Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A New Bayesian Inference Methodology for Modeling Geochemical Elements in Soil with Covariates. Characterization of Lithium in South Iberian Range (Spain)

By Sergio Meseguer, PabloJuan, Ana B. Vicente & Carlos Díaz-Avalos
University Jaume I of Castellón

Abstract- When the scientific need to model geochemical elements in soil, is using geostatistical methodologies, for instance krigings, but we can use a new possibility with Bayesian Inference. The models for the analysis were specified by the authors and estimated using Bayesian inference for Gaussian Markov Random Field (GMRF) through the Integrated Nested Laplace Approximation (INLA) algorithm. The results allow us to quantify and assess possible spatial relationships between the distribution of lithium and other possible explanatory elements. Are these other elements significant to the study? We believe the methods outlined here may help to find elements such as lithium, as well as contributing to the prediction and management of new extractions or prospection in a region in order to find each chemical element.

Keywords: *bayesian inference, iberian range, lithium, soil.*
GJCST-G Classification: *C.2.1 C.2.2*

ANEWBAYESIANINFERENCEMETHODOLOGYFORMODELINGGEOCHEMICALELEMENTSINSOILWITHCOVARIATESCHARACTERIZATIONOFLITHIUMINSOUTHIBERIANRANGESPAIN

Strictly as per the compliance and regulations of:

A New Bayesian Inference Methodology for Modeling Geochemical Elements in Soil with Covariates. Characterization of Lithium in South Iberian Range (Spain)

Sergio Meseguer [α], Pablo Juan [σ], Ana B. Vicente [ρ] & Carlos Díaz-Avalos [ω]

Abstract- When the scientific need to model geochemical elements in soil, is using geostatistical methodologies, for instance krigings, but we can use a new possibility with Bayesian Inference. The models for the analysis were specified by the authors and estimated using Bayesian inference for Gaussian Markov Random Field (GMRF) through the Integrated Nested Laplace Approximation (INLA) algorithm. The results allow us to quantify and assess possible spatial relationships between the distribution of lithium and other possible explanatory elements. Are these other elements significant to the study? We believe the methods outlined here may help to find elements such as lithium, as well as contributing to the prediction and management of new extractions or prospection in a region in order to find each chemical element. The application for the modeling is to study the spatial variation in the distribution of lithium and its relationship to other geochemical elements is analyzed in terms of the different possibilities offered by geographical and environmental factors. All in all, Lithium presents many important and meaningful uses and applications such as: ceramics and glass, electrical and electronics standing out lithium ion batteries, as well as a lubricator for greases, in metallurgy, pyrotechnics, air purification, optics, organic and polymer chemistry, and medicine. This study aims to examine the distribution of lithium in sediments from the area of Beceite, in the Iberian Range and the Catalan Coastal Range (Catalànids), within the geological context of the Iberian Plate. The Atlas Geoquímico de España (IGME, 2012) was used as the main geochemical data bank in order to carry out a statistical analysis study.

Keywords: *bayesian inference, iberian range, lithium, soil.*

I. Introduction

Lithium is a relatively rare element. Although lithium is found in many rocks and some brines, it is always present in very low concentrations. Lithium presents many important and meaningful uses and applications and the most important is electrical and electronics standing out lithium ion batteries, but not only this application, it is used as a lubricator for greases, in metallurgy, organic and polymer chemistry or medicine. The average amount in the earth's upper crust was estimated by Vine (1980) to be about 20 ppm, although others have reported values as low as 7 ppm (igneous rocks 6 ppm, sedimentary rocks 11.5 ppm) and as high as 60 ppm, this element ranking 27th on the list of abundance (Garrett et al., 2004).

Because of its small ionic radius, lithium (+1, 0.68 Å) tends to replace magnesium (+2, 0.66 Å) or iron (+2, 0.74 Å) in silicates rather than the much larger alkali ions, sodium (+1, 0.97 Å), potassium (+1, 1.33 Å), rubidium (+1, 1.47 Å) or caesium (+1, 1.67 Å), to which it is chemically more closely related (Ju-chin-Chen, 1998). Three main types of lithium ore can be differentiated: brines and related evaporites, pegmatites and sedimentary rocks. Brines containing lithium make up 66% of the world's lithium resources, pegmatites make up 26%, and sedimentary rocks account for the remaining 8% (Grubber et al., 2011). The deposits formed are related to lithium's greater capability for solubility than most other cations. Accordingly, it can sometimes be found concentrated in flowing and cooling magma and/or its associated aqueous solutions, besides also being present in evaporating brines. Lithium is associated with sedimentary rocks, including clay and lacustrine evaporites. In clay deposits, lithium is a constituent of clay minerals, such as smectite. A remarkable smectite in this environment would be Hectorite [$(Mg,Li)3Si_4O10(OH)_2$], which is rich in magnesium and lithium.

From the geochemical point of view, one of the main goals of this paper is to study the relationships of lithium with several other elements in an attempt to identify and establish trends in geochemical behaviour, within the geological context of a linking area under the influence of two tectonic domains, namely: the East Iberian range and the Catalan Coastal range. This area was remarkably tectonized by the Alpine orogeny in the Late Mesozoic and Cenozoic. Going a step further, another important aim is the study of the distribution of lithium, in relation with the potential atomic exchanges that occur in several crystalline structures – especially with the existence of Fe, Mg, K and Na. Furthermore,

Author α ω: *Department of Agricultural and Environment Sciences. Juame I University, Castellón, Spain.*
Author σ: *Mathematics Department, Statistics Area, of University Jaume I of Castellón, Spain. e-mail: juan@mat.uji.es*
Author ρ: *Instituto Investigación en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México.*

this study is a part of a more ambitious project that embraces a larger area.

In this paper is to present the prediction and the basic features of the Integrated Nested Laplace Approximation (INLA; Rue et al., 2009) approach as applied to geochemical spatial data. In previous studies of geochemical elements, it has been in Geostatistical Analysis context (Jordan et al., 2004; Juan et al., 2011), but now the INLA approach can be employed to improve the results. The INLA is designed for latent Gaussian models ranging from linear mixed to spatial and spatio-temporal models. For this reason, the INLA can be successfully used in a great variety of applications.

The paper is structured as follows. First, in Section 2 we review the main characteristics of the geochemical analysis of the data. In Section 3 the sampling data are presented, and this is followed by Section 4, where we detail the analysis of the data. We provide an overview of the theory underlying the INLA and we review the SPDE approach to dealing with spatial data in Section 5. The practical application to area-level data is presented in Section 6. After that, in Section 7, issues are discussed and some conclusions are set out.

II. Geological Background

In order to understand better the geochemical content and its arrangement throughout the range is revised in a short way the geological context. In this way, the study area lies within the geological context of the Iberian Plate, close to the Iberian Range and the Catalan Coastal Range (Catalànids), and forms a single tectonic unit (Vera, 2004). The east-central margin of the Iberian Peninsula a relief which often exceeds 1000 m. in height as the result of two alpine structures that were partially destroyed there (Fig. 1).

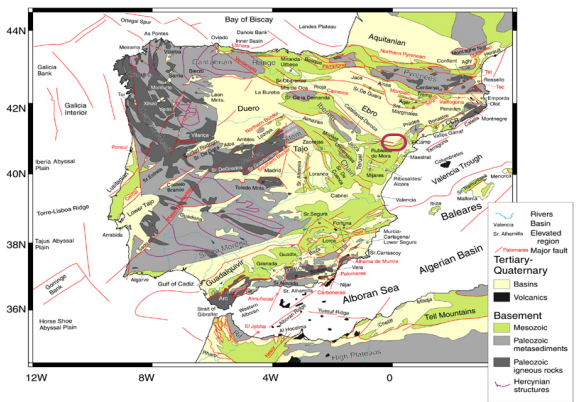

Figure 1: Geological map of the Iberian Peninsula and the western Mediterranean basins, major faults and regions (Andeweg, 2002).

The Iberian Range is over 400 km. long and extends from the Iberian Meseta to the Mediterranean Sea with a maximum width of 200 km. The Catalan Coastal Range constitutes a mountain barrier 200 km. in length and 30-40 km. wide that runs parallel to the coast and separates the Ebro basin from the Mediterranean Sea.

Both structures, the Iberian Range and the Catalan Coastal Range, converge at their eastern and southern ends to form the Maestrat, where the area known in Spanish as “zona de enlace” is located (Fig. 2).

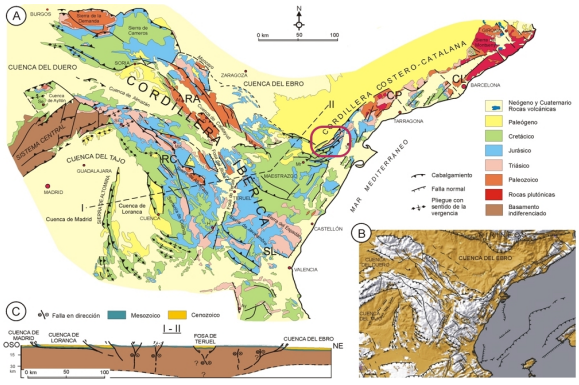

Figure 2: Geological map of the Iberian Range and Catalan Coastal Range (Vera, 2004).

It is made up of a fold-and-thrust belt running NW-SE (Iberian Range) and NE-SW (Catalànids) formed as a result of the Alpine reversion of the Mesozoic Iberian Basin. Its structure includes the fractured Hercynian basement, which has Alpine structures superimposed upon it that later underwent Neogene distension, especially in the south-east of the Iberian Range.

Most of the structural features of the Iberian Range are the result of the interaction of two factors: first, the reactivation of old compression fractures in the late Hercynian, NW-SE and NE-SW, and, secondly, the relationship between the cover and the underlying basement (Guimerà and Alvaro, 1990).

The tectonic style develops three structural levels characterized by the involvement of stratigraphic succession and deformation behaviours. The lower level consists of the Hercynian basement and overlying Permian and Triassic, while the upper level includes the Mesozoic (Jurassic and Cretaceous) and Cenozoic cover (bed) (pre- and syntectonic PaleogeneThe two are separated by a detachment level located in marls materials and evaporatic materials from the middle and upper Triassic, and hence they develop more folds and thrusts, and their main characteristic is their plastic behaviour and incompetence.

Between the Oligocene and lower Miocene a new extensional process was produced that is framed within the context of the Western European rift system. The distension mainly affects the eastern margin of the Iberian Plate and causes a complex system of horsts and grabens, parallel to the Mediterranean coast, related to the formation of the Valencia trench (Fontboté et al., 1990; Rock and Guimerà, 1992). These sedimentary basins are limited by longitudinal faults

running NE-SW and NW-SE corresponding largely to the basement fractures reactivated during the compression of the Mesozoic and the extension of the lower Tertiary (Palaeogene.

III. Sampling Information

The soil samples belong to the Geochemical Atlas of Spain, produced by the IGME in 2012, which studies the geochemical components of soils and other superficial materials in the Spanish territory. A multi-element geochemical characterization of soils and superficial materials was performed. These geo-referenced data were introduced into a high-resolution geo-database.

a) Sampling methodology

- Composite stream sediments and residual soils at two depths (0-25 cm and 25-50 cm). A total of 14,000 sampling points (sediments + 2 soil depths). A total of 36,400 samples.
- Flooding plain sediments at two depths (top and bottom) in drainage basins with an area of 3000 to 6000 km2. 332 sampling points are foreseen (664 samples).
- Supplementary data gathered at each hydrographic network and soil sampling point: pH and TOC.
- Three sampling densities were foreseen in the project, in accordance with the same number of great areas with higher or lower geological complexity, and industrial and demographic pressure (1 sample point/10 km2, 1 sample point/20 10 km2 and 1 sample point/100 km2). The spatial location of the sampling points is shown in Figure 3.

b) Sample Preparation and multi-element chemical analysis

- Sieving of sediments (150 microns), flood sediments (63 microns) and soils (2 mm).
- Total or almost total multi-element chemical analyses (63 elements) and with partial extraction (aqua regia) of samples by INAA, ICPAES and ICPMS. TOC determination.
- Analysis of organic pollutants (dioxins, pesticides, etc.) in agricultural zones, industrialized zones or areas with a high demographic density (1200 samples).

c) Quality control

- Quality control and sampling error determination (2 x 3.7% of field duplicates)
- Analytical quality control and analytical error determination (3.7% of duplicate samples)

This geochemical infrastructure complies with some basic requirements, derived from the recommend-dations of IGCP project 259 (Darnley et al., 2005) (IUGS-UNESCO). According to these requirements, sampling must satisfy the following requirements:

1. It should be based on samples of different media, which include soils (two depths), stream sediments, and floodplain sediments (2 depths). These later, in alluvial systems of a certain magnitude, will make it possible to determine the original uncontaminated values, prior to the industrialization period, and the variations that have taken place over time by comparison with present and future values.
2. It must be multi-element, including a wide range of chemical elements (63 in this case), comprising all those that are potentially dangerous for the ecosystem and those of any interest to geology and mineral resource research. They will be analysed by reliable techniques and with very low detection limits. Determination of the total or almost total contents and of the partial contents that will allow the easiest extractable or bio-available part to be estimated. This latter point is of particular interest for the environmental analysis and would allow an assessment of pollution processes to be carried out.
3. It should be available in digital form, to facilitate the integration of the geochemical information with other datasets (geology, metallogeny, land uses, animal or human morbidity and mortality).
4. It must be standardized across national boundaries, as many of the problems and concerns that arise regarding the surface make it necessary to go beyond the frontiers and perform analyses on a more global scale. The follow-up of the IGCP project 259 recommendations and the protocols derived from the FOREGS project entitled "Global Geochemical Baselines for Europe" will act as a guarantee for compliance with this requirement.

IV. Data Setting

Figure 3 shows the location of the sampling points. The study region covers an area of 20 x 30 km2 and is located in the Iberian Plate, close to the Iberian Range and the Catalan Coastal Range, the east-central margin of the Iberian Peninsula in Spain.

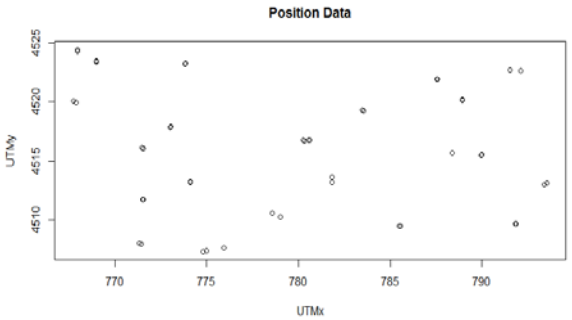

Figure 3: Beceite region with data points.

The data are distributed over a regular polygonal region, and the elements considered in our analysis are lithium, Al, K, Na, Ca, Mg, Fe, S, Mn, Ba,

Co, Ti, Zr, Nb, Ta, Ni, Cu, Pb, Zn, Sn, Rb, Cd, and Sb. Our goal is to analyse the concentration of lithium and its relationship with the concentration of the other elements considered in the study, considering those elements as covariates. The main objective is to assess how the covariates are related to the concentration of lithium in the study area and to asses which of such elements show a statistically significant relatonship with lithium.

Figure 4 shows the histogram of lithium concentrations as well as a map showing the sampling locations and the observed lithium concentrations at each site. The data show a non-symmetric distribution with a shape that not resembles the normal distribution. It shows a Gamma distribution.

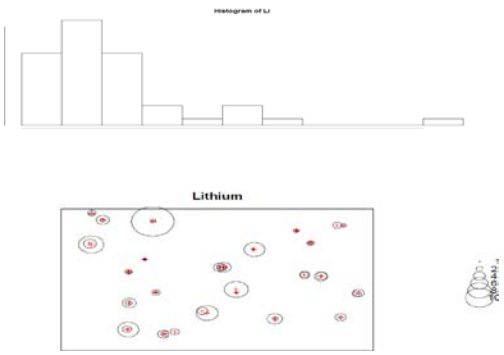

Figure 4: Top: Histogram of frequency of lithium; Bottom: Distribution of lithium in the region.

V. Methodology Statistical Inference

The basic idea behind the Bayesian approach is that in any statistical model of a random phenomenom, only one source of uncertainty exists, which can be described by suitable probability distributions of model parameters (Box and Tiao, 1973). Thus, there is no fundamental distinction between observable data or unobservable parameters, which are also considered random-vector quantity θ. .

The uncertainty about the value of the parameter vector before observing any data is described by a prior distribution $\pi(\theta)$ The inferential process combines the prior and the likelihood of observed data (x) given the model, to derive the posterior distribution $\pi(\theta|\ x)$, which is typically, but not necessarily, the objective of the inference (Bernardo and Smith, 2000; Lindley, 2006).

There are several advantages to the Bayesian approach: for instance the specification of prior distributions allows the formal inclusion of information that can be obtained through previous studies or from expert opinion; the (later) probability that a parameter does/does not exceed a certain threshold is easily obtained from the posterior distribution, thus providing a quantity that is more intuitive and interpretable than a frequentist p-value. In addition, within the Bayesian approach, it is easy to specify a hierarchical structure on the data and/or parameters, which presents the added benefit of making predictions for new observations and missing-data imputation relatively straightforward.

a) *Integrated Nested Laplace Approximation (INLA)*

The Integrated Nested Laplace Approximation (INLA; Rue et al., 2009) approach has recently been developed as a computationally efficient alternative to more commonly known methods such as Markov Chain Monte Carlo (Gelfand and Smith, 1984, Wilks, 1996). INLA is designed for latent Gaussian models, a very wide and flexible class of models ranging from generalized linear mixed models (Breslow and Clayton, 1993) to spatial and spatio-temporal models. For this reason, INLA can be successfully used in a great variety of applications (e.g. Li et al., 2012; Riebler et al., 2012; Ruíz-Cardenas et al., 2012), thanks to the availability of an R package named R-INLA (Martino and Rue, 2010). Furthermore, INLA can be combined with the Stochastic Partial Differential Equation (SPDE) approach proposed by Lindgren et al. (2011) in order to implement spatial and spatio-temporal models for point-reference data.

b) *The model*

Let Yj denote the observed data about lithium at specific points sj, j=1,…,24. Because Y>0 for all s, we may consider that Y has a Gamma distribution. The Gamma distribution for lithium was chosen because the shape for the histogram for lithium concentrations suggested this was a sensible model for such data.

Given the positiveness of the values for lithium concentration, we consider the model:

$$\begin{aligned} y \mid F_y, \beta, x_i, \theta &\sim Gamma(a_i, b_i) \\ \log(\mu_i) &= F_i^T \beta + x_i \\ x_i &\sim GF(0, \Sigma) \end{aligned} \quad (1)$$

where Fi is a vector of covariates made by the concentration of the other elements considered in this study, β a vector of coefficients and x a latent Gaussian random field , with a covariance matrix function with parameters ν, κ and σx2. With the Gamma likelihood, we consider that $E(y_i) = \mu_i = a_i / b_i$ and $V(y_i) = a_i / b_i^2 = \mu_i^2/\varphi$,, where ϕ is a precision parameter and we define a linear predictor as log(μi) (R-INLA, 2011).

In this case, for all the models, we obtained the parameters of the posterior distribution of the intercept (α), the dispersion parameter of the Gamma likelihood, the latent field parameters log(κ) and log(τ), the fixed coefficients (β) for the covariates (Al, K, Na, Ca, Mg, Fe, S, Mn, Ba, Co, Ti, Zr, Nb, Ta, Ni, Cu, Pb, Zn, Sn, Rb, Cd, Sb) and the values of (1/σe2). The nominal variance and the nominal range (distance with correlation equal to 0.1) are also presented.

For the spatial covariance structure we used the Matérn family of covariance functions and a nugget term, over j.

$$\mathrm{cov}(\varepsilon_j, \varepsilon_{j'}) = M(|j - j'|, r^2\sigma^2, \rho, \upsilon) + (1 - r^2)\sigma^2 I(j = j') \quad (2)$$

M was the Matérn function (R-INLA, 2011); σ2 denoted the sill (the total variance of the innovation process); $r^2\sigma^2$ was the variance of the spatially correlated portion of the process; $(1- r^2)\sigma^2$ corresponded to the nugget (the variability that is unique to a given station); ρ was the range of the process (the size of the region where the process was significantly correlated); and υ was the smoothness of the process (specifically, we tried $\upsilon=1$ for the model, R-INLA project). All the analyses were carried out with the freeware statistical package R (version 3.0.2) (Lindgren et al., 2011) and the R-INLA package.

c) Stochastic Partial Differential Equation Approximation

The SPDE approach allows a GF with the Matérn covariance function to be presented as a discretely indexed spatial random process (GMRF), which has significant computational advantages. Gaussian Fields are defined directly by their first- and second-order moments and their implementation is highly time-consuming and leads to the so-called "big n problem" (Lindgren et al., 2011). The idea is to construct a finite representation of a Matérn field by using a linear combination of basis functions defined in a triangulation of a given domain D. This representation leads to the SPDE approach given by (2), which is a link between the GF and the GMRF, and allows the spatio-temporal covariance function and the dense covariance matrix of a GF to be replaced with a neighbourhood structure and a sparse precision matrix, respectively, both of which are typical elements that define a GMRF. This, in turn, produces substantial computational advantages (Lindgren et al., 2011).

More particularly the SPDE approach consists in defining the continuously indexed Matérn GF X(s) as a discrete indexed GMRF by means of a basis function representation defined on a triangulation of the domain D:

$$X(s) = \sum_{l=1}^{n} \varphi_l(s)\omega_l \quad (3)$$

where n is the total number of vertices in the triangulation; {φl(s)} is the set of basis functions and {ωi} are the zero-mean Gaussian distributed weights. The basis functions are not random, but rather are chosen to be piecewise linear on each triangle:

$$\varphi_l(s) = \begin{cases} 1 & at\ vertix \\ 0 & elsewhere \end{cases} \quad (4)$$

The key is to calculate {ωi}, which reports the value of the spatial field at each vertex of the triangle. The values inside the triangle will be determined by linear interpolation (Simpson et al., 2011).

In our case, we obtain this representation of the mesh with the response variable, see Fig. 5.

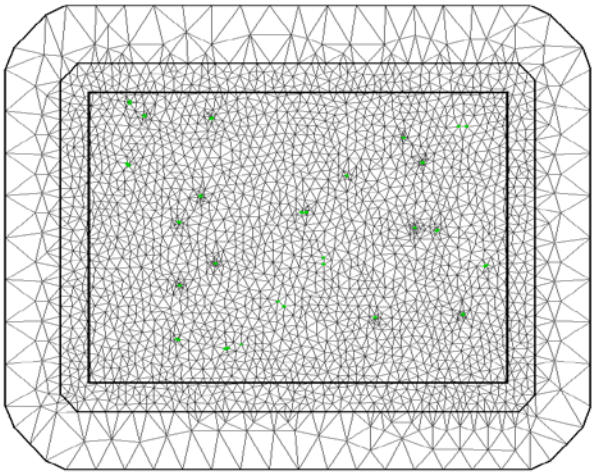

Figure 5: Triangulation mesh data for INLA.

We can see the relation between the data and the vertex for the implementation. This is the usual way of representing the regular region data when using INLA//to be able to use INLA.

VI. Results

At this point, we know the data and the statistical framework and we have presented the statistical process with the applied study. First of all we can observe the advantage of this process because, in general, the processing time is long. In our case, the time needed for pre-processing: 8.2529 seconds, for running each INLA model: 3.4478 seconds, and for post-processing: 0.4212 seconds. We have produced many previous models but the one presented here is the final version used, with this processing time.In this study, different models have been tested. The first one (M1) includes only the spatial effect, the second one (M2) includes the spatial effect and the covariates (chemical elements) and, finally, the third one (M3) has only the covariates without the spatial effect.The basic structure in INLA for each model is:

M1 =

inla(resp ~ 0 + m + f(i,model=spde5),family='gamma')

M2 =

inla(resp~0 + m + f(i,model=spde5) + Al + K + Na + Ca + Mg + Fe + S + Mn + Ba + Co + Ti + Zr + Nb + Ta + Ni + Cu + Pb+Zn+Sn+Rb+Cd+Sb, family='gamma')

M3=

inla(resp~0+m+Al+K+Na+Ca+Mg+Fe+S+Mn+Ba+Co+Ti+Zr+Nb+Ta+Ni+Cu+Pb+Zn+Sn+Rb+Cd+Sb,family='gamma')

Now, the results of the model, the fixed effects, the random effects and the comparison between the real and the predicted data are presented.

a) *Fixed effects*

In our models, the fixed effects are presented by the covariates.

Table I shows the mean, the standard deviation and the 2.5, 50 and 97.5 percentiles for the Intercept of the models (elements other than lithium).

Table 1: Intercept coefficient of the models.

	mean	sd	0.025quant	0.5quant	0.975quant
M1	3.105	0.165	2.774	3.104	3.449
M2	1.065	1.312	-1.590	1.084	3.606
M3	0.322	1.233	-2.089	0.314	2.772

For the covariates, the fixed effects are presented in models M2 and M3 (in bold, the significative covariates).

Table 2 a: Fixed effects of the model M2. Mean, sd and 2.5, 50 and 97.5 percentil.

	mean	sd	0.025quan	0.5quant	0.975quant
Al	- 0.041	0.221	-0.476	-0.04	0.39
K	1.011	0.526	-0.022	1.008	2.06
Na	- 0.720	0.914	-2.574	-0.707	1.05
Ca	0.028	0.034	-0.038	0.027	0.09
Mg	-0.021	0.060	-0.138	-0.022	0.10
Fe	0.498	0.320	-0.135	0.498	1.13
S	3.894	4.014	-3.934	3.853	11.997
Mn	-0.000	0.001	-0.002	-0.0005	0.001
Ba	-0.0033	0.008	-0.019	-0.003	0.0129
Co	-0.064	0.053	-0.172	-0.063	0.040
Ti	3.500	3.670	-3.778	3.507	10.741
Zr	0.017	0.008	-0.0006	0.017	0.034
Nb	0.128	0.064	0.004	0.127	0.259
Ta	-1.957	0.905	-3.795	-1.942	-0.207
Ni	0.023	0.034	-0.043	0.023	0.091
Cu	0.005	0.004	-0.004	0.005	0.014
Pb	0.011	0.011	-0.010	0.011	0.033
Zn	0.007	0.008	-0.008	0.007	0.024
Sn	-0.126	0.146	-0.419	-0.126	0.161
Rb	-0.012	0.014	-0.041	0.012	0.016
Cd	-0.050	0.920	-1.883	-0.047	1.760
Sb	-0.708	0.568	-1.832	-0.709	0.415

Table 2 b: Fixed effects of the model M3. Mean, sd and 2.5, 50 and 97.5 percentil.

	mean	sd	0.025quan	0.5quant	0.975quant
Al	0.011	0.214	-0.409	0.0113	0.435
K	1.156	0.500	0.177	1.152	2.153
Na	-1.051	0.925	-2.858	-1.057	0.791
Ca	0.047	0.033	-0.018	0.047	0.113
Mg	0.021	0.054	-0.085	0.021	0.128
Fe	0.490	0.311	-0.121	0.489	1.105
S	5.401	3.924	-2.256	5.372	13.220957
Mn	-0.0001	0.001	-0.002	-0.0001	0.002
Ba	-0.002	0.008	-0.018	-0.002	0.013
Co	-0.103	0.047	-0.197	-0.103	-0.009
Ti	4.833	3.303	-1.702	4.839	11.323
Zr	0.013	0.008	-0.004	0.013	0.030
Nb	0.156	0.064	0.029	0.156	0.283
Ta	-2.387	0.905	-4.163	-2.390	-0.594
Ni	0.043	0.029	-0.015	0.0438	0.102

Cu	0.003	0.004	-0.005	0.003	0.012
Pb	0.012	0.010	-0.008	0.0121	0.033
Zn	0.004	0.008	-0.011	0.0046	0.020
Sn	-0.184	0.140	-0.460	- 0.185	0.094
Rb	-0.015	0.015	-0.045	- 0.015	0.013
Cd	-0.393	0.842	-2.050	- 0.395	1.270
Sb	-0.754	0.528	-1.798	- 0.754	0.285

In Tables IIa and IIb, we can see positive values for the elements that increase the value of lithium and the negative values for the elements that decrease the amount of lithium. But the most important thing is to check which one has a significant statistical effect.

For Model M2, we can see that only Nb offers a significant result and Zr, Ta, Cu, Pb, Zn and Rb are elements close to being representative in the study of lithium.

For Model M3, we can see that K and Nb are positively significant in the study of lithium and Ta is significant in the opposite way.

b) Random effects

In this case, the hyperparameters correspond to the Gamma distribution and, since we do not know the values (Bayesian inference), they could be compared with the posterior. We have only the structure. In this case, the parameters are 1/σe2 or marginal variance nominal, marginal kappa and marginal range nominal.

MODEL 1
1/σe2: 2.229
Marginal kappa: 2.259296
Marginal variance nominal: 0.09013456
Marginal range nominal: 5.123559

MODEL 2
1/σe2: 27.419
Marginal kappa: 3.54715
Marginal variance nominal: 0.08050951
Marginal range nominal: 2.273834

If we want to know the form of the structure of the parameters, they can be seen in Fig. 6.

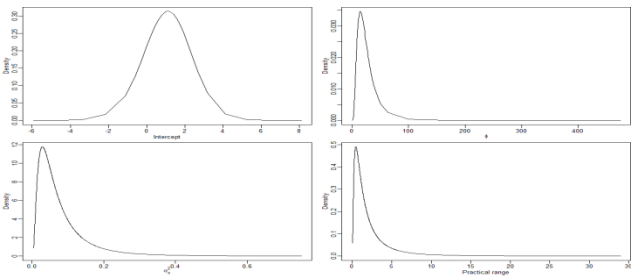

Figure 6: Marginal distribution of the parameters of M2.

For M3, the parameter associated to the Gamma likelihood has a value of 10.7411 with a standard deviation of 2.8977.

We can see the structure of this parameter in Fig. 7 below.

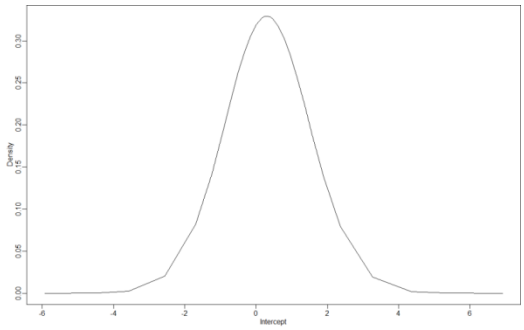

Figure 7: Marginal distribution of the parameters of M3.

Finally, in this section we can compare the Deviance Information Criterion (DIC) of the three models, where the best one is clearly the complete one, i.e. M2 (the lowest DIC value).

M1: 379.7338; M2: 286.9914; M3: 317.1496

Comparison of observed vs. predicted values and computed residuals in the same positions.

Fig. 8 shows the relation between the observed and the predicted values for the three models. The results are clearly optimum in the second one, M2. In Table III we can also see the results for the correlation coefficients and for the root-mean-square error (RMSE). In the same way, the results are satisfactory and the best model is the second one (M2).

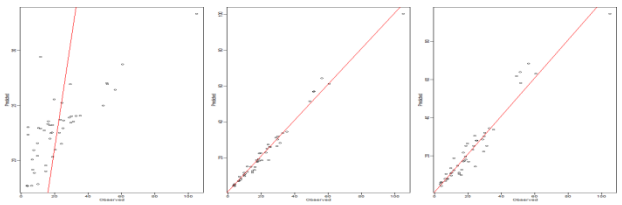

Figure 8: Correlation between Observed and Predicted values for 3 models, M1, M2 and M3.

Table 3: Correlation and RMSE for each model analyzed.

Model	Correlation coefficient	RMSE
M1	0.7833763	16.17654
M2	0.9909199	2.59907
M3	0.9667658	5.061028

c) Prediction Map

Finally, we can see the prediction maps of the latent field and the response variable for M1, with the corresponding standard deviation and the response variable for M2 (Fig. 9-11). As an example of the application of this methodology the reader is referred to Jordan et al. (2004).

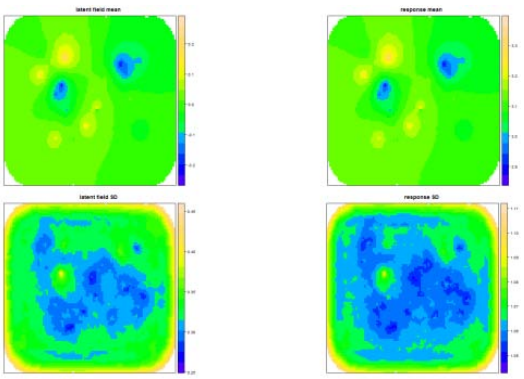

Figure 9: Prediction maps of M1. Top-left the latent field mean, the top-right the response variable mean, the down-left the latent field sd and down-right the response variable sd.

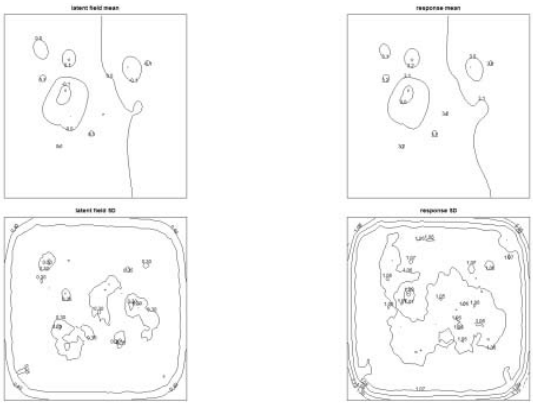

Figure 10: Contour prediction maps of M1. Top-left the latent field mean, the top-right the response variable mean, the down-left the latent field sd and down-right the response variable sd.

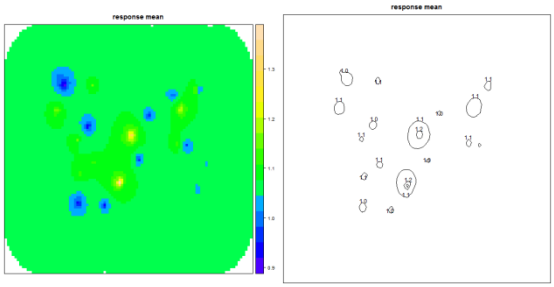

Figure 11: M2 in both cases. Left: Response variable mean. Right: Contour of response variable mean.

In all cases, the highest predicted values of lithium under these conditions are found for the North of the Beceite area.

VII. Discussion and Conclusion

This work presents the predicted values of lithium in the study region. They were obtained with a new methodology in Bayesian probabilities called INLA. The results showed that the lithium values could be studied by knowing the different covariates.

Henceforth, from the results of the model with the fixed effects, the random effects, and the comparison between real and predicted data it is possible to appreciate that lithium does not increase or decrease its concentration noticeably with any other element. On the one hand, if model 2 (M2), which fits the data better, is taken as a reference, at most it is possible to appreciate a slight increase in niobium and a minor decrease in thallium. On the other hand, model 3 (M3) showed an increase in potassium and niobium, and lithium decreased in the presence of thallium. Lithium might have been expected to connect with some specific geological materials of the area, such as clay minerals, with an important Al and K content and to a lesser extent Fe. It formed part of clay minerals, frequently appearing mixed with sandstones. In relation with the age of the materials, lithium was related mostly to Lower Cretaceous and Tertiary deposits, which cover an important surface area of the region, second to colluvial and alluvial Quaternary materials, and finally, to a much lesser extent, to Triassic formations.

In this study, we have focused in the content of Lithium at soils and sediments in relation to several elements, especially heavy metals, in order to determine its geogenic presence through the INLA modelization, taking into account that Lithium shows different geochemistry in different origin deposits, especially marine, as the main studied area.

Finally, regardless the results, the use of the model could provide new insights in order to face other geochemist studies in soils and sediments for environment, agrarian, or mining prospections.

As we known, Lithium presents many important uses and applications such as: ceramics and glass, electrical and electronics standing out lithium ion batteries, as well as a lubricator for greases, in metallurgy, pyrotechnics, air purification, optics, organic and polymer chemistry, and medicine. For all of these applications, the results we got for predicting the quantity and the location of Lithium it should be taken into account.

On the other hand, the phyllosilicates would contain potassium and would not appear with sodium clays. Furthermore, lithium would not be related with brines, for instance the marine Keuper deposits represented in the area, which contain sulphate and chlorine minerals (Fig. 12).

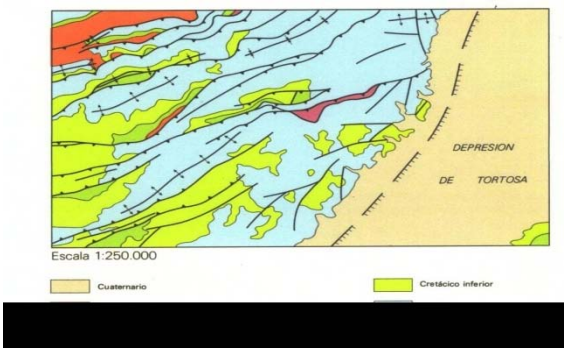

Figure 12: Tectonic map of the Beceite area in the Iberian Range and Catalan Coastal Range (IGME, 1985).

All in all, the geological context was defined during the Late Jurassic to Early Cretaceous rift stage, where several sub-basins were developed (in the Spanish provinces of Castelló, Teruel and Tarragona), which were filled by fluvial to transitional sandstones and lacustrine carbonates (Weald facies), later evolving to marine limestones (Urgon facies). These sub-basins trapped clastic sediments from the Tertiary, where potential source areas are metamorphic and pre-rift Triassic and Jurassic sedimentary cover rocks.

In summary, the most important idea is that the presence of a geological element in a region could be predicted when we know the others using the INLA procedure for the statistical inference, which is faster than the use of MCMC techniques.

References Références Referencias

1. Andeweg, B., 2002. Cenozoic Tectonic Evolution of the Iberian Peninsula: causes and effects of changing stress fields. PhD research Faculty of Earth and Life Sciences of the Vrije Universiteit in Amsterdam. Netherlands Research School of Sedimentary Geology (NSG) publication no. 20020101.
2. Bernardo, J. and Smith, A. (2000) Bayesian theory. Wiley-Blackwell.
3. Chen J., 1998. Lithium: Element and geochemistry. Springer: Encyclopedia of Earth Science, pp 369-370.
4. Cressie, N. (1993). Statistics for spatial data. Wiley.
5. Darnley, A.G., Bjorklund, A., Bolviken, B., Gustavsson, N., Koval P.V., Plant J.A., Steenfelt A., Tauchid M., Xuejing, X., Garrett, R.G., Hall, G.E.M., 2005. A Global Geochemical database for environmental and resource management: Recommendations for International Geochemical Mapping Final Report of IGCP Project 259. UNESCO Publishing.
6. Fontboté, J. M., Guimerá J., Roca, E., Sábat, F., Santanach, P., Fernández Ortigosa, F., 1990. The cenozoic geodynamic evolution of the Valencia trough (Western Mediterranean), Revista de la Sociedad Geológica de España, 3 (3 y 4): 249-259.
7. Garrett, D. E., 2004. Handbook of Lithium and Natural Calcium Chloride. Elsevier.
8. Gruber, P. W., Medina, P. A., Keoleian, G., Kesler, S. E., Everson, M.P., Wallington, T.J., 2011. Global Lithium Availability. A Constraint for Electric Vehicles? Journal of Industrial Ecology, Volume 15, Issue 5, pp 760–775, October 2011 by Yale University.
9. Guimera, J., Alvaro, M., 1990. Structure et évolution de la compression alpine dans la Chaíne ibérique et la Chaine cotiere catalane (Espagne), Bulletin de la Societé Géologique de France, VI (2): 339-348.
10. IGME, 2012. Atlas Geoquímico de España. ISBN: 978-84-7840-875-7
11. IGME..Mapa Geológico de España: Beceite. 2ª Serie MAGNA50, hoja 521.Instituto Geológico y Minero de España. I.S.S.N. 0373-2096.
12. Jordan, M.M., Navarro-Pedreño, J., García-Sánchez, E., Mateu, J. and Juan, P. (2004). Spatial dynamics of soil salinity under arid and semi-arid conditions: geological and environmental implications. Environmental Geology 45: 448–456.
13. Juan, P., Mateu, J., Jordan, M.M., Mataix-Solera, J., Meléndez-Pastor, I. and Navarro-Pedreño, J.. (2011) Geostatistical methods to identify and map spatial variations of soil salinity. Journal of Geochemical Exploration 108: 62–72.
14. Li Y, Brown P, Rue H, al Maini M, Fortin P. (2012). Spatial modelling of lupus incidence over 40 years with changes in census areas. Journal of Royal Statistic Society, serie C. 61(1):99–115.
15. Lindley, D. (2006). Understanding uncertainty. Wiley-Blackwell. Latent models. The R INLA project. http://www.r-inla.org/models/latentmodels
16. Lindgren, F., Rue, H., Lindstrom, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). Journal of the Royal Statistical Society, Series B, 73(4), pp.423-498.
17. Martino, S. and Rue, H. (2010). Implementing approximate bayesian inference using integrated nested laplace approximation: a manual for the inla program.
18. R Development Core Team. R, 2011. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/
19. R-INLA project, 2011. URL: http://www.r-inla.org/home
20. Riebler, A., Held, L. and Rue, H. (2012). Estimation and extrapolation of time trends in registry data – borrowing strength from related populations. Annals of Applied Statistics. 6(1): 304–333.

21. Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society, Series B, 71, pp.319- 392.
22. Ruiz-Cárdenas, R., Krainski, E. and Rue, H. (2012). Direct fitting of dynamic models using integrated nested laplace approximations INLA. Comput Stat Data Anal. 56(6):1808–28.
23. Simpson, D., Illian, J., Lindgren, F., Sørbye, S.H., Rue, H., 2011. Going off grid: Computationally efficient inference for log-Gaussian Cox processes, http://www.math.ntnu.no/~daniesi/S10-2011.pdf
24. USGS. Mineral commodity summaries, 2013.
25. http://minerals.usgs.gov/minerals/pubs/mcs/2013/mcs2013.pdf
26. Vera, J.A., 2004. Geología de España. Sociedad Geológica de España. Instituto Geológico y Minero de España. ISBN 84-7840-546-1.
27. Vine, J.D.,1980. Where on earth is all the Lithium? USGS.