COINS (C0ntrol of INvasive Species) – 2.0

Authors: Angela Martiradonna[1,2], Fasma Diele[1], Carmela Marangi[1], Cristiano Tamborrino[1,3]

[1] Istituto per le Applicazioni del Calcolo M. Picone, CNR, Bari
[2] Dipartimento di Matematica, Università degli Studi di Bari
[3] Dipartimento di Informatica, Università degli Studi di Bari

Baker, C. M., Blonda, P., Casella, F., Diele, F., Marangi, C., Martiradonna, A., ... & Tarantino, C. (2023). 'Using remote sensing data within an optimal spatiotemporal model for invasive plant management: the case of Ailanthus altissima in the Alta Murgia National Park.' Scientific Reports, 13(1), 14587.

The complete description of the model and numerical methods is available in the cited publication.

The procedure employs a modeling approach to achieve optimal spatiotemporal control of invasive species within natural protected areas of significant conservation importance. The model relies on diffusion equations, maintaining spatial explicitness, and incorporates a functional response (Holling Type II). This functional response models the control rate as a function of a crucial state variable—the density of the invasive species.

The control variable is expressed as the effort required for the eradication of the invasive species. Additionally, a budget constraint is imposed to limit the available effort. The species' growth is influenced by a habitat suitability function, computed internally using the land cover map of the study area and the map of the initial density of the invasive species. This function evaluates habitat suitability by analyzing the frequency of occurrence of each land cover class in the vicinity of the points where the invasive species are present.

The routine addresses a constrained optimal control problem, aiming to discover the optimal allocation of effort that minimizes the density of the invasive species across both time and space. In its current version, the model is specifically applied to the Ailanthus altissima plant species, which inhabits a National Park in the southern region of Italy (Murgia Alta).

These code revisions mark a significant advancement from the original version, characterized by an amalgamation of sophisticated modeling techniques and statistical methodologies. One notable improvement lies in the approach to estimate habitat suitability, showcasing a more nuanced understanding of ecological dynamics. The introduction of an optimal control model represents a substantial leap forward, indicating a deliberate effort to enhance the precision and efficacy of invasive species management. A commendable aspect is the judicious use of numerical methods, where optimizations have been made to improve computational efficiency. The restructuring of the code into functions not only enhances the overall clarity of the implementation but also promotes reusability and maintainability. The inclusion of external libraries, ranging from geospatial tools to statistical and plotting packages, underscores a pragmatic strategy of leveraging existing resources to streamline development. Moreover, the incorporation of visualizations throughout the simulation process is a noteworthy enhancement, providing a more intuitive understanding of the model's behavior. This not only aids in the interpretation of results but also facilitates the identification of potential issues during the simulation. Additionally, the implementation of error handling mechanisms with informative messages demonstrates a conscientious approach to software robustness. The inclusion of such features contributes to the overall reliability of the codebase, allowing for more efficient debugging and troubleshooting. In essence, these code refinements signify a holistic and thoughtful progression, reflecting an amalgamation of ecological insights, numerical finesse, and a commitment to producing a more sophisticated and reliable modeling tool.

The development and the implementation of the model and the routine have been made possible thanks to:

- Project "Modelli differenziali per la salvaguardia della biodiversità minacciata dalle specie invasive nelle aree protette. Utilizzo di dati satellitari per l'analisi di scenario e il controllo ottimo nel Parco Nazionale dell'Alta Murgia"funded by Research for Innovation' (REFIN) progamme n. 0C46E06B- Regione Puglia, Italy.

- National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union—NextGenerationEU; Award Number: Project code CN 00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP B83C22002930006, Project title "National Biodiversity Future Center—NBFC.

The package includes four main codes: *HabSuitCV.R*, *edges_n.R*, *COINSdiff.R* and *COINScontrol.R*, written in R (version 4.2.2).

HabSuitCV.R

This file contains code related to the habitat suitability function and its computation. It handles the calculation of habitat suitability based on the land cover map of the study area and the initial density map of the invasive species. The code assesses the frequency of occurrence of each land cover class in proximity to the points where the invasive species are present.

1. Loading Packages:
    - The script begins by loading two R packages: *raster* and *maptools*. These packages are commonly used for spatial data analysis and manipulation.
2. Reading Files:
    - It reads two input files: *initial_density.tif* and *land_cover.shp*.
    - The initial density file represents the initial distribution of the invasive species.
    - The land cover shapefile likely contains information about the different land cover classes in the study area.
3. Data Preprocessing:
    - The initial density raster is aggregated using the aggregate function, potentially to reduce the spatial resolution.
    - The script then creates subsets of the raster data based on different extents.
    - For each subset, it merges the raster data to create a set of rasters (rast) and a list of land cover data (uds).
4. Spatial Analysis:
    The script then iterates over these subsets (five times), performing the following steps:
        - Thresholding the initial density raster based on a predefined threshold (pres_thres).
        - Applying the function (edges_n.r) for spatial analysis.
        - Calculating scores, frequency, density, and normalized values based on land cover classes and the spatial distribution of the invasive species.
5. Output Generation:
    - Generates CSV files (scoreCV[j].csv) containing information about the frequency, normalized frequency, density, normalized density, and area for each land cover class.
    - Creates habitat suitability rasters (habitat_suitability_frequencyCV[j] and habitat_suitability_densityCV[j]) based on frequency and density calculations.
7. Summary Statistics:
    - Reads the previously generated CSV files and calculates mean and standard deviation values for each land cover class.
    - Writes a summary CSV file (HS_CV.csv) containing mean, standard deviation, and original data.
8. Habitat Suitability Raster:
    - Creates a raster (HS_CV.tif) representing habitat suitability by replacing land cover values with mean habitat suitability values.

edges_n.R

The script enhances the original land cover classification by identifying and labeling edges of certain land cover classes, making it more nuanced and providing information about the transition zones between different land cover types.

Example Modification for One Land Cover Class (Deciduous Forest):

```
ndf <- nlev[lev=='Deciduous forest']
uds_mat[find_edges(uds_mat==ndf)==1] = 47
lev[47] = "Deciduous forest edge"
```

This example finds the locations in uds_mat where the land cover is "Deciduous forest" and updates those locations to 47. It also updates the corresponding entry in the lev vector to "Deciduous forest edge." The script repeats this process for other specified land cover classes.

COINSdiff.R

The script models the spatiotemporal dynamics of invasive species using a numerical approach, considering habitat suitability and other parameters. It simulates the spread and distribution of the invasive species over time, producing output files for further analysis and visualization.

Input:

- presen_path: File path for the presence data, loaded as a raster (pres).
- suit_name: File path for habitat suitability map, loaded as a raster (HSI_rast).
- domain_path: File path for the shapefile representing the geographical domain (domain).
- uds_path: File path for the land cover shapefile (uds).
- param_path: File path for a CSV file containing model parameters (param).
- dx: Grid resolution used in the model.

Output:

- Rasters for presence, density, and density-to-suitability ratio saved in GeoTIFF format.
- Spatial plots of density over time.
- CSV files containing model results, such as habitat suitability frequency, density, and area (scoreCV files).

COINScontrol.R

The script performs a forward-backward algorithm to simulate the dynamics of the invasive species considering habitat suitability and control efforts. The output includes rasters representing the presence, density, and effort of the invasive species over time. The control effort is calculated based on the specified control parameters

Input:

1. Model parameters:
    - dx: Grid resolution.
    - out_name: Output name for saving results.
    - suit_name: File path for habitat suitability data, loaded as a raster (HSI_rast).
    - presen_path: File path for the density data of the invasive species in 2014, loaded as a raster (pres).
    - domain_path: File path for the shapefile representing the geographical domain (domain).
    - param_path: File path for a CSV file containing model parameters (param).
    - uds_path: File path for the land cover shapefile (uds).
    - Parameters (D, r, c, mu, h, nu, omega, delta, B) are extracted from the CSV file.
2. Time Settings:
    - T: Simulation time (years).
    - date0: Initial date of simulation (July 1, 2014).
    - dt: Time step for numerical integration.
    - int: Time interval for output storage.
3. Grid and Numerical Settings:
    - dxkm: Grid resolution in kilometers.

Output:

Model outputs are saved in raster files for presence, density, and effort. Each output file includes the corresponding year in its name.

- Results are saved in a directory named based on out_name.
- Output rasters for presence, density, and effort (control) saved in GeoTIFF format.
- Spatial plots of density and effort over time.
- The control effort is calculated and stored in the output directory.
- Information about the tolerance and convergence of the forward-backward algorithm.

Data:

Initial density map from: 10.5281/zenodo.2628376

Land cover map from: https://www.sit.puglia.it/portal/portale_cartografie_tecniche_tematiche/Cartografie%20Tematiche/UDS