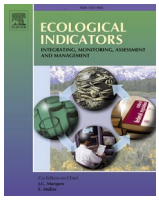
Ecological Indicators 128 (2021) 107820

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Original Articles

Delimiting the spatio-temporal uncertainty of climate-sensitive forest productivity projections using Support Vector Regression

M.A. González-Rodríguez [a,b,*], U. Diéguez-Aranda [b]

[a] CERNA Ingeniería y Asesoría Medioambental S.L., R/ Illas Cíes, 52-54-56, Ground floor, 27003 Lugo, Spain
[b] Unidade de Xestión Ambiental e Forestal Sostible, Departamento de Enxeñaría Agroforestal, Universidade de Santiago de Compostela, Escola Politécnica Superior de Enxeñaría, R/ Benigno Ledo, Campus Terra, 27002 Lugo, Spain

ARTICLE INFO

Keywords:
Pinus radiata
Site index
Extrapolation
Machine learning
Climate change
Forest growth

ABSTRACT

As climate change makes many traditional empirical growth approaches not functional for forest dynamics modelling, new climate-sensitive models are needed. However, using these newly developed models for extrapolation, such as predicting forest productivity for new areas or future scenarios is still a difficult task. In this study, we proposed a method for delimiting the uncertainty of climate-sensitive extrapolations of forest productivity (site index, *SI*) using the regularisation approach implicit in distance-based Support Vector Regression. As a case study, we predicted forest productivity with a dataset of 165 permanent research plots of radiata pine forests in Galicia (NW of Spain) as a function of bioclimatic variables from the Worldclim 2 raster datasets. The developed model was based on the radial basis kernel and, after calibrating it using cross-validation, produced adequate performance metrics, explaining up to 56% of the site index' variability. Then, we predicted forest productivity for the Galician territory basing on climate raster maps for current conditions and six future scenarios (using different Global Climate Models) and evaluated the resulting maps by delimiting the surfaces with predictions strongly regressed to the mean. This analysis revealed that the extrapolations for unseen climatic conditions were extremely regularised, even for current climate, being 60–99% of the territory regressed to the observational site index mean. In other words, the validity area delimited for the fitted model was narrow in comparison with the prediction extent. These results imply that the climatic conditions in these areas/scenarios were too different from the training datastet for making reliable predictions, at least under the optimum model setup defined by cross-validation. However, when we reduced the σ parameter, responsible for controlling distance-based regularisation, we observed a noticeable increase in validity area of the model, together with a drop in performance. This fact revealed the existence of a trade–off between highly specific models, with high performance and a small applicability area, and more generalisable models, with a broad validity area but lower performance. We concluded that the tested methodology could be a useful starting point for assessing the spatio-temporal uncertainty of forest productivity predictions in the future.

1. Introduction

Accurately predicting the growth of forest stands is currently a primary technical concern both for scientific purposes and for practical forest management. During the last decades, the uncertainty derived from climate change has intensified the need for forest growth models able to cope with future unseen climatic conditions (Bontemps and Bouriaud, 2014). This is due to the lack of *climate-sensitiveness* of many traditional growth prediction approaches, which may make them an infeasible alternative in a changing climate (Kahle et al., 2008). From the perspective of empirical growth modelling, this challenge has been usually faced by developing growth-environment models (Fontes et al., 2010), which allow for connecting growth indicators with biophysical site conditions. In most of the cases, this task has been accomplished by modelling the site index (*SI*), the most frequent empirical growth indicator (Skovsgaard and Vanclay, 2008), as a function of climate by using regression techniques (Monserud et al., 2006; Weiskittel et al., 2011; Shen et al., 2015). Usually, the applied techniques tend to be non-spatial, though geostatistical approaches, such as krigging, have also provided good results (Nothdurft et al., 2012), remaining, however, the

This article is the result of a research project partially funded by the Spanish Ministry of Science, Innovation and Universities.
* Corresponding author at: CERNA Ingeniería y Asesoría Medioambental S.L., R/ Illas Cíes ní 52-54-56, Ground floor, 27003 Lugo, Spain.
E-mail addresses: miguelangel.gonzalez.rodriguez@rai.usc.es (M.A. González-Rodríguez), ulises.dieguez@usc.es (U. Diéguez-Aranda).

https://doi.org/10.1016/j.ecolind.2021.107820
Received 1 June 2020; Received in revised form 12 January 2021; Accepted 21 May 2021
Available online 26 May 2021

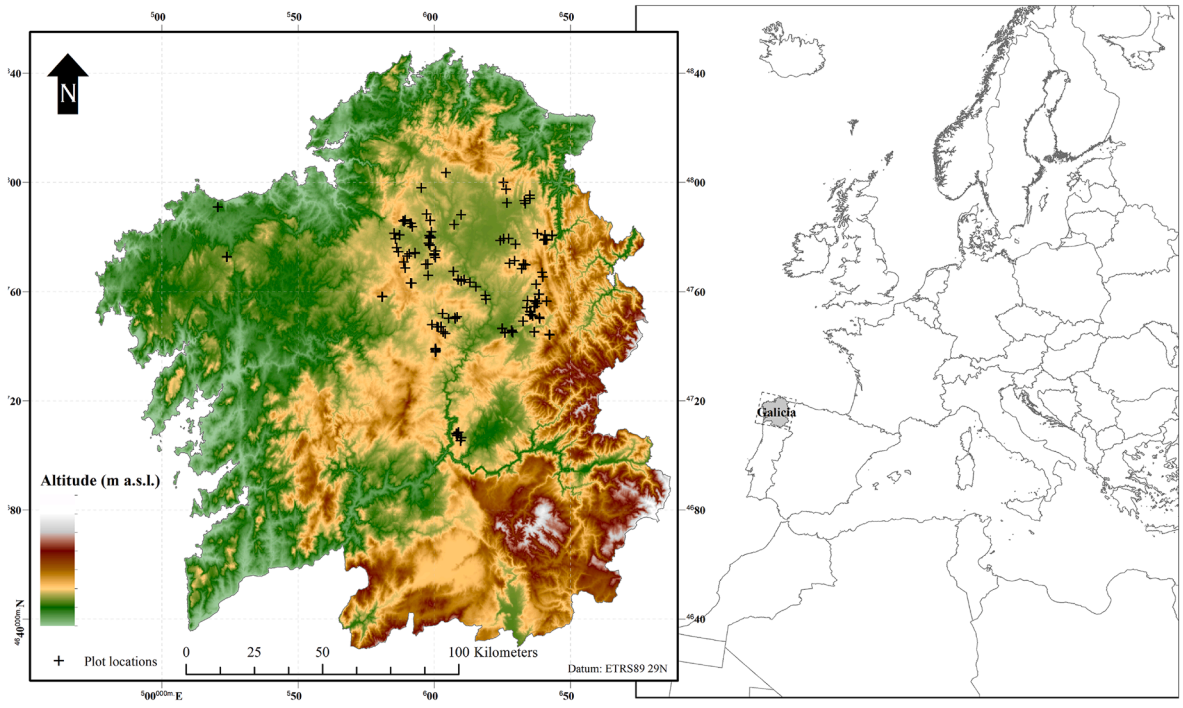


Fig. 1. Locations of the 165 research plots of pure, even-aged stands of radiata pine in Galicia, Spain.

non-parametric non-spatial alternatives the best in terms of performance (Watt et al., 2021). In this sense, machine learning techniques such as bagging, boosting or regularisation have become a frequent choice for fitting non-parametric regressions able to represent complex *SI*-environment relationships (Aertsen et al., 2010; Sabatia and Burkhart, 2014; González-Rodríguez and Diéguez-Aranda, 2020). Often, these models have been used for producing cartographic outputs that allow for "mapping" forest productivity over broad territories and future climate scenarios (Monserud et al., 2006; Jiang et al., 2015; Parresol et al., 2017). Despite the certainly powerful performance of machine learning techniques for predicting *SI*, the resulting models are frequently hard to interpret (Aertsen et al., 2010) and suffer from systematic drawbacks, such as regression to the mean (Sabatia and Burkhart, 2014). Moreover, the reliability of these *SI*–climate models is commonly evaluated only using the same original dataset (cross-validation, Hlásny et al. (2017) and González-Rodríguez and Diéguez-Aranda (2020)) or some supplementary sample statistically similar to the latter (validation *per se*, Sabatia and Burkhart, 2014). Thus, their performance regarding extrapolation errors resulting from predicting over strictly unseen climatic conditions (i.e., new areas or future scenarios) remains, to the date, unevaluated. In other words, it is not clear how reliable are predictions for climate conditions too different from the observed. From a more statistical perspective, this also means that the validation performance metrics reported when fitting these models might not be generalisable to the geographical extents where they are usually used for prediction. We believe that this lack of understanding about the behaviour of extrapolation errors is an important source of spatio-temporal uncertainty concerning forest productivity projections. Assessing this uncertainty might be a crucial forestry research objective during the following decades, as it is a key step for developing climate change-adapted management (Lindner et al., 2014).

In this study, we made a first approach to the problem of evaluating the uncertainty associated with extrapolating site index for new areas or future scenarios. The main expected outcome from this approach was to provide a method, not only for predicting site index under new climatic conditions, but also for producing a cartographic delimitation of the areas were these predictions are reliable. Specifically, our approach consisted on delimiting the validity area of a predictive model by discriminating predictions regularised basing on kernel transformations between training observations and unseen climatic conditions. For attaining this goal, we drew on the implicit advantages of Support Vector Regression, a machine learning technique barely used in forestry, for regularising predictions under climatic conditions different from the observed. As a case study, we applied this technique for developing *SI*-climate models for radiata pine (*Pinus radiata* D. Don) stands in Galicia, NW of Spain.

Table 1
Summarized statistics of the stand variables in the 165 radiata pine plot.

Variable	Mean	St. dev.	Min	Max
t(year)	26	8	12	47
N (trees ha^{-1})	796	391	187	2382
G (m^{2}ha^{-1})	36.8	11.1	9.81	68.1
d_g (cm)	26.03	7.09	11.36	44.8
H (m)	24.2	4.9	11.8	35.6
SI (m)	20.7	3.06	13.3	27.8

Note: t = stand age, N = number of stems per hectare, G = basal area, d_g = quadratic mean diameter, H = dominant height, and SI = site index.

2. Materials and methods

2.1. Site index data

The *SI* data came from a network of 165 research plots established by the Sustainable Environmental and Forest Management Unit (UXAFORES) of the University of Santiago de Compostela, Spain, whose design was described by Castedo-Dorado et al. (2007). The plots were located in pure and even-aged stands throughout the area of distribution of radiata pine in Galicia, mainly in the province of Lugo (Fig. 1). The plot size ranged from 625 to 1200 m2, depending on stand density, to achieve a minimum of 30 trees per plot. Concerning silviculture, most of the plots were never thinned since plantation while others were lightly thinned from below, so, overall, the dominant height was not affected by treatments. The measurement of these plots implied the collection of (i) diameter at breast height of all the trees, (ii) total height of a sample of

trees, and (iii) bored core samples from dominant trees for stand age estimation. The dominant height of these stands (H, m) was defined as the mean height of the 100 thickest trees in a hectare. Then, *SI* was estimated for each plot using the Generalized Algebraic Difference model developed by Diéguez-Aranda et al. (2005) for this region. A summary of stand variables for the 165 research plots used in this study is shown in Table 1.

2.2. Climate data

2.2.1. Current climate

Our source of current climatic data was the Worldclim 2 bioclimatic dataset (Fick and Hijmans, 2017), which includes a collection of raster maps of 19 bioclimatic indicators with 1 km of spatial resolution. The climatic values provided by these maps correspond to historical averages, for the period 1970–2000, of variables commonly used in research about species distribution modelling due to their ecological meaning (see the documentation of the R package *dismo* by Hijmans et al. (2017)). A summary of the current Worldclim 2 variables used in this study for climate-sensitive modelling of *SI* is shown in Appendix A (Table 4).

2.2.2. Future climate

Though in some previous studies the best accuracy of future climate projections was achieved by averaging predictions coming from a large set of different Global Climate Models (GCM) (Reichler and Kim, 2008; Flato et al., 2013), for the purpose of this study we performed *SI* predictions only for a small subset of climate models. Our criteria for selecting these models was oriented by the validation performance analysis carried out by Flato et al. (2013) for the models included in Coupled Model Intercomparison Project Phase 5 (Taylor et al., 2012). The selected models by this way were three: GFDL-CM3 (Donner et al., 2011), HadGEM2-ES (Collins et al., 2011) and MPI-ESM-LR (Giorgetta et al., 2013). In order to encompass the potential variability between the different climatic future projections we used the two most contrasted Representative Concentration Pathways, RCP 2.6 and RCP 8.5, for future *SI* prediction. The corresponding values of bioclimatic variables coming from the Worldclim 1.4 downscaled future climate projections (Hijmans et al., 2005) for these three GCM models were used for prediction of future *SI* over the Galician territory.

2.3. Site index modelling

For the development of climate-sensitive *SI* models, we tested several machine learning techniques based on Support Vector Regression (SVR) (Vapnik et al., 1997). We chose these techniques among other modelling alternatives because they provide some potential advantages for growth-environment modelling:

- they allow for representing non-linear relationships between response and predictors,
- some non-linear variants of these techniques (e.g. non-linear kernels) allow for representing interactions between predictors,
- their regularisation strategy improves model robustness and reduces the risk of over-fitting, and
- their regularisation strategy (especially for distance-based non-linear kernels) can work in a floating way, performing a stronger flattening in observations that are very different to the training sample.

In addition, to our knowledge, SVR techniques have never been used before for climate-sensitive forest growth modelling.

Support Vector Regression, and specifically the ϵ-SVR variant, is a non-parametric modelling technique that combines an L2 regularisation (Hoerl and Kennard, 1970) strategy with a loss function defined by absolute errors that exceed a certain ϵ threshold. As L2 regularisation implies that, during the model fitting stage, the square norm of all model parameters is minimised together with prediction residuals, the models resulting from this approach are intended to be "flat" in terms of the parametric space and have therefore a high generalisation ability for extrapolation purposes. Since only in some cases prediction errors are higher than ϵ, not all the observations in the training set are used for calibrating model parameters. This is the reason why this loss function is commonly named ϵ-insensitive. The observations that do contribute to model calibration are then called Support Vectors. The current formulation of the fitting problem is the so-called *dual formula*, which is derived from transforming the cited regularisation and ϵ-insensitiveness principles into a Lagrangian optimisation problem subject to Karush–Kuhn–Tucker (Karush, 1939; Kuhn and Tucker, 1951) constraints:

$$\text{maximise} \quad \begin{cases} -1\Big/2\sum_{i,j=1}^{l}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)\left\langle x_i,x_j\right\rangle \\ -\epsilon\sum_{i=1}^{l}\left(\alpha_i+\alpha_i^*\right)+\sum_{i=1}^{l}y_i\left(\alpha_i-\alpha_i^*\right) \end{cases} \tag{1}$$

$$\text{subject to} \quad \sum_{i=1}^{l}\left(\alpha_i-\alpha_i^*\right)=0 \quad \text{and} \quad \alpha_i,\alpha_i^*\in\left[0,C\right], \tag{2}$$

where x and y are respectively the predictors' matrix and response vector, α_i and α_i^* are Lagrange multipliers, l is the number of observations in the training dataset, C is the cost parameter that penalises errors that exceed the threshold ϵ, and $\langle\cdot,\cdot\rangle$ denotes a dot product.

While (1) is the general approach to the problem for building linear SVR models, a variety of *kernel* functions have been developed for generating non-linear combinations of predictors, or *features*. This can be achieved by substituting the product $\langle x_i,x_j\rangle$ in (1) by a certain $k(x_i,x_j)$ function. For instance, the homogeneous polynomial transformation of $\langle x_i,x_j\rangle$ is:

$$k(x,x')=\langle x,x'\rangle^q, \tag{3}$$

with $q\in\mathbb{N}$. Another common example is the radial basis kernel, with the following form:

$$k\left(x,x'\right)=\exp\left(-\sigma\|x-x'\|^2\right), \tag{4}$$

where σ is the dispersion parameter that controls the impact of the norm $\|x-x'\|^2$ on model flatness.

Once the optimal Lagrange multipliers are found, predictions based on the SVR model are made as follows:

$$\hat{y}=\sum_{i=1}^{l}\left(\alpha_i-\alpha_i^*\right)k\left(x_i,x\right)+b, \tag{5}$$

being b the independent parameter or model offset, which can be computed by a variety of methods, some of them described by Keerthi et al. (2001), and that tends to match the response's observational mean.

Overall, in SVR models the regularisation effect on predictions (i.e. the regression to the mean) is controlled mainly by C (L2 regularisation) and σ (distance-based regularisation). The role of C is to globally constraint, independently on the kernel transformation, the maximum values of model slopes (α_i and α_i^*) and hence the maximum variability of predictions. σ, in the case of distance-based kernels, constraints predictions based on Euclidean norms between support vectors and new conditions. In other words, when making predictions with a distance-based kernel, if the new conditions are too different from the observed in the training dataset, regularisation will consequently shrink these predictions towards the null model. Derived from this, we could consider that calibrating the intensity of this regularisation represents a trade-off between spatial specificity and generality: an intense regularisation (high values of σ) can produce models with high performance but with a narrow applicability area, while a slight regularisation (low σ)

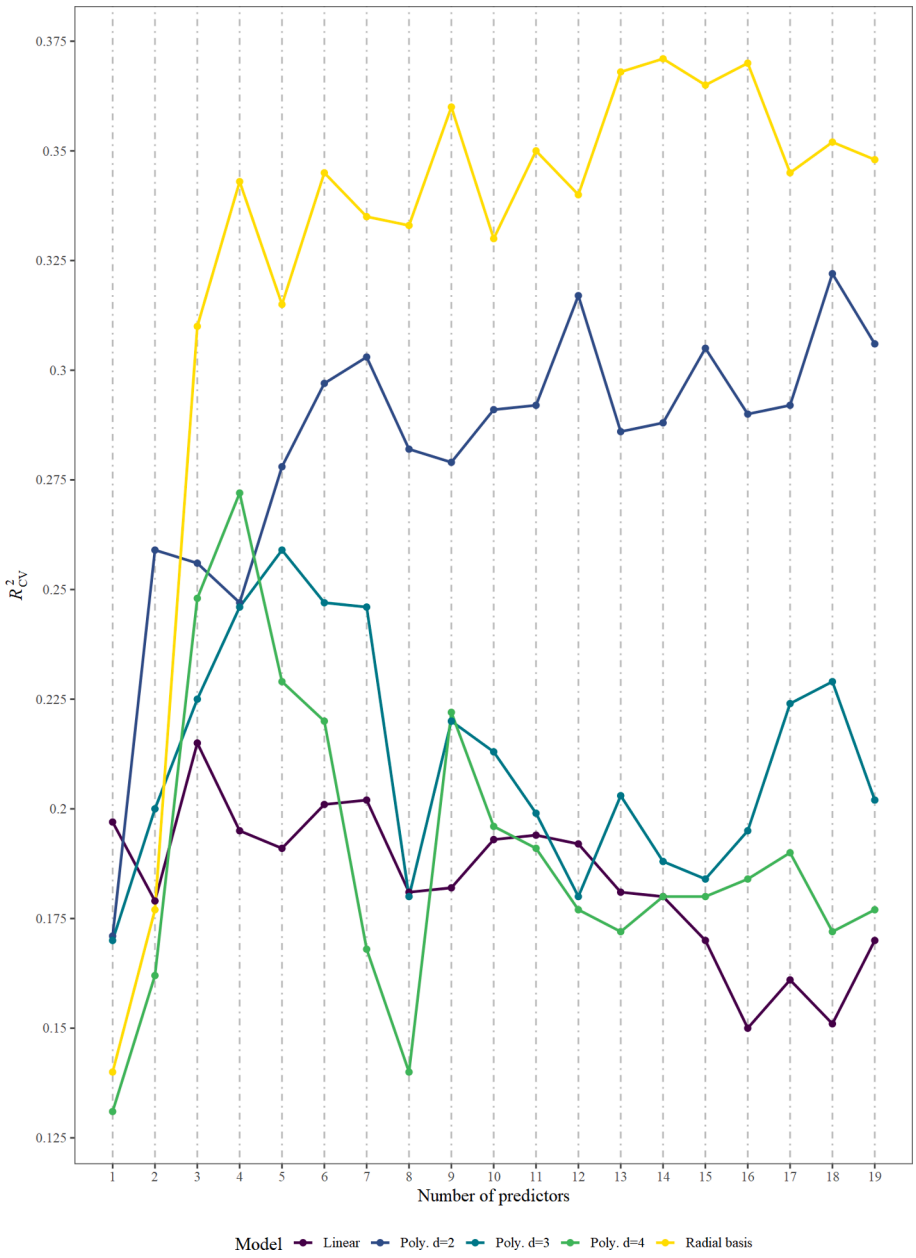


Fig. 2. Evolution of the cross-validated R^2 along the recursive variable elimination paths for the five SVR models fitted.

leads to models with an extensive validity area at the expense of having poor performance. We believe that managing this trade-off might be a useful approach for evaluating the uncertainty, in the form of predictive errors, derived from the extrapolation of ecological indicators, such as the *SI*. Hence, considering the potential practicality of using a distance-based approach for differentiating unseen climatic conditions from training conditions, we focused our *SI* modelling and uncertainty analysis on distance-based SVR methods, specifically, on the radial basis kernel. However, for the purpose of comparison, we also fitted some other SVR models, such as the linear and polynomial kernels of 2^{nd}, 3^{rd} and 4^{th} degree. Thus, we fitted a total of five different SVR models based on the implementation of the package *kernlab* (Karatzoglou et al., 2004) of the R language (R Core Team, 2020). As suggested by Karatzoglou et al. (2004), the σ parameter in the radial basis kernel was initially calibrated using the method proposed by Caputo et al. (2002), consisting on calculating the median in the 10^{th}-90^{th} percentile range of $\|x - x'\|^2$ norms. The calibration of ϵ and C constants was carried out with the R package *rminer* (Cortez, 2016) using 10-times repeated 10-fold cross-validation (CV) for ensuring the robusticity of the estimates. Variable selection was performed through a Recursive Feature Elimination (RFE) procedure (Fig. 2), in which the predictors were progressively dropped off from the set depending on their impact on model predictions. The impact of each predictor was measured using the one-dimensional sensitivity analysis implemented in the function *Importance* of *rminer* package, which provides a score based on drops in model performance after removal.

Similarly to previous studies (Weiskittel et al., 2011; González-Rodríguez and Diéguez-Aranda, 2020), the optimum number of predictors for each model was defined as the one previous to a significant decrease in the cross-validated R^2. We applied this criteria by setting a threshold value 5–10% of decrease in performance after the dropping off of each predictor. Once the number of predictors was set for each model, we carried out a second cross-validation procedure with 100 repetitions for ensuring the independence of the multipliers' estimates and goodness-of-fit statistics on the random seed. During this procedure, we re-calibrated the σ parameter in the radial basis kernel, trying values

Table 2
Summary of the number predictors and support vectors, optimum hyperparameters C, ϵ and σ, and apparent and cross-validation performance statistics of the five SVR models fitted.

Model	N° predictors	N_{sv}	Hyperparameters	NRMSE	R^2	$NRMSE_{cv}$	R^2_{cv}
SVR_{RB}	4	136	$C = 9$, $\epsilon = 0.15$, $\sigma = 0.69$	0.0979	0.559	0.118	0.386
SVR_{L}	3	124	$C = 5$, $\epsilon = 0.3$	0.132	0.1902	0.133	0.184
SVR_{P2}	5	123	$C = 7$, $\epsilon = 0.3$	0.118	0.3501	0.127	0.268
SVR_{P3}	5	139	$C = 9$, $\epsilon = 0.15$	0.114	0.398	0.125	0.288
SVR_{P4}	4	157	$C = 9$, $\epsilon = 0.05$	0.09803	0.558	0.126	0.269

along the 10th-90th percentile range and choosing the one that maximised cross-validated R^2. For subsequent analyses, we are labeling this value of σ, optimum from the perspective of cross-validation performance, as σ_{CV}. Finally, we compared the five models' apparent and CV performance statistics (Root Mean Square Error normalised by the response's mean, and R^2) and tested whether the radial basis model (hereunder, SVR_{RB}) provided better estimates than the other four models.

Considering that interpreting SVR models is currently a real technical challenge (Burns et al., 2019), for evaluating the ecological meaning of the selected model we resorted to a method specifically developed by Üstün et al. (2007) for this technique. This methods aims at recovering useful information about the modelled predictors-response relationships from the *kernel matrix* resulting from the $k(x, x')$ transformation. This is achieved by calculating the correlation between each row of this matrix and each one of the predictors included in the SVR model. For recalculating the kernel matrix we used only the N_{sv} support vectors, excluding the non-relevant observations (i.e., those with $(\alpha_i - \alpha_i^*) = 0$). Then, we sorted the N_{sv} observations by ascending value of predicted site index ($\widehat{SI}$) and calculated the kernel transformation based on the fitted SVR model. Finally, we computed the Pearson's correlation between the selected predictors and the kernel matrix generating the so-called Correlation Image. This matrix was subsequently used for assessing the role of each climatic predictor on estimated *SI*.

2.4. Delimiting uncertainty

After we fitted and compared the different support vector models, we used the radial basis model (SVR_{RB}) for generating raster maps of forest productivity predictions over the complete Galician territory for current climate and also for the six future climate scenarios considered (GFDL-CM3, HadGEM2-ES and MPI-ESM-LR models for the Representative Concentration Pathways 2.6 and 8.5). As the radial basis regularisation flattens predictions over the territory based on their dissimilarity with the training dataset (in predictor space), model predictions under unseen climatic conditions are systematically regressed to the mean. As a consequence, identifying highly regularised predictions allows for mapping the areas where the fitted model, basing on the current regularisation setup (σ), is not able to make predictions. Derived from this, we focused our uncertainty analysis on detecting and delimiting highly regularised areas. The major difficulty for performing this delimitation is to be able to discriminate highly regularised values from actual average predictions. Our approach for solving this was to identify non-regularised values in two stages: (1) firstly, we extracted predictions differing more than 5% from the observational *SI* mean (20.8 m) and (2) secondly, we repeated the "extraction" procedure after performing a smoothing operation of the raster pixels using a 3x3 moving window. Then, we classified the extracted values in this way as non-regularised and, after translating the reclassified raster pixels into spatial polygons, we obtained an estimation of regularised and non-regularised surfaces for all the *SI* raster maps generated. For improving visualisation, we also smoothed the edges of the resulting regularisation

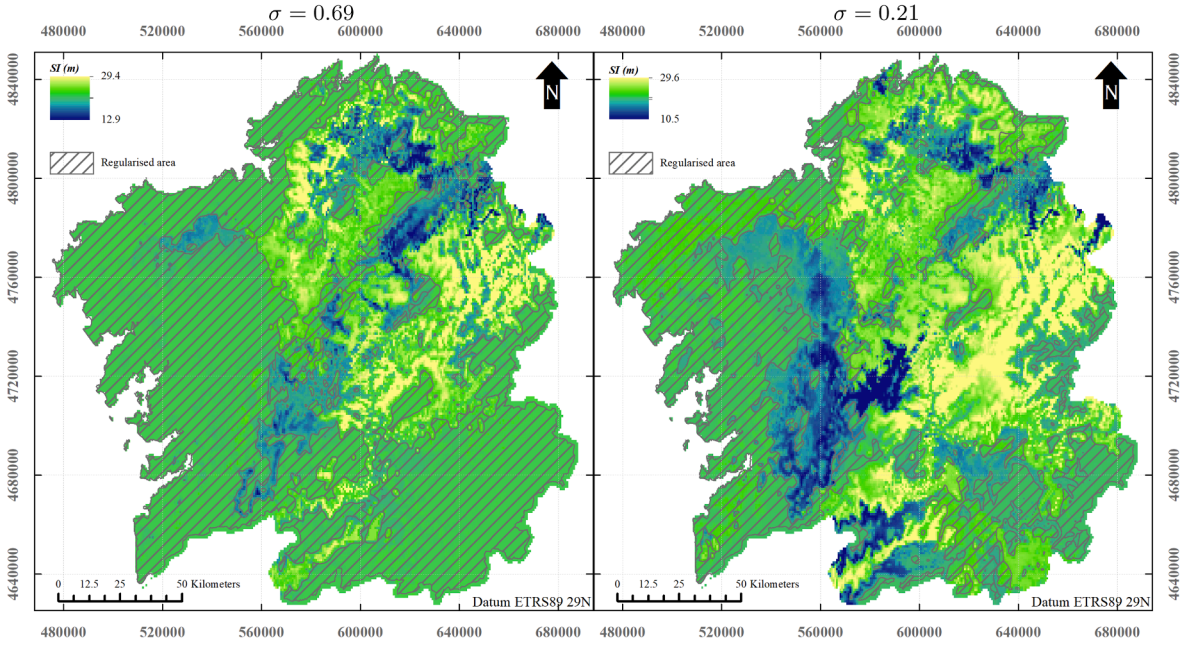


Fig. 5. Maps of site index predictions using σ_{CV} ($\sigma = 0.69$, left) and $\sigma = 0.21$ (right) for current climate. The delimited regularised areas are represented with a striped pattern.

Table 3
Summary of regularised surfaces and predicted *SI* values for the current and future climate maps with the radial basis model. The values corresponding to the reduced σ (=0.22) are presented in grey color.

	Map	Reg. (%)	SI_{min}	SI_{max}
	Current climate	0.605 0.364	11.6 10.5	28.9 29.6
RCP 2.6	GFDL-CM3	0.991 0.903	17.3 15.5	19.9 20.8
	HadGEM2-ES	0.973 0.851	16.7 16.2	23.7 25.5
	MPI-ESM-LR	0.725 0.485	16.1 11.1	23.4 29.1
RCP 8.5	GFDL-CM3	0.998 0.922	18.9 17.4	19.9 20.6
	HadGEM2-ES	0.996 0.896	17.6 16.6	19.8 23.3
	MPI-ESM-LR	0.907 0.752	14.9 14.3	26.5 27.8

surfaces using the R package *smoothr* (Strimas-Mackey, 2020). For the rest of geoprocessing operations we used the R packages *raster* (Hijmans, 2019) and *rgeos* (Bivand and Rundel, 2019).

As σ_{CV} is calibrated for maximising predictive performance exclusively, we carried out a σ reduction analysis for discussing the ability of the radial basis model for predicting *SI* along the specificity–generality trade–off. This consisted on reducing the value of σ iteratively, from σ_{CV} to the 10th percentile of the $\|x-x'\|^2$ norms range, and carrying out a performance test (both apparent and cross-validated) and a delimitation of regularised areas (for current climate) at each iteration. As a result, we obtained a path representing the variation of regularised surfaces and R^2 along the calibration range of σ. Once we had performed this analysis, we discussed potential criteria for choosing the best value of σ along the path, depending on the modelling objectives and constraints.

3. Results

3.1. Model performace

The performance statistics and calibration hyperparameters of the SVR fitted models are summarized in Table 2. The cross-validation performance of these models ran parallel to the apparent performance, being in both cases the radial basis model (SVR_{RB}) the best in terms of R^2 (0.56) and NRMSE (0.098), followed by the polynomial models (in order of decreasing degree) and, finally, the linear. The number of predictors included in the fitted models ranged from three (linear model) to five (polynomials), being four the number of predictors in the radial basis model. These predictors we, in decreasing order of variable importance score: *bio*2 (importance score = 0.273), *bio*3 (0.245), *bio*1 (0.241) and *bio*6 (0.2402).

Concerning hyperparameters, the values of ϵ tended to be close to the lower bound its variation range (minimum of 0.05 in fourth degree polynomial model) which consequently led to a high proportion of support vectors, whose number varied from 123 to 157. The optimum *C* values close to the maximum (the most frequent value was 9), which implied a low global regularisation intensity. The resulting σ hyperparameter from the cross-validation procedure (σ_{CV}) was 0.69, which is slightly higher than the mean value of the $\|x-x'\|^2$ norms range.

3.2. Regularisation analysis

The predicted map of site index for current climate using σ_{CV} (Fig. 5) yielded similar statistics than the observed *SI*, though it had a slightly higher variation range (12.9–29.4 m and 13.3–27.8 m, respectively). The regularised surface estimated for this map is approximately 60% of

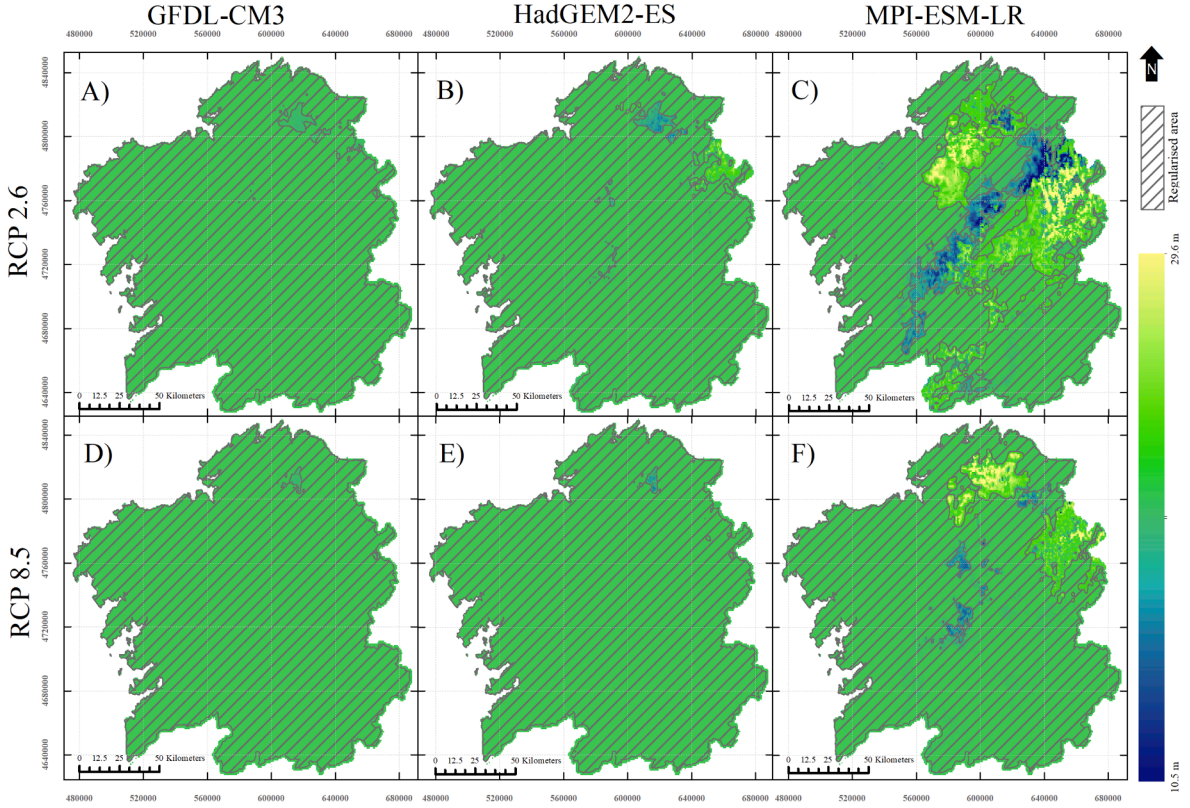


Fig. 6. Map of *SI* predictions using the radial basis model with optimum σ (σ_{CV} = 0.69) for the six climatic scenarios considered. The upper row -A, B and C plots- corresponds to the RCP 2.6, while the lower one -plots D, E and F- corresponds to RCP 8.5. The columns correspond to GFDL-CM3 (left), HadGEM2-ES (center) and MPI-ESM-LR (right) GCM models. The delimited regularised areas are represented with a striped pattern.

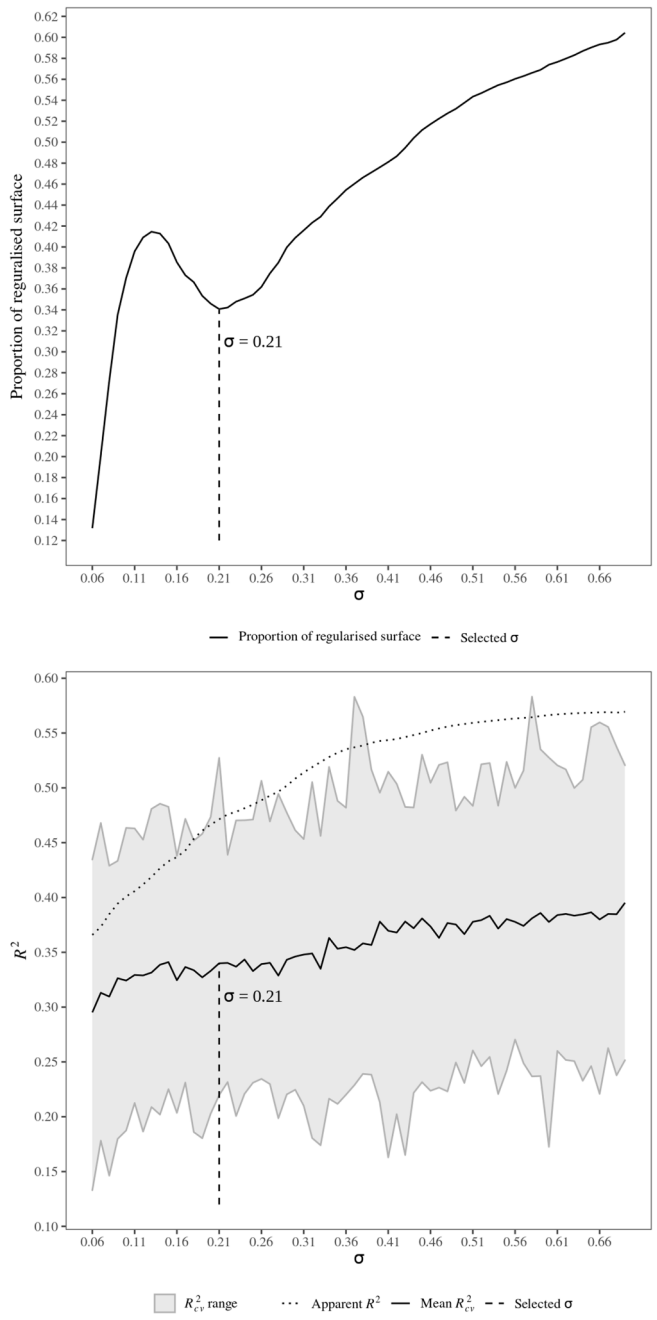


Fig. 4. Estimated proportion of regularised areas (top) and apparent and cross-validated performance (bottom) along the σ reduction path for the SVR_{RB} model.

the territory (Table 3).

Regarding future climate maps for σ_{CV} (Fig. 6), the variability of predictions was, overall, very low. For instance, GFDL-CM3 yielded values that ranged from 18.9 to 19.9 for the concentration pathway 8.5. Consequently, regularisation in these maps was very intense, but also variable across different models/scenarios. As expected, predictions for RCP 8.5 were more regularised than those for the RCP 2.6 scenario, accounting for the diverging dynamic of future climate with respect to present concentrations of greenhouse gases. GFDL-CM3 produced the most extensive regularised surfaces, accounting for 99.1% of the territory for RCP 2.6 and 99.8% for RCP 8.5. HadGEM2-ES provided very similar but slightly smaller regularised areas than GFDL-CM3. MPI-ESM-LR, on the contrary, produced significantly less regularised maps, especially for RCP 2.6 (~ 27% of the territory was not regularised).

Concerning the σ reduction path, the variation in proportion of regularised surfaces and performance metrics is shown in Fig. 4. We found a monotone drop of both apparent and cross-validated performance along this path, together with a reduction in regularised surfaces until approximately $\sigma = 0.21$. In order to provide a visual comparison of the outputs of site index maps and delimited regularised areas along the σ reduction path, we also represented the predicted *SI* maps for $\sigma = 0.21$ in Figs. 5 and 7. Under this new value of σ, the regularised surfaces for all the climate conditions considered are reduced significantly. For current climate, the proportion of regularised territory dropped from 60% to 36%, while, for the case of future climate maps, it reached a minimum value of 48% (for the MPI-ESM-LR model under RCP 2.6). However, the other two future climate models were not that sensitive to the σ reduction, being its values of regularised surfaces always above 80% of the territory. Concerning the predictive performance of the model based on $\sigma = 0.21$, the apparent R^2 dropped from 0.56 to 0.47, while the cross-validation R^2 dropped from 0.38 to 0.34.

4. Discussion

The fitted SVR *SI*-climate models explained from 20% to 56% of the response's variability (Table 2), which is in the range of previous studies of site index modelling (e.g., 24–64% in Seynave et al. (2005); 20–52% in Aertsen et al. (2010); 32–52% in González-Rodríguez and Diéguez-Aranda (2020)). However, the number of support vectors was significantly high across the different models, accounting for 75–95% of the total number of observations. Admittedly, a high number of support vectors implies that the model is balanced towards a *high bias-low variance* fitting strategy (James et al., 2013) which, in our case, might be an indicator of overfitting. Even so, the CV procedure carried out showed adequate values of performance for the models fitted ($R^2_{CV} \sim$ 18–38%), especially for radial basis model, which still remain within the range of the *SI*-climate models found in literature. According to the model performance analysis carried out, SVR_{RB} was the best alternative, providing both the highest apparent and CV R^2. Interestingly, this model also included a lesser amount of bioclimatic predictors (*bio*1, *bio*2, *bio*3, and *bio*6) than other models (e.g., the third degree polynomial).

The four predictors included in the radial basis model are temperature-related variables. The absence of rainfall-related predictors in the model may seem contrary to previous studies about radiata pine's productivity (Hunter and Gibson, 1984). However, Romanyà and Vallejo (2004) already found that *SI* for this species was not significantly sensitive to precipitation in the Atlantic areas of Spain. The analysis of the correlation image in Fig. 3 revealed, overall, a sparse and non-uniform relationship between predictors and estimated site index. This is an expected result from the used interpretation method when substantial non-linear response-predictors relationships occur (Üstün et al., 2007). In the case of *bio*1, we hypothesise that this irregular role might be due to the ambiguous relationship between this variable and other temperature variables responsible for stress factors with diverging effects on growth (i.e., frost stress or heat stress variables). The predictors *bio*2 and, to a lesser extent, *bio*6 showed a somewhat clear monotone trend in correlations with sorted support vectors (increasing and decreasing with estimated site index, respectively). In contrast, *bio*3 had its maximum near the center of the 1,2,…,N_{sv} range; showing negative correlation values towards the extremes of this range. Interestingly, the negative role of *bio*6 in high quality sites is consistent with our results in a previous article (González-Rodríguez and Diéguez-Aranda, 2020), where we found a negative influence of *mtcm* (mean temperature of

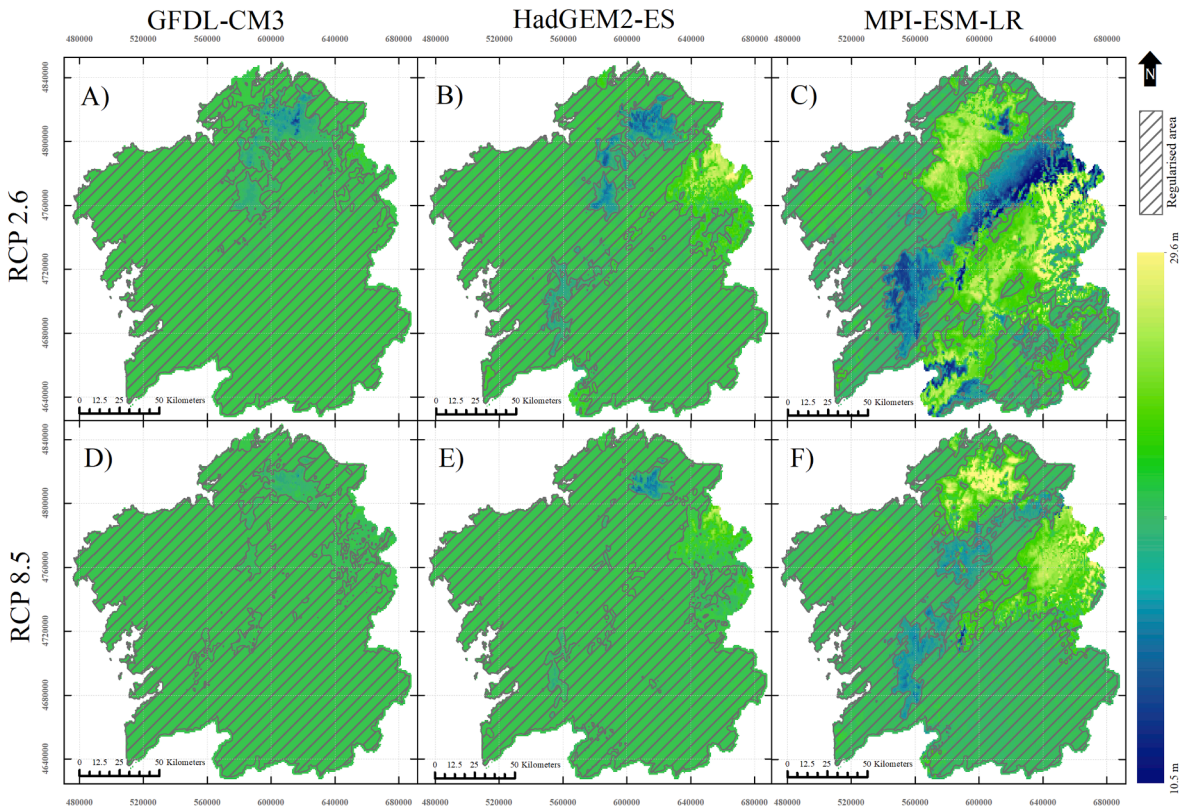


Fig. 7. Map of *SI* predictions using the radial basis model and $\sigma = 0.21$ for the six climatic scenarios considered. The upper row -A, B and C plots- corresponds to the RCP 2.6, while the lower one -plots D, E and F- corresponds to RCP 8.5. The columns correspond to GFDL-CM3 (left), HadGEM2-ES (center) and MPI-ESM-LR (right) GCM models. The delimited regularised areas are represented with a striped pattern.

coldest month) on radiata pine *SI*. In that case, we hypothesised that this could be due to the *chilling effect*, meaning that very warm winter temperatures may lead to stress on carbon balance (Smith et al., 2013, p. 116) due to high respiration rates, which has been observed in other pine species (Garber, 1983; Valkonen et al., 1990; Wu et al., 2001). The performance of *bio*2 as main driver of productivity for high *SI* values could be analogous to the role of *bio*6, implying that high diurnal ranges could reduce the stress on carbon balance related to nightly respiration, specially during the warmest months. Though the importance of night respiration on forest growth has been acknowledged in many physiological studies (Ryan, 1991; Ryan et al., 1997), we have not found any reference to this topic in studies regarding *SI* modelling. The main drawback of this interpretation could be the unexplained noise of heat stress variables (i.e., maximum temperatures in summer), necessarily correlated with *bio*6, that should affect growth negatively. However, we think that this finding may be a specific feature of the geographic extent of our study, where the humid and temperate climatic conditions (mainly Csb climate, with Csa, Cfa and Cfb local variants, according to the Köppen-Geiger classification updated by Kottek et al. (2006)) may make the high temperatures and drought-related factors in warm seasons a not usual constraint for growth. The role of *bio*3 is related to the role of *bio*2 (*bio*3 is actually derived from *bio*2), as it is reflected in the correlation image for low ranges of $\widehat{SI}$. In the alternative case, for high values of predicted *SI*, the negative influence of *bio*3 on growth may be associated with the effect of continentality (*bio*7, another Worldclim variable used for computing *bio*3), which was found an important restriction for pine growth in central Spain in a previous study (Büntgen et al., 2013).

Concerning the site index maps developed, the radial basis model produced predictions that were slightly more varied than the observed values under some conditions (mostly, current climate). This feature is interesting as it means that the model allowed for extrapolating outside of the observed range, in contrast to other tested approaches in previous studies, such as rule-based models (Weiskittel et al., 2011; Sabatia and Burkhart, 2014; Barrio-Anta et al., 2020), where predictions were constrained to the observed range or varied in a much lesser extent. Because of this, we believe that support vector regression might be considered a preferable technique to rule-based methods in many circumstances, especially when extrapolations are necessary. Besides, across the different predicted maps, highly regularised areas were mostly concentrated in the Atlantic coastal zone (very temperate and humid) and the southeastern mountainous ranges (with strong mediterranean and/or alpine influence), which is consistent with the known climatic differences between those areas and the extent encompassed by the measured research plots (Csb climate, predominantly). We believe that this fact proves the climatic coherency of the distance-based regularisation approach for delimiting homogeneous areas in terms of potential extrapolation errors.

The high proportion of regularised areas estimated for the produced maps under the σ_{CV} setup, revealed that, overall, the existing climatic conditions in most of these territories (and also across future scenarios) are too different from the observed in the training dataset for being subject of reliable site index predictions. However, the high specificity of this model, optimum in terms of cross-validation performance, can be effectively mitigated by reducing σ and, thus, obtaining a more generalisable model, able to make predictions over a broader range of climatic conditions (as shown in Figs. 5–7). Nevertheless, the outcome of this specificity-generality trade-off strongly depends on the chosen value

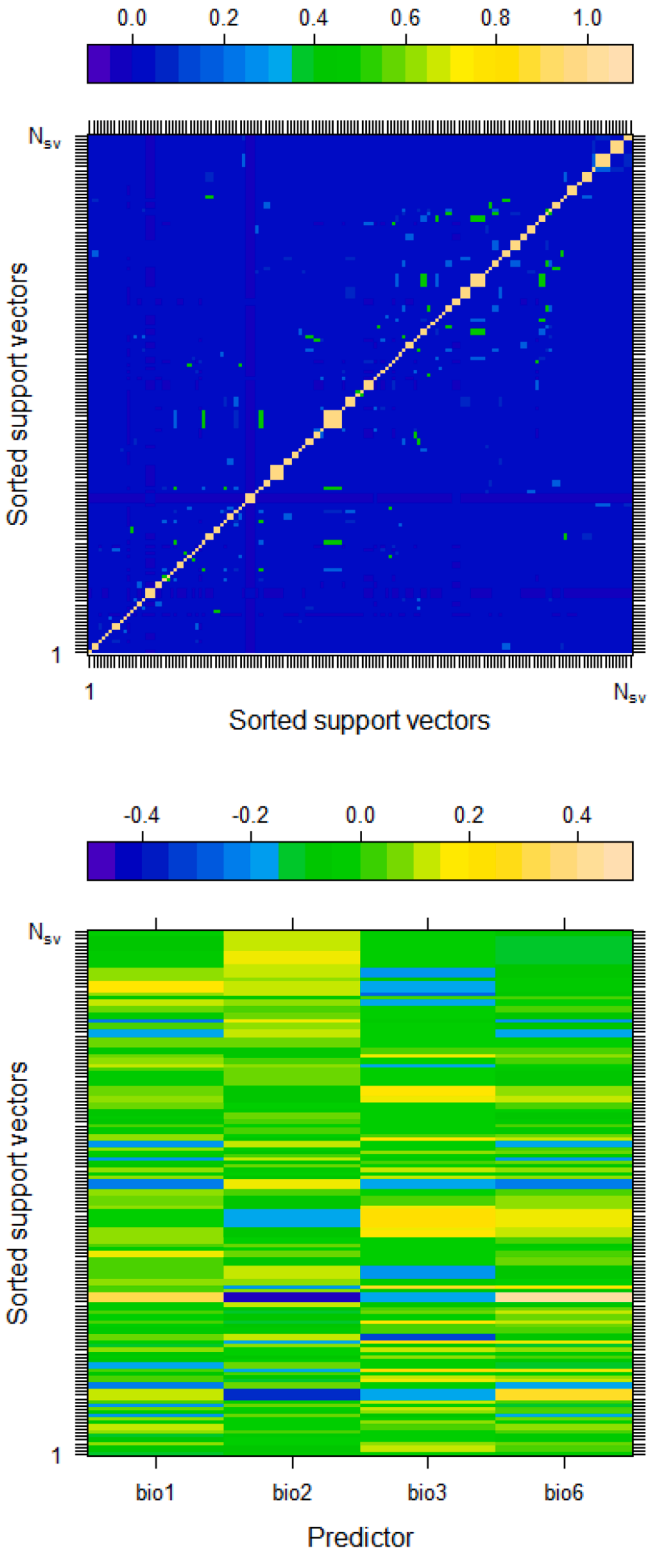


Fig. 3. Rebuilt kernel matrix (top) and Correlation Image (bottom) of the SVR_{RB} model, using the n_{sv} support vectors and the four selected bioclimatic predictors. The n_{sv} observations are sorted in increasing order of $\widehat{SI}$.

of σ, which, in our perspective, should be determined by the specific goals of the modeller. In this sense, the simplest criterion for calibrating σ could be to try to maximise the validity area of the model (i.e., to minimise the regularised areas) subject to meet a minimum performance requirement. Another alternative could be to maximise the validity area (and, hence, the dispersion of predictions) subject to keep the predictive range within reasonable bounds (for instance, by restricting predictions to maximum and minimum values registered in previous studies). Further research will be necessary for evaluating the practicality of these criteria.

Another aspect to discuss could be the adequacy of the Euclidean norm implicit in the radial basis kernel for capturing differences between climatic conditions. Other distance based-kernels have proven useful in the last decades for performing the $k(x, x')$ transformation in a variety of modelling tasks, e.g., the Laplacian kernel, the Bessel kernel, and the Pearson VII Universal kernel (Schölkopf, 2002; Cristianini and Shawe-Taylor, 2014). Testing these alternatives may be crucial development lines for improving the performance of SVR models in this subject area.

In this study, support vector regression and, in particular, the radial basis kernel variant, proved to be a useful technique for modelling site index as a function of climate. At the same time, its distance-based regularisation strategy provided an effective criterion for delimiting the actual validity area of the model, discarding all the areas with too different climatic conditions from the training dataset, and for which predictive errors could be too uncertain. We think this is an adequate starting point for assessing the spatio-temporal uncertainty of climate-sensitive forest productivity predictions, which could be also extendable to the prediction of other ecological indicators. Additionally, this approach could be a productive grounding for future research concerning climate change-adapted forest management.

5. Conclusions

In this article, we used a Support Vector Regression technique based on the radial basis kernel for predicting site index of radiata pine stands in Spain as a function of climate. The model provided an adequate performance, explaining up to 56% of the response's variability. We also found the growth-climate relationships represented in this model ecologically reasonable.

The distance-based regularisation strategy implemented in the model also allowed for delimiting the validity area of the model across different climatic scenarios. Under the optimum model setup (in terms of cross-validation performance), the climatic conditions on most of the Galician territory were considered too different from the observed in the training dataset, and hence yielded extremely regularised predictions. As a consequence, the validity area of this model delimited for these scenarios was notably narrow, especially for some future climate projections (Representative Concentration Pathway 8.5). However, reducing the σ parameter allowed for mitigating regularisation significantly, providing predictions with broader delimited validity areas, at the expense of having a lower performance. Specifically, a value of $\sigma = 0.21$, yielded a validity area of nearly 64% of the territory for current climate, while explaining up to 47% of the site index variability.

In conclusion, the proposed method allowed for successfully delimiting validity areas for the climate-sensitive site index model developed, and identified domains were predictions could be potentially too uncertain due to unseen climatic conditions. We believe that the tested methodology could be a useful starting point for understanding and delimiting the spatio-temporal uncertainty of forest productivity projections in future research.

CRediT authorship contribution statement

M.A.González-Rodríguez: Conceptualization, Formal analysis, Funding acquisition, Project administration, Software, Writing - review & editing, Investigation, Methodology, Validation, Visualization, Writing - original draft. **U. Diéguez-Aranda:** Conceptualization, Formal analysis, Funding acquisition, Project administration, Software, Writing - review & editing, Resources, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The work of the first author and main researcher of this study has been partially funded by the Spanish Ministry of Science, Innovation and Universities (DI-16–08971) and by the forest management consultancy company CERNA Ingeniería y Asesoría Medioambiental S.L. Plot data collection was carried out in the frame of two research projects (AGL2008-02259 and AGL2001-3871-C02-01) funded by the Spanish Ministry of Science and Innovation, the Spanish Interdepartmental Commission of Science and Technology and the European Commission (European Regional Development Fund).

Appendix A

Table 4
Description and statistics of the Worldclim 2 current bioclimatic variables corresponding to the locations of the 165 radiata pine plots used in this study.

Label	Variable name	Mean	St. dev.	Min	Max
*bio*1	Annual Mean Temperature (°C)	11.9	0.532	10.3	13.3
*bio*2	Mean Diurnal Range (°C)	8.61	0.5203	6.6	9.3
*bio*3	Isothermality	4.02	0.0676	3.8	4.1
*bio*4	Temperature Seasonality (100 × °C)	435.4	351	337.9	492.9
*bio*5	Max Temperature of Warmest Month (daily mean) (°C)	23.9	0.809	22.4	25.8
*bio*6	Min Temperature of Coldest Month (daily mean) (°C)	2.59	0.91	−0.5	6
*bio*7	Temperature Annual Range (°C)	21.3	1.46	16.6	23.5
*bio*8	Mean Temperature of Wettest Quarter (°C)	7.13	0.796	5.2	10.1
*bio*9	Mean Temperature of Driest Quarter (°C)	17.6	0.624	16.2	19.1
*bio*10	Mean Temperature of Warmest Quarter (°C)	17.8	0.585	16.4	19.2
*bio*11	Mean Temperature of Coldest Quarter (°C)	6.73	0.761	4.8	9.2
*bio*12	Annual Precipitation (mm)	1102	90	946	1318
*bio*13	Precipitation of Wettest Month (mm)	148	14.5	123	184
*bio*14	Precipitation of Driest Month (mm)	33	3.57	25	40
*bio*15	Precipitation Seasonality (mm)	38.8	1.95	35.3	43
*bio*16	Precipitation of Wettest Quarter (mm)	408	36.1	350	505
*bio*17	Precipitation of Driest Quarter (mm)	127	11.6	103	152
*bio*18	Precipitation of Warmest Quarter (mm)	141	13.4	109	169
*bio*19	Precipitation of Coldest Quarter (mm)	384	35.4	321	471

Note: *bio*2 is the mean difference between maximum and minimum monthly temperatures. *bio*3 is equal to $100\frac{bio2}{bio7}$, while *bio*4 is equal $100s_{t_{day}}$, with $s_{t_{day}}$ being the standard deviation of daily temperatures in a year. *bio*7 is the difference between *bio*5 and *bio*6. *bio*15 is the standard deviation of daily rainfall in a year.

References

Aertsen, W., Kint, V., van Orshoven, J., Özkan, K., Muys, B., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecol. Model.

Barrio-Anta, M., Castedo-Dorado, F., Cámara-Obregón, A., López-Sánchez, C.A., 2020. Predicting current and future suitable habitat and productivity for Atlantic populations of maritime pine (Pinus pinaster Aiton) in Spain. Ann. Forest Sci. 77 (2), 41.

Bivand, R. and Rundel, C. (2019). rgeos: Interface to Geometry Engine - Open Source ('GEOS').

Bontemps, J.D., Bouriaud, O., 2014. Predictive approaches to forest site productivity: Recent trends, challenges and future perspectives. Forestry 87 (1), 109–128.

Büntgen, U., Martínez-Peña, F., Aldea, J., Rigling, A., Fischer, E.M., Camarero, J.J., Hayes, M.J., Fatton, V., Egli, S., 2013. Declining pine growth in Central Spain coincides with increasing diurnal temperature range since the 1970s. Global Planet. Change 107, 177–185.

Burns, C., Thomason, J., and Tansey, W. (2019). Interpreting Black Box Models via Hypothesis Testing.

Caputo, B., Sim, K., Furesjo, F., and Smola, A. (2002). Appearance-based object recognition using SVMs: which kernel should I use? In Proc of NIPS workshop on Statitsical methods for computational experiments in visual processing and computer vision. Whistler.

Castedo-Dorado, F., Diéguez-Aranda, U., Álvarez-González, J.G., 2007. A growth model for Pinus radiata D. Don stands in north-western Spain. Ann. Forest Sci. 64 (4), 453–465.

Collins, W.J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C.D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S. (2011). Development and evaluation of an Earth-System model - HadGEM2. Geoscientific Model Development.

Cortez, P. (2016). rminer: Data Mining Classification and Regression Methods.

Cristianini, N. and Shawe-Taylor, J. (2014). Support Vector Machines and other kernel-based learning methods.

Diéguez-Aranda, U., Burkhart, H.E., Rodríguez-Soalleiro, R., 2005. Modeling dominant height growth of radiata pine (Pinus radiata D. Don) plantations in north-western Spain. In: For. Ecol. Manage..

Donner, L.J., Wyman, B.L., Hemler, R.S., Horowitz, L.W., Ming, Y., Zhao, M., Golaz, J.C., Ginoux, P., Lin, S.J., Schwarzkopf, M.D., Austin, J., Alaka, G., Cooke, W.F., Delworth, T.L., Freidenreich, S.M., Gordon, C.T., Griffies, S.M., Held, I.M., Hurlin, W.J., Klein, S.A., Knutson, T.R., Langenhorst, A.R., Lee, H.C., Lin, Y., Magi, B. I., Malyshev, S.L., Milly, P.C., Naik, V., Nath, M.J., Pincus, R., Ploshay, J.J., Ramaswamy, V., Seman, C.J., Shevliakova, E., Sirutis, J.J., Stern, W.F., Stouffer, R.J., Wilson, R.J., Winton, M., Wittenberg, A.T., Zeng, F., 2011. The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. J. Clim.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. (2013). Evaluation of climate models. In Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Alle, pages 741–866. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fontes, L., Bontemps, J.D., Bugmann, H., Van Oijen, M., Gracia, C., Kramer, K., Lindner, M., Rotzer, T., Skovsgaard, J.P., 2010. Models for supporting forest management in a changing environment. Forest Systems 19, 8–29.

Garber, M.P., 1983. Effects of chilling and photoperiod on dormancy release of container-grown loblolly pine seedlings. Can. J. For. Res. 13 (6), 1265–1270.

Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K.D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., Stevens, B., 2013. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. J. Adv. Modeling Earth Syst. 5 (3), 572–597.

González-Rodríguez, M., Diéguez-Aranda, U., 2020. Exploring the use of learning techniques for relating the site index of radiata pine stands with climate, soil and physiography. For. Ecol. Manage. 458.

Hijmans, R.J. (2019). raster: Geographic Data Analysis and Modeling.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol.

Hijmans, R.J., Phillips, S., Leathwick, J., and Elith, J. (2017). dismo: Species Distribution Modeling.

Hlásny, T., Trombik, J., Bošeła, M., Merganič, J., Marušák, R., Šebeň, V., Štěpánek, P., Kubišta, J., Trnka, M., 2017. Climatic drivers of forest productivity in Central Europe. Agric. For. Meteorol. 234–235, 258–273.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics.

Hunter, I.R., Gibson, A.R., 1984. Predicting Pinus radiata site index from environmental variables. NZ J. Forest. Sci. 14 (1), 53–64.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An Introduction to Statistical Learning, volume 103 of Springer Texts in Statistics. Springer, New York, New York, NY.

Jiang, H., Radtke, P.J., Weiskittel, A.R., Coulston, J.W., Guertin, P.J., 2015. Climate- and soil-based models of site productivity in eastern US tree species. Can. J. For. Res. 45 (3), 325–342.

Kahle, H.-P., Karjalainen, T., Schuck, A., Ågren, G.I., Kellomäki, S., Mellert, K.H., Prietzel, J., Rehfuess, K.E., and Spiecker, H. (2008). Causes and Consequences of Forest Growth Trends in Europe - Results of the RECOGNITION Project.

Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab – An S4 Package for Kernel Methods in R. J. Stat. Softw. 11 (9), 1–20.

Karush, W., 1939. Minima of functions of several variables with inequalities as side constraints. Master's thesis. Univ. of Chicago.

Keerthi, S.S., Shevade, S.K., Bhattacharyya, C., Murthy, K.R.K., 2001. Improvements to Platt's SMO algorithm for SVM classifier design. Neural Comput.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World Map of Köppen-Geiger Climate Classification - (updated with CRU TS 2.1 temperature and VASClimO v1.1 precipitation data 1951 to 2000). Meteorologische Zeitschrift.

Kuhn, H.W., Tucker, W., 1951. Nonlinear programming. In: Proc. 2nd Berkeley Symposium on Mathematical Statistics and Probabilistics. University of California Press, Berkeley, pp. 481–492.

Lindner, M., Fitzgerald, J.B., Zimmermann, N.E., Reyer, C., Delzon, S., van der Maaten, E., Schelhaas, M.J., Lasch, P., Eggers, J., van der Maaten-Theunissen, M., Suckow, F., Psomas, A., Poulter, B., Hanewinkel, M., 2014. Climate change and European forests: What do we know, what are the uncertainties, and what are the implications for forest management? J. Environ. Manage. 146, 69–83.

Monserud, R.A., Huang, S., Yang, Y., 2006. Predicting lodgepole pine site index from climatic parameters in Alberta. Forestry Chronicle 82 (4), 562–571.

Nothdurft, A., Wolf, T., Ringeler, A., Böhner, J., Saborowski, J., 2012. Spatio-temporal prediction of site index based on forest inventories and climate change scenarios. For. Ecol. Manage. 279, 97–111.

Parresol, B.R., Scott, D.A., Zarnoch, S.J., Edwards, L.A., Blake, J.I., 2017. Modeling forest site productivity using mapped geospatial attributes within a South Carolina Landscape, USA. For. Ecol. Manage. 406 (September), 196–207.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reichler, T., Kim, J., 2008. How well do coupled models simulate today's climate? Bull. Am. Meteorol. Soc. 89 (3), 303–311.

Romanyà, J., Vallejo, V.R., 2004. Productivity of Pinus radiata plantations in Spain in response to climate and soil. For. Ecol. Manage.

Ryan, M.G., 1991. Effects of climate change on plant respiration. Ecological. Applications.

Ryan, M.G., Lavigne, M.B., Gower, S.T., 1997. Annual carbon cost of autotrophic respiration in boreal forest ecosystems in relation to species and climate. Journal of Geophysical Research Atmospheres 102 (24), 28871–28883.

Sabatia, C.O., Burkhart, H.E., 2014. Predicting site index of plantation loblolly pine from biophysical variables. For. Ecol. Manage. 326, 142–156.

Schölkopf, B., 2002. Learning with kernels. In: In Proceedings of 2002 International Conference on Machine Learning and Cybernetics.

Seynave, I., Gégout, J.-C., Hervé, J.-C., Dhôte, J.-F., Drapier, J., Bruno, É., Dumé, G., 2005. Picea abies site index prediction by environmental factors and understorey vegetation: a two-scale approach based on survey databases. Can. J. For. Res. 35 (7), 1669–1678.

Shen, C., Lei, X., Liu, H., Wang, L., Liang, W., 2015. Potential impacts of regional climate change on site productivity of Larix olgensis plantations in northeast China. IForest 8 (OCTOBER2015), 642–651.

Skovsgaard, J.P., Vanclay, J.K., 2008. Forest site productivity: A review of the evolution of dendrometric concepts for even-aged stands. Forestry 81 (1), 13–31.

Smith, W.K., Roy, J., and Hinckley, T.M. (2013). Ecophysiology of Coniferous Forests.

Strimas-Mackey, M., 2020. smoothr: Smooth and Tidy Spatial Features.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498.

Üstün, B., Melssen, W.J., and Buydens, L.M.C. (2007). Visualisation and interpretation of Support Vector Regression models. Analytica Chimica Acta, 595(1-2 SPEC. ISS.): 299–309.

Valkonen, M.-L., Hänninen, H., Pelkonen, P., and Repo, T. (1990). Frost hardiness of Scots pine seedlings during dormancy. Silva Fennica. 1990. 24(4): 335–340, 24.

Vapnik, V., Golowich, S., Smola, A., 1997. Support vector method for function approximation, regression estimation, and signal processing. In: Mozer, M., Jordan, M., Petsche, T. (Eds.), Advances in Neural Information Processing Systems 9. MIT Press, Cambridge, MA, pp. 281–287.

Watt, M.S., Palmer, D.J., Leonardo, E.M.C., Bombrun, M., 2021. Use of advanced modelling methods to estimate radiata pine productivity indices. For. Ecol. Manage. 479, 118557.

Weiskittel, A.R., Crookston, N.L., Radtke, P.J., 2011. Linking climate, gross primary productivity, and site index across forests of the western United States. Can. J. For. Res. 41 (8), 1710–1721.

Wu, L., Hallgren, S.W., Ferris, D.M., Conway, K.E., 2001. Effects of moist chilling and solid matrix priming on germination of loblolly pine (Pinus taeda L.) seeds. New Forest. 21 (1), 1–16.