Data-driven Evidential Belief Predictive Modelling of Regional-Scale Geothermal Prospectivity in West Java (Indonesia)

Hendro Wibowo (1), Emmanuel John M. Carranza (2)

(1) Lab. Volcanology and Geothermal, Geology Department Institute of Technology Bandung (ITB), Bandung, Indonesia. E-mail: hendrohwibowo@yahoo.com
(2) International Institute for Geo-information Science and Earth Observation (ITC), Hengelosestraat 99, P.O. Box 6, 7500 AA Enschede, The Netherlands. E-mail: carranza@itc.nl. (Presenter of paper and to whom correspondence should be sent.)

KEY WORDS: Evidential belief functions, Dempster's rule of combination, spatial data integration, GIS.

INTRODUCTION

This paper presents a case demonstration of data-driven evidential belief functions (EBFs) to map regional-scale geothermal prospectivity.

EVIDENTIAL BELIEF FUNCTIONS

The Dempster-Shafer theory of evidence provides framework for EBFs (Dempster, 1967; Shafer, 1976). Estimation of EBFs for spatial evidence is always in relation to a proposition, which in the case is: "*This location contains a geothermal occurrence*". The EBFs are ***Bel*** (belief), ***Dis*** (disbelief), ***Unc*** (uncertainty) and ***Pls*** (plausibility). ***Bel*** and ***Pls*** represent, respectively, lower and upper probabilities that evidence supports the proposition. Thus, ***Pls*** is $\leq$***Bel***. ***Unc*** is equal to ***Pls***–***Bel*** and represents "doubt" on given evidence to support the proposition. If ***Bel***=***Pls***, then ***Unc***=0. ***Dis*** is belief that the proposition is false based on given evidence; it is equal to 1–***Pls***. Thus, ***Bel***+***Unc***+***Dis***=1. However, if ***Bel***=0, then ***Dis***=0; because there can be no disbelief if there is no belief, there can only be uncertainty. Interestingly, if ***Unc***=0, then ***Bel***+***Dis***=1, as in probability approach. The EBFs for each set of evidence are usually estimated based on expert knowledge (e.g., Moon, 1990; An et al., 1994). In this study, due to ''lack'' of expert knowledge, procedures for data-driven estimation EBFs by Carranza and Hale (2003) were adopted.

Suppose a region $\boldsymbol{T}$ with $\boldsymbol{N(T)}$ total number of unit cells or pixels and geothermal occurrences $\boldsymbol{D}$ in $\boldsymbol{N(D)}$ number of pixels. Suppose further that evidence maps $\boldsymbol{X_i}$ ($\boldsymbol{i}=\mathbf{1,2,\ldots,}\boldsymbol{n}$), with $\boldsymbol{C_{ij}}$ ($\boldsymbol{j}=\mathbf{1,2,\ldots,}\boldsymbol{m}$) classes of attributes, are available. Binary map of $\boldsymbol{D}$ is overlaid on each evidential map to determine number of $\boldsymbol{C_{ij}}$ pixels with $\boldsymbol{D}$ [$\boldsymbol{N(C_{ij}\cap D)}$] and number of $\boldsymbol{C_{ij}}$ pixels without $\boldsymbol{D}$ [$\boldsymbol{N(C_{ij})}-\boldsymbol{N(C_{ij}\cap D)}$]. The EBFs are then estimated as follows (Carranza and Hale 2003):

$$\boldsymbol{Bel_{C_{ij}}}=\frac{\boldsymbol{W_{C_{ij}D}}}{\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}\boldsymbol{W_{C_{ij}D}}}$$

where $\boldsymbol{W_{C_{ij}D}}=\dfrac{\dfrac{\boldsymbol{N(C_{ij}\cap D)}}{\boldsymbol{N(C_{ij})}}}{\dfrac{\boldsymbol{N(D)}-\boldsymbol{N(C_{ij}\cap D)}}{\boldsymbol{N(T)}-\boldsymbol{N(C_{ij})}}}$.

$$\boldsymbol{Dis_{C_{ij}}}=\frac{\boldsymbol{W_{C_{ij}\overline{D}}}}{\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}\boldsymbol{W_{C_{ij}\overline{D}}}}$$

where $\boldsymbol{W_{C_{ij}\overline{D}}}=\dfrac{\dfrac{\boldsymbol{N(C_{ij})}-\boldsymbol{N(C_{ij}\cap D)}}{\boldsymbol{N(C_{ij})}}}{\dfrac{\boldsymbol{N(T)}-\boldsymbol{N(D)}-[\boldsymbol{N(C_{ij})}-\boldsymbol{N(C_{ij}\cap D)}]}{\boldsymbol{N(T)}-\boldsymbol{N(C_{ij})}}}$.

Note that, if for $\boldsymbol{C_{ij}}$ estimated $\boldsymbol{W_{C_{ij}D}}=\mathbf{0}$, which means that $\boldsymbol{Bel_{C_{ij}}}=\mathbf{0}$, then the corresponding estimated $\boldsymbol{W_{C_{ij}\overline{D}}}$ is re-set to zero because if *Bel*=0, then *Dis*=0. Finally, *Unc* and *Pls* are estimated, respectively, as

$\boldsymbol{Unc_{C_{ij}}}=\mathbf{1}-\boldsymbol{Bel_{C_{ij}}}-\boldsymbol{Dis_{C_{ij}}}$; and

$\boldsymbol{Pls_{C_{ij}}}=\boldsymbol{Bel_{C_{ij}}}+\boldsymbol{Unc_{C_{ij}}}$.

Formulas, according to Dempster's (1968) rule, for combining maps of EBFs of two spatial evidence maps ($\boldsymbol{X_1}$ and $\boldsymbol{X_2}$) are as follows (adopted from Wright and Bonham-Carter, 1996):

$$\boldsymbol{Bel_{X_1X_2}}=\frac{\boldsymbol{Bel_{X_1}Bel_{X_2}}+\boldsymbol{Bel_{X_1}Unc_{X_2}}+\boldsymbol{Bel_{X_2}Unc_{X_1}}}{\beta}$$

$$\boldsymbol{Dis_{X_1X_2}}=\frac{\boldsymbol{Dis_{X_1}Dis_{X_2}}+\boldsymbol{Dis_{X_1}Unc_{X_2}}+\boldsymbol{Dis_{X_2}Unc_{X_1}}}{\beta}$$

$$\boldsymbol{Unc_{X_1X_2}}=\frac{\boldsymbol{Unc_{X_1}Unc_{X_2}}}{\beta}$$

where $\beta=\mathbf{1}-\boldsymbol{Bel_{X_1}Dis_{X_2}}-\boldsymbol{Dis_{X_1}Bel_{X_2}}$ to ensure that $\boldsymbol{Bel}+\boldsymbol{Unc}+\boldsymbol{Dis}=\mathbf{1}$. Only maps of EBFs of two spatial evidences are combined each time; other maps of EBFs representing $\boldsymbol{X_3},\ldots,\boldsymbol{X_n}$ are combined one after another by repeated applications of Dempster's rule. Final combination of maps of EBFs results in integrated ***Bel***, integrated ***Dis***, integrated ***Unc*** and integrated ***Pls*** for the proposition based on all given spatial evidences.

THE STUDY AREA

The present day volcanic arc in West Java is a superposition of Tertiary and Quaternary volcanoes. The volcanoes of geothermal interest are dominantly andesitic and, frequently, at least one geothermal field is associated with a volcanic complex, although only part of a complex is still volcanically active. Tertiary calcareous sediments underlie much of the Quaternary andesitic volcanics (Mahon et al., 2000), and thereby act as reservoirs. Several geothermal prospects had already been investigated and some have been exploited or being developed for exploitation. Nevertheless, further exploration is important to identify other prospective geothermal areas.

Several geological features are considered indicative of prospectivity for convective type of geothermal resources in West Java: (a) regional faults/fractures; (b) hot springs; (c) hydrothermal alteration; (d) volcanic rocks, as cap rocks; (e) regional negative gravity anomalies, representing volcano-tectonic depressions where volcanic heat sources and sedimentary reservoir rocks are present; and (f) shallow earthquake epicentres, associated with regional structures and geothermal fluid movements. From published and unpublished geological maps, separate map layers of regional faults/fractures, hot springs, volcanic centres, volcanic rocks, and shallow-earthquake epicentres were compiled. On shaded-relief SRTM DEMs, faults/fractures were interpreted to augment mapped faults/fractures. Using Landsat TM data, areas with hydrothermal alteration were mapped. Analysis of Bouguer anomaly map resulted in delineation of a regional volcano-tectonic depression. In addition, a map of geothermal occurrences was compiled.

GEOTHERMAL PROSPECTIVITY MODELLING

To estimate EBFs of geothermal prospectivity, a pixel size of 500×500 m was used in a raster-based GIS. This pixel size adequately represents each geothermal field occurrence as one, and only one, pixel. The pixel size used also means that regional-scale geothermal prospectivity is estimated for every 500×500 m cell in West Java.

An evidential map for each set of indicative geological features was prepared in two steps. First step is creation of map of distances to indicative geological features. Second step is creation of a map of proximity to indicative geological features partitioning a distance map into a number of proximity classes using 10-percentile intervals. Some of distance maps do not result in 10 proximity classes if mapped geological features represent more than 10 percentile of the area. For each proximity map, the map parameters defined for estimation of EBFs were determined.

The set of 53 known geothermal occurrences was split into two sub-sets by random selection. One sub-set consists about 75% of the 53 known occurrences and was used as training data for estimation of EBFs. The other sub-set consists about 25% of the 53 geothermal occurrences and was later used to validate model of geothermal prospectivity.

Estimates of EBFs are usually held in an attribute table associated with an evidential map. Due to lack of space here, only the results of data-driven estimation of EBFs for proximity to Quaternary volcanic centers are shown (Table 1). Within 8.1 km of Quaternary volcanic centers, values of ***Bel*** are higher and values of ***Unc*** are lower than beyond 8.1 km of Quaternary volcanic centres. This indicates that zones within 8.1 km of Quaternary volcanic centers have positive spatial association with known geothermal occurrences and are therefore prospective for geothermal occurrence. Geologically, the results suggest that heat sources of the Quaternary volcanoes and the geothermal occurrences are the same. Similar lines of interpretations of spatial and geological significance were made from the data-driven estimates of EBFs for the other sets of evidential maps.

*Table 1 – Estimated EBFs of proximity to Quaternary volcanic centers, West Java. [**N(T)**=187474], [**N(D)**=40].*

Proximity to Quaternary volcanic centres (km)	$N(C_{ij})$	$N(C_{ij} \cap D)$	Bel	Dis	Unc
0.0 – 8.1	18653	21	0.681	0.124	0.194
8.1 – 12.7	18502	5	0.088	0.125	0.786
12.7 – 17.7	19084	1	0.015	0.125	0.859
17.7 – 22.8	18543	2	0.033	0.125	0.842
22.8 – 28.0	18799	1	0.016	0.125	0.859
28.0 – 33.3	18953	4	0.067	0.125	0.808
33.6 – 40.3	18781	3	0.049	0.125	0.825
40.3 – 49.0	18678	3	0.049	0.125	0.825
49.0 – 60.4	18788	0	0.000	0.000	1.000
60.4 – 100.0	18836	0	0.000	0.000	1.000

Three sets of maps of integrated EBFs were created. One set of maps of integrated EBFs was created for all the seven evidential maps. A second set of maps of integrated EBFs was created using all but the evidential map of presence of or proximity to hypothesized volcano-tectonic depression. A third set of maps of integrated EBFs was created using all but the evidential layer of presence of or proximity to remotely-sensed clay alteration. The first set of maps of integrated EBFs was created to determine usefulness of integrating all relevant evidential data/information in mapping geothermal prospectivity. The second and third sets of maps of integrated EBFs were created to determine usefulness of derivative but relevant evidential information (e.g., hypothesized volcano-tectonic depression or remotely-sensed clay alteration) in

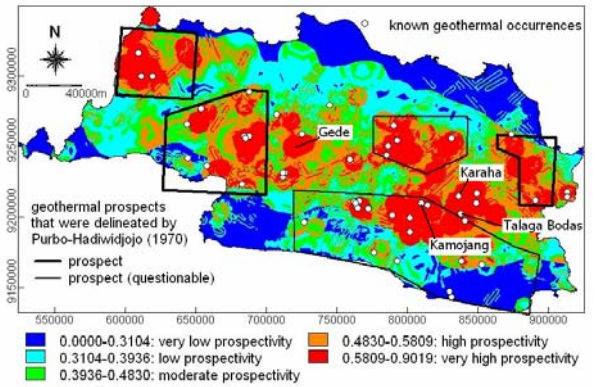

Figure 2– The best geothermal prospectivity map and boundaries of published geothermal prospect areas (Purbo-Hadiwidjojo, 1970). Locations of some well known geothermal areas are also indiated.

geothermal prospectivity mapping. Maps of integrated ***Bel*** portray spatial associations of evidences with targets (in this case geothermal occurrences), and are therefore used to classify geothermal prospectivity.

Geothermal prospectivity was categorized into five classes based on 20-percentile intervals of values of integrated ***Bel***. The 0-20 percentile of values of integrated ***Bel***. represents very low prospectivity, the 20-40 percentile represents low prospectivity, the 40-60 percentile represents intermediate prospectivity, the 60-80 percentile represents high prospectivity, and the 80-100 percentile represents very high prospectivity. The different geothermal prospectivity maps similarly indicate that zones proximal to Quaternary volcanic centres and to hot springs and zones within the hypothesized regional volcano-tectonic depression and Quaternary volcanic rocks are where geothermal fields are mostly likely to occur.

To validate the geothermal prospectivity maps, two parameters are estimated: (1) success rate; and (2) prediction rate (Table 2). Success rate is percentage of training geothermal occurrences delineated by a prospectivity class. Prediction rate is percentage of validation geothermal occurrences delineated by a prospectivity class. Validation results indicate that classification of map of integrated ***Bel*** based on all but the remotely-sensed clay alteration evidence map results in the best regional-scale geothermal prospectivity map (Figure 2).

CONCLUDING REMARKS

- Data-driven estimates of EBFs can indicate which sets of evidential data or information provide efficient and sufficient support to the proposition of regional-scale geothermal prospectivity.
- Evidential data found to provide insufficient support to the proposition of geothermal prospectivity need to be updated or methods applied to extract spatial information of interest from geoscience data need to be improved.
- The best map of regional-scale geothermal prospectivity created for West Java indicates several portions of high to very high prospective zones with no known (or reported) geothermal occurrences, which indicates usefulness of the map in guiding further geothermal exploration at higher scales.
- Results of study indicate efficacy of data-driven EBFs in regional-scale geothermal prospectivity mapping.

REFERENCES

AN, P., MOON, W. M., BONHAM-CARTER, G. F., 1994. An object-oriented knowledge representation structure for exploration data integration, Nonrenewable Resources 3, 132–145.

CARRANZA, E. J. M., HALE, M.: 2003, Evidential belief functions for data-driven geologically-constrained mapping of gold potential, Baguio district, Philippines, Ore Geology Reviews, 22, 117–132.

DEMPSTER, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics, 38: 325-339.

DEMPSTER, A.P., 1968. Generalization of Bayesian inference. Journal of the Royal Statistical Society, series B, 30: 205-247.

MAHON, T., HARVEY, C., CROSBY, D., 2000. The Chemistry of Geothermal Fluids in Indonesia and their Relationship to Water and Vapor Dominated Systems, World Geothermal Congress. International Geothermal Association, Tohoku, Kyushu, Japan.

MOON, W.M., 1990. Integration of geophysical and geological data using evidential belief function. IEEE Transactions on Geoscience and Remote Sensing, 28(4): 711-720.

PURBO-HADIWIDJOJO, M.M., 1970. A Tentative Map of the Prospective Geothermal Areas of Java and Bali, 1:1,000,000. Geological Survey of Indonesia, Bandung, Indonesia.

SHAFER, G., 1976. A Mathematical Theory of Evidence, Princeton University Press, Princeton, New Jersey, 297 pp.

Table 2 – Performace of geothermal prospectivity maps against validation geothermal occurrences ***[N(D_{val})=13]***.

Geothermal prospectivity class	Prospectivity map inclusive of all evidential maps		Prospectivity map exclusive of remotely-sensed clay alteration evidential map		Prospectivity map exclusive of hypothesized volcano-tectonic depression evidential map	
	Success rate (%)	Prediction rate (%)	Success rate (%)	Prediction rate (%)	Success rate (%)	Prediction rate (%)
Very high prospectivity	70.0	84.6	72.5	84.6	67.5	76.9
High prospectivity	20.0	7.7	12.5	15.4	20.0	23.1
Moderate prospectivity	5.0	7.7	10.0	-	7.5	-
Low prospectivity	-	-	-	-	2.5	-
Very low prospectivity	5.0	-	5.0	-	2.5	-