Instituto Dom Luiz, Faculty of Sciences
University of Lisbon
Campo Grande, 1749-016 Lisbon, Portugal
Tel: +351 927464067
E-mail: inesgmarques@fc.ul.pt

June 20th, 2019
Editorial Department of *Hydrology and Earth System Sciences*

Dear Dr. Miriam Coenders-Gerrits,

Please find enclosed the revised version of the manuscript (reference hess-2018-208) entitled "*Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area*".

We carefully considered and addressed each reviewer´s comment accordingly. In our joined letter you will find our answers and changes made, indicating the highlighted line numbers.

We are very thankful for giving us the opportunity to improve our manuscript to be accepted in your journal. To facilitate the identification of changes along the manuscript, a version of the manuscript with tracked changes was uploaded in the journal platform.

We once again declare that 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.

Thank you for considering this revised manuscript for publication. Please do not hesitate to contact me if you require further details.

With our best regards, sincerely,
Inês Gomes Marques (on behalf of all authors)

Dear Reviewer #2,

Please find enclosed the revised version of the manuscript "Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area".
We are once again very grateful for your precious and pertinent revision of our manuscript.
All yours suggestions were carefully considered and addressed. In the present letter,
you will find our answers to your comments and changes made, with corresponding lines highlighted. To ease the revision, we highlighted line numbers in yellow in our answers.

We are very thankful for your detailed assessment, which allowed a very significant improvement of the overall manuscript. To facilitate the identification of changes along the manuscript, a version of the manuscript with tracked changes was uploaded in the journal platform.

All the information included in this manuscript has been approved by all authors. The authors declare no conflict of interest.

Thank you for considering this revised manuscript for publication.
Please do not hesitate to contact us if you require further details.

With the authors best regards.

**Report #1, Reviewer #2**

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

The manuscript has been quite significantly altered from its previous version, with many relevant and good aspects added to the model development, suitability map building, map evaluation and sensitivity assessment, in terms of methodology, results and discussion. It really has become an interesting paper to read, with very good use of references throughout. I maintain my opinion that the strongest part is the regression model, and that this model can be applied very well in the future for scenario analysis within the same study area. I continue to believe that the sustainability map is less interesting, because it basically uses the regression model that was already locally optimised using GWR to calculate a new map, but it is only applicable to the study area, precisely due to the local nature of the regression coefficients. Moreover, it does not fit the original vegetation (Kernel density) map as well as the authors claim based on their validation.

***Answer: It is not surprising that the final suitability map does not exactly fit the original Kernel density. Indeed, the proxy species (Quercus suber, Quercus Ilex and Pinus pinea) can perfectly grow under more mesic Mediterranean climate conditions (sub-humid), without relying as much on groundwater to survive as in more xeric areas (semi-arid) (Abad Vinas et al., 2016). Their presence/abundance is only an indication of a possible use of groundwater. This is also why we consider that our final map obtained after the multicriteria analysis provides a more reliable indication of the higher likelihood for groundwater use by facultative deep-rooted phreatophytes species in Alentejo. We also believe our final map also provides a better estimation of the relative contribution of groundwater used by plants to remain alive, than the information given by the model alone. A paragraph was added on 617-620 to better explain the benefits of the final suitability map as compared to the model alone.***

Please allow me to start with these two main concerns I would like to see addressed in the discussion.

1. The calculation of R-squared of the GWR model provides very good results. Notwithstanding, the resulting local coefficients vary largely, from highly positive to highly negative. Moreover, the variations sometimes occur on very small distances. This means that the effect of Ai or groundwater depth on groundwater dependent vegetation can vary from highly positive to highly negative throughout the area and even within very short distances. This seems purely a statistical exercise with apparently little physical meaning and needs to be addressed in the discussion. What does it mean? How then is this method valid and applicable elsewhere?

***Answer: We agree on the fact that our modelling approach is stochastic and can be considered as a "statistical exercise". We also agree on your critic regarding the weak physical meaning of the model coefficients due to their high spatial variability. This is another argument for us not to use the model alone for prediction purposes. We are conscient that the method we developed is only locally optimized and thus difficult to apply in other regions, even under similar climate conditions, unless the methodology is fitted to local conditions/predictors. We modified the manuscript in the discussion (lines 630-632) and in the "key limitations" section (lines 742-746) to address those issues.***

2. Given the high weight of the aridity index (Ai) in the regression map, the groundwater dependent vegetation (GDV) suitability map now closely follows the Ai categorized map (Fig. B1a), as also mentioned by the authors. The good agreement observed by the authors between the

suitability map and the groundwater depth map, in my view is a coincidence, as the groundwater depth map in fact follows the aridity index map. In other words, in the more humid areas the groundwater level seems shallower, and vice-versa. In addition, the GDV suitability map does not show a good correspondence with the GDV occurrence map (Fig. 1), unlike the previous suitability map that was produced in the first version of the manuscript. In the former map (version 1 of the manuscript) soil type was the most important parameter, but that parameter was now taken out. As a direct consequence, the highest GDV density in the central north now occurs in an area of very poor to poor mapped GDV suitability, whereas in the southeastern area the GDV density is very low in an area of very good suitability. I acknowledge that the reality is always more complex and that the authors already refer to this in their discussion, but please also address the issues I have mentioned.

***Answer:***

***As explained above, the proxy species (Quercus suber, Quercus Ilex and Pinus pinea) can perfectly grow under more mesic Mediterranean climate conditions (sub-humid), without relying as much on groundwater to survive as in more xeric areas (semi-arid). Their presence is only an indication of a possible use of groundwater. The study provided by Pinto et al. (2013) have shown that Cork oak can perfectly thrive were very shallow groundwater is available while suffering drought stress were groundwater source is lower (although using groundwater in both sites). We believe this satisfactory explains the discrepancies between the GDV density and suitability maps you question. We addressed the mismatches between maps in the result section, lines 563-565 and modified a paragraph in the discussion section, in lines 643-659.***

***Abad Viñas, R., Caudullo, G., Oliveira, S., de Rigo, D., 2016. Pinus pinea in Europe: distribution, habitat, usage and threats. In: San -Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mau ri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, p. E01b4fc.***

***Pinto C., Nadezhdina N., David J. S., Kurz-Besson C., Caldeira M.C., Henriques M.O., Monteiro F., Pereira J.S., David T.S. Transpiration in Quercus suber trees under shallow water table conditions: the role of soil and groundwater. Hydrological processes, doi: 10.1002/hyp.10097, 2013.***

The fact that the suitability maps fits well with the NDWI map, could be a logical consequence of the fact that the latter represents moisture content in vegetation. Why would the highest stress be indicative for groundwater dependency? Wouldn't you expect stress to decrease if the trees have access to groundwater?

***Answer: Figure 10 does not present NDWI values, but anomalies considering the months of June, July and August of the extremely dry year of 2005, in reference to the median NDWI value of the same months over the period 1999-2009 (lines 544-545)). In June of the extreme dry year 2005, GDV vegetation experienced the highest moisture stress, as observed on Figure 10a by the negative NDWI anomaly values. GDV still contains moisture however, that changes/decreases with the onset of the summer period (aggravated by the dry winter-spring of 2005), thus reaching a point in August were the GDV has a very low water content, as expected in the end of the drought season (~null anomaly on Figure 10c). Oppositely, the vegetation over areas that do not manage to cope with summer drought (bare soil, grassland, shrubs...) uses to have the lowest moisture content since June until August with no change (null anomaly indicated in green that remains green from June to August on figure 10a-c). Therefore the GDV shows the highest absolute NDWI anomaly (highest leaf water loss), in spite of the use of groundwater to survive. Further former studies by co-authors of the present work have already shown that groundwater uptake by trees only take place in late June after the onset of the drought period (Kurz-Besson et al. 2006 & 2014, Otieno et al. 2006, David et al. 2013, Pinto et al. 2013). Those studies have also shown that trees grew new roots in deeper soil layers only after trees experienced drought stress. In extreme dry years, the piezometric drawdown is***

***expected to difficult GDV's physiological performances (Antunes et al. 2018). We are confident that those studies are in agreement with the NDWI anomaly validation maps provided. Nevertheless, we re-write the paragraph 3.5 for more clarity and added the references cited here above in the manuscript to support our arguments (Lines 547-565). We also modified changed figure 10 colours and caption in order to highlight the NDVI anomaly behavior aiming to avoid misleading issues (lines 1241-1245, 1327-1331).***

***Otieno, D.O., Kurz-Besson C., Liu J., Schmidt M.W.T., Lobo-do-Vale R., David T. S., Siegwolf R., Pereira J.S., Tenhunen J.D. (2006) Seasonal variations in soil and plant water status in a Quercus suber L. stand: roots as determinants of tree productivity and survival in the Mediterranean-type ecosystem. Plant and Soil 283: 119-13***

***Kurz-Besson C., Otieno D., Lobo-do-Vale R., Siegwolf R., Schmidt M.W.T., David T. S., Soares David J., Tenhunen J., Pereira J. S., Chaves M. (2006) Hydraulic lift in cork oak trees in a savannah-type Mediterranean ecosystem and its contribution to the local water balance. Plant and Soil 282: 361-378.***

***Pinto C., Nadezhdina N., David J. S., Kurz-Besson C., Caldeira M.C., Henriques M.O., Monteiro F., Pereira J.S., David T.S. (2013) Transpiration in Quercus suber trees under shallow water table conditions: the role of soil and groundwater. Hydrological processes.***

***David T.S. Pinto C.A. Nadezhdina N. Kurz-Besson C. Henriques M.O. Quilhó T. Cermak J. Chaves M.M. Pereira J.S., David J.S. (2013) Root functioning, tree water use and hydraulic redistribution in Quercus suber trees: A modeling approach based on root sap flow. Forest Ecology and Management 307, 136–146.***

***Kurz-Besson C., Lobo do Vale R., Rodrigues L., Almeida P., Herd A., Grant O.M., David T.S., Schmidt M., Otieno D., Keenan T., Gouveia C., Mériaux C., Chaves M.M., Pereira J.S. (2014). Cork oak physiological responses to manipulated water availability in a Mediterranean woodland. Journal of Agricultural and Forest Meteorology 184, 230-242.***

***Páscoa P. Gouveia C., Kurz-Besson C. Identificação de vegetação dependente de água subterrânea na península ibérica através de deteção remota. 10º Símposio de Meteorologia e Geofísica da APMG, Lisboa, Portugal. 2017, https://drive.google.com/file/d/0B4ZF89Veh6ziZVVCbUxBZXh1MTA/view***

Some other comments are given below:

The abstract is well written.

The introduction provides a very good overview on the need of study, but could mention the other work/studies carried out so far in the field. That is currently limited to one sentence (ln 127-129), so that the paper does not show the added value of the implemented methodology as compared to existing studies, some of which are actually referred to later on in the manuscript (e.g. Barron et al., 2014; Condesso de Melo, 2015; Costa et al., 2008; Doody et al., 2017). Therefore, no new references are needed.

**Answer:**

***The introduction section was slightly restructured. We rearranged the short overview of the studies carried out in the field (now in lines 54-78), avoiding turning the introduction any longer. We also added a new reference based on field surveys and showing that Pinus pinea relies on groundwater to cope with summer droughts. We also indicated the added value of the***

***implemented methodology in lines 132-139 and further improved the end of the introduction section in lines 140-154.***

In material and methods, section 2.3.1, attributing a low GDV suitability score to soils of high clay content can be debated. Soils of a finer texture will have large extinction depths due to an increased capacity of capillary rise. I would expect coarser soils to have vegetation of lower groundwater dependency. Please briefly elucidate on this aspect.
***Answer: We agree with your comment and the debate in the matter. Nonetheless, in this specific geographical region, deep rooting species reaching deep soil layers or groundwater are disfavored in waterlogged soils highly favored by clay content (Garcia et al. 2017; Ignacio Perez-Ramos & Marañón 2009; Dinis et al. 2014). We also believe that soils rich in clay will rather favor non-GDV species for providing more available water in shallow soil depths. This is not happening in sandy soils, therefore we gave a better score to those. We had already briefly justified this choice in our former version of the manuscript, (now in lines 208-212) and added a few more words to better justify our scoring choice.***

Garcia et al. 2017, https://ir.library.oregonstate.edu/downloads/wp988k05k;
Ignacio Perez-Ramos & Marañón 2009, https://www.researchgate.net/publication/222234643;
Dinis et al. 2014 https://core.ac.uk/download/pdf/62473102.pdf, (page 60)

In the model development (material and methods, section 2.5), how many data points are used (and what is the search radius) for the calculation of local model coefficients?

***Answer:***
***The number of points (6214) used to fit the model was already in the previous version, (now in lines 311-312). We corrected the sentence for the lector to understand that ultimately 6214 points were used to fit the model (line 311).***
***Before fitting the GWR model an Adaptive Kernel was applied to the data to find a search radius (as explained in lines 323-324 of the manuscript) that would minimize the error of the localized regressions. The adapted search radius, given locally, was obtained through minimization of the CrossValidation score. We improved the methodological explanation in lines 294-295 of the manuscript.***

In section 2.7 of material and methods briefly explain for what purpose the NDWI anomaly map was calculated.
***Answer: We added a sentence to include this missing information in paragraph 2.7 on lines 391-393.***

Please explain why you select slope (s) rather than soil thickness (S), if the latter has a higher correlation with principle component axis 2 (PC2).

***Answer: As explained in lines 318-321 of the manuscript, variables were selected under a sequential procedure. Both slope and thickness did not show correlation values higher than 0.4 and therefore where not discarded from the initial variables selection. If predictors showed correlations below 4, than the ones with the lower correlation values would be chosen. Thickness was removed from the final variable choices because it showed higher correlations with the remaining variables, as opposed to slope that showed lower correlations with the remaining variables.***

What happens to R-squared when reducing the set to four or even three variables? Given the large weight of Ai and O4 I would expect the impact to be small. Have you considered using a reduced set? This would largely facilitate the application of the method in other areas.

***Answer: The model performance assessed with global $R^2$ was little affected when only 2 or 3 predictors were used, remaining close to 0.99. Also, our modelling approach in this manuscript was only performed to provide weights for the GIS layers included in our final multicriteria analysis. On our last revision, we removed the soil type from the model equation because it drastically weakened its performance. The remaining predictors, however, did not affect the performance of the model as much, with $R^2$ remaining between the range of 0.98 to 0.99. We thus choose to keep the remaining predictors in the model (especially the groundwater depth) because of the objective of our study, in spite of their lower contribution to the model.***

Other minor comments and technical corrections:

Ln 17: delete the word "scenarios"
***Answer: Done, now line 38.***

Ln 19: delete the words "the density of"
***Answer: We improved the sentence, now line 19-20.***

Ln 25: "closely followed": this is not true. The other three parameters (groundwater depth, drainage density and slope) follow at a large distance, i.e. they are of much lower importance in the regression model.
***Answer: We corrected the sentence as "Climatic indices were the main drivers of GDV density, followed with a much lower influence by groundwater depth, drainage density and slope", now in line 25.***

Ln 28: "relative proportion". Please briefly clarify what it means. Is it the local coefficient divided by sum of local coefficients? When negative, do you use absolute values (which would make sense)? This needs to be explained in detail in section 2.6 (pg 11 ln 329-341).
***Answer: This as been clarified in lines 372-373, by adding "The relative proportion of the local coefficient x was calculated as the ratio between the modulus of the local coefficient x and the sum of the modulus of all local coefficients."***

Ln 60: include
***Answer: Corrected, Line 55***

Ln 61: "subsurface groundwater" seems a pleonasm, although I understand what you mean, when comparing it to "surface groundwater". Perhaps you could consider using the terms "emerging groundwater" vs. "resident groundwater".
***Answer: We totally agree with this suggestion. Therefore we modified the text accordingly along the manuscript (lines 56, 57, 74, 77, 120, 346).***

Ln 62: "a visible source"
***Answer: Done, line 58***

Ln 64-65: place the references after GDE
***Answer: Done, line 71***

Ln 74: "relying on", perhaps use "entirely relying on"
***Answer: Done, line 83***

Ln 76: "root system"
***Answer: Done, line 85***

Ln 115: "rising temperature"
***Answer: Done, line 125***

Ln 129-130: rephrase "coefficients proportions", e.g. to "coefficients as proportion of total sum of absolute coefficients".
***Answer: Done, line 147-148***

Ln 184: "low drainage capacity", "high clay fraction"
***Answer: Done, line 208-209***

Ln 325-328: lower drainage density leads to higher suitability, which is correct, but the explanation is incorrect, as the explanation in fact suggests the opposite, or so it seems.
***Answer: We improved the sentence, lines 355-359.***

Ln 342-343: I suggest using "representing" instead of the word "referred".
***Answer: Done, line 375-376.***

Ln 402: and in the south?
***Answer: Done in line 439.***

Ln 422: the maximum value on the map seems much higher than the value indicated in the text (0.714).
***Answer: We thank the reviewer for noticing the mistake. Indeed this was a typo, the true maximum value, excluding two outliers, is 1.166. This information was corrected in line 459.***

Ln 488: I suggest changing to: "poor suitability to GDV, corresponding to"
***Answer: Done, line 525.***

Ln 572: "did not only allow"
***Answer: Done, line 625.***

Figure 3: What are the units in this figure?
***Answer: We add a sentence to the figure legend as "The scale unit represent the number of occurrences per 10km search radius (~314 km2)" lines 1224-1225, 1294-1295. Note that ICNF forest inventory only provided information on the presence of each dominant and secondary species on 500m mesh points and their corresponding cover percentage. Therefore, on an area of 1 km2 the maximum occurrence possible is 4, thus on our map the maximum value is 4*314=1256. We therefore also modify the M&M section on heatmap accordingly, lines 176-177 .***

Figure 4: The reference to the different maps in the figure title is incorrect. Figure 4a is aridity index, not soil type, etc.
***Answer: We truly apologize for this mistake. This has been now corrected, lines 1226-1227, 1298-1301.***

Figure 10: I would not use green to indicate highest stress.
***Answer: Colors on Figure 10 have been modified, line 1326.***

Table 4: Values for slope and aridity index are incorrect in the table (the order of the scores 1-3 is inversed, as can be seen in the maps of Fig. B1, which are correct).
***Answer: Thank you for noticing these mistakes. This has been corrected in Table 4, line 1281-1282.***

Dear Reviewer #3,

Please find enclosed the revised version of the manuscript "Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area".

We did our best to carefully address all your concerns. In the present letter, you will find our responses to each comment and changes made in the manuscript. To ease the revision, we highlighted line numbers in yellow in our answers.
We also attempted to provide a better evaluation of the importance of each predictor in the final model and improved the discussion section accordingly.
To facilitate the identification of changes along the manuscript, a version of the manuscript with tracked changes was uploaded in the journal platform.
All the information included in this manuscript is completely original and has been approved by all authors.
Also, we thank you for considering this revised manuscript for publication.
Please do not hesitate to contact us for any further needed detail.

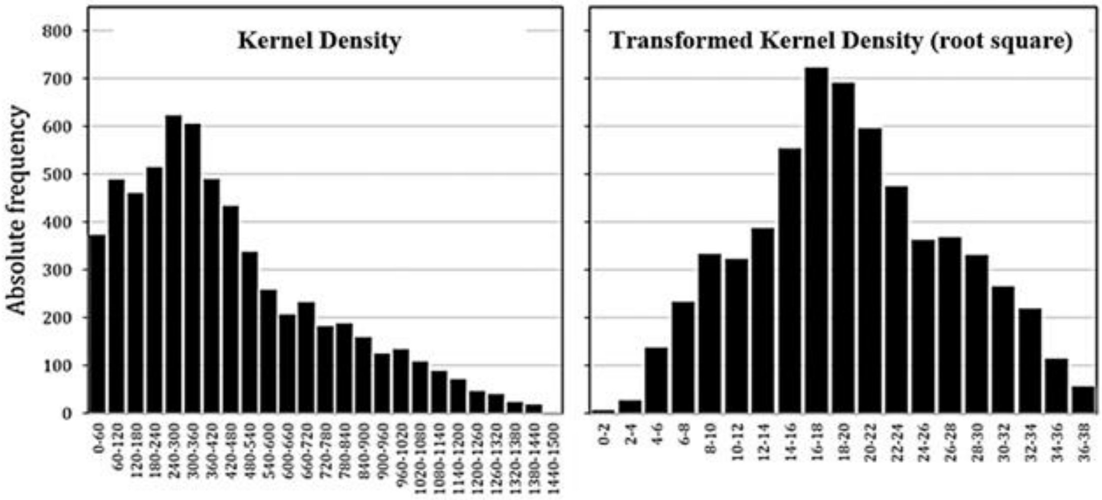
With our best regards, sincerely

---

**Report #2, Reviewer #3**
Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)
The whole paper should be condensed and restructured. The relevance of the study should first be established in introduction by presenting the field of vegetation suitability mapping in ecology (with a better review of previous research), establish the niche by indicating the gap in the present body of literature, and finally present the aim and the approach of the study.
Next, the methodology should be clearly established, starting by the choice of the modeling method which appears to be a linear regression, improved in order to take into account spatial correlation of the explaining variables.

***Answer:***
***The Geographically weighted regression (GWR) extends the ordinary least squares (OLS) regression by considering spatial nonstationarity in variable relationships and allowing the use of spatially varying coefficients in linear models while minimizing spatial autocorrelation. We added a few words in lines 307-309.***
***The normal distribution for predictors is only recommended while using the linear OLS model when the model is used for statistical inference or to calculate confidence intervals . In our study, we only used the modeling approach to provide weighting factors for the GIS multicriteria analysis performed to obtain the final suitability map. We thus assumed that for such purpose predictors' normality was not necessary.***

The data should first be plotted to illustrate their departure from Normality. In second instance, the choice of transformation methods should be justified.
***Answer:***
***Although our dependent variable (Kernel density) did not rigorously match a normal distribution after root square transformation, the distribution shape was approximated to meet the linear model assumption (see Figure below). Also we relied on the article by Li et al. 2012, (https://iovs.arvojournals.org/article.aspx?articleid=2128171), which stipulates that when the dependent variable is not distributed normally, the linear regression remains a statistically sound technique in studies of large sample sizes (i.e., >3000), which can be used anyway, even if the normality assumption is violated.***
***The square-root transformation of the response variable was already indicated in lines 297-299.***




Regarding the criterium of groundwater availability in particular:

1) Why should the soil type, aquifer permeability or aquifer transmissivity be relevant for the growth of groundwater dependent vegetation?

***Answer: We believe it would be relevant for the presence and permanence of more superficial groundwater accessible to roots.***

2) If groundwater levels need to be used as suitability criteria for a type of vegetation, the fluctuation regime need to be established (for example mean levels, 5% low and high quantile determine over a given time period)

***Answer: We think the reviewer is correct, however groundwater depth data retrieved from large diameter wells (blue triangles in figure 02) had only one single measurement. These data points covered most of the study area, thus there was not enough data to establish a temporal fluctuation regime. This weakness was already discussed in lines 660-674.***

3) The interpolation method needs to be better described. A suggestion is to follow the method used by Peterson and Barnett [2004]

***Answer: The method suggested by Peterson and Barnett [2004] (Kriging with External Drift) was also tried with the groundwater datasets used in this study. However, the resulting map of groundwater depth showed incoherent values, therefore we proceeded with the double approach: Ordinary Kriging for karts and porous aquifers and linear regression for undifferentiated geological type. We added a further explanation of the Ordinary Kriging method to lines 233-235 : "The ordinary kriging was calculated using a semivariogram in which the sill, range and nugget were optimized to create the best fit of the model to the data."***

4) Why is the drainage density relevant in the method if the water table levels are known?

***Answer: Groundwater supply at deeper levels is important for groundwater dependent vegetation survival, since there is no other source of water during the dry season. However, when a large river system is present, water will be available closer to the surface. As written in lines 238-239, the drainage density is a measure of how well the water in the basin is drained by the stream channels, thus affecting infiltration process. Therefore, this predictor provides insights on well the superficial soil layers will be fed by stream water. On another hand, the vegetation dependent on groundwater studied in this manuscript can use water from the vadose zone at a rooting depth reaching up to 15m. The depth to groundwater (piezometer level) allowed the exclusion of GDV where groundwater was deeper than 15m.***

Finally, the argumentation needs to be considerably improved. For example expressions such as 'subsurface groundwater' should be avoided and expressions such as 'surface groundwater' (line 60) or 'subsurface groundwater dependent vegetation' are meaningless.

***Answer: As suggested by both reviewers, we renamed the term "subsurface groundwater" as "resident groundwater" being the groundwater beneath the soil surface, as opposed to "emerging groundwater" being the groundwater above the soil surface. We changed the text accordingly throughout the manuscript (lines 56, 57, 74, 77, 120, 346).***

# Mapping the suitability of groundwater dependent vegetation in a semi-arid Mediterranean area

Inês Gomes Marques[1]; João Nascimento[2]; Rita M. Cardoso[1]; Filipe Miguéns[2]; Maria Teresa Condesso de Melo[2]; Pedro M. M. Soares[1]; Célia M. Gouveia[1]; Cathy Kurz Besson[1]

[1] Instituto Dom Luiz; Faculty of Sciences, University of Lisbon, Campo Grande, Ed. C8, 1749-016, Lisbon, Portugal

[2] CERIS; Instituto Superior Técnico, University of Lisbon, 1049-001, Lisbon, Portugal

*Correspondence to*: Inês Gomes Marques (icgmarques@fc.ul.pt or icgmarques@isa.ulisboa.pt)

**Abstract.**

Mapping the suitability of groundwater dependent vegetation in semi-arid Mediterranean areas is fundamental for the sustainable management of groundwater resources and groundwater dependent ecosystems (GDE) under the risks of climate change scenarios. For the present study the distribution of deep-rooted woody species in southern Portugal was modeled using climatic, hydrological and topographic environmental variables~~; and the density of~~. To do so, *Quercus suber*, *Quercus ilex* and *Pinus pinea* were used as proxy species ~~of~~ to represent the Groundwater Dependent Vegetation (GDV). Model fitting was performed between the proxy species Kernel density and the selected environmental predictors using 1) a simple linear model and 2) a Geographically Weighted Regression (GWR), to account for auto-correlation of the spatial data and residuals. When comparing the results of both models, the GWR modelling results showed improved goodness of fitting, as opposed to the simple linear model. Climatic indices were the main drivers of GDV density, ~~closely~~ followed with a much lower influence by groundwater depth, drainage density and slope. Groundwater depth did not appear to be as pertinent in the model as initially expected, accounting only for about 7% of the total variation against 88% for climate drivers

The relative proportion of model predictor coefficients was used as weighting factors for multicriteria analysis, to create a suitability map to the GDV in southern Portugal showing where the vegetation most likely relies on groundwater to cope with aridity. A validation of the resulting map was performed using independent data of the Normalized Difference Water Index (NDWI) a satellite-derived vegetation index. June, July and August of 2005 NDWI anomalies, to the years 1999-2009, were calculated to assess the response of active woody species in the region after an extreme drought. The results from the NDWI anomaly provided an overall good agreement with the suitability to host GDV. The model was considered reliable to predict the distribution of the studied vegetation.

The methodology developed to map GDV's will allow to predict the evolution of the distribution of GDV according to climate change ~~scenarios~~ and aid stakeholder decision-making concerning priority areas of water resources management.

**Keywords:** Groundwater dependent vegetation, aridity, agroforestry, suitability map, Normalized Difference Water Index

## 1 Introduction

Mediterranean forests, woodlands and shrublands, mostly growing under restricted water availability, are one of the terrestrial biomes with higher volume of groundwater used by vegetation (Evaristo and McDonnell, 2017). Future predictions of decreased precipitation (Giorgi and Lionello, 2008; Nadezhdina et al., 2015), decreased runoff (Mourato et al., 2015) and aquifer recharge (Ertürk et al., 2014; Stigter et al., 2014) in the Mediterranean region threaten the sustainability of groundwater reservoirs and the corresponding dependent ecosystems. Therefore, a sustainable management of groundwater resources and the Groundwater Dependent Ecosystems (GDE) is of crucial importance.

A widely used classification of GDE was proposed by Eamus et al. (2006). This classification distinguishes three types: 1) Aquifer and cave ecosystems, which include all subterranean waters; 2) Ecosystems reliant on emerging groundwater (e.g. estuarine systems, wetlands; riverine systems) and 3) Ecosystems reliant on resident groundwater (e.g. systems where plants remain physiologically active during extended drought periods, without a visible water source).

Mapping GDE constitutes a first and fundamental step to their active management. Several approaches have been proposed, from local field surveys measuring plant transpiration of stable isotopes (Antunes et al. 2018) up to larger spatial scales  involving ~~including~~ remote sensing techniques (e.g. Normalized Difference Vegetation Index – NDVI) (Barron et al., 2014; Eamus et al., 2015; Howard and Merrifield, 2010), remote-sensing combined with ground-based observations (Lv et al., 2013), ~~based on~~ geographic information system (GIS) (Pérez Hoyos et al., 2016a) GIS combining field surveys (Condesso de Melo et al., 2015), or even statistical approaches (Pérez Hoyos et al., 2016b). ~~An integrated multidisciplinary methodology  has also been used. A widely used classification of GDE was proposed by  that distinguishes three types: 1) Aquifer and cave ecosystems, which includes all subterranean waters; 2) Ecosystems reliant on surface groundwater (e.g. estuarine systems, wetlands; riverine systems) and 3) Ecosystems reliant on subsurface groundwater (e.g. systems where plants remain physiologically active during extended drought periods, without visible water source).~~

Despite of a wide-ranging body of literature ~~regarding~~ reviewing GDE's topics (Doody et al., 2017; Dresel et al., 2010; Münch and Conrad, 2007), most of ~~the~~ regional scale studies do not include Mediterranean regions~~ (Doody et al., 2017; Dresel et al., 2010; Münch and Conrad, 2007)~~. Moreover, studies on ecosystems relying on ~~subsurface~~ resident groundwater frequently only focused on riparian environments (Lowry and Loheide, 2010; O'Grady et al., 2006), with few examples in Mediterranean areas (del Castillo et al., 2016; Fernandes, 2013; Hernández-Santana et al., 2008; Mendes et al., 2016). There is a clear knowledge gap on the identification of phreatophyte species reliant on resident groundwater and their~~such ecosystems, their phreatophyte~~ associated vegetation (Robinson, 1958) in the Mediterranean region and the management actions that should be taken to decrease the adverse effects of climate change.

In the driest regions of the Mediterranean basin, the persistent lack of water during the entire summer periods gave an adaptive advantage to the vegetation that could either avoid or escape drought by reaching deeper stored water up to the point of entirely relying ~~in~~ on groundwater (Chaves et al., 2003; Canadell et al., 1996; Miller et al., 2010). This drought-avoiding strategy is often associated to the development of a dimorphic root system~~s~~ in woody species (Dinis 2014, David et al., 2013) or to hydraulic lift and/or hydraulic redistribution mechanisms (Orellana et al., 2012). Those mechanisms provide the ability to move water from deep soil layers, where water content is higher, to more shallow layers where water content is lower (Horton and Hart, 1998; Neumann and Cardon, 2012). Hydraulic lift and redistribution have been reported for several woody species of the Mediterranean basin (David et al., 2007; Filella and Peñuelas, 2004) and noticeably for Cork oak (*Quercus suber* L.) (David et al., 2013; Kurz-Besson et al., 2006; Mendes et al., 2016).

Mediterranean cork oak woodlands (Montados) are agro-silvo-pastoral systems considered as semi-natural ecosystems of the southwest Mediterranean basin (Joffre et al., 1999) that have already been referenced has a groundwater dependent terrestrial ecosystem (Mendes et al., 2016). Montados must be continually maintained through human management by thinning, understory use through grazing, ploughing and shrub clearing (Huntsinger and Bartolome, 1992) to maintain a good productivity, biodiversity and ecosystems service (Bugalho et al., 2009). In the ecosystems of this geographical area, the dominant tree species are the cork oak (*Quercus suber* L.) and the Portuguese holm oak (*Quercus ilex* subs *rotundifolia* Lam.*)* (Pinto-Correia et al., 2011). Additionally, stone pine (*Pinus pinea* L.) has become a commonly co-occurrent species in the last decades (Coelho and Campos, 2009). The use of groundwater has been frequently reported for both *Pinus* (Antunes et al. 2018; Filella and Peñuelas, 2004; Grossiord et al., 2016; Peñuelas and Filella, 2003) and *Quercus* genre (Barbeta and Peñuelas, 2017; David et al., 2007, 2013, Kurz-Besson et al., 2006, 2014; Otieno et al., 2006). Furthermore, the contribution of groundwater to tree physiology has been shown to be of a greater magnitude for *Quercus* sp. as compared with *Pinus* sp. (del Castillo et al., 2016; Evaristo and McDonnell, 2017).

*Q. suber* and *Q. ilex* have been associated with high resilience and adaptability to hydric and thermic stress, and to recurrent droughts in the southern Mediterranean basin (Barbero et al., 1992). In Italy and Portugal, during summer droughts *Q. ilex* used a mixture of rain-water and groundwater and was able to take water from very dry soils (David et al., 2007; Valentini et al., 1992). An increasing contribution of groundwater in the summer has also been shown for this species (Barbeta et al., 2015). Similarly, *Q. suber* showed a seasonal shift in water sources, from shallow soil water in the spring to the beginning of the dry period followed by a progressive higher use of deeper water sources throughout the drought period (Otieno et al., 2006). In addition, the species roots are known to reach depths as deep as 13m in southern Portugal (David et al., 2004). ~~Although co-occurrent to cork and holm oaks species, there is still no evidence yet that~~ *P. pinea* has been recently included in the facultative phreatophyte species (Antunes et al. 2018). Moreover, the species relies on groundwater resources during the dry season. However it shows a very similar root system (Montero et al., 2004) as compared to cork oak (David et al., 2013), with large sinker roots reaching 5 m depth (Canadell et al., 1996). Given the information available on water use strategies by the phreatophyte arboreous species of the cork oak woodlands, *Q. ilex*, *Q. suber* and *P. pinea* were considered as proxies for arboreous vegetation that belongs to GDE relying on resident groundwater (from here onwards designed as Groundwater Dependent Vegetation – GDV).

GDV of the Mediterranean basin is often neglected in research. Indeed, still little is known about the GDV distribution, but research has already been done on the effects of climate change in specific species distribution, such as *Q. suber*, in the Mediterranean basin (Duque-Lazo et al., 2018; Paulo et al., 2015). While the increase in atmospheric $CO_2$ and the r~~a~~ising temperature can boost tree growth (Barbeta and Peñuelas, 2017; Bussotti et al., 2013; Sardans and Peñuelas, 2004), water stress can have a counteracting effect on growth of both *Quercus ilex* (López et al., 1997; Sabaté et al., 2002) and *P. pinaster* (Kurz-Besson et al., 2016). Therefore, it is of crucial importance to identify geographical areas where subsurface GDV is present and characterize the environmental conditions this vegetation type is thriving in. This would contribute to the understanding of how to manage these species under unfavorable future climatic conditions.

The aim of this study was to address the mentioned gaps by creating ~~create~~ a suitability map of the ~~current distribution of the~~ arboreous phreatophyte species ~~considered here as GDV~~ in southern Portugal, traducing their potential dependency on groundwater. We used an integrated multidisciplinary methodology combining a geospatial modeling approach based on the Geographically Weighted Regression (GWR) and a GIS multicriteria analysis approach, both relying on forest inventory, edaphoclimatic conditions and topographic information. We expected this new integrated procedure to grant a more reliable estimation of the vegetation dependency on groundwater sources at the regional scale.

The Mapping methodology was based on the occurrence of known subsurface phreatophyte species and well-known environmental conditions affecting water resources availability. Several environmental predictors were selected according to their expected impact on water use, flux or ~~and~~ storage and then used in ~~a Geographically Weighted Regression (~~GWR~~)~~ to model the density of *Q. suber*, *Q. ilex* and *P. pinea* occurrence in the Alentejo region (NUTSII) of southern Portugal. ~~So far~~To our knowledge, very few applications of ~~this method~~GWR have been used to model species distribution and only recently its use has spread in ecological research (Hu et al., 2017; Li et al., 2016; Mazziotta et al., 2016). The coefficients ~~proportions~~ obtained from the model equation for each predictor and expressed as proportion of total sum of absolute coefficients were used as weights to build the suitability map with GIS multi-factor analysis, after reclassifying each relevant environmental ~~predictor~~driver. The resulting map was validated using the remote sensed vegetation index NDWI.

Based on former knowledge gathered from field surveys conducted in the region (Antunes et al. 2018, Condesso de Melo et al., 2015, Kurz-Besson et al. 2006 & 2014, Otieno et al. 2006, David et al. 2013, Pinto et al. 2013), on ~~the~~ environmental conditions ~~of the study area~~ and the species ecophysiological needs, we hypothesized that 1) groundwater depth together with climatic conditions play one of the most important environmental roles in GDV's distribution and 2) groundwater depth between 1.5 and 15 m associated with xeric conditions should favor a higher density of GDV and thus a larger use of groundwater by the vegetation.

## 2 Material and Methods

### 2.1 Study area

The administrative region of Alentejo (NUTSII) (fig01) covers an area of 31 604.9 km$^2$, between 37.22º and 39.39º N in latitude and between 9.00º and 6.55º W in longitude. This study area is characterized by a Mediterranean temperate mesothermic climate with hot and dry summers, defined as Csa in the Köppen classification (APA, n.d.; ARH Alentejo, 2012a, 2012b). It is characterized by a sub-humid climate, which has recently quickly drifted to semi-arid conditions (Ministério da Agricultura do Mar do Ambiente e do Ordenamento do Território, 2013). A large proportion of the area (above 40%) is covered by forestry systems (Autoridade Florestal Nacional and Ministério da Agricultura do Desenvolvimento Rural e das Pescas, 2010) providing a high economical value to the region and the country (Sarmento and Dores, 2013).

### 2.2 Kernel Density estimation of GDV

Presence datasets of *Quercus suber*, *Quercus ilex* and *Pinus pinea* of the last Portuguese forest inventory achieved in 2010 (ICNF, 2013) were used to calculate Kernel density (commonly called heat map) as a proxy for GDV suitability. The inventory registered the occurrence of each species on a 500m mesh grid resolution (corresponding to a maximum occurrence of 4 counts per km$^2$. Only data points with one of the three proxy species selected as primary and secondary occupation were used. The resulting Kernel density was weighted according to tree cover percentage and was calculated using a quartic biweight distribution shape, a search radius of 10 km, and an output resolution of 0.018 degrees, corresponding to a cell size of 1km. This variable was computed using QGIS version 2.14.12 (QGIS Development Team, 2017).

### 2.3 Environmental variables

Species distribution is mostly affected by limiting factors controlling ecophysiological responses, disturbances and resources (Guisan and Thuiller, 2005). To characterize the study area in terms of GDV's suitability, environmental variables expected to affect GDV's density were selected according to their constraint on groundwater uptake and soil water storage. Within possible abiotic variables, landscape topography, geology, groundwater availability and regional climate were considered to map GDV density. The twelve selected variables for modeling purposes, retrieved from different data sources, are listed in Table 1. The software used in spatial analysis was ArcGIS® software version 10.4.1 by Esri and R program software version 3.4.2 (R Development Core Team, 2016).

#### 2.3.1 Slope and soil characteristics

The NASA and METI ASTER GDEM product was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool. Spatial Analyst Toolbox was used to calculate the slope from the digital elevation model. Slope was used as proxy for the identification of shallow soil water interaction with vegetation.

The map of soil type was obtained from the Portuguese National Information System for the Environment - SNIAmb (© Agência Portuguesa do Ambiente, I.P., 2017) and uniformized to the World Reference Base with the Harmonized World Soil Database v 1.2 (FAO et al., 2009). The vector map was converted to raster using the Conversion Toolbox. To reduce the analysis complexity involving the several soil types present in the map, soil types were regrouped in three classes, according to their capacity to store or drain water (Table A1 in appendix A). The classification was based on the characteristics of each soil unit (available water storage capacity, drainage and topsoil texture) from the Harmonized World Soil Database v 1.2 (FAO et al., 2009). In the presence of dominant soil with ~~little~~ low drainage capacity, ~~mainly topsoil~~ a high clay fraction in the top soil and a high available water content ~~(AWC)~~, lower scores were given in association to decreased suitability for GDV by favoring non-GDV species. Otherwise, when soil characteristics suggested water storage at deeper soil depths, lower ~~AWC~~water content, drainage and sandy topsoil texture, higher scores were given.

Effective soil thickness (Table 1) was also considered for representing the maximum soil depth explored by the vegetation roots. It constrains the expansion and growth of the root system, as well as the available amount of water that can be absorbed by roots.

### 2.3.2 Groundwater availability

Root access to water resources is one of the most limiting factors for GDV's growth and survival, especially during the dry season. The map of depth to water table was interpolated from piezometric observations from the Portuguese National Information System on Water Resources (SNIRH) public data base (http://snirh.apambiente.pt, last accessed on March 31$^{st}$ 2017) and the Study of Groundwater Resources of Alentejo (ERHSA) (Chambel et al., 2007). Data points of large-diameter wells and piezometers were retrieved for the Alentejo region (fig02) and sorted into undifferentiated, karst or porous geological types to model groundwater depth (W). In the studied area, piezometers are exclusively dedicated small diameter boreholes for piezometric observations, in areas with high abstraction volumes for public water supply. Large diameter wells in this region are usually low yielding and mainly devoted to private use and irrigation. Due to the large heterogeneity of geological media, groundwater depth was calculated separately for each sub-basin. A total of 3158 data points corresponding to large wells and piezometers were used, with uneven measurements between 1979 and 2017. For each piezometer an average depth was calculated from the available observations and used as a single value. In areas with undifferentiated geological type, piezometric level and elevation were highly correlated (>0.9), thus a linear regression was applied to interpolate data. Ordinary kriging was preferred for the interpolation of karst and porous aquifers, combining large wells and piezometric data points. The ordinary kriging was calculated using a semi-variogram in which the sill, range and nugget were optimized to create the best fit of the model to the data. To build a surface layer of the depth to water table, the interpolated surface of the groundwater level was subtracted from the digital elevation model. Geostatistical Analyst ToolBox was used for this task.

Drainage density is a measure of how well the basin is drained by stream channels. It is defined as the total length of channels per unit area. Drainage density was calculated for a 10km grid size for the Alentejo region, by the division of the 10km square area (A) in $km^2$ by the total stream length (L) in km, as in Eq. (1).

$$D = \frac{L}{A}, \quad (1)$$

### 2.3.3 Regional Climate

Temperature and precipitation datasets were obtained from the E-OBS (http://eca.knmi.nl/download/ensembles/ensembles.php, last accessed on March 31st 2017) public database (Haylock et al., 2008). Standardized Precipitation Evapotranspiration Index (SPEI), Aridity Index ($A_i$) and Ombrothermic Indexes were computed from long-term (1951-2010) monthly temperature and precipitation observations. The computation of potential evapotranspiration (PET) was performed according to Thornthwaite (1948) and was calculated using the SPEI package (Beguería and Vicente-Serrano, 2013) in R program.

SPEI multi-scalar drought index (Vicente-Serrano et al., 2010) was calculated over a 6 month interval to characterize drought severity in the area of study using SPEI package (Beguería and Vicente-Serrano, 2013) for R program. SPEI is based on the normalization of the water balance calculated as the difference between cumulative precipitation and PET for a given period at monthly intervals. Normalized values of SPEI typically range between -3 and 3. Drought events were considered as severe when SPEI values were between -1.5 and -1.99, and as extreme with values below -2 (Mckee et al., 1993). Severe and extreme SPEI predictors were computed as the number of months with severe or extreme drought, counted along the 60 years of the climate time-series.

While the SPEI index used in this study identifies geographical areas affected with more frequent extreme droughts, the Aridity index distinguishes arid geographical areas prone to annual negative water balance (with low $A_i$ value) to more mesic areas showing positive annual water balance (with high $A_i$ value). $A_i$ gives information related to evapotranspiration processes and rainfall deficit for potential vegetative growth. It was calculated following Eq. (2) according to Middleton et al. (1992), where PET is the average annual potential evapotranspiration and P is the average annual precipitation, both in mm for the 60 years period of the climate time-series. Dry lands are defined by their degree of aridity in 4 classes: Hyperarid ($A_i$<0.05); Arid (0.05<$A_i$<0.2); Semi-arid (0.2<$A_i$<0.5) and Dry Subhumid (0.5<$A_i$<0.65) (Middleton et al., 1992).

$$A_i = \frac{P}{PET},\qquad (2)$$

Ombrothermic Indexes (O) were used to better characterize the bioclimatology of the study region (Rivas-Martínez et al., 2011), by evaluating soil water availability for plants during the driest months of the year. Four ombrothermic indexes were calculated according to a specific section of the year stated in Table 1, and following Eq. (3), where Pp is the positive annual precipitation (accumulated monthly precipitation when the average monthly mean temperature is higher than 0°C) and Tp is the positive annual temperature (total in tenths of degrees centigrade of the average monthly temperatures higher than 0°). Ombrothermic index presenting values below 2 for the analyzed months, can be considered as Mediterranean bioclimatically. For non-Mediterranean areas, there is no dry period in which, for at least two consecutive months, the precipitation is less than or equal to twice the temperature.

$$O = \frac{Pp}{Tp},\qquad (3)$$

### 2.4 Selection of model predictors

The full set of environmental variables was evaluated as potential predictors for the suitability of GDV (based on the Kernel density of the proxy species). A preliminary selection was carried out, first by computing Pearson's correlation coefficients between environmental variables and second by performing a Principal Components Analysis (PCA) to detect multicollinearity. Covariates were discarded for modeling according to a sequential procedure. Whenever pairs of variables presented a correlation value above 0.4, the variable with the highest explained variance on the first axis of the PCA was selected. In addition, selected variables had to show the lowest possible correlation values between them. Variables showing low correlations and explaining a higher cumulative proportion of variability with the lowest number of PCA axis were later selected as predictors for modeling. PCA was performed using the GeoDa Software (Anselin et al., 2006) and Pearson's correlation coefficients were computed with Spatial Analyst Tool .

### 2.5 Model development

When fitting a linear regression model based on the selected variables, the normal distribution and stationarity of the ~~model predictors~~response variable and residuals must be assured.

The Kernel density of the proxy GDV species, *Q. suber*, *Q. ilex* and *P. pinea*, showed a skewed normal distribution. Therefore, a square-root ~~normalization~~ transformation of the data was applied on the response variable, before model fitting. To be able to compare the resulting model coefficients and use them as weighting factors of the multi-criteria analysis to build the suitability map, the predictor variables were normalized using the z-score function. This allows to create standardized scores for each variable, by subtracting the mean of all data points from each individual data point, then dividing those points by the standard deviation of all points, so that the mean of each z-predictor is zero and the deviation is 1.

Spatial autocorrelation and non-stationarity are common when using linear regression on spatial data. To overcome these issues, the Geographically Weighted Regression (GWR) was used ~~to allow model coefficients to adjust to each location of the dataset, based on the proximity of sampling locations (Stewart Fotheringham et al., 1996)~~. This extension of the Ordinary Least Squares (OLS) linear regression considers the spatial non stationarity in variable relationships and allows the use of spatially varying coefficients while minimizing spatial autocorrelation (Stewart Fotheringham et al., 1996). In this study, simple linear regression and GWR were both applied to the dataset and their performances compared. Models were fitted on a 5% random subsample of the entire dataset (reaching a total of 6214 selected data points), due to computational restrictions and to decrease the spatial autocorrelation effect (Kühn, 2007). This methodology has already been applied with a subsample of 10%, with points distant 10km from each other (Bertrand et al., 2016). In spite of the subsampling, the mean and maximum distance between two random data points were, respectively, 3.6 km and 16.7 km, providing a good representation of local heterogeneity, as shown in figures 05 and 06. An additional analysis showing an excellent agreement between the two datasets is presented in FigA1 in appendix A.

Initially the model was constructed containing all selected predictors through the PCA and Pearson's correlation analysis. Afterwards, predictors were sequentially discarded to ascertain the model presenting lower second-order Akaike Information Criteria (AICc) and higher quasi-global $R^2$ chosen to predict the suitability of GDV.

Adaptive Kernel bandwidths for the GWR model fitting were used due to the spatial irregularity of the random subsample. ~~Bandwidths~~ Local search radius were obtained by minimizing the CrossValidation score (Bivand et al., 2008), thus minimizing the error of the local regressions.. To analyze the performance of the GWR model alone, the local and global adjusted R-squared were considered. To compare between the GWR model and the simple linear model, the distribution of the model residuals was used to identify clustered values as well as the AICc. The spatial autocorrelation of the models residuals was evaluated with the Moran's I test (Moran, 1950) using the Spatial Statistics Tool, and also graphically. GWR model was fitted using the *spgwr* package from R program (Bivand and Yu, 2017).

### 2.6 Suitability map building

To create the suitability map all predictor layers included in the GWR model were classified, similarly to Condesso de Melo et al. (2015) and Aksoy et al. (2017) . The likelihood of an interaction between the vegetation and groundwater resources was scored from 1 to 3 for each predictor. Scores were assigned after bibliographic review and expert opinion. The higher the score, the higher the likelihood, 1 corresponding to a weak likelihood and 3 indicating very high likelihood. Groundwater depth was divided in two classes, according to the accessibility to shallow soil water above 1.5 m and the maximum rooting depth for Mediterranean woody species reaching 13 m, reported by Canadell et al. (1996). Throughout the manuscript water between 0 and 1.5 m depth was designated as shallow soil water, while water below 1.5 m depth was considered as groundwater. 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 presents a mean rooting depth of 1.5m (Silva and Rego, 2004). 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) and not reaching levels as deep as deep-rooted species. The minimum score was given to areas where groundwater depth was too shallow (below 1.5 m) considered to belong to ~~surface~~ emerging groundwater dependent vegetation. Areas with steep slope were considered to have superficial runoff and less recharge and influence negatively tree density (Costa et al., 2008). Those areas were treated as less suitable to GDV. Values of the Ombrothermic Index of the summer quarter and the immediately previous month ($O_4$) were split in 3 classes according to Jenks natural breaks, with higher suitability corresponding to higher aridity. The higher values of $A_i$, corresponding to lower aridity had a score of 1, because a higher humid environment would decrease the necessity of the arboreous species to use deep water sources. Accordingly, an increase in aridity (lower values of $A_i$) has already been shown to increase tree decline (Waroux and Lambin, 2012) and so higher $A_i$ values corresponded to a score of 2, leaving the score 3 to intermediate values of $A_i$. Drainage density scoring was based on the drainage capability ~~of drainage~~ of the water through the hydrographical network of the river. ~~When~~ A low drainage density ~~was lower~~ (below 0.5~~),~~) implies a high loss of water through runoff along ~~a higher suitability scoring was given because the water lost from runoff through~~ the hydrographic network. This water lost for shallow soil horizons would be less available to the vegetation thus favoring a higher use of water from deep groundwater reservoirs (Rodrigues, 2011).

A direct compilation of the predictor layers could have been performed for the multicriteria analysis. However, some predictors might have a stronger influence on GDV's distribution and density than others. Therefore, there was a need to define weighting factors for each layer of the final GIS multicriteria analysis. Yet, due to the intricate relations between all environmental predictors and their effects on the GDV, experts and stakeholders suggested very different scoring for a same layer. Instead the relative proportion of each predictor was used locally, according to the GWR model (Eq. 4) as weighting factors. The final GIS multicriteria analysis was performed using the Spatial Analyst Tool by applying local model equations obtained for each of the 6214 coordinates of the Alentejo map (Eq.4),

$$S_{GDV}= Intercept + coef_{p1} * [real\ value\ X_1] + coef_{p2} * [real\ value\ X_2] + coef_{p3} * [real\ value\ X_3] + \ldots, \tag{4}$$

with $S_{GDV}$ representing the suitability to Groundwater Dependent Vegetation, brackets representing the reclassified GIS X layer corresponding to the scoring and $coef_x$ indicating the relative proportion for the predictor *x* was calculated as the ratio between the modulus of the local coefficient *x* and the sum of the modulus of all local coefficients..

According to this equation, lower values indicate a lower occurrence of groundwater use ~~referred~~ representing a lower GDV suitability while higher values correspond to a higher use of groundwater ~~referred~~ representing a higher GDV suitability. To allow for an easier interpretation, the data on suitability to GDV was subsequently classified based on their distribution value, according to Jenks natural breaks. This resulted in 5 suitability classes: “Very poor”, “Poor”, “Moderate”, “Good” and “Very Good”.

## 2.7 Map evaluation

Satellite derived remote-sensing products have been widely used to follow the impact of drought on land cover and the vegetation dynamics (Aghakouchaket al. 2015). Vegetation indexes offer excellent tools to assess and monitor plant changes and water stress (Asrar et al. 1989). The Normalized Difference Water Index (NDWI) (Gao, 1996) is a satellite-derived index that aims to estimate fuel moisture content (Maki et al., 2004) and leaf water content at canopy level, widely used for drought monitoring (Anderson et al., 2010, Gu et al., 2007; Ceccato et al., 2002a). This index was chosen to be more sensitive to canopy water content and a good proxy for water stress status in plants. Moreover, NDWI has been shown to be best related to the greenness of Cork oak woodland's canopy, expressed by the fraction of intercepted photosynthetically active radiation (Cerasoli et al., 2016).

In order to validate the GDV suitability map obtained in our study, we calculated anomalies of the Normalized Difference Water Index (NDWI) (Gao, 1996) between an extreme dry year (2005) and the median value of the surrounding 10 year period (1999-2009).

NDWI is computed using the near infrared (NIR) and the short-wave infrared (SWIR) reflectance, which makes it sensitive to changes in liquid water content and in vegetation canopies (Gao, 1996; Ceccato et al., 2002a, b). The index computation (Eq. 5) was further adapted by Gond et al. (2004) to SPOT-VEGETATION instrument datasets, using NIR (0.84 µm) and MIR (1.64 µm) channels, as described by Hagolle et al. (2005).

$$NDWI = \frac{\rho_{NIR} - \rho_{MIR}}{\rho_{NIR} + \rho_{MIR}}. \qquad (5)$$

Following Eq. (5), NDWI data was computed using B3 and MIR data acquired from VEGETATION instrument on board of SPOT4 and SPOT5 satellites. Extraction and corrections procedures applied to optimize NDWI series are fully described in Gouveia et al. (2012).

The NDWI anomaly was computed as the difference between NDWI observed in June, July and August of 2005 and the median NDWI for the considered month for the period 1999 to 2009. June was selected to provide the best signal from a still fully active canopy of woody species while the herbaceous layer had usually already finished its annual cycle and dried out. The hydrological year of 2004/2005 was characterized by an extreme drought event over the Iberian Peninsula, where less than 40% of the normal precipitation was registered in the southern area (Gouveia et al., 2009). Thus, in June 2005 the vegetation of the Alentejo region was already coping with an extreme long-term drought, which was well captured by the anomaly of the NDWI index (negative values), as formely shown by Gouveia et al. 2012.

## 2.8 Sensitivity analysis

Sensitivity analyses are conducted to identify model inputs that cause significant impact and/or uncertainty in the output. They can be used to identify key variables that should be the focus of attention to increase mode robustness in future research or to remove redundant inputs from the model equation because they do not have significant impact on the model output. Based on bootstrapping simulations (Tian et al., 2014), a sensitivity analysis was conducted on the GWR model by perturbing one input predictor at time while keeping the rest of the equation unperturbed. To simulate perturbations, 10000 values were randomly selected within the natural range of each input variable observed in the Alentejo region. Those random values were then used to run 10000 simulations of the local equation of the GWR model for each of the 6214 coordinates of the geographical area. Local outputs corresponding to the predicted GDV density were then calculated for each perturbed input variable ($A_i$, $O_4$, W, D and s). The range of output values was calculated to reflect the sensibility of the model for the perturbed input variable. The overall sensibility of the model to all input variables was estimated as the absolute difference between the minimum output value and the sum of maximum output values of all predictors, thus representing the maximum possible output range observed after perturbing all predictors.

## 3 Results

### 3.1 Kernel Density

Within the studied region of Portugal, the phreatophyte species *Quercus suber*, *Quercus ilex* and ~~the suspected phreatophyte species~~ *Pinus pinea* were not distributed uniformly throughout the territory. Areas with higher Kernel density (or higher distribution likelihood) were mostly spread between the northern part of Alentejo region and the western part close to the coast, with values ranging between 900 and 1200 (fig03). Two clusters of high density also appeared below the Tagus river. The remaining study area presented mean density values, with ~~a~~ very low ~~density~~ densities in the area of the river Tagus and in the center south.

### 3.2 Environmental conditions

The exploratory analysis of the variables performed through the PCA and Pearson correlation matrix confirmed the presence of multicollinearity. From the initial variables (Table 1), Thickness (T), number of months with severe and extreme SPEI (respectively, $SPEI_s$ and $SPEI_e$), Annual Ombrothermic Index (O), Ombrothermic Index of the hottest month of the summer quarter($O_1$) and Ombrothermic Index of the summer quarter **(**$O_3$) were discarded, while the variables slope (s), drainage density (D), soil type ($S_t$), groundwater depth (W), $A_i$ and $O_4$ were maintained for analysis (figA2 and Table A2 in appendix). A sequential removal of one predictor from the initial modeling including six variables was performed (Table 2), after which the model was reduced to 5 variables, with the highest global $R^2$ (0.99) and the lowest AICc (18050.34). Therefore, out of the initial 12 considered (fig04) were endorsed to explain the variation of the Kernel density of GDV in Alentejo the following variables: $A_i$, $O_4$, W, D and s.

In most part of the Alentejo region, slope was below 10% (fig04e) and coastal areas presented the lowest values and variability. Highest values of groundwater depth (fig04c), reaching a maximum of 255 m, were found in the Atlantic margin of the study area, mainly in Tagus and Sado river basins. Several other small and confined areas in Alentejo also showed high values, corresponding to aquifers of porous or karst geological types. Most of the remaining study area showed groundwater depths ranging between 1.5 m and 15 m. Figures 04a and 04b indicate the southeast of Alentejo as the driest area, given by minimum values of the aridity index (0.618), and much higher potential evapotranspiration that precipitation. Besides, $O_4$ presented a maximum value (~~0.714~~1.166) for this region (meaning that soil water availability was not compensated by the precipitation of the previous M-J-J-A months). This is also supported by the higher drainage density in the southeast which indicates a lower prevalence of shallow soil water due to higher stream length by area.

Combining all variables, it was possible to distinguish two sub-regions with distinct conditions: the southeast of Alentejo and the Atlantic margin. The latter is mainly distinguished by its low slope areas, shallower groundwater and more humid climatic conditions than the southeast of Alentejo.

### 3.3 Regression models

The best model to describe the GDV distribution was found through a sequential discard of each variable (Table 2) and corresponded to the model with a distinct lower AICc (18050.76) compared with the second lowest AICc (27389.74) and showed an important increase in quasi-global $R^2$ (from 0.926 for the second best model to 0.992 for the best one). The best model fit was obtained with $A_i$, $O_4$, W, D and s. This final model was then applied to the GIS layers to map the suitability of GDV in Alentejo, according to Eq. 6.

$$S_{GDV} = Intercept + A_i\, coef_p * [reclassified\ A_i\ value] + O_4\, coef_p * [reclassified\ O_4\ value] + W\, coef_p * [reclassified\ W\ value] + D\, coef_p * [reclassified\ D\ value] + s\, coef_p * [reclassified\ s\ value], \quad (6)$$

Local adjusted R-squared of the GWR model was highly variable throughout the study area, ranging from 0 to 0.99 (fig05). Also, the local $R^2$ values below 0.5 corresponded to only 0.3% of the data. The lower $R^2$ values were distributed throughout the Alentejo area, with no distinct pattern. The overall fit of the GWR model was high (Table 3). The adjusted regression coefficient indicated that 99% of the variation in the data was explained by the GWR model, while only 2% was explained by the simple linear model (Table 3). Accordingly, GWR had a substantially lower AICc when compared with the simple linear model, indicating a much better fit.

The spatial autocorrelation given by the Moran Index (Griffith, 2009; Moran 1950) retrieved from the geospatial distribution of residual values was significant for both the GWR and the linear models, indicating that observations geospatially are dependent on each other to a certain level . However, this dependence was substantially lower for the GWR model than for the linear model (z-score of 50.24 and 147.56 respectively). In the GWR model (fig06a) the positive and negative residual values were much more randomly scattered throughout the study region than in the linear model (fig06b), highlighting a much better performance of the GWR, which minimized residual autocorrelation. Indeed, in the linear model (fig06b), positive residuals were condensed in the right side of Tagus and Sado river basins, while negative values were mainly present on the left side of the Tagus river and in the center-south of Alentejo.

The spatial distribution of the coefficients of GWR predictors is presented in Fig07. They were later used for the computation of the GDV suitability score for each data point (Eq.6). The coefficient variability was three times higher for the $A_i$ as compared to $O_4$ (fig08a), reaching 66% and 22% respectively. For W, D and s, the coefficient variation was much lower, representing only about 6.2%, 3.8% and 1.2% of the total variation observed in the coefficients, respectively. The remaining variables showed a median close to 0 and the $O_4$ was the second with higher variability followed by the W. The coefficient median values were, respectively, -3.40, 0.29, -0.015, -0.018 and 0.022 for $A_i$, $O_4$, W, D and s variables.

The distributions of negative coefficients were similar for $A_i$ and the $O_4$ variables (fig07a and fig07b), with lower values in the southern coastal area, and in the Tagus river watershed. The highest absolute values were mostly found for $A_i$ in the southern area of the Alentejo region and on smaller patches in the northern region. In the center and eastern areas of Alentejo, a higher weight of the groundwater depth coefficient could be found (fig07c), approximately matching a higher influence of slope (fig07e). The groundwater depth seemed to have almost no influence on GDV density in the Tagus river watershed, expressed by coefficients mostly null around the riverbed (fig07c). The coefficient distribution of D and $O_4$ shows some similarities, mostly in the center and southeast of Alentejo (fig07d). Extreme values of $O_4$ coefficients were mostly concentrated in the eastern part of the Tagus watershed and in the southern coastal area included in the Sado watershed. Slope coefficient values showed the lowest amplitude throughout the study area (fig07e), with prevailing high positive values gathered mainly in the center of the study area and in the Tagus river watershed (northwest of the study center).

### 3.4 GDV Suitability map

The classification of the 5 endorsed environmental predictors is presented in Table 4 and their respective maps in figure B1 in appendix B. Rivers Tagus and Sado had an overall large impact on GDV's suitability for each predictor, with the exception of W. This is due to a higher water availability reflected by the values of $O_4$, D and lower slopes due to the alluvial plains of the Tagus river (figs. B1b,,d and e in appendix B). Moreover, those regions presented higher humidity conditions (through analysis of the $A_i$ in fig B1a in appendix B) and groundwater depths outside the optimum range (Fig. B1c in appendix B), therefore less suitable for GDV. Optimal conditions for groundwater access were mainly gathered in the interior of the study region (fig. B1c in appendix B), with the exception of some confined aquifers in the northeast and southeast of the study region. Favorable slopes for GDV were mostly highlighted in the Tagus river basin area, where a good likelihood of interaction between GDV and groundwater could be identified (fig. B1e in appendix B).

The final map illustrating the suitability to GDV is shown in Fig. 09. The largest classified area (8 787km$^2$) presented a very poor suitability to GDV, corresponding ~~but corresponded only~~ to approximately a quarter of the total study area (29%). This percentage was followed closely by the moderate suitability to GDV which occupied 26% (8000km$^2$). Overall, the two less suitable classes (very poor and poor) represented 47% of the study area, whilst the two best ones and the moderate class (very good, good and moderate) represented 53%. Consequently, most of the study area showed moderate to high suitability to GDV. The very good and good suitability classes cover an arch from the most south and northeastern area of the Alentejo region, passing through the Sado and southern and northern Guadiana river basins and close to the coastal line at 38ºN. Most of the center of the study area showed moderate to very good suitability ~~do~~ to GDV, while the areas corresponding to the alluvial deposits of the Tagus river showed poor to very poor suitability.

The suitability to GDV in the Alentejo region was mainly driven by $A_i$, given that the highest coefficient variability was associated to the $A_i$ predictor in the GWR model equation. This is also supported by the similar distribution pattern observed between the suitability map and the aridity index predictor (fig04a and fig09). Areas with good or very good suitability mostly matched areas of $A_i$ with score 3, corresponding to aridity index values above 0.75 (Fig. B1a in appendix B). On the other hand, the lowest suitability classes showed a good agreement with the lowest scores given to W (fig. B1c in appendix B), mostly in the coastal area and in the Tagus river basin.

### 3.5 Map evaluation

To evaluate the suitability map developed in the present study, the results were compared with the NDWI anomaly considering the month of June of the dry year of 2005 in the Alentejo area (fig10). Both maps (figs 09 and 10) showed similar patterns, with higher presence of GDV satisfactorily matching areas with the lowest NDWI anomaly. From June to September in a extremely dry year, non-DGV plants can be expected to experience a severe drought stress as in any regular summer period. Thus, those plants should show almost zero anomaly. By opposition, GDV plants coping well with usual summer drought can be expected to suffer an unusual stress under an extreme dry year even having access to groundwater (Kurz-Besson et al. 2006 & 2014, Otieno et al. 2006, David et al. 2013), with a negative impact of groundwater drawdown (Antunes et al., 2018). Therefore, GDV plants should show negative NDWI anomalies.

The NDWI anomaly was mostly negative over the Alentejo territory indicating a lower leaf water content in June and July 2005 than usual ~~water stress in the vegetation leaves~~. The loss of water attributed ~~Water stress due~~ to the extreme drought ~~was maximum (green color) in geographical areas matching~~was mostly matching geographical areas with the~~the~~ highest GDV suitability (fig09). ~~It was~~Water loss was less pronounced ~~(mostly yellowish)~~ in the central area of the Alentejo region between the Guadiana and Sado river basins, where the vegetation ~~presents a lower~~is less dense ~~density~~ (fig03). Areas with ~~positive/~~null ~~values of~~ NDWI anomaly values (indicating no NDWI change~~corresponding to geographical areas with a higher water availability~~) were mostly distributed on the coastal area of the Atlantic ocean or close to riverbeds, namely in the Tagus and Sado floodplains ~~(brown color, fig10)~~, matching areas of very poor suitability for GDV in Figure 09.

Despite an overall good agreement, the adequation between the density, suitability and NDWI maps was not perfect. Indeed, some patches showing a high vegetation occurrence/density and large NDWI anomalies also matched an area of very poor suitability for GDV.

~~Note that green and yellow areas in June 2005 (fig 10a) progressively turned to brown color in July and August 2005 (fig10c), suggesting that the corresponding vegetation recovered from the increasing water stress, despite the intensification of drought throughout the summer period.~~

### 3.6 Sensitivity analysis

The sensitivity of the model in response to the perturbation of each one of the input variables ($A_i$, $O_4$, W, D and s) is presented on Figure 11a to Figure11e. The overall sensitivity of the model is further presented on Figure 11f. For any input variable, the model sensitivity (fig11a to 11e) was higher where absolute values of local coefficients were also higher (fig07a to 07e). The maximum impact on GDV's density, corresponding to the maximum output range observed after perturbation (fig08b), was observed when perturbing the Aridity index, accounting for 66% of the total variability. The second highest impact was observed after perturbing the ombrothermic index. The variability in the model outputs observed after perturbing the remaining variables $O_4$, W, D and s accounted for 22%, 7%, 4% and 1% of the total accumulated variability, respectively (fig08b). The highest variability in the GWR model output was mostly observed in the central part of the southern half of the Alentejo region, as well as close to the main channels of the Guadiana and Tagus rivers (fig11f). Furthermore, areas with higher model sensitivity (fig11f) significantly matched higher model performance expressed by $R^2$ (fig05), assessed with a Kruskall-Wallis test ($p<0.0001$***).

## 4 Discussion

### 4.1 Modeling approach

The Geographically Weighted Regression model has been used before in ecological studies (Li et al., 2016; Mazziotta et al., 2016), but never for the mapping of GDV, to our knowledge. This approach considerably improved the goodness of fit when compared to the linear model, with a coefficient of regression ($R^2$) increasing from 0.02 to 0.99 at the global level, and an obvious reduction of residual clustering. Despite those improvements, it has not been possible to completely eliminate the residual autocorrelation after fitting the GWR model.

Kernel density for the study area provided a strong indication of presence and abundance of the tree species considered as GDV proxy for modeling. The Mediterranean cork woodlands dominate about 76% of the Alentejo region (while only 7% is covered by stone pine). In those systems, tree density is known to be a tradeoff between climate drivers (Joffre 1999, Gouveia & Freitas 2008) and the need for space for pasture or cereal cultivation in the understory (Acacio & Holmgreen 2014). In our study, the anthropologic management of agroforestry systems in the Alentejo region has not been taken into account. According to a recent study of Cabon et al. (2018) where thinning played an important role in *Q. ilex* density in a Mediterranean climate site, anthropologic management could, at least partially, explain the non-randomness of the residual distribution after GWR model fitting as well as the mismatches between the GDV and the NDWI evaluation maps.

Another explanation of the reminiscent autocorrelation after GWR fitting could be the lack of groundwater dependent species in the model. For example, *Pinus pinaster* Aiton was excluded due to its more humid distribution in Portugal, and due to conflicting conclusions driven from previous studies to pinpoint the species as a potential groundwater user (Bourke, 2004; Kurz-Besson et al., 2016). In addition, olive trees were also excluded although the use of groundwater by an olive orchard has been recently proved (Ferreira et al., 2018), however with a weak contribution of groundwater to the daily root flow, and thus with no significant impact of groundwater on the species physiological conditions.

Methods previously used by Doody et al., (2017) and Condesso de Melo et al. (2015) to map specific vegetation relied solely on expert opinion, e.g. Delphi panel, to define weighting factors of environmental information for GIS multicriteria analysis. In our study, the GWR modelling approach was used to assess weighting factors for each environmental predictor in the study area, to build a suitability map for the GDV in southern Portugal. This allowed an empirical determination of the local relevance of each environmental predictor in GDV distribution, thus avoiding the inevitable subjectivity of Delphi panels.

Also, by combining the GWR and GIS approaches we believe the final suitability map provides a more reliable indication of the higher likelihood for groundwater dependency and a safer appraisal of the relative contribution of groundwater by facultative deep-rooted phreatophytes species in the Alentejo region.

Modelling of the entire study area at a regional level did not provide satisfactory results. Therefore, we developed a general model varying locally according to local predictor coefficients. The local influence of each predictor was highly variable throughout the study area, especially for climatic predictors reflecting water availability and stress conditions. The application of the GWR model did not only allow~~ed~~ for a localized approach, by decreasing the residual error and autocorrelation over the entire studied region, but also provided insights on how GDV's density can be explained by the main environmental drivers locally.

The GWR model appeared to be highly sensitive to coefficient fitting corresponding to a good model fit, as expected in a spatially varying model. As so, high coefficients are highly reliable in the GWR model in our study. Yet, the high spatial variability of local coefficients might reflect a weak physical meaning of the GWR model that challenges its direct application in other regions, even under similar climate conditions. Predictor coefficients showed a similar behavior in the spatial distribution of the coefficients. This was noticeable for the aridity index and the groundwater depth in the Tagus and Sado river basins. Groundwater depth had no influence on GDV's density in these areas and similarly, the coefficient of aridity index showed a negative effect of increased humidity on GDV's density. In addition, a cluster of low drainage density values matched these areas. Due to the lower variability and impact of the drainage density and slope on the GDV's density, these variables might not impact significantly this vegetation density in future climatic scenarios.

### 4.2 Suitability to Groundwater Dependent Vegetation

According to our results, more than half of the study area appeared suitable for GDV. However, one quarter of the studied area showed lower suitability to GDV. The lower suitability to this vegetation in the more northern and western part of the studied area ~~can be explained by less favorable climatic and hydrological conditions, resulting from the combination of a high aridity index and low groundwater depth scores (equivalent to high shallow soil water availability)~~, ~~corresponding~~ including~~to~~ the coastal area and ~~in~~ the Tagus river basin. Those are the moist humid areas of the study area, where GDV is unlikely to rely on groundwater during the drought season because rainfall water stored in shallow soil horizons is mostly available.

The proxy species (Cork oak, Holm oak and Stone pine) can perfectly grow under sub-humid Mediterranean climate conditions, without relying as much on groundwater to survive as in more xeric semi-arid areas (Abad Vinas et al., 2016). As facultative phreatophyte species, their presence/abundance is only an indication of a possible use of groundwater. The study provided by Pinto et al. (2013) have shown that Cork oak for example can perfectly thrive were very shallow groundwater is available while suffering drought stress were groundwater source is lower but still extracted by trees. Also, former studies have shown that in the extreme dry year of 2005, Cork oak experienced a severe drought stress, close to the cavitation threshold, although its main water source was groundwater (David et al. 2013, Kurz-Besson et al. 2006, 2014). These findings can explain that part of the maximum density (Fig. 04) matches the area of very poor suitability for GDV (Fig. 09). Elsewhere, the better agreement between the two maps reflects the dominance of the aridity index on the vegetation's occurrence.

Groundwater depth appeared to have a lower influence on GDV density than climate drivers, as reflected by the relative low magnitude of the W coefficient and outputs of our model outcomes. This surprisingly disagrees with our initial hypothesis because groundwater represents a notable proportion of the transpired water of deep-rooting phreatophytes, reaching up to 86% of absorbed water during drought periods and representing about 30.5% of the annual water absorbed by trees (David et al. 2013, Kurz-Besson et al. 2014). Nonetheless, this disagreement should be regarded cautiously due to the poor quality data used and the complexity required for modelling the water table depths. Besides, the linear relationship between water depth and topography applied to areas of undifferentiated geological type can be weakened by a complex non-linear interaction between topography, aridity and subsurface conductivity (Condon and Maxell, 2015). Moreover, the high variability in geological media, topography and vegetation cover at the regional scale did not allow to account for small changes in groundwater depth (<15 m deep), which has a huge impact on GDV suitability (Canadell et al., 1996; Stone and Kalisz, 1991). Indeed, a high spatial resolution of hydrological database is essential to rigorously characterize the spatial dynamics of groundwater depth between hydrographic basins (Lorenzo-Lacruz et al., 2017). Unfortunately, such resolution was not available for our study area.

The aridity and ombrothermic indexes were the most important predictors of GDV density in the Alentejo region, according to our model outcomes. Our results agree with previous findings linking tree cover density and rooting depth to climate drivers such as aridity, at a global scale (Zomer et al., 2009; Schenk and Jackson, 2002) and specifically for the Mediterranean oak woodland (Gouveia and Freitas 2008, Joffre et al. 1999). Through previous studies showing the similarities in vegetation strategies to cope with water scarcity in the Mediterranean basin (Vicente-Serrano et al., 2013) or the relationship between rooting depth and water table depth increased with aridity at a global scale (Fan et al., 2017) we can admit that the most relevant climate drivers in this study are similarly important to map GDV in other semi-arid regions. In this study, the most important environmental variables that define GDV's density in a semi-arid region were identified, helping to fill the gap of knowledge for modelling this type of vegetation. However, the coefficients to be applied when modelling each variable need to be calculated locally, due to their high spatial variability.

Temporal data would further help discriminate areas of optimal suitability to GDV, either during the wet and the dry seasons, because the seasonal trends in groundwater depth are essential under Mediterranean conditions. Investigations efforts should be invested to fill the gap either by improving the Portuguese piezometric monitoring network, or by assimilating observations with remote sensing products focused on soil moisture or groundwater monitoring. This has already been performed for large regional scale such as GRACE satellite surveys, based on changes of Earth's gravitational field. So far, these technologies are not applicable to Portugal's scale, since the coarse spatial resolution of GRACE data only allows the monitoring of large reservoirs (Xiao et al. 2015).

### 4.3 Validation of the results

The understory of woodlands and the herbaceous layer of grasslands areas in southern Portugal usually ends their annual life cycles in June (Paço et al. 2007), while the canopy of woody species is still fully active with maximum transpiration rates and photosynthetic activities (Kurz-Besson et al. 2014, David et al. 2007, Awada et al. 2003). This is an ideal period of the year to spot differential response of the canopy of woody species to extreme droughts events using satellite derived vegetation indexes (Gouveia 2012).

The spatial patterns of NDWI anomaly in June 2005 seem to indicate that the woody canopy showed a strong loss of canopy water in the areas were tree density and GDV suitability were higher (figs03, 09 and 10). This occurred although trees minimized the loss of water in leaves with a strong stomatal limitation in response to drought (Kurz-Besson et al. 2014, Grant et al. 2010). In the most arid area of the region were Holm oak is dominant but tree density is much lower, the NDWI anomaly was generally less negative thus showing a lower water stress or higher canopy water content. Holm oak (*Quercus ilex* spp *rotundifolia*) is well known to be the most resilient species to dry and hot conditions in Portugal, due to its capacity to use groundwater, ~~and~~ associated to a higher water use efficiency (David et al. 2007). Furthermore, the dynamics of NDWI anomaly ~~spatial patterns~~ over the summer period (fig10a, b and c) pointed out that the lower water stress status on the map is progressively spreading from the most arid areas to the milder ones from June to August 2005, despite the intensification of drought conditions. This endorses the idea that trees manage to cope with drought by relying on deeper water sources in response to drought, replenishing leaf water content despite the progression and intensification of drought conditions. Former studies support this statement by showing that groundwater uptake and hydraulic lift were progressively taking place after the onset of drought by promoting the formation of new roots reaching deeper soil layers and water sources, typically ~~in~~ from July onwards, for cork oak in the Alentejo region (Kurz-Besson et al., 2006, 2014). Root elongation following a declining water table has also been reported in a review on the effect of groundwater fluctuations on phreatophyte vegetation (Naumburg et al. 2005).

Our results and the dynamics of NDWI over summer 2005 tend to corroborate the studies of Schenk and Jackson (2002) and Fan et al. (2017), by suggesting a larger/longer dependency of GDV on groundwater with higher aridity. Further investigation needs to be carried on across aridity gradients in Portugal and the Iberian Peninsula to fully validate this statement, though.

Overall, the map of suitability to GDV showed an ~~excellent~~ good agreement with the NDWI validation maps. The main areas showing good GDV suitability and highest NDWI anomalies are mostly matching in both maps. The good agreement between our GDV suitability maps, and NDWI dynamic maps opens the possibility to apply and extend the methodology to larger geographical areas such as the Iberian Peninsula~~, or~~ and to the simulation of the impact of climate changes on the distribution of groundwater dependent species in the Mediterranean basin.

Simulations of future climate conditions based on RCP4.5 and RCP8.5 emission scenarios (Soares et al., 2015, 2017) predict a significant decrease of precipitation for the Guadiana basin and overall decrease for the southern region of Portugal within 2100. Agroforestry systems relying on groundwater resources, such as cork oak woodlands, may show a decrease in productivity and ecosystem services or even face sustainability failure. Many studies carried out on oak woodlands in Italy and Spain identified drought as the main driving factor of tree die-back and as the main climate warning threatening oak stands sustainability in the Mediterranean basin (Gentilesca et al. 2017). An increase in aridity and drought frequency for the Mediterranean (Spinoni et al., 2017) will most probably induce a geographical shift of GDV vegetation toward milder/wetter climates (Lloret et al., 2004; Gonçalez P., 2001).

### 4.4 Key limitations

The GWR modelling approach used to estimate weighting factors is mostly stochastic. Consequently, the large spatial variability and symmetrical fluctuations around zero (Fig 08b) denote a weak physical meaning of the estimated coefficients, at least at the resolution chosen for the study. Also, the local nature of the regression coefficients makes the model difficult to directly apply in other regions, even with similar climate conditions, unless the methodology is properly fitted to local conditions/predictors.

With the methodology applied in this study, weighting factors can be easily evaluated solely from local and regional observations of the studied area. Nonetheless, the computation of model coefficients or expert opinion to assess weighting factors, require recurrent amendments, associated with updated environmental data, species distribution and revised expert knowledge (Doody et al., 2017).

The evolution of groundwater depth in response to climate change is difficult to model on a large scale based on piezometric observations because it requires an excellent knowledge of the components and dynamics of water catchments. Therefore, a reliable estimation of the impact of climate change on GDV suitability in southern Portugal could only been performed on small scale studies. However, the GWR model appeared to be much more sensitive to climate drivers than the other predictors, given that 88% of the model outputs variability was covered by climate indexes $A_i$ and $O_4$. Nevertheless, changes in climate conditions only represent part of the water resources shortage issue in the future. Global-scale changes in human populations and economic progresses also rules water demand and supply, especially in arid and semi-arid regions (Vörösmarty et al., 2000). A decrease in useful water resources for human supply can induce an even higher pressure on groundwater resources (Döll, 2009), aggravating the water table drawdown caused by climate change (Ertürk et al., 2014). Therefore, additional updates of the model should include human consumption of groundwater resources, identifying areas of higher population density or intensive farming. Future model updates should also account for the interaction of deep rooting species with the surrounding understory species. In particular, shrubs surviving the drought period, which can benefit from the redistribution of groundwater by deep rooted species (Dawson, 1993; Zou et al., 2005).

## 5 Conclusions

Our results show a highly dominant contribution of water scarcity of the last 30 years (Aridity and Ombrothermic indexes) on the density and suitability of deep-rooted groundwater dependent species in southern Portugal. Therefore, in geographical regions of the world with similar semi-arid climate conditions (Csa according to Köppen-Geigen classification, Peel et al. 2007) and similar physiological responses of the groundwater dependent vegetation (Vicente-Serrano et al., 2013), the use of the aridity and ombrothermic indexes could be used as first approximation to model and map deep rooted phreatophyte species and the evolution of their distribution in response to climate changes. The contribution of groundwater depth was lower than initially expected, however, this might be underestimated due to the poor quality of the piezometric network, especially in the central area of the studied region.

The current pressure applied by human consumption of water sources has reinforced the concern on the future of economic activities dependent on groundwater resources. To address this issue, several countries have developed national strategies for the adaptation of water sources for Agriculture and Forests against Climate Change, including Portugal (FAO, 2007). In addition, local drought management as long-term adaptation strategy has been one of the proposals by Iglesias et al. (2007) to reduce the climate change impact on groundwater resources in the Mediterranean. The preservation of Mediterranean agroforestry systems, such as cork oak woodlands and the recently associated *P. pinea* species, is of great importance due to their high socioeconomic value and their supply of valuable ecosystem services (Bugalho et al., 2011). Management policies on the long-term should account for groundwater resources monitoring, accompanied by defensive measures to ensure agroforestry systems sustainability and economical income from these Mediterranean ecosystems are not greatly and irreversibly threatened.

Our present study, and novel methodology, provides an important tool to help delineating priority areas of action for species and groundwater management, at regional level, to avoid the decline of productivity and cover density of the agroforestry systems of southern Portugal. This is important to guarantee the sustainability of the economical income for stakeholders linked to the agroforestry sector in that area. Furthermore, mapping vulnerable areas at a small scale (e.g.by hydrological basin), where reliable groundwater depth information is available, should provide further insights for stakeholder to promote local actions to mitigate climate change impact on GDV.

Based on the methodology applied in this work, future predictions on GDV suitability, according to the RCP4.5 and RCP8.5 emission scenarios will be shortly introduced, providing guidelines for future management of these ecosystems in the allocation of water resources.

## 6 Acknowledgements

The authors acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). The authors also wish to acknowledge the ASTER GDEM data product, a courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool. We are grateful to ICNF for sharing inventory database performed in 2010 in Portugal continental. We also thank Cristina Catita, Ana Russo and Patrícia Páscoa for the advice and helpful comments as well as Ana Bastos for the elaboration of the satellite datasets of the vegetation index NDWI and Miguel Nogueira for the insights on model sensitivity analysis. We are very grateful to Eric Font for the useful insights on soil properties. I Gomes Marques and research activities were supported by the Portuguese National Foundation for Science and Tecnhology (FCT) through the PIEZAGRO project (PTDC/AAG-REC/7046/2014). This publication was also supported by FCT- project UID/GEO/50019/2019 – Instituto Dom Luiz. The authors further thank the reviewers and editor for helpful comments and suggestions on an earlier version of the manuscript.

The authors declare that they have no conflict of interest.

## References


Acácio V. and Holmgreen M.: Pathways for resilience in Mediterranean cork oak land use systems, Annals of Forest Science, 71, 5-13, doi: 10.1007/s13595-012- 0197-0, 2009

Aghakouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D. and Hain C. R.: Remote sensing of drought: Progress, challenges and opportunities. Rev. Geophys., doi: 10.1002/2014RG000456, 2015.

Aksoy, E., Louwagie, G., Gardi, C., Gregor, M., Schröder, C. and Löhnertz, M.: Assessing soil biodiversity potentials in Europe, Sci. Total Environ., 589, 236–249, doi:10.1016/j.scitotenv.2017.02.173, 2017.

Anderson, L. O., Malhi, Y., Aragão, L. E. O. C., Ladle, R., Arai, E., Barbier, N. and Phillips, O.: Remote sensing detection of droughts in Amazonian forest canopies. New Phytologist, 187, 733–750, doi: 10.1111/j.1469-8137.2010.03355.x, 2010.

Anselin, L., Ibnu, S. and Youngihn, K.: GeoDa: An Introduction to Spatial Data Analysis, Geogr. Anal., 38(1), 5–22, 2006.

Antunes, C., Chozas, S., West, J., Zunzunegui, M., Barradas, M. C. D., Vieira,S., & Máguas, C. Groundwater drawdown drives ecophysiolog-ical adjustments of woody vegetation in a semi-arid coastal ecosystem. Global Change Biology, https://doi.org/10.1111/gcb.14403, 2018.

APA: Plano de Gestão da Região Hidrográfica do Tejo: Parte 2 - Caracterização e Diagnóstico da Região Hidrográfica, n.d.

ARH Alentejo: Plano de Gestão das Bacias Hidrográficas integradas na RH7 - Parte 2, 2012.

ARH Alentejo: Planos de Gestão das Bacias Hidrográficas integradas na RH6 - Parte 2, 2012.

Asrar, G. (Ed.): Estimation of plant-canopy attributes from spectral reflectance measurements, Theory and Applications of Optical Remote Sensing, 252–296, John Wiley, New York, 1989.

Autoridade Florestal Nacional and Ministério da Agricultura do Desenvolvimento Rural e das Pescas: 5o Inventário Florestal Nacional, 2010.

Awada T., Radoglou K., Fotelli M. N., Constantinidou H. I. A.: Ecophysiology of seedlings of three Mediterranean pine species in contrasting light regimes, Tree Physiol., 23, 33–41, doi: 2003.

Barata, L. T., Saavedra, A., Cortez, N. and Varennes, A.: Cartografia da espessura efectiva dos solos de Portugal Continental. LEAF/ISA/ULisboa. [online] Available from: http://epic-webgis-portugal.isa.utl.pt/, 2015.

Barbero, M., Loisel, R. and Quézel, P.: Biogeography, ecology and history of Mediterranean *Quercus ilex* ecosystems, in *Quercus ilex* L. ecosystems: function, dynamics and management, edited by F. Romane and J. Terradas, 19–34, Springer Netherlands, Dordrecht., 1992.

Barbeta, A. and Peñuelas, J.: Increasing carbon discrimination rates and depth of water uptake favor the growth of Mediterranean evergreen trees in the ecotone with temperate deciduous forests, Glob. Chang. Biol., 1–15, doi:10.1111/gcb.13770, 2017.

Barbeta, A., Mejía-Chang, M., Ogaya, R., Voltas, J., Dawson, T. E. and Peñuelas, J.: The combined effects of a long-term experimental drought and an extreme drought on the use of plant-water sources in a Mediterranean forest, Glob. Chang. Biol., 21(3), 1213–1225, doi:10.1111/gcb.12785, 2015.

Barron, O. V., Emelyanova, I., Van Niel, T. G., Pollock, D. and Hodgson, 0G.: Mapping groundwater-dependent ecosystems using remote sensing measures of vegetation and moisture dynamics, Hydrol. Process., 28(2), 372–385, doi:10.1002/hyp.9609, 2014.

Beguería, S. and Vicente-Serrano, S. M.: SPEI: Calculation of the Standardized Precipitation-Evapotranspiration Index. R package version 1.6., 2013.

Bertrand R., Riofrío-Dillon G., Lenoir J., Drapier J., de Ruffray P., Gégout J. C. and Loreau M.: Ecological constraints increase the climatic debt in forests, Nature Communications, 7, 12643, doi: 10.1038/ncomms12643, 2016.

Bivand, R. and Yu, D.: spgwr: Geographically Weighted Regression. [online] Available from: https://cran.r-project.org/package=spgwr, 2017.

Bivand, R. S., Pebesma, E. J. and Gómez-Rubio, V.: Applied Spatial Data Analysis with R, edited by G. P. Robert Gentleman, Kurt Hornik, Springer., 2008.

Bourke, L.: Growth trends and water use efficiency of *Pinus pinaster* Ait. in response to historical climate and groundwater trends on the Gnangara Mound, Western Australia. [online] Available from: http://ro.ecu.edu.au/theses_hons/141 (Accessed 29 January 2018), 2004.

Bugalho, M. N., Plieninger, T. and Aronson, J.: Open woodlands: a diversity of uses (and overuses), in Cork oak woodlands on the edge, edited by J. Aronson, J. S. Pereira, and J. G. Pausas, pp. 33–48, Island Press, Washington DC., 2009.

Bugalho, M. N., Caldeira, M. C., Pereira, J. S., Aronson, J. and Pausas, J. G.: Mediterranean cork oak savannas require human use to sustain biodiversity and ecosystem services, Front. Ecol. Environ., 9(5), 278–286, doi:10.1890/100084, 2011.

Bussotti, F., Ferrini, F., Pollastrini, M. and Fini, A.: The challenge of Mediterranean sclerophyllous vegetation under climate change: From acclimation to adaptation, Environ. Exp. Bot., 103(April), 80–98, doi:10.1016/j.envexpbot.2013.09.013, 2013.

Cabon, A., Mouillot, F., Lempereur, M., Ourcival, J.-M., Simioni, G. and Limousin, J.-M.: Thinning increases tree growth by delaying drought-induced growth cessation in a Mediterranean evergreen oak coppice, doi:10.1016/j.foreco.2017.11.030, 2017.

Canadell, J., Jackson, R., Ehleringer, J., Mooney, H. A., Sala, O. E. and Schulze, E.-D.: Maximum rooting depth of vegetation types at the global scale, Oecologia, 108, 583–595, doi:10.1007/BF00329030, 1996.

del Castillo, J., Comas, C., Voltas, J. and Ferrio, J. P.: Dynamics of competition over water in a mixed oak-pine Mediterranean forest: Spatio-temporal and physiological components, For. Ecol. Manage., 382, 214–224, doi:10.1016/j.foreco.2016.10.025, 2016.

Ceccato, P., Gobron, N., Flasse, S., Pinty, B. and Tarantola, S.: Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1. Theoretical approach. Remote Sens. Environ., 82, 188 – 197, 2002a

Ceccato, P., Flasse, S. and Gregoire, J.: Designing a spectral index to estimate vegetation water content from remote sensing data: Part 2. Validation and applications. Remote Sens. Environ., 82, 198 – 207, 2002b

Cerasoli S., Silva F.C. and Silva J. M. N.: Temporal dynamics of spectral bioindicators evidence biological and ecological differences among functional types in a cork oak open woodland, Int. J. Biometeorol., 60 (6), 813–825, doi: 10.1007/s00484-015-1075-x, 2016.

Chambel, A., Duque, J. and Nascimento, J.: Regional Study of Hard Rock Aquifers in Alentejo, South Portugal: Methodology and Results, in Groundwater in Fractured Rocks - IAH Selected Paper Series, pp. 73–93, CRC Press., 2007.

Chaves M. M., Maroco J. P. and Pereira J.S.: Understanding plant responses to drought — from genes to the whole plant, Funct Plant Biol, 30(3), 239 - 264, 2003.

Coelho, I. S. and Campos, P.: Mixed Cork Oak-Stone Pine Woodlands in the Alentejo Region of Portugal, in Cork Oak Woodlands on the Edge - Ecology, Adaptive Management, and Restoration, edited by J. Aronson, J. S. Pereira, J. Uli, and G. Pausas, pp. 153–159, Island Press, Washington, 2009.

Condesso de Melo, M. T., Nascimento, J., Silva, A. C., Mendes, M. P., Buxo, A. and Ribeiro, L.: Desenvolvimento de uma metodologia e preparação do respetivo guia metodológico para a identificação e caracterização, a nível nacional, dos ecossistemas terrestres dependentes das águas subterrâneas (ETDAS). Relatório de projeto realizado para a Agência P., 2015.

Condon, L. E. and Maxell, R. M.: Water resources research, Water Resour. Res., 51, 6602–6621, doi:10.1002/2014WR016259, 2015.

Costa, A., Madeira, M. and Oliveira, C.: The relationship between cork oak growth patterns and soil, slope and drainage in a cork oak woodland in Southern Portugal, For. Ecol. Manage., 255, 1525–1535, doi:10.1016/j.foreco.2007.11.008, 2008.

David, T. S., Ferreira, M. I., Cohen, S., Pereira, J. S. and David, J. S.: Constraints on transpiration from an evergreen oak tree in southern Portugal, Agric. For. Meteorol., 122(3–4), 193–205, doi:10.1016/j.agrformet.2003.09.014, 2004.

David, T. S., Henriques, M. O., Kurz-Besson, C., Nunes, J., Valente, F., Vaz, M., Pereira, J. S., Siegwolf, R., Chaves, M. M., Gazarini, L. C. and David, J. S.: Water-use strategies in two co-occurring Mediterranean evergreen oaks: surviving the summer drought., Tree Physiol., 27(6), 793–803, doi:10.1093/treephys/27.6.793, 2007.

David, T. S., Pinto, C. A., Nadezhdina, N., Kurz-Besson, C., Henriques, M. O., Quilhó, T., Cermak, J., Chaves, M. M., Pereira, J. S. and David, J. S.: Root functioning, tree water use and hydraulic redistribution in *Quercus suber* trees: A modeling approach based on root sap flow, For. Ecol. Manage., 307, 136–146, doi:10.1016/j.foreco.2013.07.012, 2013.

Dawson, T. E.: Hydraulic lift and water use by plants: implications for water balance, performance and plant-plant interactions, Oecol., 95, 565–574, 1993.

Dinis, C.O.: Cork oak (*Quercus suber* L.) root system: a structural-functional 3D approach. PhD Thesis, Universidade de Évora (Portugal), 2014

Döll, P.: Vulnerability to the impact of climate change on renewable groundwater resources: a global-scale assessment, Environ. Res. Lett., 4(4), 35006–12, doi:10.1088/1748-9326/4/3/035006, 2009.

Doody, T. M., Barron, O. V., Dowsley, K., Emelyanova, I., Fawcett, J., Overton, I. C., Pritchard, J. L., Van Dijk, A. I. J. M. and Warren, G.: Continental mapping of groundwater dependent ecosystems: A methodological framework to integrate diverse data and expert opinion, J. Hydrol. Reg. Stud., 10, 61–81, doi:10.1016/j.ejrh.2017.01.003, 2017.

Dresel, P. E., Clark, R., Cheng, X., Reid, M., Terry, A., Fawcett, J. and Cochrane, D.: Mapping Terrestrial Groundwater Dependent Ecosystems: Method Development and Example Output., 2010.

Duque-Lazo, J., Navarro-Cerrillo, R. M. and Ruíz-Gómez, F. J.: Assessment of the future stability of cork oak (*Quercus suber* L.) afforestation under climate change scenarios in Southwest Spain, For. Ecol. Manage., 409(June 2017), 444–456, doi:10.1016/j.foreco.2017.11.042, 2018.

Eamus, D., Froend, R., Loomes, R., Hose, G. and Murray, B.: A functional methodology for determining the groundwater regime needed to maintain the health of groundwater-dependent vegetation, Aust. J. Bot., 54(2), 97–114, doi:10.1071/BT05031, 2006.

Eamus, D., Zolfaghar, S., Villalobos-Vega, R., Cleverly, J. and Huete, A.: Groundwater-dependent ecosystems: Recent insights from satellite and field-based studies, Hydrol. Earth Syst. Sci., 19(10), 4229–4256, doi:10.5194/hess-19-4229-2015, 2015.

Ertürk, A., Ekdal, A., Gürel, M., Karakaya, N., Guzel, C. and Gönenç, E.: Evaluating the impact of climate change on groundwater resources in a small Mediterranean watershed, Sci. Total Environ., 499, 437–447, doi:10.1016/j.scitotenv.2014.07.001, 2014.

Evaristo, J. and McDonnell, J. J.: Prevalence and magnitude of groundwater use by vegetation: a global stable isotope meta-analysis, Sci. Rep., 7, 44110, doi:10.1038/srep44110, 2017.

Fan Y., Macho G. M., Jobbágy E. G., Jackson R. B. and Otero-Casal C.: Hydrologic regulation of plant rooting depth. Proc. Natl Acad. Sci. USA 114, 10 572–10 577, doi: 10.1073/pnas.1712381114, 2017.

FAO: Adaptation to climate change in agriculture, forestry and fisheries: Perspective, framework and priorities, Rome, 2007.

FAO, IIASA, ISRIC, ISS-CAS and JRC: Harmonized World Soil Database (version 1.1), 2009.

Fernandes, N. P.: Ecossistemas Dependentes de Água Subterrânea no Algarve - Contributo para a sua Identificação e Caracterização, University of Algarve., 2013.

Ferreira, M. I., Green, S., Conceição, N. and Fernández, J.-E.: Assessing hydraulic redistribution with the compensated average gradient heat-pulse method on rain-fed olive trees, Plant Soil, 1–21, doi:10.1007/s11104-018-3585-x, 2018.

Filella, I. and Peñuelas, J.: Indications of hydraulic lift by *Pinus halepensis* and its effects on the water relations of neighbour shrubs, Biol. Plant., 47(2), 209–214, doi:10.1023/B:BIOP.0000022253.08474.fd, 2004.

Gao, B.C.: NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ., 58, 257-266, 1996.

Gentilesca T., Camarero J. J., Colangelo M., Nolè A. and Ripullone F.: Drought-induced oak decline in the western Mediterranean region: an overview on current evidences, mechanisms and management options to improve forest resilience, iForest, 10, 796-806, doi: 10.3832/ifor2317-010, 2017.

Giorgi, F. and Lionello, P.: Climate change projections for the Mediterranean region, Glob. Planet. Change, 63(2–3), 90–104, doi:10.1016/j.gloplacha.2007.09.005, 2008.

Gond, V., Bartholome, E., Ouattara, F., Nonguierma, A. and Bado, L. Surveillance et cartographie des plans d'eau et des zones humides et inondables en regions arides avec l'instrument VEGETATION embarqué sur SPOT-4, Int. J. Remote Sens, 25, 987–1004, 2004.

Gonzalez, P.: Desertification and a shift of forest species in the West African Sahel, Clim. Res., 17, 217–228, 2001.

Gouveia A. and Freitas H.: Intraspecific competition and water use efficiency in *Quercus suber*: evidence of an optimum tree density?, Trees, 22, 521-530, 2008.

Gouveia C., Trigo R. M., DaCamara C. C.: Drought and Vegetation Stress Monitoring in Portugal using Satellite Data, Nat. Hazard. Earth Sys., 9, 1-11, doi: 10.5194/nhess-9-185-2009, 2009.

Gouveia C. M., Bastos A., Trigo R. M., DaCamara C. C.: Drought impacts on vegetation in the pre- and post-fire events over Iberian Peninsula, Nat. Hazard. Earth Sys., 12, 3123-3137, doi:10.5194/nhess-12-3123-2012, 2012.

Grant O. M., Tronina L., Ramalho J. C., Besson C. K., Lobo-do-Vale R., Pereira J. S., Jones H. G. and Chaves M. M.: The impact of drought on leaf physiology of *Quercus suber* L.

trees: comparison of an extreme drought event with chronic rainfall reduction, J. Exp. Bot., 61 (15), 4361–4371, doi: 10.1093/jxb/erq239, 2010.

Griffith, D. A. (Ed.): Spatial Autocorrelation, Elsevier Inc, Texas, 2009.

Grossiord, C., Sevanto, S., Dawson, T. E., Adams, H. D., Collins, A. D., Dickman, L. T., Newman, B. D., Stockton, E. A. and Mcdowell, N. G.: Warming combined with more extreme precipitation regimes modifies the water sources used by trees, New Phytol., doi:10.1111/nph.14192, 2016.

Gu, Y., J. F. Brown, J. P. Verdin and Wardlow, B.: A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States, Geophys. Res. Lett., 34, L06407, doi:10.1029/2006GL029127, 2007.

Guisan, A. and Thuiller, W.: Predicting species distribution: Offering more than simple habitat models, Ecol. Lett., 8(9), 993–1009, doi:10.1111/j.1461-0248.2005.00792.x, 2005.

Hagolle, O., Lobo, A., Maisongrande, P., Duchemin, B. and De Pereira, A.: Quality assessment and improvement of SPOT/VEGETATION level temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images, Remote Sens. Environ., 94, 172–186, 2005.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D. and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, J. Geophys. Res. Atmos., 113(20), doi:10.1029/2008JD010201, 2008.

Hernández-Santana, V., David, T. S. and Martínez-Ferná Ndez, J.: Environmental and plant-based controls of water use in a Mediterranean oak stand, For. Ecol. Manage., 255, 3707–3715, doi:10.1016/j.foreco.2008.03.004, 2008.

Horton, J. L. and Hart, S. C.: Hydraulic lift: a potentially important ecosystem process, Tree, 13(6), 232–235, doi:0169-5347/98, 1998.

Howard, J. and Merrifield, M.: Mapping groundwater dependent ecosystems in California, PLoS One, 5(6), doi:10.1371/journal.pone.0011249, 2010.

Hu, X., Zhang, L., Ye, L., Lin, Y. and Qiu, R.: Locating spatial variation in the association between road network and forest biomass carbon accumulation, Ecol. Indic., 73, 214–223, doi:10.1016/j.ecolind.2016.09.042, 2017.

Huntsinger, L. and Bartolome, J. W.: Ecological dynamics of *Quercus* dominated woodlands in California and southern Spain: A state transition model. Vegetation 99–100, 299–305, 1992.

ICNF: IFN6 – Áreas dos usos do solo e das espécies florestais de Portugal continental. Resultados preliminares., Lisboa, 2013.

Iglesias, A., Garrote, L., Flores, F. and Moneo, M.: Challenges to manage the risk of water scarcity and climate change in the Mediterranean, Water Resour. Manag., 21, 775–788, doi:10.1007/s11269-006-9111-6, 2007.

Joffre, R., Rambal, S. and Ratte, J. P.: The dehesa system of southern Spain and Portugal as a natural ecosystem mimic, Agrofor. Syst., 45, 57–79, doi:10.1023/a:1006259402496, 1999.

Kühn, I.: Incorporating spatial autocorrelation may invert observed patterns, Div. and Dist., 13, 66-69, doi:10.1111/j.1472-4642.2006.00293.x, 2007

Kurz-Besson, C., Otieno, D., Lobo Do Vale, R., Siegwolf, R., Schmidt, M., Herd, A., Nogueira, C., David, T. S., David, J. S., Tenhunen, J., Pereira, J. S. and Chaves, M.: Hydraulic lift in cork oak trees in a savannah-type Mediterranean ecosystem and its contribution to the local water balance, Plant Soil, 282(1–2), 361–378, doi:10.1007/s11104-006-0005-4, 2006.

Kurz-Besson, C., Lobo-do-Vale, R., Rodrigues, M. L., Almeida, P., Herd, A., Grant, O. M., David, T. S., Schmidt, M., Otieno, D., Keenan, T. F., Gouveia, C., Mériaux, C., Chaves, M. M. and Pereira, J. S.: Cork oak physiological responses to manipulated water availability in a Mediterranean woodland, Agric. For. Meteorol., 184(December 2013), 230–242, doi:10.1016/j.agrformet.2013.10.004, 2014.

Kurz-Besson, C., Lousada, J. L., Gaspar, M. J., Correia, I. E., David, T. S., Soares, P. M. M., Cardoso, R. M., Russo, A., Varino, F., Mériaux, C., Trigo, R. M. and Gouveia, C. M.: Effects of recent minimum temperature and water deficit increases on *Pinus pinaster* radial growth and wood density in southern Portugal, Front. Plant Sci, 7, doi:10.3389/fpls.2016.01170, 2016.

Li, Y., Jiao, Y. and Browder, J. A.: Modeling spatially-varying ecological relationships using geographically weighted generalized linear model: A simulation study based on longline seabird bycatch, Fish. Res., 181, 14–24, doi:10.1016/j.fishres.2016.03.024, 2016.

Lloret, F., Siscart, D. and Dalmases, C.: Canopy recovery after drought dieback in holm-oak Mediterranean forests of Catalonia (NE Spain), Glob. Chang. Biol., 10(12), 2092–2099, doi:10.1111/j.1365-2486.2004.00870.x, 2004.

López, B., Sabaté, S., Ruiz, I. and Gracia, C.: Effects of Elevated CO2 and Decreased Water Availability on Holm-Oak Seedlings in Controlled Environment Chambers, in Impacts of Global Change on Tree Physiology and Forest Ecosystems: Proceedings of the International Conference on Impacts of Global Change on Tree Physiology and Forest Ecosystems, held 26--29 November 1996, Wageningen, The Netherlands, edited by G. M. J. Mohren, K. Kramer, and S. Sabaté, pp. 125–133, Springer Netherlands, Dordrecht., 1997.

Lorenzo-Lacruz, J., Garcia, C. and Morán-Tejeda, E.: Groundwater level responses to precipitation variability in Mediterranean insular aquifers, J. Hydrol., 552, 516-531, doi:10.1016/j.jhydrol.2017.07.011, 2017.

Lowry, C. S. and Loheide, S. P.: Groundwater-dependent vegetation: Quantifying the groundwater subsidy, Water Resour. Res., 46(6), doi:10.1029/2009WR008874, 2010.

Lv, J., Wang, X. S., Zhou, Y., Qian, K., Wan, L., Eamus, D. and Tao, Z.: Groundwater-dependent distribution of vegetation in Hailiutu River catchment, a semi-arid region in China, Ecohydrology, 6(1), 142–149, doi:10.1002/eco.1254, 2013.

Maki, M., Ishiahra, M., Tamura, M.: Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data. Remote Sens. Environ, 90, 441–450, 2004.

Mazziotta, A., Heilmann-Clausen, J., Bruun, H. H., Fritz, Ö., Aude, E. and Tøttrup, A. P.: Restoring hydrology and old-growth structures in a former production forest: Modelling the long-term effects on biodiversity, For. Ecol. Manage., 381, 125–133, doi:10.1016/j.foreco.2016.09.028, 2016.

Mckee, T. B., Doesken, N. J. and Kleist, J.: The relationship of drought frequency and duration to time scales, in AMS 8th Conference on Applied Climatology, pp. 179–184., 1993.

Mendes, M. P., Ribeiro, L., David, T. S. and Costa, A.: How dependent are cork oak (*Quercus suber* L.) woodlands on groundwater? A case study in southwestern Portugal, For. Ecol. Manage., 378, 122–130, doi:10.1016/j.foreco.2016.07.024, 2016.

Middleton, N., Thomas, D. S. G. and Programme., U. N. E.: World atlas of desertification, UNEP, 1992., London., 1992.

Miller, G. R., Chen, X., Rubin, Y., Ma, S. and Baldocchi, D. D.: Groundwater uptake by woody vegetation in a semiarid oak savanna, Water Resour. Res., 46(10), doi:10.1029/2009WR008902, 2010.

Ministério da Agricultura do Mar do Ambiente e do Ordenamento do Território: Estratégia de Adaptação da Agricultura e das Florestas às Alterações Climáticas, Lisbon, 2013.

Montero, G., Ruiz-Peinado, R., Candela, J. A., Canellas, I., Gutierrez, M., Pavon, J., Alonso, A., Rio, M. d., Bachiller, A. and Calama, R.: El pino pinonero (*Pinus pinea* L.) en Andalucia. Ecologia, distribucion y selvicultura, edited by G. Montero, J. A. Candela, and A. Rodriguez, Consejeria de Medio Ambiente, Junta de Andalucia, Sevilla., 2004.

Moran, P. A. P.: Notes on continuous stochastic phenomena, Biometrika, 37(1–2), 17–23 [online] Available from: http://dx.doi.org/10.1093/biomet/37.1-2.17, 1950.

Mourato, S., Moreira, M. and Corte-Real, J.: Water resources impact assessment under climate change scenarios in Mediterranean watersheds, Water Resour. Manag., 29(7), 2377–2391, doi:10.1007/s11269-015-0947-5, 2015.

Münch, Z. and Conrad, J.: Remote sensing and GIS based determination of groundwater dependent ecosystems in the Western Cape, South Africa, Hydrogeol. J., 15(1), 19–28, doi:10.1007/s10040-006-0125-1, 2007.

Nadezhdina, N., Ferreira, M. I., Conceição, N., Pacheco, C. A., Häusler, M. and David, T. S.: Water uptake and hydraulic redistribution under a seasonal climate: Long-term study in a rainfed olive orchard, Ecohydrology, 8(3), 387–397, doi:10.1002/eco.1545, 2015.

Naumburg, E., Mata-Gonzalez, R., Hunter, R., McLendon, T., Martin, D.: Phreatophytic vegetation and groundwater fluctuations: a review of current research and application of ecosystem response modelling with an emphasis on Great Basin vegetation. Environ. Manage., 35, 726-740, 2005.

Neumann, R. B. and Cardon, Z. G.: The magnitude of hydraulic redistribution by plant roots: a review and synthesis of empirical and modeling studies, New Phytol., 194(2), 337–352, doi:10.1111/j.1469-8137.2012.04088.x, 2012.

O'Grady, A. P., Eamus, D., Cook, P. G. and Lamontagne, S.: Groundwater use by riparian vegetation in the wet–dry tropics of northern Australia, Aust. J. Bot., 54, 145–154, doi:10.1071/BT04164, 2006.

Orellana, F., Verma, P., Loheide, S. P. and Daly, E.: Monitoring and modeling water-vegetation interactions in groundwater-dependent ecosystems, Rev. Geophys., 50(3), doi:10.1029/2011RG000383, 2012.

Otieno, D. O., Kurz-Besson, C., Liu, J., Schmidt, M. W. T., Do, R. V. L., David, T. S., Siegwolf, R., Pereira, J. S. and Tenhunen, J. D.: Seasonal variations in soil and plant water status in a *Quercus suber* L. stand: Roots as determinants of tree productivity and survival in the Mediterranean-type ecosystem, Plant Soil, 283(1–2), 119–135, doi:10.1007/s11104-004-7539-0, 2006.

Paço, T.A., David, T.S., Henriques, M.O.; Pereira, J.S., Valente, F., Banza, J., Pereira, F.L., Pinto, C., David, J.S.: Evapotranspiration from a Mediterranean evergreen oak savannah: The role of trees and pasture, J. Hydrol., 369 (1-2), 98–106, doi: 10.1016/j.jhydrol.2009.02.011, 2009.

Paulo, J. A., Palma, J. H. N., Gomes, A. A., Faias, S. P., Tomé, J. and Tomé, M.: Predicting site index from climate and soil variables for cork oak (*Quercus suber* L.) stands in Portugal, New For., 46, 293–307, doi:10.1007/s11056-014-9462-4, 2015.

Peel, M.C., Finlayson, B.L. and McMahon, T.A. (2007) Updated World Map of the Köppen-Geiger Climate Classification. Hydrol. Earth Syst. Sci., 11, 1633-1644. doi: 10.5194/hess-11-1633-2007.

Peñuelas, J. and Filella, I.: Deuterium labelling of roots provides evidence of deep water access and hydraulic lift by *Pinus nigra* in a Mediterranean forest of NE Spain, Environ. Exp. Bot., 49(3), 201–208, doi:10.1016/S0098-8472(02)00070-9, 2003.

Pérez Hoyos, I., Krakauer, N., Khanbilvardi, R. and Armstrong, R.: A Review of advances in the identification and characterization of groundwater dependent ecosystems using geospatial technologies, Geosciences, 6(2), 17, doi:10.3390/geosciences6020017, 2016a.

Pérez Hoyos, I., Krakauer, N. and Khanbilvardi, R.: Estimating the probability of vegetation to be groundwater dependent based on the evaluation of tree models, Environments, 3(2), 9, doi:10.3390/environments3020009, 2016b.

Pinto C., Nadezhdina N., David J. S., Kurz-Besson C., Caldeira M.C., Henriques M.O., Monteiro F., Pereira J.S., David T.S. Transpiration in Quercus suber trees under shallow water table conditions: the role of soil and groundwater. Hydrological processes, doi: 10.1002/hyp.10097, 2013.

Pinto-Correia, T., Ribeiro, N. and Sá-Sousa, P.: Introducing the montado, the cork and holm oak agroforestry system of Southern Portugal, Agrofor. Syst., 82(2), 99–104, doi:10.1007/s10457-011-9388-1, 2011.

QGIS Development Team: QGIS Geographic Information System. Open Source Geospatial Foundation Project., 2017.

R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2016.

Rivas-Martínez, S., Rivas-Sáenz, S. and Penas-Merino, A.: Worldwide bioclimatic classification system, Glob. Geobot., 1(1), 1–638, doi:10.5616/gg110001, 2011.

Robinson, T. W.: Phreatophytes, United States Geol. Surv. Water-Supply Pap., (1423), 84, 1958.

Rodrigues, C. M., Moreira, M. and Guimarães, R. C.: Apontamentos para as aulas de hidrologia, 2011

Sabaté, S., Gracia, C. A. and Sánchez, A.: Likely effects of climate change on growth of *Quercus ilex*, *Pinus halepensis*, *Pinus pinaster*, *Pinus sylvestris* and *Fagus sylvatica* forests in the Mediterranean region, For. Ecol. Manage., 162(1), 23–37, doi:10.1016/S0378-1127(02)00048-8, 2002.

Salinas, M. J., Blanca, G. and Romero, A. T.: Riparian vegetation and water chemistry in a basin under semiarid Mediterranean climate, Andarax River, Spain. Environ. Manage., 26(5), 539–552, 2000.

Sardans, J. and Peñuelas, J.: Increasing drought decreases phosphorus availability in an evergreen Mediterranean forest, Plant Soil, 267(1–2), 367–377, doi:10.1007/s11104-005-0172-8, 2004.

Sarmento, E. de M. and Dores, V.: The performance of the forestry sector and its relevance for the portuguese economy, Rev. Port. Estud. Reg., 34(3), 35–50, 2013.

Schenk, H. J. and Jackson, R. B.: Rooting depths, lateral root spreads and belowground aboveground allometries of plants in water limited ecosystems, J. Ecol., 480–494, doi:10.1046/j.1365-2745.2002.00682.x, 2002.

Silva, J. S. and Rego, F. C.: Root to shoot relationships in Mediterranean woody plants from Central Portugal, Biologia, 59, 109–115, 2004.

Singer, M. B., Stella, J. C., Dufour, S., Piégay, H., Wilson, R. J. S. and Johnstone, L.: Contrasting water-uptake and growth responses to drought in co-occurring riparian tree species, Ecohydrology, 6(3), 402–412, doi:10.1002/eco.1283, 2012.

Soares, P. M. M., Cardoso, R. M., Ferreira, J. J. and Miranda, P. M. A.: Climate change and the Portuguese precipitation: ENSEMBLES regional climate models results, Clim. Dyn., 45(7–8), 1771–1787, doi:10.1007/s00382-014-2432-x, 2015.

Soares, P. M. M., Cardoso, R. M., Lima, D. C. A. and Miranda, P. M. A.: Future precipitation in Portugal: high-resolution projections using WRF model and EURO-CORDEX multi-model ensembles, Clim Dyn, 49, 2503–2530, doi:10.1007/s00382-016-3455-2, 2017.

Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P. and Dosio, A.: Will drought events become more frequent and severe in Europe?, Int. J. Climatol., 38(4), 1718–1736, doi:10.1002/joc.5291, 2017.

Stewart Fotheringham, A., Charlton, M. and Brunsdon, C.: The geography of parameter space: an investigation of spatial non-stationarity, Int. J. Geogr. Inf. Syst., 10(5), 605–627, doi:10.1080/02693799608902100, 1996.

Stigter, T. Y., Nunes, J. P., Pisani, B., Fakir, Y., Hugman, R., Li, Y., Tomé, S., Ribeiro, L., Samper, J., Oliveira, R., Monteiro, J. P., Silva, A., Tavares, P. C. F., Shapouri, M., Cancela da Fonseca, L. and El Himer, H.: Comparative assessment of climate change and its impacts on three coastal aquifers in the Mediterranean, Reg. Environ. Chang., 14(S1), 41–56, doi:10.1007/s10113-012-0377-3, 2014.

Stone, E. L. and Kalisz, P. J.: On the maximum extent of tree roots, For. Ecol. Manage., 46(1–2), 59–102, doi:10.1016/0378-1127(91)90245-Q, 1991.

Tian, W., Song, J., Li Z., de Wilde, P.: Bootstrap techniques for sensitivity analysis and model selection in building thermal performance, Appl. Energ., 135, 320-328, doi: 10.1016/j.apenergy.2014.08.110, 2014.

Thornthwaite, C. W.: An approach toward a rational classification of climate, Geogr. Rev., 38(1), 55–94, 1948.

Valentini, R., Scarascia, G. E. and Ehleringer, J. R.: Hydrogen and carbon isotope ratios of selected species of a Mediterranean macchia ecosystem, Funct. Ecol., 6(6), 627–631, 1992.

Vicente-Serrano, S. M., Beguería, S. and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, J. Clim., 23(7), 1696–1718, doi:10.1175/2009JCLI2909.1, 2010.

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Begueria, S., Trigo, R., Lopez-Moreno, J. I., Azorin-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Moran-Tejeda, E. and Sanchez-Lorenzo, A.: Response of vegetation to drought time-scales across global land biomes, Proc. Natl. Acad. Sci., 110(1), 52–57, doi:10.1073/pnas.1207068110, 2013.

Vörösmarty, C. J., Green, P., Salisbury, J. and Lammers, R. B.: Global water resources: Vulnerability from climate change and population growth, Science, 289, 284–288, doi:10.1126/science.289.5477.284, 2000.

Waroux, Y. P. and Lambin, E.F.: Monitoring degradation in arid and semi-arid forests and woodlands: The case of the argan woodlands (Morocco), Appl Geogr, 32, 777-786, doi: 10.1016/j.apgeog.2011.08.005, 2012.

Xiao R., He X., Zhang Y., Ferreira V. G. and Chang L.: Monitoring groundwater variations from satellite gravimetry and hydrological models: A comparison with in-situ measurements in the mid-atlantic region of the United States, Remote Sensing, 7 (1), 686–703, doi: 10.3390/rs70100686, 2015.

Zomer, R., Trabucco, A., Coe, R., Place, F.: Trees on farm: analysis of global extent and geographical patterns of agroforestry, ICRAF Working Paper-World Agroforestry Centre, 89, doi:10.5716/WP16263, 2009.

Zou, C. B., Barnes, P. W., Archer, S. and Mcmurtry, C. R.: Soil moisture redistribution as a mechanism of facilitation in savanna tree–shrub clusters, Ecophysiology, (145), 32–40, doi:10.1007/s00442-005-0110-8, 2005.

## Figure and Table Legends

Table 1: Environmental variables for characterization of the suitability of GDV in the study area.

Table 2: Effect of variable removal in the performance of GWR model linking the Kernel density of *Quercus suber*, *Quercus ilex* and *Pinus pinea* ($S_{GDV}$) to predictors Aridity Index ($A_i$); Ombrothermic Index of the summer quarter and the immediately previous month ($O_4$); Slope (s); Drainage density (D); Groundwater Depth (W) and Soil type ($S_t$). The model with all predictors is highlighted in grey and the final model used in this study is in bold.

Table 3: Comparison of Adjusted R-squared and second-order Akaike Information Criterion (AICc) between the simple regression and the GWR models.

Table 4: Classification scores for each predictor. A score of 3 to highly suitable areas and 1 to highly less suitable for GDV.

Table A1: Classification scores for soil type predictor.

Table A2: Correlations between predictor variables and principal component axis. The most important predictors for each axis (when squared correlation is above 0.3) are showed in bold. The cumulative proportion of variance explained by each principal component axis is shown at the bottom of the table.

Figure 01: Study area. On the left the location of Alentejo in the Iberian Peninsula; on the right, the elevation characterization of the study area with the main river courses from Tagus, Sado and Guadiana basins. Names of the main rivers are indicated near to their location in the map.

Figure 02: Large well and piezometer data points used for groundwater depth calculation. Squares represent piezometers data points and triangle represent large well data points.

Figure 03: Map of Kernel Density weighted by cover percentage of *Q. suber*, *Q. ilex* and *P. pinea*. The scale unit represent the number of occurrences per 10 km search radius (~314 km$^2$).

Figure 04: Map of environmental layers used in model fitting. ((a) – Aridity Index; (b) – Ombrothermic Index of the summer quarter and the immediately previous month; (c) – Groundwater Depth; (d) –Drainage density; (e) –Slope~~a) – Soil type; (b) – Slope; (c) – Groundwater Depth; (d) – Ombrothermic Index of the summer quarter and the immediately previous month and (e) – Aridity Index~~.

Figure 05: Spatial distribution of local $R^2$ from the fitting of the Geographically Weighted Regression.

Figure 06: Spatial distribution of model residuals from the fitting of the Simple Linear model (a) and Geographically Weighted Regression (b).

Figure 07: Map of local model coefficients for each variable. (a) – Aridity Index; (b) - Ombrothermic Index of the summer quarter and the immediately previous month; (c) – Groundwater Depth; (d) – Drainage density and (e) - Slope.

Figure 08: Boxplot of GWR model coefficient values for each predictor (a) and boxplot of the GWR model outputs, corresponding to GDV's density after each of the predictors was disturbed for the sensitivity analysis (b). $A_i$ stands for Aridity Index; $O_4$ for the ombrothemic index of the hottest month of the summer quarter and the immediately previous month; W for the groundwater depth; D for the drainage density and s for the slope. Error bars represent the 25$^{th}$ and 75$^{th}$ percentile while crosses indicate the 95$^{th}$ percentile.

Figure 09: Suitability map for Groundwater Dependent Vegetation.

Figure 10: Spatial patterns of NDWI anomaly values considering the months of June, July and August of the extremely dry year of 2005, in reference to the same months of the period 1999-2009, in the Alentejo region. Dark brown colors (corresponding to extreme negative NDWI anomaly values) indicate the vegetation that experienced the highest loss of

water in leaves in summer 2005 as compared to the reference period 1999-2009, while light brown colors show NDWI anomaly values very close to the usual vegetation moisture condition of the considered month. ~~NDWI anomaly considering the months of June, July and August of the extremely dry year of 2005, in reference to the same months of the period 1999-2009, in the Alentejo region. Green colors (corresponding to low NDWI values) indicates vegetation canopy undergoing a higher water stress than the average reference period 1999-2009.~~

Figure 11: Sensitivity analysis performed on the GWR model by perturbing one of the predictors, while remaining the rest of the model equation constant. Graphics present the output range of GDV's density when the aridity index (a), the ombrothermic index (b), the groundwater depth (c), the drainage density (d) or the slope variable (e) was perturbed; and the maximum possible range combining all predictors (f). The 95th percentile was used for the maximum value of the color bar for a better statistical representation of the spatial variability.

Figure A1: Boxplot of the main predictors used for the Geographically Weighted Regression model fitting (top) and the response variable (below), for the total data (left) and for the 5% subsample (right).

Figure A2: Correlation plot between all environmental variables expected to affect the presence of the Groundwater Dependent Vegetation. $O_1$, $O_3$ and $O_4$ are ombrothermic indices of, respectively, the hottest month of the summer quarter, the summer quarter and the summer quarter and the immediately previous month; O is the annual ombrothermic index, $SPEI_e$ and $SPEI_s$ are, respectively, the number of months with extreme and severe Standardized Precipitation Evapotranspiration Index; $A_i$ is Aridity index; W is groundwater depth; D is the Drainage density; T is thickness and $S_t$ refers to soil type.

Figure B1 – Predictors maps after score classification. (a) – Aridity Index; (b) – Ombrothermic Index of the summer quarter and the immediately previous month; (c) – Groundwater Depth; (d) – Drainage density and (e) – Slope.

**Table 1: Environmental variables for the characterization of the suitability of GDV in the study area.**

| Variable code | Variable type | Source | Resolution and Spatial extent |
|---|---|---|---|
| **s** | **Slope (%)** | This work | 0.000256 degrees (25m) raster resolution |
| **$S_t$** | **Soil type in the first soil layer** | SNIAmb (© Agência Portuguesa do Ambiente, I.P., 2017) | Converted from vectorial to 0.000256 degrees (25m) resolution raster |
| **T** | **Soil thickness (cm)** | EPIC WebGIS Portugal (Barata et al., 2015) | Converted from vectorial to 0.000256 degrees (25m) resolution raster |
| **W** | **Groundwater Depth (m)** | This work | 0.000256 degrees (25m) raster resolution |
| **D** | **Drainage Density** | This work | 0.000256 degrees (25m) raster resolution |
| **$SPEI_s$** | **Number of months with severe SPEI** | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **$SPEI_e$** | **Number of months with extreme SPEI** | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **$A_i$** | **Aridity Index** | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **O** | **Annual Ombrothermic Index** Annual average (January to December) | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **$O_1$** | **Ombrothermic Index of the hottest month of the summer quarter** (J, J and A) | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **$O_3$** | **Ombrothermic Index of the summer quarter** (J, J and A) | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |
| **$O_4$** | **Ombrothermic Index of the summer quarter and the immediately previous month** (M, J, J and A) | This work | 0.000256 degrees (25m) raster resolution Time coverage 1950-2010 |

**Table 2: Effect of variable removal in the performance of GWR model linking the Kernel density of *Quercus suber*, *Quercus ilex* and *Pinus pinea* ($S_{GDV}$) to predictors Aridity Index ($A_i$); Ombrothermic Index of the summer quarter and the immediately previous month ($O_4$); Slope (s); Drainage density (D); Groundwater Depth (W); and Soil type ($S_t$). The model with all predictors is highlighted in grey and the final model used in this study is in bold.**

| Type | Model | Discarded predictor | AICc | Quasi-global $R^2$ |
|---|---|---|---|---|
| GWR | $S_{GDV} \sim O_4 + A_i + s + D + W + S_t$ | | 27389.74 | 0.926481 |
| GWR | $S_{GDV} \sim O_4 + s + D + W + S_t$ | $A_i$ | 28695.14 | 0.9085754 |
| GWR | $S_{GDV} \sim A_i + s + D + W + S_t$ | $O_4$ | 28626.88 | 0.9095033 |
| GWR | $S_{GDV} \sim O_4 + A_i + s + W + S_t$ | D | 27909.86 | 0.9184337 |
| GWR | $S_{GDV} \sim O_4 + A_i + D + W + S_t$ | s | 27429.55 | 0.924176 |
| GWR | $S_{GDV} \sim O_4 + A_i + s + D + S_t$ | W | 27742.67 | 0.9208344 |
| GWR | $S_{GDV} \sim O_4 + \mathbf{A_i + s + D + W}$ | $\mathbf{S_t}$ | **18050.76** | **0.9916192** |

**Table 3: Comparison of Adjusted R-squared and second-order Akaike Information Criterion (AICc) between the simple linear regression and the GWR model.**

| Model | $R^2$ | AICc | p-value |
|---|---|---|---|
| OLS | 0.02 | 42720 | <0.001 |
| GWR | 0.99 * | 18851 | - |

*Quasi-global $R^2$

**Table 4: Classification scores for each predictor. A score of 3 was given to highly suitable areas and 1 to highly less suitable areas for GDV.**

| Predictor | Class | Score |
|---|---|---|
| Slope | 0%-5% | ~~1~~3 |
| | 5%-10% | 2 |
| | >10% | ~~3~~1 |
| Groundwater Depth | >15 m | 1 |
| | 1.5m-15m | 3 |
| | ≤1.5m | 1 |
| Aridity Index | 0.6-0.68 | ~~1~~3 |
| | 0.68-0.75 | 2 |
| | ≥0.75 | ~~3~~1 |
| Ombrothermic Index of the summer quarter and the immediately previous month | <0.28 | 1 |
| | 0.28-0.64 | 2 |
| | ≥0.64 | 3 |
| Drainage Density | ≤0.5 | 3 |
| | >0.5 | 1 |

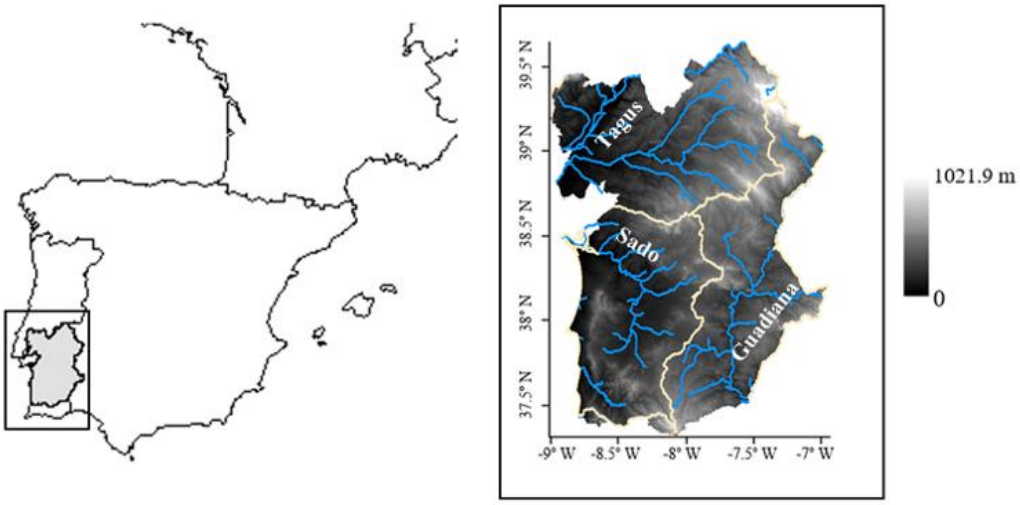



**Figure 01: Study area. On the left the location of Alentejo in the Iberian Peninsula; on the right, the elevation characterization of the study area with the main river courses from Tagus, Sado and Guadiana basins (white line) . Names of the main rivers are indicated near to their location in the map.**

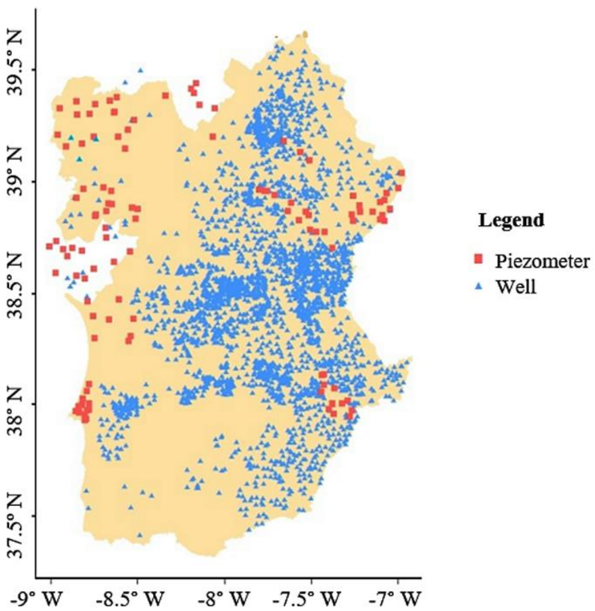



**Figure 02: Large well and piezometer data points used for groundwater depth calculation. Squares represent piezometers data points and triangle represent large well data points.**

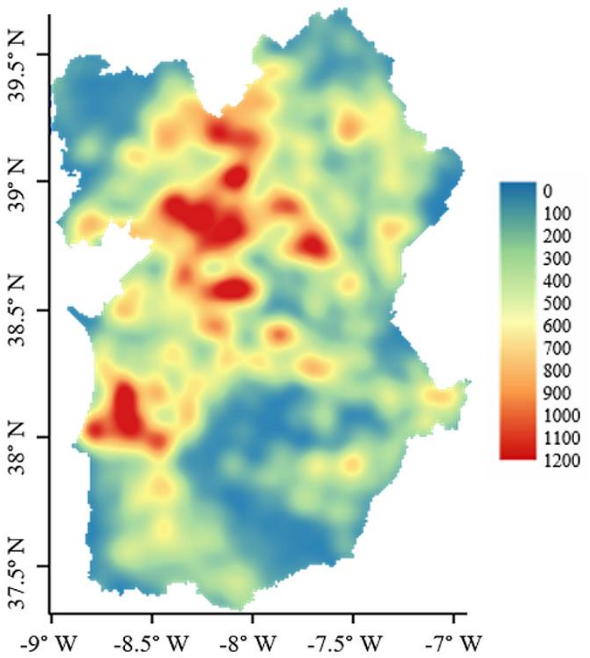



**Figure 03: Map of Kernel Density weighted by cover percentage of *Q. suber*, *Q. ilex* and *P. pinea*~~.~~. The scale unit represent the number of occurrences per 10 km search radius (~314 km2).**

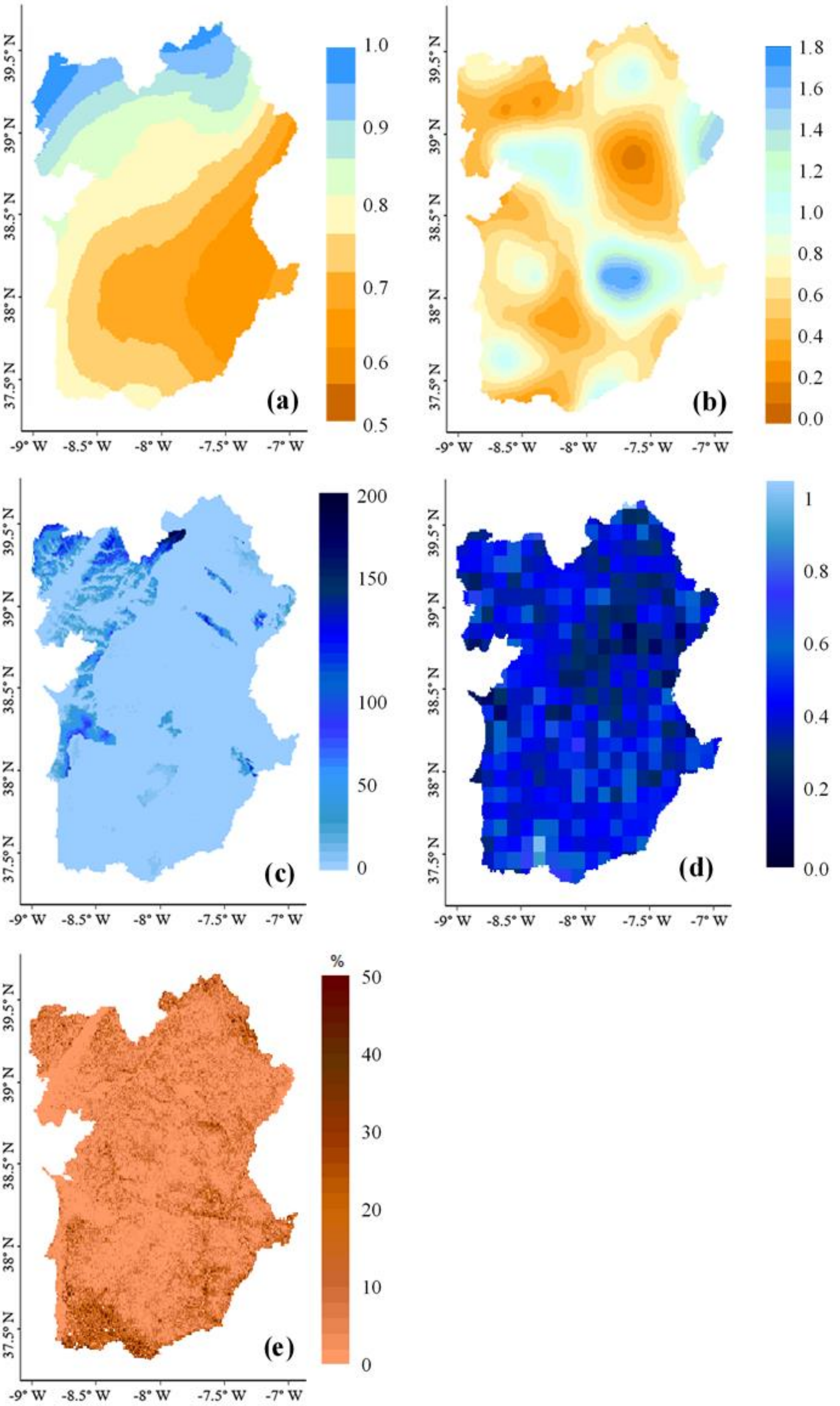

(a)
1.0
0.9
0.8
0.7
0.6
0.5
(b)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
(c)
200
150
100
50
0
(d)
1
0.8
0.6
0.4
0.2
0.0
(e)
%
50
40
30
20
10
0
39.5° N
39° N
38.5° N
38° N
37.5° N
-9° W
-8.5° W
-8° W
-7.5° W
-7° W

**Figure 04: Map of environmental layers used in model fitting. (a) – Aridity Index; (b) – Ombrothermic Index of the summer quarter and the immediately previous month; (c) – Groundwater Depth; (d) –Drainage density; (e) –Slope.) ~~– Soil type; (b) – Slope; (c) – Groundwater Depth; (d) – Ombrothermic Index of the summer quarter and the immediately previous month; (e) – Aridity Index.~~**

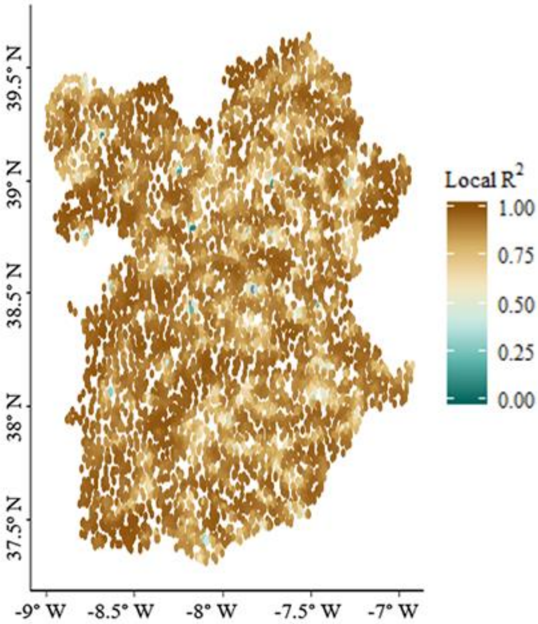



**Figure 05: Spatial distribution of local $R^2$ from the fitting of the Geographically Weighted Regression.**

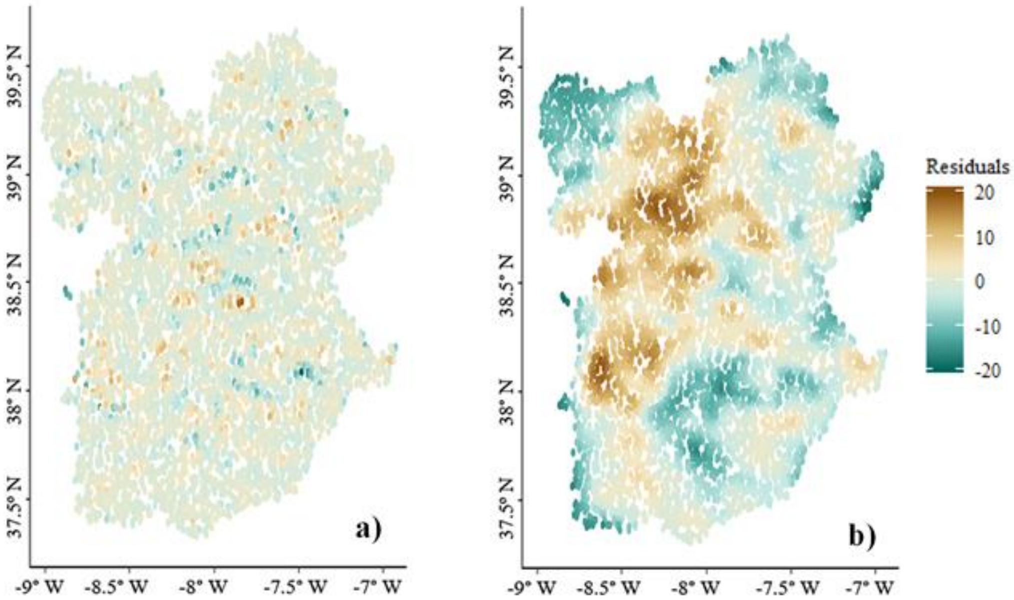



**Figure 06: Spatial distribution of model residuals from the fitting of the Geographically Weighted Regression (a) and Simple Linear model (b).**

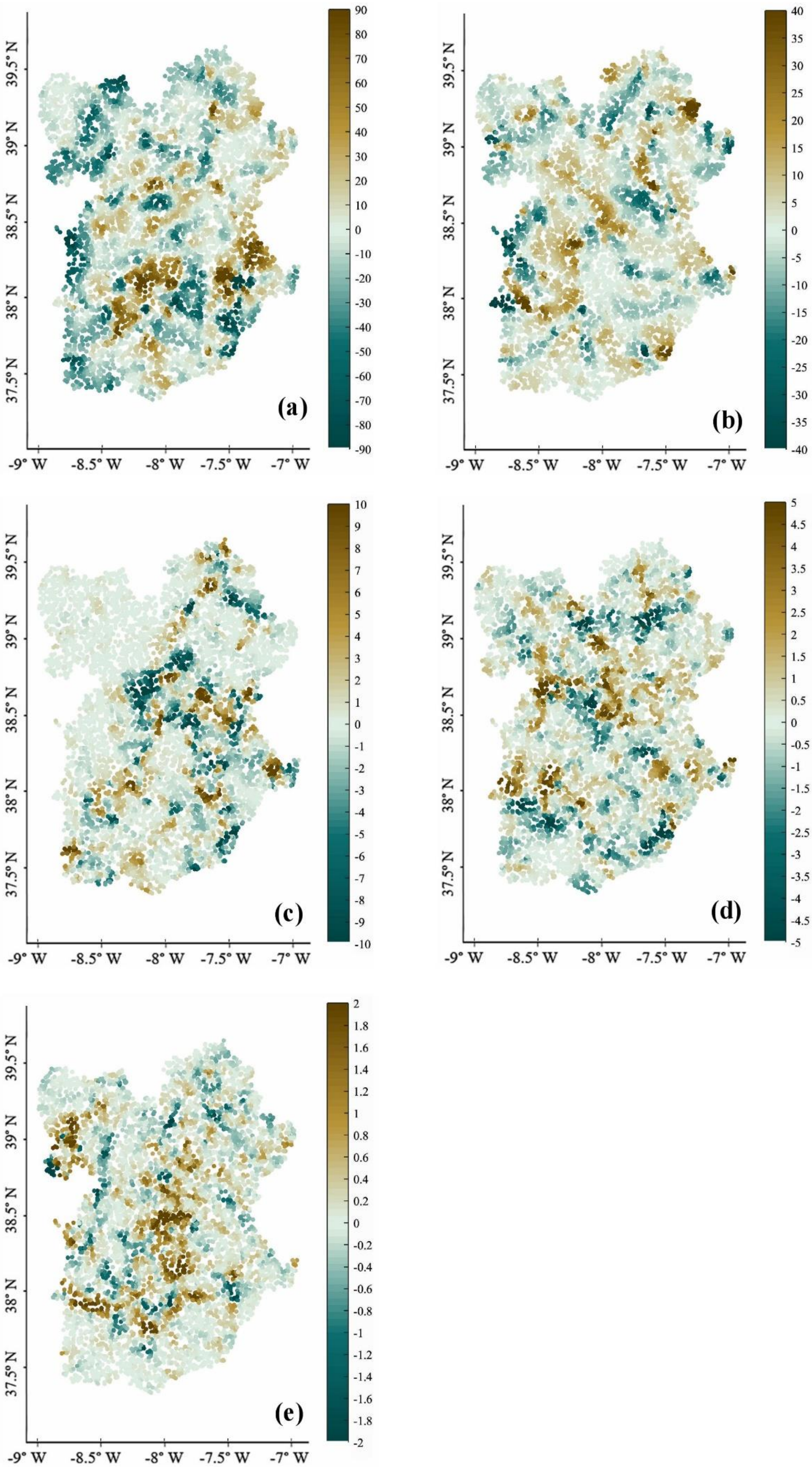

(a)
(b)
(c)
(d)
(e)
39.5° N
39° N
38.5° N
38° N
37.5° N
-9° W
-8.5° W
-8° W
-7.5° W
-7° W

**Figure 07: Map of local model coefficients for each variable. (a) – Aridity Index; (b) - Ombrothermic Index of the summer quarter and the immediately previous month; (c) – Groundwater Depth; (d) – Drainage density and (e) – Slope.**

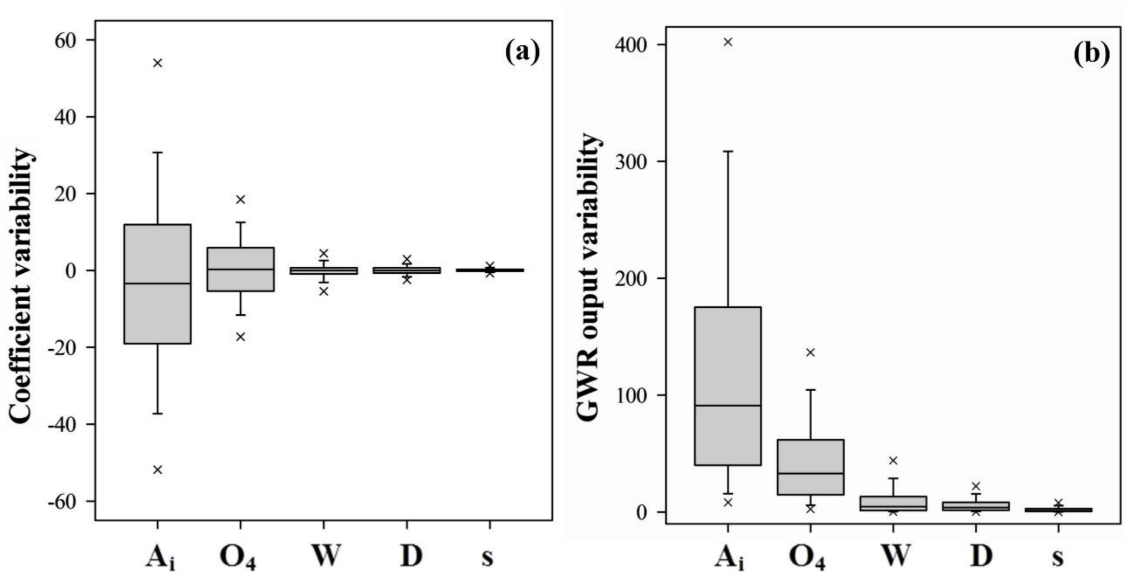



**Figure 08 – Boxplot of GWR model coefficient values for each predictor (a) and boxplot of the GWR model outputs, corresponding to GDV's density after each of the predictors was disturbed for the sensitivity analysis (b). $A_i$ stands for Aridity Index; $O_4$ for the ombrothemic index of the hottest month of the summer quarter and the immediately previous month; W for the groundwater depth, D for the drainage density and s for the slope. Error bars represent the $25^{th}$ and $75^{th}$ percentile while crosses indicate the $95^{th}$ percentile.**

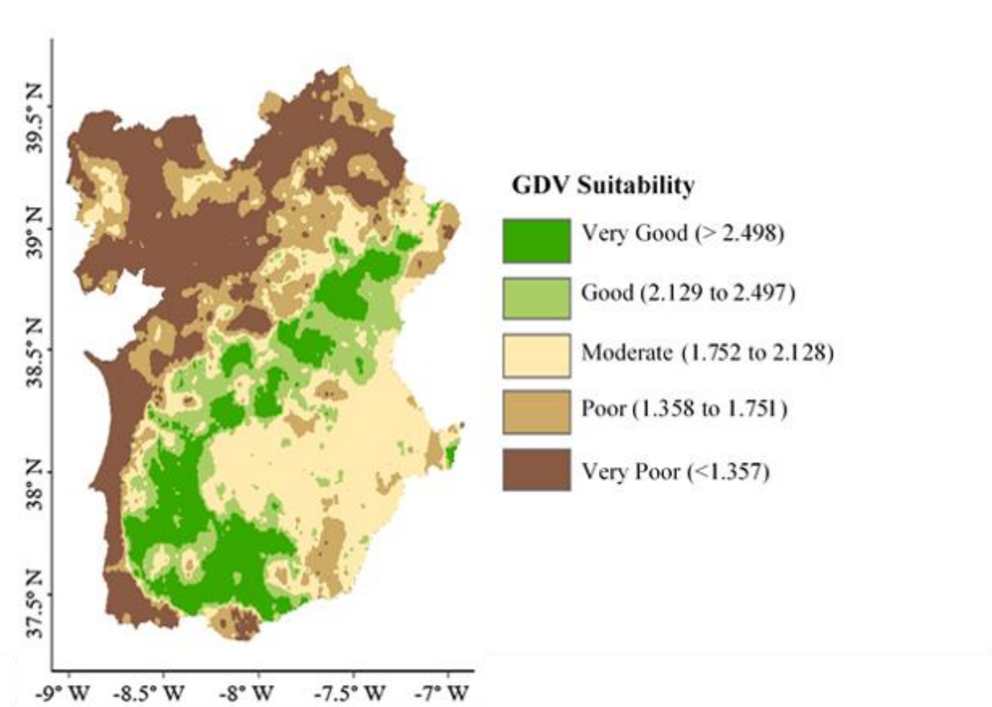



**Figure 09: Suitability map for Groundwater Dependent Vegetation.**

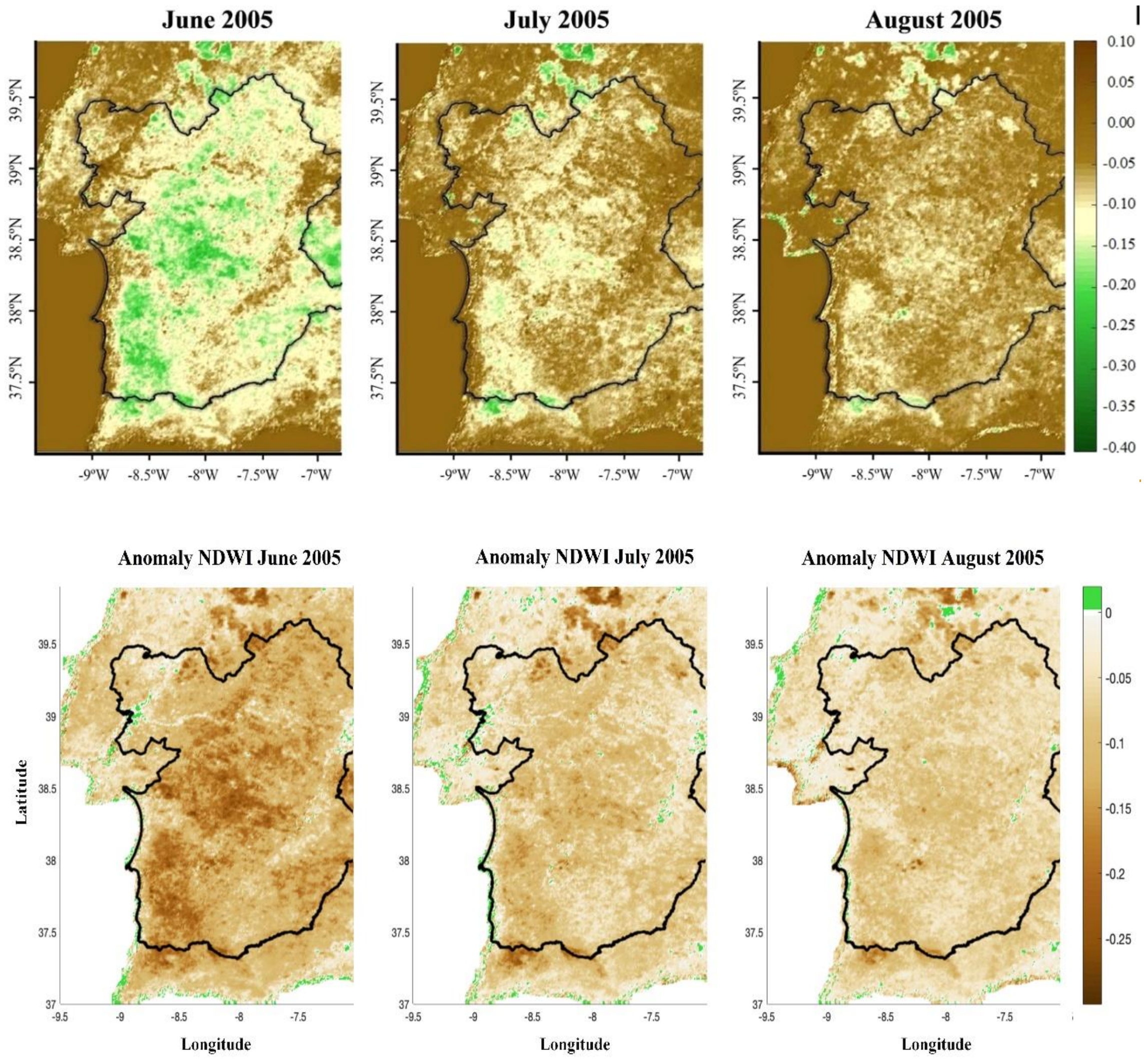



**Figure 10:** Spatial patterns of NDWI anomaly values considering the months of June, July and August of the extremely dry year of 2005, in reference to the same months of the period 1999-2009, in the Alentejo region. Dark brown colors (corresponding to extreme negative NDWI anomaly values) indicate the vegetation that experienced the highest loss of water in leaves in summer 2005 as compared to the reference period 1999-2009, while light brown colors show NDWI anomaly values very close to the usual vegetation moisture condition of the considered month ~~**NDWI anomaly considering the months of June, July and August of the extremely dry year of 2005, in reference to the same months of the period 1999-2009, in the Alentejo region. Green colors (corresponding to low NDWI values) indicates vegetation canopy undergoing a higher water stress than the average reference period 1999–2009.**~~

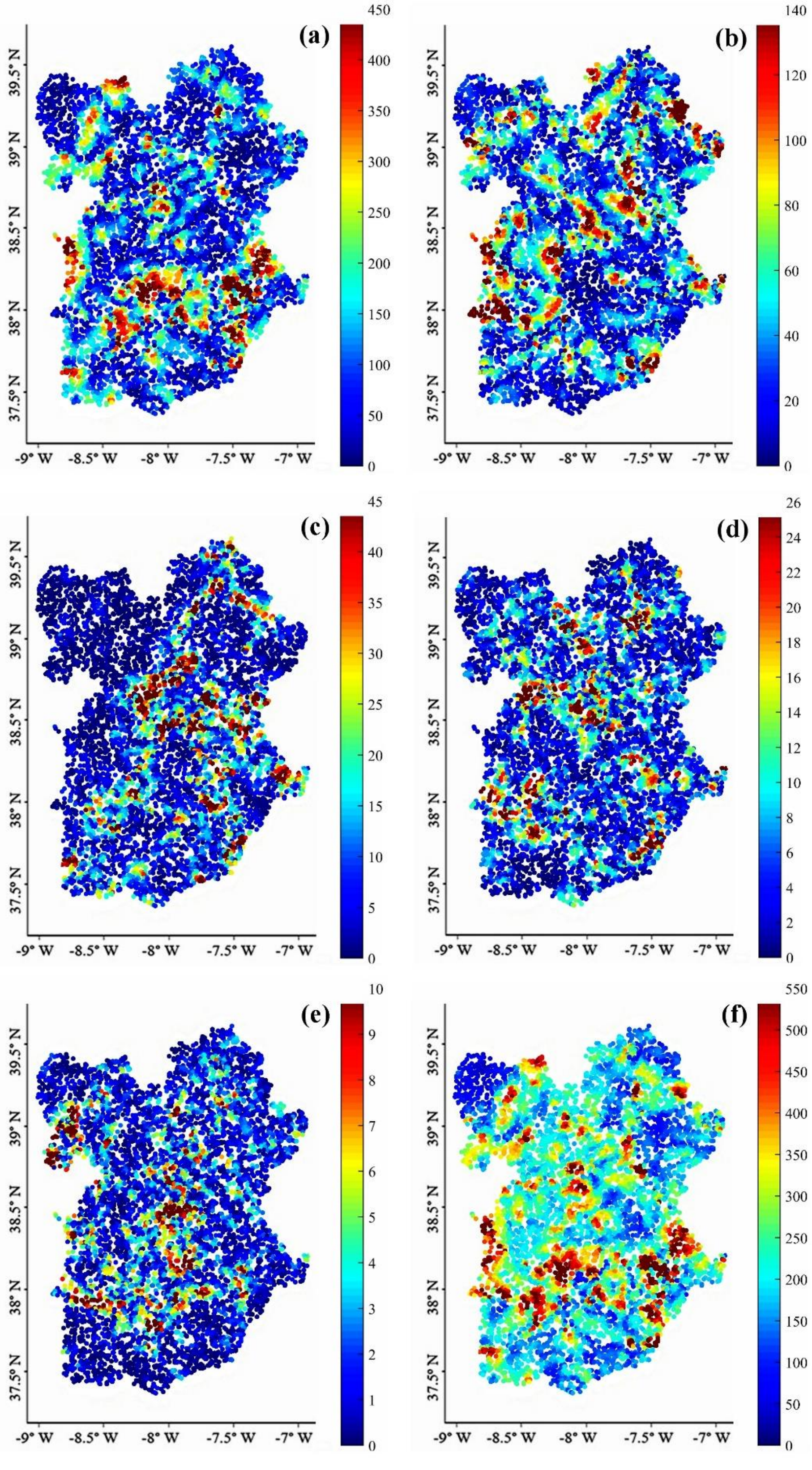



**Figure 11: Sensitivity analysis performed on the GWR model by perturbing one of the predictors, while remaining the rest of the model equation constant. Graphics present the output range of GDV's density when the aridity index (a), the ombrothermic index (b), the groundwater depth (c), the drainage density (d) or the slope variable (e) was perturbed; and the maximum possible range combining all predictors (f). The 95th percentile was used for the maximum value of the color bar for a better statistical representation of the spatial variability.**