

Article

A Linear Bayesian Updating Model for Probabilistic Spatial Classification

Xiang Huang and Zhizhong Wang *

Department of Statistics, Central South University, Changsha 410012, Hunan, China; huangxiang@csu.edu.cn

* Correspondence: wzz8713761@163.com or wangzhizhong@csu.edu.cn; Tel.: +86-731-8713-761

Academic Editor: Palmiro Poltronieri

Received: 8 September 2016; Accepted: 21 November 2016; Published: 29 November 2016

Abstract: Categorical variables are common in spatial data analysis. Traditional analytical methods for deriving probabilities of class occurrence, such as kriging-family algorithms, have been hindered by the discrete characteristics of categorical fields. To solve the challenge, this study introduces the theoretical backgrounds of the linear Bayesian updating (LBU) model for spatial classification through an expert system. The main purpose of this paper is to present the solid theoretical foundations of the LBU approach. Since the LBU idea is originated from aggregating expert opinions and is not restricted to conditional independent assumption (CIA), it may prove to be reasonably adequate for analyzing complex geospatial data sets, such as remote sensing images or area-class maps.

Keywords: expert opinions; linear Bayesian updating; spatial classification; transition probabilities

1. Introduction

Categorical spatial data, such as lithofacies, land-use/land-cover classifications, and mineralization phases, are widely investigated geographical and geological information sources. They are typically represented by mutually exclusive and collectively exhaustive classes and visualized as area-class maps [1]. In the Geo-information context, Rao's quadratic diversity was used in [2] to measure the scale-dependent landscape structure. In the geological counterpart, a spatial hidden Markov chain model was employed in [3] for estimation of petroleum reservoir categorical variables. As a geostatistical model, the Markov chain random field (MCRF) theory [4] and Markov chain sequential simulation (MCSS) algorithm [5] are common choices for the prediction of categorical spatial data. They have been widely used in spatial-related fields and gratifying results have been achieved. However, the MCRF approach is based on a conditional independent assumption (CIA), which may be inappropriate due to complex data interaction in a spatial context [6]. The Tau model [7] and Nu expression [8] introduce additional weights to relax the assumption of conditional independence. It is obvious that these power or multiplication relationships between multi-point pre-posterior probabilities and two-point conditional probabilities involve some subjective guesswork and may not be suitable in real-world spatial analysis. As for the generalized linear mixed model (GLMM) [6], where intermediate, latent, spatially correlated, normal variables are assumed for the observable non-normal responses to account for spatial dependence information. The random effects are always assumed to follow a normal distribution in the GLMM. Our concern here is whether the latent variables for different categories can be assumed to be independent of each other at the same location.

Generally speaking, the spatial classification problem can be regarded as combing two-point transition probabilities into a multi-point conditional probability. A formal introduction of most of the available approaches to aggregate probability distributions in geosciences can be found in [9]. Our task is to use the probability pooling method for spatial classification based on the pioneering work of [10,11]. We profit from the predecessor's studies and interpret transition probabilities as expert opinions. The transition probabilities are obtained by the transiogram [12] spatial measure.

The remainder of this paper is organized as follows. We begin by introducing the basic forms of the linear Bayesian updating (LBU) method in Section 2. We then make detailed proofs in Section 3 for some propositions introduced by [10,11] which have not yet been proven. A real-world case study is given in Section 4. Finally, conclusions and future challenges are discussed in Section 5.

2. Linear Bayesian Updating

Consider the spatial locations x_0, x_1, \dots, x_n in the remote sensing images or area-class maps. We use A and D_1, \dots, D_n to represent the events in sample spaces of categorical random variable $C(x_0)$ and $C(x_1), \dots, C(x_n)$ respectively and \bar{A} denotes the complementary event of A . In the case of categorical data, let \mathbf{A} be the finite set of events in the sample space Ω such that the events A_1, A_2, \dots, A_K of \mathbf{A} are mutually exclusive and collectively exhaustive. Obviously,

$$A_j = \begin{cases} 1 & \text{if } C(x_0) = j, j = 1, 2, \dots, K \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In the subsequent discussions, A will be used as a general notation for A_j . We treat the n neighboring events D_1, D_2, \dots, D_n as experts, and consider the conditional probabilities $P(A|D_i)$ as expert opinions Q_i for the occurrence of A , an event of interest. The experts' opinions are regarded as random variables Q_i whose values $q_i, 1 \leq i \leq n$, are to be revealed to the decision maker (DM). The posterior probability of A given $\mathbf{Q} = \mathbf{q}$ is then $p^*(\mathbf{q})$.

The original LBU method was firstly proposed by [10] in statistical science in the form

$$p^*(\mathbf{q}) = p + \sum_{i=1}^n \lambda_i (q_i - \mu_i) \tag{2}$$

with possibly negative weights, λ_i , expressing the amounts of correlation between each Q_i and A . μ_i denotes the mathematical expectation of Q_i . When $\mu_1 = \dots = \mu_n = p$ and $\lambda_i \geq 0$, Equation (2) yields to the linear opinion pool [13]

$$p^*(\mathbf{q}) = \lambda_0 p + \sum_{i=1}^n \lambda_i q_i \quad \text{subject to } \sum_{i=0}^n \lambda_i = 1 \tag{3}$$

where the DM is considered as one of the experts. Our LBU method follows closely to that of [10], which is proved to be the only formula satisfying $\int p^*(\mathbf{q}) dF(\mathbf{q}) = p$ for all distribution dF with mean vector μ . The highlight of our LBU model lies in the fact that the random variable Q_i has been replaced by transition probability, a measure for spatial continuity.

3. Theoretical Foundations of Linear Bayesian Updating

3.1. Parameter Ranges for Linear Bayesian Updating

Although the LBU model has been used in [10,11], it is our conviction that many theoretical challenges need to be solved to better develop this method for further use. A legitimate posterior probability can be obtained only when λ_i obeys a number of inequalities [10]. Since $p^*(\mathbf{q})$ is a probability, it must satisfy $0 \leq p^*(\mathbf{q}) \leq 1$, that is to say,

$$\begin{cases} p + \sum_{i=1}^n \lambda_i (q_i - \mu_i) \geq 0 \\ p + \sum_{i=1}^n \lambda_i (q_i - \mu_i) \leq 1. \end{cases} \tag{4}$$

Through algebra transformation, (4) can be simplified as

$$\begin{cases} 1 + \frac{\sum_{i=1}^n \lambda_i q_i}{p} \geq \frac{\sum_{i=1}^n \lambda_i \mu_i}{p} \\ \frac{\sum_{i=1}^n \lambda_i q_i - \sum_{i=1}^n \lambda_i \mu_i}{1-p} \leq 1. \end{cases} \tag{5}$$

Suppose that all λ_i are positive, since $0 \leq q_i \leq 1$, as long as

$$\begin{cases} \frac{\sum_{i=1}^n \lambda_i \mu_i}{p} \leq 1 \\ \frac{\sum_{i=1}^n \lambda_i (1 - \mu_i)}{1 - p} \leq 1, \end{cases} \quad (6)$$

(5) can be satisfied. Therefore, if the DM considers that all λ_i are positive, the most common case, then they must be chosen so that

$$\max \left\{ \sum_{i=1}^n \lambda_i \mu_i / p, \sum_{i=1}^n \lambda_i (1 - \mu_i) / (1 - p) \right\} \leq 1, \quad (7)$$

which can be regarded as a sufficient but not necessary condition of the LBU method. Only when (6), or equivalently (7), is satisfied, can Equation (2) be a valid probabilistic model.

3.2. Interpreting Parameters as Regression Coefficients for Linear Bayesian Updating

The LBU model given above has some parameters, which need to be learned or estimated by the DM. Let $\Sigma_{\mathbf{Q}}$ denote the covariance matrix of \mathbf{Q} , and $\sigma_{p^* \mathbf{Q}}$ be the vector of covariances between p^* and \mathbf{Q} . Let t denote matrix transposition and

$$\mathbf{Q} = (Q_1, Q_2, \dots, Q_n) \quad \boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n) \quad \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)^t,$$

using the definition of expectation, it is

$$E(A|\mathbf{Q}) = p^*, \quad E(E(A|\mathbf{Q})) = E(A) = E(p^*) = p.$$

In addition, Equation (2) can be given as

$$p^* = p + (\mathbf{Q} - \boldsymbol{\mu}) \cdot \boldsymbol{\lambda} \text{ and } \boldsymbol{\mu} = E(\mathbf{Q})$$

i.e.,

$$p^* = p + (Q_1 - \mu_1, Q_2 - \mu_2, \dots, Q_n - \mu_n) \cdot (\lambda_1, \lambda_2, \dots, \lambda_n)^t.$$

Taking the expectations on both sides of the equation after transformation yields

$$\begin{aligned} E[(\mathbf{Q} - \boldsymbol{\mu})^t \cdot (p^* - p)] &= E[(Q_1 - \mu_1, Q_2 - \mu_2, \dots, Q_n - \mu_n)^t \cdot (Q_1 - \mu_1, Q_2 - \mu_2, \dots, Q_n - \mu_n)] \cdot (\lambda_1, \lambda_2, \dots, \lambda_n)^t \\ &= \text{cov}(Q_1, Q_2, \dots, Q_n) \cdot (\lambda_1, \lambda_2, \dots, \lambda_n)^t \\ &= \Sigma_{\mathbf{Q}} \cdot \boldsymbol{\lambda} \end{aligned}$$

we have

$$\boldsymbol{\lambda} = \Sigma_{\mathbf{Q}}^{-1} \sigma_{p^* \mathbf{Q}}$$

provided that the covariance matrix $\Sigma_{\mathbf{Q}}$ is invertible.

Suppose we have m samples in the training set, consider the regression model

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{Y} = \begin{pmatrix} p_1^* - p \\ p_2^* - p \\ \vdots \\ p_m^* - p \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & q_{11} - \mu_1 & \cdots & q_{1n} - \mu_n \\ 1 & q_{21} - \mu_1 & \cdots & q_{2n} - \mu_n \\ & & \ddots & \\ 1 & q_{m1} - \mu_1 & \cdots & q_{mn} - \mu_n \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix},$$

$\varepsilon = (\varepsilon_1, \varepsilon_1, \dots, \varepsilon_m)^t$ denotes the random errors. Provided that the experts do not have a linearly dependent relationship, the least squares estimation of the regression coefficients yields

$$\hat{\mathbf{B}} = (\mathbf{X}^t \mathbf{X})^{-1} (\mathbf{X}^t \mathbf{Y}),$$

which can be written in its matrix form

$$\begin{aligned} \hat{\mathbf{B}} &= \left[\begin{pmatrix} 1 & 1 & \dots & 1 \\ q_{11} - \mu_1 & q_{21} - \mu_1 & \dots & q_{m1} - \mu_1 \\ q_{12} - \mu_2 & q_{22} - \mu_2 & \dots & q_{m2} - \mu_2 \\ \vdots & \vdots & \dots & \vdots \\ q_{1n} - \mu_n & q_{2n} - \mu_n & \dots & q_{mn} - \mu_n \end{pmatrix} \cdot \begin{pmatrix} 1 & q_{11} - \mu_1 & \dots & q_{1n} - \mu_n \\ 1 & q_{21} - \mu_1 & \dots & q_{2n} - \mu_n \\ \vdots & \vdots & \dots & \vdots \\ 1 & q_{m1} - \mu_1 & \dots & q_{mn} - \mu_n \end{pmatrix} \right]^{-1} \cdot \begin{pmatrix} 1 & \dots & 1 \\ q_{11} - \mu_1 & \dots & q_{m1} - \mu_1 \\ q_{12} - \mu_2 & \dots & q_{m2} - \mu_2 \\ \vdots & \dots & \vdots \\ q_{1n} - \mu_n & \dots & q_{mn} - \mu_n \end{pmatrix} \cdot \begin{pmatrix} p_1^* - p \\ p_2^* - p \\ \vdots \\ p_m^* - p \end{pmatrix} \\ &= \begin{pmatrix} m & m(\bar{Q}_1 - \mu_1) & \dots & m(\bar{Q}_n - \mu_n) \\ m(\bar{Q}_1 - \mu_1) & \sum_{i=1}^m (q_{i1} - \mu_1)^2 & \dots & \sum_{i=1}^m (q_{i1} - \mu_1)(q_{in} - \mu_n) \\ m(\bar{Q}_2 - \mu_2) & \sum_{i=1}^m (q_{i2} - \mu_2)(q_{i1} - \mu_1) & \dots & \sum_{i=1}^m (q_{i2} - \mu_2)(q_{in} - \mu_n) \\ \vdots & \vdots & \dots & \vdots \\ m(\bar{Q}_n - \mu_n) & \sum_{i=1}^m (q_{in} - \mu_n)(q_{i1} - \mu_1) & \dots & \sum_{i=1}^m (q_{in} - \mu_n)^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} mp^* - mp \\ \sum_{i=1}^m (p_i^* - p)(q_{i1} - \mu_1) \\ \sum_{i=1}^m (p_i^* - p)(q_{i2} - \mu_2) \\ \vdots \\ \sum_{i=1}^m (p_i^* - p)(q_{in} - \mu_n) \end{pmatrix}. \end{aligned}$$

When the sample size is large enough, the sample mean is approximately equal to the total expectation, we have

$$\begin{aligned} \hat{\mathbf{B}} &= \frac{1}{m} \cdot \begin{pmatrix} 1 & E(Q_1) - \mu_1 & E(Q_2) - \mu_2 & \dots & E(Q_n) - \mu_n \\ E(Q_1) - \mu_1 & E[(Q_1 - \mu_1)^2] & \dots & E[(Q_1 - \mu_1)(Q_n - \mu_n)] \\ E(Q_2) - \mu_2 & E[(Q_2 - \mu_2)(Q_1 - \mu_1)] & \dots & E[(Q_2 - \mu_2)(Q_n - \mu_n)] \\ \vdots & \vdots & \dots & \vdots \\ E(Q_n) - \mu_n & E[(Q_n - \mu_n)(Q_1 - \mu_1)] & \dots & E[(Q_n - \mu_n)^2] \end{pmatrix}^{-1} \cdot m \cdot \begin{pmatrix} E(p^*) - p \\ E[(p^* - p)(Q_1 - \mu_1)] \\ E[(p^* - p)(Q_2 - \mu_2)] \\ \vdots \\ E[(p^* - p)(Q_n - \mu_n)] \end{pmatrix} \\ &= \frac{1}{m} \cdot \begin{pmatrix} 1 & E(Q_1) - \mu_1 & E(Q_2) - \mu_2 & \dots & E(Q_n) - \mu_n \\ E(Q_1) - \mu_1 & \text{cov}(Q_1, Q_2, \dots, Q_n) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ E(Q_n) - \mu_n & \vdots & \vdots & \dots & \vdots \end{pmatrix}^{-1} \cdot m \cdot \begin{pmatrix} E(p^*) - p \\ \sigma_{p^* \mathbf{Q}} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \Sigma_{\mathbf{Q}}^{-1} \cdot \sigma_{p^* \mathbf{Q}} \end{pmatrix}_{(n+1) \times 1}. \end{aligned}$$

Therefore, our derivation gives an explanation of the parameters λ_i in the LBU model as the linear regression coefficients of $p^* - p$ with respect to $\mathbf{Q} - \boldsymbol{\mu}$ when the neighboring events are not linearly dependent. As in multiple regression, each λ_i can thus be thought of as a measure of the additional information that the i th expert provides over and above the other experts and what the DM already knows.

3.3. Invertible Conditions of Linear Bayesian Updating

We now discuss what happened when the linear systems of Equation (2) become invertible. Equation (2) can be rewritten as

$$p^*(\mathbf{q}) = p - \sum_{i=1}^n \lambda_i \mu_i + (q_1, q_2, \dots, q_n) (\lambda_1, \lambda_2, \dots, \lambda_n)^t.$$

Therefore,

$$(q_1, q_2, \dots, q_n) (\lambda_1, \lambda_2, \dots, \lambda_n)^t = p^*(\mathbf{q}) - p + (\mu_1, \mu_2, \dots, \mu_n) (\lambda_1, \lambda_2, \dots, \lambda_n)^t,$$

we have

$$(q_1, q_2, \dots, q_n) \Lambda = [p^*(\mathbf{q}) - p] (\lambda_1, \lambda_2, \dots, \lambda_n) + (\mu_1, \mu_2, \dots, \mu_n) \Lambda,$$

where

$$\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^t (\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \lambda_1^2 & \lambda_1\lambda_2 & \dots & \lambda_1\lambda_n \\ \lambda_2\lambda_1 & \lambda_2^2 & \dots & \lambda_2\lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n\lambda_1 & \lambda_n\lambda_2 & \dots & \lambda_n^2 \end{pmatrix}.$$

Since the determinant $|\Lambda| = 0$, Λ is irreversible. Therefore, q_1, q_2, \dots, q_n cannot be uniquely determined. Thus, the linear systems of Equation (2) are not invertible under these circumstances.

The case where the linear system of Equation (2) is invertible happens only when there is one expert to be consulted, i.e.,

$$p^*(q) = p + \lambda(q - \mu).$$

In this case, the left inverse of the system is

$$q = \frac{p^* - p}{\lambda} + \mu,$$

where p^* is the input. The right inverse can be given as

$$q_i = \frac{p^*(q_i) - p}{\lambda_i} + \mu_i,$$

where $p^*(q_i)$ is the desired output after consultation. The necessary q_i can be compared with the corresponding transition probability $P(A|D_i)$. If large deviation emerges, the expert Q_i may seem not to be convincing.

4. Case Study

We now present a case study to demonstrate the use of the method. The Swiss Jura data set [14] is used, where four lithology types are sampled in a 14.5 km² region. These rock types are Argovian, Kimmeridgian, Sequanian and Quaternary; corresponding class proportions of these four categories are 20.46%, 32.82%, 24.32%, 22.39% respectively. We have 259 samples in total for prediction (Figure 1).

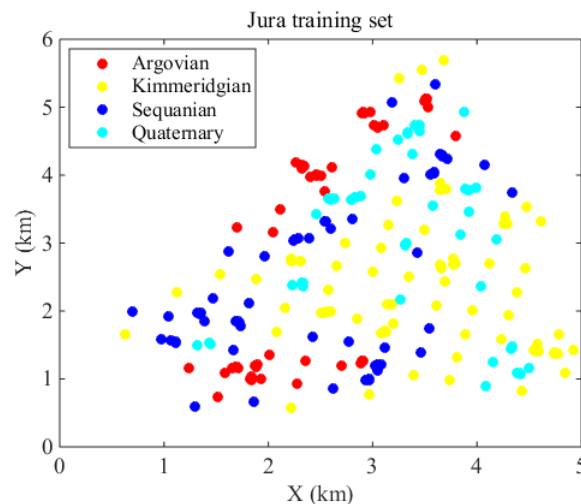


Figure 1. Jura lithology data set with four classes.

The first task we need to do is to obtain the expert opinions (i.e., transition probabilities) in spatial scenarios. We use the 10 nearest samples for prediction, thus we always get 10 experts for consultation. The detailed procedures for estimating the transition probability are beyond the scope of this work.

One can find the discussions with respect to transiogram fitting in [1,11]. We only show the descriptive statistics of transition probabilities in Table 1.

Table 1. Descriptive statistics of transition probabilities (expert opinions).

Transition Probability	Mean	Median	Maximum	Minimum	Standard Deviation
P_{11}	0.270	0.201	1.000	0.000	0.250
P_{12}	0.290	0.292	0.632	0.000	0.173
P_{13}	0.254	0.239	0.769	0.000	0.140
P_{14}	0.186	0.193	0.436	0.000	0.123
P_{21}	0.227	0.239	0.563	0.000	0.134
P_{22}	0.332	0.282	1.000	0.091	0.198
P_{23}	0.254	0.256	0.436	0.000	0.090
P_{24}	0.187	0.192	0.387	0.000	0.088
P_{31}	0.220	0.200	0.636	0.000	0.141
P_{32}	0.313	0.309	1.000	0.000	0.155
P_{33}	0.294	0.262	1.000	0.000	0.203
P_{34}	0.173	0.185	0.469	0.000	0.119
P_{41}	0.219	0.182	0.720	0.000	0.154
P_{42}	0.312	0.318	0.562	0.000	0.142
P_{43}	0.256	0.238	0.586	0.000	0.123
P_{44}	0.213	0.184	1.000	0.000	0.208

After obtaining the expert opinions, we can use the regression model represented by Equation (2) to estimate the linear weights in spatial classification. Given that multiple neighbors will be involved in spatial scenarios most of the time, the LBU should often be a multivariable linear regression model. With the estimated regression coefficients, we can use the maximum a posteriori (MAP) probability criterion for classification [11].

The final prediction results have been shown in Figure 2. We get an overall classification accuracy of 82.63% (214 out of 259). To better reflect the prediction accuracy, the precision indicator for each lithoface has also been illustrated in Figure 3.

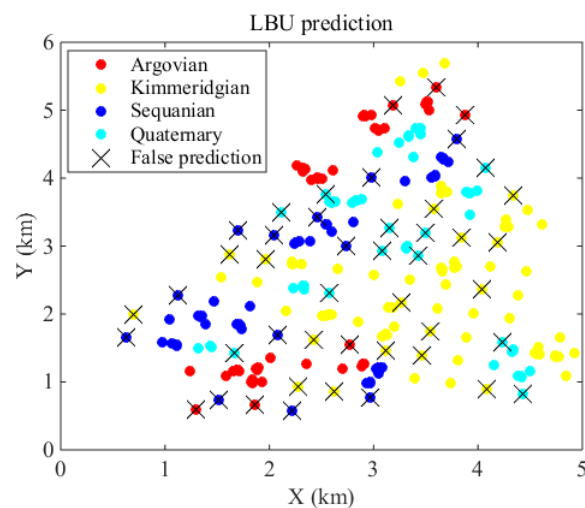


Figure 2. Lithofacies prediction of the corresponding 259 locations based on MAP probability criterion.

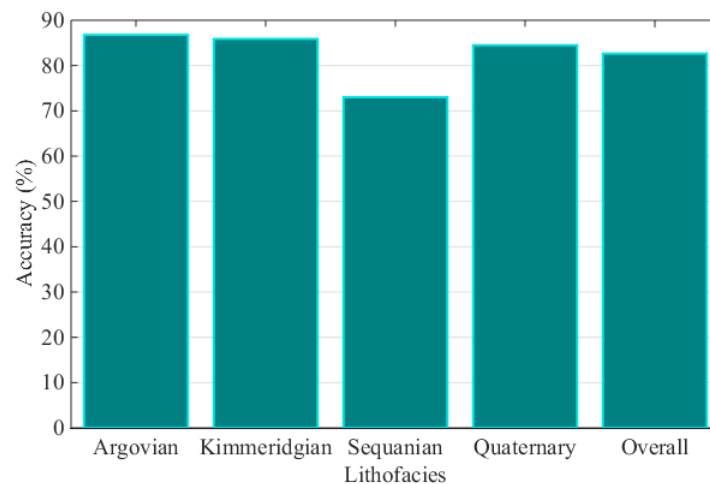


Figure 3. Prediction accuracy comparison.

5. Conclusions

In this work, we consummate the theoretical foundations of the LBU model for the prediction of categorical spatial data. We have enriched our previous findings [11] by adding some rigorous theoretical proofs of the LBU method. To show how the LBU model can work in spatial settings, a real-world case study has also been carried out. As pointed out by [11], our method can also be generalized to nonlinear systems, where more confident probability forecasting results can be obtained.

In the proposed model, the choice of the size of a neighborhood can be regarded as a variable selection problem. The involvement of more neighboring samples is likely to boost the prediction accuracy for the training set, while it may be computation-intensive and accompanied by a higher generalization error, the so-called overfitting. Challenges to determine the optimal number of neighbors will be the focus of our future works and may be addressed in our upcoming papers.

Acknowledgments: This work is supported by the National Key Research and Development Programs of China (No. 2016YFB0502601; No. 2016YFB0502303) and the Fundamental Research Funds for the Central Universities of Central South University (No. 2016zzts011). Special thanks to the editors and two anonymous reviewers for their constructive comments and suggestions. The authors are also indebted to Ying Chen, Danhua Chen, Ruizhi Zhang and Wuyue Shen for their critical reviews of the paper.

Author Contributions: Xiang Huang and Zhizhong Wang conceived and designed the proofs; Xiang Huang wrote the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cao, G.; Kyriakidis, P.C.; Goodchild, M.F. Combining spatial transition probabilities for stochastic simulation of categorical fields. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1773–1791. [[CrossRef](#)]
2. Ricotta, C.; Carranza, M.L. Measuring scale-dependent landscape structure with Rao's quadratic diversity. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 405–412. [[CrossRef](#)]
3. Huang, X.; Li, J.; Liang, Y.; Wang, Z.; Guo, J.; Jiao, P. Spatial hidden Markov chain models for estimation of petroleum reservoir categorical variables. *J. Petrol. Explor. Prod. Technol.* **2016**. [[CrossRef](#)]
4. Li, W. Markov chain random fields for estimation of categorical variables. *Math. Geol.* **2007**, *39*, 321–335. [[CrossRef](#)]
5. Li, W. A fixed-path Markov chain algorithm for conditional simulation of discrete spatial variables. *Math. Geol.* **2007**, *39*, 159–176. [[CrossRef](#)]
6. Cao, G.; Kyriakidis, P.C.; Goodchild, M.F. A multinomial logistic mixed model for the prediction of categorical spatial data. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 2071–2086. [[CrossRef](#)]

7. Krishnan, S. The tau model for data redundancy and information combination in earth sciences: Theory and application. *Math. Geosci.* **2008**, *40*, 705–727. [[CrossRef](#)]
8. Polyakova, E.I.; Journel, A.G. The Nu expression for probabilistic data integration. *Math. Geol.* **2007**, *39*, 715–733. [[CrossRef](#)]
9. Allard, D.; Comunian, A.; Renard, P. Probability aggregation methods in geoscience. *Math. Geosci.* **2012**, *44*, 545–581. [[CrossRef](#)]
10. Genest, C.; Schervish, M.J. Modeling expert judgments for Bayesian updating. *Ann. Stat.* **1985**, *13*, 1198–1212. [[CrossRef](#)]
11. Huang, X.; Wang, Z.; Guo, J. Prediction of categorical spatial data via Bayesian updating. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1426–1449. [[CrossRef](#)]
12. Li, W. Transiogram: A spatial relationship measure for categorical data. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 693–699.
13. Stone, M. The opinion pool. *Ann. Math. Stat.* **1961**, *32*, 1339–1342. [[CrossRef](#)]
14. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: New York, NY, USA, 1997.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).