Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-208-AC2, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area" *by* Inês Gomes Marques et al.

Inês Gomes Marques et al.

icgmarques@fc.ul.pt

Received and published: 25 November 2018

Dear Referee2, We are very grateful for your rigorous assessment and the valuable comments and suggestions you provided to improve our manuscript. Please find enclosed the revised version of the manuscript "Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area". We believe that all your suggestions were carefully addressed. In the present letter, you will find our responses to each comments and change made. Particularly, we have corrected the methodology to calculate drainage density. We also clarified the error you detected regarding a mistake in the map of Figure 7 where for a low aridity index (AI) we predicted a high suitability value (Figure B1b), which you attributed to the negative sign assigned to AI used as weighting factor. We evaluated the impact of each predictor on the final model,

and discovered that soil type actually considerably worsened the performance of the GWR model. We therefore decided to remove it from the final model equation selected to build the suitability map. We also attempted to provide a better evaluation of the importance of each predictor in the final model, and improved the discussion section accordingly. A version of the new manuscript was uploaded in the journal platform. All the information included in this manuscript is completely original and has been approved by all authors. The authors declare no conflict of interest. This manuscript has not been published previously or concurrently submitted for publication elsewhere. Also we thank you for considering this revised manuscript for publication. Please do not hesitate to contact us for any further needed detail.

With our best regards, sincerely

##

Referee Comments 2

General comments

The current manuscript provides an interesting insight into the use of mapping and spatial regression to assess the occurrence of groundwater dependent vegetation (GDV). Such maps can subsequently be used to predict the effect of change in any of the explanatory variables, such as climate or groundwater depth, on the spatial distribution of GDV in an area. The paper is well written and structured, and is subdivided into two parts: the first on the building of regression models for predicting GDV occurrence based on actual data, and the second part where a parameter-based index is calculated to construct so-called "suitability maps" for GDV. While I find the first part strong and with high potential of publication by adding a scenario analysis, my main concern lies with the second part. In my opinion this part is less well developed and the interpretation of the results is largely straightforward, and to a certain degree incorrect. Interpretation is largely straightforward because of the large bias in weighting of the parameters, where the contribution of soil largely exceeds all other parameters,

thus making it essentially a soil map. Additionally, interpretation is to a certain degree incorrect, due to an apparent mistake in the generation of the suitability map, where (inadvertently) a negative weighting was assigned to the aridity index, resulting in the inverse impact of this parameter on the soil map. I would therefore recommend focusing and further elaborating on the regression modelling, as I will specify in my detailed comments below.

Specific comments

Abstract

Line 13-19: The first part of the abstract is more of an introduction. I suggest starting with what was actually done (line 20). Moreover, groundwater depletion will not occur merely as a result of climate change. Finally, as groundwater level seems to have such a low impact in the regression model, the question rises to what extent groundwater depletion will play a role in the spatial distribution of GDV.

Answer: The first paragraph of the abstract was deleted. The abstract was corrected according to new results.

Line 48: When referring to climate change impact studies on recharge in Mediterranean areas add the paper of doi.org/10.1007/s10113-012-0377-3, where such a study is being reported.

Answer: Done. The suggested reference has been added in the text lines 45-46 and reference list lines 1038-1041.

PART 1: REGRESSION MODELLING

Parameter selection for the regression model

Soil type: the authors only use the first layer of the soil. To understand the importance of capillary rise feeding into the root zone, the texture of deeper soils also needs to be considered. The latter could further affect the role of groundwater depth in the model,

as fine soils have a much higher capillary (and water-holding) capacity. In the model, soil type is subdivided into two sub-parameters (2 and 3, Equation 4 and Table 2), but this is not further explained. Evidently, this increases the weight of soil type in the regression model.

Answer: The classification of soils into 3 categories was explained in lines 175-178 of the M&M section and in Table A1 in appendix A. This predictor was removed from the model fitting after revision from the authors. It has not been possible to add the texture of deeper horizon into our study because such information was only available on inaccessible printed maps. Unfortunately, no such digital data were available when the manuscript was prepared or revised.

Groundwater depth: please comment on the reliability of the results in the empty areas (areas without wells or piezometers). Can wells and piezometers be used together, in other words, are all wells installed in unconfined aquifers?

Answer: Please notice the previous answer to the reviewer 1. The region under study is an area with a very low population density, which reflected in the lack of points for piezometric level measurement, mainly in unconfined aquifers ($\sim$96% of the total area). Once the correlation between the piezometric level and the topography was successfully tested it was possible to estimate the piezometric level by kriging with external drift in areas where information was not enough. The estimation of the groundwater depth did not consider the simultaneous use of large wells and piezometers, with exception of the northwestern area, due to the lack of large wells. In the studied area, the presence of piezometers (exclusively dedicated structures for piezometric observations) is mostly associated with karst aquifers and areas with high abstraction volumes for public water supply. Oppositely, large wells are mainly devoted to private use and low volume abstractions.

Drainage density: Drainage density was calculated for six river basins. That gives little variation across the area. Is it possible to map drainage density at a higher resolution,

e.g. sub-basin scale, or a 10 km grid size? This would increase the importance of this parameter.

Answer: Indeed, there was little spatial variation of the drainage density for the studied area, therefore, as suggested by the reviewer, we recalculated this variable considering a 10km resolution grid. The methodology concerning this calculation was corrected in the ms, in lines 202-204. Due to the creation of a new drainage density map, we performed a reassessment of the multicollinearity between variables and the selection of predictors (see section 2.4). This implied recalculating Pearson's coefficients and Principal Components Analysis (PCA), presented in table A2 and figure A2 both in appendix A. It also affected predictors and coefficients in the model linking GDV density to environmental predictors). By affecting model development, model performance (Tables 2 and 3), suitability and coefficient maps were also affected (Figures 7 to 9).

Climate: The authors should provide a bit more explanation on the SPEI and particularly the ombrothermic index calculations. Please explain how/where the latter differs significantly from (and is thus not correlated to) the aridity index.

Answer: Done. Clear explanations on SPEI calculations were already provided in lines 216-223. Since the SPEI predictor was excluded from modeling further explanation would unnecessarily extend the manuscript length. We however briefly altered paragraph lines 224-227, to better explain the discrepancies between SPEI and Ios, and to clarify Ios calculation according to Table 1.

Model development

It is not clear how the parameters were normalized before entering the regression model.

Answer: The explanation of the normalization based on the z-score function was improved and changed to the M&M section, under the chapter 2.5 of Model development. Variables were standardized before entering the regression model through the calcu-

lation of a z-score. To clarify how the standardization was done the following sentence was added to lines 266-268 of the ms: "This allows to create standardized scores for each variable, by subtracting the mean of all data points from each individual data point, then divide those points by the standard deviation of all points, so that the mean of each z-predictor is zero and the deviation is 1.".

How was the soil parameter transformed into a quantitative variable? If all parameters were classified/categorized (as is often done in e.g. factorial regression analysis), this can explain the low influence of the groundwater depth parameter, as there is very little variation (in large part of the area groundwater depth is between 1.5 and 15 m). In this case, I strongly suggest increasing the number of classes for groundwater depth.

Answer: We greatly appreciate your comment. First, we would like to clarify that the main purpose of the model construction is to attribute coefficients of importance to each variable, so that these coefficients can be applied to classification scores given to each variable by expert judgement (table 4) and return a suitability map to groundwater dependent vegetation. This will allow the production of a suitability map where the coefficients of importance applied to each variable were calculated empirically. Therefore, the classification scores given to each variable were not applied in the model calculation, but rather after the local model coefficients were calculated (as a mean to construct the suitability map). The soil parameter was used has a numeric categorical variable (with the values given initially from 1 to 3), through the use of the function as.factor() in R. The usage of this function will insure that the factor is seen as nominal and not as ordinal. Because the remaining variables showed continuous values, only the soil type variable was categorized, and the remaining variables used to run the model were continuous. The scoring applied is presented is Table A1and the explanation in lines 175-178 of the methods section. The reviewer is correct about the groundwater depth variable and its very low variation above 15m. As further explained below, it has not been possible to increase the number of classes for GWDepth, for the weighting factors to be correctly applied to the GWDepth layer in the multicriteria anal-

ysis. To overcome this situation the values of water depth above 15 m were replaced by a value above 15m (15.1m), in order to emphasize the variation observed between 0 and 15m depth, which matters the most to GDV. These values were only used for the model fitting. The species used as proxies for groundwater dependent vegetation are less probable to use water at depths lower than 15m, and so all the range of values above this threshold would be considered as inaccessible by those species.

Please provide references showing that it is common practice to fit the model on a 5% random subsample. Also explain more lines 264-265.

Answer: The sub-sampling size was mostly dictated by computing limitation in the sense that the random subsample size was decreased down to 5% until the GWR model could be fitted. The mean distance between neighbor points using 5% of the original dataset was about 6 km, with a maximum distance of 15km. Nevertheless, we could find a few studies using a 10% random sub-sample of the data corresponding to a 10km resolution grid to perfrom GWR modeling (Bertrand R., 2017), as well as linear regression (Bertrand et al. 2016). The authors were using such subsampling to restrain autocorrelation issues according to Kühn (2007). We modified our text to include those references as well as the benefit for autocorrelation issues in our study, lines 272-277 and in the reference list, lines 743-745 and 900-901. In addition, we calculated basic statistic indicators for the totality of the data and compared with the random subsample. Results are presented in line 276- 279 of the ms and in Figure A1 in appendix A.

Bertrand R. Unequal contributions of speciesâĂš persistence and migration on plant communitiesâĂš response to climate warming throughout forests. bioRxiv, doi.org/10.1101/217497, 2017. Bertrand R., Riofrío-Dillon G., Lenoir J., Drapier J., de Ruffray P., Gégout J.C., Loreau M.. Ecological constraints increase the climatic debt in forests. Nature Communications 7, 12643, doi: 10.1038/ncomms12643, 2016. Kühn, I. Incorporating spatial autocorrelation may invert observed patterns. Divers. Distrib. 13, 66–69 (2007).

Results Overall section 3.2 on environmental conditions mainly consists of an explanation of each of the maps. To support the selection of the five parameters, the authors should provide all the results on correlation and PCA as supplementary material.

Answer: Done. PCA results have now been provided as supplementary material in Table A2 in appendix A and was modified according to new results, due to the construction of a new map of drainage density.

The results of the model suggest, as stated by the authors, a low importance of the groundwater depth on explaining the spatial distribution of GDV (eq. 4). However, nothing is said on how this varies locally within in the area. Are there regions where the role of groundwater is larger? Can these regions be identified?

Answer: We have plotted the local coefficients of all predictors and present it in Figures 07 and 08. In addition, we added some paragraphs with an explanation of the spatial variation of each predictor in the results section, in lines 424-443 and in the discussion section in lines 524-528.

Line 343-344: This requires quite a bit more explanation, but can be easier to follow once the calculation of Ios4 has been better explained in the methods section.

Answer: We altered lines 224-227 and 234-235 to better explain the discrepancies between SPEI and Ios and to clarify Ios calculation according to Table 1.

Line 362-364: Please elaborate on this outcome on the Moran index.

Answer: Bibliographic references for the Moran Index were added in line 416 and the respective references were added to the bibliography. In addition, we extended the results explanation on the Moran Index and the z-score, in lines 415-419.

In eq. 4 the appearance of Soil type 2 and Soil type 3 is not explained.

Answer: After revising the methodology and predictors selection, the predictor soil type was no longer included in the model.

The results would become more interesting with: 1) a more local/regional analysis of the explanatory model and the importance of each of the parameters (in particular groundwater depth);

Answer: The model equation was substituted by a local one including the proportion of the local coefficients from the total variability of all the coefficients for each local GWR model. Local relative coefficients were considered as weighting factors instead of median values (please see revised Equation 6 in the manuscript). We also added a figure corresponding to the local variation of each coefficient in Figure 07 and commented the variations in the result section, in lines 424-443 and in the discussion section, lines 543-528 and in the conclusion section lines 658-661. The relative importance of each variable in the final model is now shown in Figure 08, representing the distribution of the local coefficient values in a box plot.

2) an assessment of the use of more/less/different parameters on the final model. It seems the soil type and to a lesser extent the aridity index are the dominating parameters in the regression model. How does a model based solely on these two parameters perform? And what about including a deeper (2nd-3rd layer) soil parameter to account for water holding and capillary rise capacity? Not much can be stated on the importance of soil type for the groundwater storage (as mentioned in line 495) if only the first soil layer is assessed.

Answer: We tested the effect of removing one of the variables on the model performance and found out that the model performance increased notably when soil types was removed (AIC divided by factor 2, Table 2). The removal of any other variable however, did not seem to impact the model performance as compared to the equation including all formerly selected variables. Therefore, we excluded soil types from the final GWR model and the rest of our analyses and multicriteria analysis. Data on deeper soil parameters was not available for the study area and therefore that information could not be included in the model.

3) scenario analysis: what happens if one or more of the parameters (such as climate or groundwater level) change? You do not have to develop climate scenarios, but an assessment of the impact of a relative change in aridity index or groundwater level on the resulting map would be of high added value.

Answer: The development of climate scenarios or the assessment of the impact of aridity change on GDV suitability was out of the scope of this manuscript, since the actual manuscript is quite long already. The full assessment of climate changes impacts and corresponding uncertainties will be the focus of our next publication. We calculated preliminary results of the relative change of AI and IOS4 expected for the near future (Table a and b below). Our ongoing calculations based on scenarios RCP 4.5 and 8.5 show that AI and IOS4 climate indexes are going to decrease in the studied region (-14 to -33% within 2099), drifting from a mostly dry sub-humid climate (0.5<AI<0.65) to a mostly semi-arid one (0.2<AI<0.5) by 2099 in scenario RCP8.5, and according to the classification of Middleton et al. (1992). Ios4 is also going to suffer a huge drop (-42 to -58% within 2099). Also, while most of the territory could be considered as non-Mediterranean based on the ombrothermic index (Ios4>2) during the historical period 1971-2000, it is becoming mostly Mediterranean by 2099 in scenario RCP8.5 and according to the classification of Rivas-Martínez et al. (2011). To include such preliminary results in the M&M, result and discussion regarding climate change impact would imply to considerably increase the manuscript, while providing an incomplete picture of the changes and associated uncertainties. We therefore chose not to include the suggested assessment of the impact of a relative change in aridity index or groundwater level on the resulting map in this manuscript. Nevertheless, we know discussed the relative importance of each predictor in our final map, which give an insight of how the groundwater dependent vegetation is expected to be affected according to the predicted increased aridity, lines 641-648, 658-662 and 668-673.

Table a. Mean relative changes expected for AI and IOS4 in the near future according to climate changes scenarios RCP 4.5 and 8.5, and respective standard deviations.

Changes were computed considering 30 yr means obtained from an ENSEMBLE of eleven EU-CORDEX climate models.

Table b. Evolution of percentiles 10 and 90 values of AI and IOS4 in Alentejo from the present to the near future according to scenarios RCP 4.5 and 8.5. ECAD are observed values for the reference period 1971-2000. Historical values for the reference period 1971-2000 as well as predicted values for the future were simulated by an Ensemble of EU-Cordex models.

Discussion Much of the discussion on the modelling approach is more of a summary of the manuscript, particularly lines 425-439. I miss the interpretation of the results obtained by regression modelling, and the this could further be enriched by the discussion of the added results as proposed above.

Answer: The discussion section has been considerably modified. The dominant impact of aridity on tree density and GDV suitability is now much more discussed, as well as the lower impact of groundwater depth. The relative weight of each predictor is also discussed and considered in the key limitations and conclusions sections. (see mostly lines 443-444, 537-542, 641-646 and 658-662).

PART 2: SUITABILITY MAPPING Suitability map building The authors decide to attribute the minimum score (in terms of suitability) to areas where groundwater depth is smaller than 1.5 m, considering that vegetation extracting water from shallow depths belongs to another type of GDV. This distinction between shallow and deep groundwater dependent vegetation, which I indeed think is useful (as most vegetation can use water in the first 1.5 m if present) needs to be briefly elaborated upon.

Answer: Providing a less probable score to host the GDV to the 0-1.5m GWDepth was made to exclude riparian vegetation and shrubby species which primarily use the water from streams and the superficial soil layer. An additional explanation and references were added to the manuscript in lines 302-307: “The depth class between 0 and 1.5m was based on the riparian vegetation in semi-arid Mediterranean areas which is

mainly composed of shrub communities (Salinas et al., 2000) and present a mean rooting depths between 1 and 2m (Schenk and Jackson, 2002). The most common tree species rooting depth in riparian ecosystems is normally similar to the depth of fine sediment not reaching gravel substrates (Singer et al., 2012), but not reaching levels as deep as deep-rooted species.”.

Line 284-286. I do not understand why shallow groundwater flow would be expected at steep slopes. Normally steeper slopes are found in mountainous areas, where groundwater levels are deep.

Answer: The reviewer is correct, and we appreciate for noticing the error. The sentence was corrected in lines 308-309 and the term water flow was substituted by runoff.

Results The main finding here is that “suitability to GDV in the Alentejo region was mainly driven by soil type”. That is obvious, as the weight of this parameter is by far the largest in the suitability index (and given by two soil type variables)! The same holds for the observation “The aridity index also showed a strong influence on GDV’s suitability”, as the weight of the aridity index is highest following that of soil type. I would strongly suggest analysing alternative weights for each parameter (based for instance on the Delphi panel) and evaluating the corresponding sensitivity of the outcome, as well as the degree of success in the validation procedure.

Answer: Unfortunately, it has not been possible for us to perform this analysis within the time provided to review our manuscript. We hope that the discussion on the relative importance of each predictor in the model will be satisfactory enough for the reviewer, considering that every other request was fulfilled.

Line 395-396: “high aridity values restricted GDV’s suitability in the south”. Again, in my view it is exactly the opposite, as a high aridity is classified as class 3, i.e. of high suitability. In the south in fact aridity index is lowest, indicating the highest aridity and therefore higher suitability for GDV. I think I might have detected a mistake in the resulting map of Figure 7. Where aridity index (AI) values are low, corresponding suitability

value is high (Figure B1b), which means that overall suitability should also increase in those areas (towards the southeast). In the map of Figure 7 the values actually decrease in that area, which is contrary to what would be expected and could result from a negative weight being assigned to this parameter (as it also has a negative coefficient in the regression equation). If this is the case, the presentation and interpretation of the results on suitability mapping needs to be redone.

Answer: After thoroughly verifying the model calculations (Eq. 6) and the weighting factors used for the final multicriteria analysis (Figure 09), we must agree with the reviewer that it was a mistake to apply a negative weighting to the Aridity Index layer. Indeed, where real values of AI were low (indicating a more arid area), our scoring was high in the multicriteria analysis. To directly apply a negative weight, we should have the real predictor values and the predictor scores co-varying (or growing) accordingly. We also verified that the same logic should be applied to the other quantitative variables Slope, Ios4 and GW Depth, since scoring and real values variation were opposite. However, in the revised manuscript we have adopted different scores for the Aridity Index (scores 1, 3, 2) which were not varying linearly, and it was no longer possible to apply a linear scoring. The same was applicable in the case of GWDepth, when we came to a dead end because scoring was not varying linearly according to class values (scores 1, 3, 1). As a solution we calculated the proportion of each local coefficients from the total variability of all the coefficients for each local GWR model (Eq. 6) as a local weighting factors reflecting the relative relevance of each predictor locally. This allowed us to apply scores not varying linearly and still interpreting the results easily. This way, the weighting factor obtained in from the proportions could be directly and correctly applied to the GW Depth and Aridity Index layers.

One example of this wrong interpretation is in lines 376-378, where the authors state that the positive impact of the rivers on the GDV suitability "is due to a higher water availability reflected by the values of omborthermic and aridity indexes. In my view it should be the contrary, i.e. due to a lower water availability, indicating a higher suit-

ability for GDV. Moreover, the positive impact is not visible in the map of Figure 7. And why is there a higher groundwater depth near the river? You would expect groundwater levels to be shallowest near the river. Another example of this is in discussion section, where the authors state that "The lower suitability to this vegetation in the eastern part of the studied area can be explained by less favorable climatic and geological conditions, resulting from the combination of a high aridity index and low water retention at deep soil layers". It is again the contrary, as the ariditiy index in this (south)eastern area is lower, indicating a higher suitability and therefore higher values on the map of Figure 7. Moreover, it is not clear why the "deep soil" layer is mentioned here now, if only the first soil layer has been analysed.

Answer: We appreciate the referee comment and agree with it. Indeed, groundwater levels are expected to be higher near the river, mainly in alluvial aquifers (associated with gentle slopes). However, the opposite also occurs in areas where the rivers are associated with hard rock aquifers (generally associated with steep slopes) and where the relation surface/groundwater is more heterogeneous. The slope predictor, also considered in the presented methodology, distinguishes these occurrences.

In Figure 7 please indicate how the values were calculated.

Answer: A thorough explanation was added in the methods section, in lines 327-333. The explanation in the methods section in the ms reads: "The final GIS multicriteria analysis was performed using the Spatial Analyst Tool by applying local model equations obtained for each of the 6242 coordinates of the Alentejo map (Eq.4), Suitability = Intercept + coef1 * [real value X1] + coef2 * [real value X2] + coef3 * [real value X3] + . . ., (4) with brackets representing the reclassified GIS X layer corresponding to the scoring and coefpx indicating the relative proportion for the predictor x.". The final equation used for the calculation of the suitability map is presented in the results section, in lines 406-406, and is presented in the Equation 1 below. Suitability = Intercept + AI coefp * [reclassified AI value] + Ios4 coefp * [reclassified Ios4 value] + GWDepth coefp * [reclassified GWDepth value] + Dd coefp * [reclassified Dd value] + slope coefp

* [reclassified slope value], (1) If the authors decide to do the analysis per river basin, they should indicate the river basin boundaries in Figure 1.

Answer: As suggested by reviewer 1, we decided to use a 10 km grid mesh instead. The methodology was corrected in lines 203-204.

Line 382-383: “this high likelihood was hindered by the type of soil present in that área In terms of soil type in the Tagus basin”. That is not true, as the suitability is mostly class 3 in the Tagus river basin.

Answer: The sentence was deleted according to the new results of the revised manuscript.

Line 416-419 belongs to the discussion section, not the results section.

Answer: The paragraph was deleted according to the new validation performed in the revised manuscript.

Technical corrections Overall the text is well written and structured, the main comments above concern the content of the manuscript. Line 47: decreased precipitation

Answer: Done in line 44.

Line 56: An integrated multidisciplinary methodology

Answer: Done in line 53.

Line 63: do not include

Answer: Done in line 60.

Line 167: listed in Table 1

Answer: Done in line 158.

Line 169: 2.3.1 Slope and soil characteristics

Answer: Done in line 161.

Line 205: division of the basin area by the total stream length

Answer: Done in line 204.

Line 244: was evaluated

Answer: Done in line 246.

Line 256: based on the selected variables

Answer: Done in line 259.

Line 277: score from 1 to 3

Answer: Done in line 296.

Line 367: In the GWR model

Answer: Done in line 421.

Line 380: with the exception of

Answer: Done in line 453-454.

Line 948: Table 2: Groundwater Depth

Answer: This table was eliminated from the revised manuscript. The variable Groundwater depth was, from now on, referenced as GWDepth.

Line 956: suitable areas for GDV

Answer: Done in line 1147, in Table 4.

Figure 1: add catchment limits

Answer: Done in the new version of fig01.

Figure 4: change soil colours, or combine

Answer: The map of soil type was removed form Figure 04.

Line 990: what kind of residuals?

Answer: This was clarified in line 1171.

Figure 7: consider changing the colour coding

Answer: A new suitability map was calculated, with new colors by classes, and was added as Figure09.

Figure B1: present the maps in the same order as in Figure 4.

Answer: Done in Figure B1.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-208/hess-2018-208-AC2-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-208, 2018.

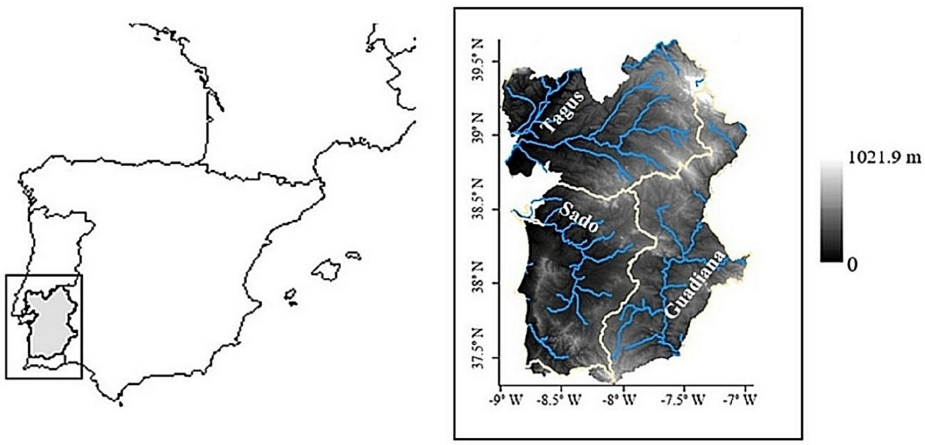

Fig. 1. Fig01

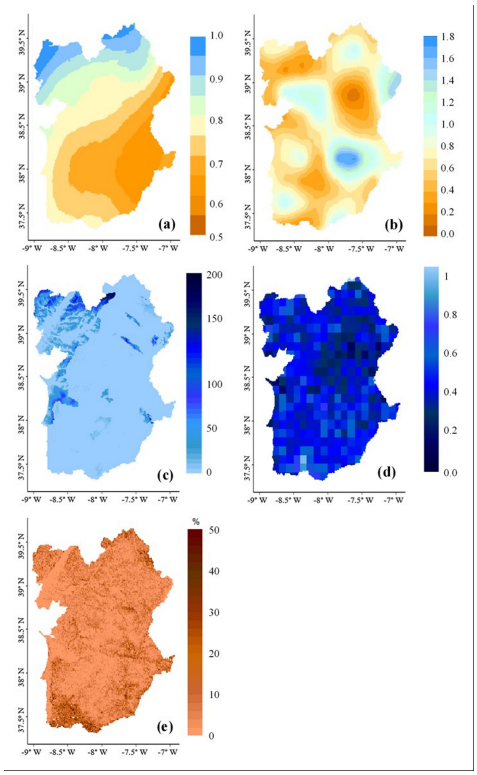

Fig. 2. Fig04

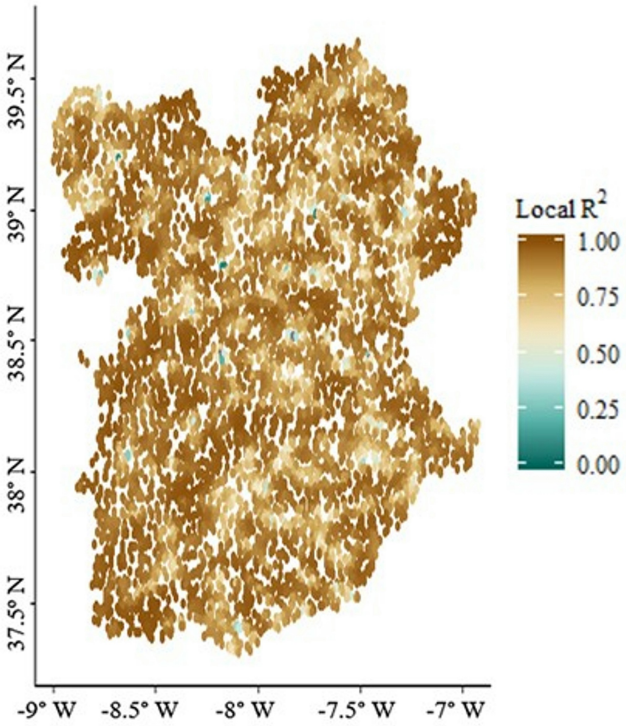

Fig. 3. Fig05

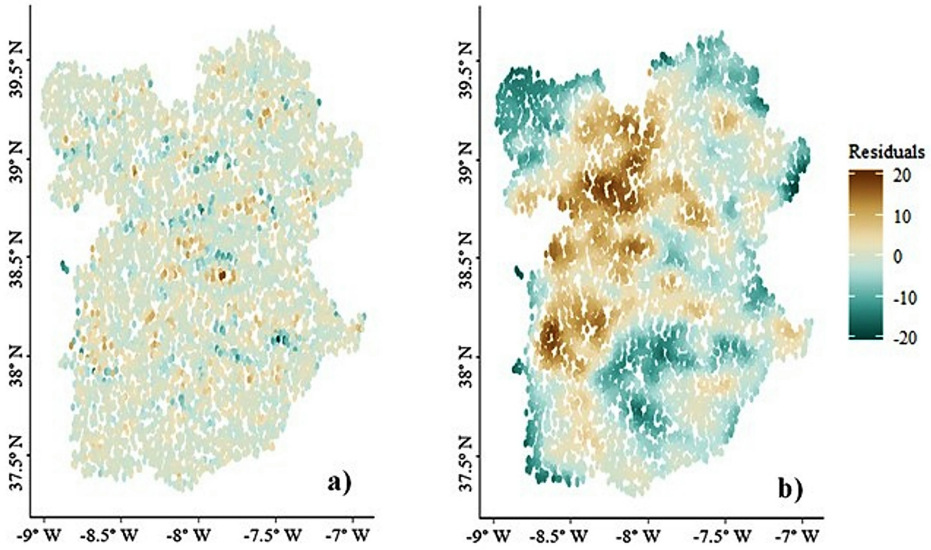

Fig. 4. Fig06

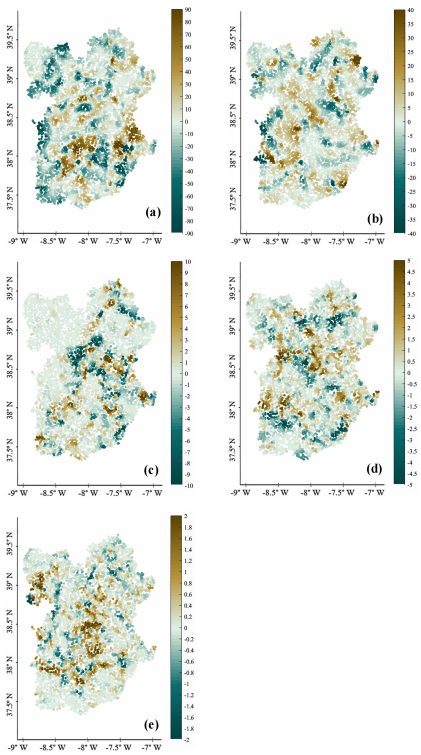

Fig. 5. Fig07

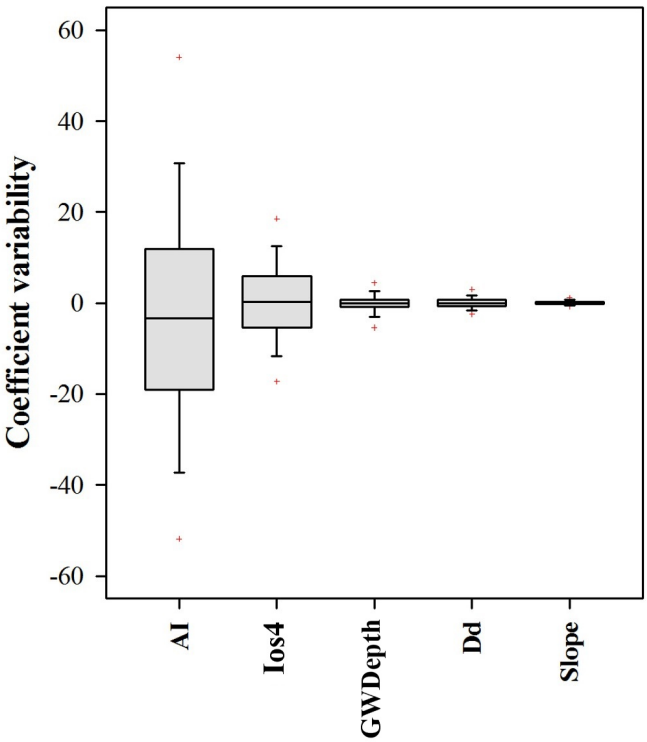

Fig. 6. Fig08

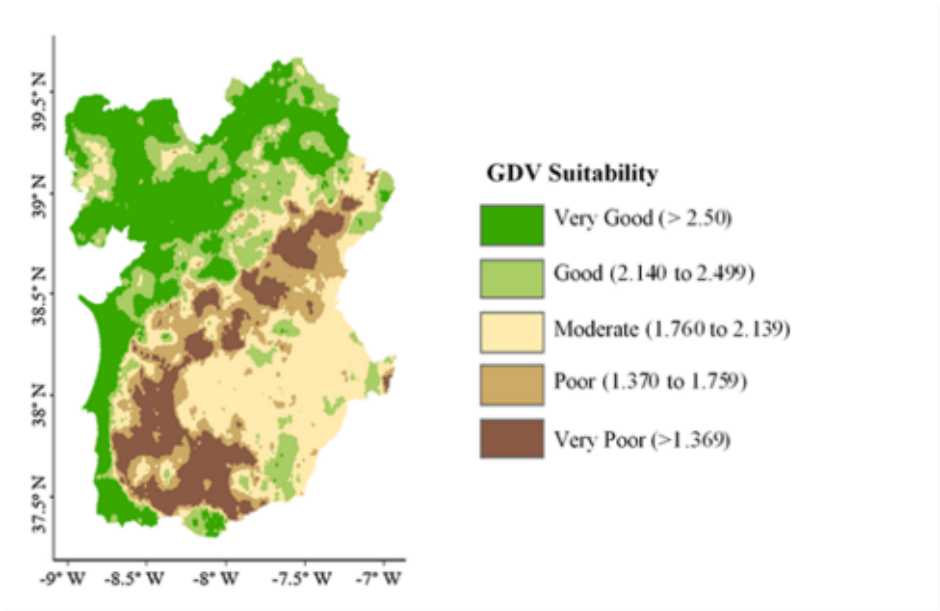

Fig. 7. Fig09

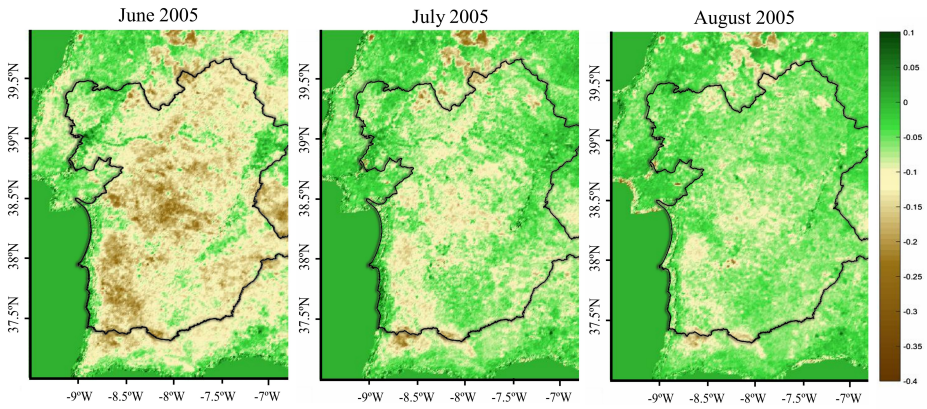

Fig. 8. Fig10

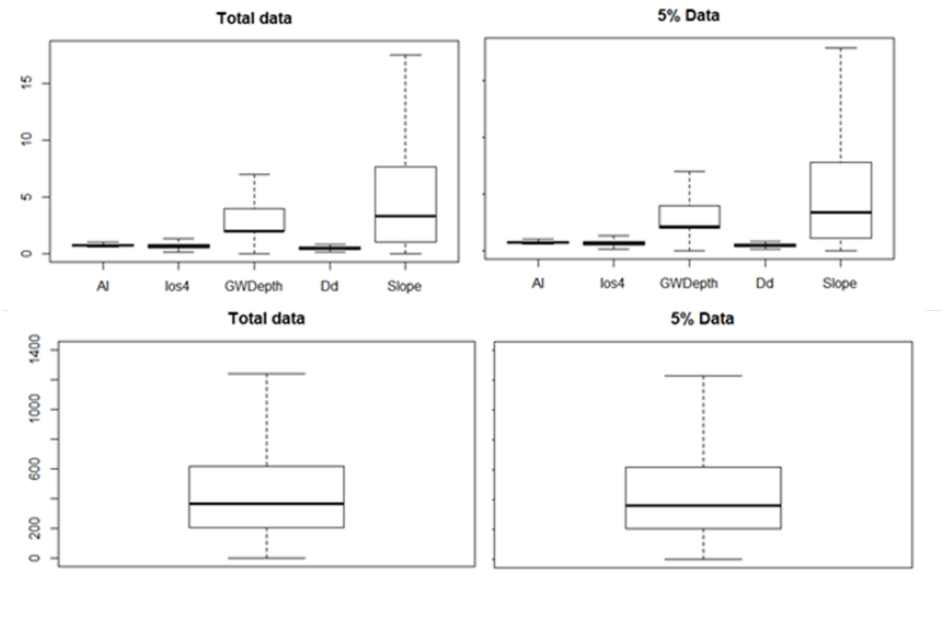

Fig. 9. FigA1

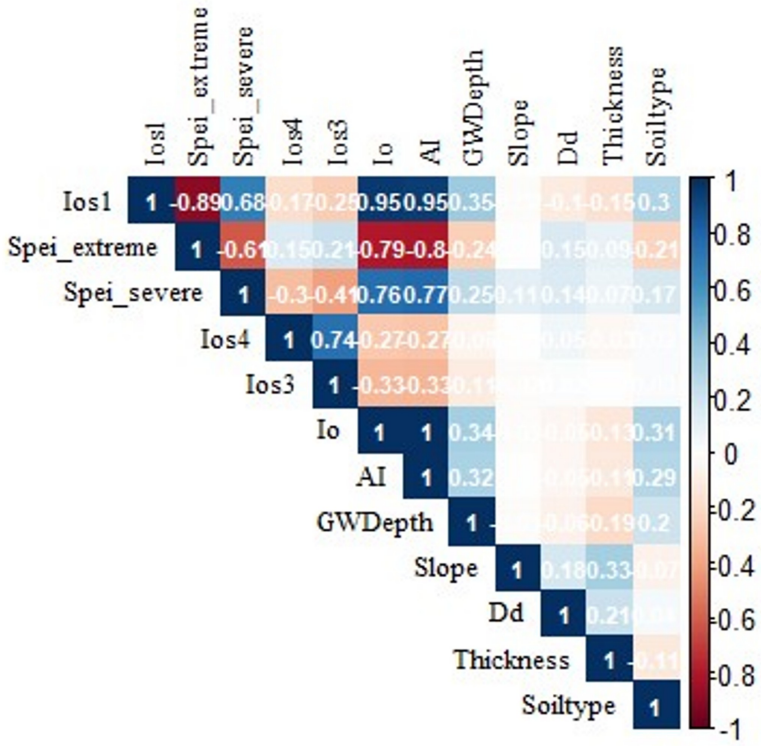

Fig. 10. FigA2

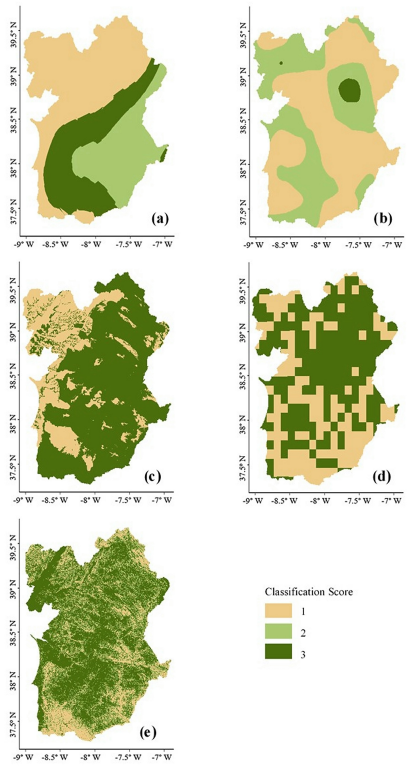

Fig. 11. FigB1