Some Misunderstandings on the Transiogram as a Spatial Correlation Measure of Categorical Data

Weidong Li & Chuanrong Zhang

Department of Geography, University of Connecticut, Storrs, CT 06269, USA

After the transiogram concept and some joint-modeling methods were proposed for implementing multidimensional Markov chain simulation of categorical spatial variables (Li 2007a, Li and Zhang 2010), some misunderstandings about the transiogram concept and transiogram models emerged in some online materials (e.g., Cao *et al.* 2013) and some manuscripts that were attempted to publish during 2016 to 2017. The misunderstandings with ambiguous statements could easily confuse many readers, including some researchers who may have interest in using transiograms in their studies. The following explanations aim to clarify the major misunderstandings.

(1) *The definition of the transiogram*

One misunderstanding on the transiogram is about its definition. During last several years, ***there were some attempts to redefine the transiogram with joint probability and/or indicator variables due to misunderstandings***. Li (2007a) defined the transiogram theoretically as a transition probability-lag function and visually as a transition probability-lag diagram. First, such a definition considers the transiogram as transition probability in the Markov chain framework. Transition probability is the fundamental element of Markov chain theory. It can be estimated directly from real data, and the estimation can be unidirectional, bidirectional, multidirectional or omnidirectional, with or without a tolerance angle and width. There is no necessity to derive transition probability values from bivariate joint probability values through the relationship between conditional probability and joint probability, while one can estimate them directly from the same real data. Although a transition probability is a two-point conditional probability, transition probability has its unique properties and meanings in Markov chain theory, and it is traditionally used within a transition probability matrix (TPM). For example, idealized transiograms can be directly computed from a TPM, but two-point joint probability has no such an advantage. Therefore, ***redefining the transiogram through a bivariate joint probability-lag function would eliminate the legitimacy of unidirectional cross transiograms and idealized transiograms***. Second, transition probabilities directly describe the dependencies of the classes (or states) of discrete/categorical variables. There is no necessity to first transform a categorical data set (with multiple classes) into multiple sets of indicator data (i.e., 0 and 1 values) and then use indicator data to estimate transition probabilities or transiograms. Carle and Fogg (1996) studied the relationships of transition probability with indicator variables, indicator covariance and indicator variograms; while the study is undoubtedly important and interesting, its major objective was using transition probability to reformulate indicator kriging models based on the relationships. In addition, the transiogram definition in Li (2007a) is also identical with the transition probability-lag functions and Markov diagrams provided in Luo (1996).

(2) *The symmetry of two-point joint probability function*

Some thought that two-point joint probabilities or joint probability functions could be asymmetric or even unidirectional. This was probably an error in some early articles related with spatial transition probability (function) (e.g., Carle and Fogg 1996). ***Transition probability, conditional probability, and joint probability are three different concepts***, although they are quantitatively related when they are used to describe the probability relationship of two events of the same random process at two spatial locations or two time points. ***A joint probability P(A∩B) (or written as P(A, B)) is always symmetric and non-unidirectional***, that is, P(A∩B) = P(B∩A). Assuming P(A∩B) to be asymmetric and estimating it asymmetrically is equal to estimating a conditional probability or transition probability. In probability theory, a conditional probability P(A|B) is defined as P(A|B) = P(A∩B)/P(B). With this definition, it seems that P(A|B) does not necessarily have the legitimacy to be unidirectional, because P(A∩B) is non-unidirectional. Transition probability P_{AB} as an element of Markov chain theory is a special conditional probability and can be unidirectional. Transition probabilities in the form of a TPM describe the properties of a Markov chain. ***If a Markov chain is irreversible, its forward transition probabilities and backward transition probabilities are not equal*** (under this situation, the forward Markov chain and the backward Markov chain may be regarded as two different Markov chains, as conventionally one TPM describes one stationary Markov chain). That is why it is not a good choice to define spatial transition probability or transiogram using spatial joint probability or joint probability-lag function.

(3) *Transiograms and the path of a Markov chain*

Some thought that spatial transition probability and transiogram should be path-dependent, because in a multidimensional space a Markov chain may have many different paths from point **u** to point **u'**, and thus if it goes through different paths the transition probabilities along different paths from **u** to **u'** should be different. This is a misunderstanding. A Markov chain as a stochastic process may have a path. However, ***transition probability and transiogram, as static measures estimated from data, have nothing to do with paths***. The **h** in transiogram is a vector variable because it contains direction, that is, **h** = (h, direction), rather than a sequence of locations. It has no meaning of path.

(4) *The validity of transiogram models*

Some though that some transiogram models (e.g., spherical and Gaussian models), except for the exponential model, are invalid ***because similar variogram models were recently thought to be invalid in indicator kriging*** under some special situations (indicator random fields, particularly excursion sets of Gaussian random fields). However, on the one hand, these variogram models have been widely used in indicator kriging for decades; on the other hand, in Markov chain geostatistics ***there is no requirement for MCRFs to be the "excursion sets of Gaussian random fields" or even "indicator random fields"***. Any reasonable models that can fit experimental transiograms effectively and meet the basic requirements of transition probabilities (e.g., summing-to-unity, non-negative) may be used to provide transition probability parameters to MCRF models, and thus may be valid transiogram models in Markov chain geostatistics (Li 2007b, Li *et al.* 2015). Transiograms include auto-transiograms and cross-transiograms, which have different curve shapes and physical meanings. Experimental transiograms have complex shapes. While idealized transiograms are smooth curves and idealized auto-transiograms tend to be exponential, studies demonstrated that some of idealized cross-transiograms are non-exponential - some have a peak in the low-lag section, and less commonly some tend to have the shape of Gaussian variogram model (Li *et al.* 2012) (Fig. 1). This means that ***even if a categorical data set is absolutely stationary Markovian, its cross-transiograms are still not all exponential***.

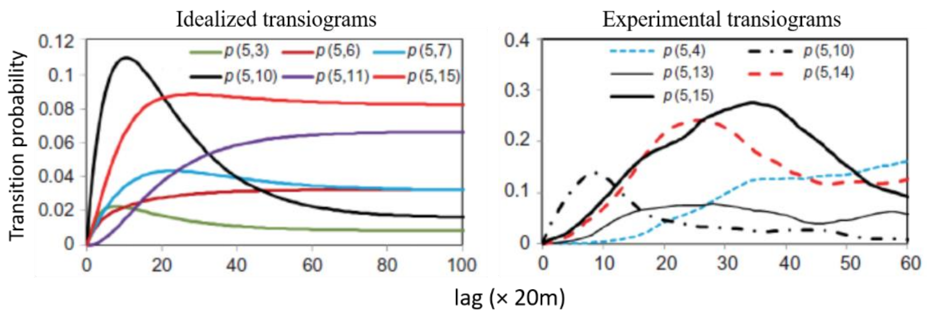

Figure 1. Some idealized (left) and experimental (right) cross-transiograms estimated from a soil map, among which some are obviously non-exponential and some are even parabolic near the origin (from Li *et al.* 2012).

The transiogram was proposed for implementing multidimensional Markov chain simulation and also as a general spatial correlation measure for categorical data, rather than specifically for implementing indicator kriging or simulating indicator random fields. Even if some variogram models were proven to be invalid for indicator random fields or indicator kriging in some situations, it does not mean that the similar math models were invalid to transiograms for categorical data or the MCRF approach. In addition, ***so far the circular and triangular variogram models have not been recommended or used by anybody for transiogram modeling***. Interestingly and ironically, Cao *et al.* (2013, figures 1 and 4) showed that some cross-transiograms computed from a TGS (truncated Gaussian simulation) realization tend to be Gaussian with parabolic features near the origin and none of the transiogram models by non-parametric transiogram modeling method is exponential, both conflicting with their assertions. At the time that the transiogram had not been widely used and the MCRF approach was still at the early stage of development, rushing to judge such a spatial correlation measure (i.e., the transiogram) in an ambiguous way could easily confuse many colleagues and potential users.

(5) *The transiogram and the energy function of Markov random fields*

Some thought that the energy function of the Markov random field (MRF) model (i.e., the Gibbs distribution) could be simplified into a spatial two-point conditional probability function in the form of an exponential function if only a two-node clique is considered; thus, the transiogram could be derived from the MRF model, and consequently the logistic function should be a transiogram model. We are afraid that this is a misunderstanding to both the MRF model/Gibbs distribution and the transiogram concept. The Gibbs distribution, as a maximum entropy distribution, include all of the cliques (from single-node cliques to multiple-node cliques) of a neighborhood (if not consider the whole random field) in its energy function, among which each clique's potential function represents the energy of the exact clique. ***There is even no reason or rationale to simplify a multiple-node clique to one or multiple two-node clique(s) while they represent energy at different levels***. Even if one might ignore all of the multiple-node cliques (e.g., three-node and four-node cliques) from the energy function of Gibbs distribution, there is no reason to consider only one two-node clique and ignore other cliques. On the contrary, as a basic element of Markov chain theory, transition probability has existed for a long time, and ***spatial transition probability had been used in geology since a long time ago*** (see Vistelius 1949, Carr *et al.* 1966). Why does one have to simplify the MRF energy function to find a transition probability while the Gibbs distribution actually contains no transition probability? In addition, as a transition probability-lag function/diagram, the transiogram does not need a partition function to normalize itself. Irrationally interpreting the MCRF model and the transiogram to confuse readers and mess up them and the related scientific area, as done by some people in a series of joking articles (e.g., Huang *et al.* 2016), cannot bring any reputation to the authors themselves.

(6) *The transiogram and pioneer studies*

As a transition probability-lag function/diagram (or curve), there is no doubt that the transiogram should cover the pertinent progress made in pioneer studies in this respect (i.e., in transition probability-lag function/diagram) (e.g., Schwarzacher 1969, Luo 1996, Carle and Fogg 1997). From the beginning, our studies attempted to do so (see Li 2007a). While ***proposing the term and concept system of "transiogram" is necessary for the MCRF approach to avoid terminological confusions*** and also for providing the convenience of description, ***it never claimed any progress made in pioneer studies as ours***. The proposition of the transiogram was a result of the long-term effort of developing 1-D Markov chain into a Markov chain geostatistical approach, which not only needed a formally-established spatial correlation measure with practical estimation methods but also needed a unique name for the measure to work within the Markov chain geostatistical approach without terminological confusion. Although we did not use the transition rate method suggested in Carle and Foggs (1997) for transiogram modeling so far, that does not mean we thought the method was not valuable. In Li (2007a), the "continuous-lag Markov chain models" generated by the transition rate method were regarded as a kind of idealized transiogram models. If one had different opinions or found that this was a mistake, he/she is welcome to discuss on that point. Idealized transiograms were thought to be important for understanding/interpreting real-data transiograms and have utilization as transiogram models mainly in subsurface characterization (see Li 2007a, Conclusions), where sample data are usually insufficient to estimate reliable experimental transiograms.

(7) *The essences of transiograms*

The transiogram has its special meanings within the Markov chain framework and in the MCRF approach. In its idealized form (i.e., either sample data for transiogram estimation are absolutely stationary Markovian, or transiograms are directly calculated from TPMs based on the stationary Markovian assumption), all of its properties are based on conventional transition probability/Markov chain theory. ***For transiograms, the essences are how to interpret their physical meanings from their features as reflections of the pattern of the***

underlying categorical field and how to infer a set of proper transiogram models to provide parameters for geospatial simulations that use spatial transition probabilities. It is obvious that the proper transiogram models should fit the reliable features of the experimental transiograms (especially their low-lag sections) as much as possible.

In general, although the transiogram concept was initially intended for Markov chain geostatistics, it may be used in other spatial statistical methods that need transition probabilities at multiple spatial/time steps if one would like, and it also can be used as an independent metric for spatial variability characterization of categorical data. As to whether other spatial/geo-statistical approaches may have some special requirements to transiogram models, it should be discussed only within the frameworks of those specific approaches. Baseless and irrational interpretations or connections with irrelevant things (e.g., energy function of MRFs) are unnecessary and could be misleading.

References

Cao, G., Kyriakidis, P. and Goodchild, M., 2013. On spatial transition probabilities as continuity measures in categorical fields. *arXiv preprint arXiv*:1312.5391.

Carle, S.F. and Fogg, G.E., 1996. Transition probability-based indicator geostatistics. *Mathematical geology*, 28(4), pp.453-476

Carle, S.F. and Fogg, G.E., 1997. Modeling spatial variability with one and multidimensional continuous-lag Markov chains. *Mathematical Geology*, 29(7), pp.891-918.

Carr, D.D., Horowitz, A., Hrabar, S.V., Ridge, K.F., Rooney, R., Straw, W.T., Webb, W. and Potter, P.E., 1966. Stratigraphic sections, bedding sequences, and random processes. *Science*, 154(3753): 1162-1164.

Huang, X., Wang, Z.Z. and Guo, J.H., 2016. Theoretical generalization of Markov chain random field from potential function perspective. *Journal of Central South University*, 23(1), pp.189-200.

Li, W., 2007a. Transiograms for characterizing spatial variability of soil classes. *Soil Science Society of America Journal*, 71(3), pp.881-893.

Li, W., 2007b. Markov chain random fields for estimation of categorical variables. *Mathematical Geology*, 39(3): 321-335.

Li, W. and Zhang, C., 2010. Linear interpolation and joint model fitting of experimental transiograms for Markov chain simulation of categorical spatial variables. *International Journal of Geographical Information Science*, 24(6), pp.821-839.

Li, W., Zhang, C. and Dey, D.K., 2012. Modeling experimental cross-transiograms of neighboring landscape categories with the gamma distribution. *International Journal of Geographical Information Science*, 26(4), pp.599-620.

Li, W., Zhang, C., Willig, M.R., Dey, D.K., Wang, G. and You, L., 2015. Bayesian Markov chain random field cosimulation for improving land cover classification accuracy. *Mathematical Geosciences*, 47(2), pp.123-148.

Luo, J., 1996. Transition probability approach to statistical analysis of spatial qualitative variables in geology. In: Forster A, Merriam DF (eds.) *Geologic modeling and mapping* (Proceedings of the 25th Anniversary Meeting of the International Association for Mathematical Geology, October 10-14, 1993, Prague, Czech Republic). Plenum Press, New York. p. 281–299.

Schwarzacher, W., 1969. The use of Markov chains in the study of sedimentary cycles. *Mathematical*

Geology, 1, pp.17–39.
Vistelius, A.B., 1949. On the question of the mechanism of formation of strata. *Dokl. Akad. Nauk SSSR*, 65(2): 191-194.

Weidong Li & Chuanrong Zhang
Department of Geography, University of Connecticut,
Storrs, CT 06269, USA
E-Mail: weidong.li@uconn.edu
& chuanrong.zhang@uconn.edu

October, 2018, Storrs, CT