil Science Society of America Journal*, 71(3): 881-893.
Li, W., and C. Zhang. 2007. A random-path Markov chain algorithm for simulating categorical soil variables from random point samples. *Soil Science Society of America Journal*, 71(3): 656-668.
Li, W., and C. Zhang. 2008. A single-chain-based multidimensional Markov chain model for subsurface characterization. *Environmental and Ecological Statistics*, 15: 157–174.
Li, W., C. Zhang, M.R. Willig, D.K. Dey, G. Wang, and L. You. 2015. Bayesian Markov chain random field cosimulation for improving land cover classification accuracy. *Mathematical Geosciences*, 47(2): 123-148.
Luo, J. 1993. Transition probability approach to statistical analysis of spatial qualitative variables in geology. In: *Proceedings of the 25th Anniversary Meeting of the International Association for Mathematical Geology*, October 10-14, 1993, Prague, Czech Republic. Forster A, Merriam DF (eds.) Geologic modeling and mapping Plenum Press, New York, 1996, p. 281–299.
Huang, X., Z. Wang, and J. Guo. 2016. Prediction of categorical spatial data via Bayesian updating. *International Journal of Geographical Information Science*, 30(7): 1426-1449.
Sartore, L. 2013. spMC: modelling spatial random fields with continuous lag spatial Markov Chains. *The R Journal*, 5: 16– 28. https://journal.r-project.org/archive/2013-2/sartore.pdf.
Sartore, L., P. Fabbri, and C. Gaetan. 2016. spMC: an R-package for 3D lithological reconstructions based on spatial Markov chains. *Computers & Geosciences*, 94: 40-47.
Schwarzacher, W. 1969. The use of Markov chains in the study of sedimentary cycles. *Mathematical Geology*, 1: 17–39.

Postscript: Since the authors have published their article in another journal (*Hydrology*) recently, I think I'd better release this review report online. Weidong Li, 03/20/2020.

Fabbri, P.; Gaetan, C.; Sartore, L.; Dalla Libera, N. Subsoil Reconstruction in Geostatistics beyond Kriging: A Case Study in Veneto (NE Italy). *Hydrology* 2020, *7*, 15.