

Multidimensional efficiency and institutional scale heterogeneity in protected areas: Evidence from Italy

Giovanni Signorello^{a,*}, Marika Cerro^a, Laura Giuffrida^a, Maria De Salvo^b

^a Department of Agricultural, Food and Environment (Di3A), University of Catania, via Santa Sofia 100, Catania 95123, Italy

^b Department of Veterinary, University of Messina, Polo Universitario dell'Annunziata, Messina 98168, Italy

ARTICLE INFO

Keywords:

Data Envelopment Analysis
Multidimensional Efficiency
Institutional Scale
Territorial Governance
National Parks
Protected Areas

ABSTRACT

Protected areas play a central role in territorial governance and the provision of environmental public goods, yet their performance is typically evaluated through effectiveness-based frameworks rather than through efficiency analysis. This study provides a multidimensional assessment of 27 Italian National Parks, conceptualized as multi-mission public organizations operating within complex territorial and institutional contexts. Four mission-specific domains - ecological performance, tourism valorization, managerial administration, and governance capacity - are evaluated using Data Envelopment Analysis within a non-radial framework that captures asymmetric inefficiencies across functions. The analysis explicitly examines the role of institutional scale, distinguishing between pure technical inefficiency and structural scale effects. Results reveal differentiated performance regimes: ecological outcomes are largely anchored in territorial endowment, whereas tourism and managerial domains exhibit both slack-driven and scale-related inefficiencies. Governance performance appears strongly conditioned by the configuration of institutional scales. These findings highlight the importance of scale alignment and mission-specific benchmarking for improving the design and coordination of protected-area governance systems.

1. Introduction

Protected areas represent one of the principal institutional instruments through which societies organize the provision of environmental public goods. National parks, in particular, are complex multi-mission public organizations entrusted with biodiversity conservation, landscape preservation, ecosystem service maintenance, recreational access, environmental education, and territorial development (Eagles et al., 2002; Watson et al., 2014). In economic terms, they operate as multi-mission public production units that transform financial, human, and territorial resources into heterogeneous environmental and institutional outputs. These outputs reflect partially overlapping conservation, recreational, managerial, and governance functions.

Despite the diversity of their mandates, the performance of protected areas is predominantly evaluated through effectiveness-based frameworks, such as the IUCN Protected Area Management Effectiveness (PAME) approach, which assess whether management objectives, governance arrangements, and planning instruments are in place and functioning. Although informative, such frameworks may underestimate

parks operating under input constraints, as they conflate goal attainment with resource adequacy. An efficiency-based perspective evaluates instead the relative capacity of parks to transform inputs into outputs, providing a complementary diagnostic tool that is particularly relevant in contexts of fiscal pressure and increasing demands for accountability.

Although a growing body of work has applied frontier methods to protected-area performance, three interrelated gaps continue to limit the empirical literature. First, evaluation has remained largely anchored in effectiveness-based logics, leaving the productive dimension of protected-area governance underexplored. Second, available efficiency studies typically aggregate heterogeneous functions into a single composite production technology, masking cross-mission disparities and obscuring the differentiated technological logics underlying ecological, tourism, managerial, and governance outcomes. Third, scale effects, when assessed at all, are rarely interpreted as structurally differentiated components of territorial environmental governance, despite their salience in multi-level institutional systems characterized by jurisdictional fragmentation and heterogeneous administrative scales. Addressing these gaps requires an empirical framework that is

* Corresponding author.

E-mail addresses: giovanni.signorello@unict.it (G. Signorello), marika.cerro@unict.it (M. Cerro), laura.giuffrida@unict.it (L. Giuffrida), maria.desalvo@unime.it (M. De Salvo).

<https://doi.org/10.1016/j.landusepol.2026.108289>

Received 24 February 2026; Received in revised form 3 July 2026; Accepted 18 August 2026

Available online 21 August 2026

0264-8377/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

simultaneously efficiency-based, multidimensional, and explicitly scale-aware. Such a framework is particularly relevant for informing policy decisions on resource allocation, inter-institutional coordination, and aligning ecological systems with administrative boundaries.

This study addresses these gaps through a multidimensional and scale-aware assessment of 27 Italian National Parks, conceptualized as multi-mission public production units operating within complex territorial and institutional contexts. Four mission-specific production frontiers - ecological performance, tourism valorization, managerial efficiency, and governance capacity - are estimated separately to preserve institutional heterogeneity. This approach contrasts with multi-criteria evaluation frameworks (e.g., (Munda, 2005)) in that it derives relative performance directly from observed input-output combinations rather than imposing *ex ante* normative weights. The analysis adopts a non-radial frontier approach (formalized in detail in Section 3) and complements efficiency estimation with statistically consistent second-stage inference to identify structural correlates of inefficiency.

The analysis addresses three interrelated questions: (i) How do Italian national parks perform across ecological, tourism, governance, and managerial domains within a non-radial frontier framework? (ii) To what extent do scale effects contribute to observed inefficiency patterns? (iii) Which structural and institutional characteristics are conditionally associated with multidimensional inefficiency once the statistical properties of DEA estimators are properly accounted for?

The study does not attempt to estimate socially optimal levels of provision for environmental public goods or to derive allocative welfare conclusions. Rather, it focuses on relative productive efficiency within the observed institutional setting, offering a useful complement to conventional effectiveness-based evaluations of protected areas.

The remainder of the paper is organized as follows. Section 2 reviews the literature on protected-area effectiveness, multidimensional efficiency assessment, and institutional scale heterogeneity, situating the present study within current debates. Section 3 develops the empirical strategy, formalizing the multidimensional production framework, presenting the non-radial Slack-Based Measure (SBM) under variable returns to scale, and outlining the bootstrap-truncated regression procedure for second-stage inference. Section 4 reports the empirical results across the four mission domains. Section 5 provides robustness and stability diagnostics. Section 6 discusses the findings in light of the institutional and governance characteristics of Italian national parks, with particular attention to policy implications for jurisdictional design and resource allocation. Section 7 concludes.

2. Literature review

2.1. From effectiveness to efficiency in protected-area evaluation

The evaluation of protected-area performance has traditionally been dominated by effectiveness-based frameworks. Building on the approach developed by Hockings (Hockings, 2006) and operationalized through the Protected Area Management Effectiveness (PAME) framework and the Global Database on Protected Area Management Effectiveness (Coad et al., 2015), these methods assess whether management objectives, governance arrangements, and planning instruments are in place and functioning. Large-scale empirical applications have further linked management effectiveness indicators to ecological outcomes (Geldmann et al., 2018; Schulze et al., 2018), contributing to global conservation policy debates and monitoring efforts within the post-2020 Global Biodiversity Framework.

By construction, however, effectiveness frameworks focus on goal attainment and institutional compliance, rather than on the productive use of resources. As a result, they may systematically favor protected areas with greater financial, human, or infrastructural endowments, without explicitly considering whether such inputs are used efficiently. In contexts with heterogeneous resource availability, this may introduce bias in performance assessment, particularly when comparing parks

operating under different budgetary or institutional constraints.

Efficiency analysis provides a complementary perspective. Following the seminal contribution of (Farrell, 1957) and the subsequent consolidation of frontier methods (Fried et al., 2008), efficiency is defined as the relative capacity of organizations to transform inputs into outputs. From this perspective, a protected area may achieve conservation targets while using more resources than comparable units or, conversely, perform relatively well despite limited inputs. Effectiveness and efficiency are therefore analytically distinct but jointly relevant, especially in policy contexts where accountability, resource allocation, and performance benchmarking are central concerns.

2.2. Efficiency analysis and multidimensional performance in protected areas

Data Envelopment Analysis (DEA), introduced by Charnes et al. (1978), provides a nonparametric framework for measuring relative technical efficiency by constructing an empirical best-practice frontier from observed input-output combinations. Its ability to accommodate multiple inputs and outputs without requiring price information has made it particularly suitable for public-sector and environmental applications (Dyckhoff and Allen, 2001; Worthington, 2000), including tourism systems (Cracolici et al., 2007; Liu et al., 2017) and regional ecological performance analysis (Bianchi et al., 2020).

Within the protected-area literature, DEA has been widely used to conceptualize parks as public production units transforming financial and human resources into environmental and recreational outcomes (Bosetti and Locatelli, 2006; Ma et al., 2009; Rusielik and Zbaraszewski, 2014). Subsequent contributions have broadened this approach across different institutional and geographical contexts. For example, Valdivieso et al. (2015) developed a DEA-based framework to assess tourism management capacity in protected areas, which was later extended to a cross-country analysis examining how policy decisions influence efficiency across 29 protected areas in five countries (Valdivieso et al., 2023). Other studies have investigated eco-efficiency and its determinants in national parks (Peng et al., 2017; Zhao et al., 2022), while applications in data-constrained environments (Jacob et al., 2018) and related domains such as urban green infrastructure (Sun et al., 2021) highlight the flexibility of frontier methods.

Despite this growing body of work, most empirical applications rely on a single aggregated production technology, implicitly assuming that protected areas pursue a unified objective function. This assumption is difficult to reconcile with the inherently multi-mission nature of these institutions. National parks simultaneously perform ecological, recreational, managerial, and governance functions, each characterized by distinct yet partially overlapping technological and institutional logics (Eagles et al., 2002; Lockwood, 2010). Aggregating these heterogeneous outputs into a single efficiency score risks masking cross-dimensional disparities and limiting the interpretability of results for policy purposes. Alternative evaluation frameworks, such as multi-criteria decision analysis (Munda, 2005), explicitly incorporate normative weighting schemes across dimensions, with applications ranging from sustainability indices for managed marine protected areas (Avelino et al., 2019) to multi-objective fishery management (Pascoe et al., 2023).

A second limitation concerns the treatment of inefficiency. Much of the literature relies on radial models, which assume proportional adjustment across outputs. In multi-objective public organizations, however, performance gaps are often unevenly distributed across dimensions. The non-radial Slack-Based Measure (SBM) proposed by Tone (2001) directly incorporates input excesses and output shortfalls into the efficiency score, thereby offering a more comprehensive representation of performance gaps in settings where different outputs reflect distinct policy goals. Pascoe et al. (2023) further argue that frontier methods provide a flexible alternative to multi-criteria approaches, as they generate “benefit-of-the-doubt” weights endogenously rather than

imposing them *ex ante*.

More broadly, the outputs generated by protected areas are intrinsically heterogeneous, spanning ecological outcomes, tourism services, and institutional achievements. This diversity reinforces the need for multidimensional performance frameworks that avoid *ex ante* aggregation and instead evaluate each mission domain within its own production space. Frontier methods are particularly well suited to this task, as they derive relative performance directly from observed input–output combinations without imposing normative weighting schemes. Recent work in adjacent domains - including the analysis of determinants of inefficiency in public parks and gardens services (Benito et al., 2025) - further supports this direction, highlighting the value of combining frontier estimation with rigorous second-stage diagnostics.

Despite these advances, three limitations remain in the empirical literature on protected-area efficiency. First, most studies continue to estimate aggregated frontiers, thereby conflating distinct production processes. Second, asymmetric inefficiencies across mission domains are rarely explicitly addressed through non-radial specifications such as SBM (Tone, 2001). Third, second-stage analyses of inefficiency determinants often rely on conventional econometric approaches (e.g., OLS or Tobit) that do not fully account for the statistical properties of frontier estimators, rather than on bootstrap-based methods that correct for the serial dependence inherent in DEA scores (Simar and Wilson, 2007).

2.3. Institutional scale and governance in protected-area systems

While the use of efficiency analysis in protected-area research has expanded, the role of institutional scale remains comparatively underexplored. In many applications, differences in scale are treated primarily as a technical modeling choice - typically through alternative assumptions on returns to scale - rather than as a substantive feature of environmental governance systems.

In practice, protected areas operate within multi-level institutional arrangements, often characterized by overlapping competencies, jurisdictional fragmentation, and coordination requirements across national, regional, and local authorities. These features imply that scale is not merely a statistical attribute but a structural dimension influencing organizational performance, transaction costs, and governance effectiveness.

Such dynamics are particularly relevant in the Italian context, where the national park system operates within a multi-level governance framework defined by Law n. 394/1991. Park authorities interact with regional governments, local municipalities, and other institutional actors through formal coordination mechanisms, including the mandatory institutional concertation with the Regions concerned and the Park Community Council, a statutory representative body of municipalities and regional authorities. This configuration corresponds to a shared-governance model in the IUCN typology (Borrini-Feyerabend et al., 2013), in which multiple actors contribute to decision-making processes and resource management.

These institutional arrangements introduce potential scale mismatches between ecological systems and administrative boundaries, as well as coordination costs associated with fragmented governance structures. From an efficiency perspective, such features may generate performance differences that are not attributable to managerial inefficiency alone, but rather to the structural configuration of institutional scale.

Despite their relevance, these aspects have received limited attention in empirical efficiency studies. Scale effects are often reported but rarely interpreted in relation to governance structures, territorial complexity, or policy design. Moreover, robustness diagnostics and statistically consistent second-stage inference are seldom integrated within a unified empirical framework.

2.4. Synthesis and contribution

The literature reviewed above highlights three interrelated gaps. First, the dominance of effectiveness-based approaches has left the productive dimension of protected-area performance underexplored, particularly regarding how resources are transformed into environmental and institutional outcomes. Second, within the efficiency literature, the widespread use of aggregated production technologies has obscured the multidimensional nature of protected-area functions, limiting the ability to identify mission-specific performance patterns. Third, the role of institutional scale in shaping efficiency outcomes has been insufficiently conceptualized, often treated as a technical modeling issue rather than as a core dimension of territorial governance.

This study addresses these gaps by adopting a multidimensional and scale-aware efficiency framework. Specifically, it evaluates protected areas across distinct mission domains - ecological performance, tourism valorization, managerial administration, and governance capacity - while explicitly examining the role of institutional scale and its interaction with performance outcomes. By integrating frontier analysis with scale-sensitive diagnostics and consistent second-stage inference based on the Simar–Wilson bootstrap truncated regression procedure (Simar and Wilson, 2007), the study provides an empirically grounded approach to understanding how governance structures and territorial configurations influence the efficiency of protected-area systems.

3. Empirical strategy: multidimensional frontier estimation and bootstrap inference

The empirical strategy adopted in this study is structured into two sequential, conceptually distinct phases. In the first phase, we estimate multidimensional technical efficiency scores using a nonparametric frontier approach. In the second phase, we analyze the structural and institutional correlates of inefficiency using a bootstrap-truncated regression model consistent with the statistical properties of nonparametric frontier estimators. This two-stage framework allows us to separate the measurement of relative performance from the identification of systematic conditional associations, while ensuring statistical coherence between the two steps.

The methodological choice of a non-radial frontier specification is motivated by the multi-mission nature of the production technology under analysis. National parks generate heterogeneous outputs whose performance gaps are unlikely to be proportional: a park may operate close to the ecological frontier while exhibiting substantial slack in governance outcomes, or vice versa. Radial models, which assume proportional expansion of all outputs, would obscure such asymmetric inefficiencies. The non-radial Slack-Based Measure (SBM) proposed by Tone (Tone, 2001) directly incorporates input excesses and output shortfalls into the efficiency score, thereby providing a more comprehensive representation of performance gaps in multi-objective public organizations. Scale heterogeneity is examined through systematic comparisons of constant- and variable-returns-to-scale specifications, enabling a structured decomposition of pure technical inefficiency and institutional scale effects. Finally, statistically consistent inference on the structural correlates of inefficiency is obtained through the Simar–Wilson bootstrap truncated regression procedure, which corrects for the serial dependence inherent in DEA efficiency scores.

3.1. Conceptual framework and production technology

Italian national parks are modeled as multi-mission public production units operating within a multidimensional technological space. As discussed in the Introduction, their statutory mandates combine ecological conservation, tourism valorization, managerial administration, and institutional governance. These mission domains generate heterogeneous, largely non-market outputs that cannot be meaningfully aggregated through observed prices.

From a production-theoretic perspective, each park transforms a vector of inputs $x_i \in R_+^m$ into a vector of outputs $y_i \in R_+^s$.

Because no market price system exists to construct a scalar performance index, efficiency is defined relative to an empirical production possibility set. This set is assumed to satisfy standard regularity conditions - convexity, free disposability of inputs and outputs, and closedness. Efficiency is therefore interpreted as the distance of an observed input–output combination from the boundary of the production set.

To preserve institutional heterogeneity and avoid dimensionality distortions, four mission-specific technologies are estimated separately: ecological performance, tourism valorization, managerial efficiency, and governance capacity. This structure reflects the multidimensional nature of public production and ensures that efficiency comparisons remain domain-specific rather than artificially aggregated.

Formally, let parks be indexed by $i = 1, \dots, N$. Each park employs an input vector x_i in R_+^m to generate an output vector y_i in R_+^s . The production possibility set is defined as

$$\mathcal{T} = \{(x, y) \in R_+^m \times R_+^s : x \text{ can produce } y\}$$

The technology is assumed to satisfy standard regularity conditions - convexity, free disposability, and closedness - consistent with the nonparametric production framework developed in Färe et al. (1994). These assumptions ensure the existence of a well-defined empirical frontier against which observed production plans can be evaluated. Efficiency is therefore interpreted as the distance of an observed input–output combination from the boundary of $\mathcal{T}$. The empirical frontier is defined as the smallest convex set that envelops the observed input–output combinations, thereby defining efficiency in purely relative terms within the institutional sample. The baseline technology is estimated under variable returns to scale, allowing the empirical frontier to adapt to heterogeneous institutional scale configurations.

Given the multi-mission structure of protected areas and the moderate sample size ($N = 27$), estimating a single high-dimensional frontier would substantially increase total dimensionality ($m + s$) and potentially weaken discriminatory power in a nonparametric setting. Instead, we estimate four conceptually distinct production technologies corresponding to ecological performance, tourism valorization, managerial efficiency, and governance capacity. This mission-specific specification preserves institutional heterogeneity while maintaining proportionality between sample size and dimensionality, thereby supporting statistical reliability of the frontier estimator.

3.2. Data and variable construction

The empirical analysis is based on a cross-sectional dataset covering the Italian national-level protected-area system, comprising 27 major parks operating under multi-mission mandates of conservation, tourism management, and institutional governance.¹ All units share comparable territorial scale, ecological relevance, managerial complexity, and tourism pressure. The dataset reflects the full set of operational Italian National Parks, thereby representing the complete national production system rather than a statistical subsample. The number of decision-making units is therefore institutionally determined.

Fig. 1 illustrates the geographic distribution of the 27 Italian National Parks, highlighting the territorial heterogeneity underlying the national protected-area system.

From a production-theoretic perspective, these protected areas operate under structurally similar technological and institutional conditions. The resulting cross-sectional variation in environmental endowment, infrastructure, financial resources, and governance arrangements supports robust frontier estimation.

¹ The dataset comprises all operational Italian national parks and three regional parks characterized by national-equivalent territorial scale and governance mandates.

Data were compiled from official administrative reports, park management plans, and institutional documentation published by each park authority. Additional sources include reports issued by the Italian Court of Auditors and official publications from national public agencies, including ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale (2023)) and ISTAT (2023). Data were also retrieved from the official national portal of Italian protected areas (Italian Parks Portal, www.parks.it). Supplementary information was obtained from international environmental databases, including UNEP-WCMC and IUCN (UNEP-WCMC and IUCN, 2026), as well as the Digital Observatory for Protected Areas (DOPA; (Dubois et al., 2019)).

Table 1 reports descriptive statistics for all variables included in the analysis, including both the input–output variables used in the DEA models and the structural covariates employed in the second stage truncated regression. The dispersion patterns reveal substantial heterogeneity across territorial scales, institutional configurations, and operational capacities, thereby justifying the adoption of a multidimensional, scale-aware efficiency framework.

The mission-specific production technologies are summarized in Table 2. Each domain is modeled as a distinct input–output transformation process, reflecting the multidimensional mandate of national parks.

The ecological performance model captures the transformation of territorial endowment and management-mediated accessibility into measurable environmental outcomes. Park surface, environmental infrastructure expenditure, and trail network length jointly represent the structural and managerial inputs that generate carbon stock, forest cover, and biodiversity levels.

The tourism valorization model conceptualizes tourism performance as the activation of environmental and infrastructural capital. Accommodation facilities, visitor centers, and trail infrastructure enable the production of visitor flows and perceived recreational quality.

The management efficiency model focuses on the administrative transformation of financial and human resources into operational outputs, measured through visitor numbers and commercial revenues.

Finally, the governance capacity model isolates the institutional dimension, assessing how human resources translate into governance quality and formal compliance mechanisms.

All models are estimated using a non-oriented Slack-Based Measure (SBM) under variable returns to scale (VRS), allowing simultaneous proportional and non-proportional adjustments in inputs and outputs.

3.3. Mission-specific DEA specifications

To operationalize the multidimensional production framework outlined above, four mission-specific DEA models are estimated separately. Rather than aggregating heterogeneous public objectives into a single high-dimensional technology, each mission domain is modeled as an autonomous production space defined by its own input–output configuration. This specification preserves institutional heterogeneity, reduces dimensionality bias in a moderate sample, and ensures the interpretability of efficiency scores within each functional domain.

The mission-specific input–output configurations are presented in Table 2. Each domain reflects a distinct technological logic consistent with the statutory mandate of national parks.

All models are estimated using a non-oriented Slack-Based Measure (SBM) under variable returns to scale (VRS), as introduced in Section 2.2. The non-oriented formulation is appropriate because both input excesses and output shortfalls are economically meaningful in the context of public environmental production. Parks may simultaneously exhibit underutilized resources and unrealized output potential; the SBM framework captures these non-proportional inefficiencies through slack adjustments.

The separation between domains is not merely descriptive but reflects structurally distinct production processes. Ecological performance represents the transformation of territorial endowment and

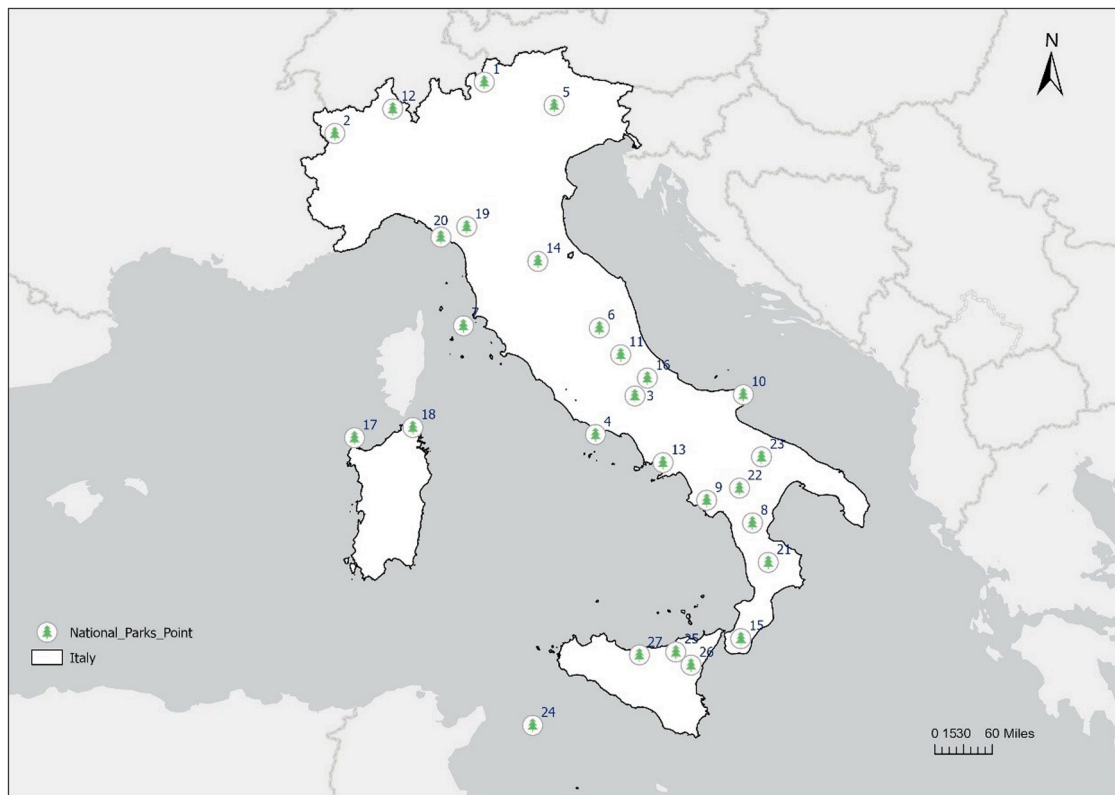


Fig. 1. Geographic distribution of the 27 Italian National Parks included in the analysis.

management-mediated accessibility into measurable environmental outcomes. Tourism valorization captures the activation of environmental and recreational infrastructure into visitor performance and perceived quality. Management efficiency concerns the administrative transformation of financial and human resources into operational outputs. Governance capacity isolates the institutional dimension, assessing how organizational structures translate into formal compliance and governance quality.

In the ecological production model, park surface captures the natural endowment dimension, while environmental infrastructure expenditure and trail network length reflect management intensity and anthropogenic accessibility. Although trail infrastructure is commonly associated with tourism, it also constitutes an ecologically relevant management variable. Trail density influences habitat exposure, spatial fragmentation, distribution of visitor pressure, and monitoring capacity. Well-designed and regulated trail systems may mitigate uncontrolled access and reduce diffuse ecological disturbance, whereas poorly managed or excessive trail networks may increase pressure on ecosystems (Buckley, 2009; Farrell and Marion, 2001; Lynn and Brown, 2003; Monz et al., 2013). Including trail length in the ecological specification, therefore, allows the frontier to account not only for passive territorial endowment but also for the interaction between conservation management and controlled accessibility. Ecological production is thus modeled as the transformation of both natural capital and management-mediated access into environmental outcomes.

Governance outputs are partly institutional and discrete. The combination of a continuous governance index and a binary indicator of approved management plans reflects formal compliance and institutional activation rather than physical production. This hybrid output structure introduces potential scale sensitivity under constant returns to scale, further justifying the adoption of a VRS specification.

Because each mission is evaluated within its own technological space, efficiency scores are not directly comparable across domains in absolute magnitude. They represent relative performance within each

mission-specific frontier. Cross-domain comparisons, therefore, concern structural patterns rather than ordinal ranking equivalence.

To ensure adequate discriminatory power and mitigate frontier inflation, Table 3 reports the dimensional structure of each model. All specifications satisfy the proportionality condition discussed in Section 3.7,

$$N \geq \max\{m \times s, 3(m+s)\}.$$

where m denotes the number of inputs and s the number of outputs. The resulting ratios confirm that the sample size is sufficient relative to dimensionality, supporting the stability and credibility of the estimated frontiers.

3.4. Radial DEA and the output distance function

Classical Data Envelopment Analysis (DEA) measures efficiency through radial scaling of inputs or outputs. The output distance function is defined as

$$D^{out}(x, y) = \sup\{\theta > 0 : (x, \theta y) \in \mathcal{T}\}.$$

The Farrell output-oriented efficiency measure is

$$\theta_i = 1/D^{out}(x_i, y_i).$$

Under variable returns to scale (VRS), the technology is the convex hull of observed input–output combinations. The envelopment form of the output-oriented BCC model (Banker et al., 1984) solves

$$\max_{\theta, \lambda} \theta$$

subject to:

$$X\lambda \leq x_i, \quad Y\lambda \geq \theta y_i, \quad 1^\top \lambda = 1, \quad \lambda \geq 0$$

Under constant returns to scale (CRS), the convexity constraint $1^\top \lambda = 1$ is removed.

Table 1
Descriptive statistics.

Panel A – DEA Variables				
Variable	Mean	SD	Min	Max
Total cost (€)	4,660,379	3,848,587	964,857	16,099,109
Total employees (number)	25.000	36.530	6.000	165.000
Number of accommodations (number)	210.000	166.000	1.000	631.000
Km of trails (km)	323.560	315.330	1.200	1500.000
Visitor centers (number)	10.000	9.630	1.000	41.000
Attendances (thousand visitors)	506.037	853.464	16.659	3136.502
Average rating (count)	129.490	200.610	2.400	794.000
Mean carbon stock (Mg C km ⁻²)	19,740.74	5239.96	10,523.08	29,842.02
Forest cover (%)	42.070	21.460	3.300	78.600
Biodiversity index (Shannon)	1.660	0.460	0.390	2.450
Governance index	1.480	1.014	0.000	4.000
Park plan (dummy)	0.290	0.460	0.000	1.000
Total visitors (thousand visitors)	119.787	175.078	5.540	765.910
Sale of commercial items (€)	299,701	561,280	0.000	1948,454
Panel B – Regression Variables				
Variable	Mean	SD	Min	Max
Park surface (km ²)	634.90	538.40	39.00	1822.22
Management plan maturity (years since approval)	5.660	6.010	0.000	17.000
Number of regions involved	1.410	0.690	1.000	3.000
Park age (years)	37.290	25.270	7.000	101.000
Park brand (dummy)	0.370	0.490	0.000	1.000
Accessibility (50 km index)	5019.407	3435.674	64.000	12,882.000
LIFE participation (count)	1.740	2.280	0.000	9.000
UNESCO Biosphere Reserve (dummy)	0.370	0.490	0.000	1.000
Territorial location (macro-area dummy)	0.220	0.420	0.000	1.000

Notes: Descriptive statistics are based on the full sample of 27 Italian national parks. Monetary variables are expressed in euros (€). Visitor-related variables are expressed in thousands of visitors. Mean carbon stock is expressed in Mg C km⁻². Dummy variables take 0–1. All statistics are rounded to three decimal places.

Radial models assume proportional expansion of all outputs. In multi-objective public organizations, however, inefficiencies are unlikely to be uniformly distributed across dimensions. A park may be close to the ecological frontier while exhibiting substantial slack in governance outcomes. The proportionality restriction of radial models would obscure such asymmetric performance gaps.

Radial efficiency measures collapse multidimensional performance gaps into a single proportional expansion factor.

Table 2
Input–output structure of mission-specific DEA models.

Mission Domain	Inputs (unit of measurement)	Outputs (unit of measurement)	Technological Interpretation
Ecological Performance	Park surface (km ²); Environmental infrastructure expenditure (€); Trail network length (km)	Mean carbon stock (Mg C km ⁻²); Forest cover (%); Biodiversity index (Shannon)	Transformation of territorial endowment and management-mediated accessibility into ecological outcomes
Tourism Valorization	Accommodation facilities (number); Visitor centers (number); Trail network length (km)	Attendances (thousand visitors); Average online reviews (count)	Activation of environmental and infrastructural capital into tourism performance
Management Efficiency	Operating expenditure (€); Total employees (number)	Visitors (thousand visitors); Sale of commercial items (€)	Administrative transformation of financial and human resources into operational outputs
Governance Capacity	Total employees (number)	Governance index (composite score); Approved management plan (dummy: 0/1)	Institutional configuration and compliance within governance architectures

Notes: Monetary variables are expressed in euros (€). Mean carbon stock is expressed in Mg C km⁻². The trail network variable enters both the ecological and tourism models to capture its dual role as management-mediated ecological access and tourism infrastructure. All mission-specific models are estimated using a non-oriented Slack-Based Measure (SBM) under variable returns to scale (VRS).

3.5. Slack-based measure (SBM) under variable returns to scale

To relax the proportionality assumption, we adopt the non-oriented Slack-Based Measure (SBM) proposed by Tone (Tone, 2001). The non-oriented specification is preferred because both input excesses and output shortfalls are economically relevant in public environmental production. National parks may exhibit resource misallocation (input slack), for instance, in staffing or infrastructure intensity, while simultaneously underperforming in specific output dimensions such as governance capacity or ecological outcomes. A purely output-oriented model would focus exclusively on output expansion and ignore potential input inefficiencies, whereas an input-oriented model would disregard output shortfalls. The non-oriented SBM allows inefficiency to emerge along both dimensions, providing a more comprehensive representation of performance gaps in multi-mission public organizations.

The SBM directly incorporates input excesses and output shortfalls - referred to as slacks - into the efficiency score. Let

$$X = [x_1, \dots, x_N], \quad Y = [y_1, \dots, y_N].$$

For a given park o , define input slack variables $s^- \in R_+^m$ and output slack variables $s^+ \in R_+^s$.

The non-oriented SBM under VRS solves the fractional program:

$$\min_{\lambda, s^-, s^+} \rho_o = \left(1 - \frac{1}{m} \sum_{j=1}^m \frac{s_j^-}{x_{jo}} \right) / \left(1 + \frac{1}{s} \sum_{r=1}^s \frac{s_r^+}{y_{ro}} \right)$$

subject to

$$x_o = X\lambda + s^-, y_o = Y\lambda - s^+, 1'\lambda = 1, \quad \lambda \geq 0, \quad s^- \geq 0, \quad s^+ \geq 0.$$

Efficiency satisfies $0 < \rho_o \leq 1$, and $\rho_o = 1$ if and only if all slack variables are zero, corresponding to Pareto efficiency.

The fractional objective function can be transformed into an equivalent linear program via the Charnes–Cooper transformation, ensuring computational tractability.

Because the SBM objective function contains ratios of the form s_r^+/y_{ro} , strictly positive output values are required for all evaluated units. For outputs that may take zero values, a small positive constant $\varepsilon > 0$ is added to ensure computational feasibility. The constant $\varepsilon > 0$ was set to a negligible positive value and verified not to affect efficiency rankings. Where needed, we replaced y_{ro} with $y_{ro} + \varepsilon$ in the normalized

Table 3
Dimensional structure of mission-specific DEA models.

Dimension	Number of Inputs (m)	Number of Outputs (s)	m + s	N / (m + s)
Ecological	3	3	6	4.50
Tourism	3	2	5	5.4
Management	2	2	4	6.75
Governance	1	2	3	9.0

slack ratios s_r^+ / y_{ro} (equivalently, outputs were shifted by ε before estimation).

3.6. Returns to scale and scale efficiency

Under VRS, the convexity constraint $1'\lambda = 1$ allows for increasing, constant, or decreasing returns to scale. Removing this constraint yields the CRS specification.

Let ρ_i^{VRS} and ρ_i^{CRS} denote the SBM efficiency scores. Scale efficiency is defined as

$$SE_i = \rho_i^{CRS} / \rho_i^{VRS}$$

If $SE_i < 1$, part of the observed inefficiency reflects scale distortions rather than purely technical inefficiency.

3.7. Dimensionality and finite-sample considerations

As nonparametric boundary estimators, DEA efficiency scores are subject to well-known dimensionality and finite-sample considerations. Under standard regularity conditions - smooth frontier, independence of observations, and compact support - DEA estimators are consistent but converge at rate

$$N^{(-2/(m+s))}$$

where $m+s$ denotes total dimensionality. This convergence rate deteriorates rapidly as dimensionality increases, implying the well-known “curse of dimensionality” in nonparametric frontier estimation (Kneip et al., 1998; Simar and Wilson, 2008).

To preserve discriminatory power in moderate samples, we follow commonly adopted proportionality guidelines in the DEA literature and ensure that

$$N \geq \max\{m \times s, 3(m+s)\}.$$

Each mission-specific frontier satisfies this proportionality rule. Estimating four parsimonious frontiers instead of a single high-dimensional technology preserves discriminatory power and mitigates finite-sample frontier inflation.

3.8. Structural correlates of inefficiency: bootstrap truncated regression

The second phase of the empirical strategy investigates the structural and institutional correlates of inefficiency across the four mission-specific production technologies. Conceptually, this stage moves from measuring relative technical performance to examining whether systematic differences in organizational, territorial, and institutional characteristics are conditionally associated with observed efficiency gaps.

The first-stage SBM estimator provides a nonparametric measure of relative technical efficiency with respect to the observed production frontier. However, efficiency scores do not explain why deviations from best practice arise. The second stage addresses this issue by modeling inefficiency as a function of exogenous structural characteristics.

Following Simar and Wilson (2007), we estimate a truncated-normal regression model for the bias-corrected inefficiency scores. Specifically, inefficiency is defined as

$$\delta_i = 1 / (\hat{\rho}_i^{bc} - 1), \delta_i \geq 0$$

where $(\hat{\rho}_i^{bc})$ denotes the bias-corrected efficiency estimate. This transformation maps efficiency scores bounded in $(0,1]$ into a non-negative inefficiency measure, where zero corresponds to full efficiency and positive values indicate increasing technical inefficiency.

The second-stage model is specified as:

$$\delta_i = \alpha + Z_i\beta + \varepsilon_i$$

where Z_i is a vector of structural and institutional characteristics, and $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. The regression is estimated using a likelihood function truncated at zero to respect the non-negativity of δ_i . Under the regularity conditions discussed by Simar and Wilson (2007), the bootstrap procedure provides consistent inference in finite samples and corrects for the serial dependence inherent in DEA efficiency scores.

The regression does not model production directly; rather, it models deviations from the estimated frontier. In economic terms, the parameters β capture conditional associations between structural characteristics and relative technical inefficiency. A negative coefficient indicates that higher values of the corresponding structural variable are associated with lower inefficiency (i.e., greater relative efficiency), holding the frontier structure constant. The specification is therefore comparative and reduced form: it identifies systematic correlates of inefficiency rather than structural production parameters or causal effects.

A central identifying assumption of the Simar–Wilson framework is that the explanatory variables Z_i are exogenous with respect to the data-generating process of inefficiency, conditional on the production frontier. More formally, the identifying assumption is that Z_i affects the conditional distribution of inefficiency but does not shift or alter the underlying production frontier itself. The frontier is treated as common across units in the first stage, and the second-stage regression explains deviations from that frontier rather than structural differences in technology.

We do not formally test the separability condition between Z_i and the production technology, but adopt the standard reduced-form interpretation consistent with the Simar–Wilson framework. Conditional nonparametric frontier models, as developed by Daraio and Simar (2007), provide an alternative route to incorporate environmental variables directly into the frontier; the present study, however, retains the two-stage formulation for analytical transparency and direct comparability with prior protected-area applications. In other words, Z_i may influence inefficiency but are not themselves determined by contemporaneous random shocks affecting the production process.

The set of explanatory variables differs across mission domains to reflect the structural specificity of each production technology.

Table 4 summarizes the structural covariates included in the second-stage truncated regression models. Variable selection follows a domain-specific logic consistent with the technological structure of each mission-specific frontier.

This domain-specific specification ensures that structural covariates are aligned with the institutional logic of each production technology and avoids imposing a uniform set of determinants across heterogeneous mission domains.

In the context of national parks, variables - such as territorial extent,

Table 4

Structural variables included in the Simar–Wilson bootstrap truncated regression models.

Variable (Code)	Description	Expected Association	Domains Included
Planning maturity (piano_c)	Recency of approved management plan	Negative (↓ inefficiency)	Management; Governance
Governance fragmentation (enti_c)	Number of regions involved in administration	Positive (↑ inefficiency)	Management
Park age (eta_c)	Years since establishment	Ambiguous / Positive	Tourism
Territorial extension (superf_c)	Park surface	Scale-related	Management
UNESCO Biosphere MAB (biosphere_MAB)	UNESCO designation	Negative	Ecological

Notes: Domain-specific covariates explaining bias-corrected SBM inefficiency scores (δ_i). Expected signs indicate the direction of conditional association with inefficiency.

the number of regions involved in governance, the park's historical age, or macro-territorial location - can reasonably be treated as pre-determined structural characteristics. These variables are institutionally or geographically fixed and plausibly exogenous to short-run managerial performance.

Other institutional variables, such as certifications or participation in European programs, may partly reflect accumulated historical performance. For this reason, the second-stage analysis is explicitly interpreted as identifying conditional correlates rather than structural causal determinants. The regression framework does not claim that changes in each institutional characteristic would mechanically cause efficiency improvements. Rather, it reveals systematic associations between organizational features and relative frontier distance.

This distinction is particularly important in public-sector settings, where institutional attributes may be jointly determined with performance over longer time horizons. The truncated regression is therefore interpreted as a reduced-form characterization of inefficiency heterogeneity, not as a structural production equation.

Since DEA efficiency scores are estimated and serially correlated through their dependence on a common frontier, conventional two-stage OLS or Tobit procedures may yield inconsistent or biased inference (McDonald, 2009; Simar and Wilson, 2011).

Following Simar and Wilson (2007), Algorithm 2, we implement a bootstrap-truncated regression procedure. Under the regularity conditions discussed by Simar and Wilson (2007), the bootstrap procedure provides consistent inference in finite samples and improves the statistical reliability of second-stage estimation.

The robustness and numerical stability of the second-stage results are subsequently evaluated through dedicated diagnostics presented in Section 4.

All efficiency estimations were performed in R. Mission-specific SBM-VRS and SBM-CRS frontiers were estimated using the *rDEA* package (Bogetoft and Otto, 2010), while bootstrap bias correction and second stage truncated regressions were implemented using the *FEAR* package within the Simar and Wilson framework (Wilson, 2008). These tools provide established and transparent implementations of nonparametric frontier estimation and inference procedures.

3.9. Robustness and stability analysis

Building on the two-stage framework described above, robustness diagnostics are implemented to evaluate the stability of both the frontier construction and second-stage inference. Given the institutional scale of the national system and the multidimensional specification of the production technologies, nonparametric efficiency analysis may be sensitive to influential observations, dimensionality constraints, assumptions regarding returns to scale, and resampling variability in second-stage inference. The robustness strategy is therefore designed to evaluate each potential source of fragility.

First, frontier stability is examined through a leave-one-out (LOO) procedure. Each park is sequentially excluded, and SBM efficiency scores are recomputed under the baseline VRS specification. Rank correlations and absolute deviations relative to baseline estimates are used to assess whether frontier construction is disproportionately influenced by individual benchmark units.

Second, the distribution of efficient units is analyzed to evaluate the discriminatory power of each mission-specific frontier. In moderate samples, excessive clustering at an efficiency of 1 may indicate dimensionality-induced frontier inflation. The share of fully efficient parks is therefore examined for each specification.

Third, sensitivity to returns-to-scale assumptions is assessed by comparing efficiency scores under variable returns to scale (VRS) and constant returns to scale (CRS). While VRS isolates pure technical inefficiency at the observed scale, CRS captures both technical and scale inefficiency. The comparison allows evaluation of the extent to which relative performance rankings depend on the imposed scale structure.

Fourth, the numerical stability of the second-stage Simar–Wilson truncated regression is verified by re-estimating each model using alternative numbers of bootstrap replications ($B = 1000, 2000$, and 5000). The stability of coefficient estimates and confidence intervals across replication sizes is used to assess sensitivity to resampling variability.

Taken together, these diagnostics are designed to verify that the multidimensional efficiency patterns identified in the baseline specification are not artifacts of influential observations, dimensionality constraints, scale assumptions, or computational instability. While robustness checks cannot eliminate all finite-sample limitations inherent to nonparametric frontier models, they provide a structured framework for assessing the empirical stability of the results.

Fig. 2 summarizes the empirical workflow adopted in the study, from frontier estimation to second-stage inference and robustness diagnostics.

4. Results

4.1. Multidimensional efficiency under the SBM–VRS specification

We begin by presenting the baseline efficiency estimates obtained under the non-oriented Slack-Based Measure (SBM) with variable returns to scale (VRS). These scores represent relative technical efficiency within each mission-specific production technology and therefore provide the reference structure for all subsequent comparisons.

Fig. 3 presents the distribution of SBM–VRS efficiency scores across the four mission domains. The boxplot representation allows a direct comparison of dispersion, median performance, and frontier

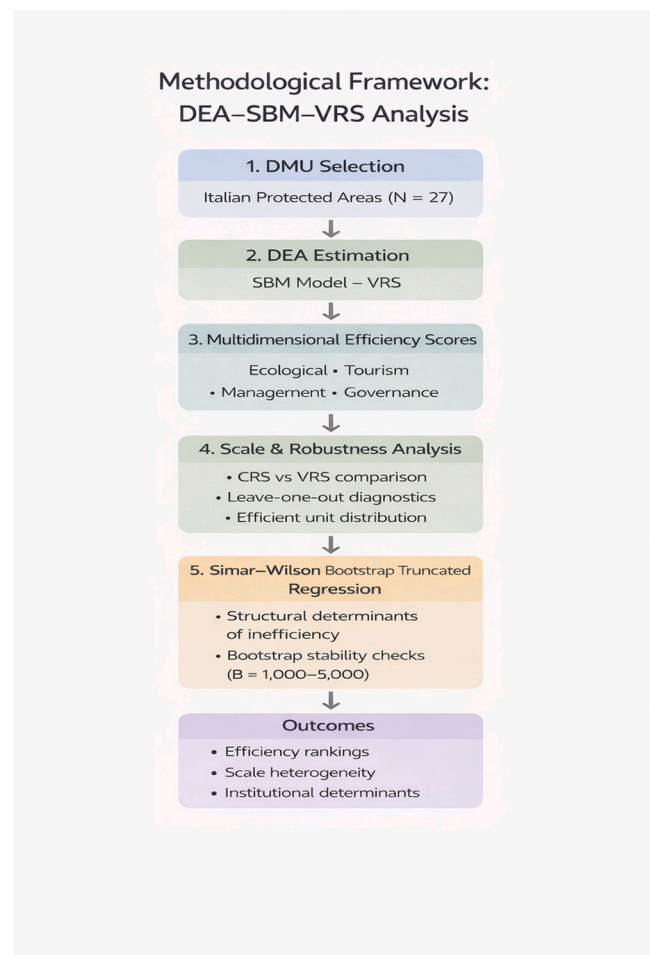


Fig. 2. Empirical workflow of the multidimensional DEA–SBM analysis.

concentration patterns within each technological space.

Table 5 summarizes the comparison between VRS and CRS specifications across mission domains. The complete park-level VRS and CRS efficiency scores are reported in Appendix **Table A1**.

Ecological efficiency appears largely scale-robust: mean efficiency declines only marginally under CRS, and scale efficiency remains close to unity. Tourism and management display moderate reductions in mean efficiency and lower shares of frontier units under CRS, indicating the coexistence of technical and scale distortions. Governance presents a qualitatively different structure. While mean efficiency under VRS is relatively high and over half of parks are efficient, CRS estimation yields a substantial decline in both mean efficiency and the share of frontier units. The low Spearman rank correlation confirms that governance rankings are highly sensitive to scale assumptions, consistent with a structurally scale-dominated production process.

4.1.1. Ecological performance

Ecological efficiency emerges as comparatively high and exhibits a pronounced concentration near the frontier. Under VRS, 16 out of 27 parks (59.3%) are classified as fully efficient. While this proportion exceeds that observed in other mission domains, it does not signal dimensionality-induced frontier inflation. As discussed in **Section 5**, the proportionality condition is satisfied ($N \geq 3(m+s)$), and leave-one-out diagnostics confirm strong ranking stability. The observed frontier concentration therefore reflects structural similarity in ecological endowment and conservation performance rather than mechanical overfitting of the nonparametric estimator.

Once territorial surface, environmental infrastructure, conservation expenditure, and management-mediated accessibility (including trail infrastructure) are incorporated into the production set, cross-sectional dispersion remains limited. This suggests that ecological production is structurally anchored in territorial and environmental endowment. While managerial decisions and infrastructure configuration may introduce marginal variation, the dominant determinant of ecological efficiency appears to be the underlying natural capital base rather than short-run institutional differentiation.

Table 5

Summary comparison of VRS and CRS efficiency across mission domains.

Mission Domain	Mean VRS	Mean CRS	Mean SE	Share Eff. VRS	Share Eff. CRS	Spearman ρ
Ecological	0.959	0.933	0.973	0.593	0.296	0.751
Tourism	0.925	0.779	0.842	0.407	0.148	0.868
Management	0.914	0.882	0.966	0.407	0.333	0.777
Governance	0.749	0.498	0.665	0.519	0.185	0.126

Notes: Efficiency scores are estimated using the non-oriented Slack-Based Measure (SBM) under variable returns to scale (VRS). Values range from 0 to 1, where 1 denotes full technical efficiency relative to the mission-specific production frontier.

The ecological frontier reflects a technologically cohesive space characterized by a strong concentration of parks near their structural potential.

4.1.2. Tourism valorization

In contrast, tourism efficiency exhibits greater heterogeneity. Although approximately 41 % of parks lie on the frontier under VRS, the distribution of scores is more dispersed than in the ecological domain.

Inefficiency in tourism production frequently reflects output shortfalls - such as lower attendance or weaker visitor satisfaction - rather than excessive infrastructure inputs. The non-radial SBM specification is particularly relevant here, as it captures asymmetric slack patterns across outputs and inputs. Some parks display underutilized tourism potential despite adequate infrastructure endowment, indicating that performance gaps are more closely related to activation capacity, positioning strategies, or managerial alignment than to infrastructural scarcity per se.

Tourism production, therefore, appears to combine structural characteristics with managerial responsiveness, generating a more articulated efficiency structure than the relatively endowment-driven ecological domain.

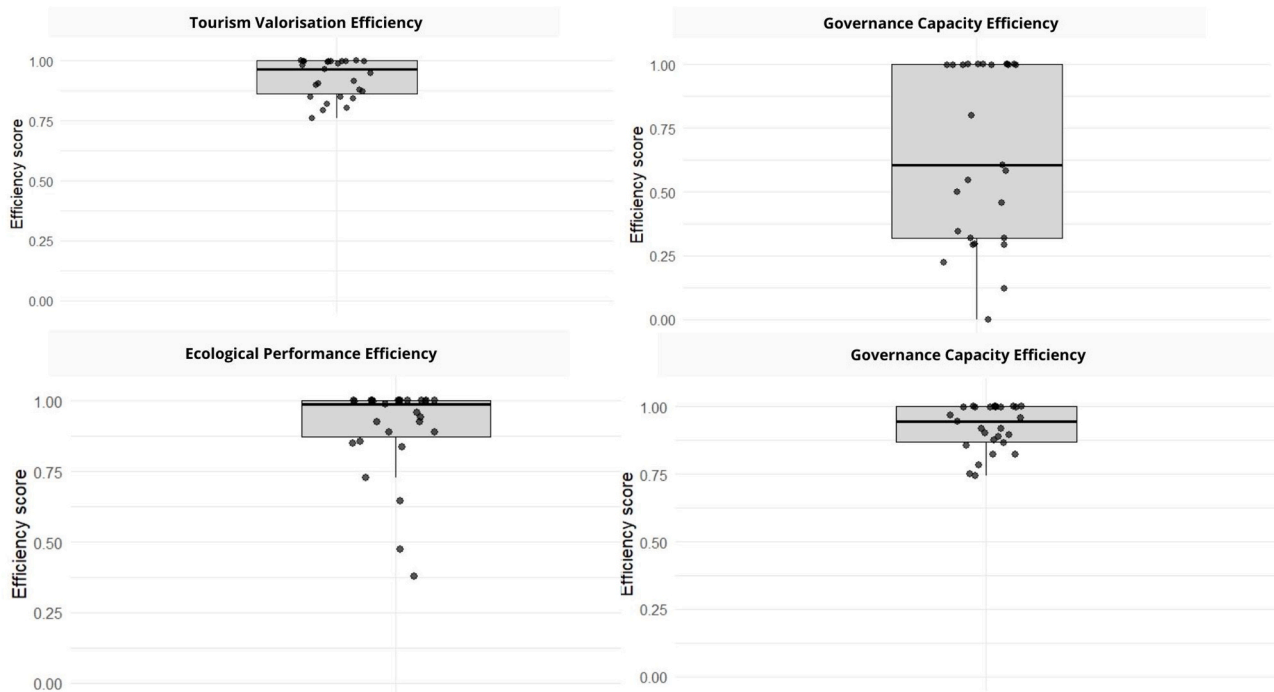


Fig. 3. Distribution of SBM-VRS efficiency scores across mission domains. Notes: Boxplots of SBM-VRS efficiency scores for ecological performance, tourism valorization, management efficiency, and governance capacity. Boxes represent the interquartile range; horizontal lines indicate medians; whiskers denote the range excluding outliers. Efficiency values are bounded between 0 and 1.

4.1.3. Management efficiency

Management efficiency occupies an intermediate position. Under VRS, mean efficiency equals 0.914, with approximately 41 % of parks operating at the frontier.

Because the SBM integrates both input excesses and output shortfalls, inefficiency in this domain reflects either overutilization of financial and human resources relative to outcomes or insufficient operational outputs given observed inputs. From an administrative perspective, this identifies potential margins for organizational alignment between resource allocation and mission delivery.

Compared with ecological production, managerial performance appears more sensitive to internal administrative configuration and less tightly constrained by structural endowment. At the same time, dispersion is not extreme, suggesting that managerial technologies are relatively comparable across the Italian park system.

4.1.4. Governance capacity

Governance efficiency exhibits a distinctive configuration. Under VRS, mean governance efficiency equals 0.749, and approximately 52 % of parks are fully efficient. This concentration near the frontier is higher than in other domains, but it does not reflect complete efficiency saturation.

Governance outputs - such as approved management plans, certifications, and coordination instruments - are institutional and partly discrete in nature. As a result, variation under flexible scale assumptions may be limited once minimum compliance thresholds are met. Parks operating at comparable administrative scale tend to cluster near the frontier when evaluated relative to peers.

However, interpreting governance efficiency requires an explicit comparison with CRS results, as scale effects play a central role in this domain.

Across all four dimensions, it is important to emphasize that efficiency scores represent relative positions within the Italian national park system. They should not be interpreted as absolute measures of environmental or institutional quality, but rather as distances from observed best practice within each mission-specific technological space.

Multidimensional efficiency profiles for each park are reported in Figure A1 (Appendix). The radar representation highlights the domain-specific nature of performance and confirms that efficiency leadership in one mission does not systematically translate into dominance across other production spaces.

4.2. Scale effects: VRS versus CRS

To assess the role of scale heterogeneity, efficiency scores were recomputed under constant returns to scale (CRS). Under CRS, the production technology imposes proportional scaling between inputs and outputs, thereby combining pure technical inefficiency with scale inefficiency.

As reported in Table 5, ecological, tourism, and management domains exhibit moderate reductions in mean efficiency under CRS relative to VRS. These differences indicate the presence of scale distortions but do not fundamentally alter relative performance ordering.

This stability is formally verified through the Spearman rank correlations reported in Table 5. The correlations between VRS and CRS efficiency scores equal 0.751 for ecological performance, 0.868 for tourism, and 0.777 for management. These positive and substantial correlations indicate that comparative benchmarking in these domains is not structurally driven by the scale assumption.

Ecological performance appears comparatively robust to scale adjustments. The modest divergence between VRS and CRS reinforces the interpretation that ecological outcomes are largely determined by territorial characteristics rather than by scale misalignment.

Tourism and management domains display more visible CRS–VRS divergence, indicating the coexistence of slack-driven and scale-driven inefficiencies. Some parks operate efficiently at their observed scale

but would not be efficient under proportional scaling, suggesting that institutional size and resource configuration contribute to observed performance gaps.

The governance dimension exhibits a markedly different pattern. Under CRS, mean efficiency declines sharply from 0.749 to 0.498, and the share of efficient parks falls from 52 % (14 out of 27) under VRS to 18.5 % (5 out of 27) under CRS. The Spearman rank correlation between VRS and CRS governance scores is low ($\rho \approx 0.126$), indicating substantial reordering when scale proportionality is imposed.

This divergence confirms that governance inefficiency is predominantly scale-related rather than purely slack-driven. When parks are evaluated relative to peers operating at comparable institutional scale (VRS), measured slack is limited. Once proportional scaling is imposed (CRS), structural heterogeneity in territorial complexity, administrative burden, and coordination requirements becomes visible. Governance production, therefore, appears shaped primarily by institutional scale configuration rather than marginal input misallocation.

Fig. 4 visually illustrates the comparison between variable- and constant-returns-to-scale specifications across mission domains. In the ecological, tourism, and management dimensions, most parks cluster close to the 45-degree line, indicating moderate scale distortions and relatively stable rankings. By contrast, the governance panel reveals substantial vertical dispersion, with numerous parks exhibiting high VRS efficiency but markedly lower CRS scores. This graphical evidence reinforces the interpretation that governance inefficiency is predominantly scale-related rather than slack-driven.

4.3. Determinants of inefficiency

The second-stage Simar–Wilson bootstrap truncated regression results are reported in Table 6. Inefficiency is defined as the bias-corrected transformation

$$\delta_i = 1 / (\hat{\rho}_i^{bc} - 1)$$

so that higher values correspond to greater distance from the frontier.

In the management domain, planning maturity is negatively and significantly associated with inefficiency. Parks with more recently approved management plans operate closer to the frontier, suggesting that institutional planning coherence enhances operational alignment. Governance fragmentation is positively associated with managerial inefficiency, consistent with coordination-cost arguments in multi-level governance contexts. Territorial extension does not display a statistically significant effect, indicating that managerial inefficiency is not mechanically driven by size.

In the tourism domain, park age is positively associated with inefficiency, potentially reflecting institutional inertia or legacy infrastructure mismatches. Branding status and accessibility indicators do not exhibit robust statistical associations, suggesting that structural characteristics may matter less than activation strategies.

In governance, planning maturity shows a directionally consistent association with lower inefficiency, although overall explanatory power remains limited. Given the strong scale sensitivity documented in the CRS comparison, governance inefficiency appears primarily structural rather than driven by observable managerial covariates.

In the ecological domain, institutional variables - including park age, UNESCO Biosphere designation, and LIFE participation - do not display statistically significant associations with inefficiency. This absence of robust institutional determinants reinforces the interpretation that ecological performance is predominantly shaped by territorial and environmental endowment within the observed cross-section.

4.4. Synthesis of multidimensional patterns

Taken together, the results reveal systematic and economically interpretable cross-domain heterogeneity.

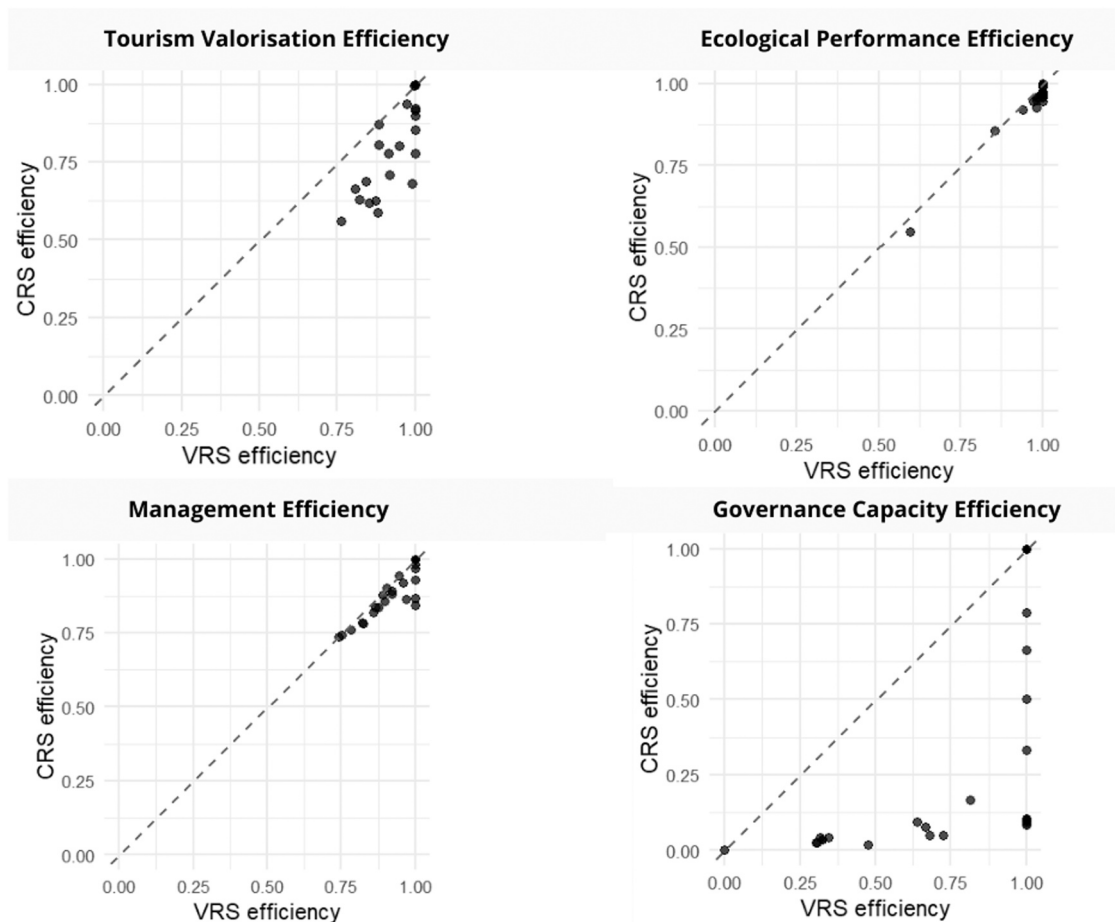


Fig. 4. VRS–CRS efficiency comparison across mission domains. Notes: VRS and CRS efficiency scores for each park and mission domain. The dashed 45-degree line represents equality between VRS and CRS efficiency. Observations below the diagonal indicate the presence of scale inefficiency. Each panel corresponds to one mission-specific production frontier estimated using the non-oriented SBM specification.

First, efficiency is mission-specific. Parks performing strongly in ecological transformation are not necessarily leaders in tourism, management, or governance, underscoring the need for mission-specific frontier estimation.

Second, scale effects are domain-dependent. Ecological efficiency appears relatively scale-robust. Tourism and management combine slack-driven and scale-driven inefficiencies. Governance efficiency is structurally dominated by scale heterogeneity.

Third, institutional planning maturity emerges as the most consistent correlate of reduced inefficiency in operational domains, while ecological performance remains largely endowment-driven.

The robustness diagnostics presented in the following section verify that these patterns are structurally stable and not artifacts of influential observations, dimensionality distortions, scale specification, or bootstrap variability.

5. Robustness and stability analysis

Building on the robustness strategy outlined in Section 3.9, this section reports the empirical diagnostics implemented to assess the structural stability of the estimated frontiers and the credibility of second-stage inference. Given the moderate sample size and the multi-dimensional specification of the production technologies, particular attention is devoted to frontier stability, dimensionality considerations, scale sensitivity, and bootstrap convergence. The following subsections present the results of these complementary diagnostics.

5.1. Leave-one-out frontier stability

The first diagnostic examines whether the empirical frontier is driven by a small number of extreme observations. In nonparametric models, units located at the boundary of the production set serve as reference points for the entire sample. If such units exert disproportionate leverage, efficiency rankings may be artificially distorted.

To evaluate this issue, a leave-one-out (LOO) procedure was conducted for each mission-specific SBM model under the baseline VRS specification. Each park was sequentially excluded from the sample, and efficiency scores were recomputed for the remaining units. For each replication, we calculated the Spearman rank correlation between baseline and LOO efficiency scores, the mean absolute change in efficiency, and the maximum absolute change in efficiency. These indicators allow a direct assessment of the sensitivity of frontier construction and ranking stability to potentially influential observations.

This procedure directly tests whether the frontier is structurally anchored to specific benchmark parks.

Table 7 reports the leave-one-out (LOO) stability diagnostics for the SBM–VRS efficiency scores across the four mission-specific production technologies. The table reports the minimum, median, and maximum Spearman rank correlations between baseline and leave-one-out rankings, as well as the mean and maximum absolute deviations in efficiency scores. High rank correlations and limited deviations indicate structural stability of the estimated frontiers. The procedure sequentially excludes each park from the sample and recomputes efficiency scores for the remaining units, thereby allowing an assessment of the frontier's

Table 6

Simar–Wilson bootstrap truncated regression results (bias-corrected SBM inefficiency).

Variable	Coef.	Boot. SE	95% CI (low)	95% CI (high)
Panel A. Governance				
(Intercept)*	1.287	125.511	0.520	449.373
Planning maturity (piano_c)*	−0.059	6.448	−0.243	−0.012
Governance fragmentation (enti_c)	0.486	225.784	−11.949	768.079
Center	1.571	291.754	−249.538	840.410
South	−0.465	95.649	−317.232	0.507
Islands	−0.527	87.231	−311.184	0.153
sigma	0.295	0.114	0.109	0.405
Panel B. Tourism				
(Intercept)*	1.151	11.072	0.971	8.410
Park age (eta_c)*	0.012	0.454	0.003	0.302
Park branding (marchio parco)	−0.003	10.054	−0.299	1.369
Accessibility (acc_c)	1.74e−05	0.001	−5.97e−06	0.000
sigma	0.097	0.198	0.029	0.501
Panel C. Management				
(Intercept)*	1.170	29.061	0.956	90.310
Territorial extension (superf_c)	−0.143	2.377	−2.005	0.503
Planning maturity (piano_c)*	−0.025	2.040	−0.881	−0.007
Governance fragmentation (enti_c)*	0.469	57.764	0.140	200.755
sigma	0.097	0.227	0.045	0.654
Panel D. Ecological				
(Intercept)	2.434	5.209	−7.830	15.594
Park age (eta parco)	0.210	0.243	−0.016	0.757
UNESCO Biosphere MAB (biosphere_MAB)	−4.190	12.237	−15.367	0.983
LIFE participation (Life_1)	4.621	3.902	−0.530	12.564
sigma	0.827	0.310	0.022	1.155

Notes: The dependent variable is the bias-corrected inefficiency measure, Boot. SE and confidence intervals are obtained via bootstrap (Simar–Wilson Algorithm 2). * indicates 95% bootstrap confidence interval excludes zero.

Table 7

Leave-one-out stability diagnostics for SBM–VRS efficiency scores.

Mission Domain	Spearman ρ (min)	Spearman ρ (median)	Spearman ρ (max)	Mean $ \Delta $	Max $ \Delta $
Ecological	0.75	1.00	1.00	0.00	0.44
Touristic	0.97	1.00	1.00	0.00	0.22
Management	0.97	1.00	1.00	0.00	0.32
Governance	0.96	1.00	1.00	0.00	0.50

Notes: Leave-one-out (LOO) diagnostics are computed under the baseline SBM–VRS specification. In each replication, one park is excluded, and efficiency scores are recomputed for the remaining $N-1$ units ($N = 27$). Spearman's rank correlation refers to the association between baseline and LOO efficiency rankings. Mean $|\Delta|$ and Max $|\Delta|$ denote, respectively, the average and maximum absolute differences between baseline and LOO efficiency scores. Values rounded to two decimal places; mean deviations smaller than 0.005 appear as 0.00 due to rounding.

sensitivity to potentially influential observations.

Across all mission domains, ranking stability proves remarkably high, indicating that frontier construction is not driven by isolated or extreme observations.

In the management dimension, the minimum Spearman rank correlation across replications equals 0.97, with a median equal to one. Mean absolute deviations are negligible, and even the maximum deviation remains limited relative to the dispersion of baseline scores. These results indicate that the managerial frontier is not anchored to a single benchmark park but reflects a structurally distributed configuration of peers.

The tourism model exhibits even stronger robustness. The minimum

rank correlation reaches 0.97, while both mean and median absolute deviations are effectively zero. The tourism frontier is therefore highly stable and does not depend on dominant reference units.

The governance dimension, despite its pronounced scale sensitivity under CRS discussed later, remains structurally stable under VRS. The minimum correlation ($\rho = 0.96$) and median equal to one confirm that frontier construction is not driven by institutional outliers. Although a single replication yields a relatively large absolute deviation, this effect remains localized and does not alter the overall ranking structure. Governance scale effects identified in subsequent sections, therefore, reflect genuine technological structure rather than frontier instability.

The ecological model shows slightly greater sensitivity, with a minimum correlation of 0.75. Nevertheless, mean deviations remain small, and no systematic ranking reversals occur. This moderate sensitivity is economically interpretable: parks characterized by particularly large territorial endowments or distinctive ecological profiles may serve as strong reference units. Importantly, their exclusion does not fundamentally modify the ranking distribution.

Taken together, the leave-one-out diagnostics provide strong evidence that none of the mission-specific frontiers is artificially constructed around extreme or dominant units. The empirical frontier in each domain is structurally supported by multiple benchmark parks.

Consequently, the multidimensional efficiency patterns documented in the baseline specification cannot be attributed to influential-observation bias. They reflect stable technological configurations within each mission-specific production space.

5.2. Distribution of efficient units and frontier discriminatory power

To further assess discriminatory power and potential dimensionality effects, Table 8 reports structural diagnostic indicators for each mission-specific SBM–VRS frontier, including the number of inputs and outputs, total dimensionality ($m + s$), the proportionality ratio $N/(m+s)$, and the number and share of fully efficient units.

Across all domains, the proportionality condition for nonparametric frontier estimation is satisfied. The ratio $N/(m+s)$ ranges from 4.5 (ecological model) to 9.0 (governance model), exceeding conventional thresholds used to mitigate dimensionality-induced frontier inflation. This confirms that each specification retains adequate discriminatory capacity given the sample size ($N = 27$).

The ecological frontier exhibits the highest share of efficient units, with 16 out of 27 parks (59.3%) classified as fully efficient under VRS. In moderate samples, such concentration may raise concerns regarding potential dimensionality effects. However, the dimensional structure of the ecological model ($m = 3$; $s = 3$) implies $m + s = 6$ and therefore $3(m+s) = 18$. With $N = 27$, the proportionality condition $N \geq 3(m+s)$ is satisfied, yielding a dimensionality ratio $N/(m+s) = 4.5$. Moreover, leave-one-out diagnostics presented in Section 5.1 confirm strong

Table 8

Frontier Diagnostic Indicators (SBM–VRS) Diagnostic indicators of dimensional structure and discriminatory capacity for each mission-specific frontier estimated using the non-oriented Slack-Based Measure (SBM) under variable returns to scale (VRS).

Mission Domain	N	m	s	$m + s$	$N/(m+s)$	Efficient Units (VRS)	Share Efficient
Ecological	27	3	3	6	4.50	16	0.593
Tourism	27	3	2	5	5.40	12	0.444
Management	27	2	2	4	6.75	11	0.407
Governance	27	1	2	3	9.00	14	0.519

Notes: N denotes the number of decision-making units (parks), m the number of inputs, s the number of outputs, and $(m + s)$ total dimensionality. The ratio $N/(m+s)$ assesses proportionality conditions for nonparametric frontier estimation. “Efficient Units” refers to parks with an efficiency score equal to 1 under VRS.

ranking stability. The concentration of efficient units, therefore, reflects structural similarity in ecological endowment and environmental performance across parks rather than estimator distortion or frontier inflation.

Tourism and management display lower shares of frontier units (44.4% and 40.7%, respectively), consistent with greater dispersion in operational performance and a stronger role for managerial heterogeneity. These proportions fall comfortably within ranges typically observed in moderate-dimensional DEA applications and indicate satisfactory discriminatory power.

The governance frontier presents a distinct configuration. Under VRS, 51.9% of parks are fully efficient, suggesting limited pure technical slack when units are evaluated relative to peers operating at comparable institutional scale. However, as demonstrated in Section 4.2, this concentration is largely absorbed under CRS, where the share of efficient units declines sharply. This divergence confirms that governance inefficiency is predominantly scale-driven rather than a consequence of frontier saturation.

Taken together, the distributional diagnostics indicate that none of the mission-specific frontiers suffers from dimensionality failure. Variations in the share of efficient units reflect structural differences in technological dispersion across domains rather than artifacts of sample size or model specification.

5.3. Sensitivity to returns-to-scale assumptions

To further assess robustness, efficiency rankings under VRS were systematically compared with those obtained under CRS.

For the tourism, ecological, and management dimensions, Spearman rank correlations between VRS and CRS scores remain positive and substantial (approximately 0.868, 0.751, and 0.777, respectively). Although mean efficiency levels decline under CRS - indicating the presence of scale distortions - the ordering of parks remains broadly preserved. Relative performance comparisons are therefore largely invariant to scale assumptions in these domains.

The governance dimension exhibits a markedly different pattern. The Spearman rank correlation between VRS and CRS governance scores is approximately 0.126. Rankings change substantially when proportional scaling is imposed. This divergence confirms that governance inefficiency is structurally dominated by scale effects.

Importantly, this result does not undermine the credibility of the governance frontier. Rather, it reveals a distinctive technological structure. The VRS specification isolates pure technical inefficiency at the observed scale, while CRS exposes scale distortions embedded in the institutional configuration. The divergence between the two models, therefore, provides complementary insight into governance production.

5.4. Bootstrap stability of second-stage inference

The final robustness layer concerns the numerical stability of the second-stage Simar–Wilson bootstrap truncated regression. Because inference relies on resampling, estimated standard errors and confidence intervals may, in principle, depend on the number of bootstrap replications. Ensuring convergence of inferential results across alternative replication sizes is therefore essential.

To assess numerical stability, each second-stage model was re-estimated using 1000, 2000, and 5000 bootstrap replications. Across all mission domains, coefficient estimates remain invariant up to the third decimal place. Bootstrap standard errors exhibit only minor fluctuations as the number of replications increases, and percentile confidence intervals display negligible numerical variation. Importantly, no coefficient changes sign or statistical significance across replication sizes.

In the management and tourism domains, the statistically significant associations identified in the baseline specification remain robust across all bootstrap intensities. Confidence intervals remain stable and well-

defined, reinforcing the structural credibility of these relationships. In the ecological and governance models, where explanatory power is comparatively limited, the absence of statistically significant associations remains stable across replication sizes. This invariance indicates that weak relationships reflect structural characteristics of the data rather than insufficient resampling precision.

Overall, the bootstrap diagnostics suggest numerical convergence of the second-stage estimates and confirm that the reported inferential results are not driven by resampling variability. Detailed replication tables for alternative bootstrap sizes are available from the authors upon request.

5.5. Overall assessment of robustness

The combined diagnostics provide strong evidence supporting the credibility of the multidimensional efficiency assessment.

First, frontier construction is not driven by influential units. Second, the distribution of efficient parks remains within acceptable bounds across domains. In the governance dimension, efficiency under VRS exhibits a strong concentration near the frontier, yet without full saturation. This pattern is structurally interpretable and does not indicate frontier inflation. Third, relative rankings are largely invariant to scale assumptions in ecological, tourism, and management domains, while governance efficiency exhibits pronounced scale sensitivity consistent with its institutional configuration. Fourth, bootstrap inference in the second stage proves numerically stable across resampling intensities.

The multidimensional efficiency patterns documented in the baseline specification therefore represent systematic, domain-specific performance differences rather than artifacts of modeling choices, dimensionality constraints, or numerical instability.

6. Discussion

The multidimensional efficiency assessment reveals systematic heterogeneity across mission domains in Italian national parks. Consistent with their characterization as multi-mission public production units, performance does not conform to a unified technological logic. Ecological transformation, tourism valorization, managerial administration, and governance capacity operate within partially differentiated production spaces. The results, therefore, support a production-theoretic interpretation of protected areas as institutions embedded in heterogeneous technological regimes rather than a single aggregated performance framework.

Ecological performance emerges as comparatively homogeneous once the territorial surface and environmental endowment are incorporated into the production set. Efficiency dispersion is limited in relative terms, although the frontier exhibits a pronounced concentration of parks operating at full efficiency under VRS. Differences between variable- and constant-returns-to-scale specifications remain limited, indicating that ecological inefficiency - when present - reflects relatively small deviations from the frontier. Within the Italian cross-section, ecological outcomes appear primarily anchored in territorial and environmental characteristics, with management-mediated accessibility (e.g., trail networks) playing a secondary role in shaping dispersion.

Tourism valorization displays greater dispersion and a more articulated efficiency structure. While a substantial share of parks operates at or near the frontier under VRS, inefficiency in this domain reflects both output shortfalls and scale-related distortions. Relative rankings remain broadly preserved under CRS, suggesting that comparative positioning is not exclusively scale-driven. Tourism production appears sensitive to managerial activation, infrastructure configuration, and historical development trajectories. Unlike ecological performance, tourism efficiency is more responsive to strategic and organizational alignment.

Governance efficiency exhibits a distinctive configuration. Under VRS, mean governance efficiency equals 0.749, and approximately 52 % of parks are fully efficient, indicating a strong concentration near the

frontier but not complete saturation. Under CRS, mean efficiency declines markedly, and the share of efficient units falls to 12.5 %. The divergence between the two specifications confirms that governance inefficiency is predominantly scale-related rather than purely slack-driven. Governance outputs - such as management plans, certifications, and coordination mechanisms - are institutional and partly discrete. When parks are evaluated relative to peers operating at comparable administrative scale, measured slack is limited; once proportional scaling is imposed, structural heterogeneity in territorial complexity and multi-level coordination becomes visible. Governance performance appears structurally conditioned by institutional scale configuration and territorial complexity. This suggests that reforms aimed at improving governance efficiency may require adjustments in jurisdictional alignment and coordination structures rather than marginal increases in administrative inputs.

Management efficiency occupies an intermediate position. Inefficiency reflects both slack-driven and structural components, and rankings remain stable under alternative specifications. Managerial performance thus appears to be partially amenable to organizational reform yet also influenced by scale alignment and governance fragmentation.

The second-stage analysis provides additional nuance. Institutional planning maturity is consistently associated with lower inefficiency in management and, to a lesser extent, governance domains. Governance fragmentation correlates positively with managerial inefficiency, consistent with coordination-cost arguments in multi-level governance settings. In tourism, park age is positively associated with inefficiency, suggesting possible institutional inertia. By contrast, ecological inefficiency shows no systematic association with institutional covariates, reinforcing the interpretation that ecological performance is largely endowment-driven within the observed cross-section.

Taken together, the findings indicate that inefficiency in protected areas cannot be attributed to a single structural mechanism. Some domains are predominantly endowment-driven, others reflect managerial alignment, and others are primarily shaped by institutional scale heterogeneity. The multidimensional frontier framework, therefore, uncovers differentiated technological logics that would remain obscured under an aggregated production representation. From a broader ecological economics perspective, these findings reinforce the view that environmental public goods are produced within institutionally layered and scale-dependent governance systems. The multidimensional nature of protected-area performance reflects the coexistence of biophysical endowment constraints, organizational capacity, and institutional coordination costs. Efficiency, therefore, cannot be interpreted as a purely managerial attribute but emerges from the interaction between ecological scale, administrative design, and mission-specific public-good provision.

More broadly, the results suggest that public environmental institutions cannot be evaluated through single-dimensional efficiency paradigms. Multidimensional frontier analysis reveals that different missions follow distinct technological and institutional regimes, implying that performance benchmarking must be domain-contingent rather than globally aggregated.

6.1. Policy implications: scale alignment, jurisdictional design, and resource allocation

Although the following implications are discussed with reference to the Italian institutional framework, they are broadly applicable to protected-area systems operating under multi-level governance arrangements and jurisdictional fragmentation.

The empirical evidence on the structural role of institutional scale - particularly in the governance domain - carries direct implications for institutional design, incentive structures, and jurisdictional configuration. Three policy directions emerge from the multidimensional, scale-aware framework adopted in this study, each grounded in the

recognition that performance gaps in fragmented governance systems reflect structural transaction costs rather than coordination deficits amenable to marginal procedural adjustments.

First, the sharp decline in mean governance efficiency from 0.749 under VRS to 0.498 under CRS, combined with the statistically significant positive association between governance fragmentation and managerial inefficiency in the management domain (0.469*), indicates that jurisdictional overlap generates structural institutional friction - not merely insufficient communication among agencies. Within the analytical framework of transaction cost economics, this friction represents a measurable cost of operating across misaligned administrative boundaries: each additional layer of jurisdictional involvement increases the costs of negotiation, monitoring, and enforcement associated with collective decision-making. In such settings, adding generic coordination layers - standing committees, memoranda of understanding, joint consultation forums - without altering the underlying incentive structures is unlikely to produce lasting efficiency gains, because sectionalism and jurisdictional protectionism remain rational behaviors as long as the boundary structures and property rights over territorial governance remain unchanged. The policy prescription, therefore, is not more coordination but rather incentive realignment: institutional mechanisms that make it individually rational for each administrative actor to internalize the costs currently externalized onto neighboring jurisdictions. In the Italian context, the Environmental Economic Zones (Zone Economiche Ambientali, ZEA) introduced by Law 141/2019 (implementing art. 7 of Law 394/1991) represent precisely this type of hard incentive instrument. Unlike traditional coordination mechanisms, the ZEA framework creates differentiated fiscal regimes tied to territorial boundaries, thereby altering the economic calculus of actors operating within and around protected areas. The present results suggest that the ZEA model could be extended in two concrete directions to address the scale-driven inefficiencies documented in this study. First, fiscal incentives within ZEA could be made conditional on mission-specific efficiency benchmarks derived from frontier analysis, rewarding parks (and the municipalities within their perimeters) that demonstrably reduce slack in governance or managerial domains, rather than distributing benefits based on aggregate indicators. Second, ZEA fiscal privileges could be modulated by the degree of inter-jurisdictional integration: municipalities participating in formalized cross-boundary governance arrangements - such as joint management plans, shared monitoring systems, or unified visitor management protocols - would receive enhanced fiscal benefits, thereby creating a direct pecuniary incentive for institutional convergence rather than relying on voluntary cooperation.

Second, jurisdictional design should be reconsidered where persistent scale-driven inefficiency signals a fundamental misalignment between ecological systems and administrative boundaries. The finding that ecological efficiency is largely endowment-driven, whereas governance efficiency is severely penalized by scale misalignment under CRS, implies that the two domains respond to qualitatively different institutional logics. Where this divergence is persistent, marginal managerial adjustments are structurally inadequate: what is required is a hard realignment of functional jurisdictional boundaries. In operational terms, this may involve the transfer of specific governance competencies - planning approval, environmental monitoring, visitor management - from fragmented municipal or regional authorities to the park authority or to purpose-built inter-park bodies operating at the ecological rather than the administrative scale. The critical insight from the empirical evidence is that such realignment need not be uniform across the entire park system: mission-specific diagnostics allow policymakers to identify which domains are structurally penalized by scale misalignment and to target jurisdictional reforms accordingly, avoiding costly and politically contentious system-wide restructuring in favor of domain-specific functional realignments.

Third, resource allocation mechanisms should be redesigned to reflect the heterogeneous responsiveness of mission domains to inputs.

The finding that ecological efficiency is largely driven by territorial endowment implies that additional financial resources alone are unlikely to yield substantial short-term improvements in ecological outcomes. By contrast, managerial and governance domains appear more responsive to institutional factors such as planning maturity, administrative stability, and reduced inter-jurisdictional transaction costs. This differential responsiveness argues against uniform budget increases and in favor of incentive-compatible allocation formulae that channel resources toward the institutional determinants of inefficiency rather than toward the symptoms. Concretely, funding schemes should incorporate mission-specific and scale-adjusted performance indicators, replacing aggregate metrics that conflate endowment conditions, managerial performance, and structural scale effects. The ZEA framework again provides a natural vehicle for this differentiation: by tying area-specific fiscal incentives to measurable mission-level efficiency gains, the instrument can function as a performance-based transfer mechanism that rewards institutional reform rather than merely compensating for resource scarcity.

Taken together, these implications point toward a governance-sensitive model of protected-area policy design. The evidence presented in this study suggests that efficiency gains in multi-mission public organizations are unlikely to emerge from increased funding or from the addition of generic coordination mechanisms alone. Instead, they require reengineering at an institutional scale and of incentive structures: aligning jurisdictional boundaries with ecological rather than administrative logics, embedding hard fiscal incentives into the governance architecture, and targeting institutional reforms to the specific mission domains where scale misalignment generates measurable transaction costs. The shift from resource-based to incentive-based policy design represents, in the authors' view, the most actionable contribution of the evidence on multidimensional efficiency to the ongoing reform of protected-area governance systems.

Interpretation must nonetheless remain bounded by the cross-sectional and nonparametric nature of the estimation strategy. Efficiency is measured relative to the observed Italian frontier rather than an external benchmark. Although dimensionality was carefully controlled and bootstrap procedures ensure inferential consistency, finite-sample properties remain inherent to DEA. Moreover, the second-stage regression identifies conditional associations rather than causal mechanisms. Within these limits, however, the evidence provides a coherent representation of mission-specific efficiency patterns in public environmental governance.

7. Conclusions

This study provides a multidimensional assessment of 27 Italian National Parks, conceptualized as multi-mission public organizations embedded in territorially differentiated governance systems. By evaluating ecological performance, tourism valorization, managerial administration, and governance capacity as distinct mission domains, the analysis demonstrates that protected-area performance is structurally multidimensional and cannot be meaningfully summarized through aggregated indicators.

Three central insights emerge.

First, performance is mission-contingent rather than uniform. Ecological, tourism, managerial, and governance domains operate under partially differentiated technological and institutional logics. Parks that perform strongly in ecological outcomes are not systematically leaders in tourism activation, administrative efficiency, or governance capacity. This heterogeneity implies that policy evaluation and benchmarking should be domain-specific rather than based on composite rankings. Aggregated indicators risk obscuring mission trade-offs and masking structural strengths or weaknesses within specific functional areas.

Second, institutional scale plays a differentiated and structurally significant role. Ecological performance appears relatively robust to scale assumptions, suggesting that biophysical endowment and

territorial characteristics largely anchor environmental outcomes within the observed cross-section. In contrast, governance capacity is highly sensitive to the configuration of institutional scales. When proportional scaling is imposed, substantial reordering of performance emerges. This indicates that governance inefficiency is predominantly scale-related rather than slack-driven. This finding underscores the importance of jurisdictional alignment, inter-regional coordination, and administrative architecture in shaping governance performance. In multi-level territorial systems, performance gaps may reflect structural scale misalignment rather than managerial underperformance.

Third, planning maturity consistently correlates with lower inefficiency in operational domains, particularly management and governance. Parks with more consolidated planning instruments operate closer to the frontier, suggesting that institutional coherence and procedural stability contribute to improved resource transformation. By contrast, ecological performance remains largely endowment-driven within the Italian context, reinforcing the interpretation that environmental outcomes are more tightly linked to territorial characteristics than to short-run managerial variation.

Taken together, these results suggest that improving protected-area performance requires differentiated policy strategies. Ecological outcomes may primarily benefit from long-term conservation and territorial planning, whereas governance and management domains may respond more directly to institutional reform, coordination mechanisms, and capacity-building interventions. Efficiency improvements in governance, in particular, appear unlikely to arise solely from additional inputs; instead, they may require adjustments in institutional scale and intergovernmental coordination structures.

More broadly, the findings indicate that performance assessment in protected-area systems should integrate multidimensional and scale-aware perspectives. Treating parks as homogeneous conservation entities may obscure the interaction between territorial complexity, institutional design, and mission-specific delivery. A mission-specific benchmarking framework can therefore support more targeted policy design, allowing decision-makers to distinguish between endowment constraints, organizational slack, and structural scale effects.

The analysis is subject to the limitations inherent in cross-sectional nonparametric estimation. Efficiency is measured relative to the observed Italian frontier rather than an external benchmark, and second-stage associations are interpreted as conditional rather than causal relationships. Nevertheless, the robustness diagnostics confirm that the multidimensional patterns identified are structurally stable within the national system.

For the Italian national park system, three concrete policy directions follow from the analysis. First, the statistically significant association between governance fragmentation and managerial inefficiency should be interpreted as a structural transaction cost arising from jurisdictional overlap, not as a coordination deficit that can be solved through additional consultation mechanisms. Policy responses should therefore focus on incentive realignment rather than procedural layering. The Environmental Economic Zones (Law 141/2019) provide a concrete instrument for embedding hard fiscal incentives into the governance architecture, conditioning area-specific fiscal benefits on measurable mission-level efficiency improvements and on the degree of inter-jurisdictional integration. Second, where persistent scale-driven inefficiency signals a fundamental misalignment between ecological systems and administrative boundaries, the policy response should involve functional jurisdictional realignment - transferring specific governance competencies to entities operating at the ecological rather than the administrative scale - guided by mission-specific diagnostics that identify which domains are structurally penalized by scale misalignment. Third, resource allocation should shift from uniform budget increases toward incentive-compatible formulae that channel resources toward the institutional determinants of inefficiency, incorporating mission-specific and scale-adjusted performance indicators rather than aggregate metrics that conflate endowment, managerial slack, and structural

scale effects.

Future research could extend this framework by using longitudinal data to assess dynamic performance trajectories, incorporating spatial interdependencies among neighboring parks, and exploring institutional reforms that alter jurisdictional scale. Such extensions would further clarify how territorial governance architectures shape the transformation of public resources into environmental and institutional outcomes over time.

CRedit authorship contribution statement

Maria De Salvo: Visualization, Validation, Supervision. **Laura**

Giuffrida: Validation, Investigation, Data curation. **Marika Cerro:** Methodology, Investigation, Formal analysis, Data curation. **Giovanni Signorello:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of Competing Interest

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

Appendix

Table A1

Complete SBM Efficiency Scores under VRS and CRS Specifications

Park	ECO VRS	ECO CRS	GOV VRS	GOV CRS	TOUR VRS	TOUR CRS	MGT VRS	MGT CRS
Stelvio National Park	1.0	0.991	1.0	0.1	1.0	0.778	1.0	0.842
Gran Paradiso National Park	1.0	0.974	1.0	0.099	1.0	0.921	1.0	0.929
Abruzzo, Lazio and Molise NP	1.0	0.957	1.0	0.092	1.0	0.917	1.0	1.0
Circeo National Park	1.0	0.958	1.0	1.0	1.0	0.898	1.0	0.983
Dolomiti Bellunesi NP	1.0	0.946	1.0	0.5	1.0	1.0	1.0	1.0
Monti Sibillini NP	1.0	1.0	1.0	0.082	1.0	1.0	1.0	1.0
Tuscan Archipelago NP	1.0	1.0	1.0	0.105	1.0	1.0	1.0	0.966
Pollino National Park	1.0	0.975	1.0	0.333	1.0	1.0	1.0	1.0
Cilento and Vallo di Diano NP	1.0	1.0	1.0	0.664	1.0	1.0	1.0	0.869
Gargano National Park	1.0	1.0	1.0	1.0	1.0	0.852	1.0	1.0
Gran Sasso and Monti della Laga NP	1.0	0.967	1.0	0.787	1.0	1.0	1.0	1.0
Val Grande National Park	1.0	1.0	1.0	1.0	1.0	1.0	0.969	0.865
Vesuvius National Park	1.0	1.0	1.0	1.0	0.996	0.996	0.958	0.92
Casentinesi Forests NP	1.0	0.966	1.0	1.0	0.99	0.681	0.944	0.944
Aspromonte National Park	1.0	0.991	0.813	0.167	0.972	0.937	0.92	0.891
Maiella National Park	1.0	1.0	0.726	0.05	0.95	0.8	0.92	0.883
Asinara National Park	0.998	0.97	0.679	0.047	0.917	0.707	0.902	0.902
La Maddalena Archipelago NP	0.997	0.968	0.665	0.075	0.915	0.779	0.896	0.857
Tuscan-Emilian Apennines NP	0.995	0.965	0.638	0.095	0.884	0.87	0.889	0.879
Cinque Terre National Park	0.985	0.957	0.477	0.019	0.882	0.806	0.875	0.838
Sila National Park	0.982	0.948	0.343	0.042	0.879	0.588	0.866	0.837
Lucanian Apennines NP	0.982	0.927	0.325	0.033	0.873	0.624	0.857	0.819
Alta Murgia National Park	0.981	0.957	0.325	0.033	0.85	0.618	0.825	0.782
Pantelleria Island NP	0.971	0.944	0.317	0.043	0.843	0.688	0.822	0.785
Nebrodi Park	0.941	0.921	0.307	0.025	0.821	0.629	0.784	0.761
Etna Park	0.855	0.855	0.303	0.024	0.808	0.664	0.751	0.744
Madonie Park	0.597	0.546	0.0	0.0	0.762	0.558	0.743	0.734

Notes: Efficiency scores are estimated using the non-oriented Slack-Based Measure (SBM). Values range from 0–1, where 1 denotes full technical efficiency within each mission-specific frontier.

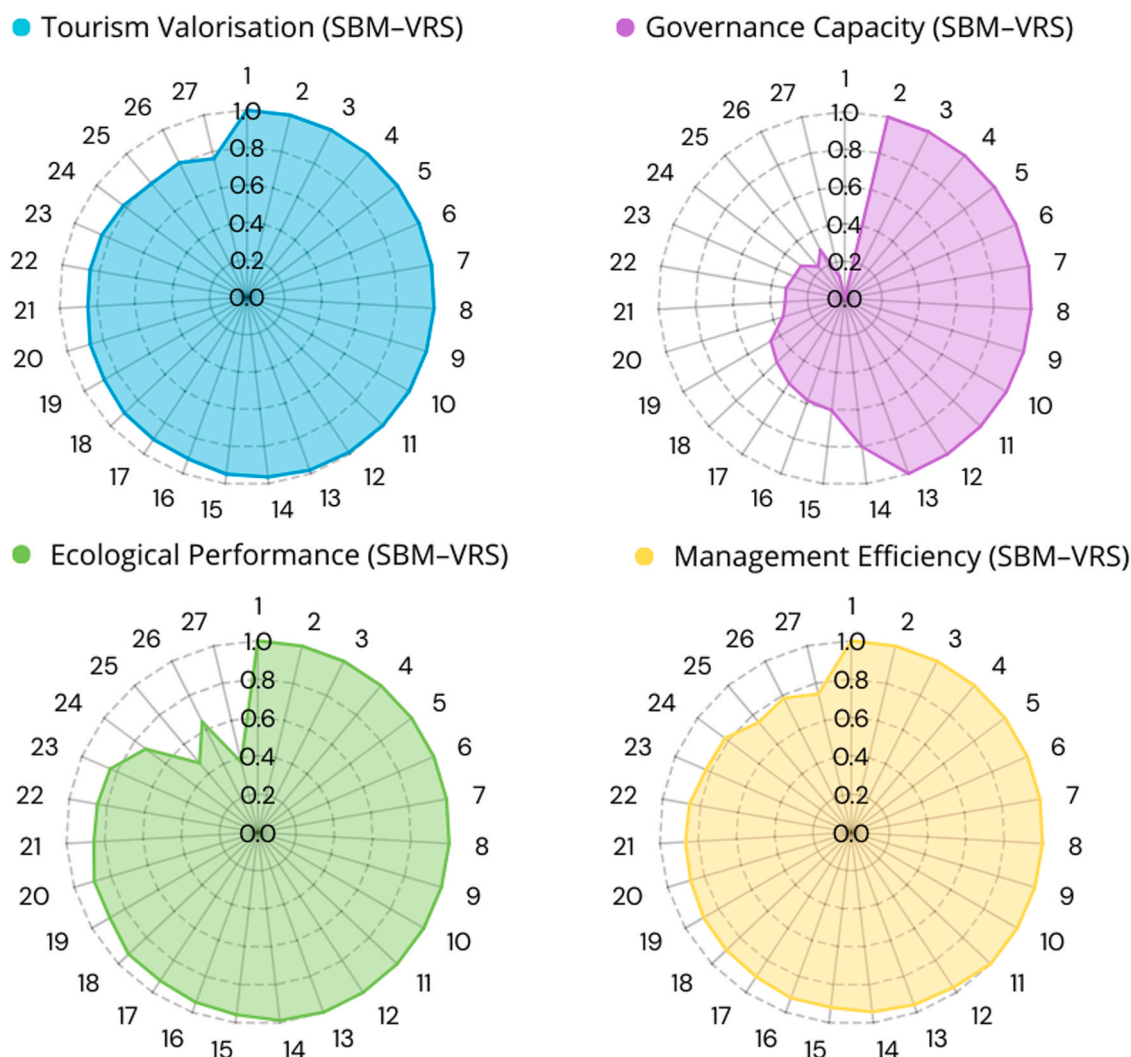


Figure A1. Multidimensional SBM-VRS efficiency profiles by park. Notes: Radar plots display park-level efficiency scores under the SBM-VRS specification for each mission domain. Scores are plotted on a 0–1 scale, where 1 denotes frontier performance. The figure highlights multidimensional heterogeneity across parks and the absence of uniform dominance across domains

Data availability

Data will be made available on request.

References

- Avelino, J.E., Sasaki, J., Esteban, M., Salah, P., Jamero, M.L., Valenzuela, V.P., 2019. Sustainability Evaluation of Marine Protected Areas Index (SEMPAI): a multi-criteria decision-making method to determine the effectiveness of the El Nido-Taytay Managed Resource Protected Area. *Ocean. & Coast. Manag.* <https://doi.org/10.1016/j.ocecoaman.2019.104891>.
- Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models to estimate technical and scale inefficiencies in data envelope analysis. *Manag. Sci.* 30 (9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>.
- Benito, B., del Carmen Conesa-Pérez, M., Parra-Merono, M.C., 2025. Determinants of inefficiency in the provision of public parks and gardens services. *Cities* 163, 106031. <https://doi.org/10.1016/j.cities.2025.106031>.
- Bianchi, M., del Valle, I., Tapia, C., 2020. Measuring eco-efficiency in European regions: evidence from a territorial perspective. *J. Clean. Prod.* 276, 123246. <https://doi.org/10.1016/j.jclepro.2020.123246>.
- Bogetoft, P., Otto, L., 2010. Benchmark DEA SFA R. 157. <https://doi.org/10.1007/978-1-4419-7961-2>.
- Borrini-Feyerabend, G., Dudley, N., Jaeger, T., Lassen, B., Pathak Broome, N., Phillips, A., Sandwith, T., 2013. Governance of protected areas: from understanding to action. *Best Practice Protected Area Guidelines Series*. IUCN, Gland, Switzerland.
- Bosetti, V., Locatelli, G., 2006. A data envelopment analysis approach to the assessment of natural parks' economic efficiency and sustainability. The case of Italian national parks. *Sustain. Dev.* 14 (4), 277–286. <https://doi.org/10.1002/sd.288>.
- Buckley, R., 2009. Assessing the net effects of ecotourism on the environment: a framework, a first assessment and future research. *J. Sustain. Tour.* 17 (6), 643–672. <https://doi.org/10.1080/09669580902999188>.
- Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* 2 (6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8).
- Coad, L., Leverington, F., Knights, K., Geldmann, J., Eassom, A., Kapos, V., Kingston, N., de Lima, M., Zamora, C., Cuadros, I., Nolte, C., Burgess, N.D., Hockings, M., 2015. Measuring impact of protected area management interventions: current and future use of the global database of protected area management effectiveness. *Philos. Trans. R. Soc. B Biol. Sci.* 370 (1681), 20140281. <https://doi.org/10.1098/rstb.2014.0281>.
- Cracolici, M.F., Nijkamp, P., Cuffaro, M., 2007. Efficiency and productivity of Italian tourist destinations: a quantitative estimation based on data envelopment analysis and the malmquist method. In: Matias, In. A., Nijkamp, P., Neto, P. (Eds.), *Advances in Modern Tourism Research*. Physica-Verlag, Heidelberg, pp. 325–343. https://doi.org/10.1007/978-3-7908-1718-8_17.
- Daraio, C., Simar, L., 2007. Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *J. Product. Anal.* 28 (1), 13–32.
- Dubois, Grégoire, Mandrici, Andrea, Delli, Giacomo, Battistella, Luca, García Bendito, Eduardo, 2019. Digital observatory for protected areas 4. Joint Research Centre (JRC) [Dataset]. European Commission. PID: <http://data.europa.eu/89h/b207a3ae-8b18-453b-8234-8cfd1d4c805>.
- Dyckhoff, H., Allen, K., 2001. Measuring ecological efficiency with data envelopment analysis (DEA). *Eur. J. Oper. Res.* 132 (2), 312–325. [https://doi.org/10.1016/S0377-2217\(00\)00154-5](https://doi.org/10.1016/S0377-2217(00)00154-5) Get rights and content.

- Eagles, P.F., McCool, S.F., Haynes, C.D., 2002. Sustainable Tourism in Protected Areas: Guidelines for Planning and Management (n. 8. IUCN).
- Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. *Production Frontiers*. Cambridge University Press.
- Farrell, M.J., 1957. The measurement of productive efficiency. *J. R. Stat. Soc. Ser. A Stat. Soc.* 120 (3), 253–281. <https://doi.org/10.2307/2343100>.
- Farrell, T.A., Marion, J.L., 2001. Trail impacts and trail impact management related to visitation at Torres del Paine National Park, Chile. *Leis. /Loisir* 26 (1–2), 31–59. <https://doi.org/10.1080/14927713.2001.9649928>.
- Geldmann, J., Coad, L., Barnes, M.D., Craigie, I.D., Woodley, S., Balmford, A., Brooks, T. M., Hockings, M., Knights, K., Mascia, M.B., McRae, L., Burgess, N.D., 2018. A global analysis of management capacity and ecological outcomes in terrestrial protected areas. *Conserv. Lett.* 11 (3), e12434. <https://doi.org/10.1111/conl.12434>.
- Hockings, M., 2006. *Effectiveness Assessment: A Framework for Assessing the Effectiveness of Protected Area Management*. IUCN.
- ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) 2023. Turismo nei parchi. <https://indicatoriambientali.isprambiente.it/it/turismo/turismo-nei-parchi>.
- ISTAT (Istituto Nazionale di Statistica), 2023. Presenze turistiche. <https://www.istat.it/tag/presenze/>.
- Jacob, D., Onadeko, S., Nelson, I., Shotuyo, A., 2018. Evaluation of old Oyo National Park efficiency using DEA approach. *Econ. Environ. Stud.* 18 (1), 193–210. <https://doi.org/10.25167/ees.2018.45.12>.
- Kneip, A., Park, B.U., Simar, L., 1998. A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econom. Theory* 14 (6), 783–793. <https://doi.org/10.1017/S0266466698146042>.
- Liu, J., Zhang, J., Fu, Z., 2017. Tourism eco-efficiency of Chinese coastal cities—Analysis based on the DEA-Tobit model. *Ocean. & Coast. Manag.* 148, 164–170. <https://doi.org/10.1016/j.ocecoaman.2017.08.003>.
- Lockwood, M., 2010. Good governance for terrestrial protected areas: a framework, principles and performance results. *J. Environ. Manag.* 91 (3), 754–766. <https://doi.org/10.1016/j.jenvman.2009.10.005>.
- Lynn, N.A., Brown, R.D., 2003. Effects of recreational use impacts on hiking experiences in natural areas. *Landsc. Urban. Plan.* 64 (1–2), 77–87. [https://doi.org/10.1016/S0169-2046\(02\)00202-5](https://doi.org/10.1016/S0169-2046(02)00202-5).
- Ma, X.L., Ryan, C., Bao, J.G., 2009. Chinese National parks-resource usage efficiencies, spatial proximity and roles: an application of data envelopment analysis. *Tourism in China*. Routledge, pp. 67–87.
- McDonald, J., 2009. Use of least squares and tobits in second-stage DEA efficiency analyzes. *Eur. J. Oper. Res.* 197 (2), 792–798. <https://doi.org/10.1016/j.ejor.2008.07.039>.
- Fried, H.O., Lovell, C.K., Schmidt, S.S. (Eds.), 2008. *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.
- Monz, C.A., Pickering, C.M., Hadwen, W.L., 2013. Recent advances in recreation ecology and the implications of different relationships between recreation use and ecological impacts. *Front. Ecol. Environ.* 11 (8), 441–446. <https://doi.org/10.1890/120358>.
- Munda, G., 2005. *Multi-criteria decision analysis and sustainable development. Multiple Criteria Decision Analysis: Avant-garde Investigations*. Springer, New York, NY, pp. 1235–1267 (New York. Portale dei parchi italiani. <https://www.parks.it>).
- Pascoe, S., Cannard, T., Dowling, N.A., Dichmont, C.M., Asche, F., Little, L.R., 2023. Use of Data Envelopment Analysis (DEA) to assess management alternatives in the presence of multiple objectives. *Mar. Policy* 148, 105444. <https://doi.org/10.1016/j.marpol.2022.105444>.
- Peng, H., Zhang, J., Lu, L., Tang, G., Yan, B., Xiao, X., Han, Y., 2017. Eco-efficiency and its determinants at a tourism destination: a case study of Huangshan National Park, China. *Tour. Manag.* 60, 201–211. <https://doi.org/10.1016/j.tourman.2016.12.005>.
- Rusielik, R., Zbaraszczyński, W., 2014. The efficiency of scientific and tourism activity of Polish National Parks with use DEA method. *Econ. Environ. Stud.* 14 (3 (31)), 283–297.
- Schulze, K., Knights, K., Coad, L., Geldmann, J., Leverington, F., Eassom, A., Marr, M., Butchart, S.H.M., Hockings, M., Burgess, N.D., 2018. An assessment of threats to terrestrial protected areas. *Conserv. Lett.* 11 (3), e12435. <https://doi.org/10.1111/conl.12435>.
- Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage semiparametric models of production processes. *J. Econ.* 136 (1), 31–64. <https://doi.org/10.1016/j.jeconom.2005.07.009>.
- Simar, L., Wilson, P.W., 2008. Statistical inference in nonparametric frontier models: recent developments and perspectives. In *The Measurement of Productive Efficiency and Productivity Growth* (Fried, Lovell, Schmidt, eds.), Oxford University Press.
- Simar, L., Wilson, P.W., 2011. Two-stage DEA: caveat emptor. *J. Product. Anal.* 36 (2), 205–218. DOI 10.1007/s11123-011-0230-6. DOI: 10.1007/s11123-011-0230-6.
- Sun, Y., Gao, C., Li, J., Gao, M., Ma, R., 2021. Assessing the cooling efficiency of urban parks using data envelopment analysis and remote sensing data. *Theor. Appl. Climatol.* 145 (3), 903–916. <https://doi.org/10.1007/s00704-021-03665-2>.
- Tone, K., 2001. A Slack-based measure of efficiency in data envelope analysis. *Eur. J. Oper. Res.* 130 (3), 498–509. [https://doi.org/10.1016/S0377-2217\(99\)00407-5](https://doi.org/10.1016/S0377-2217(99)00407-5).
- UNEP-WCMC and IUCN, 2026. Protected planet: the world database on protected areas (WDPA), Cambridge, UK. <https://www.protectedplanet.net/en/thematic-areas/wdpa?tab=WDPA>.
- Valdivieso, J.C., Eagles, P.F., Gil, J.C., 2015. Efficient management capacity evaluation of tourism in protected areas (<https://doi.org/10.1016/j.jenvman.2014.09.079>). *J. Environ. Plan. Manag.* 58 (9), 1544–1561. <https://doi.org/10.1080/09640568.2014.937479>.
- Valdivieso, J.C., Eagles, P.F.J., Gil, J.C., 2023. Evaluation of the impact of policymakers' decisions in the management capacity of protected areas: efficiency evidence from five countries. *Environ. Manag.* 71 (2), 274–284. <https://doi.org/10.1007/s00267-022-01729-1>.
- Watson, J.E., Dudley, N., Segan, D.B., Hockings, M., 2014. The performance and potential of protected areas. *Nature* 515 (7525), 67–73. <https://doi.org/10.1038/nature13947>.
- Wilson, P.W., 2008. FEAR: a software package for frontier efficiency analysis with R. *Socio-Econ. Plan. Sci.* 42 (4), 247–254. <https://doi.org/10.1016/j.seps.2007.02.001>.
- Worthington, A.C., 2000. Cost efficiency in Australian local government: a comparative analysis of mathematical programming and econometrical approaches. *Financ. Account. & Manag.* 16 (3), 201–223. <https://doi.org/10.1111/1468-0408.00105>.
- Zhao, M., Dong, S., Xia, B., Li, Y., Li, Z., Chen, W., 2022. Effective and sustainable managed protected areas: evaluation and driving factors of eco-efficiency of China's forest parks. *Forests* 13 (9), 1406. <https://doi.org/10.3390/f13091406>.