



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA

Finanziato nell'ambito del Piano Nazionale di Ripresa e Resilienza PNRR. Missione 4, Componente 2, Investimento 1.3 Creazione di "Partenariati estesi alle università, ai centri di ricerca, alle aziende per il finanziamento di progetti di ricerca di base"



**GRINS**  
FOUNDATION

## Spoke 7 "Territorial Sustainability"

### WP3 – Area Specific Planning

#### Deliverable 7.2.3

*Decision Support System (DSS) for policy makers and territorial planners*

<b>Document data</b>	
Decision Support System (DSS) for policy makers and territorial planners	Spoke 7 Work Package 2 D 2.3 - Decision Support System (DSS) for policy makers and territorial planners
Owner	UNIBA_DEMDI
Contributor/s	POLIMI_DIG and UNIBA_DEMDI
Document version	D 2.3 - v.2.0
Last version date	10/12/2025

## CREDITS

### Attributions:

This report was coordinated and written by the University of Bari "Aldo Moro" and Politecnico of Milan.

### Composition of the research groups:

UNIBA\_DEMDI Research group: Angela Stefania Bergantino (Coordinator and Spoke 7 leader), Mario Intini, Alessandro Buongiorno, Giulio Fusco (GRINS researchers), Elsa Amaddeo, Gianluca Monturano (GRINS post-doc).

POLIMI Research group: Marika Arena (Coordinator), Giovanni Bonaccorsi, Francesco Scotti, Valeria Maria Urbano (GRINS researchers).

# Executive summary

This report presents the development of a Decision Support System (DSS) designed to evaluate and predict the resilience and attractiveness of Italian territories, with the overarching goal of supporting sustainable territorial development. The DSS allows decision-makers to explore territorial dynamics, understand the drivers of resilience and attractiveness, and simulate alternative scenarios to inform the decision-making process. The report explains how the system operates, illustrates its main steps and components.

The development of the DSS follows a structured construction framework based on four key stages and begins with an assessment of decision-makers' needs. Local governments were identified as the primary stakeholders, given their central role in territorial planning and public administration. DSS have been shown to enhance decision-making quality, improve organizational performance, and strengthen management control in public sector contexts. In this regard, the proposed DSS aims to offer local authorities a data-driven tool to support evidence-informed decisions.

The DSS provides a comprehensive overview of territorial attractiveness and resilience across territories. By examining the relationship between these outcome variables and the underlying territorial features, the system enables users to identify relevant patterns and drivers. To operationalize these capabilities, two main applications were developed: "Exploration" and "Prediction".

The Exploration functionality focuses on analyzing the level of association between selected territorial characteristics and the outcome variables. This feature supports decision-makers in identifying which factors are most strongly related to territorial resilience and attractiveness.

The Prediction functionality enables users to estimate how variations in input variables may affect future performance. To validate and demonstrate the predictive capacity of the DSS, two applications were implemented, focusing respectively on PM2.5 concentration levels and tourism flows. These applications provide concrete examples of how the DSS can support informed decision making in different domains.

Finally, the report presents the applications developed on the Amelia platform, illustrating their operational steps and user interface, and demonstrating how decision-makers can interact with the system to support informed territorial strategies.

## TABLE OF CONTENTS

<b>Executive summary</b> .....	<b>3</b>
<b>1. Spoke 7 Territorial Sustainability</b> .....	<b>5</b>
1.1 Deliverable Objective .....	5
<b>2. DSS Framework</b> .....	<b>5</b>
2.1 Knowledge Phase .....	6
2.2 Modelling Phase .....	7
2.3 Output of Modelling .....	7
2.4 Scenario Analysis and Planning .....	7
<b>3. DSS Function</b> .....	<b>8</b>
3.1 Exploration .....	8
3.2 Prediction .....	15
<b>References</b> .....	<b>25</b>

# 1. Spoke 7 Territorial Sustainability

## 1.1 Deliverable Objective

The objective of this report is to outline the development of a Decision Support System (DSS), aiming to enhance territorial sustainability by evaluating and predicting the resilience and attractiveness of Italian territories. Building upon the foundational work established in Deliverable 1.1, “Dashboard of geo-localised and real-time territorial indicators”, which outlined the dimensions of analysis and indicator selection for a comprehensive assessment of territorial resilience and attractiveness, and Deliverable 2.1, “A Decision Support System for the analysis of the attractiveness and resilience of Italian territories”, which outlined the design of the Decision-Support System describing the general functions, this report highlights the functionalities of the DSS. More specifically, it aims to illustrate how the system works in practice and how its components support the analysis and prediction of territorial resilience. Furthermore, it clarifies how, operationally, the DSS allows to assess the relationship among alternative variables characterizing the analyzed territories, offering insights into the drivers of resilience and attractiveness, and enabling scenario analysis to guide sustainable territorial development.

## 2. DSS Framework

This chapter recaps on the framework we adopted for the design of a comprehensive DSS specifically tailored for territorial analysis and intervention planning, as presented in Deliverable 2.1, “A Decision Support System for the analysis of the attractiveness and resilience of Italian territories”. The system addresses the unique challenges faced by local governments in assessing territorial resilience and attractiveness while providing analytical tools for intervention design and impact forecasting. By integrating data management capabilities, advanced analytical models and user-friendly interfaces, the proposed DSS framework supports evidence-based decision-making in territorial governance.

The system's construction framework, depicted in Figure 1.1, is divided into four key phases.

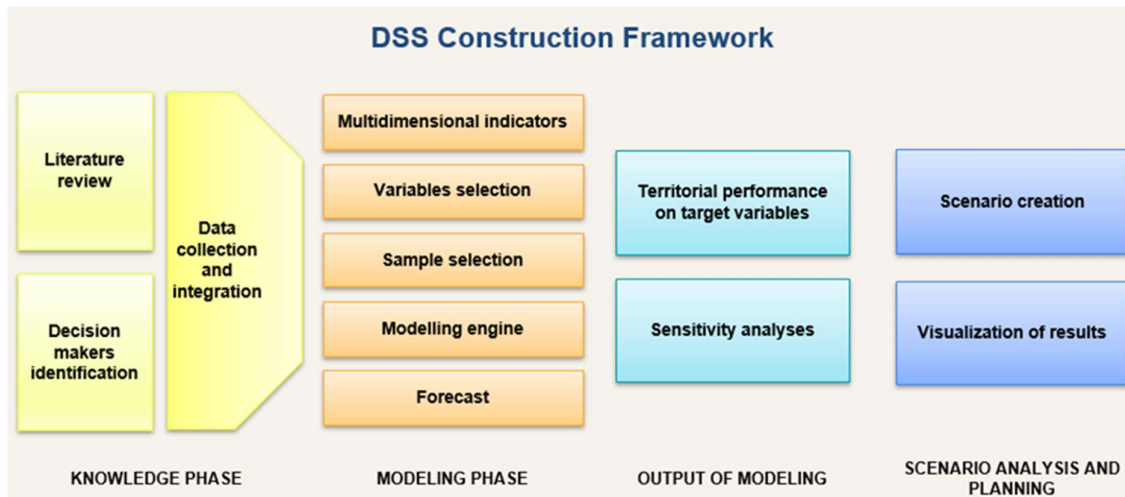


Figure 2.1 DSS Framework.

## 2.1 Knowledge Phase

The Knowledge Phase lays the groundwork for the DSS by collecting essential information and developing a solid knowledge base. A literature review was conducted to clarify the concepts of resilience and attractiveness and to build a model for measuring them, with particular attention to the role of infrastructure and services. This review proposes a conceptual framework linking drivers of resilience and outcomes, identifying key indicators for assessing territorial resilience and attractiveness. The framework defines primary dimensions of territorial capital: economic, social, human, and physical capital. This framework enables a structured assessment of the specific drivers that contribute to a territory's resilience and attractiveness. Each dimension is supported by distinct sub-dimensions and associated indicators:

1. **Economic Capital** includes resources that support financial stability, like income and wealth. It is further segmented into areas such as economic health (e.g., GDP), innovation capacity, and labor market dynamics.
2. **Social Capital** encompasses societal norms and values that facilitate social cohesion. This dimension includes aspects like civil society engagement and government quality.
3. **Human Capital** addresses factors contributing to personal development and societal health, including demographics, educational attainment, health infrastructure, and living conditions.
4. **Physical Capital** covers essential infrastructure and environmental assets supporting livelihoods, including built structures, natural resources, and facilities for utilities and services.

A second step in the Knowledge Phase is the identification of the DSS's final users. While several categories of decision-makers may benefit from the system, including policymakers, urban planners, and private-sector actors, we identified **local governments** as the primary stakeholders. This choice reflects the central role that DSS plays in supporting public administration: research shows that well-designed DSS can improve decision-making quality, enhance organizational

performance, and strengthen management control in local government contexts (Matheus et al., 2020). At the same time, local authorities often face challenges such as limited technical expertise, data integration difficulties, and the need for tools that deliver clear and easily interpretable outputs (Peignot et al., 2013).

The final step of the knowledge phase, the **data collection and integration** step. An opendata repository is proposed to store and monitor indicators over time, supporting the project's broader objectives (Amaddeo et al., 2024). Indicators for each dimension are sourced from reputable national and European databases, including ISTAT, EUROSTAT, and various government agencies. This robust data integration framework is crucial for constructing a detailed, multi-layered understanding of resilience and attractiveness across territories.

## 2.2 Modelling Phase

The Modelling Phase establishes an integrated analytical framework designed to evaluate territorial resilience and attractiveness using carefully selected metrics. This phase includes five interconnected components.

Firstly, to reflect the multifaceted nature of resilience and attractiveness, we define a set of multidimensional indicators that reflect the framework defined in the knowledge phase. Building on the multidimensional indicators, this step identifies specific variables that embody the core dimensions of resilience and attractiveness established in the initial knowledge phase. After selecting the variables, the sample selection process defines the model's spatial and temporal scope, identifying the regions, provinces, or cities to be analysed. The Modeling Engine then serves as the system's computational core, integrating the selected variables to generate predictive insights. Through statistical, econometric, or machine-learning methods, it produces scenarios that estimate how resilience and attractiveness may evolve under different conditions. The resulting forecasts provide decision-makers with a forward-looking view, supporting proactive planning and informed resource allocation.

## 2.3 Output of Modelling

The Output of Modelling phase translates model results into actionable insights through two main components. Sensitivity analyses examine how changes in specific variables influence resilience and attractiveness. These analyses highlight which factors exert the strongest effects, thus guiding decision-makers in prioritizing interventions and allocating resources more effectively.

## 2.4 Scenario Analysis and Planning

The final phase, Scenario Analysis and Planning, supports decision-makers in developing strategic interventions. The **Scenario Creation** component involves developing possible future states or intervention pathways. By testing these scenarios in the DSS, decision-makers can explore the probable outcomes of various strategies, compare their impacts on resilience and attractiveness, and make informed decisions about which paths to pursue.

## 3. DSS Function

Based on this construction framework and the identification of decision-makers' needs, the DSS is designed to provide a comprehensive overview of the attractiveness and resilience of Italian territories at various geographical scales. By **exploring** the relationship between attractiveness, resilience, and other territorial attributes, the DSS empowers decision-makers to identify key drivers of these outcomes and **predict** how changes in input variables may impact future performance.

### 3.1 Exploration

The DSS serves the purpose of exploring the level of association between variables. Specifically, the Exploration phase aims to analyze the relationships between territorial resilience/attractiveness and observable characteristics. The process involves five main steps (Figure 1.2).

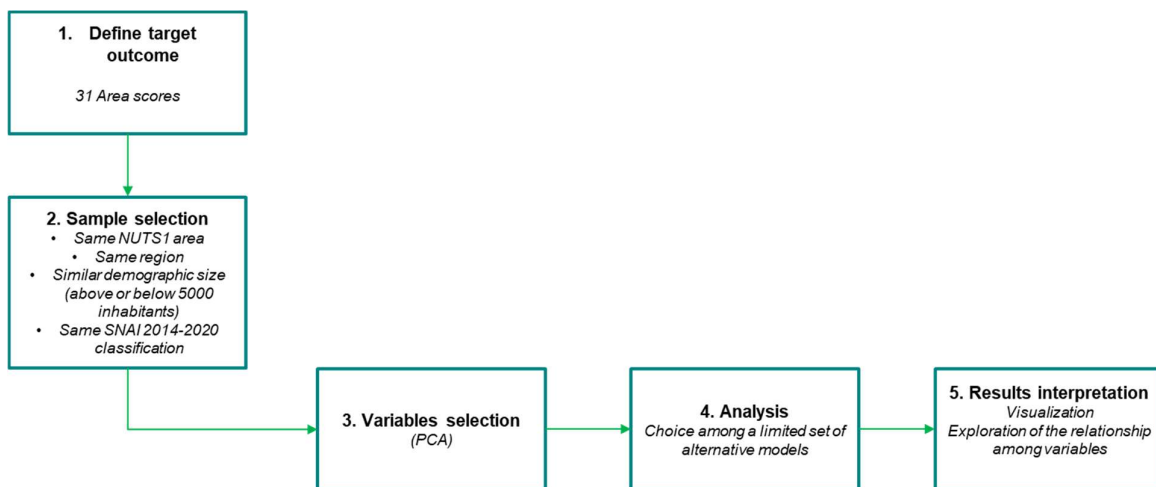


Figure 3.2: Workflow of the DSS aiming to explore the relationship among variables.

1. **Define Target Outcome:** In this phase, the user selects a target outcome variable. The DSS offers the possibility to choose it from among 31 indicators available in the data repository. Such variables represent “area scores” which capture key aspects related to the Economic, Human, Physical, and Social capital of Italian territories (see Table 1). Furthermore, they represent different dimensions of territorial resilience or attractiveness. Such scores are obtained based on multi-step aggregation strategy that includes Principal Component Analysis (PCA) and non-linear weighting to aggregate the specific indicators underlying each Area (Figure 1.3). For further details on the aggregation method and on the single indicators pertaining to each area, please refer to Amaddeo et al. (2024).

Table 1: The structure of the Area score.

Capital	Macroarea	Area
Economic capital	Economy	Availability of financial markets Disposable Income pc
	Labour market	Employment
Human capital	Demography	Population
	Education	University
	Health	Health services
Physical capital	Environment	Emissions
		Greenery/park coverage
		Water capacity
	Housing and household assets	Housing stock construction quality
	Infrastructure and services	Transportation access Density of water supply Tourist capacity Urban road per capita
Social capital	Ethnic integration	Foreign population
	Political participation	Voter turnout

## Step 2 - Seleziona una categoria

Selezionare una variabile nel albero. È possibile scegliere le variabili a tre livelli di dettaglio crescente: capitali, macroree ed aree.

- Capitale economico
  - Economia
    - Disponibilità di mercati finanziari
    - Reddito disponibile pro capite
  - Mercato del lavoro
- Capitale fisico
- Capitale sociale
- Capitale umano

Conferma categoria selezionata

Figure 1.3 Target variable selection.

**2. Sample Selection:** Subsequently, the user can select the sample that will be the subject of the analysis. In particular, the study can either include the entire dataset or be restricted to a sub-group of observations (Figure 1.4). For a given target municipality, the user can decide to perform a sample restriction as follows:

- Same NUTS1 area: municipalities belonging to the same macro-regional area of the target territory. NUTS1 category in Italy are North-West, North-East, Centre, South, Islands (example: Turin, Milan, and Genoa all fall within the North-West NUTS1 area).
- Same region: municipalities located within the same administrative region of the target territory (Example: Florence, Pisa, and Siena can be grouped because they are all in Tuscany).
- Similar areas based on simple clustering: the user will obtain a sample of territories that exhibit similar characteristics of the target territory based on three variables used for the cluster analysis (income per capita, land extension and total number of employed individuals).

This flexibility allows for targeted analysis within specific clusters or geographical areas, potentially improving the internal validity of the analysis.

Figure 1.4. Sample selection.

**3. Variables Selection:** Then, the DSS relies on Recursive Feature Elimination with Cross-Validation (RFECV) to isolate the subset of observable characteristics that most effectively explain territorial resilience or attractiveness (Figure 1.5). In this implementation, RFECV is performed with a Random Forest classifier and a stratified 5-fold cross-validation scheme, ensuring that feature selection is grounded in predictive performance rather than variance-based reduction. To avoid leakage, the procedure first splits the dataset into training and test sets and performs standardization strictly on the training data before fitting the selector. RFECV then iteratively removes the least informative variables, one at a time, evaluating each reduced subset through cross-validated accuracy. The process yields an optimal set of predictors that maximize classification performance on held-out folds, while also exporting the selected features, their rankings, and diagnostic plots. As a result, the analysis focuses only on variables that genuinely enhance the model's ability to discriminate territorial profiles, improving both interpretability and robustness.

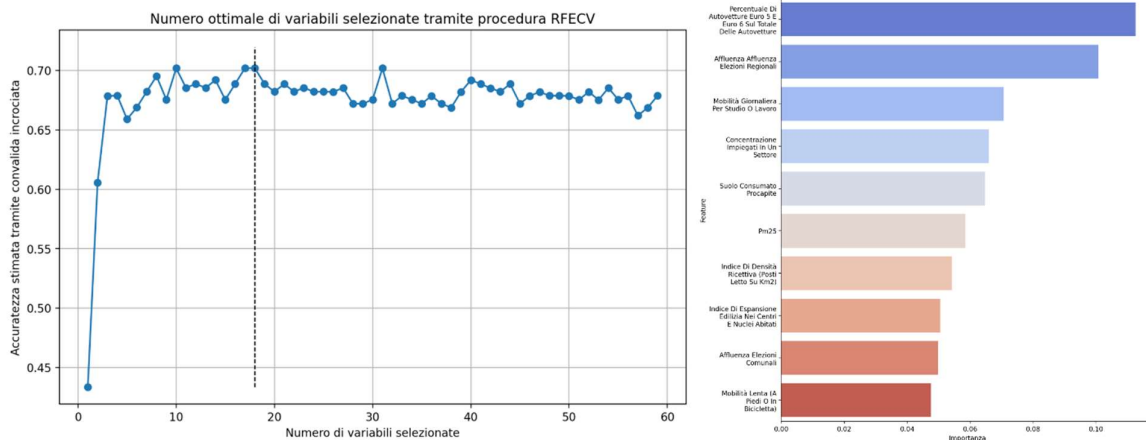


Figure 3.5: Results of RFECV the feature selection procedure. Left plot: accuracy improvement obtained by selecting features. Right plot: select features and importance.

**4. Machine learning model:** After identifying the optimal feature subset through RFECV, the analysis proceeds by training a Random Forest classifier using only the selected variables to predict four distinct classes, each representing a quartile of the target variable. This setup transforms a continuous territorial indicator into four ordered categories, enabling the model to distinguish territories positioned in the lowest, intermediate, and highest resilience or attractiveness groups. The classifier is trained on the standardized training data with 100 trees and a fixed random seed to ensure reproducibility. Once fitted, the model computes Gini-based feature importance scores that quantify how strongly each predictor contributes to separating the four quartile-based classes. Model performance on the unseen test set is assessed through class-specific precision, recall, and F1-scores, along with a confusion matrix that details how well the model assigns cases to the correct quartile. Together, these steps verify whether the RFECV-derived feature subset supports accurate and stable four-class predictions.

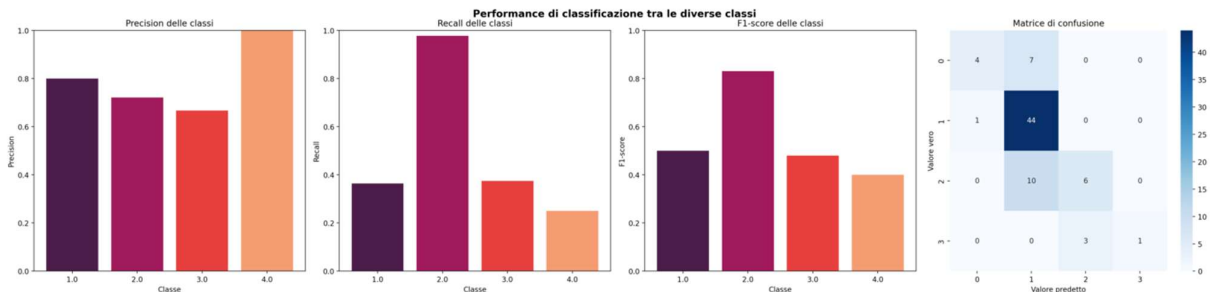


Figure 1.6: Classification performance across different classes.

**5. Results Interpretation:** Results are visualized, and the DSS explores the relationships between the outcome and explanatory variables. This includes examining the direction and magnitude of their interactions, allowing users to interpret how various characteristics impact resilience or attractiveness. The analysis pipeline generates a comprehensive set of outputs that provide complementary perspectives on territorial classification (Figure 1.6):

- **Feature Selection Plot:** Displays cross-validation accuracy across different feature subset sizes, identifying the optimal number of predictors needed to maximize classification performance while avoiding overfitting. The plot marks the point where adding more features yields diminishing returns.
- **Feature Importance Rankings:** the relationship between the explanatory and dependent variables can be explored based on Feature Importance. This measure can be interpreted as the mean decrease in impurity that each feature introduces across all the decision trees while constructing them. Impurity can be measured through a Gini Impurity index ( $G(S)$ ):

$$G(S) = \sum_{k=1}^K p_k * (1 - p_k)$$

where  $p_k = \frac{S_k}{S}$  representing the fraction of inputs in S with label k.

The plots report horizontal bar charts showing Gini-based importance scores from the Random Forest classifier, quantifying which variables most effectively distinguish territories across quartiles. These rankings reveal both primary discriminators (typically >0.10 importance) and secondary supporting features (0.05-0.10 range).

- **Classification Performance Report:** A visualization presenting per-class precision, recall, and F1-scores alongside a confusion matrix heatmap. This reveals the model's predictive accuracy for each quartile, highlights systematic misclassification patterns (e.g, adjacent quartile confusion), and identifies which classes are most reliably predicted versus those with ambiguous boundaries.
- **Feature Distribution Analysis:** A grid of kernel density estimates showing how the top 9 features distribute across quartiles. These winsorized distributions expose whether high-performing territories cluster around specific value ranges, whether variables exhibit monotonic trends across quartiles, or whether distinct subgroups emerge within classes—revealing skewness, overlaps, and threshold effects.
- **Class Feature Profiles:** A multiplot displaying boxplot distributions of the top 10 features for each quartile (Figure 1.7). These profiles provide distributional "fingerprints" that characterize typical feature value ranges per class, highlight interquartile spread and outlier presence, and enable identification of features whose discriminative power stems from median shifts versus distributional shape differences.
- **SHAP Analysis Visualizations:** Shapley Values constitute the average contribution of a feature value to the prediction in different coalitions, where a coalition is a combination of alternative features included in the model (Figure 1.8). The contribution of a feature to a coalition is estimated by computing the difference in the prediction of the outcome variable when the underlying feature is included or excluded from the set of features employed in the model.
  - **Summary Plots:** Beeswarm charts showing feature-level SHAP values colored by feature magnitude (red = high, blue = low), revealing directional relationships (positive/negative/non-linear) between features and quartile predictions.

- **Importance Bar Charts:** Mean absolute SHAP values ranked by feature, providing an instance-based complement to Gini importance that captures average prediction impact rather than tree-splitting utility.

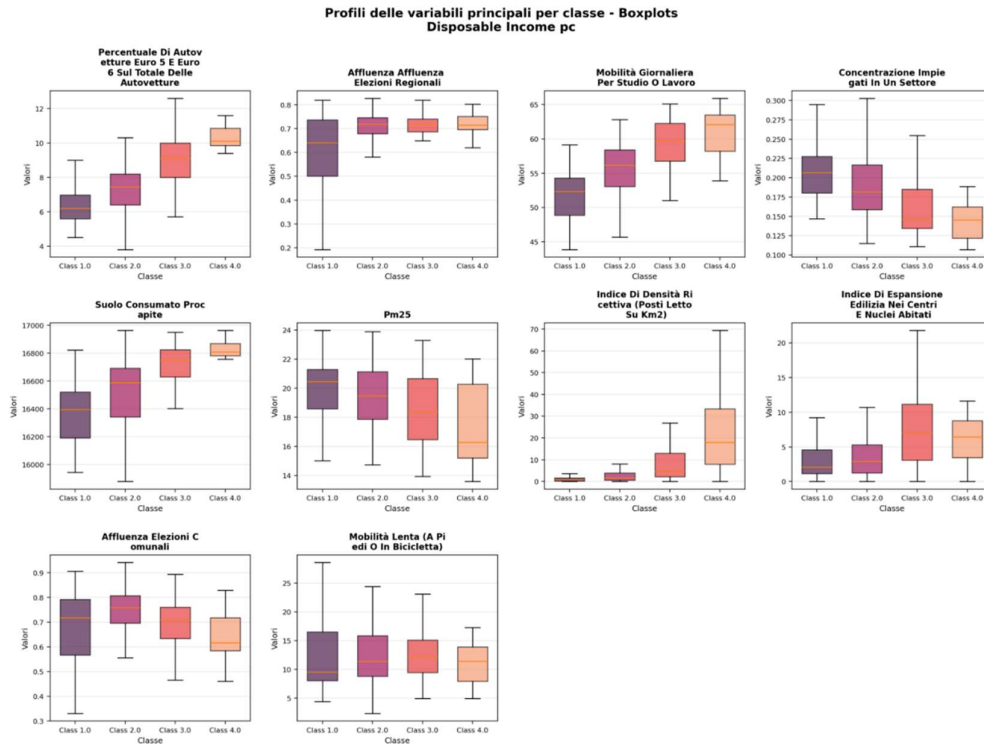


Figure 1.7: Class profiles.

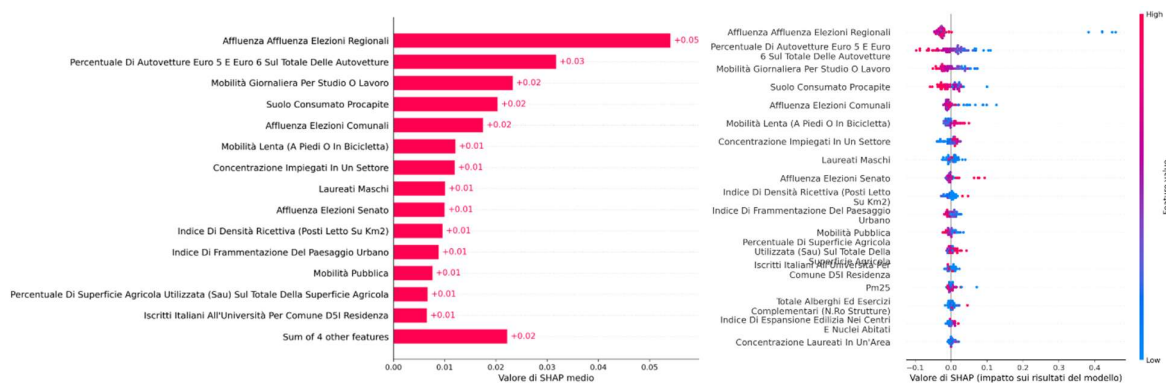


Figure 1.8: The left panel shows Feature Importance for the most important features used in the analysis. The right panel exhibits a Shapley Value of the most important features estimated through a Random Forest model.

- **Conclusive Report:** As final output the DSS will produce an automatically generated text document providing a comprehensive interpretation of results in three main sections:
  - Class Characterizations: Statistical profiles for each quartile showing percentile ranges (10th-90th), medians, and means for top features, with comparative rankings across classes to identify which quartiles exhibit the highest/lowest values.
  - Decision Rules: Human-readable classification pathways extracted from a simplified decision tree (max depth 5, entropy criterion), presenting conditions like "Feature  $\leq$  threshold" that lead to specific quartile predictions, along with rule purity (classification confidence) and coverage metrics.
  - Performance Assessment: Overall model quality rating (Excellent/Good/Acceptable/Needs Improvement based on accuracy thresholds), per-class precision/recall/F1-score analysis with actionable diagnostics (e.g., identifying false positive/negative patterns), key insights on discriminative features, class separation quality assessment, and operational recommendations regarding feature set efficiency, data collection priorities, and class imbalance mitigation strategies.



Figure 1.9: Left plot shows general information contained in the exploration report. Right plot shows class features.



Figure 1.10: Left plot shows decision rules obtained from the machine learning model. Right plot shows evaluation of the model performance.

### INTUIZIONI CHIAVE E RACCOMANDAZIONI

1. **CARATTERISTICHE DISCRIMINANTI CRITICHE:** Le prime 3 caratteristiche che meglio distinguono i livelli di Reddito disponibile pro capite sono:
  1. *Percentuale Di Autovetture Euro 5 E Euro 6 Sul Totale Delle Autovetture* (importanza: 11.3%)
  2. *Affluenza Affluenza Elezioni Regionali* (importanza: 10.1%)
  3. *Mobilità Giornaliera Per Studio O Lavoro* (importanza: 7.1%)
2. **SEPARAZIONE DELLE CLASSI:** La separazione delle classi è robusta. Le classi sono ben differenziate sulla base delle caratteristiche chiave.
3. **RACCOMANDAZIONI OPERATIVE:**
  - o ✓ Il modello è affidabile per identificare territori a diversi livelli di Reddito disponibile pro capite e può informare decisioni politiche.
  - o ✓ Un insieme compatto di 18 variabili è sufficiente, rendendo la raccolta dei dati efficiente.

FINE DEL REPORT

Figure 1.11: Final section of the report with general intuition of the analysis and policy conclusions.

These outputs collectively enable analysts to identify decisive predictors and more moderate contributors through importance rankings, while distributional visualizations such as density plots and boxplots highlight how territorial indicators vary across the four quartile-based classes. SHAP's local explanations further uncover non-linearities, interaction effects, and the specific mechanisms driving individual predictions, and the performance metrics provide a clear assessment of the model's reliability and recurring misclassification patterns. Complementing these analytical elements, the automatically generated report synthesizes the results into an accessible interpretive document: it outlines the statistical profiles that differentiate each quartile, presents human-readable decision rules derived from a simplified decision tree to clarify how the model arrives at specific predictions, and evaluates overall and class-specific performance with concrete diagnostic insights and operational recommendations (Figures 1.9, 1.10, 1.11). Overall, these outputs provide a coherent picture of how the model prioritizes variables, distinguishes between the quartile-based classes, and explains its own predictions through performance diagnostics and synthetic summaries. Taken together, these components help clarify the underlying patterns that drive class separation and shape the model's predictive distinctions, while also offering a structured interpretation to guide policy analysis and decision-making.

## 3.2 Prediction

The DSS also serves the purpose of identifying the primary drivers of selected outcome variables. In addition, the Prediction service aims to forecast changes in the outcome variable based on input variations. Two applications were developed to test this functionality, focusing on PM2.5 concentrations and tourism flows. The workflow illustrated in Figure 1.12 applies to both the PM2.5 and the tourism prediction services, as they share the same analytical structure and sequence of operational steps. For simplicity, the application related to PM2.5 is presented in detail. The tourism prediction service follows a similar DSS logic, allowing users to select tourism-related outcome variables, such as total, domestic, or international tourist presences, and to apply the same spatial regression and scenario analysis techniques.

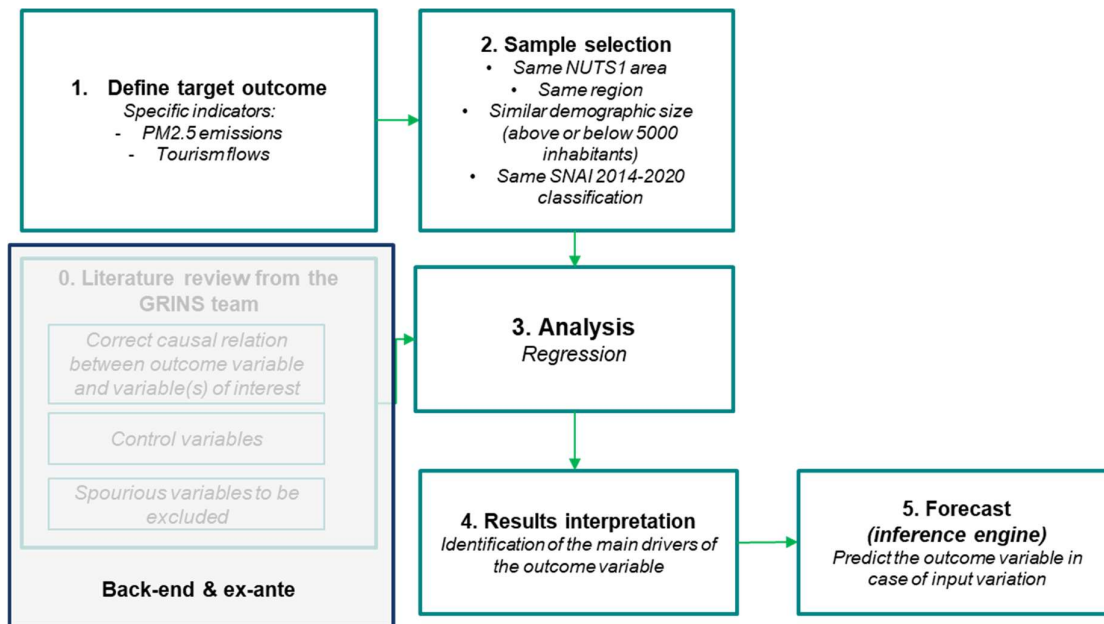


Figure 3.12: Workflow of the DSS aiming to identify the main drivers of the outcome variable and predict how variation of such factors may affect the variation of the dependent variable.

- 1. Define Target Outcome:** The user can select a target outcome variable. In this case, the selection is restricted to two indicators: PM2.5 emissions and tourism at the local level. The definition of the outcome variables available in the prediction function is based on prior literature review which informs these indicators, ensuring they have causal links to significant drivers. This background research helps maintain the econometric model's focus on relevant causal factors and avoids spurious variables. In this phase, the system asks for confirmation to load the input data (Figure 1.13). Both datasets are already stored in Amelia and only need to be imported into the DSS, not uploaded by the user.

## Servizio 1 – Predizione

### Step 1 – Carica i dati di input

Carica i dati di input per iniziare l'analisi. Premi il pulsante qui sotto per procedere.

---

### Stato del Servizio

Di seguito è riportato lo stato di completamento di ciascun passaggio del servizio.

Completato | 
  Non completato

Step	1	2	3	4	5	6	7	8
Stato	✗	✗	✗	✗	✗	✗	✗	✗

Figure 1.13 – DSS: Input data upload

2. **Sample Selection:** The user can decide to perform a sample restriction either based on different grouping criteria (Figures 1.14 and 1.15). Municipalities can be grouped according to the following:

- Same NUTS1 area: municipalities belonging to the same macro-regional area. NUTS1 category in Italy are North-West, North-East, Centre, South, Islands (example: Turin, Milan, and Genoa all fall within the North-West NUTS1 area).
- Same region: municipalities located within the same administrative region (Example: Florence, Pisa, and Siena can be grouped because they are all in Tuscany).
- Similar demographic size: municipalities with comparable population levels, above or below 5,000 inhabitants. (Example: Small towns such as Cogne and Castel del Monte can be grouped together; similarly, larger municipalities over 5,000 inhabitants form another group).
- Same SNAI 2014–2020 classification: municipalities sharing the same territorial role and degree of accessibility. SNAI categories in Italy are: *Urban pole (Polo)*, *Inter-municipalities pole (Polo intercomunale)*, *Surrounding area (Cintura)*, *Intermediate area (Area intermedia)*, *Peripheral area (Area periferica)*, *Ultra-peripheral area (Area ultraperiferica)* (Example: Matera and Gangi can be grouped as “peripheral areas”, while Bologna and Bari belong to the “urban poles”).

For the tourism outcome selected, sample restrictions are applied to account for the more limited coverage of municipal-level tourism data available from official statistics. As tourism indicators are available for a smaller subset of municipalities compared to environmental data, only macro-regional groupings (NUTS1, with North-West, North-East,

Centre, South, Islands) and SNAI classifications are employed, to ensure sufficient sample size and statistical reliability.

The screenshot shows the 'Servizio 1 - Predizione' interface. It has a 'Reset Servizio' button in the top right. The main heading is 'Step 1 - Carica i dati di input', followed by the instruction 'Carica i dati di input per iniziare l'analisi. Premi il pulsante qui sotto per procedere.' and a 'Carica dati di input' button. Below this is 'Step 2 - Seleziona una categoria', with the instruction 'Scegli un'opzione:'. A dropdown menu is open, showing four options: 'Analisi su 1 regione' (selected), 'Analisi su macro-area (NUTS1)', 'Analisi per dimensione demografica (< o > 5000 abitanti)', and 'Analisi per classificazione SNAI 2014-2020'.

Figure 1.14 DSS: Sample selection criteria

This screenshot shows the same 'Servizio 1 - Predizione' interface. The 'Step 1' section is identical. In 'Step 2 - Seleziona una categoria', the dropdown menu is closed, and 'Analisi su 1 regione' is selected. Below it, a new dropdown menu is labeled 'Seleziona la regione' and has 'Lombardia' selected. At the bottom of the step, there is a 'Conferma categoria selezionata' button.

Figure 1.15: DSS: Sample selection

This flexibility allows for targeted analysis within specific clusters or geographical areas, potentially improving the internal validity of the analysis.

Given the nature of tourism variables, the tourism module subsequently allows users to select the specific outcome measure among total, domestic, or international tourist presences.

At the end of this phase, the DSS provides an overview of the selected sample (Figure 1.16).



**Step 1 — Carica i dati di input**

Carica i dati di input per iniziare l'analisi. Premi il pulsante qui sotto per procedere.

Carica dati di input

**Step 2 - Selezione di una categoria**

Scegli un'opzione:

Analisi su 1 regione

Seleziona la regione:

Lombardia

Conferma categoria selezionata

**Step 3 — Riepilogo dati selezionati**

Hai selezionato: **Analisi su 1 regione**

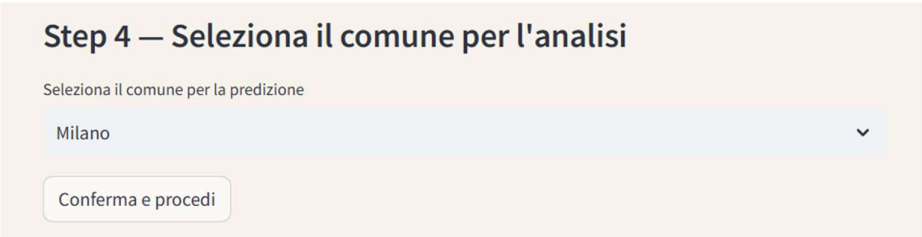
Hai scelto regione: **Lombardia**

Numero di osservazioni selezionate: **5996**

Procedi al passo successivo

Figure 1.16 DSS: Recap of selected sample

The user selects the municipality of interest for the analysis from the sample defined above (Figure 1.17).



**Step 4 — Selezione del comune per l'analisi**

Seleziona il comune per la predizione

Milano

Conferma e procedi

Figure 1.17 DSS: Selection of the municipality for the prediction.

In this phase the DSS in both modules provides an overview of the analyzed area as it was selected in the previous phase (Figure 1.18).

### Step 5 — Visualizzazione dell'area analizzata e stima del modello ↔

Esegui predizione

Comune selezionato per l'analisi: Milano

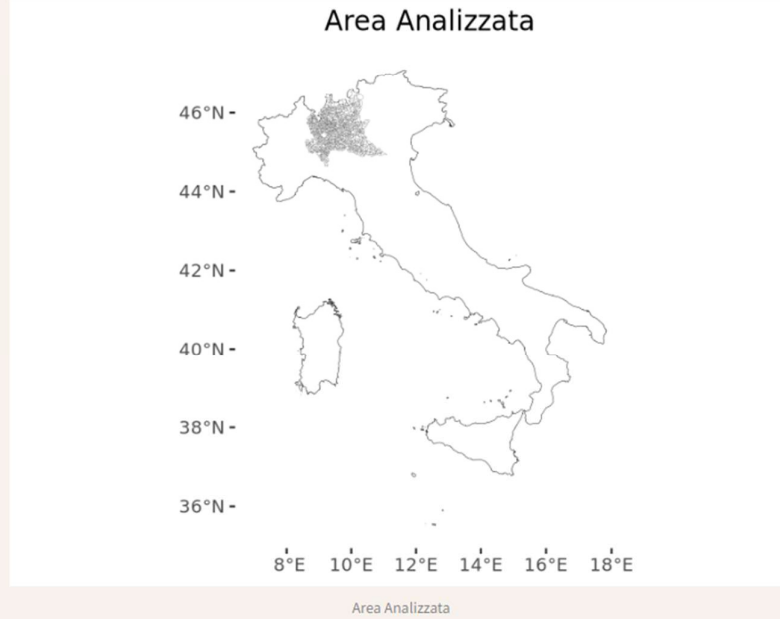


Figure 1.18 DSS: Overview of the selected area

3. **Model Selection:** The DSS relies on the following Spatial Durbin Model (SDM):

$$Y_{i,t} = \lambda \sum_j w_{i,j} * Y_{j,t} + \alpha + \beta X_{i,t} + \gamma \sum_j w_{i,j} * X_{j,t} + \delta M_{i,t} + \varphi S_{i,t} + \varepsilon_{i,t}$$

where  $Y_{i,t}$  is the PM2.5 variation observed in municipality  $i$  in year  $t$ ,  $X_{i,t}$  is the vector of our variables of interest including EU green funds, as well as current expenditures and investments in energy diversification at the municipality level. We also include the autoregressive term  $\sum_j w_{i,j} * Y_{j,t}$  which allows us to capture the fact that the variation of PM2.5 in a municipality may correlate with PM2.5 dynamics in neighbor administrative units. Moreover, we include the spatial spillovers of our variables of interest ( $\sum_j w_{i,j} * X_{j,t}$ ) since the amount of EU green funds, current expenditures and investments in energy diversification in a municipality may affect air quality in neighbor administrative units. Moreover, we encompass a set of control variables to raise the robustness of our findings. In particular,  $M_{i,t}$  is a vector of meteorological variables including wind speed, humidity, rainfall and temperature. Finally,  $S_{i,t}$  is a vector of socio-economic variables encompassing employees in the manufacturing sector, population density, and income per capita.  $\varepsilon_{i,t}$  is the idiosyncratic term. The models is estimated over the timeframe 2017-2020.

We test the robustness of our results through the spatial exponential decay matrix  $W = \{w_{i,j}\}_{i,j=1}^n$  (with  $n$  being the number of municipalities in our model):

$$w_{i,j} = \exp(-\alpha d_{i,j})$$

where  $d_{i,j}$  is the distance in kilometers between municipality  $i$  and  $j$  and  $\alpha$  is the spatial

decay parameter. We use a value equal to 2 for  $\alpha$  since it is the value minimizing the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for most analysed samples.

For the tourism application, the same Spatial Durbin Model specification is adopted, with the outcome variable  $Y_{i,t}$  representing the number of tourist presences observed in municipality  $i$  in year  $t$ . Depending on the user's selection, the dependent variable refers to total, domestic, or international tourist presences. The vector  $X_{i,t}$  includes tourism-related expenditure variables under the control of local governments, such as municipal tourism expenditure, together with their spatial spillovers, capturing the effect of tourism-related spending in neighboring municipalities. As in the PM2.5 application, the model accounts for spatial dependence through the inclusion of both the spatially lagged dependent variable and the spatially lagged explanatory variables. Additional control variables are included to ensure robustness of the estimated relationships by accounting for structural, infrastructural, and cultural factors that influence tourism demand. Specifically, the model controls for accessibility in minutes for airports and ports, for cultural and territorial attractiveness through UNESCO heritage sites, officially recognized villages, quality or environmental labels, and for environmental and landscape features through the share of municipal territory located within protected natural park areas. These variables capture exogenous determinants of tourism flows, allowing the model to better isolate the effect of tourism-related public expenditure and its spatial spillovers on tourist presences.

**4. Results Interpretation:** The interpretation of the results relies on the estimated coefficients, which primarily inform the direction of the association between each predictor and PM2.5 concentrations. Positive coefficients indicate that an increase in the corresponding variable is associated with higher PM2.5 levels, whereas negative coefficients suggest a reduction in PM2.5. Moreover, the magnitude of the coefficients provides insight into the extent to which each variable contributes to variations in PM2.5. These results are translated into natural language to ensure they are easily understandable to all users. Overall, this step provides clear insights into how different factors influence the target outcome, enabling users to make informed, evidence-based decisions. Examples of the results provided include the following:

- “The variable Spillover EU funds expenditure has a negative impact on PM2.5. The coefficient is: -351.1217.”
- “The variable Spillover current expenditure for energy diversification has a positive impact on PM2.5. The coefficient is: 0.4468.”
- “The variable Spillover investment in energy diversification has a negative impact on PM2.5. The coefficient is: -0.2443.”
- “The variable Employees in the manufacturing sector has a negative impact on PM2.5. The coefficient is: -0.0032.”
- “The variable Population density has a positive impact on PM2.5. The coefficient is: 0.0113.”
- “The variable Wind speed has a positive impact on PM2.5. The coefficient is: 0.0326.”
- “The variable Humidity has a negative impact on PM2.5. The coefficient is: -0.1006.”
- “The variable Temperature has a positive impact on PM2.5. The coefficient is: 0.1304.”

- “The variable Precipitation has a positive impact on PM2.5. The coefficient is: 0.1592.”

For the tourism application, results are interpreted following the same logic. Estimated coefficients indicate the direction and strength of the association between predictor and tourist presences. Positive coefficients signal an increase in tourism flows, while negative coefficients indicate a reduction. As in the PM2.5 case, results are automatically translated into natural language to ensure accessibility for non-technical users. The DSS highlights statistically significant drivers and provides qualitative interpretations of their effects on tourism dynamics. Examples of the output generated for the tourism application include the following:

- “Municipal tourism expenditure has a significant positive effect on tourist presences.”
- “Tourism expenditure in neighboring municipalities has a significant effect on local tourist presences, capturing spatial spillover dynamics.”
- “The presence of cultural and territorial assets, such as UNESCO heritage sites and recognized historic villages, is associated with higher tourist presences.”

- 5. Forecast:** The DSS enables users to assess how variations in selected input variables influence the outcome variable. Predictions can be generated by allowing multiple variables—provided they are under the decision-maker’s direct control and were found to be statistically significant in the preceding estimation step—to vary simultaneously.

In the case of PM2.5, the three input expenditure variables considered are all under the control of the decision-maker (i.e., the local government): current expenditure, investment or EU funding. The user selection option is activated when at least one of the three expenditure variables (current expenditure, investment, or EU funding) is statistically significant. If more than one variable is significant, the DSS allows the user to decide freely whether to adjust all significant variables, or only one of them, according to their needs. An interactive panel opens allows the user to set the desired change in expenditure by specifying the numerical value for the selected variable. The user can choose to input these variations either as absolute values or as percentage changes (Figure 1.19).

In the tourism application, the forecasting logic follows the same structure, with predictions generated by varying only statistically significant expenditure variables under the control of local decision-makers. In this case, following the literature, a single aggregated category of tourism-related public expenditure is considered, combining municipal spending on tourism, environmental and cultural protection, and local public transport, including both current expenditure and investments. When significant, the DSS allows users to introduce percentage variations in aggregated municipal tourism-related expenditure as well as in corresponding expenditure of neighboring municipalities, capturing both direct and spatial spillover effects. Unlike the PM2.5 application, variations are specified exclusively in percentage terms, reflecting the nature and scale of tourism-related expenditure data. The resulting forecasts are translated into natural language, enabling users to evaluate the expected impact of alternative tourism policy scenarios.

## Step 7 – Inserisci le variazioni per le variabili significative

Inserisci variazioni per tutte le variabili significative ^

Per ciascuna variabile significativa, inserisci il valore e seleziona il tipo di variazione:

Valore per 'Investimenti diversificazione energetica'	Tipo di variazione per 'Investimenti diversificazione energetica'
0,00 - +	<input type="text" value="0,00"/> <input type="text" value="-"/> <input type="text" value="+"/>
Valore per 'Spesa corrente diversificazione energetica'	<input type="text" value="0,00"/> <input type="text" value="-"/> <input type="text" value="+"/>
Valore per 'Spesa fondi EU'	Tipo di variazione per 'Spesa fondi EU'
0,00 - +	<input type="text" value="0,00"/> <input type="text" value="-"/> <input type="text" value="+"/>

OK per procedere

Figure 1.19 DSS: Input selection for the prediction

The DSS presents the forecasting results by showing how changes in the input variables affect the outcome variable. In the case of PM2.5, the predicted variation resulting from changes in expenditure levels is expressed in natural language to facilitate comprehension (Figure 1.20). This forecasting capability allows stakeholders to assess the potential effects of different scenarios, thereby supporting strategic and evidence-based decision-making.

Similarly, for the tourism application, the predicted variation in tourist presences resulting from changes in tourism expenditure is expressed in natural language, enabling users to evaluate the effects of alternative tourism policy scenarios.

## Step 7 – Inserisci le variazioni per le variabili significative

Inserisci variazioni per tutte le variabili significative

Per ciascuna variabile significativa, inserisci il valore e seleziona il tipo di variazione:

Valore per 'Investimenti diversificazione energetica'	Tipo di variazione per 'Investimenti diversificazione energetica'
10,00 - +	Percentuale
Valore per 'Spesa corrente diversificazione energetica'	Tipo di variazione per 'Spesa corrente diversificazione energetica'
0,00 - +	Percentuale
Valore per 'Spesa fondi EU'	Tipo di variazione per 'Spesa fondi EU'
0,00 - +	Percentuale

OK per procedere

## Step 8 – Applica la variazione e visualizza il risultato

Applica variazioni

La variazione del PM2.5 è diminuita di 0.03 %

Figure 1.20 DSS: Predicted variation of PM2.5

## References

Amaddeo, E., Arena, M., Bergantino, A.S., Bonaccorsi, G., Buongiorno, A., Clemente, A., Flori, A., Intini, M., Scotti, F., Urbano, V.M. and Vitagliano, M. (2024). A set of multidimensional indicators to assess the resilience and attractiveness of Italian provinces and municipalities (2010–2022 panel data). *Data in Brief*, 111042.

Matheus, R., Janssen, M., & Maheshwari, D. (2020). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*, 37(3), 101284.

Peignot, J., Peneranda, A., & Amabile, S. (2013). Strategic decision support systems for local government: A performance management issue? The use of information systems on the decision-making and performance management of local government. *International Business Research*, 6(2), 92.

Pierri, F., Scotti, F., Bonaccorsi, G., Flori, A., & Pammolli, F. (2023). Predicting economic resilience of territories in Italy during the COVID-19 first lockdown. *Expert Systems with Applications*, 232, 120803.