ORIGINAL PAPER

Combining remote sensing, habitat suitability models and cellular automata to model the spread of the invasive shrub *Ulex europaeus*

Tobias Gränzig · **Anne Clasen** ·
Fabian Ewald Fassnacht · **Anna Cord** ·
Michael Förster

Abstract Modeling the past or future spread patterns of invasive plant species is challenging and in an ideal case requires multi-temporal and spatially explicit data on the occurrences of the target species as well as information on the habitat suitability of the areas at risk of being invaded. Most studies either focus on modeling the habitat suitability of a given area for an invasive species or try to model the spreading behavior of an invasive species based on temporally or spatially limited occurrence data and some environmental variables. Here we suggest a workflow that combines habitat suitability maps, occurrence data from multiple time steps collected from remote sensing data, and cellular automata models to first reconstruct the spreading patterns of the invasive shrub *Ulex europaeus* on the island Chiloé in Chile and then make predictions for the future spread of the species. First, *U. europaeus* occurrences are derived for four time steps between 1988 and 2020 using remote sensing data and a supervised classification. The resulting occurrence data is combined with occurrence data of the native range of *U. europaeus* from the GBIF database and selected environmental variables to derive habitat suitability maps using Maxent. Then, cellular automata models are calibrated using the occurrence estimates of the four time steps, the suitability map, and some additional geo-layer containing information about soils and human infrastructure. Finally, a set of calibrated cellular automata models are used to predict the potential spread of *U. europaeus* for the years 2070 and 2100 using climate scenarios. All individual steps of the workflow where reference data was available led to sufficient results (supervised classifications Overall Accuracy > 0.97; Maxent AUC > 0.85; cellular automata Balanced Accuracy > 0.91) and the spatial patterns of the derived maps matched the experiences collected during the field surveys. Our model predictions suggest a continuous expansion of the maximal potential range of *U. europaeus*, particularly in the

T. Gränzig · A. Clasen · M. Förster (✉)
Geoinformation for Environmental Planning Lab,
Technische Universität Berlin, Straße des 17. Juni 145,
10623 Berlin, Germany
e-mail: michael.foerster@tu-berlin.de

A. Clasen
Department of Forest Planning/Forest Research/
Information Systems, Landesforstanstalt Mecklenburg-
Vorpommern, Zeppelinstraße 3, 19061 Schwerin,
Germany

F. E. Fassnacht
Institute of Geography and Geoecology (IfGG),
Karlsruhe Institute of Technology (KIT), Kaiserstr. 12,
76131 Karlsruhe, Germany

F. E. Fassnacht
Remote Sensing and Geoinformatics, Institute
of Geographical Sciences, Freie Universität Berlin,
Malteserstr. 74-100, 12249 Berlin, Germany

A. Cord
Computational Landscape Ecology, Institute of Geography,
Dresden University of Technology, Helmholtzstr. 10,
01069 Dresden, Germany

Eastern and Northern part of Chiloé Island. We deem the suggested workflow to be a good solution to combine the static habitat suitability information—representing the environmental constraints—with a temporally and spatially dynamic model representing the actual spreading behavior of the invasive species. The obtained understanding of spreading patterns and the information on areas identified to have a high invasion probability in the future can support land managers to plan prevention and mitigation measures.

Keywords Chile · Landsat · Google earth · GBIF · Maxent · Cellular automata

Introduction

Understanding the spread dynamics of invasive alien plants (IAP) requires spatially explicit occurrence data of the IAP species at multiple time steps. Collecting such occurrence data in the field is time- consuming and often restricted by the accessibility of sites. This can lead to a spatial collection bias with a higher proportion of occurrence records along roads (Bradley 2014). Further, depending on the dispersal characteristics of the target species and the frequency and extent of environmental disturbances, invasion processes may last several years or decades. Such time periods are rarely covered in field records.

For some IAP, multi-temporal remote sensing data can be a promising alternative data-source to reconstruct, model and predict invasion dynamics over longer time periods. However, the successful assessment of invasions patterns from current and historic remote sensing data depends on some requirements. Most importantly, the targeted IAP has to be identifiable in the remote sensing data. This requires that the size of individuals or stands of the target species is sufficiently large to cause a clear spectral, textural or structural signal in the remote sensing data set. This signal also has to be unique enough to be separable from co-occurring native plant species (Müllerová et al. 2017; Bradley 2014; Huang and Asner 2009).

Such a unique spectral signature of an IAP can for example be caused by physiological differences between native species and the IAP that result in distinct optical traits (different pigment levels, different morphology, etc.). A unique spectral signal of an IAP can also be phenology-related. For example an IAP may differ from native species in their date of leaf emergence, senescence initialization or other phenological stages. Such phenological differences can in some cases be detected with remote sensing and were exploited to successfully map invasive species in several vegetation communities (e.g., Hoyos et al. 2010; Somers and Asner 2012). For some IAP flowering events may also facilitate their remote detection (e.g., Somodi et al. 2012). Related discussions about the importance of specific optical traits as well as spatial, spectral and temporal resolution of remote sensing data for identifying IAP can be found in several reviews (Bradley 2014; Huang and Asner 2009; Rocchini et al. 2015).

One key challenge for identifying occurrences of a target species in historic remote sensing data is the often missing corresponding reference data. However, the increasing number of historic high-resolution data sets, e.g. from orthophotos and very high spatial resolution satellite image-archives, may provide information that can replace field-collected reference data for some IAP. This has been demonstrated in some studies that successfully reconstructed invasion dynamics of IAP from multi-temporal remote sensing datasets and also incorporated high resolution imagery as validation data. For example, Liu et al. (2016) applied a combination of digital orthophotos and Landsat satellite images to successfully reconstruct the invasion dynamics of *Phragmites* along the Detroit river across a time period of 11 years.

However, most earlier studies also had field data available for at least some years (e.g., Espinar et al. 2015; Mao et al. 2019; Ren et al. 2019; Wang et al. 2015).

One of the most popular sources of freely available high-resolution and multi-temporal remote sensing data is the Google Earth Pro (GEP) platform. Since its launch in 2005, the amount of available data (e.g., from DigitalGlobe, NASA, ESA, GeoEye-1, IKONOS,,SPOT etc.) and their quality have steadily increased (Lesiv et al. 2018). However, because the data in GEP cannot be accessed in its original format, many studies use the data only for visual interpretation, to verify their results, or to delineate reference data (Yu and Gong 2012; Visser et al. 2014).

These studies studies showed that it is possible to map the invasion state of some IAP across several time steps from historic remote sensing data, even if some time steps lack field reference data. Some

studies also model the invasion dynamics between the reconstructed invasion state maps and provide estimates for the future spread of the targeted IAP. Cellular automata (CA) models are often used for the prediction step. Cellular automata are well-established in invasion ecology and ecology in general. Their ease of implementation in combination with their ability to simulate spatial patterns make them suitable to address many ecological questions which contain a spatial dimension (Breckling et al. 2011). Good introductions to CA models in the context of ecology are for example given in Breckling et al. (2011), Dunkerley (1999), and Ghosh et al. (2017). The basic structure of a CA model consists of a grid with many individual cells in a defined state. The state of a cell can change over time depending on its properties and its neighborhood (adjacent pixels). Whether a cell changes its state or not depends on a set of transitional rules that is defined by the modeler, either by using expert knowledge only or by combining the latter with calibration data.

Cellular automata models have been applied successfully to depict and predict invasions pattern of IAP (e.g., Barbosa et al. 2018; Merow et al. 2011) including also examples where the CA model based on remotely sensed occurrence maps of IAP. For example, Barona and Mena (2014) used Landsat and SPOT imagery to map an invasive shrubland species (*Psidium guajava*) on the Galapagos Islands at three time steps covering the period between 1980 and 2009. Using these maps, they trained a CA-Markov chains GEOMOD model to predict the spread of the species for the year 2030 based on the predictor variables humidity and altitude. Huang et al. (2008) followed a similar approach and mapped the spread of a grassland IAP (*spartina alterniflora*) in a Marshland environment by initiating a CA model with a remote sensing-based land-cover classification map from the start of the considered time period. Then, they used yearly land-cover classification maps for a 6 years time period to validate the CA model and the established rule-set. Using the same data set as Huang et al. (2008), Zheng et al. (2015) suggest that adding constraints to a CA model can improve the accuracy for modeling the spread of IAP. In their case, the additional consideration of environmental variables describing the wetland area notably improved the model accuracy in comparison to CA models exclusively based on neighborhood relations. Lu et al.

(2013) modeled the spread of an invasive weed (*Leucaena leucocephal*) in Taiwan with a CA model and remote sensing data. They first classified infested areas in SPOT satellite images for three time points between 1988 and 2007 including also high resolution imagery as reference data. Then, they trained a logistic regression model to explain the spread of the IAP between the first two time points with a set of environmental variables (e.g., elevation, slope, distance of IAP patches to roads, forest edges and farmlands). The probability maps of the logistic regression model were then fed into a CA model to successfully predict the spatial occurrence of the IAP at the third time-point.

In summary, the combination of IAP occurrence maps derived from remote sensing data and process-based spatially dynamic CA models has been successfully used in several studies to reconstruct the spreading of IAP and may also hold potential for estimating future invasion patterns. However, the number of studies following such an approach is still limited and so far, hardly any study attempted to predict their CA models to time-points in the more distant future. Moreover, the generally rule-based nature of CAs is in most cases still driven by expert knowledge, which is known to be biased by personal experience, accessibility of areas, as well as the often limited knowledge about the processes influencing the spreading of invasive species (Ghosh et al. 2017).

In contrast to CAs, correlative species distribution models (SDMs) have been used frequently to model the current and also the future potential distribution of IAPs without including potentially biased expert knowledge (Cao et al. 2021). However, correlative SDMs face some other limitations, partly related to the basic assumptions underlying correlative SDM. These include (1) that the distribution of a species is limited wholly or partly by aspects of climate and (2) that most species are in an equilibrium with climate, that is, the species occurs in all reachable and suitable habitats (Araújo and Peterson 2012). The latter is unrealistic as constraints related to dispersal as well as biotic interactions and human influence prevent many species from occurring in theoretically suitable habitats (Araújo and Peterson 2012). In the context of biological invasions, another assumption of correlative SDMs is that the niche observed for the species in its native range will be the same as in the new

region where it occurs as IAP (Wiens and Graham 2005; Wiens et al. 2010).

This so called niche conservatism is a crucial requirement if a SDM trained with species occurrence data from the native range is applied to a new area where a species acts as an IAP. However, niche shifts and also expansions have been observed for several IAP (Pena-Gomez et al. 2014). Solely relying on training data from the native range of the IAP may thus lead to an underestimation of the invasive potential of the species. Contrarily, training a SDM based on occurrence data only from the non-native range is very likely to violate the equilibrium assumption as the species in most cases did not have the time yet to reach all the areas in which it could in theory survive (Pili et al. 2020). Again there is a notable risk that the area of suitable habitat for the IAP will be underestimated. Despite the drawbacks of correlative SDMs, they have been widely applied as tools for predicting the potential distribution of IAP under current and future climate conditions and been proven useful for assessing invasions risks. Methodical developments, particularly approaches that combine multiple SDMs (Gallien et al. 2012) or combine SDMs with other model types, representing for example spreading properties of the IAP, may further improve the performance of SDM in the context of invasion ecology.

In this study we want to show the advantage of a combined model at the example of *U. europaeus*, the common gorse (Rees and Hill 2001). The native range of the shrub is the Atlantic coast of Western-Europe. After its introduction to other continents, it became invasive in numerous coastal regions worldwide, for example in North and South America and Australia (Hornoy et al. 2011). In Central-Chile *U. europaeus* is considered as a highly damaging IAP with negative effects on agriculture and plantation forestry (Norambuena et al. 2000) and on the unique endemic flora (Altamirano et al. 2016). Therefore, some studies have developed methods to map the recent spread of *U. europaeus* based on remote sensing data (e.g., Kattenborn et al. 2019; Gränzig et al. 2021) or modelled its invasion dynamics and distribution (e.g., Sorbe et al. 2023; Bateman and Vitousek 2018; Altamirano et al. 2016). However, no study to date has used historical and current presence data in combination with future climate data, as well as additional geo-data, to reconstruct and predict the invasion pattern of *U. europaeus*.

The objectives of this study are:

1. To develop a combined workflow of correlative SDMs and CAs that allows to reconstruct the historic spread of *U. europaeus* as observed from multi-temporal satellite data,
2. To use this retrospective data set to calibrate CA models predicting the future spread under different climate change scenarios and to assess the continued invasion risk of *U. europaeus* on Chiloé Island.
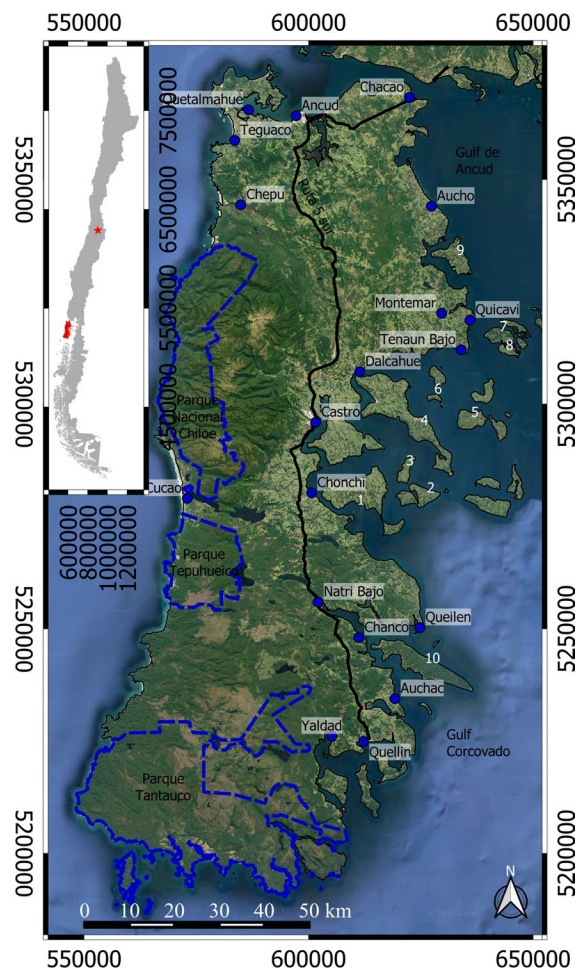
Fig. 1 Chile (The inset map with Chiloe Island is indicated in red. The red star indicates the capital of Chile, Santiago de Chile) and Chiloé Island including National parks (blue lines), Ruta 5 (black line) and settlements (blue points) from Open Street Map (OSM). The white numbers indicate the islands at the surrounding gulfs: 1 = Isla Lemuy, 2 = Isla Quehui, 3 = Isla Chelin, 4 = Isla Quinchao, 5 = Isla Quenac, 6 = Isla Llingua, 7 = Isla Mechuque, 8 = Isla Anihue, 9 = Isla Caucahue, 10 = Isla Tranqui

Data and methods

Study site and target species

This study focuses on Chiloé Island in south-central Chile (see Fig. 1). The local climate has a strong oceanic influence with a corresponding high air humidity, high precipitation rates and low temperature amplitudes (annual average rainfall of 2090 mm, mean annual temperature of 12 °C). Chiloé Island extends over an area of approx. 9180 km^2 and is mainly covered by native north-Patagonian rainforests (66.9%) and agricultural land (27.4%), with the most intensive agricultural use taking place in the north-eastern and central parts of the island (Barrena et al. 2014).

The north-eastern and the central part of Chiloé Island represent the main distribution area of *U. europaeus* today Gränzig et al. (2021). Once introduced as "natural" fence due to its dense and spiky foliage, it quickly spread into meadows, agricultural land, and plantations. This process was supported by the favourable climatic conditions on the island and the high adaptability of *U. europaeus* to disturbances. *U. europaeus* is a fast growing leguminous shrub which can be up to 7 m tall (but more often between 2 and 4 m). It shows a very unique yellow flower during spring (Clements et al. 2001). Seeds of *U. europaeus* are able to germinate after fire events. In addition, the species has a taproot system with lateral spreading capacities, which supports rapid re-sprouting after disturbances. Thus, removing *U. europaeus* after establishment is very expensive and difficult to undertake (Norambuena et al. 2000).

Data

Generally, the spread of *U. europaeus* is assumed to depend on various environmental factors which can be summarized by a range of geodatasets including (historic and future) climate data, relief and altitude-related parameters (sun azimuth angle), infrastructure data, as well as soil variables (see Table 1). For modeling the past and predicting the future invasion patterns we used environmental and remote sensing data for both, the native (Western Europe) and the invasive range (Chiloé Island) of *U. europaeus* following the recommendation of Clements et al. (2001). Furthermore, we included infrastructure and soil data for modeling the spreading patterns of *U. europaeus*.
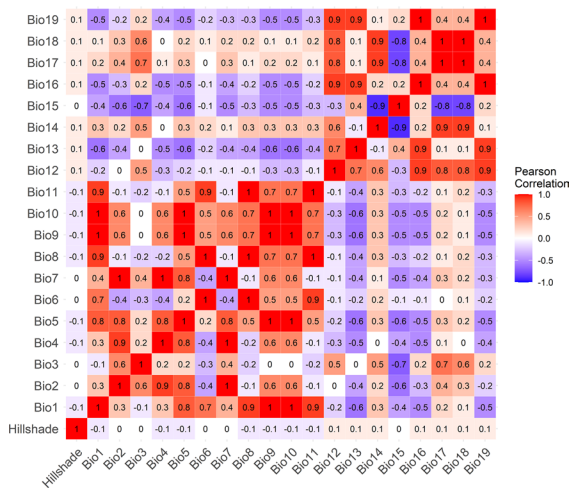
Environmental data

WorldClim (version 2.1) data (30 s resolution) was acquired for the study area of Chiloé Island (tiles 33 and 43) and for Western-Europe (Fick and Hijmans 2017). WorldClim data represents average climate data for the time-period 1970–2000 and includes 19 bioclimatic variables.

For predicting the future invasion patterns, general circulation models (GCM) were acquired from the Coupled Model Inter-comparison Project 5 (CMIP5, https://worldclim.org/data/v1.4/cmip5_30s.html). Each GCM dataset used in our study consists of the same bioclimatic variables as the Worldclim 2.1 data. We used CMIP5 climate data, because of its availability in 30 s resolution. The CMIP5 data were downloaded for two representative concentration pathways (RCP). RCP 45 is considered as an intermediate scenario, whereas the RCP 85 is the worst-case scenario with a continuous cumulative rise of the CO_2 emissions (IPCC 2014). The 30 s resolution CIMP5 data were available for two time periods 2050 (averaged for 2041–2060) and 2070 (averaged for 2061–2080) (Taylor et al. 2012). Although the WorldClim data was already provided as averaged values for both time-periods, we used them, but averaged the different models provided by WorldClim for each time-period. All GCMs (17 total, see Table 7) were then averaged for the 19 bioclimatic variables and for each RCP.

We used the TanDEM-X DEM standard product with a spatial resolution of 12 m and a vertical accuracy of 2 m (Esch et al. 2012) to derive topographic information for Chiloé Island. For Central-Europe, we used Shuttle Radar Topography Mission data (SRTM-3) with a 90 m resolution. From both DEMs, hillshade images were calculated taking into account aspect and slope as well as the solar azimuth angle at the highest zenith state.

From the total set of available environmental variables, we selected a subset representing the most important environmental requirements of *U. europaeus* in its native and non-native range. We furthermore ensured that the selected variables have low to medium inter-correlations (see Fig. 2). During this expert selection, the following earlier reported environmental requirements of *U. europaeus* were considered:

Fig. 2 Correlation between the environmental variables. A list of the abbreviations of the bioclimatic variables of Worldclim can be found in Table 8. Variables that had a correlation of less than ± 0.5 were used in this study

– The distribution of *U. europaeus* depends primarily on temperature (cannot survive very high and very low temperatures) (Zabkiewicz 1976).
– *Ulex europaeus* prefers maritime conditions with a distinct seasonality in precipitation (Clements et al. 2001).
– *Ulex europaeus* is a light demanding plant (intolerant against heavy shade) (Hackwell 1980).

Based on these assumptions and after dropping highly correlated variables, five variables (Hillshade, Annual Mean Temperature (Bio1), Temp. Annual Range (Bio7), Annual Precipitation (Bio12), Precipitation Variability (Bio15)) were selected.

To obtain a temporally consistent data-set, the four selected bioclimatic variables were linearly interpolated to 5-years intervals for the future climate conditions. Thus, the 5-years interpolation starts in 2000 based on the historical WorldClim 2.1 data, continues with the CMIP5 data for 2050 (10 interpolation steps), and ends in 2070 (mid-term forecast) (four interpolation steps between CMIP 5 for 2050 and 2070). Climate data from 2070 were constantly used for the long-term forecast (to 2100) because no climate forecasts beyond 2070 are available, as climate forecasts for such large time periods are very imprecise.

Additionally we used two geo-layers to account for abiotic and anthropogenic factors in the CA models. We used a soil-type data set for Chile from CIREN (Chilean Natural Resources Information Centre—status 2010–2012, (CIREN 2003)) to include soil suitability information. The data (scale of 1:10 000) was aggregated to eight suitability classes based on expert knowledge. Furthermore infrastructure data for Chiloé Island was downloaded from the OpenStreet-Map platform (OSM, wiki.openstreetmap.org). We focused on layers representing the road network as roads are known to be a primary propagation route for *U. europaeus* (Moss 1960). Each pixel within 100 m of the nearest traffic route was considered part of the road-affected area.

Figure 3 shows maps of the selected climate and terrain variables as well as of the additional geo-layer for soil and roads.
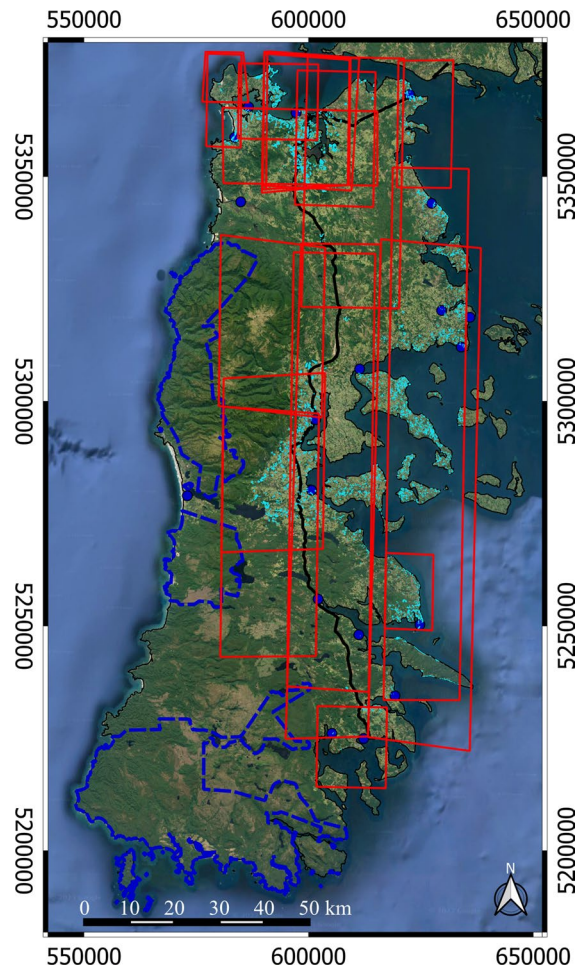
Remote sensing data

For the long-term calibration of the CA models, we used four Level 2A Landsat images with very low cloud cover acquired on October 8, 1988 (Landsat 4), October 23, 1999 (Landsat 7), October 27, 2015 and November 09, 2020 (both Landsat 8). All scenes were acquired during the flowering period of *U. europaeus*.

To train the classification algorithm for mapping *U. europaeus* occurrences in the Landsat scenes, high resolution RGB Google Earth Pro (GEP) imagery were used to identify retro-perspective and actual presence points. The first available high resolution image during the flowering phase was from 2009 and the latest from 2020. In total, 13 suitable dates were available. We intensively scanned these images for larger (at least the extent of one Landsat pixel) *U. europaeus* patches and digitized them. Patches from 2009 to 2015 were then merged for the 2015 Landsat image and from 2009 to 2020 for the 2020 Landsat image. We assume that once a site has been colonized by *U. europaeus*, this status remains unchanged even if *U. europaeus* has been removed. This was necessary due to the strong dynamic between removal and regrowth of *U. europaeus*. More than 4500 patches of *U. europaeus* were detected (See Fig. 4). Based on the aggregated layer for 2020, we traced these patches back to get presence points for the years 1999 and 1988 (See Fig. 5) via an onscreen backward analysis. For this step

(a) Hillshade   (b) Bio 01   (c) Bio 07

(d) Bio 12   (e) Bio 15   (f) Soil

(g) Roads
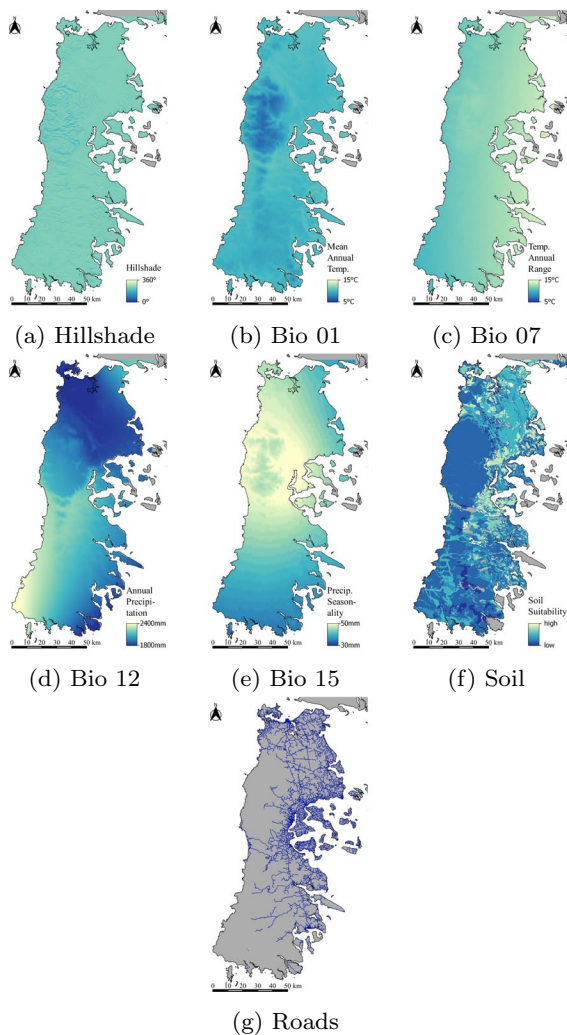
Fig. 3 Environmental and infrastructure variables. A list of the abbreviations of the bioclimatic variables of Worldclim can be found in Table 8 and in Table 1. A deeper blue indicates lower values for maps (**a–f**). The range of values is indicated in the legends inserted in Figures **a–f**


Fig. 4 Aggregated *U. europaeus* patches (cyan) detected in Google Earth Pro (GEP) in the time period 2009–2020. The red squares indicates the footprints of the high-resolution images available in GEP. The blue lines indicates the National parks, the black line indicates the Ruta 5 sur and the blue points indicate settlements (Source: Open Street Map)

we used the downloaded Landsat images. Already GEP-digitized patches which show the unique yellow flower of *U. europaeus* in the historic Landsat images before 2009 were considered as *U. europaeus* patch also in the backward time-steps where no high resolution reference data was available. If there was was no indication of a yellow flowering in the Landsat images, we omitted the patch. Within the patches which we assigned to the *U. europaeus* class for all Landsat images, 2000 random points

were selected. These occurrence points were then used as reference points for the classification of *U. europaeus* in " Mapping the occurrence of Ulex (Step 1)" Section.

Methods

The methodological workflow (see Fig. 6) consists of 3 major steps: (1) Mapping past (1988, 1999, 2015) and recent (2020) occurrences of *U. europaeus* based on multi-temporal remote sensing data and Random Forest classifications. (2) Modelling the past and
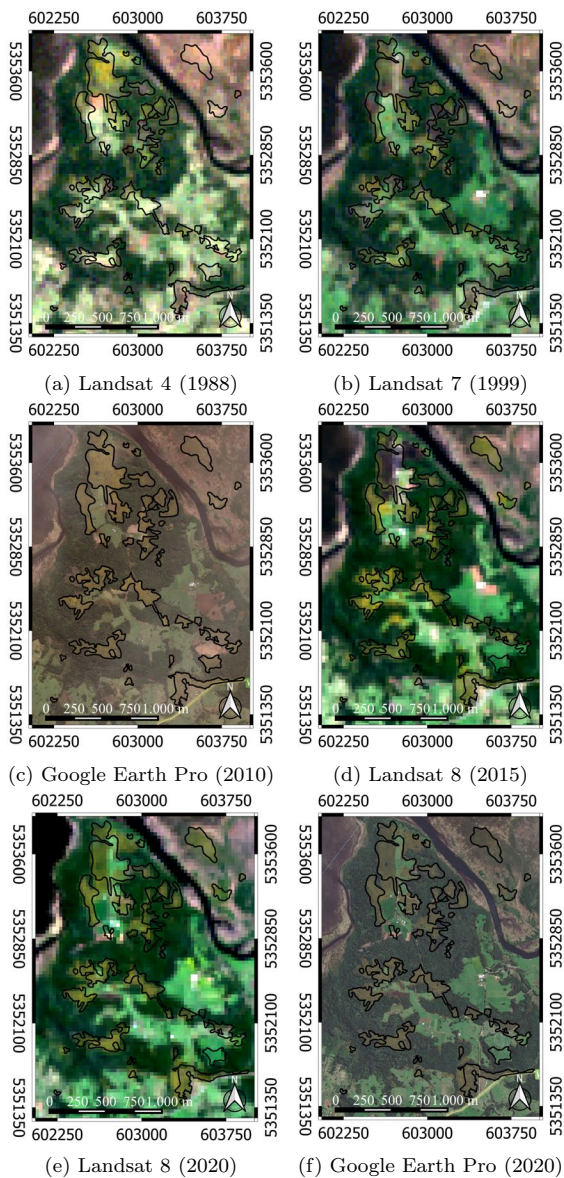
Fig. 5 Delineated *U. europeaus* patches based on the last available high-resolution Google Earth Pro image for this site from 2010 with background images: **a** Landsat 4 (year 1989), **b** Landsat 7 (year 1999), **c** Google Earth Pro (2010), **d** Landsat 8 (year 2015), **e** Landsat 8 (year 2020), **e** Google Earth Pro (2020)

predict the future habitat suitability for *U. europeaus* with the help of a SDM based on environmental variables. (3) Modelling the past (starting in 1988) and predict the future invasion pattern of *U. europaeus* with CA models based on occurrence maps, habitat suitability layers derived from the SDM and

additional geo-layers (roads and soil). Before predicting the future invasion pattern, the CA models were calibrated based on the past (1999 and 2015) and recent occurrence maps (2020) to establish a set of optimal transition rules. Based on this set of transition rules the invasion pattern is predicted for a mid-term period (until the year 2070) and a long-term period (until the year 2100) and finally the invasion probability is calculated. The latter prediction period (2100) is slightly longer than the period of the climate model (2070), because the spreading process might need more time to reach its maximum extend, especially in the southern part of the island even under no further changing climate conditions.

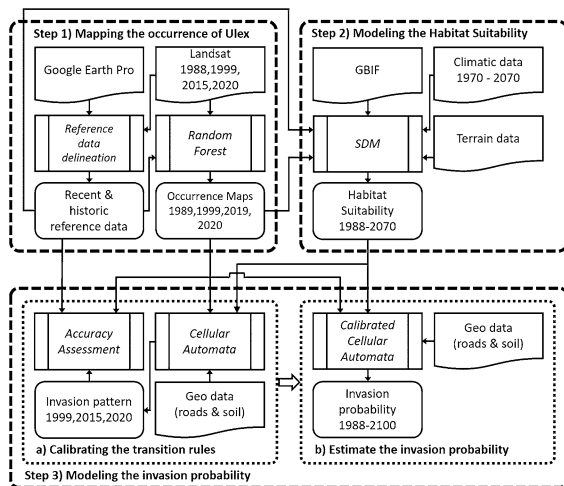*Mapping the occurrence of Ulex (Step 1)*

For the retrospective CA model calibration, four occurrence maps were prepared, representing the spatial distribution of *U. europaeus* on Chiloé Island in the years 1988, 1999, 2015 and 2020. The occurrences of *U. europaeus* were classified based on the Landsat images and the reference points described in the previous "Remote sensing data" Section . The presence points of each time-stamp were divided into 50% calibration and 50% validation for assessing the classification quality. The Random Forest algorithm (Breiman 2001) was used as a one-class-classifier to derive the probability values of potential *U. europaeus* occurrences. Optimal hyper-parameters (mtry = Randomly Selected Predictors, split-rule = Splitting Rule, min.node.size = Minimal Node Size) of the RF algorithm (Ranger algorithm of the caret package in R) were optimized with a grid-search approach based on a 5-fold cross-validation of the calibration points. The calibrated RF models were then applied to the entire Landsat images to estimate *U. europaeus* occurrence probability values. Estimated occurrence probabilities above 0.8 were considered as occurrences, and were compared to the validation points for calculating the Overall Accuracy (OA).

Based on the resulting occurrence maps, reference data-sets for the CA models were derived for each time-stamp. All classified *U. europaeus* patches in the occurrence map of 1988 were used as initial reference data. For the other three time stamps, the occurrence maps were aggregated for all previous time stamps, always starting in 1988. For the time-stamps in 2015 and 2020, the digitized *U. europaeus* patches,

Table 1 Utilized variables for the modelling process including their data-source and resolution. CI = Chiloé Island, WE = Western-Europe

| Region	Data-source	Variable	Res.
WE	GBIF	Occurance points Ulex	–
CI	Google earth pro	Reference data for LCLU + Ulex	–
WE/CI	WorldClim CMIP5	BIO1: Annual mean temperature	30 s
WE/CI	WorldClim CMIP5	BIO7: Temperature annual range (BIO5-BIO6)	30 s
WE/CI	WorldClim CMIP5	BIO12: Annual precipitation	30 s
WE/CI	WorldClim CMIP5	BIO15: Precipitation seasonality	30 s
WE	SRTM	Hillshade	90 m
CI	TandDEM-X	Hillshade	12 m
CI	NASA USGS	Landsat 4 (1988)	30 m
		Landsat 7 (1999)	
		Landsat 8 (2015)	
		Landsat 8 (2020)	
CI	CIREN	Soil quality (1–8)	–
CI	Open street Map	Roads	–

Fig. 6 Workflow of the study. The displayed steps are described in "Mapping the occurrence of Ulex (Step 1)" Section for Step 1, "Remote sensing data" Section for Step 2, "Modelling the Invasion probability with transition rules (Step 3a)" Section for Step 3a, and "Estimating the invasion probability (Step 3b)" Section for Step 3b

scanned in Google Earth Pro, were additionally added to increase the amount of reference data. In this way, all classified and digitized occurrences of *U. europaeus*, since 1988, are included in the latest reference data of 2020. That is, our reference data for each time step represents all areas were *U. europaeus* has occurred at least once until a given time step (maximal distribution).

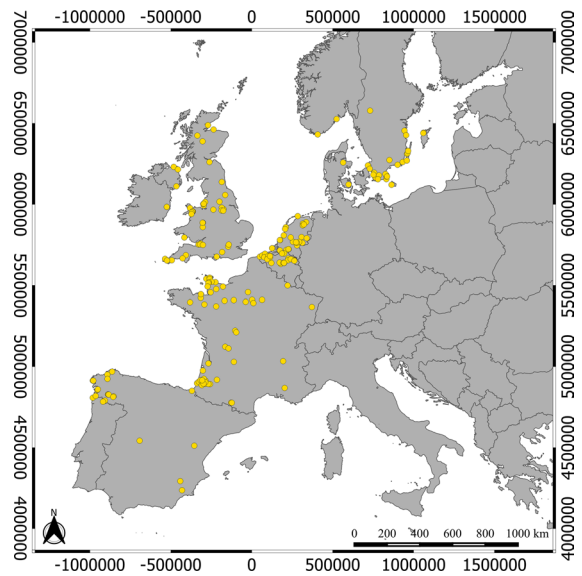*Modeling the habitat suitability (Step 2)*

To model the current and estimate the future habitat suitability for *U. europaeus* on Chiloé Island, the Maxent species distribution algorithm (Phillips et al. 2004) was used as implemented in the dismo package in R (Hijmans et al. 2017). In this study, the estimated potential spatial distribution provided by Maxent is translated to an habitat suitability ranging from 0 (low suitability) to 1 (high suitability).

The occurrence points of *U. europaeus* on Chiloé Island were derived from the reference data for the year 2020 originating from the classification of the Landsat images (see "Mapping the occurrence of Ulex (Step 1))" Section and the digitized patches in GEP (see "Remote sensing data") Section. Based on the aggregated occurrence map, we randomly selected 200 pixels using a stratified distance sampling within *U. europaeus* patches to avoid auto-correlation effects within the Maxent models (see Fig. 7). The occurrence points from the native range were downloaded from the Global Biodiversity Information Facility archive (Ulex L. in GBIF Secretariat 2022). To obtain a balanced set of invasive and native occurrence points we also randomly selected 200 GBIF points from Europe (see Fig. 8). In Fig. 9 the range of the selected environmental variables is shown for the selected occurrence points for Chiloé Island and Western-Europe. As background data for the Maxent model,
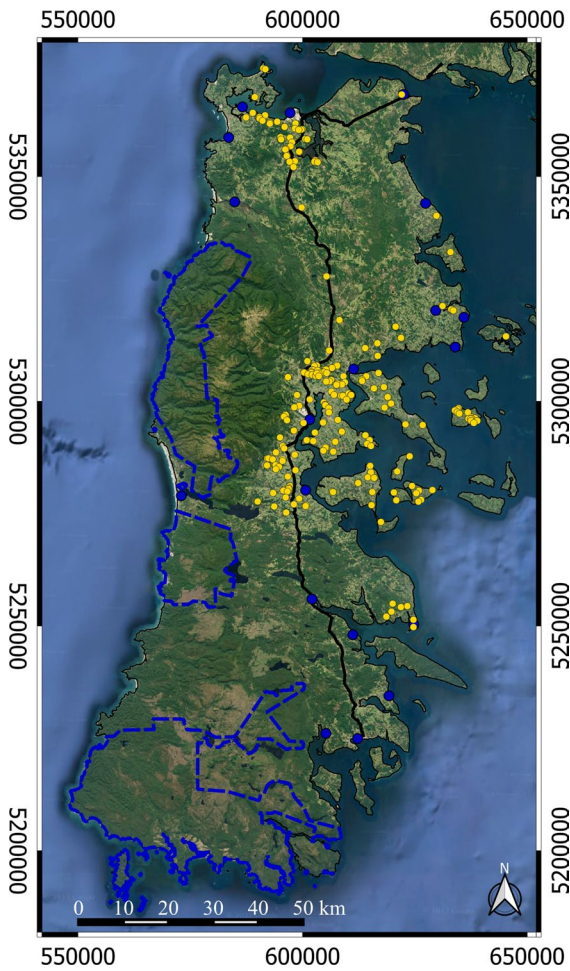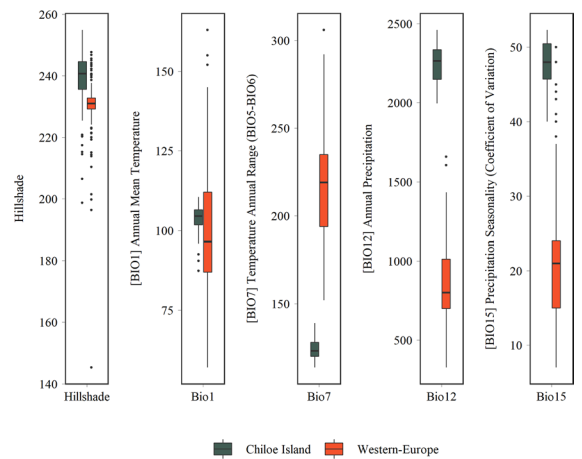
Fig. 7 Selected invasive occurrence points (yellow points) on Chiloé Island including national parks (blue dashed line), Ruta 5 sur (black line) and settlements (blue points) from Open Street Map

Fig. 8 Selected native occurrence points from Western-Europe (GBIF)

Fig. 9 Comparison of the selected environmental variables for the occurrence points of *U. europaeus* on Chiloé Island and Western-Europe (GBIF)

20.000 random points were selected for Chiloé Island as well as for Western-Europe. The background points for the native range are distributed over entire Western-Europe. In the first step of the SDM, the Maxent model was calibrated to the historic and recent (1988–2020) (see "Environmental data" Section) environmental variables based on the presence points describing the native and non-native realized niche. The presence points were divided into 75% for calibration and 25% for AUC

calculation to check the model quality. The calibrated SDM was then applied to the interpolated future (2020–2070) environmental variables (see "Environmental data" Section) to predict the habitat suitability in 5-years steps. This was done for both RCP scenarios separately.

Modelling the invasion probability with transition rules (Step 3a)

The last part of the modelling of the invasion probability consists of two steps. In step 3a ("Modelling the Invasion probability with transition rules (Step 3a)" Section) transitions rules for a Cellular Automaton model are identified based on past and recent Occurrence Maps, Habitat Suitability information (1988–2020) as well as information on roads and soil. These calibrated transition rules are used in step 3b ("Estimating the invasion probability (Step 3b)" Section) to estimate the future invasion probability based on optimal transition rules, the so-called level-sets.

The two main components of CA modeling are the applied transition rules and the considered neighborhood. Depending on the transition rules the state of the cells will be updated in each iteration (year). Here, a cell can either be colonized by *U. europaeus* (state = 1) or remain un-colonized (state = 0). Since the aim of the CA models is to predict the potential distribution of *U. europaeus* and not the actual invasion status, the state of a cell is not reversible once colonized by *U. europaeus*. In this study, the habitat suitability was used as baseline information to describe the likelihood of a cell of being invaded. However, since the invasion probability of an area by *U. europaeus* may not only depend on suitable climatic and topographical conditions but also on local soil conditions and spreading vectors, we altered the SDM-based habitat suitability based on the Open Street Map infrastructure layer and the soil map. This was done by assigning both geo-layers a specific habitat suitability value. Since *U. europaeus* is able to grow on many different soil types, only unfavourable soil types were considered as constraints (low values). After integrating the unfavourable soil types in the habitat suitability maps, roads (represented by a 100 m buffer around linear OSM road features) were additionally integrated (high values). The optimal values for soil and the roads are tested in "Estimating the invasion probability (Step 3b)" Section.

This refined environmental variability map builds the basic "landscape" for modeling the spread of *U. europaeus* with the CA models. The second important parameter to model the actual spread over time, is the neighborhood of each individual cell in the landscape. Depending on the number of neighbors that are already invaded by *U. europaeus*, the invasion pressure is calculated. The spatial definition of the neighborhood of an individual cell is another parameter that can be optimized (Ghosh et al. 2017). Following Li and Yeh (2000) a circular shape was used in this study instead of a rectangular, to equally consider the invasion state of all neighboring cells, independent from directions. Only pixels whose pixel centers were within the distribution radius were considered. In addition to the shape, the size of the neighborhood is also important, since it directly regulates the spreading distance of the target plant. *U. europaeus* can spread by growth several meters per year (up to 5 m), but can bridge far larger distances through the transport of seeds by animals (ants and birds) (Moss 1960) as well as seed transport in streams and by vehicles on roads (Hill 1949). Here, we tested several neighborhood sizes during the calibration task (see "Estimating the invasion probability (Step 3b)" Section). Based on the habitat suitability and the invasion pressure calculated from the neighborhood of each cell, the following two general transitions rules were derived:

1. Depending on the habitat suitability of a cell (from low = 0 to high = 1) a certain
2. invasion pressure (from low = 0% to high = 100%) is required to change the state from un-colonized (state = 0) to colonized (state = 1).

To consider the full gradual range of both parameters, we parameterized them and grouped them into four levels (low, medium-low, medium-high, high). The general assumption was, that a cell with a high habitat suitability value (e.g. > 0.6) requires only a low invasion pressure (e.g. > 1% of the surrounding pixel are already colonized) to change the state from 0 to 1 and vice-versa. Different levels of both parameters are tested in "Estimating the invasion probability (Step 3b)" Section to find the optimal combination to supplement the two general transitional rules by additional rules. These optimal transition rules are called level-sets. They were used to model the spread of *U. europaeus* from the initial state in 1988 to the recent state in 2020 and to forecast the future invasion pattern.

The CA modeling task is separated into two individual steps, (a) reconstruction of the invasion dynamics between the years 1988 and 2020 and (b) prediction of the future invasive pattern until 2100.

Table 2 Levels-sets of the environmental habitat suitability (HS) and invasion pressure (IP) values

Parameter	ID	Low	Medium-low	Medium-high	High
Env. suitability	HS1	≥ 0.1	≥ 0.3	≥ 0.5	≥ 0.7
	HS2	≥ 0.2	≥ 0.4	≥ 0.6	≥ 0.8
	HS3	≥ 0.3	≥ 0.5	≥ 0.7	≥ 0.9
	HS4	≥ 0.1	≥ 0.2	≥ 0.4	≥ 0.8
	HS5	≥ 0.2	≥ 0.3	≥ 0.5	≥ 0.9
	HS6	≥ 0.1	≥ 0.5	≥ 0.7	≥ 0.8
	HS7	≥ 0.2	≥ 0.6	≥ 0.8	≥ 0.9
Invasion pressure	IP1	≥ 1%	≥ 5%	≥ 10%	≥ 15%
	IP2	≥ 1%	≥ 10%	≥ 20%	≥ 30%
	IP3	≥ 1%	≥ 15%	≥ 30%	≥ 45%
	IP4	≥ 1%	≥ 20%	≥ 40%	≥ 60%
	IP5	≥ 1%	≥ 25%	≥ 50%	≥ 75%
	IP6	≥ 1%	≥ 30%	≥ 60%	≥ 90%

Table 3 Size of the circular neighborhood in pixel (NS), the extent of the neighborhood in pixel and the invasion distance in meter

Parameter	ID	Extent	Distance (m)
Neighborh. size	NS1	3 pix	30
	NS2	5 pix	60
	NS3	7 pix	90
	NS4	9 pix	120
	NS5	11 pix	150
	NS6	13 pix	180
	NS7	15 pix	210

Table 4 Habitat suitability (HS) values for soil (SO) and roads (RO)

Parameter	ID	HS values
Roads	RO1	0.6
	RO2	0.8
	RO3	1.0
Soil	SO1	0.0
	SO2	0.2
	SO3	0.4
	SO4	0.6

In the first iteration of the retrospective modeling, the CA models use the 30 m grid of the occurrence map of 1988 as starting point. The invasion pressure is always calculated first in each iteration based on the occurrence map. The occurrence map will then be updated according to the transition rules in each iteration. Finally, the CA models stop in 2020 and the modelled invasion status is compared to the mapped *U. europaeus* occurrences in 1999, 2015 and 2020 (see "Mapping the occurrence of Ulex (Step 1)" Section) to assess the model performance. For the latter two, the digitized patches from GEP (see "Remote sensing data" Section) were included.

The CA model performance was optimized in an exhaustive parameter tuning for all possible combinations (see next "Estimating the invasion probability (Step 3b)" Section). Once all tested transition rules, defined in Tables 2, 3 and 4, had been

evaluated, the upper 95% percentile (best 5%) of all transition rules were used to predict the invasion pattern of *U. europaeus* on Chiloé Island for the entire period between 1988 and 2070 (midterm forecast) and until the year 2100 (long-time forecast). For this purpose, the habitat suitability maps of the past (1988–2020) and the future period (2020–2070) were used. For the long-term forecast, the latest habitat suitability map from 2070 was used for all forecasts after 2070. The decision to also include a long-term forecast for 2100 (for which no climate scenario is available) was motivated by our interest in understanding whether the somewhat untouched areas of the two national parks in the south and west of the island are likely to be invaded by *U. europaeus* in the long-term. For the long-term forecasts, no changes in climate were considered because there were no suitable climate models for this time period. Therefore, the

advancing spread of Ulex in the long-term modeling Ulex is solely due to changes in invasion pressure, as habitat suitability remained constant.

In a last step, the predictions (0 = un-colonized; 1 = colonized) were aggregated and divided by the number of selected transition rules to receive the invasion probability of *U. europaeus* on Chiloé Island.

Estimating the invasion probability (Step 3b)

Based on the assumptions in "Modeling the Habitat Suitability (Step 2) and Modelling the Invasion probability with transition rules (Step 3a)" Sections about the spreading ability and the environmental requirements of *U. europaeus*, several parameter level-sets were tested to establish optimal transition rules for the CA models. Seven levels-sets of the habitat suitability (HS) and six levels-sets of the invasion pressure (Table 2) were tested in the CA model. Both parameters were always separated into four levels but the way the four levels were determined varied for HS. We tested three linear (HS1–HS3), two exponential (HS4 & HS5) and two exponential leveling (HS6 & HS7) assumptions (which drive the ways of how the HS from 0 to 100 % is split into four classes) to simulate the processes of invasion dynamics. The levels of invasion pressure on the other hand, were considered to be only a linear phenomenon. Since, the invasion pressure is related to the neighborhood size (NS) we additionally examined seven circular neighborhood sizes (Table 3). Furthermore, four habitat suitability values for the unfavourable soil types (SO) and three for roads (RO) (Table 4) were tested for the habitat suitability map aggregation. The aggregated suitability maps were included in the CA model to describe the properties of the cells.

As already mentioned, all parameter level-sets of HS, IP, NS, RO and SO were combined into different transition rules and tested with an exhaustive parameter tuning approach. In total 1848 unique transition rules were evaluated. For each transition rule set the Balanced Accuracy (BA) was calculated by comparing the predicted invasion pattern for the years 1999, 2015 and 2020 with the corresponding aggregated reference data (occurrence maps) of the same year (see "Mapping the occurrence of Ulex (Step 1)" Section). The BA was used because of the

imbalance between the amount of pixels representing presence or absence of *U. europaeus* within the reference data. The BA averages (Eq. 1) the True Positive Rate (TPR or Sensitivity) (Eq. 2) and the True Negative Rate (TNR or Specificity) (Eq. 3). Finally, the BA values of the three time stamps were averaged for each parameter setting.

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2} \qquad (1)$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \qquad (2)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \qquad (3)$$

Results

Past and recent occurrence maps

The past and recent occurrence maps (see Step 1 in Fig. 6) generated with the Random Forest algorithm achieved a very high predictive accuracy for all four time-stamps. The Overall Accuracy for the validation points ranges between 0.975 and 0.993 (Table 5). However, there were some mis-classifications between *U. europaeus* and some smaller land cover classes which were not included in the classification. This was particularly the case for the typical Magellan peat-bogs on Chiloé Island, which can show similar yellow color as *U. europaeus*. This was also recognized by Gränzig et al. (2021). Thus all occurrences maps were visually checked and clearly wrong pixels were manually corrected.

Table 5 Accuracy assessment of the Random Forest classifications performed with recent and historic reference data as seen in Figs. 4 and 5 as part of the workflow Step 1 shown in Fig. 6

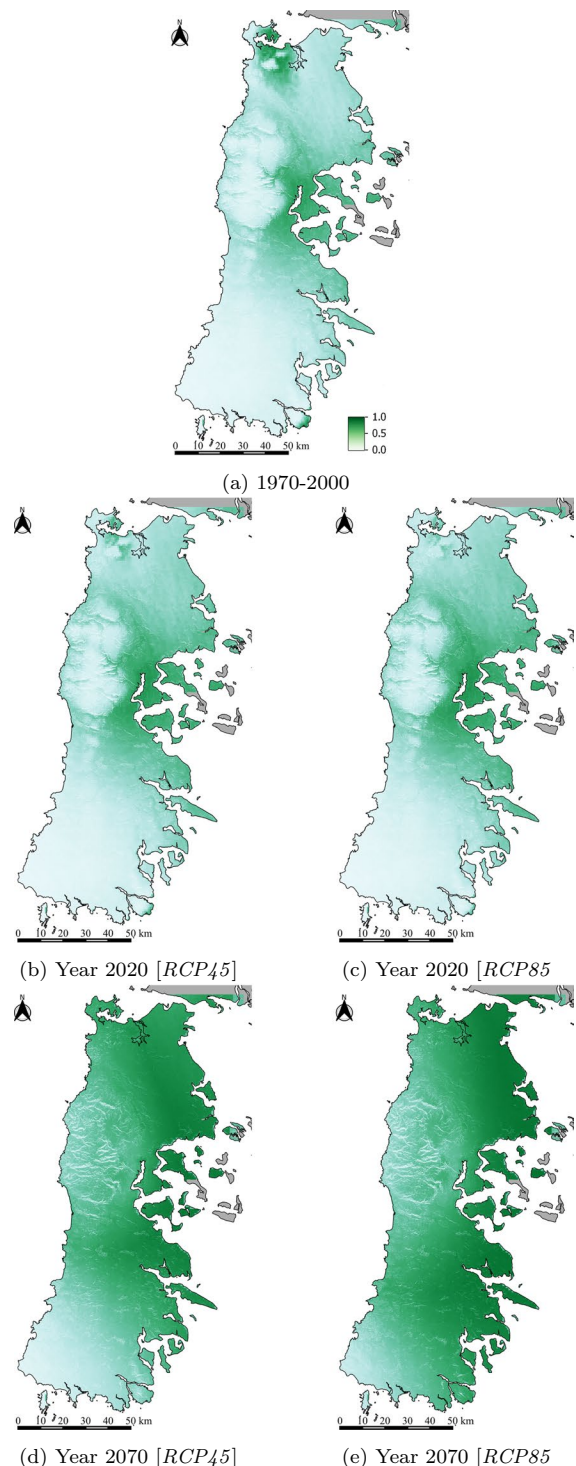| Year	Overall accuracy	Balanced accuracy	Kappa
1988	0.975	0.912	0.889
1999	0.980	0.938	0.914
2015	0.993	0.976	0.971
2020	0.979	0.933	0.910

Habitat suitability maps

The Maxent species distribution model (SDM) was used to estimate the past (1988–2020) and the future (2020–2070) habitat suitability for *U. europaeus* on Chiloé Island. The retrospective SDM reached an accuracy (here defined as AUC value) of 0.888 for training (75% of the presence points) and 0.851 for test data (25% of the presence points). According to the jackknife test, within the Maxent algorithm, the Annual Temperature Range (Bio 7) has the most unique information of all selected environmental variables. The precipitation seasonality (Bio15) was the second most important variable, by some distance. Hillshade, Annual mean temperature (Bio1) and annual precipitation (Bio12) show comparably low unique information. In Fig. 10 the habitat suitability maps of the past and recent time periods and for the year 2070 are shown for both climate scenarios.

Calibration of the cellular automata

The overall performance of the unique transition rule sets tested during the exhaustive parameter tuning ranges between 0.622 and 0.933 Balanced Accuracy (BA). However, some noticeable differences between the included parameter level-sets are obvious. The boxplots in Fig. 11 show the BA of all Cellular Automata models, depending on whether a certain level-set, neighborhood-size or habitat suitability value for soil and roads are included in the transition rule or not. Thus, the influence of the individual parameter level-sets are highlighted.
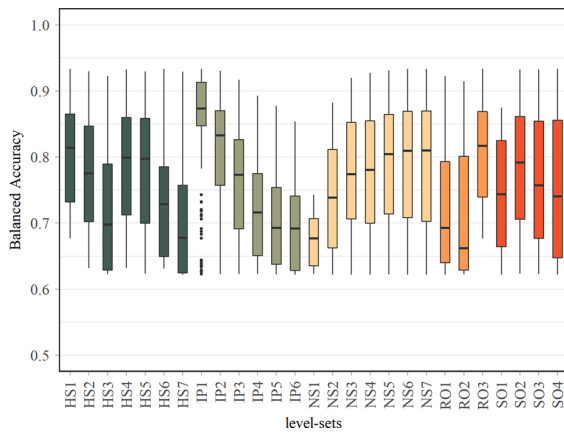
The BA values of the six invasion pressure (IP) level-sets (see Table 2 for definition) show the strongest differences among each other. This indicates, that this parameter has a strong influence on the performance of the models. IP1 shows the overall best results, since it reaches the above mentioned maximal BA value and an average of 0.855 BA. In comparison, models including IP6 achieve only a maximal BA value of 0.854 with an average of 0.692. The maximal BA value of the other invasion pressure level-sets is gradually decreasing step-wise from 0.930 (avg. BA 0.813) to 0.877 (avg. BA 0.706).

The circular neighborhood size (NS) also shows some differences in BA values between the examined sizes. NS1 shows the lowest overall maximal BA (0.742) and the lowest average BA value (0.672)

(a) 1970-2000

(b) Year 2020 [*RCP45*] (c) Year 2020 [*RCP85*

(d) Year 2070 [*RCP45*] (e) Year 2070 [*RCP85*

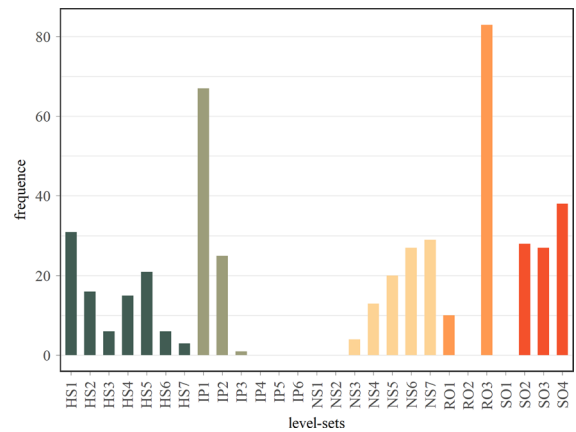**Fig. 10** Habitat suitability maps

Fig. 11 Balanced accuracy of all CA models according to whether a certain level-set (habitat suitability = HS, invasion pressure = IP, neighborhood size = NS, roads = RO, soils = SO) was included in the transition rules or not. More information on the definition level-sets can be found in Tables 2, 3 and 4. Colours indicate just different groups of level-sets

Fig. 12 Contribution of the level-sets (habitat suitability = HS, invasion pressure = IP, neighborhood size = NS, roads = RO, soils = SO) on the selected upper 95%-percentile of all tested transition rules (93 cases). A higher frequency shows that this level-set was more often included in a model of a very high accuracy. More information on the definition level-sets can be found in Tables 2, 3 and 4. Colours indicate just different groups of level-sets

of all parameter levels. This means that all models including a small neighborhood size of 30 m (NS1) perform by far the worst. Even IP1 is showing negative outliers when combined with NS1. Accordingly, the BA values increase with increasing neighborhood size. Models using neighborhood sizes of 180–210 m (13–15 pixels) produce the maximal (0.933) BA value and had similar average BA values (0.789 and 0.787).

The influence of the habitat suitability (HS) is smaller in comparison to the invasion pressure level-sets. The maximal BA values of the seven different level-sets show only small differences. They range between 0.922 and 0.933 BA. However, the average BA values show some differences. HS1 achieves the highest average BA value (0.804), whereas HS6 shows the lowest (0.704). The average BA values of HS2, HS4 and HS5 range around 0.77. HS3 and HS7 reached average BA values around 0.72. Interestingly, HS1 in combination with NS1 shows slightly better results than all other parameters, with the exception of R03.

The tested habitat suitability (HS) values for roads (RO) also show some influence on the model performance. If the influence of roads on habitat suitability is considered very high (RO3—see Table 4 for definition), the model achieves the maximum BA value (avg. BA 0.809), but when a lower habitat suitability for roads is assumed (RO1 or RO2), the maximal BA

value drops to 0.922 (avg. BA 0.721) and 0.915 (avg. BA 0.713), respectively.

A reverse trend can be seen for the constrains associated with soil (SO). All models that assumed no influence of soils on habitat suitability (SO1—see Table 4 for definition) show a comparatively lower BA value (max BA 0.874; avg BA 0.746) than the other three tested habitat suitability values, which gradually increase the influence of soil. All other values of habitat suitability for soil (SO1–SO3) showed similar maximum BA values around 0.93, but slightly different average values (0.783, 0.765, and 0.755, respectively).

If only the upper 95% percentile of all tested transition rules sets are considered, 93 different sets remain. The range of BA values is then concentrated between 0.916 and 0.933. Figure 12 depicts the frequency of the individual parameter level-sets on the upper 95%-percentile. In most of the cases HS1 and IP1 for habitat suitability and invasion pressure are included in the best performing transition rules with a large circular neighborhood (NS7) and well as high influence of roads and soils (RO3 and SO4). All selected 93 transition rules sets were used to estimate the invasion probability of *U. europaeus* on Chiloé Island for the entire period between 1988 and 2070 (mid-term forecast) and until the year 2100 (long-time forecast).

Historical invasion patterns

Based on the upper 95%-percentile of all tested parameter-sets, the historical (1988–2015), the actual (2020), the mid-term future (until 2070), and the long-term future (until 2100) potential maximal spread of *U. europaeus* was estimated. The resulting predictions were summed for each time-stamp and divided by the number of CA models performed.

In Fig. 13 the historical and the actual invasion probabilities are shown. In the year 1990 (2 years after the initial start of the CA models) the extent of the possible invasion of *U. europaeus* is limited to small local patches in the North and in Central-Chiloé. More information on the local development of the spread of *U. europaeus* can be found in the Appendix. By 2020, most of the small isolated patches of high probability have merged into larger contiguous areas. However, clear differences in the extent of areas with high and low invasion probabilities are apparent for the north and Central-Chiloé. In 2010 and 2020, the areas showing a the high invasion probability increase strongly. A continuous invasion corridor is now visible with high invasion probabilities along the "Ruta 5 sur" from Ancud to Central-Chiloé. However, this corridor is much wider in 2020 than in 2010. The highest invasion probabilities are still observed in the North and along the roads. It is is very likely that *U. europaeus* could occur almost anywhere on the island on the East-coast. The enclaves more in the South are still isolated, however, due to their increasing expansion, they are almost connected to the main invasion areas in 2020.

Future invasion pattern

In Fig. 14 the potential invasion patterns of *U. europaeus* is illustrated for the mid-term and the long-term forecast and for both RCP scenarios. Until the year 2070 the potential invasion will spread much stronger into the eastern, western and southern parts of the island. The North will be very likely completely invaded by *U. europaeus*. In the year 2100, both scenarios shows only little differences. The only un-invaded areas will be the main parts of the Parque Nacional Chiloé, the Parque Tepuhueico at the west-coast and of the Parque Tantauco at the south-coast of Chiloé Island. Table 6 shows the estimated spatial extent of the historical and future invasion of *U.*

Table 6 Estimated spatial extent of historical and future *U. europaeus* invasion

Year	Scenario	Spatial extent (km^2)
1990	Historical	309
2000	Historical	1383
2010	Historical	1993
2020	Historical	2565
2070	RCP45	4800
2100	RCP45	5519
2070	RCP85	4877
2100	RCP85	5801

europaeus. While the area in 1990 was quite small (309 km^2), by 2020, the invaded area had increased more than eight-fold (2565 km^2). For the mid-term forecast for the year 2070, an area of 4800 km^2 for RCP45 and 4877 km^2 for RCP85 was estimated. In the long-term forecast, the area for RCP45 increases to 5519 km^2 and for RCP85 even up to 5801 km^2.

Discussion

In this study we introduced a workflow combining correlative SDMs and CAs that allows to reconstruct the historic spread and the modelling of the future spread of *U. europaeus*. This general workflow is based on three main parts:

– Mapping of the occurrences of *U. europaeus* between 1988 and 2020
– Modelling the habitat suitability of *U. europaeus* with a correlative SDM
– Estimating the invasion probability with calibrated CA models

In the following sections we will discuss each of this parts individually and then end with a paragraph on the synergistic use of these methods.

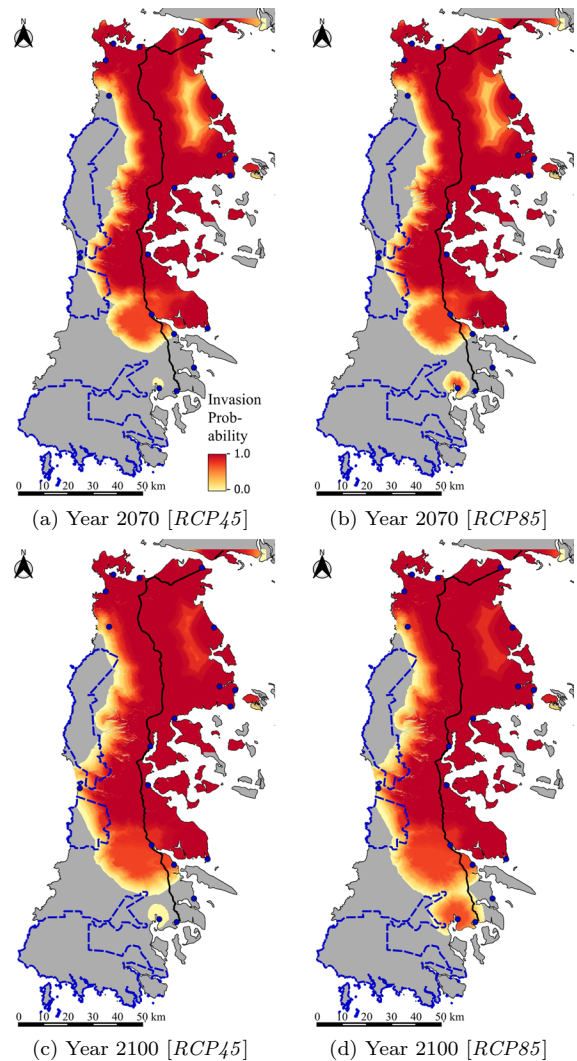## Mapping occurrences between 1988 and 2020

For mapping the historic to recent occurrences of *U. europaeus*, the overall accuracies reached very high values (between 0.975 and 0.993), which is in
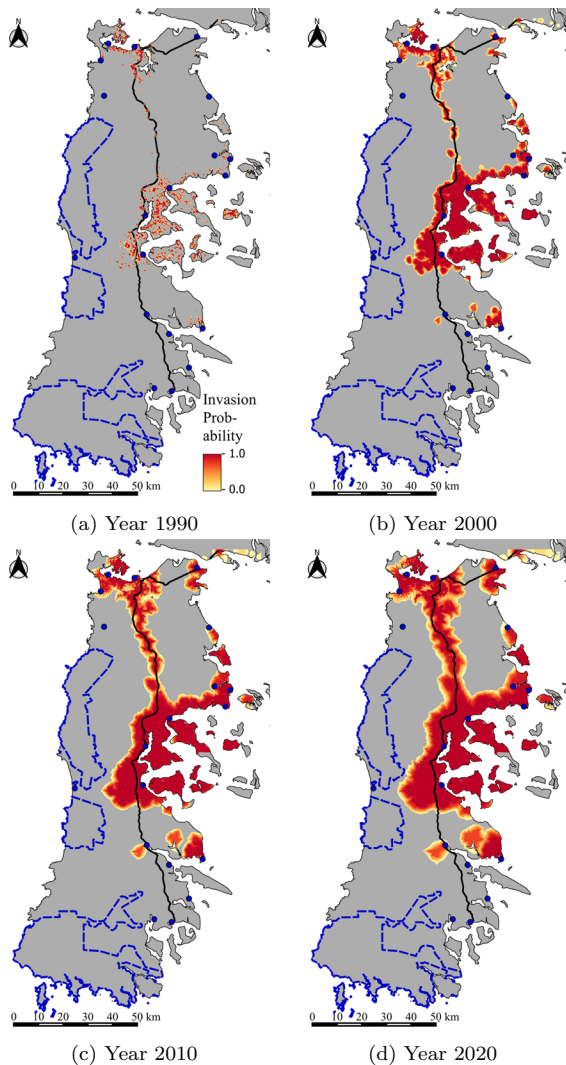
Fig. 13 Historical (1990, 2000 and 2010) and actual (2020) estimation of the invasion pattern of *U. europaeus*

Fig. 14 Mid-term forecast (2070) and long-term forecast (2100) of the invasion pattern of *U. europaeus* for the scenarios RCP45 and RCP85

accordance to other studies for this IAP (Kattenborn et al. 2019; Gränzig et al. 2021). These results are related to the distinctive yellow flowering of *Ulex euopaeus*, which makes it comparably easy to detect its presence, even in the past, if a suitable remote sensing dataset is available.

Nevertheless, it remains a challenge to identify sufficient historic presence points of *U. europaeus* to serve as reference data. Only with an appropriate number and, even more important, with a representative distribution of presence points, the historic and the future invasion pattern can be estimated

realistically (Elith et al. 2011). To achieve this, we collected more than 4500 individual occurrences from high-resolution GEP imagery available for the time period between 2009 and 2020. A drawback of this approach is that the distribution of the identified patches depends strongly on the spatial coverage and even more on the acquisition time of the high-resolution images in GEP. Although the coverage of high-resolution imagery was fairly well distributed for the main inhabited parts of the island (see Fig. 4), it cannot be excluded that patches of *U. europaeus* were missed due to a lack of spatial coverage or a temporal

offset to the main flowering phase of *U. europaeus*. For example, there were no high-resolution images available for the area around the village of Chepu along the Rio Chepu. The same applies for the very south of Chiloé Island. It can not fully be clarified whether *U. europaeus* already has invaded that area. Concerning the problem of temporal offset of GEP images with the flowering phase, it must be considered that the flowering phase of *U. europaeus* on Chiloé Island can have a temporal delay between the north and the south, due to differing climatic conditions (Rees and Hill 2001). Therefore, images showing the flowering phase in one part of the island may be too early or too late in other parts. Individuals of *U. europaeus* which had already withered or had not yet started to flower. would then be difficult to distinguish from native shrubs and other vegetation. Moreover, in the our approach for collecting reference data, we assumed that a site stayed colonized once *U. europaeus* was established. We deemed this to be a reasonable approach knowing that the successful management of *U. europaeus* was limited in the recent decades and since we were most interested in determining the invasion potential.

A further challenge was to derive reliable training points for the time period before 2009 where no high resolution data was available. We traced back points identified between 2009 and 2020 but for 1988 and 1999 the confirmation of the training points relies only on Landsat data at 30 m resolution. This approach could be highly error-prone if the target species does not have a unique bloom or if only a small number of training points were derived. Since the yellow flower of *U. europaeus* is visible even on Landsat's 30 m images (see Altamirano et al. 2016; Shepherd and Lee 2002), it can be assumed with high confidence that the patches identified in the GEP images and backtracked in the Landsat data indeed represented *U. europaeus*. Even if some patches were incorrectly identified as *U. europaeus*, the impact on the quality of the classification maps would be small thanks to the large amount of patches reliably identified in GEP. The spatial offsets between GEP high-resolution images and other remote sensing data, can be a source of uncertainties as discussed by Goudarzi and Landry (2017) and Yu and Gong (2012). Depending on the location on the globe the horizontal error can be so high (Potere 2008), that manual geometric correction is required to remove offsets (Yu and

Gong 2012). Since, no high precision ground control points were available to estimate the spatial offsets of GEP's high-resolution images for our study site, the spatial offsets were evaluated by the ruler tool available in GEP. With only one exception, all high-resolution images were found to have a spatial offset less than 30 m resolution to the applied Landsat images. These small offsets seem acceptable for the purpose of modelling invasive species with medium resolution images as stated by Visser et al. (2014) and are in accordance with the findings of other studies (Goudarzi and Landry 2017; Pulighe et al. 2016). The one outlier image showed a very strong offset of around 300 m to all other high-resolution images. All patches delineated in this image were manually corrected to match the previous and following high-resolution images.

In summary, the workflow for obtaining spatially explicit occurrence maps for *U. europaeus* across longer time periods using Landsat satellite images and supervised classifications trained with reference data from high-resolution images works reasonably well. The expected limitations and uncertainties were all considered acceptable in the given case due to the specific spectral properties of *U. europaeus* which enabled a reliable identification, both during the visual interpretation phase as well as during the supervised classification (the latter reached very high overall and kappa accuracies).

Mapping the habitat suitability with a SDM

In the second part of the methodical workflow we calculated habitat suitability maps using Maxent. We used *U. europaeus* presence data from both, the native range and the non-native range on Chiloé Island. Including presence data from the native range of an invasive alien species (IAP) for modelling their distribution is crucial for taking into account that some niche shifts and also expansions have been observed for IAP (Pena-Gomez et al. 2014). Despite Chiloé Island and Western-Europe both having a maritime climate, Fig. 9 indicates that presence points from Chiloé Island and from Western-Europe show different ranges for some of the included variables. Whereas the illumination conditions (hillshade, CI = 239/WE = 230) and the annual mean temperature (Bio1) are very similar across both regions (CI = 10.4/WE = 10.0), a

strong difference in the temperature annual range (Bio7) is apparent (CI = 124, WE = 214). Chiloé Island has an on average lower difference between winter and summer temperatures. Since the presence points from GBIF are distributed over the entire range of Western-Europe (from South-Spain to South-Sweden), the range of the annual temperature range is large. However, even for the strongly oceanic influenced Mediterranean climate of Galicia in Spain, where very mild winter and summer temperatures occur, the range is still higher than on Chiloé Island. Even larger differences can be seen for the annual precipitation (Bio12). The annual amount of precipitation is almost 1400 mm higher for Chiloé Island than for the presence points from Western-Europe (CI = 2243/WE = 858). There is also a stronger variation of precipitation during the year on Chiloé Island (Bio15, CI = 47/WE = 21). This underscores that the *U. europaeus* population on Chiloé Island is on the edge or even outside of its native range. Considering only the presence points of Chiloé Island or only from Western-Europe in our SDM would not correspond to this extended niche of *U. europaeus*. The differing conditions on Chiloé Island seem to create an advantage for *U. europaeus*. Potentially this is related to missing competitors or seed predators that may normally occur in regions with less extreme precipitation and which may have more difficulties thriving under the conditions on Chiloé Island. As one example, Norambuena et al. (2000) analysed the potential of seed predators (gorse spider mite) to control the invasion of *U. europaeus* in Chile. However, the establishment of some of the mite strains was hampered by the high precipitation rates in some parts of Chile.

Overall, the results of the Maxent habitat suitability models presented here were acceptable with an AUC of > 0.85. We deem that the corresponding habitat suitability maps provide useful information about the areas with climatic conditions that are suitable for *U. europaeus* to establish and grow. However, modeling the species distribution based only on climate variables is not adequate for species where anthropogenic and biotic factors may contribute to the invasion success and the spreading behavior (Elliott-Graves 2016). To account for this, we included a proxy for anthropogenic drivers in the CA models by implementing

infrastructure-related data and considering the dispersion distance which is influenced by human activities. The CA models are discussed with more detail below.

Using cellular automata for estimating future distributions of *U. europaeus*

In the last step of our workflow we implemented CA models to combine the habitat suitability maps with a spatially dynamic approach which allows to capture the spreading behavior of *U. europaeus* on Chiloé Island. Finding reasonable transition rules for the CA models was essential to re-produce the historic and forecast the future spread of *U. europaeus*. To account for the complex dynamics of invasion, we parameterized the main factors (e.g., habitat suitability based on climate and terrain data, as well as infrastructure and abiotic factors) and conducted an exhaustive search for the most appropriate transition rules for the CA models. In general, those rules that allow a rapid spread of *U. europaeus* were identified as the best possible options to estimate the historical invasion probability. IP1 was the most used level-set in the upper 95%-percentile. This indicates that only a low invasion pressure is needed to change the state of a cell from un-colonized to colonized for a given habitat suitability (HS) value. This underlines the high invasion success of *U. europaeus*. Since beside IP1 one of the two largest neighborhood sizes (NS6 and NS7) was frequently used in the best performing transition rules, a high dispersal speed (up to 180 m or even 210 m per year) can be assumed. Furthermore, the fact that HS1 (low HS values for the four levels) was the most used level-set for the habitat suitability, reinforces the observation that only comparably low invasion pressure and habitat suitability is needed to change the state of a cell to "invaded".

The trend that the most suitable rule-sets had high HS values for the unfavorable soil types (e.g., SO4) and for the roads (RO3), indicates that *U. europaeus* is hardly affected by spatial constraints with respect to soil types and is also able to spread faster along roads. All these trends stress the enormous invasion potential of *U. europaeus* mentioned by Clements et al. (2001) and stand in contrast to the study of (Zheng et al. (2015)) who found clear improvements of the CA models by including invasion constrains. In our case, including the soil layer only marginally

improved the CA model's performance. Nevertheless, we cannot exclude the possibility that invasion constrains based on other factors, for example based on LULC, could improve our CA model results. Some studies have also included a random spread component to their CA to account for dispersion related to human activities (Barbosa et al. 2018). In our approach human activities were taken into account by assuming a higher habitat suitability for roads and their surrounding area (100 m buffer) as well as by considering comparably large dispersion distances of *U. europaeus*. We assumed that the different neighborhood sizes (NS) we considered relate to different dispersion processes for *U. europaeus* (vegetative or by the transportation of seeds by animals, humans or streams). We did not consider the possibility that *U. europaeus* could have still been planted on Chiloé Island during the time period between 1988 and 2020. Since we do not have any information on this, we restrained from adding this factor to the CA models.

Furthermore, we are aware that 1988 does not represent the first occurrence of *U. europaeus* on Chiloé Island. We could already identify numerous *U. europaeus* patches in 1988. We therefore have to assume that the invasion of *U. europaeus* started several years or even decades before 1988. The invasion probability of some regions must have been already high in the initial year of our study. We can also not fully exclude the possibility that *U. europaeus* had a stronger presence on Chiloé Island in 1988 than our derived historical reference data indicates. If this would be the case, the transition rules detected for the CA models may have been different and rules that indicate a slower dispersion of *U. europaeus* might have prevailed.

One notable restriction of our CA models relates to the uncertainties of the historic occurrence data which we obtained from remote sensing data. For example, *U. europaeus* patches can be seen in Google Earth Pro for the area east of Quellin, but those points are not included in the aggregated reference points until the year 2020. One possible explanation could be, that *U. europaeus* was introduced in this area through water transport by small boats. This assumption is supported by Google Earth Pro images in which *U. europaeus* patches can be seen on the coastline of Isla Tranqui. We hypothesize that *U. europaeus* migrated from the coast near Quellin across the sea to Isla

Tranqui and from there back to the main island near the village of Auchac. However, it is unclear whether this movement started from some seeds transported by the stream to the Gulf of Corcovado or from some small boats docked at Isla Tranqui, transporting seeds to until then un-colonized areas. Simulating this movement or behavior of *U. europaeus* would be too complex for the comparably simple approach to simulate spread of the IAP with the CA model presented here but could be addressed in future studies.

While the CA models were able to reproduce the invasion patterns observed in the occurrence maps available for four dates between 1988 and 2020 with good balanced accuracies (>0.91), we are aware that we still do not capture all relevant processes. It is important to stress that the high accuracies obtained for the CA model notably depended on a good parameterization and that the exhaustive search for suitable parameters showed clear tendencies. This is important as it strengthens our confidence that the spread behavior of *U. europaeus* can actually be captured with our approach. If very differing parameter sets would have led to comparable accuracies, the meaningfulness of the selected parameterization sets would have been hard to interpret.

Based on this, we assume that using the CA models to predict the future invasion state of *U. europaeus* is acceptable, even though we are fully aware that the further we reach into the future, the less reliable the predictions are likely to be for obvious reasons. As many transition rule sets led to comparable accuracies and it was hard to identify a single clearly best-performing set, we used the average of the 5% best models to predict the future state of the *U. europaeus* invasion. We think this is a good compromise to integrate the variability of the different parameter sets with comparable model accuracies in the predictions and thereby potentially increase the robustness of the predictions.

Overall, the prediction maps obtained for 2070 and 2100 showed reasonable patterns which represented a continuation of the spatial developments of the spread of *U. europaeus* observed between 1988 and 2020. We believe that if no unexpected changes are introduced to the ecosystem (e.g., a new seed predator or other biotic antagonists for *U. europaeus*), the predicted invasion patterns may give a realistic

estimation of the future state of *U. europaeus* invasion on Chiloé Island.

Reflection on the suggested combined approach

Our workflow which combines presence data of *U. europaeus* obtained from remote sensing images, habitat suitability maps and CA models performed well according to the sparse validation data that we had available. Reflecting on the limitations of applying either of the three steps in isolation, our approach provides several obvious advantages. By applying remote sensing data instead of field data, we obtain a spatially and temporally more complete dataset. This is particularly relevant for understanding the spatial invasion dynamics and for avoiding well-known observer biases of field-collected data. By using this spatially and temporally more complete information to train a habitat suitability model which is then fed into CA models, we manage to simultaneously consider the environmental base-constrains and the spread dynamics which are also influenced by anthropogenic factors, which were also integrated into the CA models. These advantages are contrasted by several challenges associated with the approach. Using multiple models which each contain a certain level of error or uncertainty may lead to cumulative errors which may be particularly problematic for the last step of the workflow where we predict the invasion dynamics to the future. Here, the climate scenario data further add to the uncertainty. Tracing all mentioned uncertainties throughout the workflow may be possible but typically requires sophisticated statistical approaches which are not readily available. In our study, we assume that uncertainty problems were less pronounced because each of the models in the individual steps, which were based on actual data, performed reasonably well and led to plausible results. It is, however, clear that this problem may be a lot more pronounced in other use-cases. Some key aspects which may hamper the transferability of the presented workflow to other species and regions include: 1. the spectral properties and the resulting detectability of the target species from remote sensing data, 2. availability of high resolution historic images, 3. key-factors driving the invasion success and potentially missing data on these drivers.

Conclusions

In this study we present a workflow that allowed us to reconstruct the historic spread of *U. europaeus* on Chiloé Island in the South of Chile. We used a combination of multispectral satellite data, GBIF occurrence data and climate data from the World-Clim database to derive habitat suitability maps which were then fed into CA models along with soil and infrastructure information to first model the spread of *U. europaeus* starting from the initial situation in 1988 to the current state in 2020. The derived model was then applied to predict the future spread of *U. europaeus* making use of climate scenario data. All modeling steps resulted in plausible spatial patterns and reasonable accuracies which indicate that the overall approach is sound.

The key innovation of the approach lies in the combination of spatially dynamic CA models with the habitat suitability products, representing the full ecological niche. Only the combination of habitat suitability and the spatial modeling of the invasion process enables a reasonable prediction of future invasion spread patterns. For the latter, the availability of high quality spatially continuous occurrence maps for the invasive species at different points in time is a key-requirement. The purely remote sensing-based solution suggested here to provide such maps may be suitable not only for this case but also for other cases where the species is clearly identifiable in the remote sensing data.

Our results predict an ongoing spread of *U. europaeus* over large parts of Chiloé Island until 2100 with just slight difference for the examined climate scenarios. However, we have to point out that this is a prediction of the potential maximal spread of *U. europaeus* based on the historic invasion dynamics. Changes in environmental, anthropogenic or ecological factors may affect the invasion dynamics of *U. europaeus* in the future. We nevertheless believe that our prediction maps indicating the current and future spread of the species can support local land managers in their efforts to prevent further spread of *U. europaeus* and potentially also help to plan mitigation measures. Moreover, it might be worthwhile to dedicate more research to develop improved control and management techniques to contain the still ongoing invasion of *U. europaeus*.

Declarations

Conflict of 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.

Appendix

See Tables 7 and 8.

Table 7 Utilized general circulation models (GCM) of the coupled model intercomparison Project 5 (CMIP5)

GCM	Modeling center (or Group)
ACCESS1-0	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia
BCC-CSM1-1	Beijing Climate Center, China Meteorological Administration
CCSM4	National Center for Atmospheric Research
CNRM-CM5	Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory
GISS-E2-R	NASA Goddard Institute for Space Studies
HadGEM2-AO	National Institute of Meteorological Research/Korea Meteorological Administration
HadGEM2-ES	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)
HadGEM2-CC	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)
INMCM4	Institute for Numerical Mathematics
IPSL-CM5A-LR	Institut Pierre-Simon Laplace
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine- Earth Scienceand Technology
MRI-CGCM3	Meteorological Research Institute
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MPI-ESM-LR	Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)
NorESM1-M	Norwegian Climate Centre

Table 8 Abbreviations of the Worlclim bioclimatic variables

Code	Bioclimatic variable
BIO1	Annual mean temperature
BIO2	Mean diurnal range (mean of monthly (max temp–min temp))
BIO3	Isothermality (BIO2/BIO7) (times 100)
BIO4	Temperature seasonality (standard deviation times 100)
BIO5	Max Temperature of warmest month
BIO6	Min Temperature of coldest month
BIO7	Temperature annual range (BIO5-BIO6)
BIO8	Mean temperature of wettest quarter
BIO9	Mean temperature of driest Quarter
BIO10	Mean temperature of warmest quarter
BIO11	Mean temperature of coldest quarter
BIO12	Annual precipitation
BIO13	Precipitation of wettest month
BIO14	Precipitation of driest month
BIO15	Precipitation seasonality
BIO16	Precipitation of wettest quarter
BIO17	Precipitation of driest quarter
BIO18	Precipitation of warmest quarter
BIO19	Precipitation of coldest quarter

Description of the local modelling results

The invasion patterns of *U. europaeus* were described in a summarized way in the result section. In the following, the spatial patterns are explained in more detail.

At the beginning of the modelled time period in 1989 *U. europaeus* is present around the town of Ancud in the North of the Island. Patches of high probability also appear more frequently along the coastline and extend to Punta Quetalmahue and end at the village of Teguaco on the Playa Mar Brava. Also the headland in the very north of the island shows a high probability to be invaded at some small locations in 1990. In the North-East, there are some patches with high invasion probability around the small town of Chacao, were the "Ruta 5 sur" enters the island. Starting from Ancud some isolated patches of high probability spread along the "Ruta 5 sur" in southern direction to the densely populated area around the towns of Castro, Chonchi and Dalcahue. This area shows a very dense pattern of high probabilities. The islands of Quinchao, Quenac, Lemuy, Chelín and Quehui, between the Gulf of Ancud and the Gulf of Corcovado, were also most likely already colonized by *U. europaeus* in 1989. From this probably heavily

infested area, the possible occurrences extend further eastward along the coast towards the villages of Tenaún Bajo, Quicavi and Montemar. Also on the island of Caucahué some small patches of *U. europaeus* have very likely been present in 1989. A little further north of Isla Caucahué, on the mainland of Chiloé Island, near the village of Aucho, a small patch of *U. europaeus* also seems very likely. Some small enclaves of *U. europaeus* can also be observed around the town of Queilén and the village of Natri Bajo in the South of the Island. These patches could represent the most remote and southern occurrences of *U. europaeus* in 1989.

By 2020, there is a large and almost contiguous patch of high probability in Central-Chiloé with only lower values at the narrow invasion borders. However, there are also some clearly visible core areas with high probabilities and broad invasion borders with lower values in the north. Especially the roads show much higher values than their surrounding areas. The extent of the remote enclaves between the city of Queilén and the village of Natri Bajo has also increased.

Until the year 2070 the potential invasion area will spread much stronger into the eastern, western and southern parts of the island. The North will

be very likely completely invaded by *U. europaeus*. The area between the "Ruta 5 sur" and the east-cost will be invaded too. However, in the interior of these two regions lower probability can be observed. A long modelled invasion front begins from the village of Chepu at the Rio Chepu and runs to the small village of Chanco on the east-coast. This invasion front runs around the higher regions of the Parque Nacional Chiloé (Chiloé National Park) with the dense Valdivian rain forest on the west coast. In Central-Chiloé this front almost reaches the west coast at the village of Cucao after following Lago Huillinco and Lago Cucao. The small enclaves of high probability between the town of Queilén and the village of Natri Bajo will be reached by the main invasion area in 2070. Starting from this point the potential invasion will spread stronger into the south according to the model. However the model shows only medium probabilities. A new enclave of a potential *U. europaeus* infestation was modelled in the south at the village of Yaldad next the the town Quellón. Here we can observe the strongest difference between the two RCP scenarios. In the RCP45 scenario, the enclave seems to be in its initial stage, with a very limited extent and a low invasion probability. In the RCP85 scenario, on the other hand, the enclave shows a higher probability and already starts to spread into the surrounding areas.

In the year 2100, both scenarios shows only little differences. The invasion front will move stronger into the Parque Nacional Chiloé and will finally reach the west-coast at the village of Chanco. In the RCP45 scenario, the isolated enclave at the village of Yaldad increases just slowly and the probability remains low. However, in the RCP85 scenario, this area is much stronger infested by *U. europaeus* and the city of Quellón shows high invasion probabilities. It is even possible that the main invasion area will connect with this small enclave in the year 2100. The only un-invaded areas will be the main parts of the Parque Nacional Chiloé, the Parque Tepuhueico at the west-coast and of the Parque Tantauco at the south-coast of Chiloé Island.

References

Altamirano A, Cely JP, Etter A, Miranda A, Fuentes-Ramirez A, Acevedo P, Salas C, Vargas R (2016) The invasive species *Ulex europaeus* (Fabaceae) shows high dynamism in a fragmented landscape of south-central Chile. Environ Monitor Assess. https://doi.org/10.1007/s10661-016-5498-6

Araújo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. Ecology 93(7):1527–1539

Barbosa NP, Ferreira JA, Nascimento CA, Silva FA, Carvalho VA, Xavier ER, Ramon L, Almeida AC, Carvalho MD, Cardoso AV (2018) Prediction of future risk of invasion by limnoperna fortunei (dunker, 1857)(mollusca, bivalvia, mytilidae) in brazil with cellular automata. Ecol Ind 92:30–39

Barona PC, Mena C (2014) Using remote sensing and a cellular automata-markov chains-geomod model for the quantification of the future spread of an invasive plant: a case study of psidium guajava in isabela island, galapagos. Int J Geoinform 10(3):23–30

Barrena J, Nahuelhual L, Báez A, Schiappacasse I, Cerda C (2014) Valuing cultural ecosystem services: agricultural heritage in Chiloé island, Southern Chile. Ecosyst Serv 7:66–75

Bateman JB, Vitousek PM (2018) Soil fertility response to *Ulex europaeus* invasion and restoration efforts. Biol Invasions 20(10):2777–2791

Bradley BA (2014) Remote detection of invasive plants: a review of spectral, textural and phenological approaches. Biol Invasions 16(7):1411–1425

Breckling B, Pe'er G, Matsinos YG (2011) Cellular automata in ecological modelling. In: Modelling complex ecological dynamics, Springer, pp 105–117

Breiman L (2001) Random forests. Mach Learn 45:5–32

Cao J, Xu J, Pan X, Monaco TA, Zhao K, Wang D, Rong Y (2021) Potential impact of climate change on the global geographical distribution of the invasive species, Cenchrus spinifex (Field sandbur, Gramineae). Ecol Indicat. https://doi.org/10.1016/j.ecolind.2021.108204

CIREN (2003) Descripciones de suelos, materiales y simbolos. CIREN

Clements DR, Peterson DJ, Prasad R (2001) The biology of Canadian weeds: 112–*Ulex europaeus* L. Can J Plant Sci 81(2):325–337

Dunkerley D (1999) Cellular automata. Machine learning methods for ecological applications. Springer, pp 145–183

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of maxent for ecologists. Divers Distrib 17(1):43–57

Elliott-Graves A (2016) The problem of prediction in invasion biology. Biol Philos 31(3):373–393. https://doi.org/10.1007/s10539-015-9504-0

Esch T, Taubenböck H, Roth A, Heldens W, Felbier A, Schmidt M, Mueller AA, Thiel M, Dech SW (2012) Tandem-x mission-new perspectives for the inventory and monitoring of global settlement patterns. J Appl Remote Sens 6(1):061702

Espinar JL, Díaz-Delgado R, Bravo MA, Vilà M (2015) Linking azolla filiculoides invasion to increased winter temperatures in the doñana marshland (sw spain). Aquat Invasions 10(1):17

Fick SE, Hijmans RJ (2017) Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37(12):4302–4315

Gallien L, Douzet R, Pratte S, Zimmermann NE, Thuiller W (2012) Invasive species distribution models-how violating the equilibrium assumption can create new insights. Glob Ecol Biogeogr 21(11):1126–1136

Ghosh P, Mukhopadhyay A, Chanda A, Mondal P, Akhand A, Mukherjee S, Nayak S, Ghosh S, Mitra D, Ghosh T et al (2017) Application of cellular automata and markov-chain model in geospatial environmental modeling-a review. Remote Sens Appl Soc Environ 5:64–77

Goudarzi MA, Landry RJ (2017) Assessing horizontal positional accuracy of Google Earth imagery in the city of Montreal, Canada. Geodesy Cartography 43(2):56–65. https://doi.org/10.3846/20296991.2017.1330767

Gränzig T, Fassnacht FE, Kleinschmit B, Förster M (2021) Mapping the fractional coverage of the invasive shrub *Ulex europaeus* with multi-temporal Sentinel-2 imagery utilizing UAV orthoimages and a new spatial optimization approach. Int J Appl Earth Observ Geoinform 96(August 2020):102281. https://doi.org/10.1016/j.jag.2020.102281

Hackwell K (1980) Gorse: a helpful plant for regenerating native forest. Forest Bird 215:25–28

Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) dismo: species distribution modeling. https://CRAN.R-project.org/package=dismo, r package version 1.1-4

Hill DD (1949) Gorse control. Oregon Agricultural Experimental Station Bulletin 450

Hornoy B, Tarayre M, Hervé M, Gigord L, Atlan A (2011) Invasive plants and enemy release: evolution of trait means and trait correlations in *Ulex europaeus*. PLoS ONE 6(10):1–10. https://doi.org/10.1371/journal.pone.0026275

Hoyos LE, Gavier-Pizarro GI, Kuemmerle T, Bucher EH, Radeloff VC, Tecco PA (2010) Invasion of glossy privet (ligustrum lucidum) and native forest loss in the sierras chicas of córdoba, argentina. Biol Invasions 12(9):3261–3275

Huang Cy, Asner GP (2009) Applications of remote sensing to alien invasive plant studies. Sensors 9(6):4869–4889

Huang H, Zhang L, Guan Y, Wang D (2008) A cellular automata model for population expansion of Spartina alterniflora at Jiuduansha Shoals, Shanghai, China. Estuar Coast Shelf Sci 77(1):47–55

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, https://doi.org/10.1046/j.1365-2559.2002.1340a.x

Kattenborn T, Lopatin J, Förster M, Braun AC, Fassnacht FE (2019) Uav data as alternative to field sampling to map woody invasive species based on combined sentinel-1 and sentinel-2 data. Remote Sens Environ 227:61–73. https://doi.org/10.1016/j.rse.2019.03.025

Lesiv M, See L, Laso Bayas JC, Sturn T, Schepaschenko D, Karner M, Moorthy I, McCallum I, Fritz S (2018) Characterizing the spatial and temporal availability of very high resolution satellite imagery in google earth and microsoft bing maps as a source of reference data. Land 7(4):118

Li X, Yeh AGO (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int J Geogr Inf Sci 14(2):131–152. https://doi.org/10.1080/136588100240886

Liu X, Zhang A, Wang H, Liu H (2016) Using multi-remote sensing data to assess phragmites invasion of the detroit river international wildlife refuge. World J Eng 13:44

Lu ML, Huang JY, Chung YL, Huang CY (2013) Modelling the invasion of a central american mimosoid tree species (leucaena leucocephala) in a tropical coastal region of taiwan. Remote sensing letters 4(5):485–493

Mao D, Liu M, Wang Z, Li L, Man W, Jia M, Zhang Y (2019) Rapid invasion of spartina alterniflora in the coastal zone of mainland china: spatiotemporal patterns and human prevention. Sensors 19(10):2308

Merow C, LaFleur N, Silander JA Jr, Wilson AM, Rubega M (2011) Developing dynamic mechanistic species distribution models: predicting bird-mediated spread of invasive plants across northeastern north america. Am Nat 178(1):30–43

Müllerová J, Brůna J, Bartaloš T, Dvořák P, Vítková M, Pyšek P (2017) Timing is important: Unmanned aircraft vs. satellite imagery in plant invasion monitoring. Frontiers in plant science 8:887

Moss GR (1960) Gorse. a weed problem on thousands of acres of farm laud. New Zealand Journal of Agriculture 100(6):561–7

Norambuena H, Escobar S, Rodriguez F (2000) The Biocontrol of Gorse, *Ulex europaeus*, in Chile: A Progress Report. Proc of the International Symposium on Biological Control of Weeds 961(July 1999):955–961, https://www.researchgate.net/profile/Fernando_Rodriguez12/publication/237442332_The_Biocontrol_of_Gorse_Ulex_europaeus_in_Chile_A_Progress_Report/links/575ab24008aec91374a614e5.pdf

Pena-Gomez FT, Guerrero PC, Bizama G, Duarte M, Bustamante RO (2014) Climatic niche conservatism and biogeographical non-equilibrium in eschscholzia californica (papaveraceae), an invasive plant in the chilean mediterranean region. PloS one 9(8)

Phillips SJ, Dudík M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004:655–662. https://doi.org/10.1145/1015330.1015412

Pili AN, Tingley R, Sy EY, Diesmos MLL, Diesmos AC (2020) Niche shifts and environmental non-equilibrium undermine the usefulness of ecological niche models for invasion risk assessments. Sci Rep 10(1):1–18

Potere D (2008) Horizontal positional accuracy of google earth's high-resolution imagery archive. Sensors 8(12):7973–7981. https://doi.org/10.3390/s8127973

Pulighe G, Baiocchi V, Lupia F (2016) Horizontal accuracy assessment of very high resolution Google Earth images in the city of Rome. Italy. International Journal of Digital Earth 9(4):342–362

Rees M, Hill RL (2001) Large-scale disturbances , biological control and the. Journal of Applied of Ecology 38(Wilson 1994):364–377

Ren GB, Wang JJ, Wang AD, Wang JB, Zhu YL, Wu PQ, Ma Y, Zhang J (2019) Monitoring the invasion of smooth cordgrass spartina alterniflora within the modern yellow river delta using remote sensing. J Coastal Res 90(sp1):135–145

Rocchini D, Andreo V, Förster M, Garzon-Lopez CX, Gutierrez AP, Gillespie TW, Hauffe HC, He KS, Kleinschmit B, Mairota P et al (2015) Potential of remote sensing to predict species invasions: A modelling perspective. Prog Phys Geogr 39(3):283–309

Shepherd J, Lee W (2002) Satellite mapping of gorse at regional scales. New Zealand Plant Protection 55:95–98, https://doi.org/10.30843/nzpp.2002.55.3926

Somers B, Asner GP (2012) Invasive species mapping in hawaiian rainforests using multi-temporal hyperion spaceborne imaging spectroscopy. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6(2):351–359

Somodi I, Čarni A, Ribeiro D, Podobnikar T (2012) Recognition of the invasive species robinia pseudacacia from combined remote sensing and gis sources. Biological conservation 150(1):59–67

Sorbe F, Gränzig T, Förster M (2023) Evaluating sampling bias correction methods for invasive species distribution modeling in maxent. Ecological Informatics p 102124, https://doi.org/10.1016/j.ecoinf.2023.102124, https://www.sciencedirect.com/science/article/pii/S157495412300153X

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteor Soc 93(4):485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Ulex L in GBIF Secretariat (2022) Gbif backbone taxonomy. checklist dataset. https://doi.org/10.15468/39omei

Visser V, Langdon B, Pauchard A, Richardson DM (2014) Unlocking the potential of Google Earth as a tool in invasion science. Biol Invasions 16(3):513–534. https://doi.org/10.1007/s10530-013-0604-y

Wang A, Chen J, Jing C, Ye G, Wu J, Huang Z, Zhou C (2015) Monitoring the invasion of spartina alterniflora from 1993 to 2014 with landsat tm and spot 6 satellite data in yueqing bay, china. PloS one 10(8)

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst 36:519–539

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Jonathan Davies T, Grytnes JA, Harrison SP et al (2010) Niche conservatism as an emerging principle in ecology and conservation biology. Ecol Lett 13(10):1310–1324

Yu L, Gong P (2012) Google Earth as a virtual globe tool for Earth science applications at the global scale: progress and perspectives. Int J Remote Sens 33(12):3966–3986. https://doi.org/10.1080/01431161.2011.636081

Zabkiewicz JA (1976) The ecology of gorse and its relevance to new zealand forestry. In: Proceedings of the F.R.I. Symposium No. 18: The Use of Herbicides in Forestry in New Zealand, 1975., pp 63–68

Zheng Z, Tian B, Zhang L, Zou G (2015) Simulating the range expansion of spartina alterniflora in ecological engineering through constrained cellular automata model and gis. Mathematical Problems in Engineering 2015