Methods for Integrating Conditional Probabilities for Geostatistical Modeling

Sahyun Hong and Clayton V. Deutsch

Centre for Computational Geostatistics
Department of Civil and Environmental Engineering
University of Alberta

The data available for geostatistical modeling can often be divided into (1) direct measurements of the primary variable being predicted, and (2) secondary data sources such as geological features and geophysical measurements that are of different variables at different scale. The secondary data are calibrated to the variable being predicted. Often, there is no explicit multivariate distribution of the primary variables with all secondary data; thus, there is a need to directly integrate prior global information, sparse primary data and calibrated secondary data. This paper presents methodologies to integrate conditional probabilities from multiple data sources for categorical variable estimation. Consider the unknown data event A given information sources B and C. The ultimate goal could be summarized as the estimation of P(A|B,C) over all grids and then P(A|B,C) is used for simulation of event A. Each conditional probabilities, P(A|B) and P(A|C), can be evaluated instead of jointly modeling of P(A|B,C). These two conditional probabilities are partially correlated each other since data source B and C are informative (correlated) to the unknown data event A. Plausible combining method must consider data redundancy inherent among data sources. A new combining algorithm of P(A|B) and P(A|C) is developed and evaluated by synthetic examples. Comparing with traditional integration ways, test example shows that the proposed method has better performance than traditional combining methods and considering data redundancy is critical to final estimation results

Introduction

Subsurface models of lithology are often poorly constrained due to lack of dense well control. Although limited in vertical resolution, exhaustive secondary data usually provide valuable information regarding the lateral variations of lithology. Co-kriging approach relies on a generalized linear regression model, which is inadequate when combining lithology indicator variables and continuous secondary attributes. Instead, a new method uses to combine each calibrated conditional probability to construct a posteriori distribution of lithology at each location. The posterior distribution combines a local prior distribution obtained by indicator kriging or training images with a function representing the secondary likelihood of the lithofacies.

Our objective is to construct lithologic subsurface models by combining observations of lithology in wells with secondary attributes data, assumed to be related to lithology. To simplify our discussion, we consider only two lithoclasses and refer to them as sand and shale. We define a binary lithology indicator variable:

$$k(\mathbf{u}) = \begin{cases} 1, \text{ if location } \mathbf{u} \in \text{ sand} \\ 2, \text{ if location } \mathbf{u} \in \text{ shale} \end{cases}$$

A sand/shale indicator samples are specified by the vector $\vec{i} = (i_1, ..., i_n)$, where random variable i has either 1 or 2 at sampled location and n is the number of neighborhood indicator samples. Associated with each location, there is a secondary attribute data y_j, j = 1,…,m which is assumed to provide indirect and imperfect information about the local lithology. For the entire model secondary data vector is denoted by $\vec{y} = (y_1(\mathbf{u}), ..., y_m(\mathbf{u}))$, $\mathbf{u} \in$ area and m is the number of secondary data variables.

Incorporating of several secondary variables requires identification of the local posterior distributions at estimation location $\mathbf{u}$, $P(k(\mathbf{u}) \mid \vec{i}, y_1(\mathbf{u}), ..., y_m(\mathbf{u}))$ involving the secondary data vector ($y_1(\mathbf{u})$,…,$y_m(\mathbf{u})$)

at co-location **u**. Only co-located secondary variables are retained to be used as secondary information and this is equivalent to Markov-type screening assumption. Traditionally, some form of co-kriging ha been used to estimate these distributions. Unfortunately, the linear data combination underlying the co-kriging technique is inadequate when mixing discrete and continuous variables, such as lithology indicator data and seismic attributes. We decomposed the required posterior distribution at each location through Bayesian relations:

$$P^* = P(k(\mathbf{u}) \mid \vec{i}, y_1(\mathbf{u}), ..., y_m(\mathbf{u}))$$
$$= P(k(\mathbf{u}) \mid \vec{i}) P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}), \vec{i}) \frac{P(\vec{i})}{P(\vec{i}, y_1(\mathbf{u}), ..., y_m(\mathbf{u}))}$$

The second term $P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}), \vec{i})$ can be summarized as $P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}))$ which explains the estimated lithology at estimation location **u** mostly influences the co-located secondary attributes. It is fairly reasonable summation that secondary attributes at location **u** is closely related to lithology type at location **u** rather than related to surrounding $\vec{i}$.

Thus, the posterior distribution follows as:

$$P^* = P(k(\mathbf{u}) \mid \vec{i}) P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u})) C$$

where, $C = P(\vec{i}) / P(\vec{i}, y_1(u), ..., y_m(u))$

Unknown constant term C is independent of lithology estimation $k(\mathbf{u})$ and is therefore not required. The first term $P(k(\mathbf{u}) \mid \vec{i})$ is called a prior distribution of $k(\mathbf{u})$ which reflects the spatial interdependence between the lithologic variables and can be calculated using either indicator kriging or training image. The second term $P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}))$ is referred to secondary likelihood distribution. Specifically, $P(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u})|k(\mathbf{u})=1)$ and $P(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u})|k(\mathbf{u})=2)$ gives the likelihood of observing sand and shale respectively at location **u**, where the measured secondary attributes is equal to $(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u}))$.

Methodology

Provided that secondary variable $(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u}))$ follows jointly Gaussian distribution, this likelihood term can directly calculated from the lithoclass-condtional normal distributions:

$$P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}) = 1) = \exp\left(-\frac{1}{2} (\mathbf{Y}(\mathbf{u}) - \boldsymbol{\mu}_{k=1})^t \sum\nolimits_{k=1}^{-1} (\mathbf{Y}(\mathbf{u}) - \boldsymbol{\mu}_{k=1}) \right)$$

and

$$P(y_1(\mathbf{u}), ..., y_m(\mathbf{u}) \mid k(\mathbf{u}) = 2) = \exp\left(-\frac{1}{2} (\mathbf{Y}(\mathbf{u}) - \boldsymbol{\mu}_{k=2})^t \sum\nolimits_{k=2}^{-1} (\mathbf{Y}(\mathbf{u}) - \boldsymbol{\mu}_{k=2}) \right)$$

$\mathbf{Y}(\mathbf{u})$ vector is secondary attributes vector $(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u}))$ at estimation location **u**. $\Sigma_{k=1}$ and $\mu_{k=1}$ are covariance and mean vector of secondary variables where sampled as k=1. As a same way, $\Sigma_{k=2}$ and $\mu_{k=2}$ are covariance and mean vector of secondary variables where sampled as k=2. Likelihood probability can be calculated under joint Gaussian distribution, however, secondary variables rarely show jointly Gaussian distribution in practice. Thus, we viewed the estimation of the likelihood $P(y_1(\mathbf{u}),\ldots,y_m(\mathbf{u})|k(\mathbf{u}))$ as the combination of each conditional probability through the integration model:

$$\text{Likelihood} = \Phi[P(y_1(\mathbf{u})|k(\mathbf{u})),\ldots, P(y_m(\mathbf{u})|k(\mathbf{u}))]$$

where $\Phi(\bullet)$ is the integration model.

In this section, we introduced the traditional integration model and proposed a new model to combine conditional probability.

Permanence of Ratios (PR-model)

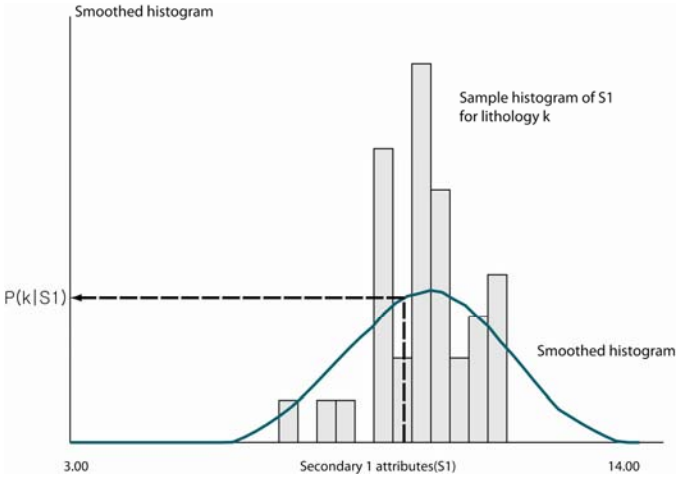
This method is based on the fact that ratios of information increments is more stable than the increments themselves. To simplify the notations, we assumed two secondary attributes at estimation location u simply as S1 and S2, and lithology as k at location u.

Previous notation	Changed notation
$y_1(\mathbf{u})$ and $y_2(\mathbf{u})$	S_1 and S_2
$k(\mathbf{u})$	k

Permanence of Ratios (noted as PR from here) gives us the following:

$$a = \frac{1-P(k)}{P(k)}, b = \frac{1-P(k \mid S_1)}{P(k \mid S_1)}, c = \frac{1-P(k \mid S_2)}{P(k \mid S_2)}$$
$$x = \frac{1-P(k \mid S_1, S_2)}{P(k \mid S_1, S_2)}$$

,where $P(k)$ is a global proportions and $P(k|S_1)$ and $P(k|S_2)$ are litho-class conditional probability. One simple way to calculate conditional probability is for using histogram smoothing technique. Below figure illustrates how to obtain litho-class conditional probability. Sample histogram of S_1 is first built for the lithology k and smoothed histogram is estimated based on sample histogram. For a specific location, probability of lithology k given S_1 attribute at the location is obtained using smoothed line.



The PR amounts to assume

$$\frac{x}{b} \simeq \frac{c}{a}$$

which is interpreted as that information incremental contribution of data S_2 to knowledge of lithology k is the same after or before knowing S_1 (see the details in reference [4]). PR-model provides the conditional probability of lithology k given all secondary data such as:

$$P_{PR}(k \mid S_1, S_2) = \frac{1}{1+x} = \frac{\dfrac{1-P(k)}{P(k)}}{\dfrac{1-P(k)}{P(k)} + \left(\dfrac{1-P(k \mid S_1)}{P(k \mid S_1)}\right)\left(\dfrac{1-P(k \mid S_2)}{P(k \mid S_2)}\right)}$$

$$= \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$$

Interestingly, estimated probability $P_{PR}(k|S_1,S_2)$ using permanence of ratio is exactly same as the estimated probability using conditional independence assumption, referred to $P_{CI}(k|S_1,S_2)$. This equivalence of PR and CI is verified in Appendix A.

Tau-model

Tau-model is proposed to consider the dependence among data source. Tau-model introduce some S_1-dependence is to set the prior-to-S1 contribution c/a to a power τ, which depends on both S_1 and S_2:

$$\frac{x}{b} \simeq \left(\frac{c}{a}\right)^{\tau(S_1,S_2)}$$

where τ weight control the contribution of S_2 to S_1. Tau-model can be viewed as a permanence of ratios model that imposes τ exponent on each conditional probability.

$$x \simeq b\left(\frac{c}{a}\right)^{\tau}$$

where, $x = P(\bar{k} \mid S_1, S_2) / P(k \mid S_1, S_2)$

The conditional probability $P(k|S_1,S_2)$ by Tau-model is following,

$$P_{Tau}(k \mid S_1, S_2) = \frac{1}{1+x}$$

$$= \frac{1}{1+b\left(\dfrac{c}{a}\right)^{\tau}} = \frac{1}{1+\dfrac{P(\bar{k} \mid S_1)}{P(k \mid S_1)}\left(\dfrac{\dfrac{P(\bar{k} \mid S_2)}{P(k \mid S_2)}}{\dfrac{P(\bar{k})}{P(k)}}\right)^{\tau}} = \frac{1}{1+\dfrac{P(\bar{k} \mid S_1)}{P(k \mid S_1)} \cdot \dfrac{P(\bar{k} \mid S_2)^{\tau} P(k)^{\tau}}{P(k \mid S_2)^{\tau} P(\bar{k})^{\tau}}}$$

$$= \frac{1}{1+\dfrac{P(\bar{k} \mid S_1)}{P(k \mid S_1)} \cdot \dfrac{P(\bar{k} \mid S_2)^{\tau} P(k)^{\tau}}{P(k \mid S_2)^{\tau} P(\bar{k})^{\tau}}} = \frac{1}{1+\dfrac{P(S_1 \mid \bar{k})P(\bar{k})/P(S_1)}{P(S_1 \mid k)P(k)/P(S_1)} \cdot \dfrac{P(k)^{\tau} P(S_2 \mid \bar{k})^{\tau} P(\bar{k})^{\tau} / P(S_2)^{\tau}}{P(\bar{k})^{\tau} P(S_2 \mid k)^{\tau} P(k)^{\tau} / P(S_2)^{\tau}}}$$

$$= \frac{P(k)P(S_1 \mid k)P(S_2 \mid k)^{\tau}}{P(k)P(S_1 \mid k)P(S_2 \mid k)^{\tau} + P(\bar{k})P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})^{\tau}}$$

This probability induced by Tau-model is exactly same as $P_{PR}(k|S_1,S_2)$ unless τ weight is considered (τ = 1).

$$P_{PR}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$$

$$P_{Tau}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)^{\tau}P(k)}{P(S_1 \mid k)P(S_2 \mid k)^{\tau}P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})^{\tau}P(\bar{k})}$$

This result implies that data inter-dependency effect can be considered as assigning exponential weights onto each conditional probability. In tau-model, τ weights are independent of lithology type k which indicates $P(S_2 \mid k)$ and $P(S_2 \mid \bar{k})$ has same τ weight such as $P(S_2 \mid k)^{\tau}$ and $P(S_2 \mid \bar{k})^{\tau}$. Simple way to find appropriate τ weight is to use linear correlation of secondary S_1 and S_2, which estimates τ as,

$$\tau = 1 - \rho(S_1, S_2)$$

S_2 information is completely ignored if $\rho(S_1,S_2)=1.0$ ($\tau=0 \rightarrow P(S_2|k)^{\tau}=1$), which indicates full dependence between S_1 and S_2. Thus, $P_{\text{Tau}}(k|S_1,S_2)$ becomes $P(k|S_1)$ and there is no information update through incorporating S_2. S_2 information is completely utilized if $\rho(S_1,S_2)=0$ ($\tau=1 \rightarrow P(S_2|k)^{\tau}= P(S_2|k)$), which means full independence between S_1 and S_2. In this case, S_2 plays as completely new information and $P_{\text{tau}}(k| S_1,S_2)$ reverts to $P_{\text{PR}}(k|S_1,S_2)$.

Lamda-model

A new method that combines conditional probability is proposed and referred as lamda-model. Lamda-model incorporates data dependence weights, say, (λ_1, λ_2):

$$P(S_1, S_2 \mid k) \simeq P(S_1 \mid k)^{\lambda_1} P(S_2 \mid k)^{\lambda_2}$$

Above decomposition with λ weights produces the below relations,

$$P_{\text{lamda}}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)^{\lambda_1} P(S_2 \mid k)^{\lambda_2} P(k)}{P(S_1 \mid k)^{\lambda_1} P(S_2 \mid k)^{\lambda_2} P(k) + P(S_1 \mid \bar{k})^{\lambda_1} P(S_2 \mid \bar{k})^{\lambda_2} P(\bar{k})}$$

Or (λ_1, λ_2) weights can be decided to be dependent on lithology type k, $(\lambda_1^k, \lambda_2^k)$ then,

$$P_{\text{lamda}}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)^{\lambda_1^k} P(S_2 \mid k)^{\lambda_2^k} P(k)}{P(S_1 \mid k)^{\lambda_1^k} P(S_2 \mid k)^{\lambda_2^k} P(k) + P(S_1 \mid \bar{k})^{\lambda_1^{\bar{k}}} P(S_2 \mid \bar{k})^{\lambda_2^{\bar{k}}} P(\bar{k})}$$

PR-model, Tau-model and Lamda-model have similar forms except how to consider data redundancy among data sources. See the below table.

Model	Probabilistic form of combining probability
PR-model	$P_{PR}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$
Tau-model	$P_{Tau}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)^{\tau}P(k)}{P(S_1 \mid k)P(S_2 \mid k)^{\tau}P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})^{\tau}P(\bar{k})}$
Lamda-model	$P_{\text{lamda}}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)^{\lambda_1^k} P(S_2 \mid k)^{\lambda_2^k} P(k)}{P(S_1 \mid k)^{\lambda_1^k} P(S_2 \mid k)^{\lambda_2^k} P(k) + P(S_1 \mid \bar{k})^{\lambda_1^{\bar{k}}} P(S_2 \mid \bar{k})^{\lambda_2^{\bar{k}}} P(\bar{k})}$

In lamda-model, data dependence weights, (λ_1,λ_2) or $(\lambda_1^k, \lambda_2^k)$ measures redundancy inherent among the secondary data sources S_1 and S_2. Choosing appropriate redundancy weights is an essential part since redundancy measures, (λ_1,λ_2), have a great impact on the secondary likelihood distribution and consequently the final updated probability.

Method 1

Let us consider two secondary variable, S_1 and S_2, and two lithology k=1(sand) and =2(shale). In the method 1, redundancy weights $(\lambda_1^k, \lambda_2^k)$ are estimated by total probability theorem such as,

$$\sum_{k=1}^{2} P(S_1, S_2 \mid k) P(k) = P(S_1, S_2)$$

$$\sum_{k=1}^{2} P(S_1 \mid k)^{\lambda_1^k} P(S_2 \mid k)^{\lambda_2^k} P(k) = P(S_1, S_2)$$

Left-hand-side term is iteratively estimated with $(\lambda_1^k, \lambda_2^k)$ and then optimal $(\lambda_1^k, \lambda_2^k)$ is kept when the difference from $P(S_1,S_2)$ falls in the threshold. Obtained redundancy weights are category dependent, $(\lambda_1^k, \lambda_2^k)$, k = 1,2.

Method 2

In the method 2, data redundancy weights are estimated using primary samples and optimality criterion which is to minimize the square errors. We know the true probability of lithology k= 1(sand) at the primary sample location where true lithology is k=1 is exactly 1. Probability of the lithology (k=1) conditioned to all secondary variables, therefore, should be close to 1 and $(\lambda_1^k, \lambda_2^k)$ are estimated to minimize the difference between true probability and approximated probability with $(\lambda_1^k, \lambda_2^k)$. Figure-1 illustrates simple example of exact probability at data locations.

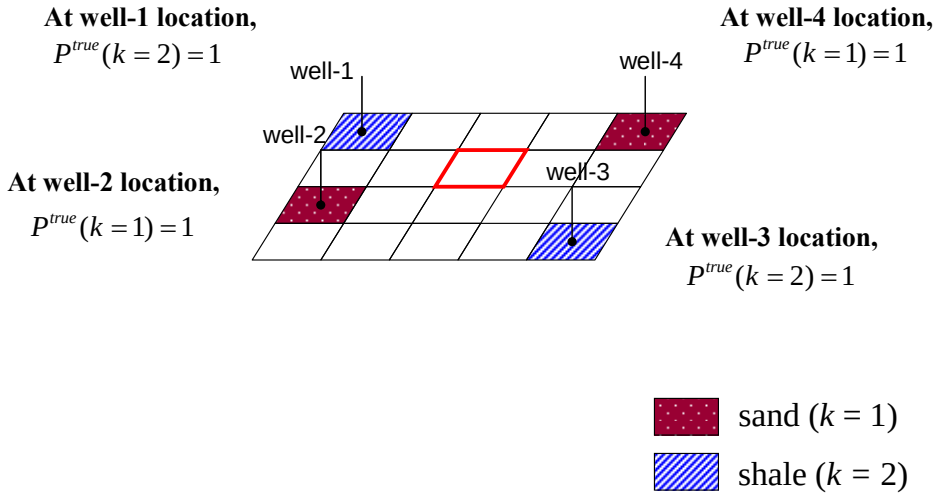


Figure-1: Simple example for explaining method 1.

Estimated probability of lithology k=1(sand) at well-2 and -4 should be close or equal to 1.0 since lithology at well-2 and -4 is sampled as k=1(sand).

$$P(k=1 \mid S_1(\mathbf{u}), S_2(\mathbf{u})) = 1.0$$
$$\simeq P(S_1(\mathbf{u}) \mid k=1)^{\lambda_1^{k=1}} P(S_2(\mathbf{u}) \mid k=1)^{\lambda_2^{k=1}} P(k=1) \frac{1}{P(S_1(\mathbf{u}), S_2(\mathbf{u}))} \quad (1)$$

,where location **u** represents well-2($\mathbf{u}_2$) and well-4($\mathbf{u}_4$) location.

As a same way, estimated probability of lithology *k*=2(shale) at well-1 and -3 locations should be close or equal to 1.0.

$$P(k=2 \mid S_1(\mathbf{u}), S_2(\mathbf{u})) = 1.0$$
$$\simeq P(S_1(\mathbf{u}) \mid k=2)^{\lambda_1^{k=2}} P(S_2(\mathbf{u}) \mid k=2)^{\lambda_2^{k=2}} P(k=2) \frac{1}{P(S_1(\mathbf{u}), S_2(\mathbf{u}))} \quad (2)$$

,where location **u** includes well-1($\mathbf{u}_1$) and well-3($\mathbf{u}_3$) location.

$S_1(\mathbf{u})$ and $S_2(\mathbf{u})$ are secondary attributes at sample location **u**. To estimate ($\lambda_1^{k=1}, \lambda_2^{k=1}$), equation (1) is built as log-linear form by taking logarithm for both sides then,

$$\begin{pmatrix} \log\left(P(S_1(\mathrm{u}_2) \mid k=1\right) & \log\left(P(S_2(\mathrm{u}_2) \mid k=1\right) \\ \log\left(P(S_1(\mathrm{u}_4) \mid k=1\right) & \log\left(P(S_1(\mathrm{u}_4) \mid k=1\right) \end{pmatrix} \begin{pmatrix} \lambda_1^{k=1} \\ \lambda_2^{k=1} \end{pmatrix} = \begin{pmatrix} \log\left(\frac{P(S_1(\mathrm{u}_2), S_2(\mathrm{u}_2))}{P(k=1)}\right) \\ \log\left(\frac{P(S_1(\mathrm{u}_4), S_2(\mathrm{u}_4))}{P(k=1)}\right) \end{pmatrix}$$

The first two lines in the matrix involve exactitude of probability at well-2 and -4. The above matrix form is shown as matrix,

$$\mathbf{L} \bullet \boldsymbol{\lambda} = \mathbf{D}$$

One way to solve the matrix equation is least square solution which is

$$\mathbf{L} \bullet \boldsymbol{\lambda} = \mathbf{D}$$
$$\boldsymbol{\lambda} = \arg\min_{\boldsymbol{\lambda}} \| \mathbf{L} \bullet \boldsymbol{\lambda} - \mathbf{D} \|^2$$

Thus, lamda for lithology *k*=1 is $\boldsymbol{\lambda}^{k=1} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{D}$.

To estimate ($\lambda_1^{k=2}, \lambda_2^{k=2}$), equation (2) is taken by logarithm as well,

$$\begin{pmatrix} \log\left(P(S_1(\mathrm{u}_1) \mid k=2\right) & \log\left(P(S_2(\mathrm{u}_1) \mid k=2\right) \\ \log\left(P(S_1(\mathrm{u}_3) \mid k=2\right) & \log\left(P(S_1(\mathrm{u}_3) \mid k=2\right) \end{pmatrix} \begin{pmatrix} \lambda_1^{k=2} \\ \lambda_2^{k=2} \end{pmatrix} = \begin{pmatrix} \log\left(\frac{P(S_1(\mathrm{u}_1), S_2(\mathrm{u}_3))}{P(k=2)}\right) \\ \log\left(\frac{P(S_1(\mathrm{u}_1), S_2(\mathrm{u}_3))}{P(k=2)}\right) \end{pmatrix}$$

Lamda weights for lithology *k*=2 is calculated using least square solution,

$$\boldsymbol{\lambda}^{k=2} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{D}$$

Synthetic Examples

Synthetic test data is applied to evaluate the considered methods. We generated two secondary data sets that have 100 × 100 exhaustively sampled values. Different correlation between secondary 1 and secondary 2 is tested; high linear correlation, low linear correlation and non-linear correlation. In this test,

primary indicator information is not integrated because our objective is to evaluate the discussed methodologies combining secondary data sets. The integrated results are evaluated based on quantitative goodness measure; closeness to true lithology and entropy. Closeness to true value is summarized by:

$$C_k = E\{P(\mathbf{u}_\alpha;k)| \text{ true} = k\},\ k = 1,2$$

The closeness measures C_k, k = 1,2 are easily interpreted relative to the global proportions $P(k)$, k = 1,2 i.e.,

$$C_k^{rel} = \frac{C_k - P(k)}{P(k)}, k = 1,2$$

Classically defined Shannon entropy is used as another measure of goodness:

$$S_k = E\left\{-\sum_{k'}^{2} P_{k'}^{*} \ln(P_{k'}^{*}) \mid \text{true} = k\right\}, k = 1,2$$

where $P_{k'}^{*}$ is integrated probability of lithology k'.

Entropy measure can be interpreted as the uncertainty of the predicted probability. The entropy would be 0.0 in the ideal case of complete information.

Figure-2 illustrates correlation between synthetic secondary 1 and secondary 2 data sets. Scatter plot of exhaustive data shows high linear correlation above 0.95. 700 secondary data is extracted from 10000 (100×100 image) data values and they are set to indicate lithology 1. 600 secondary data is extracted from 10000 data values and they are set to indicate lithology 2. Thus, we prepared global proportion of lithology 1 is 700/(700+600)=0.54 and global proportion of lithology 2 is 600/(700+600)=0.46. Right scatter plot of Figure-2 represents the correlation between extracted secondary 1 and secondary 2 corresponding to the lithology types. Low values in secondary data are closely related to lithology 1 and high values in secondary data sets closely related to lithology 2. We set secondary data to be highly redundant each other.

First of all, conditional probabilities given secondary 1 and secondary 2, $P(k|S_1)$ and $P(k|S_2)$, are calibrated separately. This calibration was performed using histogram smoothing technique. Probability maps are shown in the upper part in Figure-3 and calibrated probabilities looks similar each other since secondary 1 and secondary 2 data sets are highly correlated. Integrated probability maps are shown in the lower part in Figure-3 with the measurement of closeness and entropy. In lamda-model, two approaches to estimate λ weights are applied. In tau-model, τ weights are extracted from the linear correlation between secondary 1 and 2, τ = 1.0 – 0.939 for lithology 1 and τ = 1.0 – 0.959 for lithology 2. Permanence of ratios model produced the worst result and lamda-model provides the best integrated result in terms of quantified goodness. Integrated result using tau-model shows smaller closeness and larger entropy than the result using lamda-model. However, tau-model shows much better performance than PR-model (158% improvement in closeness and 83% improvement in entropy) just as considering linear correlation between secondary data sets.

Secondly, low correlated secondary data sets are tested. Figure-4 represents low correlation between synthetic secondary 1 and 2. In this second example, we prepared two secondary data sets which show little redundancy. Integrated results are shown in the Figure-4. Although PR-model has the worst result and lamda-model has the best integrated result, one cannot notice significant improvements among three integrated results in terms of goodness measure.

Another example is the case of nonlinear relations between secondary data as shown in Figure-6. Scatter plot between exhaustive secondary 1 and 2 shows non-linear relations. Extracted secondary 1 and 2 samples corresponding to lithology 1 and 2 has non-linear relations as well. For the lithology 1, linear correlation between secondary 1 and 2 is 0.239 and for the lithology 2, linear correlation between 1 and 2 is 0.659. These linear correlations are used in the tau-model to estimate τ weights. Integrated results are shown in the Figure-7 and closeness and entropy are quantified. Tau-model does not show better performance than PR-model even though tau-model considered redundancy weight τ. Lamda-model gives us the best integrated results in the case of non-linear secondary data relations.

Discussions and Conclusions

Bayesian theorem with Markov-type screening assumption allows a local posterior probability to be decomposed into the product of prior and likelihood distribution. This paper focuses on how to obtain likelihood distribution because it is not easy to estimate likelihood unless all secondary variables follow jointly Gaussian distributions. We interpreted the estimation of likelihood as the combination of single conditional probability through the plausible integration model. Permanence of ratios and tau-model are introduced as the plausible integration model. Lamda-model is proposed as a new integration model and compared with PR and tau-model. Synthetic test data is applied for the evaluation and we observed choosing the appropriate redundancy weights have a great impact on the integrated results. In terms of quantitative goodness measures, lamda-model is the best integration model.

References

C. V. Deutsch. Geostatistical reservoir modeling. Oxford University Press, New York, 2002.

C. V. Deutsch. A Short Note on Cross Validation of Facies Simulation Methods. In Report 1, Centre for Computational Geostatistics, Edmonton, AB, Canada, 1998.

R. L. Winkler. The consensus of subjective probability distributions. Management Science, 15(2):B-61 to B75, 1968.

A. G. Journel. Combining Knowledge from Diverse Sources: An Alternative to Traditional Data Independence Hypotheses. Mathematical Geology, 34(5):573-596, 2002

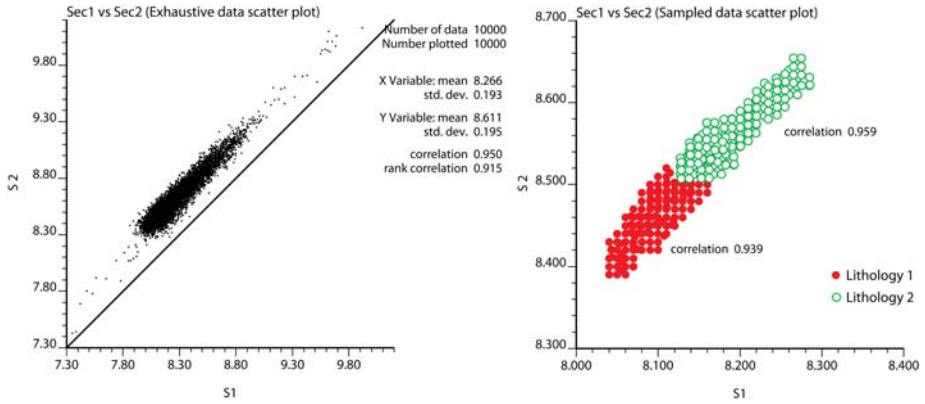


Figure-2: Correlation between secondary 1 and secondary 2 synthetic data (left) and correlation between sampled secondary 1 and secondary 2 (right). Sampled secondary values are colored according to lithology type (• lithology 1 o lithology 2).

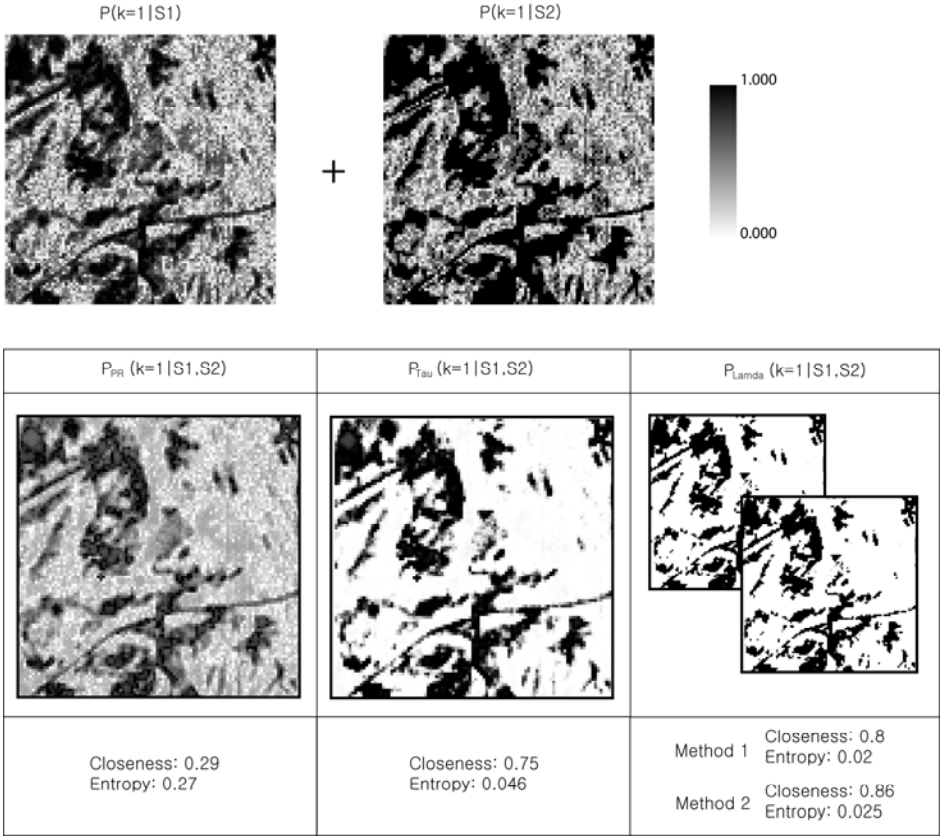


Figure-3: Integrated probability of lithology 1 using permanence of ratios, tau-model, and lamda-model with method 1 and method 2. Three integrated methods are evaluated by closeness and entropy as shown in the bottom of the maps. Closeness and entropy are average values over lithology 1 and 2.

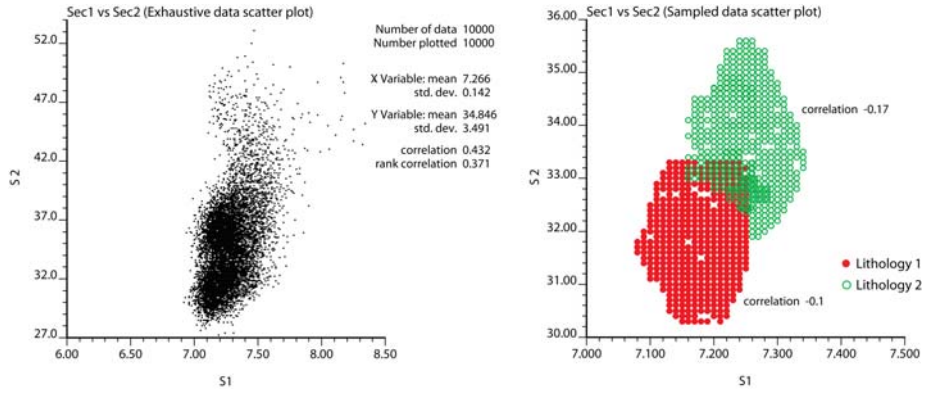


Figure-4: Correlation between secondary 1 and secondary 2 synthetic data (left) and correlation between sampled secondary 1 and secondary 2 (right). Sampled secondary values are colored according to lithology type (•: lithology 1, o: lithology 2).

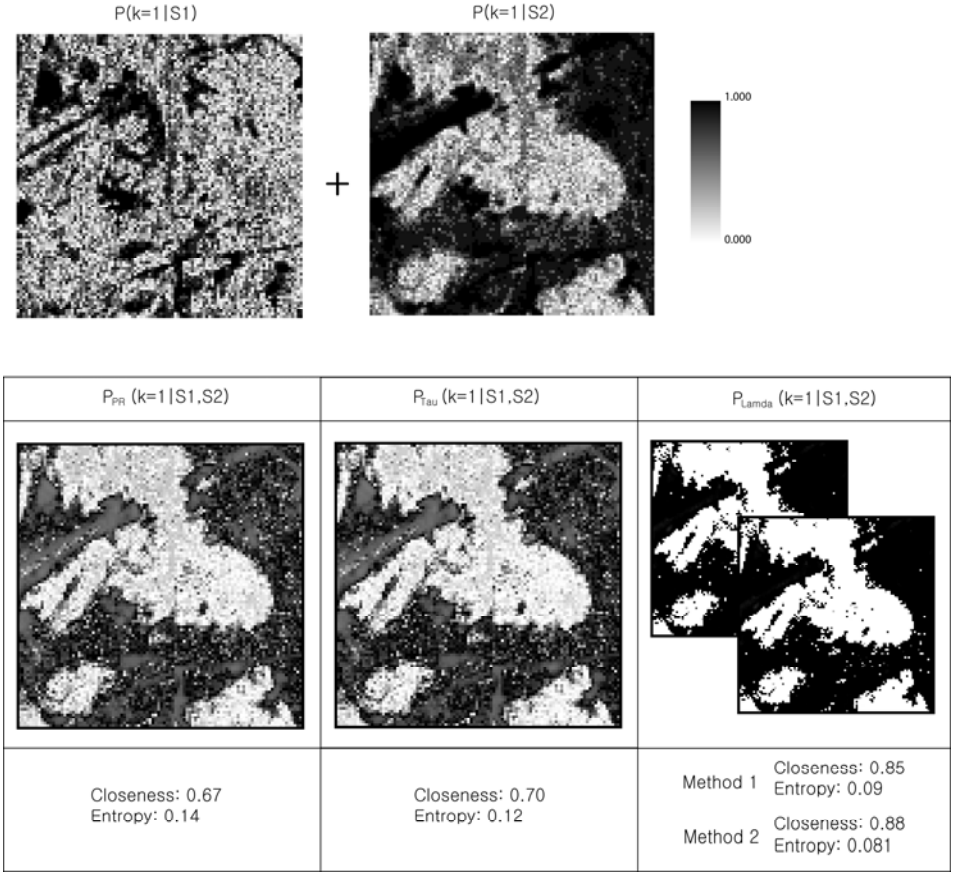


Figure-5: Integrated probability of lithology 1 using permanence of ratios, tau-model, and lamda-model with method 1 and method 2. Three integrated methods are evaluated by closeness and entropy as shown in the bottom of the maps. Closeness and entropy are average values over lithology 1 and 2.

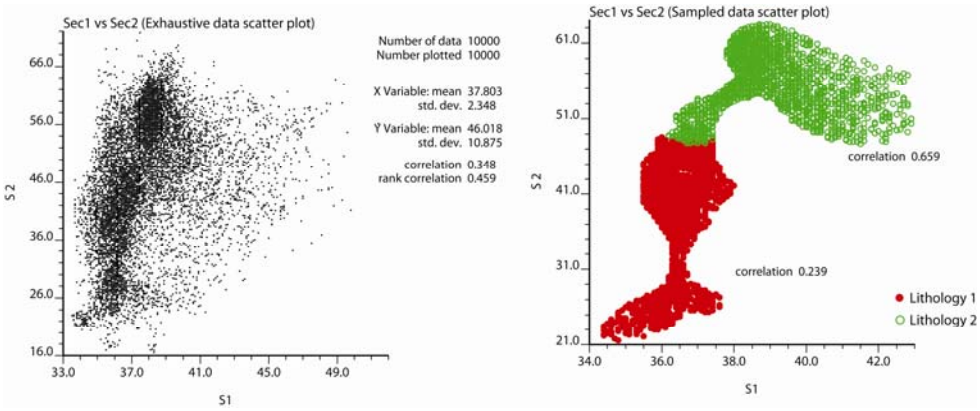


Figure-6: Correlation between secondary 1 and secondary 2 synthetic data (left) and correlation between sampled secondary 1 and secondary 2 (right). Sampled secondary values are colored according to lithology type (•: lithology 1, o: lithology 2).

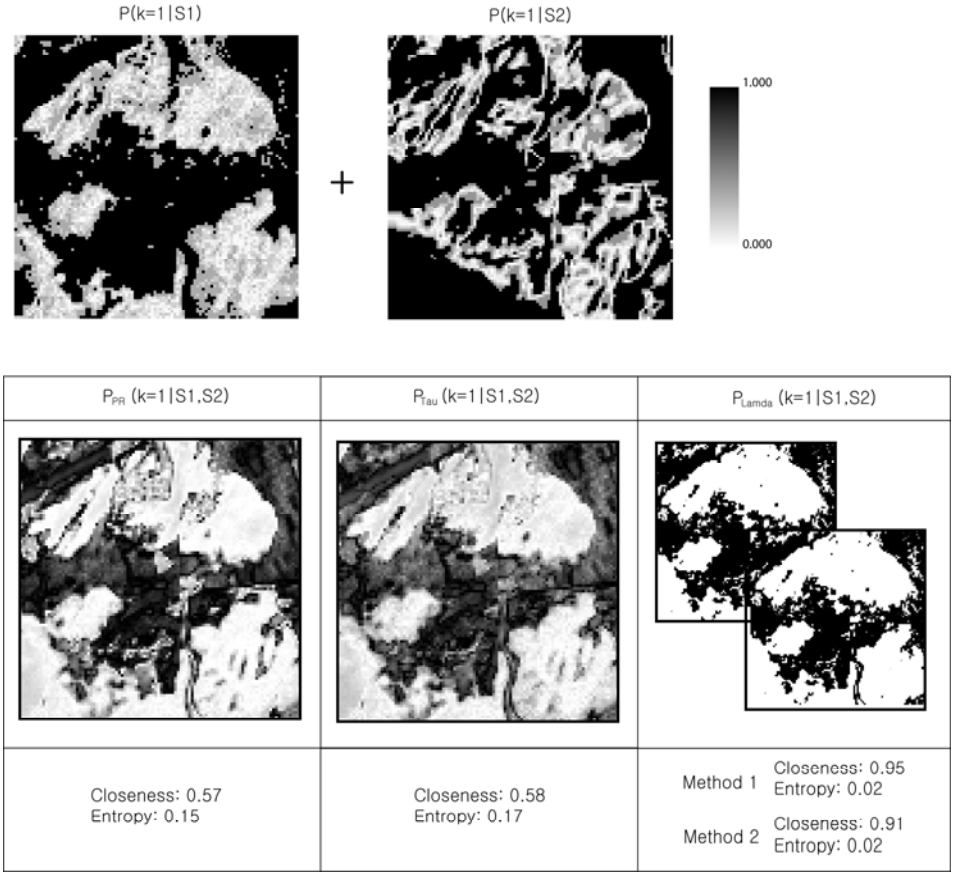


Figure-7: Integrated probability of lithology 1 using permanence of ratios, tau-model, and lamda-model with method 1 and method 2. Three integrated methods are evaluated by closeness and entropy as shown in the bottom of the maps. Closeness and entropy are average values over lithology 1 and 2.

Appendix

Conditional probability P(k|S_1,S_2) is shown below using permanence of ratios approximation.

$$P(k \mid S_1, S_2) = \frac{\dfrac{P(\bar{k})}{P(k)}}{\dfrac{P(\bar{k})}{P(k)} + \dfrac{P(\bar{k} \mid S_1)}{P(k \mid S_1)} \dfrac{P(\bar{k} \mid S_2)}{P(k \mid S_2)}}$$

Bayesian relations make the following,

$$P(k \mid S_1, S_2) = \frac{\dfrac{P(\bar{k})}{P(k)}}{\dfrac{P(\bar{k})}{P(k)} + \dfrac{\dfrac{P(S_1 \mid \bar{k})P(\bar{k})}{P(S_1)}}{\dfrac{P(S_1 \mid k)P(k)}{P(S_1)}} \dfrac{\dfrac{P(S_2 \mid \bar{k})P(\bar{k})}{P(S_2)}}{\dfrac{P(S_2 \mid k)P(k)}{P(S_2)}}} = \frac{\dfrac{P(\bar{k})}{P(k)}}{\dfrac{P(\bar{k})}{P(k)} + \dfrac{P(S_1 \mid \bar{k})P(\bar{k})}{P(S_1 \mid k)P(k)} \dfrac{P(S_2 \mid \bar{k})P(\bar{k})}{P(S_2 \mid k)P(k)}}$$

$$= \frac{1}{1 + \dfrac{P(S_1 \mid \bar{k})}{P(S_1 \mid k)} \dfrac{P(S_2 \mid \bar{k})}{P(S_2 \mid k)} \dfrac{P(\bar{k})}{P(k)}}$$

Thus, the estimated probability P(k|S_1,S_2) using permanence of ratios is summarized as,

$$P_{PR}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$$

For m secondary variables, the conditional probability given all secondary variables is

$$P_{PR}(k \mid S_1, ..., S_m) = \frac{P(k)\prod_{i=1}^{m} P(S_i \mid k)}{P(k)\prod_{i=1}^{m} P(S_i \mid k) + P(\bar{k})\prod_{i=1}^{m} P(S_i \mid \bar{k})}$$

Now, let us derive P(k|S_1,S_2) using conditional independence assumption.

Conditional independence assumption enables us to decompose P(k|S_1,S_2) into

$$P(k \mid S_1, S_2) = \frac{P(S_1, S_2 \mid k)P(k)}{P(S_1, S_2)} = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1, S_2)}$$

Let us denote 1/P(S_1,S_2) as constant unknown term C that is independent of lithology type k then we have,

$$P(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1, S_2)} = P(S_1 \mid k)P(S_2 \mid k)P(k)C$$

Also, probability $P(\bar{k} \mid S_1, S_2)$ is obtained by conditional independence,

$$P(\bar{k} \mid S_1, S_2) = \frac{P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}{P(S_1, S_2)} = P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})C$$

Unknown constant term C is the same when estimating both $P(\bar{k} \mid S_1, S_2)$ and $P(k \mid S_1, S_2)$. The basic probability property that sum of $P(\bar{k} \mid S_1, S_2)$ and $P(k \mid S_1, S_2)$ is equal to 1 gives us,

$$P(k \mid S_1, S_2) + P(\bar{k} \mid S_1, S_2) = 1$$
$$P(S_1 \mid k)P(S_2 \mid k)P(k)C + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})C = 1$$
$$C = \frac{1}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$$

Finally, substitute C term into $P(S_1 \mid k)P(S_2 \mid k)P(k)C$ then we have,

$$P_{CI}(k \mid S_1, S_2) = \frac{P(S_1 \mid k)P(S_2 \mid k)P(k)}{P(S_1 \mid k)P(S_2 \mid k)P(k) + P(S_1 \mid \bar{k})P(S_2 \mid \bar{k})P(\bar{k})}$$

For m multiple secondary variables, the probability $P(k \mid S_1, ..., S_m)$ and $P(\bar{k} \mid S_1, ..., S_m)$ can be shown as,

$$P(k \mid S_1, ..., S_m) = P(k)P(S_1 \mid k) \cdots P(S_m \mid k)C'$$
$$P(\bar{k} \mid S_1, ..., S_m) = P(\bar{k})P(S_1 \mid \bar{k}) \cdots P(S_m \mid \bar{k})C'$$

And unknown constant term C' is given by $P(k \mid S_1, ..., S_m) + P(\bar{k} \mid S_1, ..., S_m) = 1$.

$$C' = \frac{1}{P(S_1 \mid k) \cdots P(S_m \mid k)P(k) + P(S_1 \mid \bar{k}) \cdots P(S_m \mid \bar{k})P(\bar{k})}$$

Substitute C' into equation $P(k \mid S_1, ..., S_m) = P(k)P(S_1 \mid k) \cdots P(S_m \mid k)C'$ then we have

$$P_{CI}(k \mid S_1, ..., S_m) = \frac{P(k)\prod_{i=1}^{m} P(S_i \mid k)}{P(k)\prod_{i=1}^{m} P(S_i \mid k) + P(\bar{k})\prod_{i=1}^{m} P(S_i \mid \bar{k})}$$

Therefore, $P_{PR}(k|S_1,\ldots,S_m)$ is equivalent to $P_{CI}(k|S_1,\ldots,S_m)$.