Environmetrics

WILEY

RESEARCH ARTICLE OPEN ACCESS

Estimating the Impact of Socioeconomic Drivers on Land Degradation in Italy via Spatio-Temporal Additive Expectile Regression

Luca Merlo[1] | Beatrice Foroni[2] | Ioannis Konaxis[3] | Lea Petrella[2] | Luca Salvati[2]

[1]Department of Human Sciences, Link Campus University, Rome, Italy | [2]MEMOTEF Department, Sapienza University of Rome, Rome, Italy | [3]Department of Tourism Studies, University of Piraeus, Piraeus, Greece

Correspondence: Luca Merlo (l.merlo@unilink.it)

Received: 31 May 2025 | **Revised:** 9 February 2026 | **Accepted:** 13 February 2026

Keywords: boosting | desertification risk | ESAI | Italy | spatio-temporal expectile regression | stability selection

ABSTRACT

Climate changes, soil degradation, and desertification increasingly threaten the entire territory of Italy due to the complex interplay between natural processes and anthropogenic forces. Motivated by this pressing issue, this paper investigates how land-use and socioeconomic drivers have shaped desertification dynamics across Italian provinces between 1960 and 2020 using an additive expectile regression model for spatio-temporal data. The effects of continuous covariates are modeled as the sum of a linear part and a nonlinear smooth function using penalized B-splines, while tensor-product B-splines based on coordinate information are used to capture spatial dependencies. To promote sparsity in the model, estimation is carried out via component-wise boosting combined with a stability selection approach, retaining only the most dominant factors. The results identify industrialization and tourism development activities as the primary amplifiers of desertification risk in vulnerable regions. Our analysis also uncovers strong spatial heterogeneity of desertification processes related to province characteristics, which is particularly acute in southern Italy and in the major islands.

1 | Introduction

Environmental variables not only trigger adverse effects but also create significant obstacles for statistical modeling because of their empirical characteristics. Alongside its environmental dimension, land degradation—taken as a representative socioecological process—has significant economic, demographic, and territorial implications (Akhtar-Schuster et al. 2017). Degraded land because of a progressive depletion of soil quality exerts a detrimental influence on regional economies and local development processes (Romm 2011). More generally, desertification risk undermines the ability of communities to sustainably exploit natural resources and limits opportunities for long-term growth (Lanfredi et al. 2022). While reflecting the interplay between the economic sphere and ecological systems, mitigation of (and adaptation to) land degradation is a direct challenge outlined in the Zero Net Land Degradation strategy and Sustainable Development Goals promoted by the United Nations (Dasgupta et al. 2006; Cowie et al. 2018; Elnashar et al. 2022). Understanding land degradation risk and the underlying desertification processes requires an accurate (spatially explicit) assessment of specific land characteristics, including agricultural practices, territorial constraints, and place-specific socioeconomic characteristics (Vu et al. 2014), with the final aim at identifying vulnerable areas and delineating appropriate mitigation and adaptation policies (Hill et al. 2008; Gupta et al. 2020; Imbrenda et al. 2021). The importance of the spatial dimension also depends on the established risk of environmental pressures, mainly of human

origin (Benassi et al. 2020). While most of the identified drivers were demonstrated to act through indirect (off-site) transmission channels, few of them were regarded as direct (on-site) factors of land degradation (e.g., Cowie et al. 2018; Smiraglia et al. 2019; Morianou et al. 2021), making impact assessment analysis a particularly challenging task for land degradation processes. In this perspective, it should be noted how the empirical distributions of many relevant variables, especially socioeconomic indicators deemed critical in shaping the complex geography of land degradation in high-income countries, are often heavily skewed, leptokurtic, and thus exhibiting extreme values (Salvati et al. 2011). Such a process may also bias any statistical procedure aimed at summarizing the intrinsic content of non-redundant indicators into a unique, composite index of land vulnerability to degradation (e.g., Zambon et al. 2018). A correct application of a theoretical and empirical framework stemming from indication theory and composite information summary provides the necessary knowledge base to a particularly complex, multi-dimensional process, such as land degradation (e.g., Cecchini et al. 2019), frequently assumed as nonlinear in the original biophysical (and especially socioeconomic) drivers (Zambon et al. 2017).

Moreover, spatial interactions among regions and locations possibly follow complex, nonlinear dependencies, and temporal heterogeneity adds to this complexity due to the interplay of anthropogenic activities and atmospheric dynamics (Hill et al. 2008). Taken together, these aspects suggest considering a very articulated research design based on the interplay of time and space dimensions (Canfora et al. 2017). Appropriate spatial scales, mimicking different policy levels, and representative time points of different socioeconomic settings should illuminate such a theoretical perspective with the necessary results of an empirical analysis of relevant (stylized) facts—also in light of global change and climate warming (e.g., Biasi et al. 2019). While several causal chains with direct (or indirect) impact on land quality have been hypothesized and empirically tested, advanced econometric and non-traditional (spatially and/or temporally explicit) approaches (e.g., based on expectile regressions) have been rarely applied to socioeconomic drivers (Gupta et al. 2020). Our study sheds specific light on this literature, exploiting the potential of such econometric approaches to identify the most significant drivers of land degradation, and to determine their (point or latent) effect to guide policies towards greater effectiveness of intervention (Lanfredi et al. 2022).

From a methodological standpoint, traditional approaches to vulnerability assessment tend to focus on shifts in the conditional mean of a climate indicator—for instance, changes in average temperature to gauge desertification risk (Seifollahi-Aghmiuni et al. 2022). However, such measures of central tendency can miss critical behavior in the tails of the distribution, where the most severe hazards arise (Colantoni et al. 2015). Indeed, community characteristics may exert their strongest influence (in terms of magnitude, sign or significance) not on average conditions (like frequently explored in recent literature: See, for instance, Salvati and Zitti 2008; Ferrara et al. 2016; Francaviglia et al. 2019), but on extreme exposures—that is, on the upper tails of climate variables that drive disproportionate impacts (e.g., Incerti et al. 2007). To address these limitations, one possible approach is to model the entire conditional distribution of a response variable as a function of covariates through expectile regression. Expectile regression, introduced by Newey and Powell (1987), is a generalization of standard mean regression based on an asymmetric squared loss function, and offers richer insights on how their effect varies at different parts (the specific expectile) of the response distribution. Although expectiles have a less intuitive interpretation than conventional quantiles (Jones 1994), they offer additional substantial benefits. In particular, the asymmetric squared loss yields a distribution-sensitive characterization of the response that is especially informative when attention focuses on the tails of the conditional density (Tyralis et al. 2023). Moreover, because this loss function is continuously differentiable, parameter estimates can be obtained efficiently by iterative weighted least squares algorithms (Waldmann et al. 2017). From a computational perspective, expectile regression ensures uniqueness of the Maximum Likelihood solutions, and provides greater stability, which is of particular value in complex regression specifications including nonlinear, random, or spatial effects (Sobotka and Kneib 2012). These advantages have spurred a wide range of applications in different fields, such as longitudinal data analysis (Tzavidis et al. 2016), life expectancy (Nigri et al. 2022), economics and finance (Gerlach and Chen 2015; Taylor 2008; Kim and Lee 2016; Foroni et al. 2024). In a spatial context, Sobotka and Kneib (2012) introduced geoadditive expectile regression in an additive regression setting that includes spatial information. More recently, to capture both smooth spatial and temporal variations, Spiegel et al. (2020) proposed a semiparametric spatio-temporal extension with interaction terms based on P-splines (Eilers and Marx 1996).

Building on this literature, the novelty of the present study is to link the spatio-temporal distribution of a policy-relevant index of land sensitivity, called the Environmental Sensitive Area Index (ESAI), across all Italian provinces over three epochs (1960, 1990 and 2020) to a rich set of socioeconomic attributes, with emphasis on the territories more exposed to desertification risk (Salvati et al. 2015). Although Italy's Mediterranean climate is rich in natural resources (Kairis et al. 2022), the country is not immune to desertification risk (Colantoni et al. 2015), but it remains particularly vulnerable to land degradation as underscored by its inclusion in Annex IV of the Global Convention of Desertification. To disentangle these complexities and account for the characteristics of such complex data structure, using the ESAI framework, we consider a semiparametric additive spatio-temporal expectile regression model using P-splines. Because nonlinear relationships between covariates and the response variable occur (e.g., Salvati et al. 2011), continuous covariates are represented as the sum of a linear component and a smooth nonlinear deviation. The spatial effects are modeled using bivariate smoothing B-splines expressed in terms of the tensor product of the marginal spline basis of coordinate information, with a second-order difference penalty to promote smoothness.

Estimation proceeds through an iterative component-wise gradient boosting procedure (Bühlmann and Yu 2003). The algorithm involves fitting sequentially simpler models that depend on one or more features (referred to as base learners), and selecting those that contribute the most to reducing the loss with each iteration. This strategy is computationally efficient and proves particularly well-suited to capture intricate patterns, relationships, and nonlinear interactions among variables in high-dimensional settings with a large number of features. By selecting a subset of features

at each step, component-wise boosting inherently encourages sparse solutions. To formally identify only the most important factors fostering the evolution of land vulnerability in Italy, we integrate this approach with the stability selection procedure proposed by Meinshausen and Bühlmann (2010) to reduce the risk of selecting noise variables. Finally, inference on the partial effects of the predictors and spatial components included in the model is carried out via point-wise 95% confidence bands using a nonparametric bootstrap approach.

To our knowledge, this is the first time an additive space-time expectile model has been used for the analysis of desertification dynamics in Italian provinces. This approach allows to identify neighborhood-specific factors that influence vulnerability and resilience, particularly in areas where desertification severity is most acute, offering targeted guidance for adaptive land management strategies.

The remainder of this paper is organized as follows. In Section 2, we illustrate the ESAI framework and the data in detail. We formally introduce the model considered and the estimation procedure in Section 3. In Section 4, we analyze the risk of desertification in Italian provinces over the period 1960–2020 and discuss the results. Finally, Section 5 concludes.

2 | Operational Framework

2.1 | Study Area

The study area selected in this study coincides with the Italian territory (nearly 302,000 km^2, of which 23% lowlands, 42% uplands, and 35% mountains), which reflects a diverse mix of complex metropolitan structures, rural districts, economic patterns, sociodemographic dynamics, and environmental conditions. The economic gap between northern-central and southern regions reflects a long-established industry-service dichotomy (Salvati et al. 2011). Northern Italy includes accessible and flat areas within the Po Valley (Sallustio et al. 2018), displaying a particularly high level of per-capita income (significantly higher than the corresponding European Union average) and an efficient (and highly diversified) mix of primary, secondary, and tertiary activities, giving rise to a considerable (capital and labor) productivity, especially in industrial districts of Lombardy, Veneto and Emilia-Romagna regions. The Apennines mountain range separates northern Italy from central Italy, a region showing a polarization in urban and rural districts and a diversified structure centered on small-scale manufacturing, tourism, and high-quality agriculture (Coluzzi et al. 2022). These regions play a significant role in the country's economy thanks to a balanced mix of agglomeration and vertical integration (Quaranta et al. 2020), competition and collaboration, trust relationships over formal contracts, effective production systems, innovation, and improved accessibility (Lanfredi et al. 2022). Touristic poles, such as Rome and Florence, but also the landscape of Tuscany, and the medieval villages of Umbria and Marche regions—renowned tourism attractors all over the world—are located in this macro-region. Conversely, southern Italy is a marginal context with a slightly younger population than the rest of the country, and a productive structure focused on low-income agriculture and traditional tertiary activities (i.e., commerce) concentrated in urban centers (Imbrenda et al. 2021). The Mediterranean climate, characterized by prolonged summer droughts, locally constrained availability of water for irrigation, poor soil quality in some locations and a relatively large amount of steep land—less suitable for high-value, intensive agricultural activities, as well as industry and services, because of low accessibility from Europe—may reduce the attractiveness of this macro-region to economic capital and foreign mass tourism. Personal wealth, economic capital accumulation, and local unemployment levels in the three Italian macro-regions described above clearly reflect such divides (Salvati and Zitti 2009).

2.2 | Evaluating Land Degradation

The Environmentally Sensitive Areas (ESA) is one of the most widely used frameworks to assess the level of land vulnerability in Europe, North Africa, and the Middle East (Salvati et al. 2011; Ferrara et al. 2020; Morianou et al. 2021). Being used both for empirical studies and for ecological reporting (Elnashar et al. 2022), it produces a composite index of land exposure to degradation known as the Environmental Sensitive Area Index (ESAI). Taken as representative dynamics in the study area, climate change, landscape transformations and population growth, together with unsustainable land management and agricultural practices (Smiraglia et al. 2019), are responsible for triggering large-scale processes of land degradation in Italy, and are expected to drastically augment pressure on lands (Lanfredi et al. 2022). The ESA approach assesses land vulnerability to degradation, integrating four "environmental quality" dimensions regarded as the most significant factors associated with early processes of desertification, namely referring to climate, soil, vegetation, and land cover (Kairis et al. 2014). Such dimensions were assessed combining 11 variables (3 for climate quality: Aridity index, annual precipitation rate, aspect; 4 for soil: Parent material, soil depth, soil texture, slope; 4 for vegetation: Plant cover, erosion risk, fire risk, crop intensification) in a composite index, the so called ESAI; see Ferrara et al. (2020) for further details.

This specification was widely used at both local, regional, and national scales in southern Europe and, more recently, in many other countries all over the world (see Ferrara et al. 2020 for a review). Referring uniquely to the ecological dimension of land degradation, the index was extensively tested for statistical stability, reporting reliability, and coherence with field measures (Lanfredi et al. 2022) and is adopted here to avoid multi-collinearity with socioeconomic predictors when applying econometric techniques (Tomao et al. 2017). The composite index assigns a sensitivity score to each analysis domain (Coluzzi et al. 2022), ranging from 1 (the lowest land vulnerability to degradation) to 2 (the highest vulnerability to degradation). The average value of the ESAI has been calculated for the three time points (1960, 1990, 2020) at the provincial level (Salvati et al. 2011). Figure 1 shows the province-level distribution of the ESAI for each epoch, and the ESAI growth rate over the interval 1960–2020 (bottom right). Over 60 years, the average vulnerability to land degradation increased from 1.34 to 1.36, rising only 1.5%, albeit pockets of high risk remain. The most sensitive regions are found in southern Italy (Sicily and Apulia). In both regions, a high but stable proportion of critical land was observed during the study period, with an average southern Italy ESAI of 1.39 compared to that of

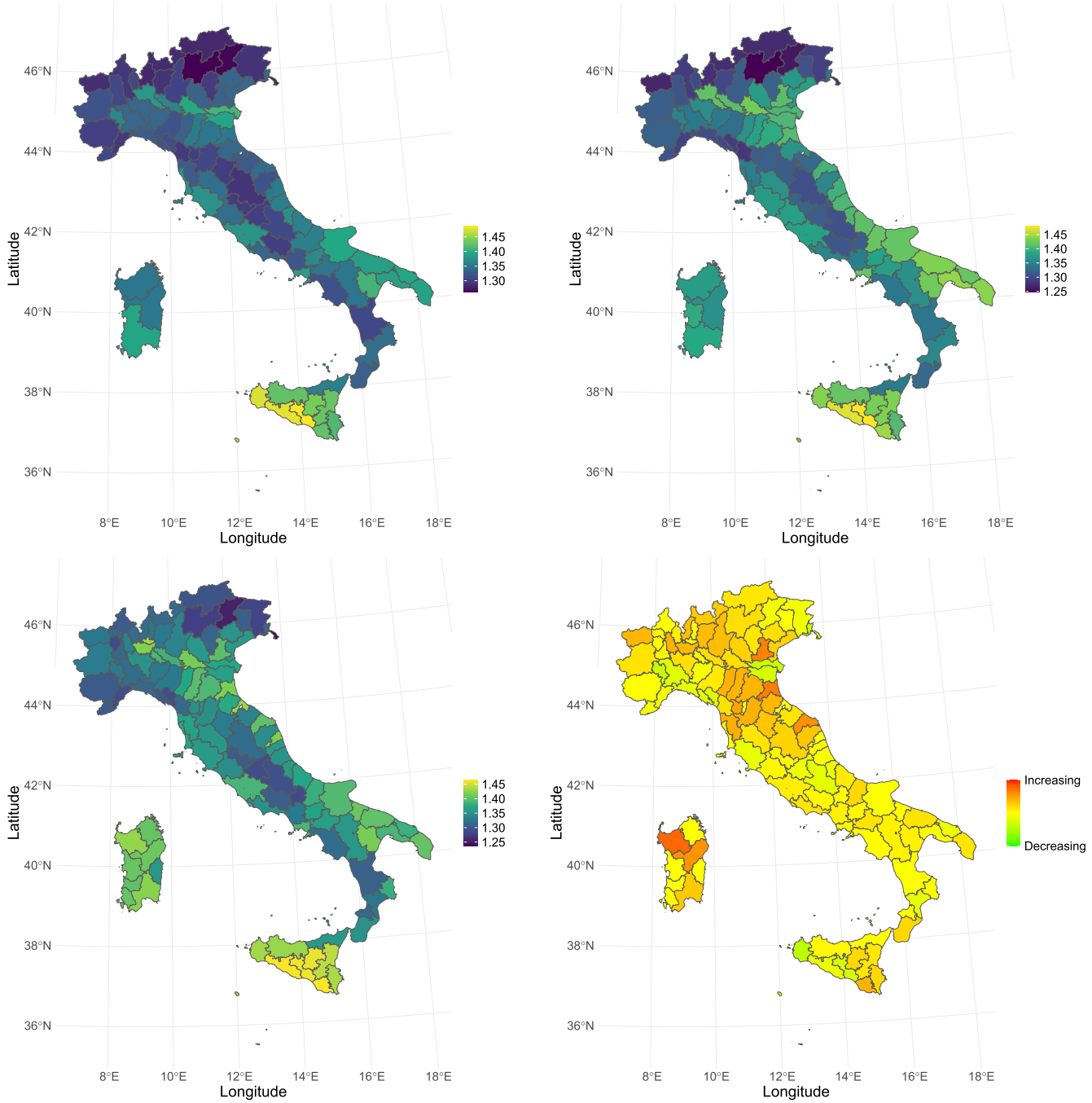

FIGURE 1 | From left to right and from top to bottom, province-level distribution of the ESAI scores observed all over Italy in 1960, 1990, and 2020. The graph on the bottom-right represents the ESAI growth rate over the interval 1960–2020.

northern Italy (1.33). On average, flat areas are the most sensitive land to desertification and showed the highest observed increase in the ESAI between 1960 and 2020. The ESAI varied in time and space, showing the highest growth rate in the Po valley, along the Adriatic coasts, in Umbria, flat areas of Tuscany, and in northern Sardinia (~ 4%), while being stable or even decreasing (−2.8%) in south-western parts of Sicily. The level of sensitivity to land degradation was found to be low and stable in mountain areas, while weakly increasing in uplands (1.2%). Overall, the index showed a very modest increase of 0.98% between 1960 and 2020, on average.

The spatial scale of the ESAI is aligned with the geographical detail of the collected variables tested as factors influencing land degradation (see below).

2.3 | Socioeconomic Drivers of Land Degradation

The dataset considered in our analysis is based on Italian provinces (NUTS-3), which combine sufficient geographical detail, avoiding a purely local approach, while evidencing the

TABLE 1 | Description of the variables in the sample.

Variable	Description
elev_median_km	Median elevation at the sea level (km)
cv_alt	Coefficient of variation of the altimetry
sup_km2	Surface area (km^2)
dipend_strut	Structural dependency ratio
pop_cap_percent	% of population in the provincial capital
dens_pro	Population density (inhabitants/km^2)
diff_dens_rate	Difference in population density in the provincial capital and that of the rest
dens_add_tur_pop_res	Density of population employed in the tourism sector
add_ulind	Share of workers in local industrial units (%)
add_ulser	Share of workers in local service units (%)
prod_lav_ser_media_percent	Avg. marginal service-sector labor productivity (%)
prod_serv_ind	Service versus industry productivity ratio
prod_ter	Indicator of land productivity
percent_va_agr	Agriculture's share of value added (%)
percent_va_ind	Industry's share of value added (%)
percent_va_ser	Services' share of value added (%)
contr_val_agg_percent	Province contribution to the national total value added (%)
vaproc_percent_media	Avg. per capita provincial value added (%)
sat_percent	Share of agricultural area in the surface area (%)
sat_aziende_agr	Agricultural area over total number of agricultural enterprises
sat_popres	Agricultural area over resident population
indice_di_vecchiaia_p_65_p_14_100	% of people 65+ over people aged ≤ 14
dens_post_letto_es_ric_sup_ha_100	% of beds in lodging facilities over surface area (hectare)
n_post_letto_es_alb	Number of hotel beds over number of hotel structures
add_cred_totale_pop	Density of population employed in the quaternary sector

spatial relationship between local districts, which may have a role in improving or slowing down the effectiveness of land degradation mitigation (and/or adaptation) strategies (Salvati et al. 2016). We focus on three years — 1960, 1990, and 2020 — from the past to the present, thereby capturing six decades of temporal evolution in land degradation and desertification dynamics. The provinces of Italy comprised 92 (1960), 95 (1990), and 110 (2020) administrative boundaries in each year of the study period. The final sample size for analysis is 297 observations. Drawing on the prevailing literature on desertification drivers, our model integrates 25 indicators spanning socioeconomic, demographic, and land-use dimensions. These indicators were selected considering the most recent evidence from land degradation studies in southern Europe, trying to keep ex-ante the intrinsic redundancy characteristic of any econometric specification under control.

The description of all the variables is summarized in Table 1, while summary statistics of continuous and discrete variables are provided in Table 2 for 1960, 1990, and 2020, as well as for the entire dataset. Topographic variables, such as median elevation and variation of altitude, are obviously time-invariant, and any variation stems solely from the redrawing of provincial boundaries. On the contrary, demographic and socioeconomic variables capture six decades of profound structural change. Population density and the aging index of the population surge drastically, posing pressure on both land and welfare systems. A similar pattern can be observed for the share of value-added services, which has become the dominant sector. The supply of tourist accommodation more than triples both in absolute and density terms, reflecting Italy's shift toward a service- and tourism-oriented economy.

In light of these comments, we applied additive explicit expectile regressions to evaluate each predictor's marginal impact on ESAI scores across Italian provinces, by year, incorporating both spatial and temporal dependencies. The study was carried out in a purely explorative framework, based on previous econometric specifications of the socioeconomic drivers of land degradation and the empirical results of earlier studies performing more traditional spatial approaches.

3 | Methodology

In this section, we briefly review expectile regression and formally introduce the considered additive spatio-temporal expectile model applied to the data illustrated above. Then, the estimation and stability selection procedures are presented.

TABLE 2 | Summary statistics of the variables over all Italy in 1960, 1990, 2000, and in the sample. For binary variables, we report the absolute and percentage frequencies.

	Epoch			
Variable	**1960**	**1990**	**2020**	**Full data**
elev_median_km	0.386 (0.392)	0.386 (0.388)	0.379 (0.376)	0.383 (0.384)
cv_alt	0.823 (0.328)	0.823 (0.327)	0.803 (0.324)	0.816 (0.325)
sup_km2	3270 (1850)	3170 (1710)	2740 (1590)	3040 (1720)
dipend_strut	0.500 (0.0726)	0.473 (0.039)	0.544 (0.041)	0.508 (0.0604)
pop_cap_percent	0.257 (0.161)	0.263 (0.140)	0.269 (0.141)	0.263 (0.147)
dens_pro	215 (272)	233 (317)	259 (366)	237 (323)
diff_dens_rate	0.311 (0.249)	0.280 (0.221)	0.307 (0.221)	0.300 (0.230)
dens_add_tur_pop_res	0.009 (0.005)	0.004 (0.003)	0.006 (0.002)	0.006 (0.004)
add_ulind	6.85 (3.37)	5.29 (1.21)	5.21 (1.19)	5.74 (2.24)
add_ulser	2.46 (0.548)	2.84 (0.511)	2.81 (0.392)	2.71 (0.510)
prod_lav_ser_media_percent	1.02 (0.121)	1.03 (0.121)	1.03 (0.170)	1.03 (0.141)
prod_serv_ind	0.549 (0.123)	0.384 (0.0564)	0.539 (0.107)	0.493 (0.125)
prod_ter	0.155 (0.116)	2.80 (2.48)	3.81 (4.20)	2.36 (3.29)
percent_va_agr	0.190 (0.097)	0.0564 (0.0366)	0.0268 (0.018)	0.0867 (0.091)
percent_va_ind	0.339 (0.100)	0.301 (0.0834)	0.251 (0.073)	0.294 (0.093)
percent_va_ser	0.472 (0.070)	0.642 (0.0753)	0.722 (0.068)	0.619 (0.126)
contr_val_agg_percent	0.011 (0.003)	0.0108 (0.003)	0.0108 (0.003)	0.011 (0.003)
vaproc_percent_media	0.989 (0.294)	0.973 (0.231)	0.923 (0.224)	0.960 (0.251)
sat_percent	0.876 (0.069)	0.728 (0.123)	0.539 (0.165)	0.704 (0.189)
sat_aziende_agr	6.91 (3.92)	9.13 (5.32)	15.3 (11.7)	10.7 (8.79)
sat_popres	0.692 (0.464)	0.594 (0.449)	0.440 (0.388)	0.567 (0.444)
indice_di_vecchiaia_p_65_p_14_100	52.9 (21.7)	119 (47.7)	159 (34.1)	113 (56.8)
dens_post_letto_es_ric_sup_ha_100	3.62 (5.43)	13.1 (15.8)	18.7 (26.2)	12.2 (19.5)
n_post_letto_es_alb	23.0 (8.66)	51.1 (20.1)	71.1 (30.5)	49.8 (29.8)
add_cred_totale_pop	0.002 (0.001)	0.007 (0.003)	0.006 (0.004)	0.005 (0.004)

3.1 | Background on Expectile Regression

For a continuous response variable, expectile regression provides a much more flexible approach and a complete picture of the conditional distribution of the response than classical regression models targeting conditional averages. This method, proposed by Newey and Powell (1987), can be thought of as a generalization of mean regression based on asymmetric least squares estimation. More generally, expectiles can be embedded in a common framework within the wider class of generalized quantiles defined as the minimizers of an asymmetric power loss function of order 2 (Chen 1996). Formally, let $\tau \in (0,1)$ and consider the following asymmetric loss function

$$\omega_\tau(u) = u^2 \cdot |\tau - \mathbf{1}(u < 0)|, \quad u \in \mathbb{R} \tag{1}$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. Given a set of covariates $\mathbf{X} = \mathbf{x} \in \mathbb{R}^p$, the τ-th conditional expectile of Y, $\mu_\tau(\mathbf{x})$, can be expressed as the solution of the following minimization problem:

$$\mu_\tau(\mathbf{x}) = \arg\min_{m \in \mathbb{R}} \mathbb{E}[\omega_\tau(Y - m(\mathbf{x}))]. \tag{2}$$

where $m(\cdot)$ is a predictor depending on the covariates $\mathbf{x}$. Specifically, setting $\tau = \frac{1}{2}$ makes $\mu_{1/2}(\boldsymbol{x})$ equal to the conventional conditional mean of Y given $\mathbf{X} = \mathbf{x}$; choosing any $\tau \neq \frac{1}{2}$ instead allows us to target any part of the conditional distribution of the response. In practice, standard quantiles have a more intuitive interpretation than expectiles, even if they target essentially the same part of the distribution of interest. However, despite the popularity and easy interpretability of the former, the latter offer some advantages: (a) we gain uniqueness of the Maximum Likelihood solutions which is not guaranteed in the quantile context; (b) from a computational standpoint, since the squared loss function $\omega_\tau(\cdot)$ is differentiable, $\mu_\tau(\mathbf{x})$ can be estimated by efficient iterative reweighted least squares, in contrast to algorithms used for fitting quantile regression models.

In the next section, following the work of Spiegel et al. (2020), we apply expectile regression to spatio-temporal data with an additive regression modeling structure.

3.2 | Additive Spatio-Temporal Expectile Regression Model

Let $Y_t(\mathbf{s})$ be an absolutely continuous response variable and $\mathbf{X}_t(\mathbf{s})$ be a p-dimensional vector of explanatory variables observed at time $t = 1, \ldots, T$ and location $\mathbf{s} \in S \subseteq \mathbb{R}^2$. In our application, $Y_t(\mathbf{s})$ represents the ESAI in year t and location $\mathbf{s} = (s_1, s_2)$ represents the spatial coordinates of the centroid of an Italian province. The generic realization of $Y_t(\mathbf{s})$ and $\mathbf{X}_t(\mathbf{s})$ at time t and $\mathbf{s}$ is given by $y_t(\mathbf{s})$ and $\mathbf{x}_t(\mathbf{s})$, respectively.

We assume that the τ-th conditional expectile of $Y_t(\mathbf{s})$ given $\mathbf{x}_t(\mathbf{s})$ at level $\tau \in (0, 1)$ can be described using the following spatio-temporal additive regression model:

$$\mu_\tau(\mathbf{x}_t(\mathbf{s})) = \beta_{\tau 0} + \sum_{j=1}^{p} f_{\tau k}(x_{tj}) + f_\tau^{(spat)}(\mathbf{s}), \tag{3}$$

where $\beta_{\tau 0}$ is an intercept term, $f_{\tau 1}(x_{t1}), \ldots, f_{\tau p}(x_{tp})$ are expectile-specific unknown functions of the covariates and $f_\tau^{(spat)}(\mathbf{s})$ is a bivariate smooth function of coordinates information capturing spatial dependence in the model. To capture potential nonlinear effects, the former include linear terms as well as smooth, nonlinear functions based on penalized B-splines (P-splines, Eilers and Marx 1996). More specifically, by a P-spline decomposition (see Kneib et al. 2009; Hofner et al. 2011), the effect of each predictor $f_{\tau j}(x_{tj})$ is decomposed into a linear polynomial (without intercept) and a smooth deviation from this linear term. For each predictor $j = 1, \ldots, p$, we thus get:

$$f_{\tau j}(x_{tj}) = \beta_{\tau j} x_{tj} + f_{\tau j}^{(centered)}(x_{tj}). \tag{4}$$

The nonlinear functions $f_{\tau j}^{(centered)}(\cdot)$ are approximated using p cubic B-spline basis functions $\mathbf{B}_1, \ldots, \mathbf{B}_p$ so that $f_{\tau j}^{(centered)}(\cdot) = \mathbf{B}_j \boldsymbol{\gamma}_{\tau j}$, where $\boldsymbol{\gamma}_{\tau j}$ is the τ-dependent vector of basis coefficients corresponding to the jth predictor. Similarly, to model the spatial effects we express $f_\tau^{(spat)}(\cdot)$ in terms of the tensor product of cubic marginal B-spline basis functions for smooth surfaces $\mathbf{Z}$, so that $f_\tau^{(spat)}(\cdot) = \mathbf{Z}\boldsymbol{\theta}_\tau$, with $\boldsymbol{\theta}_\tau$ being a vector of unknown coefficients. All effects can be integrated within a unifying framework, where the predictor $\mu_\tau(\mathbf{x}_t(\mathbf{s}))$ in Equation (3) can be represented as the product of a design matrix $\mathbf{K}$ obtained by stacking all basis functions and the complete vector of regression coefficients $\boldsymbol{\Phi}_\tau = (\beta_{0\tau}, \ldots, \beta_{p\tau}, \boldsymbol{\gamma}_{\tau 1}, \ldots, \boldsymbol{\gamma}_{\tau p}, \boldsymbol{\theta}_\tau)'$. The complete predictor in Equation (3) can then be written as $\mu_\tau(\cdot) = \mathbf{K}\boldsymbol{\Phi}_\tau$. To prevent overfitting, we control the trade-off between smoothness and model fit by adding a penalty term associated with each vector of spline coefficients. The penalty is expressed as $\boldsymbol{\Phi}_\tau' \mathbf{P} \boldsymbol{\Phi}_\tau$ where $\mathbf{P} = \mathbf{D}'\mathbf{D}$ and $\mathbf{D}$ is a matrix of second-order differences, with a second-order difference penalty with regularization parameter vector, $\boldsymbol{\lambda}$.

The model parameters in $\boldsymbol{\Phi}_\tau$ are estimated using component-wise gradient descent boosting (Bühlmann and Hothorn 2007) as described in the following section.

3.3 | Estimation

Boosting is an ensemble method used in machine learning that involves fitting numerous penalized simpler models (also known as base learners) sequentially, and then pooling them together to improve model accuracy. Throughout this article, we refer to a base learner as a regression estimator that uses as input variables the subset of columns of the design matrix $\mathbf{K}$ associated with a given predictor x_{tj}, $j = 1, \ldots, p$. For instance, in the case of a one-dimensional B-spline, a base learner is the design matrix representing the basis functions of the spline with a linear link function. Therefore, the aim is to estimate the functions $f_{\tau 1}, \ldots, f_{\tau p}, f_\tau^{(spat)}$ specified by the base learners optimizing a differentiable and convex loss function. Here, we consider the minimization of the following loss function:

$$\mathcal{R} = \sum_{t=1}^{T} \sum_{\mathbf{s} \in S} \omega_\tau \left(y_t(\mathbf{s}) - \mu_\tau(\mathbf{x}_t^\star(\mathbf{s})) \right), \tag{5}$$

where $\omega_\tau(\cdot)$ is the asymmetric squared loss in Equation (1) and $\mathbf{x}_t^\star(\mathbf{s})$ refers to the complete vector of covariates, including continuous, categorical covariates, as well as spatial information. The estimation procedure is summarized in Algorithm 1.

The key aspect of component-wise boosting is that during each iteration, the algorithm evaluates all candidate base learners (one per covariate), and updates the one that maximally reduces the variance of the gradient residuals, rather than updating the model as a whole. This reduces computational costs and, at the same time, selects only relevant features for predictions, inherently promoting sparsity in the model. The size of the update in each step is controlled by the step size ν. A smaller ν slows convergence, requiring more iterations, but avoids too large steps along the gradient. The algorithm terminates when either a maximum number of iterations is reached or when the performance of

ALGORITHM 1 | Component-wise boosting algorithm for spatio-temporal additive expectile regression.

Step 1. Set the iteration index $m = 0$. Initialize the value of β_0 using the empirical τ-th expectile of the response and set all base learners $\hat{f}_{\tau j}^{[0]}(\cdot) = 0$, for $j = 1, \ldots, r$.

Step 2. Set $m = m + 1$ and calculate the negative gradient residuals of the empirical risk $\mathcal{R}$ in Equation (5). This yields the negative gradient vector $\mathbf{u}^{[m]}$ with generic element

$$u_t^{[m]}(\mathbf{s}) = \frac{\partial}{\partial \mu} \omega_\tau \left(y_t(\mathbf{s}) - \mu_\tau^{[m-1]}(\mathbf{x}_t^\star(\mathbf{s})) \right), \tag{6}$$

for a given time t and location $\mathbf{s}$.

Step 3. Choose the predictor $j^\star$ that minimizes the squared loss:

$$j^\star = \arg\min_j \sum_{t=1}^{T} \sum_{\mathbf{s} \in S} (u_t^{[m]}(\mathbf{s}) - \hat{h}_j(\mathbf{x}_t^\star(\mathbf{s}))^2, \tag{7}$$

where $\hat{h}_j = \mathbf{H}_j \mathbf{u}$, with $\mathbf{H}_j = \mathbf{K}_j(\mathbf{K}_j'\mathbf{K}_j + \lambda_j \mathbf{K}_j)^{-1}\mathbf{K}_j'$ being the hat matrix.

Step 4. Update the $j^\star$-th base learner to $\hat{f}_{\tau j}^{[m]} = \hat{f}_{\tau j}^{[m-1]} + \nu \hat{h}_j^\star$, where $\nu \in (0, 1]$ is the step size.

Keep iterating by repeating steps 2 and 4 for a predefined number of iterations m_{stop}.

the model stops improving significantly. The choice of the stopping iteration m_{stop} is a crucial issue in boosting and must be carefully chosen. A large number of iterations likely overfit the data, resulting in suboptimal prediction accuracy (Bühlmann and Hothorn 2007). Therefore, early stopping, that is, halting iterations before overfitting, is a common strategy to estimate the optimal m_{stop}. Based on our data, we set $\nu = 0.1$ and determine m_{stop} based on 10-fold cross-validation.

The impact of the penalty on the base learners can be determined via the degrees of freedom defined as the trace of the hat matrix, $df(\lambda_j) = df_j = \text{tr}(\mathbf{H}_j)$, which controls its smoothness. A base learner with greater degrees of freedom, that is, less penalization, offers greater flexibility than a base learner with smaller degrees of freedom, that is, more penalization, and therefore has a greater chance of being selected by the boosting algorithm. As suggested by Kneib et al. (2009) and Hofner et al. (2011), to ensure that the complexity of each base learner is comparable, we fix the degrees of freedom of all penalized base learners for each continuous predictor variable at the same value df. Different degrees of freedom are specified for the spatial base learner $f_\tau^{(spat)}(\cdot)$. To tune the smoothing parameter df, we employ a grid search over a candidate set of values for the degrees of freedom via a 10-fold cross-validation.

3.4 | Stability Selection

Given the high dimensionality of potential predictors and the complexity of the phenomenon under study, understanding the relationships between factors contributing to the evolution over time of land vulnerability scores in Italy may require including a large number of variables, which, however, can potentially muddy the interpretation of the results. To formally distinguish meaningful predictors of land degradation dynamics in Italy from noise variables, we integrate the boosting framework with the stability selection procedure proposed by Meinshausen and Bühlmann (2010). Stability selection works by leveraging on repeatedly subsampling a fraction of the original data and fitting the model on each subset. This method leads to consistent variable selection and provides an upper bound (which is fixed by the researcher) to control the expected number of noninformative predictors in the final model. In our analysis, for each subsample the model is fitted using the component-wise boosting algorithm in Section 3.3. After boosting estimation, it is then possible to compute the proportion of subsamples in which each base learner is selected, called selection probability, which reflects the predictor's stability across subsamples. The final model comprises only those base learners that exceed a fixed threshold π_{thr}. Higher thresholds reduce false positives but may miss weak signals. To include those edges that are present in a "sufficiently large number" of estimated models, we set $\pi_{\text{thr}} = 0.75$. Finally, to bound the number of falsely selected base learners that is tolerated, we specified a per-family error rate of 10%.

4 | Results and Discussion

The empirical findings of this study are presented in two separate sections. Section 4.1 illustrates the effect of included socioeconomic predictors on ESAI scores in provinces in normal and risk areas, while Section 4.2 discusses the spatial distribution of the ESAI across Italian provinces.

4.1 | Analysis of Socioeconomic Factors on ESAI Scores

We fitted the additive spatio-temporal expectile model considered at levels $\tau = (0.50, 0.90)$. This allows us to link the ESAI scores to the variables listed in Table 2 in areas of low/medium sensitivity and areas at high risk of desertification. Prior to fitting, continuous predictors were standardized to have zero mean and unit variance, and a three-level categorical covariate, *year*, indicating the three epochs—1960, 1990, 2020—was included. Due to multicollinearity, we also removed *percent_va_agr* and *vaproc_percent_media*. For continuous covariates, we used the decomposition in Equation (4) to separate the linear and nonlinear parts, which allowed for a data-driven decision on linearity versus nonlinearity of the effects. In practice, the model was implemented using the `mboost` package (Hofner et al. 2014) where linear and nonlinear base learners were specified using the functions `bols()` and `bbs()`, respectively. The province-level smooth spatial effect was implemented using the function `bspatial()`. We tuned the degrees of freedom and optimal stopping iteration using a 10-fold cross-validation. This procedure selected 3 degrees of freedom for all continuous covariates and 9 degrees of freedom for the spatial term. We also investigated a 5-fold cross-validation to select df and m_{stop}, and found comparable results.

We start by commenting on the results of the stability selection described in Section 3.4, repeating the estimation procedure over 500 subsamples from the original dataset, and retaining only those base learners (one for each predictor) that appeared in at least 75% of the estimated models. Figure 2 displays the top 20 base learners at expectile levels $\tau = 0.50$ (left) and $\tau = 0.90$ (right) with the highest selection probabilities, respectively. Unsurprisingly, the spatial component was selected by stability selection as the most informative in both models. This confirms the evident spatial heterogeneity of desertification processes in Italy's provinces, reflecting mixed and largely differentiated patterns across local territories. Likewise, median elevation (*elev_median_km*) enters each model as the combination of both linear and smooth nonlinear terms. With regard to differences, the relevance of some covariates is distinctly expectile-dependent, with several drivers selected only at one expectile, highlighting mechanisms that operate differently across the distribution. In particular, at $\tau = 0.90$, *pop_cap_percent* and *n_post_letto_es_alb* are uniquely selected, whereas at $\tau = 0.50$, *n_post_letto_es_alb*, *prod_serv_ind* and *dens_add_tur_pop_res* are included. Only the provincial land devoted to agriculture (*sat_percent*) appears in both models, emphasizing that factors driving desertification in highly vulnerable areas differ compared with regions of low or moderate sensitivity. It is also interesting to note that the variable *year* was rarely chosen by stability selection, confirming the negligible effect of temporal dependence on the ESAI scores observed in the exploratory analysis.

We thus refitted the proposed spatio-temporal expectile regression, incorporating only those base learners selected by stability

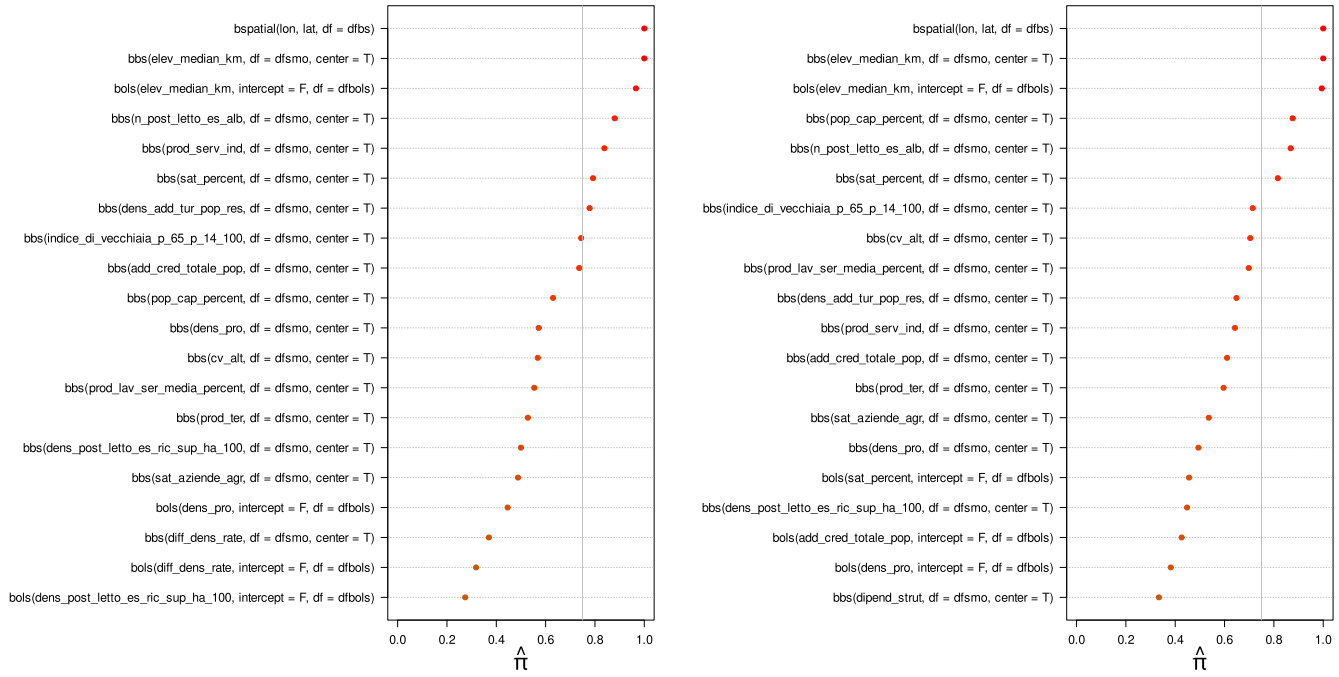

FIGURE 2 | Selection probabilities for each base learner at expectile levels $\tau = 0.50$ (left) and $\tau = 0.90$ (right). The first 20 base learners with the highest selection probabilities are plotted. The vertical gray line denotes the selected threshold $\pi_{\text{thr}} = 0.75$.

TABLE 3 | Variable importance and selection frequencies from boosting estimation of the fitted models at $\tau = 0.50$ (left) and $\tau = 0.90$ (right). Variable importance is measured as the % reduction in the objective function $\mathcal{R}$.

	$\tau = 0.50$				$\tau = 0.90$		
Variable	**Type**	**Selection freq.**	**Reduction (%)**	**Variable**	**Type**	**Selection freq.**	**Reduction (%)**
spatial	smooth	0.86	72.09	spatial	smooth	0.79	81.88
elev_median_km	smooth	0.02	16.16	elev_median_km	smooth	0.04	10.97
elev_median_km	linear	<0.01	6.23	elev_median_km	linear	0.01	2.12
n_post_letto_es_alb	smooth	0.03	2.50	n_post_letto_es_alb	smooth	0.02	1.78
sat_percent	smooth	0.04	1.55	dens_add_tur_pop_res	smooth	0.07	1.75
pop_cap_percent	smooth	0.05	1.46	prod_serv_ind	smooth	0.03	1.08
—	—	—	—	sat_percent	smooth	0.04	0.42

selection. Table 3 reports the most significant variables in the fitted models at $\tau = 0.50$ (left) and $\tau = 0.90$ (right), ranked by selection frequency and variable importance scores, measured as the % reduction in the objective function $\mathcal{R}$ in Equation (5). Echoing Figure 2, the spatial component dominates model fit, accounting for 70% of $\mathcal{R}$ reduction at the mean and 82% in the upper tail. Among covariates, median elevation (*elev_median_km*) emerges as the single most influential driver, which contributes to 22.39% at $\tau = 0.50$ and 13.09% at $\tau = 0.90$, respectively. Socioeconomic factors, though still meaningful, explain roughly half as much as elevation, indicating that while human pressures amplify desertification risk, topography features remain the primary determinants, especially under extreme conditions.

We continue the analysis by commenting on the estimated effects of covariates retained in the final model. Since our main focus is to analyze provinces in the upper tail of the ESAI distribution, which are the most exposed to severe desertification, in the following, we assess the covariate effects at the 90th expectile model. Uncertainty in expectile regression parameter estimates was quantified with a nonparametric bootstrap. Specifically, the model was refitted on 500 resamples, and pointwise $(1-\alpha)\%$ confidence intervals were calculated as $[\hat{q}_{j\alpha/2}, \hat{q}_{j1-\alpha/2}]$, where $\hat{q}_{j\alpha/2}$ denotes the estimated $\alpha/2 \times 100\%$ quantile of the j-th estimate with $j = 1, \ldots, p$. Figure 3 overlays the fitted curves (red) with their 95% bootstrap confidence intervals (gray bands).

With the exception of *sat_percent*, all continuous variables show a statistically significant effect. Median province elevation shows the largest absolute effect size, exhibiting a non-monotonic relationship. The linear and nonlinear components are both positive and decreasing at low elevations, but the direction of the latter is reversed at higher altitudes and displays a pronounced U-shape trajectory. This suggests that low-altitude provinces

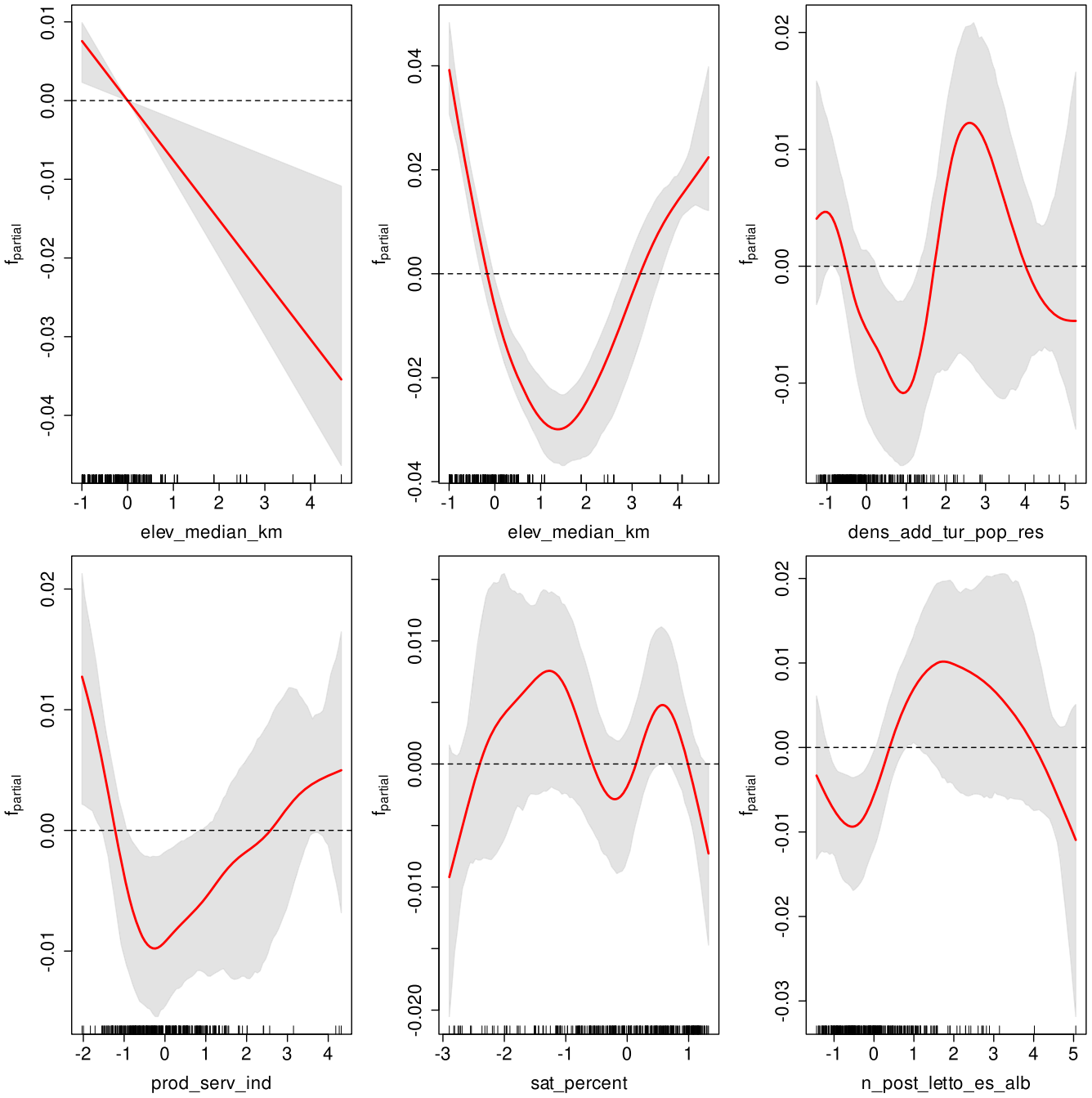

FIGURE 3 | Estimated linear and nonlinear effects of continuous covariates at expectile level $\tau = 0.90$ (red lines) with 95% bootstrap confidence intervals in parentheses obtained over 500 resamples (gray shaded areas).

face the greatest desertification pressure, the risk then declines at mid-elevations before increasing sharply at high elevations. The estimated effect of *dens_add_tur_pop_res* and *prod_serv_ind* oscillates between a positive effect at low values, then declines sharply and becomes negative. Modest values dampen desertification, but high intensities reverse the effect, reflecting the shift from moderate to heavy land and natural resources exploitation. A similar threshold pattern appears for *n_post_letto_es_alb*: Limited tourist capacity appears to bolster resilience, whereas extensive infrastructure sharply elevates risk. Overall, these findings suggest that these factors are mildly protective against desertification at low and moderate values; however, intensive industrialization and tourism development become dominant stressors in provinces facing the gravest desertification threats.

The statistical results of expectile regressions basically confirm the empirical findings of previous studies for Italy and southern Europe at large (e.g., Salvati et al. 2011; Colantoni et al. 2015; Lanfredi et al. 2022). Land degradation in the northern Mediterranean region, and especially in Italy, as one of the reference countries in Annex 4 of the global UNCCD convention to combat desertification, is driven by a combination of socioeconomic pressures, including intensive agricultural practices, urban sprawl, and unsustainable land use (Kairis et al. 2014; Kosmas et al. 2015;

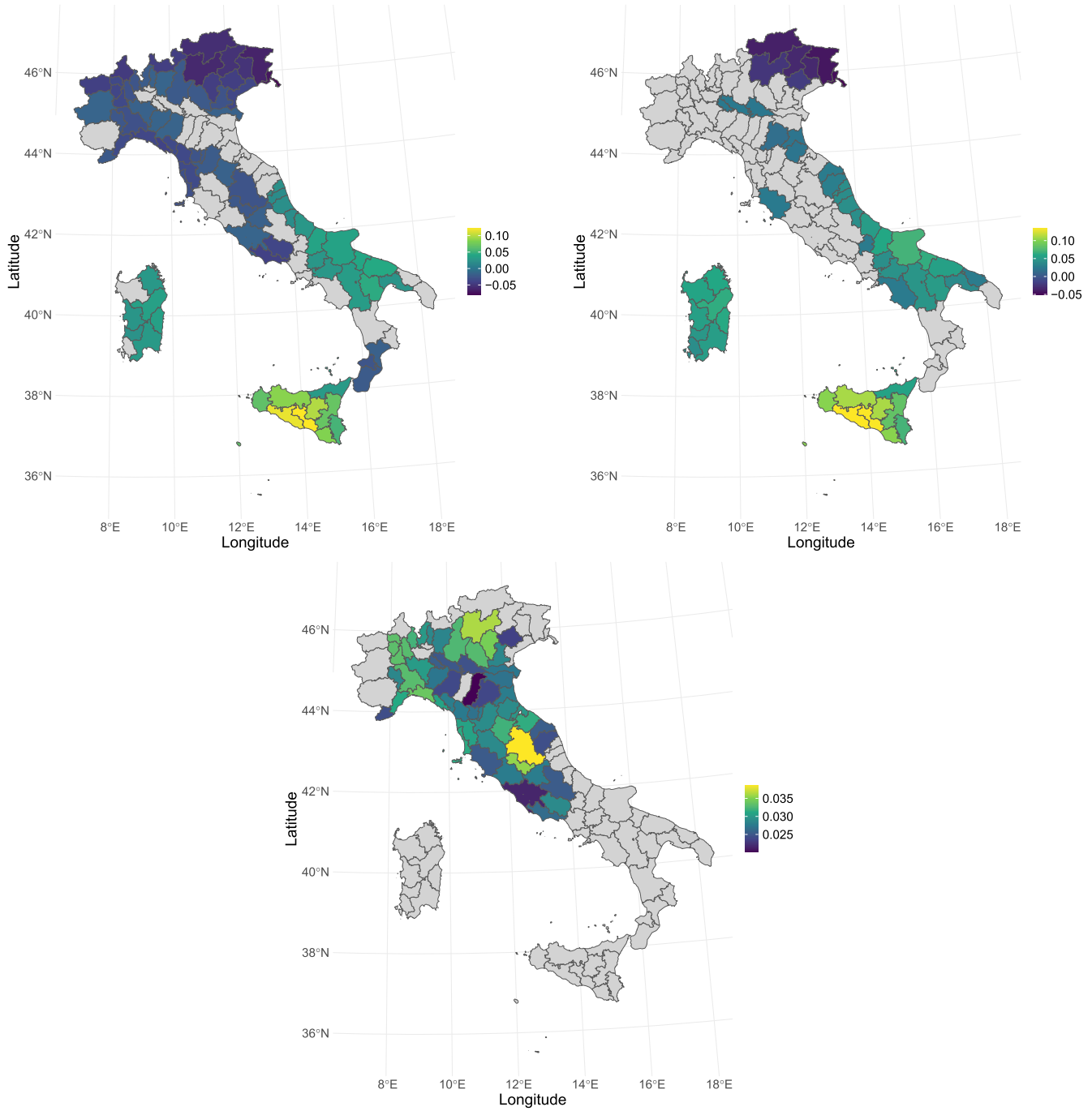

FIGURE 4 | Estimated spatial effects at $\tau = 0.50$ (left) $\tau = 0.90$ (right) for Italian provinces in 2020. The lower panel represents the 90–50 expectile range calculated as the difference between the estimated spatial effects at the 90th and 50th expectile.

Smiraglia et al. 2019). Indirect impacts stemming from the level of economic activity have also been evident from expectile models' outcomes (Hill et al. 2008; Vu et al. 2014; Salvati et al. 2015). Rapid urbanization and tourism development along coastal zones exacerbate land degradation through habitat destruction and increased pressure on water resources (Salvati et al. 2011; Cowie et al. 2018; Coluzzi et al. 2022). Among the latent drivers of land degradation, rural depopulation and economic decline in many inland areas have led to land abandonment, reducing land management and increasing vulnerability to wildfires (Tomao et al. 2017). The dependence on monoculture farming and overgrazing, often incentivized by agricultural subsidies, was assumed as a further contribution to soil erosion, biodiversity loss, and ultimately, land degradation (Quaranta et al. 2020), and requires a more intense monitoring effort. These interconnected drivers highlight the complex linkage between economic activity, population dynamics, and environmental sustainability Dasgupta et al. (2006), and justify the use of refined, nonlinear approaches of statistical estimation focusing on the spatial prediction of extreme values (e.g., identified in this study as the 90th expectile of the statistical distribution of vulnerability levels reflected in the ESAI scores). This perspective aligns with

a broad statistical literature advocating tail-specific methods for assessing risk and vulnerability under heterogeneous conditions (Reich 2012). In this respect, expectile-based approaches effectively summarize upper-tail behavior by capturing the severity of unfavorable outcomes, a feature that is particularly valuable in environmental systems characterized by strong spatial and structural heterogeneity (Tyralis et al. 2023). For stakeholders and authorities, our framework is intended to support evidence-based decision-making by providing an interpretable and comparable measure of territorial vulnerability across different areas.

A direct comparison of the models' outcomes considering together average ESAI and 90th expectile ESAI scores as the dependent variable clearly encourage further studies evaluating the spatial distribution of extreme—and not average—vulnerability levels across the selected analysis' spatial domain (Francaviglia et al. 2019), to clarify the intrinsic spatial patterns and trends characteristic of desertification risk in Mediterranean countries, which is demonstrated to be strongly associated with extremely high ESAI scores.

4.2 | Analysis of Spatial Effects

We conclude the analysis by evaluating the spatial heterogeneity of desertification processes across Italian provinces (e.g., Lamonica et al. 2020). Given the negligible temporal effect on ESAI scores, we here focus on 2020 only. Figure 4 maps the estimated spatial effects for Italian provinces at the expectile level $\tau = 0.50$ (left) and $\tau = 0.90$ (right). Provinces with non-significant effects based on 95% bootstrap confidence intervals are shaded in gray.

As expected, positive spatial effects—indicating elevated desertification risk—are concentrated in southern regions, notably Sicily, Apulia, and Sardinia. In contrast, several provinces in central (Lazio, Umbria, Liguria) and northern Italy exhibit significant negative effects. Furthermore, on average, darker colors appear along the northern Apennines and the northern end of the country, whereas they are more concentrated in Northeast Italy at the 90th expectile. The lower panel of Figure 4 portrays the 90–50 expectile range, defined as the difference between the 90th and 50th expectile of the estimated spatial effects. This can be thought of as a measure of tail variability that provides critical insights into the spatial heterogeneity of desertification across the Italian territory. The plot reveals that central and northern provinces are associated with higher spreads than those of the south. Markedly greater spatial dispersion can be observed across lowland and inland uplands, such as Perugia, Alessandria, and Trento. By contrast, mountain and Alpine provinces tend to exhibit non-significant spatial variability.

The spatial structure identified by the estimate of individual territorial effects allows us to outline a geography of complexity in the land degradation vulnerability index. This spatial structure tends to become progressively more complex over time, from 1960 to 2020, and defines a greater geographical polarization between vulnerable areas and non-exposed areas, in line with the empirical results of other works. At the same time, the empirical results denote an increase in heterogeneity at the local scale (e.g., Kosmas et al. 2015; Jakubínskỳ et al. 2021; Coluzzi et al. 2022), which probably reflects the spatial complexity of the mix of triggering factors whose importance, in compositional terms, varies greatly over time (Salvati et al. 2012).

5 | Conclusion

Climate change, soil degradation, and desertification increasingly threaten Mediterranean landscapes through complex interactions of natural and human forces. The goal of this work is to quantify the impact of an extensive set of land-use and socioeconomic factors on a policy-relevant desertification index for Italy's provinces over a long time window, with particular emphasis on regions experiencing the most severe degradation. To this end, we develop an apposite semiparametric additive spatio-temporal expectile regression model leveraging P-splines to flexibly capture nonlinear trends and spatially heterogeneous effects. Model fitting relies on an iterative, component-wise gradient boosting algorithm, which we integrate with stability selection to perform variable selection and promote sparsity. Our framework not only accommodates heterogeneity across space and time but also highlights how drivers of desertification differ across the distribution of the ESAI, offering actionable insights for targeted land-management strategies. From a policy-relevant perspective, it can assist authorities, planners, and environmental agencies in prioritizing interventions, allocating resources, and monitoring high-risk areas, thereby strengthening strategic planning and actions aimed at mitigating severe land degradation. The "applied" perspective of the proposed framework, emphasizing its function as a monitoring and assessment tool designed for analysts and institutions engaged in environmental evaluation, is particularly indicated as a basic knowledge support in any spatial planning and local development policy decision. Being spatially explicit, the empirical results of this study can be immediately incorporated in any national and regional strategy aimed at fighting the early desertification process, in line with the general statements of the UNCCD and the specific guidelines of the Italian National Plan to Combat Desertification. Policy prescriptions that are intrinsically tuned with territorial heterogeneity and the fine-grained characteristics of local districts are particularly appropriate not only in the fight of early desertification process, but also in the containment of any phenomenon of environmental degradation demanding place-specific approaches.

With this general perspective in mind, the study advances desertification analysis across the Italian provinces, overcoming critical shortcomings in the extant literature. With respect to previous studies, this work offers several innovative contributions. First, rather than modeling measures of central tendency, we directly model the entire conditional distribution of the desertification index, with particular emphasis on provinces exposed to "high risk" circumstances. Second, we employ a flexible nonparametric modeling framework for regression expectiles capable of capturing complex, nonlinear interdependencies among covariates, and accommodating both spatial heterogeneity and temporal dynamics. Third, this approach is integrated with stability selection to identify the most important factors that are truly affecting desertification, while reducing the risk of selecting spurious features due to noise or overfitting.

From an applied standpoint, the results of the expectile model are in line with the empirical evidence already outlined in previous field and modeling studies. The innovative contribution of this work lies in a more precise analysis of the importance of the different drivers considered at a highly detailed spatial scale, thus allowing an estimate of both the factors that act at a typically local scale, and of dynamics that are articulated on larger, regional or homogeneous district territories. The intrinsic relationship between local dynamics and regional spatial trends over time is clearly explained by the expectile model, which also provides an explicit prediction of the impact of the different socioeconomic factors of landscape transformation (e.g., estimated jointly, over a sufficiently long time span, considering climatic, soil, and vegetation quality).

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The dataset analyzed in this study is available from the authors upon request.

References

Akhtar-Schuster, M., L. C. Stringer, A. Erlewein, et al. 2017. "Unpacking the Concept of Land Degradation Neutrality and Addressing Its Operation Through the Rio Conventions." *Journal of Environmental Management* 195: 4–15.

Benassi, F., S. Cividino, P. Cudlin, A. Alhuseen, G. R. Lamonica, and L. Salvati. 2020. "Population Trends and Desertification Risk in a Mediterranean Region, 1861-2017." *Land Use Policy* 95: 104626.

Biasi, R., E. Brunori, C. Ferrara, and L. Salvati. 2019. "Assessing Impacts of Climate Change on Phenology and Quality Traits of *Vitis Vinifera* L: The Contribution of Local Knowledge." *Plants* 8, no. 5: 121.

Bühlmann, P., and T. Hothorn. 2007. "Boosting Algorithms: Regularization, Prediction and Model Fitting." *Statistical Science* 22, no. 4: 477–505.

Bühlmann, P., and B. Yu. 2003. "Boosting With the L 2 Loss: Regression and Classification." *Journal of the American Statistical Association* 98, no. 462: 324–339.

Canfora, L., L. Salvati, A. Benedetti, and R. Francaviglia. 2017. "Is Soil Microbial Diversity Affected by Soil and Groundwater Salinity? Evidences From a Coastal System in Central Italy." *Environmental Monitoring and Assessment* 189, no. 7: 319.

Cecchini, M., I. Zambon, A. Pontrandolfi, et al. 2019. "Urban Sprawl and the Olive Landscape: Sustainable Land Management for Crisis Cities." *GeoJournal* 84, no. 1: 237–255.

Chen, Z. 1996. "Conditional L_p-Quantiles and Their Application to the Testing of Symmetry in Non-Parametric Regression." *Statistics & Probability Letters* 29, no. 2: 107–115.

Colantoni, A., C. Ferrara, L. Perini, and L. Salvati. 2015. "Assessing Trends in Climate Aridity and Vulnerability to Soil Degradation in Italy." *Ecological Indicators* 48: 599–604.

Coluzzi, R., L. Bianchini, G. Egidi, et al. 2022. "Density Matters? Settlement Expansion and Land Degradation in Peri-Urban and Rural Districts of Italy." *Environmental Impact Assessment Review* 92: 106703.

Cowie, A. L., B. J. Orr, V. M. C. Sanchez, et al. 2018. "Land in Balance: The Scientific Conceptual Framework for Land Degradation Neutrality." *Environmental Science & Policy* 79: 25–35.

Dasgupta, S., K. Hamilton, K. D. Paudey, and D. Wheeler. 2006. "Environment During Growth: Accounting for Governance and Vulnerability." *World Development* 34, no. 9: 1597–1611.

Eilers, P. H., and B. D. Marx. 1996. "Flexible Smoothing With B-Splines and Penalties." *Statistical Science* 11, no. 2: 89–121.

Elnashar, A., H. Zeng, B. Wu, T. G. Gebremicael, and K. Marie. 2022. "Assessment of Environmentally Sensitive Areas to Desertification in the Blue Nile Basin Driven by the MEDALUS-GEE Framework." *Science of the Total Environment* 815: 152925.

Ferrara, A., C. Kelly, G. A. Wilson, et al. 2016. "Shaping the Role of 'Fast' and 'Slow' Drivers of Change in Forest-Shrubland Socio-Ecological Systems." *Journal of Environmental Management* 169: 155–166.

Ferrara, A., C. Kosmas, L. Salvati, A. Padula, G. Mancino, and A. Nol. 2020. "Updating the MEDALUS-ESA Framework for Worldwide Land Degradation and Desertification Assessment." *Land Degradation & Development* 31, no. 12: 1593–1607.

Foroni, B., L. Merlo, and L. Petrella. 2024. "Quantile and Expectile Copula-Based Hidden Markov Regression Models for the Analysis of the Cryptocurrency Market." *Statistical Modelling* 25: 1471082X241279513.

Francaviglia, R., C. Di Bene, R. Farina, L. Salvati, and J. L. Vicente-Vicente. 2019. "Assessing 4 Per 1000 Soil Organic Carbon Storage Rates Under Mediterranean Climate: A Comprehensive Data Analysis." *Mitigation and Adaptation Strategies for Global Change* 24: 795–818.

Gerlach, R., and C. W. S. Chen. 2015. "Bayesian Expected Shortfall Forecasting Incorporating the Intraday Range." *Journal of Financial Econometrics* 14, no. 1: 128–158.

Gupta, A. K., M. Negi, S. Nandy, et al. 2020. "Mapping Socio-Environmental Vulnerability to Climate Change in Different Altitude Zones in the Indian Himalayas." *Ecological Indicators* 109: 105787.

Hill, J., M. Stellmes, T. Udelhoven, A. Roder, and S. Sommer. 2008. "Mediterranean Desertification and Land Degradation: Mapping Related Land Use Change Syndromes Based on Satellite Observations." *Global and Planetary Change* 64, no. 3–4: 146–157.

Hofner, B., T. Hothorn, T. Kneib, and M. Schmid. 2011. "A Framework for Unbiased Model Selection Based on Boosting." *Journal of Computational and Graphical Statistics* 20, no. 4: 956–971.

Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid. 2014. "Model-Based Boosting in R: A Hands-On Tutorial Using the R Package Mboost." *Computational Statistics* 29: 3–35.

Imbrenda, V., G. Quaranta, R. Salvia, et al. 2021. "Land Degradation and Metropolitan Expansion in a Peri-Urban Environment." *Geomatics, Natural Hazards and Risk* 12, no. 1: 1797–1818.

Incerti, G., E. Feoli, L. Salvati, A. Brunetti, and A. Giovacchini. 2007. "Analysis of Bioclimatic Time Series and Their Neural Network-Based Classification to Characterise Drought Risk Patterns in South Italy." *International Journal of Biometeorology* 51: 253–263.

Jakubínskỳ, J., M. Prokopova, P. Raška, et al. 2021. "Managing Floodplains Using Nature-Based Solutions to Support Multiple Ecosystem Functions and Services." *Wiley Interdisciplinary Reviews: Water* 8, no. 5: e1545.

Jones, M. C. 1994. "Expectiles and M-Quantiles are Quantiles." *Statistics & Probability Letters* 20, no. 2: 149–153.

Kairis, O., A. Karamanos, D. Voloudakis, et al. 2022. "Identifying Degraded and Sensitive to Desertification Agricultural Soils in Thessaly, Greece, Under Simulated Future Climate Scenarios." *Land* 11, no. 3: 395.

Kairis, O., C. Kosmas, C. Karavitis, et al. 2014. "Evaluation and Selection of Indicators for Land Degradation and Desertification Monitoring: Types of Degradation, Causes, and Implications for Management." *Environmental Management* 54: 971–982.

Kim, M., and S. Lee. 2016. "Nonlinear Expectile Regression With Application to Value-At-Risk and Expected Shortfall Estimation." *Computational Statistics & Data Analysis* 94: 1–19.

Kneib, T., T. Hothorn, and G. Tutz. 2009. "Variable Selection and Model Choice in Geoadditive Regression Models." *Biometrics* 65, no. 2: 626–634.

Kosmas, C., V. Detsis, M. Karamesouti, K. Kounalaki, P. Vassiliou, and L. Salvati. 2015. "Exploring Long-Term Impact of Grazing Management on Land Degradation in the Socio-Ecological System of Asteroussia Mountains, Greece." *Land* 4, no. 3: 541–559.

Lamonica, G. R., M. C. Recchioni, F. M. Chelli, and L. Salvati. 2020. "The Efficiency of the Cross-Entropy Method When Estimating the Technical Coefficients of Input–Output Tables." *Spatial Economic Analysis* 15, no. 1: 62–91.

Lanfredi, M., G. Egidi, L. Bianchini, and L. Salvati. 2022. "One Size Does Not Fit All: A Tale of Polycentric Development and Land Degradation in Italy." *Ecological Economics* 192: 107256.

Meinshausen, N., and P. Bühlmann. 2010. "Stability Selection." *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 72, no. 4: 417–473.

Morianou, G., N. Kourgialas, V. Pisinaras, G. Psarras, and G. Arambatzis. 2021. "Assessing Desertification Sensitivity Map Under Climate Change and Agricultural Practices Scenarios: The Island of Crete Case Study." *Water Supply* 21, no. 6: 2916–2934.

Newey, W. K., and J. L. Powell. 1987. "Asymmetric Least Squares Estimation and Testing." *Econometrica* 55: 819–847.

Nigri, A., E. Barbi, and S. Levantesi. 2022. "The Relationship Between Longevity and Lifespan Variation." *Statistical Methods & Applications* 31, no. 3: 481–493.

Quaranta, G., R. Salvia, L. Salvati, et al. 2020. "Long-Term Impacts of Grazing Management on Land Degradation in a Rural Community of Southern Italy: Depopulation Matters." *Land Degradation & Development* 31, no. 16: 2379–2394.

Reich, B. J. 2012. "Spatiotemporal Quantile Regression for Detecting Distributional Changes in Environmental Processes." *Journal of the Royal Statistical Society. Series C, Applied Statistics* 61, no. 4: 535–553.

Romm, J. 2011. "Desertification: The Next Dust Bowl." *Nature* 478: 450–451.

Sallustio, L., D. Pettenella, P. Merlini, et al. 2018. "Assessing the Economic Marginality of Agricultural Lands in Italy to Support Land Use Planning." *Land Use Policy* 76: 526–534.

Salvati, L., A. Mancini, S. Bajocco, R. Gemmiti, and M. Carlucci. 2011. "Socioeconomic Development and Vulnerability to Land Degradation in Italy." *Regional Environmental Change* 11: 767–777.

Salvati, L., A. Mavrakis, A. Colantoni, G. Mancino, and A. Ferrara. 2015. "Complex Adaptive Systems, Soil Degradation and Land Sensitivity to Desertification: A Multivariate Assessment of Italian Agro-Forest Landscape." *Science of the Total Environment* 521: 235–245.

Salvati, L., L. Perini, A. Sabbi, and S. Bajocco. 2012. "Climate Aridity and Land Use Changes: A Regional-Scale Analysis." *Geographical Research* 50, no. 2: 193–203.

Salvati, L., and M. Zitti. 2008. "Regional Convergence of Environmental Variables: Empirical Evidences From Land Degradation." *Ecological Economics* 68, no. 1–2: 162–168.

Salvati, L., and M. Zitti. 2009. "Substitutability and Weighting of Ecological and Economic Indicators: Exploring the Importance of Various Components of a Synthetic Index." *Ecological Economics* 68, no. 4: 1093–1099.

Salvati, L., M. Zitti, and L. Perini. 2016. "Fifty Years on: Long-Term Patterns of Land Sensitivity to Desertification in Italy." *Land Degradation & Development* 27, no. 2: 97–107.

Seifollahi-Aghmiuni, S., Z. Kalantari, G. Egidi, L. Gaburova, and L. Salvati. 2022. "Urbanisation-Driven Land Degradation and Socioeconomic Challenges in Peri-Urban Areas: Insights From Southern Europe." *Ambio* 51, no. 6: 1446–1458.

Smiraglia, D., I. Tombolini, L. Canfora, S. Bajocco, L. Perini, and L. Salvati. 2019. "The Latent Relationship Between Soil Vulnerability to Degradation and Land Fragmentation: A Statistical Analysis of Landscape Metrics in Italy, 1960–2010." *Environmental Management* 64: 154–165.

Sobotka, F., and T. Kneib. 2012. "Geoadditive Expectile Regression." *Computational Statistics & Data Analysis* 56, no. 4: 755–767.

Spiegel, E., T. Kneib, and F. Otto-Sobotka. 2020. "Spatio-Temporal Expectile Regression Models." *Statistical Modelling* 20, no. 4: 386–409.

Taylor, J. W. 2008. "Estimating Value at Risk and Expected Shortfall Using Expectiles." *Journal of Financial Econometrics* 6, no. 2: 231–252.

Tomao, A., V. Quatrini, P. Corona, A. Ferrara, R. Lafortezza, and L. Salvati. 2017. "Resilient Landscapes in Mediterranean Urban Areas: Understanding Factors Influencing Forest Trends." *Environmental Research* 156: 1–9.

Tyralis, H., G. Papacharalampous, and S. Khatami. 2023. "Expectile-Based Hydrological Modelling for Uncertainty Estimation: Life After Mean." *Journal of Hydrology* 617: 128986.

Tzavidis, N., N. Salvati, T. Schmid, E. Flouri, and E. Midouhas. 2016. "Longitudinal Analysis of the Strengths and Difficulties Questionnaire Scores of the Millennium Cohort Study Children in England Using M-Quantile Random-Effects Regression." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179, no. 2: 427–452.

Vu, Q. M., Q. B. Le, E. Frossard, and P. L. Vlek. 2014. "Socio-Economic and Biophysical Determinants of Land Degradation in Vietnam: An Integrated Causal Analysis at the National Level." *Land Use Policy* 36: 605–617.

Waldmann, E., F. Sobotka, and T. Kneib. 2017. "Bayesian Regularisation in Geoadditive Expectile Regression." *Statistics and Computing* 27: 1539–1553.

Zambon, I., A. Benedetti, C. Ferrara, and L. Salvati. 2018. "Soil Matters? A Multivariate Analysis of Socioeconomic Constraints to Urban Expansion in Mediterranean Europe." *Ecological Economics* 146: 173–183.

Zambon, I., A. Colantoni, M. Carlucci, N. Morrow, A. Sateriano, and L. Salvati. 2017. "Land Quality, Sustainable Development and Environmental Degradation in Agricultural Districts: A Computational Approach Based on Entropy Indexes." *Environmental Impact Assessment Review* 64: 37–46.