Comments on "Theoretical Generalization of Markov Chain Random Field from Potential Function Perspective"

Weidong Li and Chuanrong Zhang
Department of Geography, University of Connecticut, Storrs, CT 06268

Huang et al. (2016) claimed that (1) they generalized the Markov chain random field (MCRF) model (Li 2007) using potential functions, (2) they found the internal relationship between Markov random fields (MRF) and MCRF, (3) they solved the highly time-consuming problem of MRF, (4) they proved that MCRF could be derived from MRF, and (5) they implemented a 3D MCRF model in a case study. In their paper, they copied many pieces of sentences from our earlier papers, and wrote lots of complex equations to puzzle readers. However, after carefully reading their paper, we found that it was just a trick that could effectively mislead many people.

The second section of Huang et al. (2016) is "Markov chain model". However, they focused on introducing the unconditional coupled Markov chain (CMC) model of Elfeki (1996). Given the fact that both the unconditional CMC model and the conditional CMC model (Elfeki and Dekking 2001) underestimates small classes and have other deficiencies, it is difficult to say the CMC model represents "Markov chain model". Further extensions of the CMC model by Li et al. (2004) and Li and Zhang (2006) made it more flexible and useful, but only under the situations of very dense sample data (e.g., dense borehole or survey line data) or lacking small classes. We do recognize the contribution of the CMC model, but this does not mean the model is theoretically sound. The small class underestimation problem of the CMC model had been proved and solved many years ago (see Li 2007). It may not be an interesting issue to always introduce it.

The third section of the paper is "Localized MRF approach". However, the authors came to reinterpret irrationally the figure presented in Li (2007), which represents the MCRF defined on sparse sample data. The nearest data, the location being estimated and the transition probability directions in the figure together were surprisingly interpreted as various "cliques" (with one-edge, two-edges, three-edges and so on) by the authors. They stated that arrows (i.e., transition probability head-tail directions) "represent energy transfer direction". They did not forget to mention that "MRF is traditionally defined as undirected graph in graph theory". Here our questions are: 1) who defined MRF on sparse sample data? 2) what is the reason for an unsampled location to transfer energy to a nearest datum? and 3) if MRF (note that it should be "Markov network") is an undirected graph, why did it become a directed graph here?

Then the authors considered site s with three neighbors - s_1, s_2 and s_3, and provided all the potential functions V_n. Each of those potential functions for their "cliques" was decomposed into a sum of logarithms and negative logarithms of two-point and multiple-point conditional probabilities. For example, their "three-edge clique" potential function was decomposed as

$$\begin{aligned} V_3\left(f_s, f_{s_1}, f_{s_2}, f_{s_2}\right) = & -\ln \Pr\left(f_s \mid f_{s_1}, f_{s_2}, f_{s_3}\right) + \\ & \ln \Pr\left(f_s \mid f_{s_1}, f_{s_2}\right) + \ln \Pr\left(f_s \mid f_{s_1}, f_{s_3}\right) + \\ & \ln \Pr\left(f_s \mid f_{s_2}, f_{s_3}\right) - \ln \Pr\left(f_s \mid f_{s_1}\right) - \\ & \ln \Pr\left(f_s \mid f_{s_2}\right) - \ln \Pr\left(f_s \mid f_{s_3}\right) \end{aligned}$$

Then the authors obtained their Equation (14), and claimed that "Here, the joint probability of an MRF can be expressed by the product of local conditional probabilities. The complicated parameter estimation algorithm and iteration process can be avoided by computing local conditional probabilities. So, the highly time-consuming problem of MRF mentioned in Ref. [21] has been solved." (note that here the Ref. [21] means Stien and Kolbjørnsen (2011)). Here our questions are: 1) how can you solve the multiple-point conditional probabilities in your clique potential functions? and 2) if you cannot solve the multiple-point conditional probability terms, how can you solve the localized MRF model? The truth is that the authors solved nothing!

The fourth section of the paper is "Relationship between MRF and MCRF". However, what the authors did are as follows: 1) wrote the simplified MCRF model based on the conditional independence assumption into a different style (i.e., an exponent-logarithm form); 2) directly put the new form of MCRF model into their energy function; and 3) then changed the two-point conditional probability term and those two-point likelihood function terms into transition probabilities with distance lags. Then they claimed that they derived the joint probability of an MRF exactly as the MCRF model. How did they derive it? Use a trick – use the MCRF model in a different form to derive the MCRF model! They used the likelihood functions in the MCRF model to replace their "clique" conditional probability functions in their MRF model, without any reasons!

They did not forget to mention the conditional independence assumption. Our question is: how could they use the conditional independence assumption to a multiple-point conditional probability such as $\Pr\left(f_s \mid f_{s_1}, f_{s_2}, f_{s_3}\right)$ in their localized MRF model to get a two-point transition probability? Apparently, the authors could not understand that the conditional independence assumption for nearest data within a neighborhood around an unobserved location, which was defined in Li (2007), is only applicable to the multiple-point likelihood functions such as $\Pr(f_2 \mid f_s, f_1, f_3)$ with f_s representing the label of unobserved location, not applicable to the ordinary multiple-point conditional probabilities such as $\Pr\left(f_s \mid f_{s_1}, f_{s_2}, f_{s_3}\right)$. One cannot get $\Pr\left(f_s \mid f_{s_1}, f_{s_2}, f_{s_3}\right) = \Pr(f_2 \mid f_s)$ or $\Pr(f_s \mid f_2)$ with the conditional independence assumption. In the MCRF model (Li 2007), the multiple-point conditional probabilities that were simplified into transition probabilities using the conditional independence assumption are essentially multiple-point likelihood functions. That is why the directions of those transition probabilities are from the unobserved location to nearest data, not the inverse. Apparently, because they could not understand and pass through this step, the authors had to play a trick. Therefore, the authors did not derive the simplified MCRF model from MRF model, nor generalized it by potential functions, because the MCRF model is already generalized, let alone that MRF was never a spatial model defined on sparse sample data.

There is no necessity to further examine their MRF model anymore. Now let's check their case study. They have only four well logs, but they estimated both vertical and lateral transiogram models from them. The 3-D model they used for their case study (i.e., Equation (28) in their paper) is exactly the 3-D MCRF model provided in Li (2007) for conditional simulation. They claimed they did their simulation "by using MCRF algorithm", but they did not mention what and where the MCRF algorithm they used is. They provided "four realizations of MCRF lithofacies stochastic simulation" and three maximum occurrence probability maps of mudstone generated by MCRF algorithm, conditioned on 1765 samples". The authors did not tell where and how they suddenly got 1765 samples. But one thing we can see is that none of the result maps was from the real MCRF model. Apparently they did not develop a suitable simulation method to condition the simulation of a grid cell to surrounding nearest data.

After their "MCRF simulation", they made a comparison with TMC simulation, and this time they provided a simulation procedure which has no way to work. Did they use the TMC fixed-path simulation algorithm? What is surprising is that they only used a simple local conditional probability $\Pr(x_{i,j,k} = l \mid x_{i-1,j,k} = l_1, x_{i,j-1,k} = l_2, x_{i,j,k-1} = l_3)$, which even cannot be conditioned to surrounding sample data or nearest data in six cardinal directions, for performing simulation. Isn't it suitable to compare a simulation by a 3-D MCRF model, which is conditional to nearest data in six cardinal directions, with a simulation by a TMC model that is only conditional to three immediately adjacent cells? Does this mean that their MCRF conditional simulation was also a simulation conditional to only three immediately adjacent cells? Is such a spatial simulation proper?

In the second section, they mentioned that "Liang et al. (2014) proposed the triplex Markov chain (TMC) in 3-D space" with showing an unconditional 3-D CMC model. In the Introduction section, they stated that “Liang et al. (2014) used MRF tool to simulate lithofacies distribution and proposed a Gibbs distribution to characterize reservoir heterogeneity”. In their eyes, the unconditional CMC model is a MRF and their 3-D CMC model was always unconditional (to borehole data ahead), but in the TMC method (Li et al. 2004) both CMCs are conditional to sample data ahead in two directions. They knew that the TMC model is composed of two CMCs, but it seems they did not understand that the purpose of the TMC using two CMCs was to overcome the directional effect (essentially the unilateral correlation problem of simulated data) occurring in simulated realizations of the CMC model, as they stated that “The fully independent assumption caused directional effect (i.e., pattern inclination or diagonal trend) and the small-class underestimation problem.” The directional effect is irrelevant with the full independence assumption.

Finally they concluded that “The focus of this work is the theoretical foundation of Markov chain random field (MCRF) and relationship between MRF and MCRF.” In Li (2007), the MCRF model was clearly derived using Bayes’ theorem and the derivation process was simple and rational. However, in Huang et al. (2016), the authors threw away the real theoretical foundation of the MCRF model – Bayes’ theorem, and they assigned a new theoretical foundation to the MCRF model.

Then what is the contribution of their paper and what scientific problem did they solve in their paper, except for playing a trick with misleading statements? Claiming that they derived the MCRF model from the MRF model by playing a trick and performing wrong simulation case studies to get wrong results to mislead others are improper.

References

[1] HUANG X, WANG ZZ, GUO JH. Theoretical generalization of Markov chain random field from potential function perspective. Journal of Central South University, 2016, 23(1): 189–200.

[2] LI W. Markov chain random fields for estimation of categorical variables. Mathematical Geology, 2007, 39(3): 321-335.

[3] ELFEKI AM. Stochastic characterization of geological heterogeneity and its impact on groundwater contaminant transport. Ph.D. diss. Delft University of Technology, Balkema Publisher, The Netherlands. ISBN 90-5410-666-2. 1996.

[4] ELFEKI AM and DEKKING FM. A Markov chain model for subsurface characterization: Theory and applications. Mathematical Geology, 2001, 33: 569–589.

[5] LI W, ZHANG C, BURT JE, ZHU A, FEYEN J. Two-dimensional Markov chain simulation of soil type spatial distribution. Soil Science Society of America Journal, 2004, 68(5): 1479-1490.

[6] LI W, ZHANG C. A generalized Markov chain approach for conditional simulation of categorical variables from grid samples. Transactions in GIS, 2006, 10(4): 651-669.

[7] STIEN M, KOLBJØRNSEN O. Facies modeling using a Markov mesh model specification. Mathematical Geosciences, 2011, 43(6): 611–624.

[8] LIANG YR, WANG ZZ, GUO JH. Reservoir lithology stochastic simulation based on Markov random fields. Journal of Central South University, 2014, 21: 3610-3616.

Weidong Li Email: weidong.li@uconn.edu or weidongwoody@gmail.com Tel: 1-860-486-2728
Chuanrong Zhang Email: chuanrong.zhang@uconn.edu Tel: 1-860-486-2610

May 31, 2016, Storrs.