Remote Sensing Applications: Society and Environment 33 (2024) 101121

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

A hybrid modelling approach for detecting seasonal variations in inland Green-Blue Ecosystems

Bruna Almeida [a, c, *], Pedro Cabral [a, b, **]

[a] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312, Lisboa, Portugal
[b] School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[c] Centre for Ecology, Evolution and Environmental Changes (cE3c), Azorean Biodiversity Group (GBA), University of the Azores (UAc), Faculty of Sciences and Technology (FCT-UAc), Rua Mãe de Deus, Campus Universitário de Ponta Delgada, Edifício do Complexo Científico, 9500-321, Ponta Delgada, Portugal

ARTICLE INFO

Keywords:
Remote sensing
Machine learning
Land use classification
Aquatic ecosystems
Terrestrial ecosystems
Spatiotemporal analysis

ABSTRACT

Deforestation, environmental pollution, and the overexploitation of resources, in addition to the Earth's natural cycles, are scaling up the impacts of climate change in the provision of Ecosystem Services (ES). Green-Blue Ecosystems (GBE) are impacted by climatic conditions, topography, and water presence. Data-driven modelling techniques may effectively capture the effects of seasonal variations while modelling natural ecosystems. This research proposes a hybrid modelling approach that combines Deep Learning and traditional Machine Learning, Sensitivity Analysis and Feature Importance Evaluation (FIE) to investigate seasonality effects on mapping GBE. The models, built using satellite imagery from the Spring and Summer seasons of the Mediterranean climate zone, included spectral indices, topography (DEM), and groundwater depth (GD). The model that best suited the analysis was selected using sensitivity tests and hyperparameter optimization. The study shows that land cover classes of transitional woodland shrubs, inland marshes, cultivated land parcels, and watercourses are better classified in the Spring, with an accuracy of 0.814. FIE indicates that spectral indices are the most important predictors for detecting green ecosystems in both seasons. Additionally, DEM and GD are the most relevant predictors to classify watercourses in the Summer. An analytical examination of the input data and hyperparameter settings facilitates understanding of models' behaviour while improving models' prediction. This research provides an advanced understanding of the effects of seasonal variations on the status of GBE and enhances understanding of modelling ES in areas with a growing need for changes in land use and high water supply demand.

1. Introduction

To ensure the long-term viability of Ecosystem Services (ES), it is critical to identify drivers of change and to gather information for their assessment (Meddens et al., 2022). The value of ES is favourably connected with the extent of vegetation and water but negatively correlated with the area of bare land (Wang et al., 2021a, b). As competition for land use increases, natural resource manage-

* Corresponding author. Centre for Ecology, Evolution and Environmental Changes (cE3c), Azorean Biodiversity Group (GBA), University of the Azores (UAc), Faculty of Sciences and Technology (FCT-UAc), Rua Mãe de Deus, Campus Universitário de Ponta Delgada, Edifício do Complexo Científico, 9500-321, Ponta Delgada, Portugal.
** Corresponding author. School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, 210044, China.
E-mail addresses: balmeida@novaims.unl.pt (B. Almeida), cabral@nuist.edu.cn (P. Cabral).

https://doi.org/10.1016/j.rsase.2023.101121
Received 28 March 2023; Received in revised form 29 October 2023; Accepted 3 December 2023
Available online 10 December 2023

ment faces new challenges in maintaining GBE functioning and conditions (Doody et al., 2017). Urban sprawl, agriculture practices, and climate change put continuous pressure on natural systems around the world, with ongoing biodiversity losses (Skidmore et al., 2021).

Terrestrial and Aquatic Ecosystems, such as forests, croplands, wetlands, rivers, and lakes provide the most important services in terms of ecosystem provisioning and regulation (Mugiraneza et al., 2019). Commonly referred to as Green-Blue Ecosystems (GBE), the status and functioning of terrestrial and aquatic ecosystems depend on climatic variables, such as precipitation, evapotranspiration, solar radiation, wind, and other variables included in the carbon and water cycles (Gwal et al., 2020). GBE health and behaviour are also influenced by terrain conditions and the presence of groundwater, such as phreatophyte vegetation that requires water on a permanent or episodic basis (Dwire et al., 2018). Seasonality has a substantial impact on these ecosystems during dry periods with little rain, high temperatures and evapotranspiration. As a result, seasonal wetlands dry out while permanent wetlands decrease in extent, vegetation heterogeneity and crop production diminish as well as water yield (Kundu et al., 2022).

The association between aquatic and terrestrial ecosystems are usually observed within rivers that are surrounded by vegetation (Bertrand et al., 2012), and where there is recharge and discharge of groundwater from and to water bodies (Eamus and Ray, 2006). It is strongly influenced by biophysical parameters, such as climate, hydrology, topography, and human activities (Yang and Liu, 2020). Groundwater over-abstraction, land-use conversion, climate change, and climate variability are important drivers to explain changes in GBE (Wang et al., 2021a, b).

Satellite Earth Observation (SEO), Geographical Information Systems (GIS) and Machine Learning (ML) provide environmental managers and decision-makers with a wide range of tools to assess and monitor ES (Alarcon Blazquez et al., 2023). SEO technology gives a more in-depth perspective for studying land use and land cover change on the earth's surface, as well as a big source of data for evaluating ES values, which is critical for environmental protection and national ecological security (Wang et al., 2021a, b). The integration of SEO data with ML is one of the most promising approaches to developing timely and accurate environmental monitoring systems (Reddy, 2021). Traditional ML offers effective techniques for modelling relationships in data with vast feature spaces (Domingos, 2012), but Deep Learning (DL) has been suggested as a better option when it comes to the extraction of significant characteristics from complex datasets (Chang et al., 2019).

The research on the changes in the spatiotemporal aspects of land use and landscape patterns can provide a thorough understanding of the dynamics of the Earth's surface caused by human activities, natural forces, and biological variables over time and space (Chen et al., 2021). Therefore, the lack of guidelines for implementing multi-season satellite-based environmental models is a challenge for those who support decision-making with geospatial information (Doody et al., 2017). More research is needed to determine the association between GBE and biophysical factors utilizing remote sensing data since environmental variables influence the spectral response of ecosystems (Sahana et al., 2022). By reviewing the literature, we identified a common concern about advancing the knowledge of methodologies, tools and strategies that improve the development of seasonal satellite-based data-driven models of GBE, as most research has been focused on specific ecosystem types (green or blue) but not on understanding their relationship.

Recent studies combined SEO and ML to delineate GBE while investigating seasonality effects. Barron et al. (2014) proposed a multi-season method to predict vegetation and waterbodies that can persist through a prolonged dry period, in the Ellen Brook region, an important groundwater basin in Western Australia's south-west. The region is characterized by a Mediterranean climate with GBE associated with localised groundwater, diffuse discharge zones, and riparian vegetation. Mpakairi et al. (2022a, b) investigated the effects of land use conversions and climate change on vegetation heterogeneity in the Khakea-Bray, a transboundary aquifer primarily recharged by precipitation, located between southwestern Botswana and South Africa. The temperature and rainfall in the region are typical of semi-arid savannahs, with high evapotranspiration rates, minimal rainfall during the wet season, and high temperatures during the dry season. Kundu et al. (2022) developed an approach to explore seasonality effects in wetlands in a transboundary river basin (Punarbhaba River) including India and Bangladesh by assessing hydro-period, water depth, and water presence consistency. The study region has a subtropical monsoon climate with seasonal wet and dry spells as well as hot and cold thermal spells. The annual rainfall averages 1500 mm, with 82% falling during the monsoon season (June to September), resulting in the greatest spatial extension of wetlands found during this season. Zhang et al. (2022) classified GBE to assess the status of saline wetlands in the groundwater discharge zones of the Great Plains Aquifer system in eastern Nebraska. In this area, the growing season of wetland vegetation is defined as continuous temperatures above 0 °C from April to October, with the highest water availability in the Autumn and the lowest in the Summer. However, to the best of our knowledge, there is no research investigating the seasonal influence in mapping GBE and estimating their magnitude and relationship in the Lower Tagus Aquifer system. We argue that adopting a hybrid modelling strategy that includes DL and classical ML, as well as hyperparameter optimization, sensitivity analysis, and Feature Importance Estimates (FIE), could close this gap.

The rationale of this study relies on what previous research noted concerning dealing with large datasets and combining techniques to take advantage of each method. When handling large datasets, it is crucial to employ strategies that ensure efficient management and analysis and select optimized algorithms that allow for more efficient processing. Thus, the main aim of this research is to overcome these concerns by developing a hybrid modelling approach to map GBE and estimate their relationship considering seasonality. We propose a two-step modelling strategy that helps to manage and analyze effectively large geospatial datasets while enhancing overall efficiency in GBE mapping workflows.

In the context of climate change, Portugal is recognized as a hotspot among the most vulnerable European countries (Novo et al., 2018). Recent studies have shown evidence of climatic changes, such as the long periods of drought recorded in 1990, 2004/2005 and 2012 (Coelho et al., 2013; de Lima et al., 2015; Almeida and Cabral, 2021). The more frequent occurrence of these events is increasing the severity of seasonality effects on GBE and compromising the provision of services such as freshwater supply, and consequently crop production, and carbon storage and sequestration (Rijal et al., 2021). The selected study area is within a vulnerable groundwater

body (Novo et al., 2018), the Lower Tagus Aquifer system located in the surroundings of Lisbon, the most populated and dynamic area of mainland Portugal (Eurostat, 2023).

Findings from this research can contribute to improving understanding of how these natural systems can be modelled while considering land use dynamics, groundwater depth, and topography to investigate the impact of seasonality on mapping GBE. The purpose of this study is to unveil the potential of SEO and ML in monitoring the response of GBE to seasonal changes and assessing land cover dynamics.

2. Material and methods

2.1. Study area

Climate change, drought intensification, deforestation, erosion, desertification, urbanization, rural abandonment, and uncontrolled tourist development are jeopardizing Mediterranean ecosystems (Chrysafis et al., 2020). The Mediterranean climate is known for its particular regional characteristics: large seasonal contrast in temperature and rainfall (hot dry summers and mild winters), strong wind systems, intense precipitation, and Mediterranean cyclones (Li et al., 2012).

Like other Mediterranean regions, Portugal is susceptible to extreme climatic events, such as droughts and desertification effects and is one of the European countries that are most reliant on water resources that originate outside their territory (Almeida and Cabral, 2021). The Lower Tagus Aquifer (study area) is in the downstream area of the international river basin districts PTRH5A in Portugal and its correspondence ES030 in Spain (Fig. 1). These catchments constitute an important source of water for both Mediterranean countries.

With a length of roughly 1000 km and a catchment area of 80,630 km^2, the Tagus River begins at a height of about 1600 m in eastern Spain and flows into the Atlantic Ocean near Lisbon (Vis and Kasse, 2009). The river is known for its large size, productivity, quality of its waters, and being responsible for the region's economic and demographic development. The river's proximity to agricultural land facilitates water supply for agricultural intensification (Almeida and Cabral, 2023).

Temporary and permanent crops such as rice fields, and other inundated croplands, vineyards, fruit trees and olive groves are well-developed in the region. Forest and seminatural areas are mostly represented by mixed forest and transitional woodland/shrub, and water and groundwater interactions support the maintenance of ephemeral and permanent wetlands (Ramos et al., 2017).

Precipitation in the Tagus basin has a seasonal cycle, large inter-annual variability, and periodic episodes of extremely wet or extremely dry years (Vis et al., 2010). Spring (March through May) contributes 24% of annual precipitation, Summer (June through August) 6%, Autumn (September through November) 28%, and Winter (December through February) 42% (Belo-Pereira et al., 2011). This strong seasonality in precipitation results in maximum rainfall extreme occurrences (potentially leading to floods) and relatively frequent extended drought spells (Espírito Santo et al., 2014). Long drought episodes typically cause agricultural damage and alter water resources and availability for many uses (Oliveira et al., 2005).

2.2. Data

Sentinel-2 satellite imagery data with a spatial resolution of 10 m were downloaded from the Copernicus Open Access Hub. The modelling process had as input two multispectral images from the Spring (March 26, 2018) and Summer (August 23, 2018) seasons of the Mediterranean climate zone.

Spring and Summer were selected since they are typically the seasons with the lowest average contributions of precipitation, and as a result, less soil water-holding capacity (Belo-Pereira et al., 2011). This is required to distinguish waterbodies from flooded areas (Barron et al., 2014), as major floods that have occurred in the study area since 1980 were throughout the Autumn-Winter season (November to February) (Espírito Santo et al., 2014), and detect the seasonal variation in the greenness of vegetated land (particularly in the Summer) (Doody et al., 2017). The dates were selected based on the percentage of cloud cover, which was less than 5%.

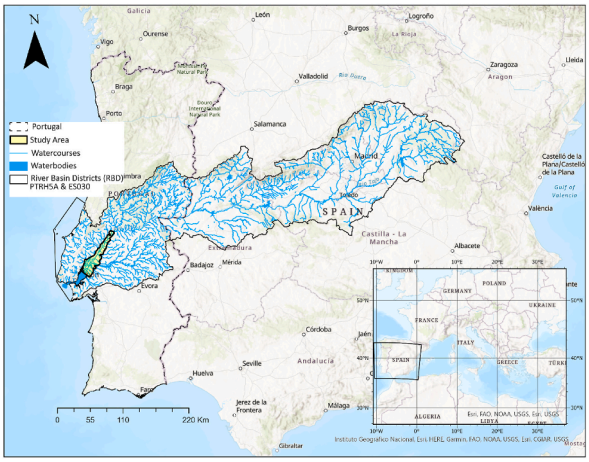

Fig. 1. Study area within the transboundary river basins – PTRH5A (Portugal) and ES030 (Spain).

Satellite-derived indices such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), and Normalized Difference Water Index (NDWI), were calculated. The NDVI was included as a predictor to highlight and distinguish vegetation of different types and at different stages of its evolutionary seasonal cycle (Rouse et al., 1973). The NDMI was employed as it is an indicator of forest change and water stress that enables capturing the moisture differences between land use types (Silleos et al., 2008). Finally, NDWI was used to distinguish water and soil (Gao, 1996) capturing water only in the areas of the deepest channels and with the cleanest water (Ji et al., 2015). The Sentinel-2 derived indices were computed using the following arithmetic functions of bands: (1), (2), and (3):

$$NDVI = \frac{Band\ 8 - Band\ 4}{Band\ 8 + Band\ 4} \tag{1}$$

$$NDMI = \frac{Band\ 8 - Band\ 11}{Band\ 8 + Band\ 11} \tag{2}$$

$$NDWI = \frac{Band\ 3 - Band\ 8}{Band\ 3 + Band\ 8} \tag{3}$$

The reference data for the first-stage modelling is based on the Corine Land Cover (CLC) map of 2018 (spatial resolution of 100 m) (Copernicus Programme, 2023), considering subclasses (CLC Level 3) of agriculture, forest and seminatural areas, wetlands, and water bodies of the CLC Level 1. Surface water and groundwater bodies datasets are from the European Environment Agency – EEA (EEA, 2023) in a version released in the same year. The Digital Elevation Model (DEM), with a spatial resolution of 30 m, is provided by NASA (NASA, 2023) through the Shuttle Radar Topography Mission (SRTM), and the depth of groundwater is derived from the National Information System of Hydric Resources - SNIRH (SNIRH, 2023).

Min-max normalization was applied to the original data to scale the predictors in the range (0, 1) while keeping the relationships between them (4):

$$Xs = \frac{X - Xmin}{Xmax - Xmin} \tag{4}$$

where Xs represents the scaled value of X, $Xmin$ represents the minimum value of X, and $Xmax$ represents the highest value of X.

2.3. Methods

It is crucial when handling large datasets to employ strategies that ensure efficient management and analysis and implement optimized algorithms that allow for more efficient processing (Nikparvar and Thill, 2021). Besides, instead of loading the entire dataset into memory at once, a good practice is to break it down into manageable parts, or sub-setting the study area.

Automatic classifications with DL can take several hours to complete because training entails the generation of Convolutional Neural Networks (CNN) from the training and validation datasets, requiring the implementation of parallel processing techniques to speed up the analysis. For that, an NVIDIA Graphics Processing Unit (GPU) was configured before training to distribute the workload across multiple cores allowing for faster computation and analysis of the data. The graphic card used in this task is the GEFORCE RTX30 with 16 GB of dedicated memory.

The present approach proposes a hybrid method combining DL and traditional ML algorithms in a two-step process to take advantage of the strengths of each technique. Fig. 2 shows the flowchart diagram of the modelling process.

The goal of the first stage of modelling is to classify land use by seasons with maximum accuracy. Multiclass and binary classifications will be performed using Sentinel-2 bands (2, 3, 4 and 8) and DL. Sensitivity analysis tests and hyperparameter tuning will be used to choose the best models.

In the second stage, the main objective is to classify green and blue ecosystems (GE and BE) by seasons, using as input data the binary classification of GE and BE, spectral indices (NDVI, NDWI and NDMI), topography and depth of groundwater (GD). The final models will be presented and discussed with a performance comparison between ML methods. Exploratory analysis and FIE will be carried out to estimate the relationship between predictors of GE and BE by seasons, and then assess which variables contribute the most to classifying GBE. This strategy allows to work with subsets of the data, reducing memory usage, and improving processing speed, and overall efficiency.

2.3.1. First-stage: DL modelling

DL via neural networks uses stacked layers to enhance learning by building up representations of the input data in consecutive layers (Guo et al., 2021). To deal with data that needs to preserve its sequence, CNN is the recommended method (Zeiler and Fergus, 2013). This is a technique that convolves the input data to combine successive values to extract higher-level features (LeCun et al., 2015). It has also shown excellent performance in extracting objects from images through down-sampling (Lobert et al., 2021). Utilizing deeper and, therefore, more powerful networks only became possible in recent years through methodological and technical advancements, such as better training algorithms and more powerful computers (A and S, 2023).

In supervised learning, the training data is known and labelled, allowing for direct improvement of the learner's performance. Labels for the training dataset were created based on the CLC 2018 nomenclature guidelines, considering the land use classes of agriculture, forest and seminatural areas, wetlands, and water bodies included in the CLC level one. The artificial surface areas were used as masks.

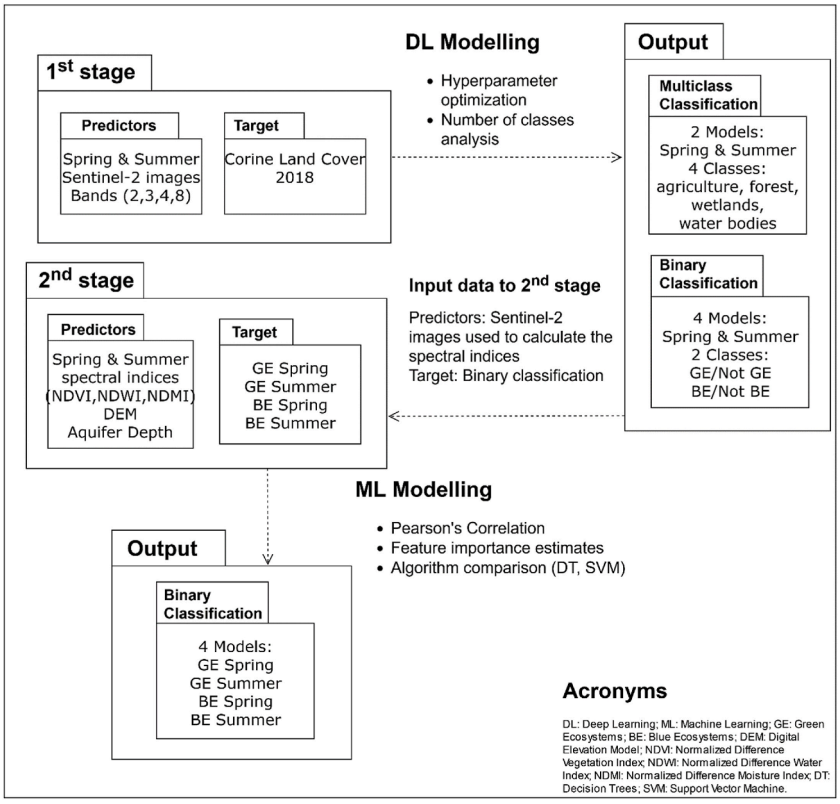

Fig. 2. Overview of the methodology development.

Sensitivity analysis is carried out to understand the effects and behaviours in the outputs, testing different input values and hyperparameters of the model. The task started creating the training datasets in ArcGIS Pro 2.9.0 (ESRI, 2023) using the tool *Export Training Data for Deep Learning* to convert the data into a format that can be used in the DL packages. The tool converted the labelled vector data (land use classes) into classified image chips using the metadata format of *Classified Tiles*. The multispectral images (a composite of four bands 2, 3, 4 and 8) generated image chips containing the respective class samples from the labelled vector. A few values for image chip size, *Tile Size* and *Stride* (with and without) were tested to find the best model. This tiling strategy is important for large images, to keep the context in the border of the tiles (Zeiler and Fergus, 2013).

The tool *Train Deep Learning Model*, which accepts as input the output folder of the preceding program, was used for training. Model parameters, such as *Max Epochs* and *Batch sizes* were tested with few values to fit the model. The *Model Type* is the U-Net approach used for pixel-based classification. The U-Net architecture works as an encoder network followed by a decoder (Ronneberger et al., 2015), with the learning occurring at different stages of the encoder (Zeiler and Fergus, 2013). The selected *Backbone model* is the pre-trained classification network *ResNet-34* trained on the Imagenet Dataset (Krizhevsky et al., 2012). The dataset contains more than 1 million images and is 34 layers deep. The use of a pre-training dataset is a method known as Transfer Learning. The *Learning Rate* is set to be extracted from the learning curve during the training process, and the *Validation size* is 20%.

Classification models need to be evaluated regarding quality and accuracy (Stehman and Foody, 2019). The following metrics are calculated to validate the prediction power of all classification tasks: accuracy, recall, precision, and F1-score. The accuracy is the fraction of correct prediction, recall measures how many of the true positive samples were correctly predicted by the model, while precision measures how many negative samples were not labelled as positive. As a weighted harmonic mean factor, F1-score is associated with precision and recall, and was the guide to selecting the best model, as it is the best option to measure model quality in imbalanced classification problems (Zhang and Urbanowicz, 2020). The model's output files are the trained model saved in a PyTorch format, a file containing the model's definition (tile size, classes, and model type), and an HTML file with details about the model's performance.

2.3.2. Second-stage: ML modelling

The DL model output is utilized as reference data for the second stage of modelling to differentiate between aquatic and terrestrial ecosystems (GBE) by seasons, with spectral indices (NDVI, NDWI, NDMI), elevation (m), and aquifer depth (GD) (m) as predictors.

This process includes data preparation, exploratory analysis, cross-validation (CV) partitioning, FIE, and accuracy evaluation. Decision Trees (DT) and Support Vector Machines (SVM) were implemented using scikit-learn (Pedregosa et al., 2011). DT divides a dataset into increasingly smaller subgroups using the same splitting decision iteratively (Zhang et al., 2019). It can handle nonlinear relationships and mixed predictor categories by default, and it is resistant to outliers and collinearity effects (Osborne and Alvares-

Sanches, 2019). The disadvantage of DT is related to accuracy and instability as a result of error propagation down via consecutive splits in the tree (Breiman, 1984).

The SVM algorithm classifies data by establishing an optimal hyperplane that can divide training data into a predefined number of classes (Boser et al., 1992). The algorithm is based on the principle of structural risk minimization overcoming the problem of overfitting (Wang et al., 2021a, b). SVM draw attention due to their robustness, accuracy, and suitability for small training data. However, SVM take longer to perform, particularly with large datasets (Khwarahm, 2021).

To avoid any data leakage before completing other processes, a 10-fold stratified data partition (CV) is carried out for both ML models. Two filter algorithms, the mutual information (MI) (Peng et al., 2005) and MultiSURF (Urbanowicz et al., 2018) (a Relief-based feature selection algorithm) are used to score the importance of a feature based on how well a feature's value distinguishes samples that are like one another but belong to distinct classes (Pudjihartono et al., 2022). MI is proficient at analysing univariate associations between a feature and the target, while MultiSURF is sensitive to feature interactions without evaluating combinations of features and examining only individuals very near and very far from one another (Urbanowicz et al., 2022).

The same set of functions used to assess models' accuracy in the first stage is used to evaluate predictions in this step, except the ROC-AUC which is the metric added in this stage. A Receiver Operating Characteristic (ROC) curve is a graph that depicts a classification model's performance across all categorization levels. The AUC (area under the ROC curve) is a function of recall and translates models' accuracy. AUC values close to one mean better-performing classifiers and higher discrimination ability (Mouta et al., 2021).

Statistical comparisons of algorithm performance applied Kruskal-Wallis one-way analysis of variance to determine if any ML algorithm or datasets yielded significantly different performance than the others for each evaluation metric (Urbanowicz et al., 2022). For any metric where a significant difference is observed, Mann-Whitney U-tests and Wilcoxon Rank test will be calculated to identify which algorithm yielded significantly better or worse performance. The Kruskal-Wallis (Kruskal and Wallis, 1952) is a nonparametric statistical test that compares the differences in a single, non-normally distributed continuous variable between three or more independently sampled groups. The Mann-Whitney *U* test compares two groups based on a single ordinal variable with no specified distribution (Mann and Whitney, 1947; Wilcoxon, 1945). The null hypothesis states that the n populations sampled have all the same average, whereas the alternative hypothesis states that at least one sample comes from a distribution with a different average.

3. Results and analysis

3.1. First-stage modelling

A land cover classification mapping GBE in the study area is the main outcome of this first-stage modelling. Sensitivity analysis is an integral part of the models' development, involving an analytical examination of the model's settings such as *Max Epochs*, *Batch size, Tile size, Stride,* and the number of classes.

The first sensitivity tests considered the maximum number of land use classes presented in the study area (17 classes), maximum values for *Tile size* and *Stride*, and minimum values for *Max Epochs* and *Batch size.* The best results were obtained when the image chip size in both dimensions (x, y) (*Tile size*) and the distance that separate tiles (*Stride*) were kept to a minimum, combined with a maximum number of iterations (*Max Epochs)* and *Batch size*. Keeping the *Stride* equal to the *Tile size* guarantees no overlapping in the convolutions. Considering all the existing classes in the study area (a total of 17), the model detected six of them. The other 11 classes did not have a representative number of images and features to be modelled with good accuracy. Table 1 shows the land cover classes considered in the first classification task, and those final classes selected by the best sensitivity test.

Table 1
Land cover classes used in the classification task, grouped by sensitivity tests. Legend: CLC – Corine Land Cover.

Sensitivity Test	CLC Level 1	CLC Level 3	Label	Detected (*)
First	2	211	Non-irrigated arable land	
		212	Permanently irrigated land	*
		213	Rice fields	
		221	Vineyards	
		231	Pastures	
		242	Complex cultivation patterns	*
		243	Land principally occupied by agriculture, with significant areas of natural vegetation	
		244	Agro-forestry areas	
	3	311	Broad-leaved forest	*
		312	Coniferous forest	
		313	Mixed forest	
		321	Natural grasslands	
		324	Transitional woodland-shrub	*
		331	Beaches, dunes, sands	
	4	411	Inland marshes	*
	5	511	Watercourses	*
		512	Water bodies	
Best	2	242	Complex cultivation patterns	*
	3	324	Transitional woodland-shrub	*
	4	411	Inland marshes	*
	5	511	Watercourses	*

Fig. 3 shows the ground truth (first column) and predictions (second column) of the summer and spring models' output. These are the results of the first set of sensitivity analyses classifying 17 land cover classes, generating image chips with a size of 224 × 224 px, number of iterations equal to 40 and 100, and batch sizes of 16 and 32.

Decreasing the size of the tiles did not improve the classification accuracy, considering 17 classes. To overcome this, we tested excluding the small classes keeping the most representative samples. After testing with 12, 8, 6, and 4 classes, the best results were considering the following: 242 - complex cultivation patterns, 324 - transitional woodland shrubs, 411 - inland marshes, and 511 - watercourses. Land cover class 242 includes cultivated land parcels with different cultivation types (temporary and permanent crops and pastures) referred to as agricultural areas; 324 represents areas of the natural development of forest formations, transitional bushy and herbaceous vegetation, referred to as woodlands; 411 refer to the wetlands including herbaceous and high floating vegetation in areas usually flooded in winter, or saturated by fresh water all the year; and 511 represents natural or artificial surface watercourses serving as water drainage channels. Fig. 4 highlights the influence of tile size (64 and 32 px) by comparing ground truth and predictions for both seasons, considering four land cover classes, and keeping the maximum number of iterations (100) and batch size (64).

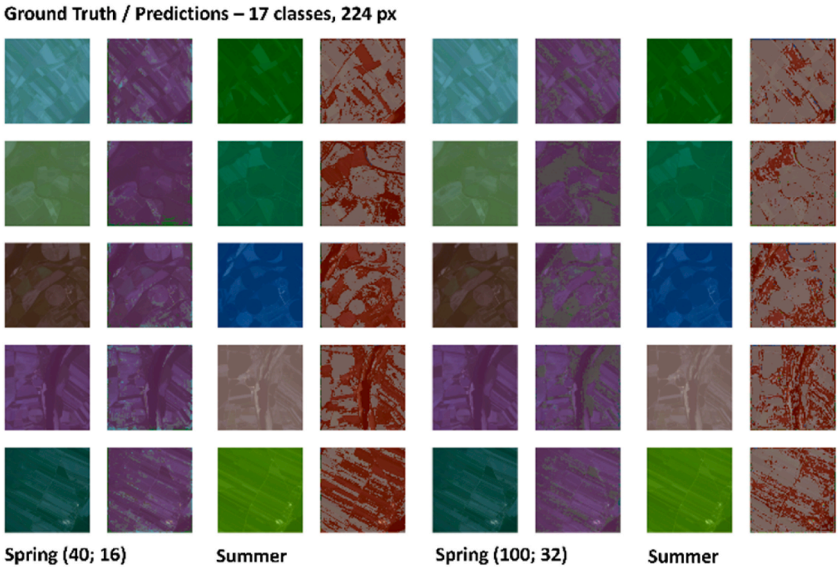

Fig. 3. Ground truth and predictions of two different models for both seasons. The models were tested with 17 classes, image size of 224 × 224 px, number of iterations equal to 40 and 100, and batch size of 16 and 32. Legend: Season (number of iterations; batch size).

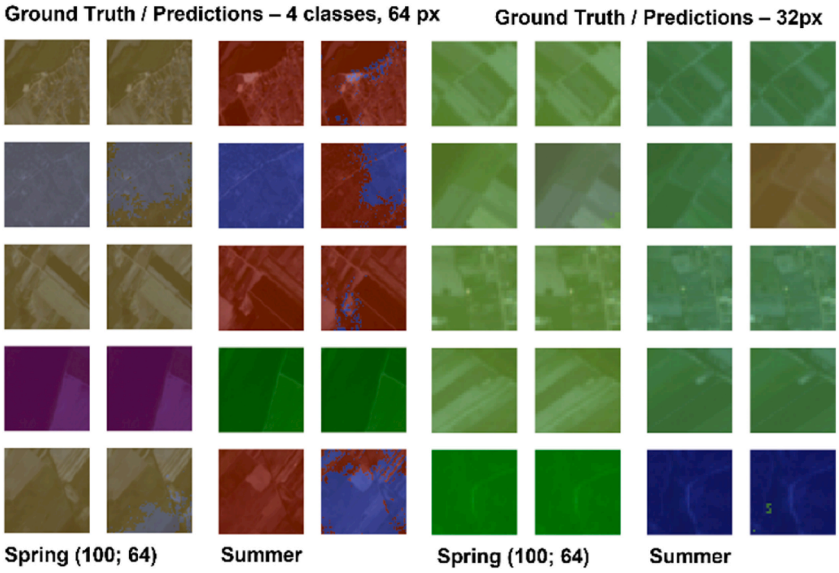

Fig. 4. Ground truth and predictions of two different models for both seasons. The models were tested with 4 classes, image sizes of 64 × 64 px and 32 × 32 px, number of iterations equal to 100, and a batch size of 64.

The accuracy of the best multiclass models (four classes, 32 × 32 px image size, 100 epochs and 64 batches, both seasons) is presented in.

Table 2. The models' accuracy was greatest for the spring model, with agricultural areas and watercourses showing the highest results in all metrics, followed by forests and wetlands. The model for the Summer season did not detect wetlands, but classified forests with greater precision than in Spring. The LR slice notation (low; high) represents the combination of low and high values referring to the LR for the first layer and the remaining trainable part of the model, respectively. The LR changes with every iteration. The low and high LR were the same for both seasons, explained by both models' settings configured to complete the iterations, even when the system stopped learning.

The final procedure of this stage was building a binary model mapping GBE. A binary classification task is performed to classify complex cultivation patterns, transitional woodland shrubs, inland marshes, and watercourses into GE and BE by seasons to be used as reference data for the second-stage modelling. The models included a set of 7086 images (32 × 32 px) with 139,156 features, *Tile size* (32 px) and *Stride* (32 px). The best models were selected after testing the number of iterations and batch size. Table 3 shows the best results for the binary classifications. The Spring model (100; 64) showed a better performance than the Summer model for all metrics and tests. The accuracy for the best Spring model (100; 64) is 0.876 and 0.866 for the Summer model (100; 64).

Sensitivity analysis showed that the Spring model did not have a considerable improvement after the 20th iteration, even after increasing the Batch size. While the Summer model needed five times more iterations to show some improvements. Nevertheless, the optimal LR extracted from the learning curve during training was the same for both Summer models. Fig. 5 shows the evolution through the iterations of the two models' evaluation functions, accuracy, and F1-score. After the 10th epoch, the accuracy for both seasons did not improve much and had very small variations, while the F1-score kept increasing, and did not stabilize before the 80th iteration.

The training and validation losses are shown in Fig. 6. The graph displays the amount of error as the model trained over time. Even though the Spring model has more deviation losses for training and validation (0.050; 0.055) than the Summer (0.037; 0.054), the mean value of losses in the Spring model was (0.390; 0.343) against (0.445; 0.368) for the Summer model. The larger values for the latter are explained by the higher losses verified at the beginning of modelling.

Table 2
DL models' performance evaluation summary, considering seasons, the maximum number of iterations and batch sizes (Max Epochs; Batch size). Legend: Land cover classes: 242 – agricultural areas, 324 – woodlands, 411 – wetlands, and 511 – watercourses, LR – Learning Rate.

Type	Model (100; 64)	Class	Accuracy	F1-score	Recall	Precision	LR (low; high)
Multiclass	**Spring**	**242**	0.814	0.872	0.932	0.819	1.0965e-05; 1.0965e-04
		324		0.627	0.518	0.795	
		411		0.481	0.354	0.752	
		511		0.785	0.770	0.801	
	Summer	**242**	0.790	0.862	0.950	0.789	1.0965e-05; 1.0965e-04
		324		0.540	0.402	0.818	
		411		0	0	0	
		511		0.740	0.717	0.766	

Table 3
DL models' performance evaluation summary, considering seasons, number of iterations and batch size (Max Epochs; Batch size). Legend: LR - Learning Rate.

Type	Model	**Accuracy**	**F1-score**	**Recall**	**Precision**	**LR (low; high)**
Binary	Spring (100; 64)	0.876	0.650	0.576	0.740	9.120e-06; 9.120e-05
	Summer (100; 64)	0.866	0.622	0.543	0.729	7.586e-06; 7.586e-05
	Spring (20; 8)	0.876	0.648	0.575	0.743	2.754e-05; 2.754e-04
	Summer (20; 8)	0.846	0.529	0.434	0.677	7.586e-06; 7.586e-05

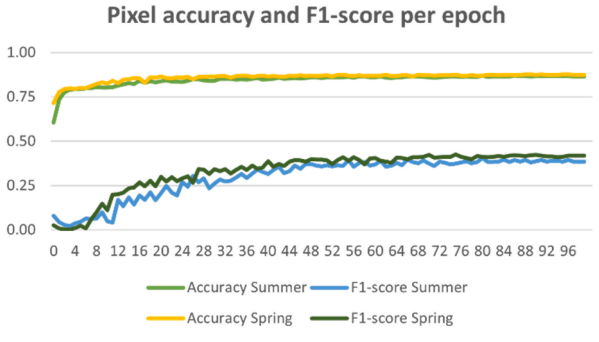

Fig. 5. Evolution of models' performance through iterations for both seasons.

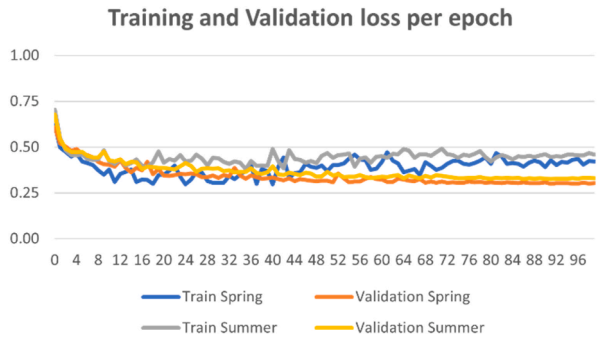

Fig. 6. Training and validation loss per epoch for the Spring and Summer models.

3.2. Second-stage modelling

3.2.1. Exploratory analysis

The exploratory analysis draws evidence about the relationship between predictors by calculating Pearson's correlation coefficient. The analysed datasets have 14,385 instances, five numerical features (NDMI, NDVI, NDWI, DEM and GD), and no missing values. A scatterplot matrix with correlations and variable distribution for each ecosystem type by season can be found in (Appendix A. Supplementary Material - Fig. S1). Table 4 summarizes the results of Pearson's correlation by ecosystem type (GE and BE) by season.

For all ecosystem types and seasons, NDMI showed a high positive correlation with NDVI, and NDWI showed a high negative correlation with NDVI and NDMI. The values for DEM followed a pattern of negative correlation with NDMI, and NDVI, and positive with NDWI, except for BE in the Summer, where the associations oppositely changed. The relationship between DEM and GD was negative for all seasons but for BE in the Summer, the value was twice higher (−0.747). The correlation between GD and the spectral indices also followed a pattern, except for BE in Summer, where the values reversely changed.

The GE Spring model revealed a strong positive association between NDMI and NDVI (0.803), but negative correlations between NDWI and NDMI (−0.732), and NDWI and NDVI (−0.969). The correlation of NDMI and NDWI with topography was at least three times more in the Summer than in the Spring, and NDVI double increased between these seasons.

3.2.2. Estimates of feature importance

Fig. 7 compares the FIE between seasons and ML algorithms, for the GE. In the Spring, the DT model estimated NDMI (0.19) as the most representative feature, while SVM predicted NDVI (0.23), but also NDWI (0.18) and NDMI (0.17). In the Summer, SVM and DT estimated NDWI as the most relevant feature and all the other predictors were also considerably weighted, except DEM.

For aquatic ecosystems in the Spring season (Fig. 8), SVM highlighted NDWI (0.18) and NDVI (0.16) as the most relevant variables, while DT homogeneously weighted all predictors, except aquifer depth. In the SVM Summer model, elevation (0.30) and groundwater depth (0.25) were the most important features estimated by those two algorithms. DT excluded NDWI as a predictor in the classification of the BE and did not give much consideration to NDMI and NDVI.

Table 4
Pearson's correlation coefficient by ecosystem type (GE and BE) by season. Legend: GE – Green Ecosystems, BE – Blue Ecosystems, DEM – Digital Elevation Model, GD – groundwater depth.

Ecosystem/Season	Predictor	Pearson Correlation				
GE Spring		**NDMI**	**NDVI**	**NDWI**	**DEM**	**GD**
	NDMI		0.803	−0.732	−0.255	0.123
	NDVI			−0.969	−0.468	0.196
	NDWI				0.467	−0.204
	DEM					−0.388
GE Summer		**NDMI**	**NDVI**	**NDWI**	**DEM**	**GD**
	NDMI		0.795	−0.717	−0.617	0.185
	NDVI			−0.964	−0.900	0.293
	NDWI				0.966	−0.334
	DEM					−0.386
BE Spring		**NDMI**	**NDVI**	**NDWI**	**DEM**	**GD**
	NDMI		0.791	−0.738	−0.120	0.048
	NDVI			−0.963	−0.305	0.132
	NDWI				0.326	−0.133
	DEM					−0.383
BE Summer		**NDMI**	**NDVI**	**NDWI**	**DEM**	**GD**
	NDMI		0.753	−0.733	0.118	−0.096
	NDVI			−0.950	0.066	−0.049
	NDWI				−0.080	0.079
	DEM					−0.747

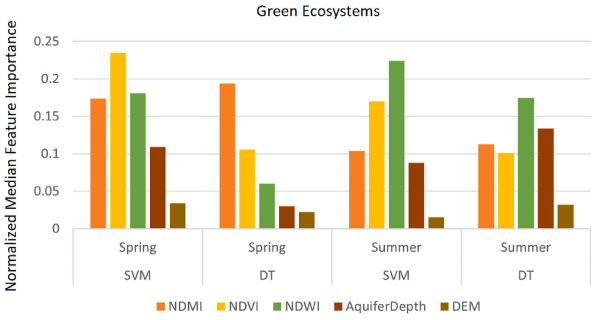

Fig. 7. Green Ecosystems feature importance estimates.

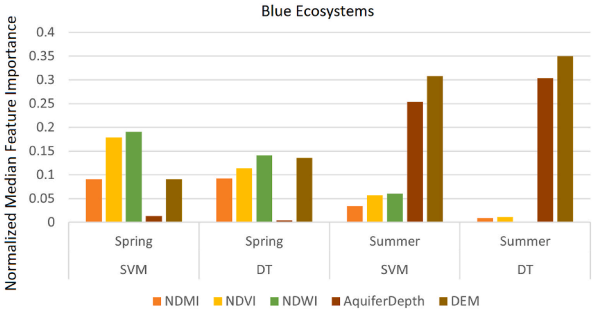

Fig. 8. Blue Ecosystems feature importance estimates.

3.2.3. Models' performance comparison

Table 5 displays the comparative performance between SVM and DT models. The overall results point out that SVM performed best, with all values greater than those obtained using the DT algorithm. The GE Spring model overcomes the Summer results regarding precision. The GE Summer model shows higher values for accuracy, recall, and F1-score. The BE Summer model performed better than the BE Spring in all accuracy functions.

Statistical comparisons of algorithm performance applying Kruskal-Wallis one-way analysis of variance to determine if the algorithms yielded significantly different performance for each evaluation metric. The Kruskal-Wallis significance results compare algorithms for each metric/dataset combination, showing that accuracy, F1-score and recall were significantly different in the GE Spring and Summer models. ROC-AUC is significantly different for GE Spring, and BE Spring and Summer models, and F1-score and precision for BE Spring.

Where these differences were observed, the Mann-Whitney U-tests and Wilcoxon Rank test were calculated to identify which algorithm yielded significantly better or worse performance, but any of these tests was significant to point out a single 'winner' algorithm across all metrics.

4. Discussion

Overall, this research was motivated by limitations acknowledged in previously published studies regarding the deployment of multi-season satellite-based models to categorize GBE and establish the link between drivers. Few studies have investigated the response of GBE to seasonality, such as Barron et al. (2014), Mpakairi et al. (2022a, b), Kundu et al. (2022) and Zhang et al. (2022), but none have gone so far as this research to develop a hybrid approach to classify GBE and understand the relationship between ecosystems, seasonal climate, and groundwater presence. The comparison between seasons captured the effects of seasonal climate variation and relationships between GE and BE. In dry periods, forests, lakes and rivers, and wetlands tend to show water stress, but if they have access to water sources, such as groundwater, the impacts between seasons tend to be less.

Seasonal GBE mapping has been studied in various parts of the world, often limited to local-scale investigations or specific climatic zones. Our study focused on analysing similar approaches carried out in river basins or aquifer systems that commonly experience dry events and understanding and overcoming their main challenges. The findings agreed with previous research as vegetation uses groundwater more during the driest season of the year when alternative water sources become depleted and evapotranspiration demand increases (Aldous and Gannett, 2021).

During the first stage of modelling, a set of sensitivity analysis tests was performed to understand which hyperparameters influence models' accuracy and to identify important uncertainty sources. Sensitivity analysis involves an analytical examination of the input data and hyperparameter settings. The sensitivity tests increased our understanding of models' behaviours and improved models' predictions. The use of this technique helped to provide guidance for future research and produce more accurate models. The outcomes of this step can stand alone as an efficient method to classify land cover (multiclass and binary) using DL in ArcGIS Pro but also can be used in an integrated approach, as proposed in this research.

Table 5
Models' performance values. Legend: Significance level set at 0.05 – Sig (*).

Model	Algorithm	Accuracy	F1-score	Recall	Precision	ROC-AUC
GE Spring	**SVM**	0.781	0.870	0.781	0.984	0.861
	DT	0.738	0.842	0.739	0.979	0.788
	Statistic	3.857	3.857	3.857	2.333	3.857
	P-Value	0.049	0.049	0.049	0.126	0.049
	Kruskal-Wallis Sig(*)	*	*	*		*
	Whitney U-tests Sig(*)					
	Wilcoxon Rank Sig(*)					
GE Summer	**SVM**	0.807	0.886	0.812	0.976	0.832
	DT	0.774	0.864	0.776	0.975	0.826
	Statistic	3.857	3.857	3.857	0.428	0.428
	P-Value	0.049	0.049	0.049	0.512	0.512
	Kruskal-Wallis Sig(*)	*	*	*		
	Whitney U-tests Sig(*)					
	Wilcoxon Rank Sig(*)					
BE Spring	**SVM**	0.800	0.366	0.785	0.238	0.875
	DT	0.781	0.334	0.747	0.216	0.825
	Statistic	1.190	3.857	0.428	3.857	3.857
	P-Value	0.275	0.049	0.512	0.049	0.049
	Kruskal-Wallis Sig(*)		*		*	*
	Whitney U-tests Sig(*)					
	Wilcoxon Rank Sig(*)					
BE Summer	**SVM**	0.935	0.689	0.980	0.532	0.982
	DT	0.929	0.665	0.965	0.507	0.975
	Statistic	0.784	1.190	2.333	0.428	3.857
	P-Value	0.375	0.275	0.126	0.512	0.049
	Kruskal-Wallis Sig(*)					*
	Whitney U-tests Sig(*)					
	Wilcoxon Rank Sig(*)					

The results also showed that the DL classification models (Spring and Summer) performed better considering a smaller number of classes, and a higher number of iterations and batch sizes. These models performed slightly better using the Spring dataset to classify transitional woodland shrubs, wetlands, agricultural areas, and watercourses than using the Summer dataset, which did not capture wetlands.

A good example of wetlands in the study region is the protected area Paul do Boquilobo Nature Reserve, a Biosphere Reserve recognized by UNESCO since 1981. It includes inland marshes with areas saturated with water long enough to promote aquatic processes. The DL Spring model was able to delineate seasonal wetlands, discriminating the inland marshes class. The DL Summer model could not differentiate wetlands from the other classes. Instead, it classified seasonal wetlands as a mix of herbaceous vegetation and surface water. Barron et al. (2014) also reported difficulties in the delineation of wetlands in a groundwater basin in the south-west of Western Australia, as these ecosystems are mostly dependent on long-term hydrological regimes. In their study, it is claimed that their approach is better suited for mapping GBE in places with a distinct and extended dry season, with an overall classification accuracy varying from 59% to 91%, which is in line with our results, ranging from 73.80% to 93.50%.

The second-stage modelling provides advanced understanding through exploratory data analysis, assessing the spatial-temporal distribution of GBE and estimating the relationship between ecosystem types, predictors, and seasons. FIE demonstrates the novelty of combining satellite-derived indices, topography, and aquifer depth to improve the model's quality. The comparison between ML algorithms expanded our knowledge and pointed out baselines for future research.

Spectral indices (NDVI, NDWI, and NDMI) showed high positive and negative associations, in all models. Topography presented relevant associations with the spectral indices particularly when modelling GE in the Summer season. In the BE Summer model, aquifer depth and elevation were highly negatively correlated with each other. This contrasting relationship was expected in the study area in this season, as the surface water yield depends on the groundwater level (Almeida and Cabral, 2023). This relationship between drivers aids in the delineation of areas with a higher moisture signature at the surface and the detection of thermal anomalies, which are often linked to a higher level of water availability in groundwater discharge zones. Large changes in climatic conditions, along with continued groundwater abstraction, may result in large variations in groundwater discharges to rivers, springs, and wetlands, causing impacts on the extent of GBE.

FIE evidenced that satellite-derived indices are efficient in delineating GE in both seasons and BE in the Spring, and topography and the aquifer depth for the BE Summer model. The results are compatible with those presented by Barron et al. (2014) in the Mediterranean climate zone of Australia when interpreting the land surface response to dry from combined changes in spectral indices (NDVI and NDWI). Their findings revealed that GBE are characterized by high NDVI and NDWI values at the start of the dry season, but rapidly decrease as vegetation dries up and the moisture content of the soil falls below the soil field capacity.

The high values of NDVI are also an indication of vegetation heterogeneity as demonstrated in the research carried out by Mpakairi et al. (2022a, b). Their findings revealed that vegetation heterogeneity in a transboundary aquifer located between Botswana and South Africa was higher during the wet season and lower during the dry season, and places with a high water table

(e.g., a spring or groundwater seepage) are more ecologically varied than areas with a low water table in arid conditions. These findings also justify the relationship between NDVI and GD in our results. Kundu et al. (2022) utilized NDWI to map water depth, hydroperiod, and consistency of water presence in a transboundary river (Punarbhaba River) that connects India and Bangladesh. The highest positive NDWI values suggest greater water thickness, which confirms our results regarding the importance of NDWI in mapping GBE in the Spring.

The elevation and groundwater depth are some of the important variables in hydro-ecological models (O'Grady et al., 2011), which was also verified in our results. The topography creates specific habitats that influence the occurrence of certain plant species, and the groundwater depth is a proxy for detecting whether ecosystems can freely access water, periodically or not (Eamus et al., 2006). The water presence has the potential to influence the status of GBE, therefore groundwater draw-down and unsustainable groundwater abstraction are compromising ecological integrity at all levels (Eamus and Froend, 2006). Kundu et al. (2022) found that water depth is a factor of ecosystem resilience, as many species that like to reside at different depths benefit from deeper water depth.

Performance comparison between algorithms indicates that SVM was better than DT in mapping inland GE and BE in both seasons. These methods are the most used to model complex class signatures (Maxwell et al., 2018). SVM may perform better with imbalanced data, due to its stronger generalization capacity, even with small training sample sizes (Tarantino et al., 2021). While DT has the disadvantage of instability due to the propagation of errors down through subsequent splits in the tree (Breiman, 1984). Zhang et al. (2022) tested six ML algorithms including SVM and DT to classify GBE in the groundwater discharge zones of the Great Plains Aquifer system in eastern Nebraska. Their optimal model (SVM) achieves an overall accuracy of 99.95% for BE classification, against 93.50% obtained in our BE Summer model. For GE classification, their gradient tree boost model achieves an overall accuracy of 94.07% while here the best accuracy for GE is 80.70%, also with SVM.

Our findings require testing other algorithms and methods, in locations with similar environmental characteristics. The statistical tests did not allow us to choose the best algorithm but evidenced that the performance comparisons within the selected metrics were significantly different. Barron et al. (2014), Mpakairi et al. (2022a, b), and Kundu et al. (2022) refer as a limitation that the large-scale technique may not discover seasonal GBE smaller than 30 × 30 m, suggesting efforts for model development using a finer scale, as the present study does, using satellite imagery with 10 m spatial resolution.

Seasonal changes in the greenness of vegetated land can be linked to phenological patterns of vegetation and climate-vegetation interactions. Because of limited rainfall, semi-arid and arid environments are water-scarce, and vegetation relies on surface or subterranean water (Sahana et al., 2022). The GBE are potentially exposed to climate change because of this (Pandey et al., 2023). Apart from climate and lowering of groundwater level, land cover changes also drive changes in GBE (Mpakairi et al., 2022a, b). Urban sprawl, population growth, and increased agricultural area demand are the predominant drivers of land conversion (Kundu et al., 2022).

Future work can expand the methodology to assess GBE at the national scale and explore other spectral indices, landscape metrics, and contextual data to flag the location and characteristics of GBE. Classifying GBE by seasons in a two-step modelling approach has the benefit of getting the most out of those techniques and adds value to the modelling process by working with subset types of the data. An alternative way to overcome those concerns and handle those large datasets would be using cloud-based storage and computing resources. Even though depending on spatial and spectral resolutions, the process would also require sub-setting.

Moreover, more research is needed to discriminate seasonal wetlands from flooded areas and shrublands, which may require different approaches as they have different responses in terms of reflectance, water requirements, phenology, and spatial and temporal patterns. Concerns came out under this scenario and must be explored in future studies, such as 1) Would it be necessary to include other bands that are specifically able to capture those unique ecosystems in the Summer, or to use other spectral indices or landscape metrics to highlight them? 2) Which algorithms would better to capture these seasonal variations in wetlands? and, 3) Which should be a reasonable scale of modelling to detect them?

5. Conclusions

The overall aim of the study is to have a better understanding of how GBE systems should be modelled considering the dynamics of land use and seasonal climate effects. The approach uses SEO products to detect seasonal climate variations and relationships in the implementation of seasonal satellite-based data-driven models. The sensitivity analysis advanced the knowledge in applying DL models in complex landscapes using GIS tools. The sensitivity tests showed that the hyperparameter settings are key to building accurate models and maximizing the benefits of each technique. FIE demonstrated the novelty of applying satellite products with contextual data to empower the model's quality. The findings of this study indicate the feasibility of using an observation strategy to continuously monitor the GBE of the Lower Tagus Aquifer system, as well as in other parts of the world.

Authorship contributions

Bruna Almeida: Conceptualization; Data curation; Formal analysis; Methodology; Investigation; Software; Validation; Visualization; Writing - original draft; Writing - review & editing. Pedro Cabral: Conceptualization, Methodology, Investigation, Writing - review and editing, Project administration, Supervision.

Ethical Statement

Hereby, I Bruna Almeida, consciously assure that for the manuscript "A hybrid modelling approach for detecting seasonal variations in inland green-blue ecosystems" the following is fulfilled:

1) This material is the authors' own original work, which has not been previously published elsewhere.
2) The paper is not currently being considered for publication elsewhere.
3) The paper reflects the authors' own research and analysis in a truthful and complete manner.
4) The paper properly credits the meaningful contributions of co-authors and co-researchers.
5) The results are appropriately placed in the context of prior and existing research.
6) All sources used are properly disclosed (correct citation). Literally copying of text must be indicated as such by using quotation marks and giving proper references.
7) All authors have been personally and actively involved in substantial work leading to the paper and will take public responsibility for its content.

I agree with the above statements and declare that this submission follows the policies of Remote Sensing Applications: Society and Environment as outlined in the Guide for Authors and in the Ethical Statement.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This study was supported by the research project MaSOT – Mapping Ecosystem Services from Earth Observations, funded by the Portuguese Science Foundation – FCT [EXPL/CTA-AMB/0165/2021]. The authors gratefully acknowledge the financial support of the FCT, through the MagIC Research (Centro de Investigação em Gestão de Informação - UIDB/04152/2020). We are grateful for the constructive remarks from two anonymous reviewers.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rsase.2023.101121.

References

A, A.S., S, A.A., 2023. Land-cover classification with hyperspectral remote sensing image using CNN and spectral band selection. Remote Sens. Appl. 31, 100986. https://doi.org/10.1016/J.RSASE.2023.100986.

Alarcon Blazquez, M., van der Veeren, R., Gacutan, J., James, P., 2023. Compiling preliminary SEEA Ecosystem Accounts for the OSPAR regional sea: experimental findings and lessons learned. One Ecosyst. 8, e108030. https://doi.org/10.3897/oneeco.8.e108030.

Aldous, A.R., Gannett, M.W., 2021. Groundwater, biodiversity, and the role of flow system scale. Ecohydrology 14. https://doi.org/10.1002/eco.2342.

Almeida, B., Cabral, P., 2023. Data-driven modelling of freshwater ecosystems: a multiscale framework based on global geospatial data. In: Proceedings of the 9th International Conference on Geographical Information Systems Theory, Applications and Management. SCITEPRESS - Science and Technology Publications, pp. 104–111. https://doi.org/10.5220/0012037800003473.

Almeida, B., Cabral, P., 2021. Water yield modelling, sensitivity analysis and validation: a study for Portugal. ISPRS Int. J. Geo-Inf. 10. https://doi.org/10.3390/ijgi10080494.

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

Belo-Pereira, M., Dutra, E., Viterbo, P., 2011. Evaluation of global precipitation data sets over the Iberian Peninsula. J. Geophys. Res. Atmos. 116. https://doi.org/10.1029/2010JD015481.

Bertrand, G., Goldscheider, N., Gobat, J.M., Hunkeler, D., 2012. Review: from multi-scale conceptualization to a classification system for inland groundwater-dependent ecosystems. Hydrogeol. J. https://doi.org/10.1007/s10040-011-0791-5.

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. pp. 144–152.

Breiman, L., 1984. Classification and Regression Trees, first ed. Routledge, New York. https://doi.org/10.1201/9781315139470.

Chang, T., Rasmussen, B., Dickson, B., Zachmann, L., 2019. Chimera: a multi-task recurrent convolutional neural network for forest classification and structural estimation. Rem. Sens. 11, 768. https://doi.org/10.3390/rs11070768.

Chen, H., Chen, C., Zhang, Z., Lu, C., Wang, L., He, X., Chu, Y., Chen, J., 2021. Changes of the spatial and temporal characteristics of land-use landscape patterns using multi-temporal Landsat satellite data: a case study of Zhoushan Island, China. Ocean Coast Manag. 213, 105842. https://doi.org/10.1016/J.OCECOAMAN.2021.105842.

Chrysafis, I., Korakis, G., Kyriazopoulos, A.P., Mallinis, G., 2020. Predicting tree species diversity using geodiversity and sentinel-2 multi-seasonal spectral information. 2020. Sustainability 12, 9250. https://doi.org/10.3390/SU12219250. 12, 9250.

Coelho, M.F., Ramos, A., de Lima, I., Trigo, R., 2013. Seasonal changes in daily precipitation extremes in mainland Portugal from 1941 to 2007. Reg. Environ. Change 14. https://doi.org/10.1007/s10113-013-0515-6.

Copernicus Programme, 2023. CLC 2018 — Copernicus Land Monitoring Service. [WWW Document]. URL. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018. 1.24.23.

de Lima, M.I.P., Santo, F.E., Ramos, A.M., Trigo, R.M., 2015. Trends and correlations in annual extreme precipitation indices for mainland Portugal, 1941–2007. Theor. Appl. Climatol. 119, 55–75. https://doi.org/10.1007/s00704-013-1079-6.

Domingos, P., 2012. A few useful things to know about machine learning. Commun. ACM. https://doi.org/10.1145/2347736.2347755.

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

Dwire, K.A., Mellmann-Brown, S., Gurrieri, J.T., 2018. Potential Effects of Climate Change on Riparian Areas, Wetlands, and Groundwater-dependent Ecosystems in the Blue Mountains. Clim Serv, Oregon, USA. https://doi.org/10.1016/j.cliser.2017.10.002.

Eamus, D., Froend, R., 2006. Groundwater-dependent ecosystems: the where, what and why of GDEs. Aust. J. Bot. 54, 91–96. https://doi.org/10.1071/BT06029.

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

Eamus, D., Ray, F., 2006. Groundwater-dependent ecosystems : the where , what and why of GDEs. Aust. J. Bot. 54, 91–96.

EEA, 2023. WISE Water Framework Directive Database. — European Environment Agency [WWW Document]. URL. https://www.eea.europa.eu/data-and-maps/data/wise-wfd-4. 1.24.23.

Espírito Santo, F., Ramos, A.M., de Lima, M.I.P., Trigo, R.M., 2014. Seasonal changes in daily precipitation extremes in mainland Portugal from 1941 to 2007. Reg. Environ. Change 14, 1765–1788. https://doi.org/10.1007/S10113-013-0515-6/FIGURES/14.

ESRI, 2023. ArcGIS Pro - ESRI. Environmental Systems Research Institute.

Eurostat, 2023. Database - Population and Demography. [WWW Document]. URL. https://ec.europa.eu/eurostat/web/population-demography/demography-population-stock-balance/database. 1.17.23.

Gao, B.C., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

Guo, M., Yu, Z., Xu, Y., Huang, Y., Li, C., 2021. Me-net: a deep convolutional neural network for extracting mangrove using sentinel-2A data. Rem. Sens. 13, 1292. https://doi.org/10.3390/rs13071292.

Gwal, S., Singh, S., Gupta, S., Anand, S., 2020. Understanding forest biomass and net primary productivity in Himalayan ecosystem using geospatial approach. Model Earth Syst Environ 6, 2517–2534. https://doi.org/10.1007/s40808-020-00844-4.

Ji, L., Geng, X., Sun, K., Zhao, Y., Gong, P., 2015. Target detection method for water mapping using landsat 8 OLI/TIRS imagery. Water (Basel) 7, 794–817. https://doi.org/10.3390/w7020794.

Khwarahm, N.R., 2021. Spatial modeling of land use and land cover change in Sulaimani, Iraq, using multitemporal satellite data. Environ. Monit. Assess. 193. https://doi.org/10.1007/s10661-021-08959-6.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 84–90. https://doi.org/10.1145/3065386.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47, 583–621. https://doi.org/10.1080/01621459.1952.10483441.

Kundu, S., Pal, S., Mandal, I., Talukdar, S., 2022. How far damming induced wetland fragmentation and water richness change affect wetland ecosystem services? Remote Sens. Appl. 27, 100777. https://doi.org/10.1016/j.rsase.2022.100777.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Li, L., Casado, A., Congedi, L., Dell'Aquila, A., Dubois, C., Elizalde, A., L'Hévéder, B., Lionello, P., Sevault, F., Somot, S., Ruti, P., Zampieri, M., 2012. Modeling of the mediterranean climate system. The Climate of the Mediterranean Region: Future Times 419–448. https://doi.org/10.1016/B978-0-12-416042-2.00007-0.

Lobert, F., Holtgrave, A.K., Schwieder, M., Pause, M., Vogt, J., Gocht, A., Erasmi, S., 2021. Mowing event detection in permanent grasslands: systematic evaluation of input features from Sentinel-1, Sentinel-2, and Landsat 8 time series. Remote Sens. Environ. 267. https://doi.org/10.1016/j.rse.2021.112751.

Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 18, 50–60. https://doi.org/10.1214/aoms/1177730491.

Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int. J. Rem. Sens. https://doi.org/10.1080/01431161.2018.1433343.

Meddens, A.J.H., Steen-Adams, M.M., Hudak, A.T., Mauro, F., Byassee, P.M., Strunk, J., 2022. Specifying geospatial data product characteristics for forest and fuel management applications. Environ. Res. Lett. 17. https://doi.org/10.1088/1748-9326/ac5ee0.

Mouta, N., Silva, R., Pais, S., Alonso, J.M., Gonçalves, J.F., Honrado, J., Vicente, J.R., 2021. 'The best of two worlds'—combining classifier fusion and ecological models to map and explain landscape invasion by an alien shrub. Rem. Sens. 13, 3287. https://doi.org/10.3390/rs13163287.

Mpakairi, Kudzai S., Dube, T., Dondofema, F., Dalu, T., 2022a. Spatio–temporal variation of vegetation heterogeneity in groundwater dependent ecosystems within arid environments. Ecol. Inf. 69, 101667. https://doi.org/10.1016/j.ecoinf.2022.101667.

Mpakairi, Kudzai Shaun, Dube, T., Dondofema, F., Dalu, T., 2022b. Spatial characterisation of vegetation diversity in groundwater-dependent ecosystems using in-situ and sentinel-2 MSI satellite data. Rem. Sens. 14, 2995. https://doi.org/10.3390/rs14132995.

Mugiraneza, T., Ban, Y., Haas, J., 2019. Urban land cover dynamics and their impact on ecosystem services in Kigali, Rwanda using multi-temporal Landsat data. Remote Sens. Appl. 13, 234–246. https://doi.org/10.1016/j.rsase.2018.11.001.

NASA, 2023. Earthdata | Earthdata. [WWW Document]. URL. https://www.earthdata.nasa.gov/. 3.2.23.

Nikparvar, B., Thill, J.C., 2021. Machine learning of spatial data. 2021. ISPRS Int. J. Geo-Inf. 10, 600. https://doi.org/10.3390/IJGI10090600. 10, 600.

Novo, M.E., Oliveira, M., Martins, T., Henriques, M.J., 2018. Projecto Bingo: O Impacto das Alterações Climáticas na Componente Subterrânea do Ciclo Hidrológico. Revista Recursos Hídricos 39, 59–74. https://doi.org/10.5894/rh39n2-cti3.

O'Grady, A.P., Carter, J.L., Bruce, J., 2011. Can we predict groundwater discharge from terrestrial ecosystems using existing eco-hydrological concepts? Hydrol. Earth Syst. Sci. 15, 3731–3739. https://doi.org/10.5194/hess-15-3731-2011.

Oliveira, M.M., Novo, M.E., Paulo, J., Ferreira, L., 2005. Models to Predict the Impact of the Climate Changes on Aquifer Recharge. p. 310.

Osborne, P.E., Alvares-Sanches, T., 2019. Quantifying how landscape composition and configuration affect urban land surface temperatures using machine learning and neutral landscapes. Comput. Environ. Urban Syst. 76, 80–90. https://doi.org/10.1016/j.compenvurbsys.2019.04.003.

Pandey, H.K., Kumar Singh, V., Kumar Singh, S., Kumar Sharma, S., 2023. Mapping and validation of groundwater dependent ecosystems (GDEs) in a drought-affected part of Bundelkhand region, India. Groundw Sustain Dev 23, 100979. https://doi.org/10.1016/J.GSD.2023.100979.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Peng, H., Long, F., Ding, C., 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1226–1238. https://doi.org/10.1109/TPAMI.2005.159.

Pudjihartono, N., Fadason, T., Kempa-Liehr, A.W., O'Sullivan, J.M., 2022. A review of feature selection methods for machine learning-based disease risk prediction. Frontiers in Bioinformatics 2, 72. https://doi.org/10.3389/FBINF.2022.927312.

Ramos, T.B., Horta, A., Gonçalves, M.C., Pires, F.P., Duffy, D., Martins, J.C., 2017. The INFOSOLO database as a first step towards the development of a soil information system in Portugal. Catena 158, 390–412. https://doi.org/10.1016/j.catena.2017.07.020.

Reddy, C.S., 2021. Remote sensing of biodiversity: what to measure and monitor from space to species? Biodivers. Conserv. 30, 2617–2631. https://doi.org/10.1007/s10531-021-02216-5.

Rijal, S., Rimal, B., Acharya, R.P., Stork, N.E., 2021. Land use/land cover change and ecosystem services in the Bagmati River Basin, Nepal. Environ. Monit. Assess. 193, 651. https://doi.org/10.1007/s10661-021-09441-z.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional Networks for Biomedical Image Segmentation.

Rouse, J., Haas, R., Schell, J., Deering, D., 1973. Monitoring Vegetation Systems in the Great Plains with ERTS. Third ERTS Symposium, NASA.

Sahana, M., Saini, M., Areendran, G., Imdad, K., Sarma, K., Sajjad, H., 2022. Assessing Wetland ecosystem health in Sundarban Biosphere Reserve using pressure-state-response model and geospatial techniques. Remote Sens. Appl. 26, 100754. https://doi.org/10.1016/J.RSASE.2022.100754.

Silleos, N.G., Alexandridis, T.K., Gitas, I.Z., Perakis, K., 2008. Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years. https://doi.org/10.1080/10106040608542399. 10.1080/10106040608542399,21-28.

Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali, A., Schaepman, M.E., Paganini, M., Kissling, W.D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., Wegmann, M., Wingate, V., 2021. Priority list of biodiversity metrics to observe from space. Nat Ecol Evol 5, 896–906. https://doi.org/10.1038/s41559-021-01451-x.
SNIRH, 2023. Sistema Nacional de Informação de Recursos Hídricos [WWW Document]. URL. https://snirh.apambiente.pt/. 3.2.23.
Stehman, S.V., Foody, G.M., 2019. Key issues in rigorous accuracy assessment of land cover products. Remote Sens. Environ. 231, 111199. https://doi.org/10.1016/j.rse.2019.05.018.
Tarantino, C., Forte, L., Blonda, P., Vicario, S., Tomaselli, V., Beierkuhnlein, C., Adamo, M., 2021. Intra-annual sentinel-2 time-series supporting grassland habitat discrimination. Rem. Sens. 13, 1–29. https://doi.org/10.3390/rs13020277.
Urbanowicz, R.J., Meeker, M., La Cava, W., Olson, R.S., Moore, J.H., 2018. Relief-based feature selection: introduction and review. J. Biomed. Inf. 85, 189–203. https://doi.org/10.1016/J.JBI.2018.07.014.
Urbanowicz, R.J., Zhang, R., Cui, Y., Suri, P., 2022. STREAMLINE: A Simple, Transparent, End-To-End Automated Machine Learning Pipeline Facilitating Data Analysis and Algorithm Comparison. https://doi.org/10.48550/arxiv.2206.12002.
Vis, G.J., Bohncke, S.J.P., Schneider, H., Kasse, C., Coenraads-Nederveen, S., Zuurbier, K., Rozema, J., 2010. Holocene flooding history of the lower Tagus valley (Portugal). J. Quat. Sci. 25, 1222–1238. https://doi.org/10.1002/jqs.1401.
Vis, G.J., Kasse, C., 2009. Late quaternary valley-fill succession of the lower Tagus valley, Portugal. Sediment. Geol. 221, 19–39. https://doi.org/10.1016/j.sedgeo.2009.07.010.
Wang, L., Chen, C., Xie, F., Hu, Z., Zhang, Z., Chen, H., He, X., Chu, Y., 2021a. Estimation of the value of regional ecosystem services of an archipelago using satellite remote sensing technology: a case study of Zhoushan Archipelago, China. Int. J. Appl. Earth Obs. Geoinf. 105, 102616. https://doi.org/10.1016/J.JAG.2021.102616.
Wang, X., Han, J., Wang, Xia, Yao, H., Zhang, L., 2021b. Estimating soil organic matter content using sentinel-2 imagery by machine learning in shanghai. IEEE Access 9, 78215–78225. https://doi.org/10.1109/ACCESS.2021.3080689.
Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics Bull. 1, 80. https://doi.org/10.2307/3001968.
Yang, X., Liu, J., 2020. Assessment and Valuation of Groundwater Ecosystem Services: A Case Study of Handan City, China. Water (Switzerland), vol. 12. https://doi.org/10.3390/w12051455.
Zeiler, M.D., Fergus, R., 2013. Visualizing and Understanding Convolutional Networks.
Zhang, J., Okin, G.S., Zhou, B., 2019. Assimilating optical satellite remote sensing images and field data to predict surface indicators in the Western U.S.: assessing error in satellite predictions based on large geographical datasets with the use of machine learning. Remote Sens. Environ. 233, 111382. https://doi.org/10.1016/j.rse.2019.111382.
Zhang, L., Hu, Q., Tang, Z., 2022. Assessing the contemporary status of Nebraska's eastern saline wetlands by using a machine learning algorithm on the Google Earth Engine cloud computing platform. Environ. Monit. Assess. 194, 193. https://doi.org/10.1007/s10661-022-09850-8.
Zhang, R.F., Urbanowicz, R.J., 2020. A scikit-learn compatible learning classifier system. In: GECCO 2020 Companion - Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion. pp. 1816–1823. https://doi.org/10.1145/3377929.3398097.