

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

Pablo Juan¹

¹ University Jaume I of Castellón

Received: 15 December 2015 Accepted: 2 January 2016 Published: 15 January 2016

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.

Index terms— bayesian inference, iberian range, lithium, soil.

1 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 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, 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.

2 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). 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).

3 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 Paleogene). The two are separated by a detachment level located in marls materials and evaporitic 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; Fock and Guimerà, 1992).

4 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 multielement geochemical characterization of soils and superficial materials was performed. These georeferenced data were introduced into a high-resolution geo-database. 1. It should be based on samples of different media, which

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 :

where n is the total number of vertices in the triangulation; $\{\phi_i(s)\}$ is the set of basis functions and $\{\omega_i\}$ are the zero-mean Gaussian distributed weights. The basis functions are not random, but rather are chosen to be piecewise linear on each triangle: $\phi_i(s) = 1$ if s is a vertex of the triangle, and 0 elsewhere.

The key is to calculate $\{\omega_i\}$, 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). Year 2016 (G © 2016 Global Journals Inc. (US))

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

8 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. M was the Matérn function (R-INLA, 2011); τ^2 denoted the sill (the total variance of the innovation process); σ^2 was the variance of the spatially correlated portion of the process; $(1 - \rho^2)\tau^2$ 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 κ was the smoothness of the process (specifically, we tried $\kappa=1$ 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. Now, the results of the model, the fixed effects, the random effects and the comparison between the real and the predicted data are presented.

9 a) Fixed effects

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

Table ?? 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). In Tables ??Ia 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.

10 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 $1/\sigma^2$ or marginal variance nominal, marginal κ and marginal range nominal. ??ODEL If we want to know the form of the structure of the parameters, they can be seen in Fig. 6. 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. 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 ??II 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). In all cases, the highest predicted values of lithium under these conditions are found for the North of the Beceite area.

11 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). (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.

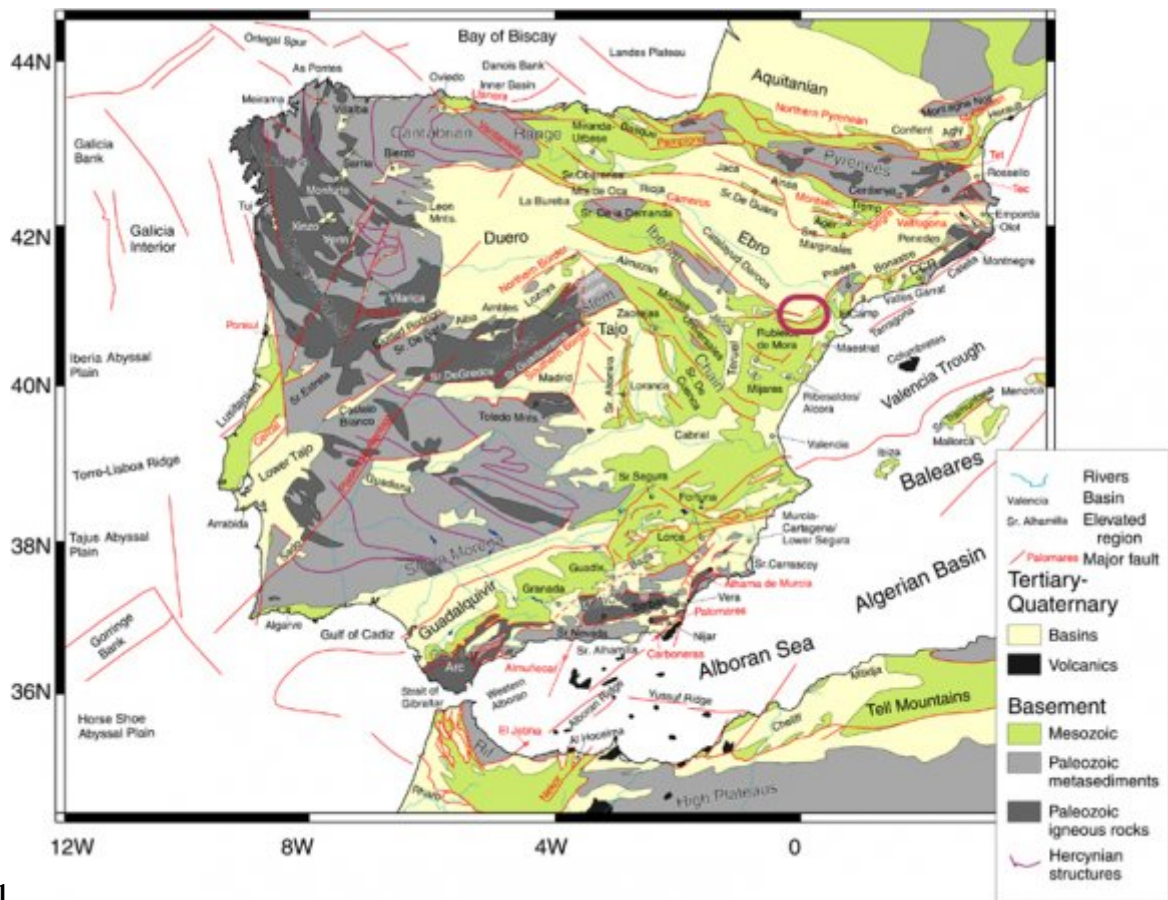


Figure 1: Figure 1

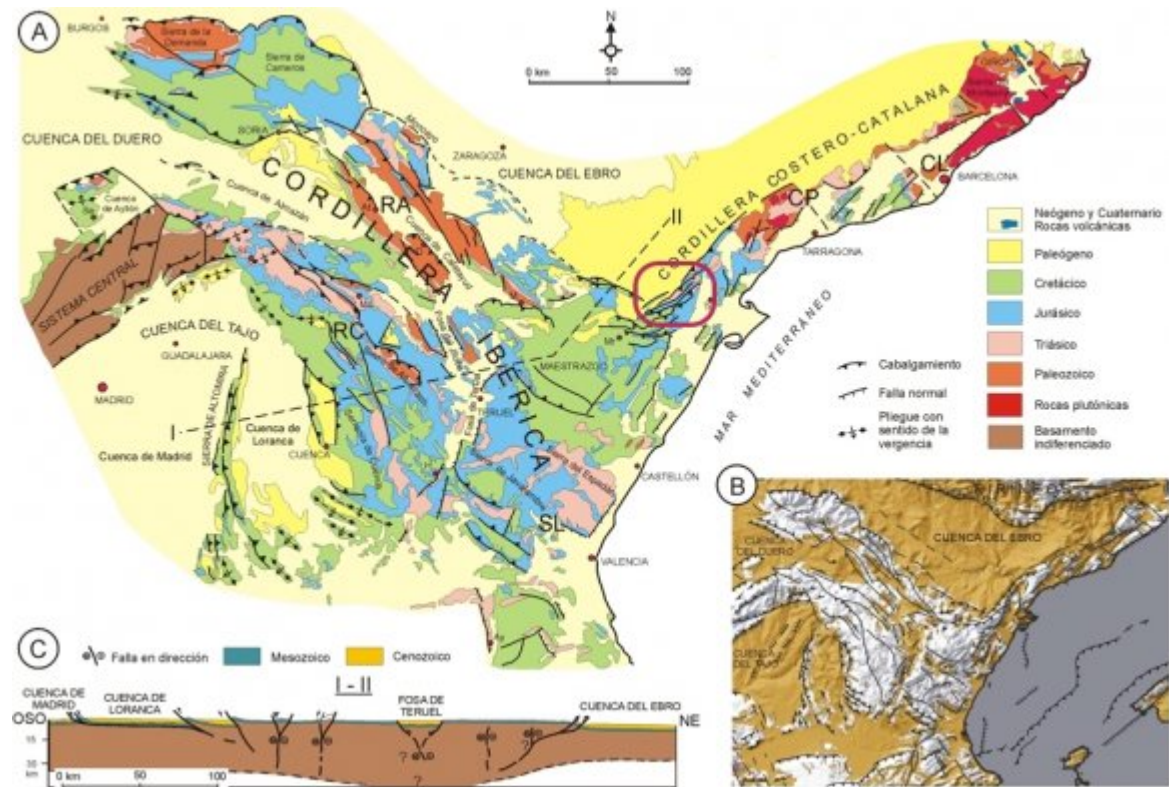


Figure 2:

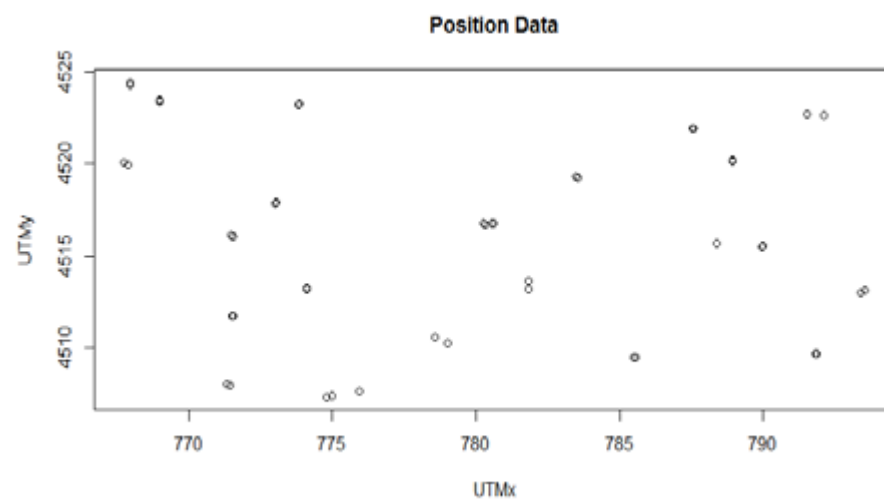


Figure 3:

3

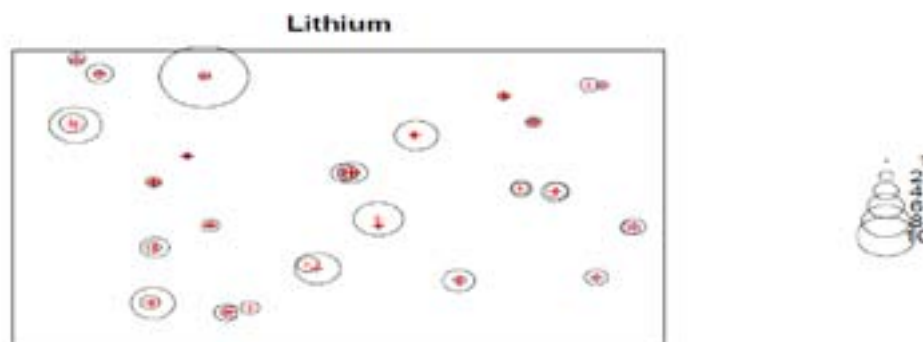


Figure 4: Figure 3

4



Figure 5: Figure 4

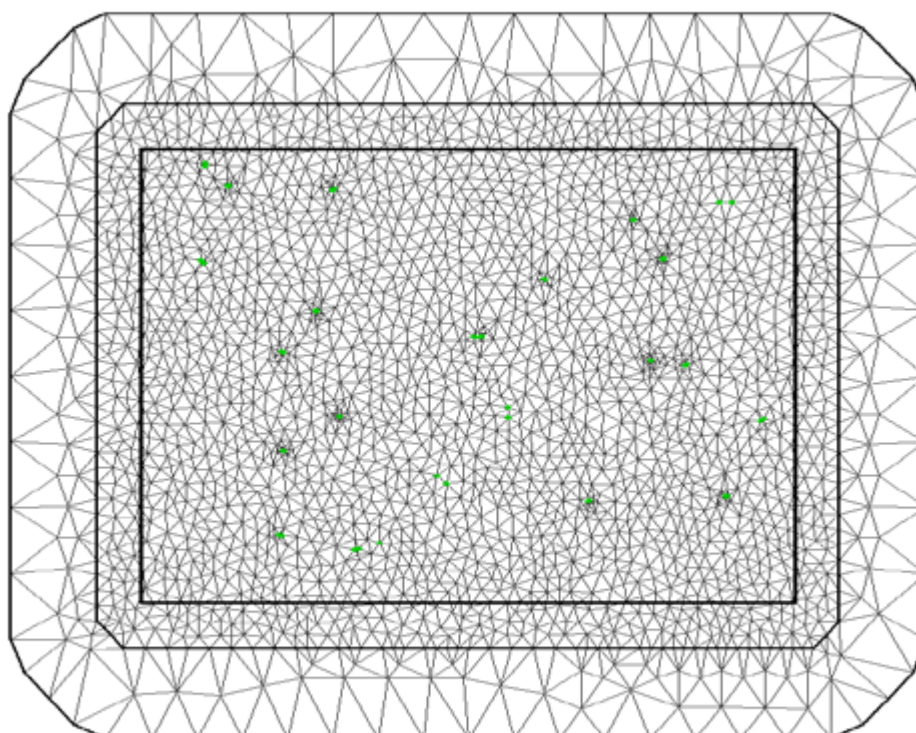


Figure 6:

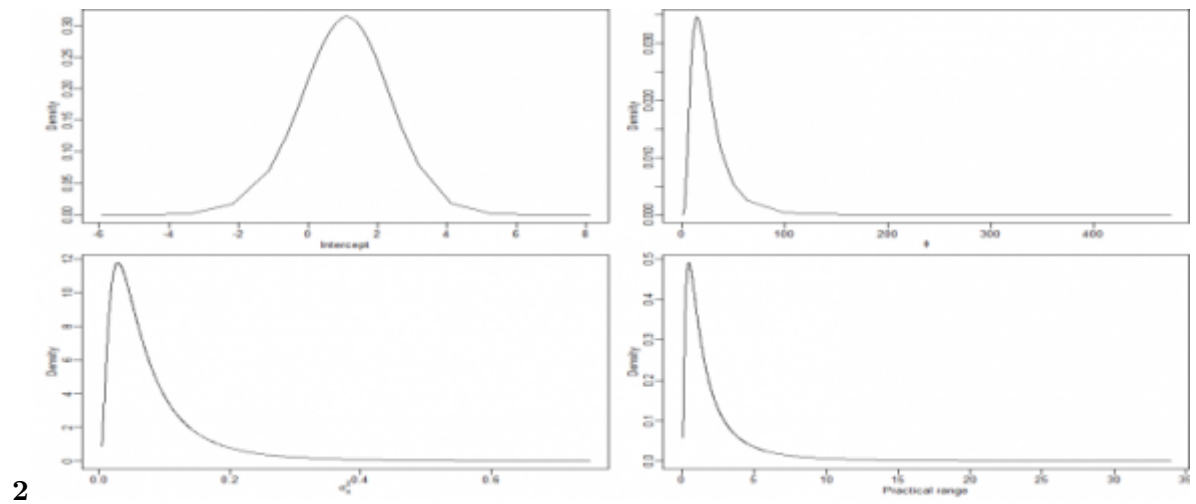


Figure 7: $A = \sigma^2 / s$

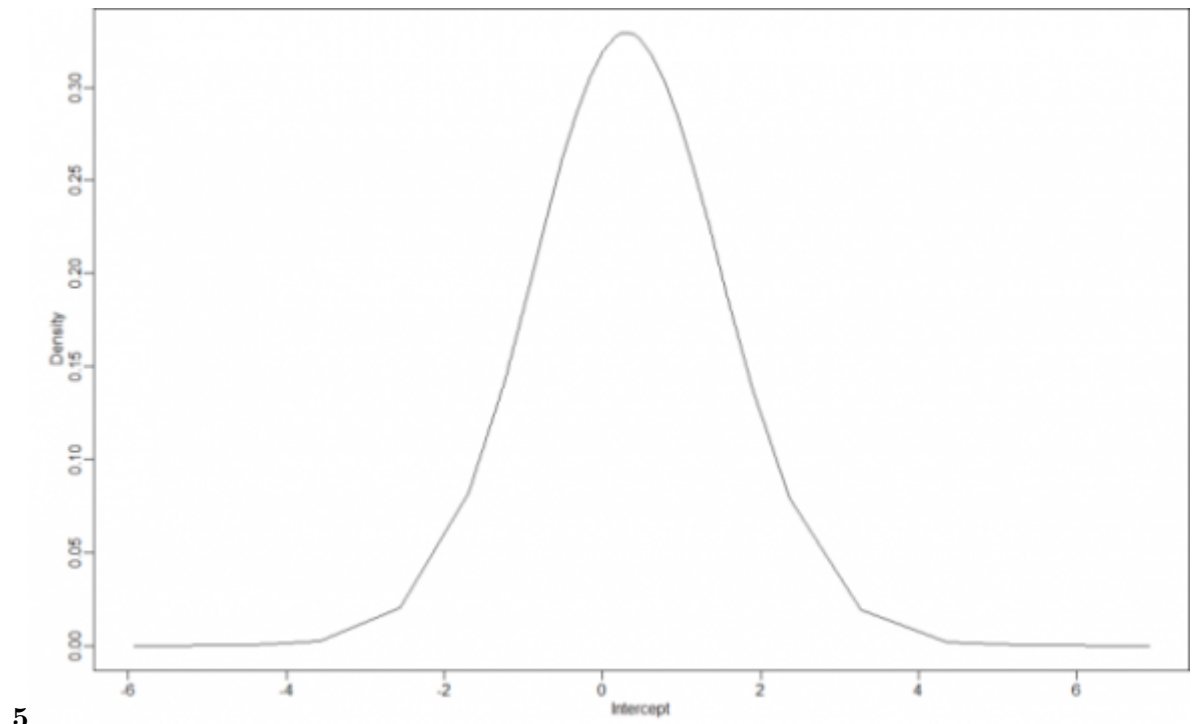


Figure 8: Figure 5

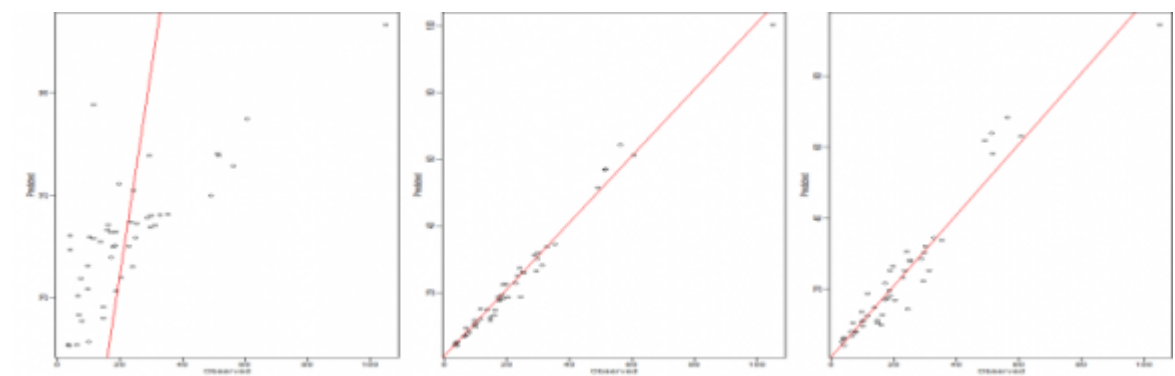
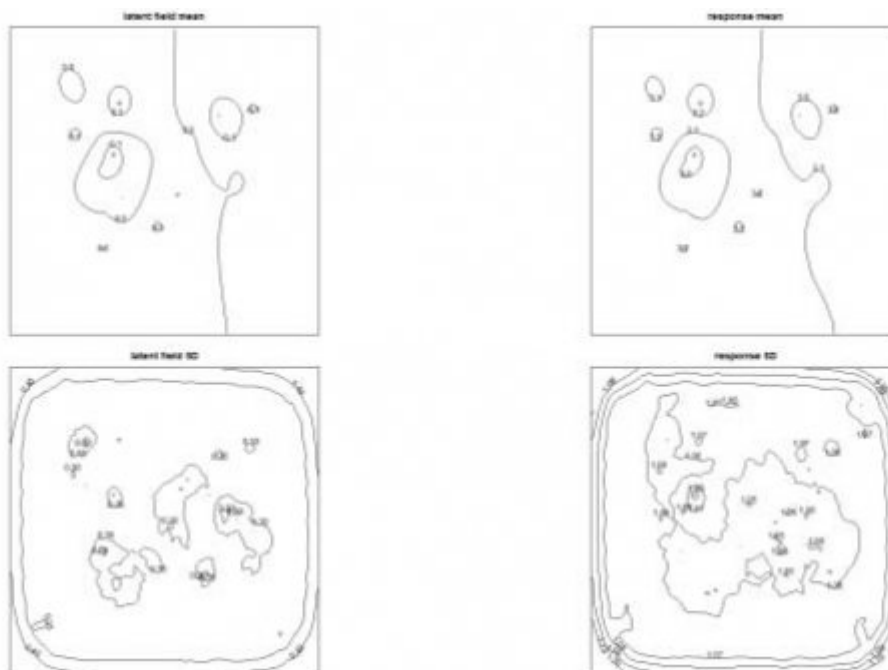
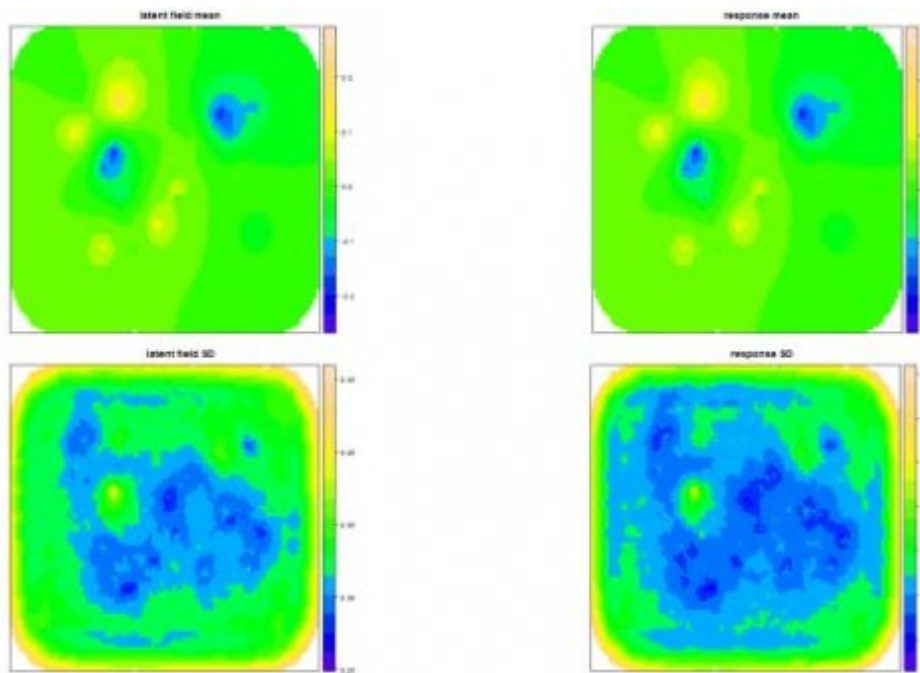
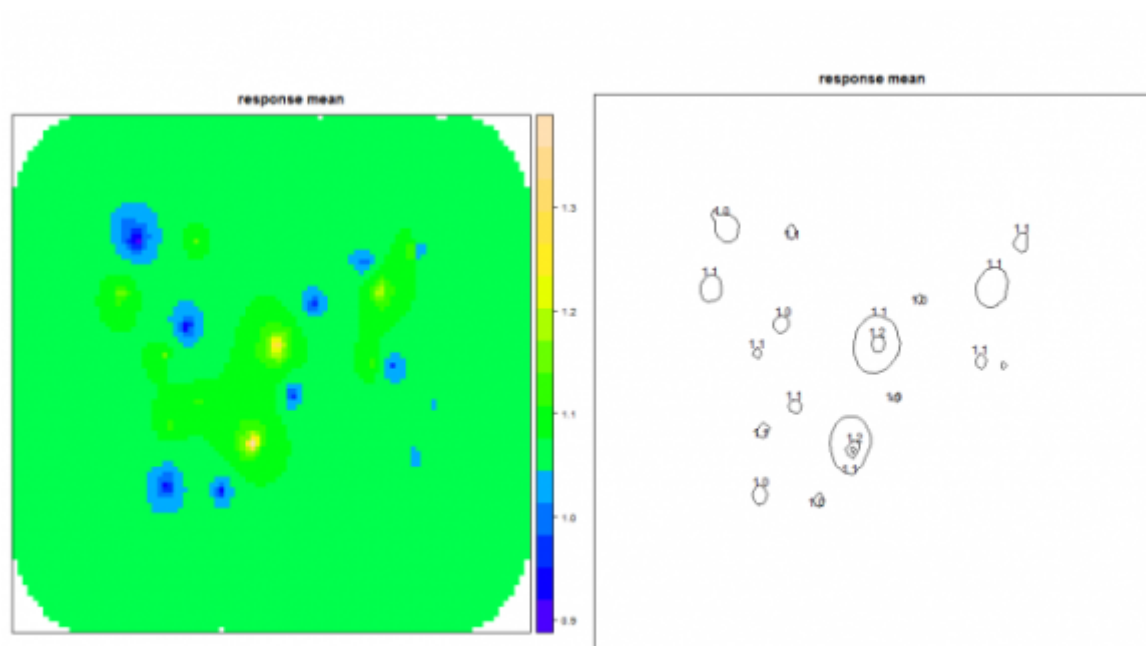


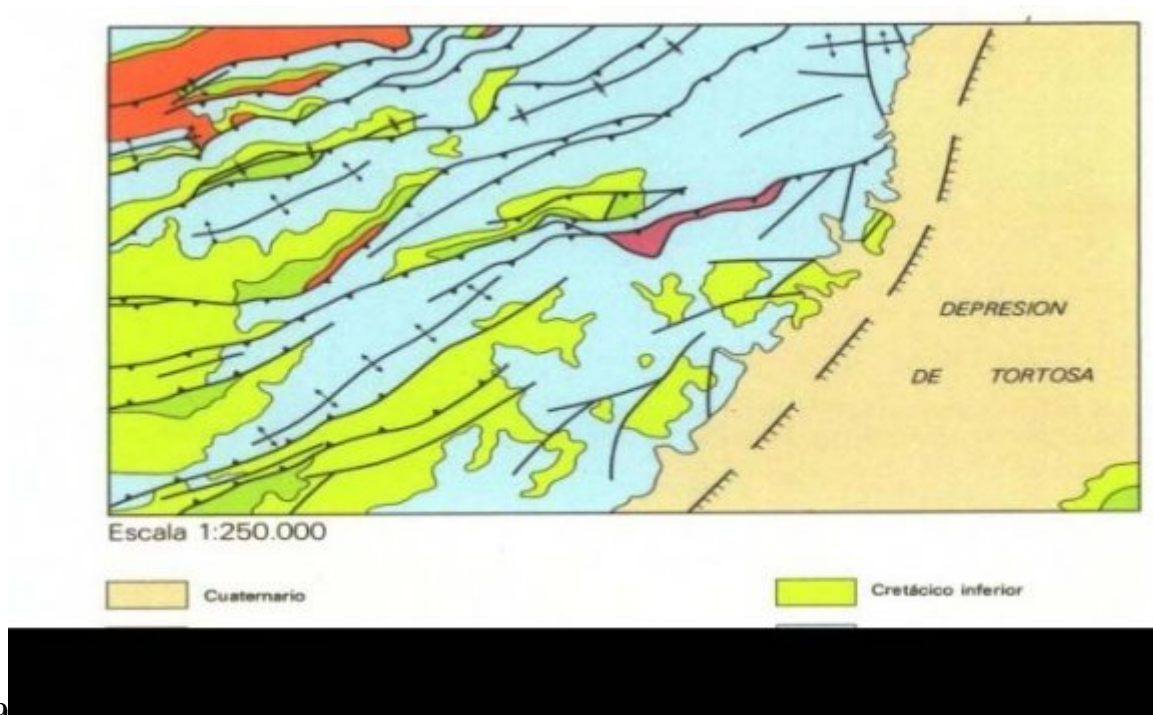
Figure 9:





8

Figure 12: Figure 8



9

Figure 13: Figure 9

1

	sd	0.025quant
M1	3.1050.165	2.774 3.1043.449
M2	1.0651.312	- 1.0843.606
		1.590
Year M3	For the covariates, the fixed effects are presented in models M2 and M3 (in bold, the significative C	
2016		
44	Fixed effects of the model M2. Mean, sd and 2.5, 50 a	

[Note: G b:A]

Figure 14: Table 1 :

2

Figure 15: Table 2

3

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

[Note: c) Prediction MapFinally, 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][10][11]. As an example of the application of this methodology the reader is referred to Jordan et al.(2004).A]

Figure 16: Table 3 :

[Igme] , Igme .

[Project ()] , R-Inla Project . <http://www.r-inla.org/home> 2011.

[Darnley et al. ()] *A Global Geochemical database for environmental and resource management: Recommendations for International Geochemical Mapping Final Report of IGCP Project 259*, A G Darnley , A Bjorklund , B Bolviken , N Gustavsson , P V Koval , J A Plant , A Steinfeld , M Tauchid , X Xuejing , R G Garrett , G E M Hall . 2005. UNESCO Publishing.

[R Development Core Team. R (ed.) ()] *A language and environment for statistical computing. R Foundation for Statistical Computing*, <http://www.R-project.org/> R Development Core Team. R (ed.) 2011. Vienna, Austria.

[Lindgren et al. ()] ‘An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion)’. F Lindgren , H Rue , J Lindstrom . *Journal of the Royal Statistical Society, Series B* 2011. 73 (4) p. .

[Rue et al. ()] ‘Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion)’. H Rue , S Martino , N Chopin . *Journal of the Royal Statistical Society, Series B* 2009. 71 p. .

[Atlas Geoquímico de España ()] *Atlas Geoquímico de España*, 2012. IGME

[Bernardo and Smith ()] *Bayesian theory*, J Bernardo , A Smith . 2000. Wiley-Blackwell.

[Andeweg ()] ‘Cenozoic Tectonic Evolution of the Iberian Peninsula: causes and effects of changing stress fields’. B Andeweg . *Netherlands Research School of Sedimentary Geology* 2002. 20020101. (PhD research Faculty of Earth and Life Sciences of the Vrije Universiteit in Amsterdam)

[Ruiz-Cárdenas et al. ()] ‘Direct fitting of dynamic models using integrated nested laplace approximations INLA’. R Ruiz-Cárdenas , E Krainski , H Rue . *Comput Stat Data Anal* 2012. 56 (6) p. .

[Riebler et al. ()] ‘Estimation and extrapolation of time trends in registry databorrowing strength from related populations’. A Riebler , L Held , H Rue . *Annals of Applied Statistics* 2012. 6 (1) p. .

[Vera ()] *Geología de España. Sociedad Geológica de España. Instituto Geológico y Minero de España*, J A Vera . 2004.

[Geológico De España] Mapa Geológico De España y Minero de España. I.S.S.N. . *Beceite. 2ª Serie MAGNA50, hoja 521. Instituto Geológico*, p. .

[Juan et al. ()] ‘Geostatistical methods to identify and map spatial variations of soil salinity’. P Juan , J Mateu , M M Jordan , J Mataix-Solera , I Meléndez-Pastor , J Navarro-Pedreño . *Journal of Geochemical Exploration* 2011. 108 p. .

[Gruber et al. (2011)] ‘Global Lithium Availability. A Constraint for Electric Vehicles?’. P W Gruber , P A Medina , G Keoleian , S E Kesler , M P Everson , T J Wallington . 10.1111/jiec.2011.15.issue-5/issuetoc. *Journal of Industrial Ecology* 2011. October 2011. 15 (5) p. . by Yale University

[Simpson et al. ()] *Going off grid: Computationally efficient inference for log-Gaussian Cox processes*, D Simpson , J Illian , F Lindgren , S H Sørbye , H Rue . <http://www.math.ntnu.no/~daniesi/S10-2011.pdf24>. **USGS.Mineralcommoditysummaries** 2011. 2013.

[Garrett ()] *Handbook of Lithium and Natural Calcium Chloride*, D E Garrett . 2004. Elsevier.

[Martino and Rue ()] *Implementing approximate bayesian inference using integrated nested laplace approximation: a manual for the inla program*, S Martino , H Rue . 2010.

[Lindley ()] *Latent models. The R INLA project*, D Lindley . <http://www.r-inla.org/models/latentmodels> 2006. Wiley-Blackwell. (Understanding uncertainty)

[Chen ()] ‘Lithium: Element and geochemistry’. J Chen . *Encyclopedia of Earth Science*, 1998. Springer. p. .

[Jordan et al. ()] ‘Spatial dynamics of soil salinity under arid and semi-arid conditions: geological and environmental implications’. M M Jordan , J Navarro-Pedreño , E García-Sánchez , J Mateu , P Juan . *Environmental Geology* 2004. 45 p. .

[Li et al. ()] ‘Spatial modelling of lupus incidence over 40 years with changes in census areas’. Y Li , P Brown , H Rue , M Maini , P Fortin . *Journal of Royal Statistic Society, serie C* 2012. 61 (1) p. .

[Cressie ()] *Statistics for spatial data*, N Cressie . 1993. Wiley.

[Guimera and Alvaro ()] *Structure et évolution de la compression alpine dans la Chaîne ibérique et la Chaîne cotière catalane (Espagne)*, J Guimera , M Alvaro . 1990. VI: Bulletin de la Société Géologique de France. p. .

[Fontboté et al. ()] ‘The cenozoic geodynamic evolution of the Valencia trough’. J M Fontboté , J Guimerá , E Roca , F Sábato , P Santanach , F Fernández Ortigosa . *Revista de la Sociedad Geológica de España* 1990. 3 p. .

[Vine ()] *Where on earth is all the Lithium? USGS*, J D Vine . 1980.