Ref. Ares(2026)6578274 - 30/06/2026

TOURAL

Multidimensional model of tourism verticals driving the sustainable balanced growth among rural & remote grids and urban clusters of rural regions, fostering macro-regional cooperation

101132489, HORIZON-CL2-2023-HERITAGE-01

Deliverable D5.1:

Report on Sustainability Simulator & Predictive Advisor

Deliverable Version:	D5.1, v1
Document Identifier:	TOURAL_D5.1_Sustainability Simulator and Predictive Advisor_CERTH_v1
Preparation Date:	March 5, 2026
Document Status:	Final
Author(s):	Aliki Stefanopoulou, Alexis Papaioannou, Nikolaos Gkaripis, Iakovos Michailidis, Eleni Tsompanidou, Evdoxia Eirini Lithoxoidou, Dimosthenis Ioannidis - CERTH
Dissemination Level:	PU - Public

Funded by the European Union

HORIZON-CL2-2023-HERITAGE-01-05

Deliverable SUMMARY SHEET

Deliverable Details	
Type of Document:	Deliverable
Document Reference #:	D5.1
Title:	Report on Sustainability Simulator & Predictive Advisor
Version Number:	1
Preparation Date:	March 5, 2026
Delivery Date:	June 30, 2026
Author(s):	Aliki Stefanopoulou, Alexis Papaioannou, Nikolaos Gkaripis, Iakovos Michailidis, Eleni Tsompanidou, Evdoxia Eirini Lithoxoidou, Dimosthenis Ioannidis - CERTH
Contributor(s)	Marija Dušević (ZDRN), Roxana Iacob (GALDD), Alessandro Perfetti (Itaca)
Document Identifier:	TOURAL_D5.1_Sustainability Simulator and Predictive Advisor_v1
Document Status:	Final
Dissemination Level:	PU - Public

Project Details	
Project Acronym:	TOURAL
Project Title:	Multidimensional model of tourism verticals driving the sustainable balanced growth among rural & remote grids and urban clusters of rural regions, fostering macro-regional cooperation
Project Number:	101132489
Call Identifier:	HORIZON-CL2-2023-HERITAGE-01
Call Theme:	**HORIZON-CL2-2023-HERITAGE-01-05**
Project Coordinator:	Centre for Research and Technology Hellas (CERTH)
Participating Partners:	CERTH – Centre for Research and Technology Hellas – Greece MoN – NESEBAR MUNICIPALITY – Bulgaria GALDD – ASOCIATIA GAL DELTA DUNARII – Romania KIN – CIVIL ORGANIZATION KINBURN – Ukraine BSEC – ORGANIZATION OF THE BLACK SEA ECONOMIC COOPERATION – Türkiye ICBSS – INTERNATIONAL CENTER FOR BLACK SEA STUDIES – Greece MoS – DIMOS SKOPELOU – Greece ITACA – IL TERRITORIO AL CENTRO DELLATTENZIONE - ASSOCIAZIONE PER LO SVILUPPO LOCALE – Italy PoT – PROVINCIA DI TERAMO – Italy

	ABR – REGIONE ABRUZZO – Italy ATL – ATLANTIS SYMVOULEFTIKI ANONYMI ETAIREIA ATLANTIS CONSULTING SA – Greece IHU – DIETHNES PANEPISTIMIO ELLADOS – Greece
Contract Start Date:	January 1, 2024
Duration:	36 Months

Deliverable D5.1: Short Description
This deliverable consolidates the results of Tasks T5.1 and T5.2 into a single integrated decision-support framework. Task T5.1 focuses on the Tourism Flows Simulation and **Flow** Capacity Tool, which estimates expected tourism flows for a selected destination and future month or season. The tool is designed to assess how incoming tourism demand interacts with destination capacity and sustainability-related indicators. It uses heterogeneous data sources, including tourism statistics, spatial characteristics, population density, pollution levels, transport accessibility, and hospitality infrastructure. By combining historical and current datasets, it supports AI-based simulation and digital-twin-oriented destination modelling. Task T5.2 builds on the simulation tool by introducing a Predictive Advisor for capacity assessment and scenario evaluation. The advisor analyses different tourism flow configurations across regions and seasons. It is designed to evaluate the pressure exerted by incoming tourism on environmental, social-cultural, and economic dimensions. The system applies KPI-based scoring and advanced multi-criteria decision-making methods to compare alternative scenarios. Explainable AI techniques are used to derive actionable recommendations on desirable tourism volumes and visitor-type mixes. The outputs support destination managers, policy makers, and local authorities in making evidence-based decisions. Overall, the deliverable provides an integrated methodological and technical foundation for sustainable tourism planning and capacity management in TOURAL pilot regions. **Keywords**: Tourism Flows, Capacity Management, Sustainable tourism, Decision support system

Deliverable D5.1: Revision History			
Version:	**Date:**	**Status:**	**Comments**
0.1	5/3/2026	draft	First ToC structure by CERTH
0.2	27/03/26	draft	ToC of Predictive Advisor & capacity tool
0.3	22/04/26	draft	Addition of Technical Requirements for the Digital Platform
0.4	22/05/26	draft	Addition of Predictive Advisor & capacity tool content
0.5	22/06/26	Semi-final	Ready for peer review
1.0	30/06/26	Final	Final Deliverable

Copyright notices

© 2026 TOURAL Consortium Partners. All rights reserved. TOURAL is a Horizon Europe Project supported by the European Commission under contract #101132489. For more information on the project, its partners, and contributors please see http://www.toural-project.eu/. You are permitted to copy and distribute verbatim copies of this document, containing this copyright notice, but modifying this document is not allowed. All contents are reserved by default and may not be disclosed to third parties without the written consent of the TOURAL partners, except as mandated by the European Commission contract, for reviewing and dissemination purposes. All trademarks and other rights on third party products mentioned in this document are acknowledged and owned by the respective holders. The information contained in this document represents the views of TOURAL members as of the date they are published. The TOURAL consortium does not guarantee that any information contained herein is error-free, or up to date, nor makes warranties, express, implied, or statutory, by publishing this document.

Table of Contents

1 Introduction 1
1.1 Purpose of the Deliverable 1
1.2 Relation to other activities 1
1.3 Contribution to TOURAL Objectives 1
1.4 Structure of the deliverable 2
2 Conceptual Framework 3
2.1 Overview 3
2.2 Conceptual Principles 3
2.2.1 Tourism Destinations as Dynamic Systems 3
2.2.2 Sustainability as a Multi-Dimensional Concept 5
2.2.3 Research Gap and Positioning of the TOURAL Framework 7
2.3 Integrated Framework for Simulation, Optimisation and Decision Support 8
2.3.1 Knowledge Acquisition Layer 10
2.3.2 Tourism Knowledge Base Layer 10
2.3.3 Destination Digital Twin Layer 10
2.3.4 Flow Capacity Tool Layer (T5.1) 11
2.3.5 Scenario Evaluation and Predictive Advisor Layer (T5.2) 11
2.3.6 Decision Support and Validation Layer 12
2.4 Framework Workflow 12
3 Data Framework 14
3.1 Data Sources 14
3.1.1 External datasets and benchmarking frameworks 14
3.1.2 Definition of the TOURAL Set of KPIs 15
3.1.3 Description of the selected KPIs 16
3.2 Data Limitations and Assumptions 19
4 Tourism Flows Impact Analysis and Capacity Control Tool 22
4.1 Purpose and Functional Scope of the Flow Capacity Tool 22
4.1.1 Objective of the Tool 22
4.1.2 Main Outputs of the Simulation 23
4.2 Digital Representation Model of the Destination 23
4.2.1 Tourism Flow Layer 24
4.2.2 Destination Capacity Layer 24
4.2.3 Sustainability Performance Layer 24
4.3 Simulation Inputs and Feature Space 25
4.3.1 Scenario Inputs 26
4.3.2 Features for Tourism-Flow Prediction 26

4.3.3 Dynamic KPI Calculation and Feature Construction ... *27*
4.4 Simulation Pipeline and Tourism Flow Prediction ... 28
4.4.1 Preparing the Scenario Inputs ... *29*
4.4.2 AI-Based Tourism Flow Prediction ... *29*
4.4.3 Flow Classification and Dashboard Outputs ... *30*
4.5 Impact Modelling, Capacity Monitoring and Pressure Detection ... 32
4.5.1 Sustainability Impact and Capacity Monitoring by Dimension ... *33*
4.5.2 KPI Scoring, Status Classification and Pressure Signals ... *34*
4.6 Impact Modelling per Sustainability Dimension ... 36
4.7 Tourism Flow Simulation and Destination State Estimation ... 37
4.8 Evaluation & Results ... 38
4.8.1 Tourism-Flow Prediction Performance ... *38*
4.8.2 KPI-based Impact Results ... *41*
5 Predictive Advisor on Areas Capacity VS Incoming Tourism Flows ... 45
5.1 Purpose and Functional Scope ... 45
5.2 Data Models ... 47
5.2.1 Scenario Assembly ... *53*
5.2.2 Derived Metric Pipeline ... *54*
5.3 KPI scoring methods ... 55
5.3.1 Forward Utilization (positive indicators — higher is better) ... *55*
5.3.2 Inverse Utilization (pressure indicators — lower is better) ... *56*
5.3.3 Composite and Category Scores ... *56*
5.4 KPI Indicator Specifications ... 57
5.5 Optimization Engine ... 65
5.5.1 Search Space ... *65*
5.5.2 Objective Function ... *67*
5.5.3 Trial Budget and Runtime ... *67*
5.5.4 Result Extraction ... *68*
5.6 Action Plan Generation ... 68
5.7 Results ... 70
6 Implementation ... 79
6.1 TOURAL ICT Suite Overview ... 79
6.2 Technical Architecture and Stack ... 80
6.2.1 Frontend Presentation ... *80*
6.2.2 Backend Infrastructure ... *81*
6.2.3 Deployment & Process Management ... *82*
6.2.4 Architecture & Stack of Flow Capacity Tool (T5.1) & Predictive Advisor (T5.2) ... *82*

6.3 API and Interoperability Layer 82
6.3.1 Flow Capacity Tool (T5.1) API Specification 82
6.3.2 Predictive Advisor (T5.2) API Specification 84
7 Conclusions 89
8 References 90
9 Annex – Flow Capacity Tool 92
10 Annex – Predictive Advisor Tool 96
Predictive Advisor – Results from the rest of the pilots 100

List of Figures

Figure 1: ICT-driven Conceptual Framework of the TOURAL Flow Capacity and Predictive Advisor tools 9
Figure 2: Flow Capacity Tool overview dashboard 22
Figure 3: Main system components of the digital-twin representation 24
Figure 4: Simulation pipeline and Tourism flow prediction tool 28
Figure 5: Flow Capacity Tool front-end interface for scenario input selection, including destination and time-period parameters. 29
Figure 6: Flow Capacity Tool dashboard showing flow-class summary and indicator-level sustainability outputs. 32
Figure 7: Predictive Advisor Architecture 46
Figure 8: Initial configuration for the North Sporades scenario 71
Figure 9: Baseline Results for the North Sporades scenario 72
Figure 10: Optimization results for the North Sporades scenario 73
Figure 11: Optimization results for the North Sporades scenario 74
Figure 12: TOURAL Digital Platform 79
Figure 13: Toural platform architecture 80

List of Tables

Table 1: Limitations of Existing Approaches 8
Table 2: Overview of Data Sources and Categories 14
Table 3: Feature groups by modelling function 25
Table 4: Sustainability impact logic by dimension 33
Table 5: KPI status categories and dashboard interpretation 36
Table 6: Sustainability dimensions, representative indicators, and example impacts considered by the Flow Capacity Tool 37
Table 7: Evaluation setup for tourism-flow prediction 38
Table 8: Comparison of tourism-flow prediction models for North Sporades 39
Table 9: CNN-BiLSTM feature-space comparison for North Sporades 40
Table 10: Environmental Conditions KPI Results for the North Sporades (Alonnisos) Pilot - June 2026 41
Table 11: Social & Cultural KPI Results for the North Sporades (Alonnisos) Pilot - June 2026 42
Table 12: Economic Aspects KPI Results for the North Sporades (Alonnisos) Pilot - June 2026 43
Table 13: Static destination parameters: field definitions and data types 47
Table 14: Seasonal flow parameters: field definitions and data types 48
Table 15: Destination-specific optimal KPI targets across the five TOURAL pilot sites 49
Table 16: Per-archetype fields of the tourist type mix data model 51
Table 17: Behavioural and environmental parameters of the eight tourist archetypes 52
Table 18: KPI specification and scoring method - Pillar A: Destination Management 57
Table 19: KPI specification and scoring method - Pillar B: Financial and Economic Performance 58
Table 20: KPI specification and scoring method - Pillar C: Social and Cultural Impact 59
Table 21: Impact scaling for Pillar D 62
Table 22: KPI specification and scoring method - Pillar D: Environmental Metrics 62
Table 23: Infrastructure parameter search domains for the optional optimization mode 66
Table 24: Prompt template for tourist type mix action generation 69
Table 25: Language model inference parameters for action plan generation 69
Table 26: Action plan output format for tourist type mix and KPI improvement recommendations 70
Table 27: KPI utilization scores before and after optimization — North Sporades, mid season 75
Table 28: Optimized tourist type mix and LLM-generated policy actions — North Sporades, mid season 76
Table 29: KPI Improvement Actions 77
Table 30: Frontend technologies used 81

Table 31: Backend technologies used **81**
Table 32: Core API Endpoint for the Flow Capacity Tool **83**
Table 33: Request Body Parameters for /predict-flow **83**
Table 34: Example JSON Request Payload **83**
Table 35: Example JSON Response Payload **84**
Table 36: Init Scenario Endpoint **84**
Table 37: Request Payload Structure for Init Scenario **85**
Table 38: Example JSON Request Payload (Init Scenario) **85**
Table 39: Example JSON Response (Immediate Acknowledgment) **86**
Table 40: Webhook Receiver Endpoint **86**
Table 41: Example Webhook Request Payload (From External Engine) **86**
Table 42: Polling Endpoint **87**
Table 43: Query Parameters for Polling Endpoint **87**
Table 44: Example JSON Response (Pending State) **87**
Table 45: Example JSON Response (Completed State) **88**
Table 46: Comparison of tourism-flow prediction models for Danube Delta **92**
Table 47: CNN-BiLSTM feature-space comparison for Danube Delta **92**
Table 48: Environmental Conditions KPI Results for the Danube Delta Pilot - June 2026 ..... **93**
Table 49: Social & Cultural KPI Results for Danube Delta Pilot - June 2026 ..... **94**
Table 50: Economic Aspects KPI Results for the Danube Delta Pilot - June 2026 ..... **95**
Table 51: Summary of environmental and economic estimation functions. **96**
Table 52: KPI utilization scores before and after optimization — Teramo, High season **101**
Table 53: Optimized tourist type mix and LLM-generated policy actions — Teramo, High season **101**
Table 54: KPI Improvement Actions **103**
Table 55: KPI utilization scores before and after optimization — Island of Pag , Medium season **105**
Table 56: Optimized tourist type mix and LLM-generated policy actions — Island of Pag , Medium season **106**
Table 57: *KPI Improvement Actions* **107**
Table 58: KPI utilization scores before and after optimization — Nessebar, High season **109**
Table 59: Optimized tourist type mix and LLM-generated policy actions — Nessebar, High season **110**
Table 60: KPI Improvement Actions **111**
Table 61: KPI utilization scores before and after optimization — Danube Delta, Low season **114**
Table 62: Optimized tourist type mix and LLM-generated policy actions — Danube Delta, Low season **115**
Table 63*: KPI Improvement Actions* **116**

Executive Summary

This deliverable presents the tools developed under Tasks “T5.1 Tourism Flows Impact Analysis and Capacity Control Tool” and “T5.2 Predictive Advisor on Areas Capacity VS Incoming Tourism Flows” of the TOURAL project. It describes the Flow Capacity Tool and the Predictive Advisor, highlighting their main functionalities, their interconnection, and their contribution to sustainable tourism planning and evidence-based destination management.

The tool developed under Task 5.1 is designed to simulate tourism flows for a selected destination and future time period, and to assess their implications for destination capacity and sustainability performance. By combining heterogeneous datasets, AI-based modelling approaches, and a digital-twin-oriented representation of the destination, it generates forecasts, impact estimates, KPI-based assessments, pressure signals, alerts, and dashboard outputs. The tool developed under Task 5.2 builds upon these outputs and serves as the decision-support layer of the framework, enabling the evaluation, comparison, and optimisation of alternative tourism scenarios through KPI scoring, multi-criteria analysis, and explainable recommendation mechanisms.

The Flow Capacity Tool was evaluated using pilot-specific monthly datasets comparing CNN-BiLSTM against Random Forest, XGBoost and HistGradientBoosting models. The Predictive Advisor was applied across five pilot destinations and generated optimized scenarios, KPI-based sustainability scores and action-plan recommendations. Results demonstrate that the framework can forecast tourism flows, detect sustainability pressure conditions and support evidence-based planning through optimization and scenario comparison.

The deliverable also presents the concept of the integrated TOURAL Digital Platform, developed to provide a common front-end entry point for the ICT tools of the project. Rather than offering separate interfaces for individual applications, the TOURAL Digital Platform is intended to function as a unified digital environment through which users can access tools developed across the project, including tools originating from both WP3 and WP5. In this context, the present deliverable focuses on the detailed presentation of two of these tools, namely the Flow Capacity Tool and the Predictive Advisor.

List of Symbols and Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
API	Application Programming Interface
CCT	Cultural & Creative Tourism
CH	Cultural Heritage
CNN-BiLSTM	Convolutional Neural Network - Bidirectional Long Short-Term Memory
CST	Cultural Science Tourism
ICT	Information and Communication Technology
KAC	Knowledge Awareness Centre
KPI	Key Performance Indicator
LLM	Large Language Model
MDM	Multidimensional Model
MRC	Macro-Regional Cooperation
RRG	Rural and Remote Grids
SBG	Sustainable Balanced Growth
SLT	Silver Tourism
TV	Tourism Verticals
UCCST	UCNHT-CCT-CST-SLT
UCH	Underwater Cultural Heritage
UCNHT	Cultural & Nature Heritage Tourism
UCRR	Urban Clusters of Rural Regions
UW	Underwater
VR	Virtual Reality
WP	Work Package

1 Introduction

1.1 Purpose of the Deliverable

Deliverable 5.1 ‘’Report on Sustainability Simulator & Predictive Advisor’’ presents the design and functionality of the Flow Capacity Tool (T5.1) and the Predictive Advisor tool (T5.2). Together, these tools form an integrated decision-support system for analyzing tourism dynamics and supporting sustainable destination management. The deliverable explains the role of each tool, the way they interact within a common analytical workflow, and their contribution to evidence-based destination management. It also documents their integration within the TOURAL Digital Platform as a common front-end environment for accessing project tools and services.

1.2 Relation to other activities

This deliverable, *‘’D5.1 Report on Sustainability Simulator Predictive Advisor”*, is closely related to the activities developed across WP2, WP3, WP4, and WP5 of the TOURAL project. It reports the outcomes of Task *‘’T5.1 Tourism Flows Impact Analysis and Capacity Control Tool”* and *‘’T5.2 Predictive Advisor on Areas Capacity VS Incoming Tourism Flows”*, which together form two complementary components of the project’s ICT ecosystem. Their conceptual and methodological foundations build on the work carried out in WP3, particularly ‘’*D3.1 TOURAL UCCST Models”*, ‘’*D3.2 TOURAL Optimisation Methods for UCCST Models – Sustainability & Barriers”*, and ‘’*D3.3 Action Plans, KACs UCCST Model Implementation”,* from which the sustainability dimensions, Key Performance Indicator (KPI) logic, and optimisation perspective are derived. The framework is also informed by WP2 activities, especially ‘’*D2.2 Policy Pathways for TOURAL Regions Macro-Regions”* and ‘’*D2.5 Preliminary Analysis of TOURAL Co-Work Processes”*, which provide policy, stakeholder, and co-design input relevant to the development of the tools. Within WP5, T5.1 provides the simulation and impact-assessment layer, while T5.2 builds upon these outputs to compare scenarios, assess trade-offs, and generate actionable recommendations. The outputs of both tools are intended to support WP4 activities, in particular ‘’*T4.2 Co-design Business Model Individualisation (region/macro-region) & Resiliency Assessment”*, ‘’*T4.3 Co-design of TOURAL Tourism Products Reinforced by Marketing Strategies”*, and ‘’*T4.5 Small-Scale UCCST Model Services Validation”*, as well as the related deliverables ‘’*D4.2 Business Models, Tourism Products, Marketing Strategies Resiliency Assessment for TOURAL Regions*” and *‘’D4.4 Report on Small-Scale UCSST Model Services Validation”*. These interconnections are further elaborated in Chapter 2, where the conceptual framework supporting the integration of the relevant TOURAL activities is presented.

1.3 Contribution to TOURAL Objectives

The tools developed under Tasks T5.1 and T5.2, and analyzed in this deliverable, contribute directly to the overall objectives of the TOURAL project by transforming the conceptual models, sustainability frameworks, policy pathways, and stakeholder knowledge developed throughout the project into operational decision-support tools for sustainable tourism management. In particular, the deliverable operationalizes the project’s ambition to support sustainable and balanced tourism development in rural, remote, and coastal regions through simulation, impact assessment, KPI-based evaluation, scenario comparison, and recommendation generation.

More specifically, the Flow Capacity Tool (T5.1) supports the analysis of tourism flows, destination capacity, and sustainability pressures, while the Predictive Advisor (T5.2) enables the evaluation and optimisation of alternative tourism scenarios across environmental, social-cultural, and economic dimensions. In this way, D5.1 strengthens evidence-based decision making, supports resilience and adaptive planning, facilitates stakeholder-oriented regional planning processes, and contributes to the digital transformation of tourism governance through its integration within the TOURAL Digital Platform.

1.4 Structure of the deliverable

This deliverable is organised into seven main chapters. Following this introductory chapter, Chapter 2 presents the conceptual framework of the Flow Capacity and Predictive Advisor Tools, including the main principles and the relations among knowledge acquisition, destination modelling, simulation, decision support, and validation. Chapter 3 describes the data framework, the selected KPIs, and the main assumptions and limitations of the analysis.

Chapter 4 focuses on the Flow Capacity Tool developed under Task T5.1, while Chapter 5 presents the Predictive Advisor developed under Task T5.2. Chapter 6 outlines the implementation aspects of the TOURAL Suite, including the technical architecture and the interoperability layer. Finally, Chapter 7 provides the conclusions of the deliverable, followed by the references and the annex.

2 Conceptual Framework

2.1 Overview

The Flow Capacity Tool (T5.1) and Predictive Advisor (T5.2) have been designed as complementary components of an integrated decision-support ecosystem aiming to support sustainable tourism planning, capacity management, and evidence-based decision-making across TOURAL pilot regions.

The conceptual foundation of the framework is based on the assumption that tourism destinations constitute complex socio-ecological systems where tourism demand, infrastructure capacity, environmental resources, local communities, and economic activities continuously interact. Consequently, tourism development cannot be evaluated solely through visitor volumes or economic performance, but must instead be assessed through a multidimensional sustainability perspective encompassing environmental, economic, social, and cultural dimensions.

The design of the TOURAL framework is informed by established research on tourism carrying capacity, sustainability assessment, destination management, tourism-flow modelling, tourist segmentation, and optimisation-based decision support. These research streams provide the scientific foundations upon which the Flow Capacity Tool and Predictive Advisor have been developed.

Building upon the UCCST framework developed within WP3, the Flow Capacity Tool and Predictive Advisor operationalize the conceptual models, sustainability assessments, optimisation methods, and regional action plans developed throughout the project. The framework transforms the static knowledge generated within previous work packages into dynamic simulations, scenario evaluations, and actionable recommendations capable of supporting tourism stakeholders during planning and decision-making activities.

The resulting framework establishes a continuous workflow linking tourism knowledge acquisition, destination modelling, simulation, optimisation, policy support, and pilot validation.

This chapter presents the conceptual and scientific foundations underpinning the TOURAL framework. It first reviews the key concepts and methodological approaches that informed the design of the tools and subsequently introduces the integrated framework through which these concepts are operationalised within the Flow Capacity Tool and Predictive Advisor.

2.2 Conceptual Principles

2.2.1 Tourism Destinations as Dynamic Systems

The conceptualisation of tourism destinations has undergone a substantial transformation, moving away from static, linear representations towards a view of destinations as complex, dynamically evolving systems. Early destination theory was dominated by the Tourism Area Life Cycle model, which depicted destinations as progressing through a characteristic sequence of stages, exploration, involvement, development, consolidation, stagnation, and ultimately rejuvenation or decline, and which framed the resources that enable a region to become a tourist destination as finite and ultimately exhaustible [1]. Although foundational, such stage-based models captured the temporal dimension of destination change only at an aggregate, descriptive level and offered limited insight into the underlying mechanisms, the interacting components, feedback structures, and non-linear responses, that drive that change. Recent scholarship has reinforced this reorientation, reframing destination management itself as the steering of a dynamic, evolving system rather than the administration of a static product [2].

A foundational reorientation in this regard is the reconceptualisation of tourism as a complex adaptive system, which argued that destinations should be understood through the integration of human and natural subsystems, non-linear dynamics, and the principles of resilience and adaptation rather than through reductionist cause-and-effect models [3]. This perspective was subsequently formalised through complexity-science methods, providing quantitative evidence that a tourism destination behaves as a complex system characterised by non-linear stakeholder interactions, feedback loops, emergent structure, and limited predictability [4]. More recent work has consolidated the complex-adaptive-systems lens as an analytical framework for destination governance, drawing explicit attention to the structure-agency interface, the process of adaptation through which systems survive in times of change, and the often-undervalued role of time, and arguing for a shift from reactively fixing problems towards proactively building adaptive capacity [5]. Building on the same lens, destination governance models have begun to treat coordination deficiencies and unplanned growth as systemic entropy to be managed, integrating public, private, and community actors, and increasingly the digital traces of visitors themselves, as interacting elements of a single adaptive system [6]. A practical implication common to this body of work is that interventions do not produce proportional, predictable outcomes. Small changes in demand or its composition may propagate disproportionately, while large interventions may be absorbed with little net effect.

A direct consequence of this systems-oriented understanding is the reinterpretation of tourism carrying capacity. Early formulations treated carrying capacity as a fixed threshold, an upper bound on the number of visitors a destination could absorb without unacceptable degradation of its physical, economic, or socio-cultural environment [7]. More recent research has reconceptualised carrying capacity as a dynamic and multidimensional construct that varies with destination type, seasonality, infrastructure quality, and the composition of the tourist population, and that must therefore be modelled and managed rather than fixed once and for all [8]. Within this view, the state of a destination, expressed across environmental, social-cultural, and economic dimensions, is not a constant property but an emergent outcome of the continuous interaction between incoming tourism demand and the destination's finite capacities. This implies that the same destination may be sustainable under one demand configuration and over-pressured under another, and that meaningful assessment requires representing the destination's evolving state over time rather than a single static measurement. It also implies that capacity is not a single number but a profile: a destination may retain ample physical and economic capacity while its social or ecological capacity is already exceeded, so that aggregate visitor counts conceal precisely the imbalances that matter most for sustainable management.

To capture this temporal dimension, a growing body of work has applied simulation to tourism destinations, representing them through stocks, flows, and feedback structures in order to explore how a destination might evolve under alternative development trajectories. System-dynamics models have been used to test tourism development scenarios on island and coastal destinations, showing how unchecked growth can erode carrying capacity through water shortages, pollution, and overcrowding, and how different policy scenarios produce divergent long-term sustainability outcomes [9]. More recently, the same prospective ambition has been pursued through agent-based models that conceptualise the tourist city as a complex system and reproduce the spatial and temporal interactions of heterogeneous residents and visitors, allowing emergent patterns of tourism intensity, and the effectiveness of candidate management strategies, to be examined under different scenarios [10]. In parallel, digital-twin approaches have been proposed as virtual, continuously updated representations of a destination that integrate heterogeneous real-time data to support scenario simulation and data-driven decision-making, marking a shift away from management based on historical or static data towards anticipatory, state-aware destination

management [11]. Across these approaches, the common thread is a move from describing what a destination is towards anticipating how it would respond to a contemplated change in demand, infrastructure, or policy.

Taken together, these contributions establish the systems perspective as the dominant scientific basis for understanding destinations, while also exposing a persistent limitation in practice. Much of the complexity-science and simulation literature remains methodologically sophisticated but highly bespoke: models are typically developed for individual destinations, require substantial data and modelling effort, and seldom translate their outputs into the routine, indicator-level information used by destination managers in everyday decision-making. As a result, the explicit representation of a destination's state, and of its evolution under alternative tourism scenarios, remains relatively under-operationalised in tools designed for practical destination management. This gap between a mature conceptual understanding of destinations as dynamic systems and the scarcity of operational tools that make this understanding actionable motivates the systems-oriented design adopted in the present framework. In this framework, the destination is modelled as a dynamic system whose state is estimated from the interaction between predicted tourism flows and finite capacities, and is subsequently used as the basis for scenario evaluation and recommendation.

2.2.2 Sustainability as a Multi-Dimensional Concept

The Predictive Advisor (T5.2) is grounded on established sustainability assessment, carrying-capacity management and tourism optimization literature. This section summarizes the main conceptual and methodological foundations that informed the design of the KPI framework, tourist segmentation model and optimization engine implemented in the TOURAL Predictive Advisor.

The measurement and management of sustainable tourism at the destination level has been a subject of sustained scholarly and policy interest since the early 1990s, when the World Tourism Organization first formalized the concept of tourism carrying capacity as the maximum visitor load a destination can sustain without unacceptable degradation of its physical, economic, or socio-cultural environment[1]. Early frameworks treated carrying capacity largely as a static threshold, but subsequent research progressively recognized it as a dynamic, multidimensional construct dependent on destination type, seasonal variability, infrastructure quality, and the composition of the tourist population itself [12] [13]**.** Modelling a destination as a system has since emerged as a promising approach to supporting tourism management for sustainability, enabling the estimation of optimal visitor numbers in line with the specific characteristics of a given territory and the simulation of tourism impacts in terms of costs and benefits across different visitor segments [14].

The institutionalization of sustainability indicator frameworks at the European level represents a significant milestone in operationalizing these concepts for destination managers. The European Tourism Indicator System (ETIS)[2], developed by the European Commission, is a management, information, and monitoring tool specifically intended for tourism destinations, designed as a locally owned and led process for collecting and analyzing data with the overall objective of assessing the impact of tourism on a destination. ETIS is structured around 43 indicators organized into four categories, destination management, social and cultural impact, economic value, and environmental impact, and was designed to encourage destination management organizations (DMOs) to look beyond traditional statistics such as visitor numbers and length of stay, instead reflecting on the broader impacts of tourism on local economies, communities, and environments.

[1] https://www.e-unwto.org/toc/unwtoecd/1992/1

[2]https://ec.europa.eu/docsroom/documents/6657/attachments/1/translations/en/renditions/native

While ETIS has been widely piloted across European destinations, empirical assessments of its implementation have consistently highlighted challenges in data availability, local adaptability, and the translation of indicator outputs into actionable policy decisions [15] [16]. More recent conceptual proposals have expanded this four-pillar approach to encompass visitor economy, resident support, workforce satisfaction, and environmental health, arguing that the high cost and complexity of ongoing data collection have historically impeded the broad adoption of sustainability initiatives at the community level [17].

The question of how the composition of tourist types influences destination sustainability outcomes has received comparatively less formal treatment in the optimization literature, despite its practical importance. Environmental impact at a destination varies not only with the total volume of tourism but also with the type of tourist and the intensity of site use, a finding that motivates the need for visitor segmentation approaches that go beyond aggregate flow statistics [18]. Research on small island destinations has been particularly instructive in this regard: tourism development on small islands presents acute sustainability challenges due to ecological fragility, spatial constraints, and reliance on limited natural and social resources, conditions under which the differentiation of tourist segments by expenditure, length of stay, carbon footprint, and water consumption becomes a meaningful lever for destination management [19] [20]. Studies on archipelago ecotourism have demonstrated that different tourism types impose substantially different carrying capacity constraints, reinforcing the argument that heterogeneous tourist mix modelling is a prerequisite for meaningful sustainability optimization.

The application of mathematical optimization and operations research methods to tourism planning has grown considerably over the past two decades, though the dominant focus has been on itinerary planning and routing rather than destination-level strategy. Metaheuristic approaches, including genetic algorithms, simulated annealing, and ant colony optimization have been widely applied to tourist trip planning and the Orienteering Problem variants that underlie personalized route construction [21] [22]. Hybrid metaheuristics combining simulated annealing with genetic algorithms have been shown to effectively optimize urban travel routes by intelligently selecting destinations and sequencing visits based on user preferences and geographic constraints [23]. More recently, IoT-augmented approaches have begun applying population-based metaheuristics to resource allocation management at tourism hotspots, exploiting the ability of these algorithms to traverse large search spaces and balance exploration and exploitation in the distribution of infrastructure, personnel, and facilities. However, the majority of these contributions address operational-level planning problems rather than the strategic-level question of how destination-wide KPI profiles respond to changes in tourist composition and infrastructure policies [22].

System dynamics modelling has emerged as a complementary paradigm for capturing the feedback structure of tourism destinations as complex adaptive systems. Integrated system-dynamics and multi-objective evolutionary algorithm frameworks have been proposed to jointly optimize economic, environmental, and social objectives, with case studies demonstrating portability across destinations with different data availability and capacity constraints [14] [24]. While such approaches offer theoretical completeness, their data requirements and computational complexity can pose barriers to operational deployment in resource-limited destination management contexts [24].

Sequential model-based optimization, and in particular the Tree-structured Parzen Estimator (TPE) algorithm, offers a computationally tractable alternative for black-box optimization over mixed continuous-discrete search spaces [25]. Operationalized within the Optuna framework, TPE implements Bayesian optimization using a surrogate model that iteratively refines probability density estimates over the parameter space, concentrating sampling in regions

associated with high objective values [26]. This approach is particularly well-suited to settings where evaluating the objective function is expensive or time-consuming, balancing exploration of novel parameter combinations with exploitation of promising regions already identified [25] [26]. While TPE has been extensively applied in machine learning hyperparameter tuning, its use as a general-purpose optimizer for simulation-based policy search, including the kind of scenario-based KPI optimization described in this work, represents an emerging and relatively underexplored application domain.

Taken together, the literature identifies a clear gap at the intersection of destination-level sustainability indicator frameworks and data-driven optimization: existing indicator systems such as ETIS[3] provide structured measurement but lack optimization capabilities [15], while existing optimization methods for tourism focus predominantly on routing and itinerary problems rather than strategic destination management [21] [23]. The framework proposed in this work addresses this gap by coupling a multi-pillar KPI evaluation model, grounded in ETIS principles and extended with environmental and societal indicators, with a TPE-based optimization engine that searches over tourist type mix and infrastructure parameters to maximize overall KPI utilization.

2.2.3 Research Gap and Positioning of the TOURAL Framework

The management of sustainable tourism destinations has traditionally relied on sustainability indicator frameworks, tourism demand forecasting models, and tourism impact assessment methodologies. While these approaches provide valuable insights into tourism dynamics, several limitations remain.

First, existing sustainability assessment frameworks, such as ETIS and related indicator-based systems, primarily focus on monitoring and measuring tourism performance. Although they support the evaluation of environmental, economic, social, and cultural sustainability dimensions, they offer limited capabilities for exploring alternative future development strategies or identifying optimal intervention scenarios. Consequently, they function mainly as descriptive assessment tools rather than proactive decision-support systems.

Second, tourism forecasting and optimisation research has predominantly focused on operational problems such as tourist route planning, itinerary optimisation, resource allocation, and visitor demand prediction. Existing approaches rarely address the strategic challenge of balancing tourism demand with destination capacities, sustainability objectives, and long-term regional development goals. Moreover, many forecasting models require extensive historical data and are primarily designed for urban contexts, limiting their applicability in rural and regional destinations characterised by sparse and heterogeneous datasets.

Third, despite increasing recognition of tourism destinations as complex socio-ecological systems, relatively few frameworks integrate simulation, sustainability assessment, optimisation, and policy recommendation generation within a unified decision-support environment. Existing approaches often stop at the analysis stage and provide limited support,

Table 1, for translating analytical outputs into actionable planning recommendations for destination managers and policymakers.

[3]https://ec.europa.eu/docsroom/documents/6657/attachments/1/translations/en/renditions/native

Table 1: Limitations of Existing Approaches

Approach	Focus	Main Limitation
ETIS and sustainability indicator frameworks	Sustainability monitoring	Measure sustainability but do not optimise future tourism strategies
Carrying-capacity assessment models	Impact Assessment	Primarily evaluate impacts and thresholds rather than identify optimal interventions
Route-planning and tourism optimisation methods	Demand Prediction	Focus on operational routing problems rather than destination-level sustainability management
System-dynamics tourism models	Operational decisions	Require substantial data and modelling effort, limiting operational adoption

The TOURAL framework addresses these limitations by combining tourism-flow simulation, capacity-aware sustainability assessment, scenario evaluation, optimisation, and evidence-based decision support within a single integrated framework. Building upon the project objective of balancing tourism flows with regional capacities and sustainability impacts, the framework models tourism destinations as dynamic systems in which visitor demand interacts with environmental resources, infrastructure capacities, local communities, and economic activities.

Unlike traditional indicator systems, the proposed framework enables stakeholders to evaluate alternative tourism development scenarios before implementation. Unlike conventional optimisation approaches, it focuses on destination-level sustainability management rather than operational routing problems. Finally, by integrating simulation outputs with optimisation mechanisms and recommendation generation, the framework supports a transition from descriptive tourism analytics towards predictive and prescriptive decision-support capabilities.

The resulting framework therefore positions TOURAL at the intersection of sustainable tourism assessment, tourism simulation, optimisation, and policy support, providing destination managers with a practical mechanism for exploring future scenarios, identifying sustainable tourism configurations, and supporting evidence-based decision-making across diverse regional contexts.

2.3 Integrated Framework for Simulation, Optimisation and Decision Support

The proposed framework is organised into six interconnected layers that collectively support tourism simulation, scenario evaluation, optimisation, and decision support.

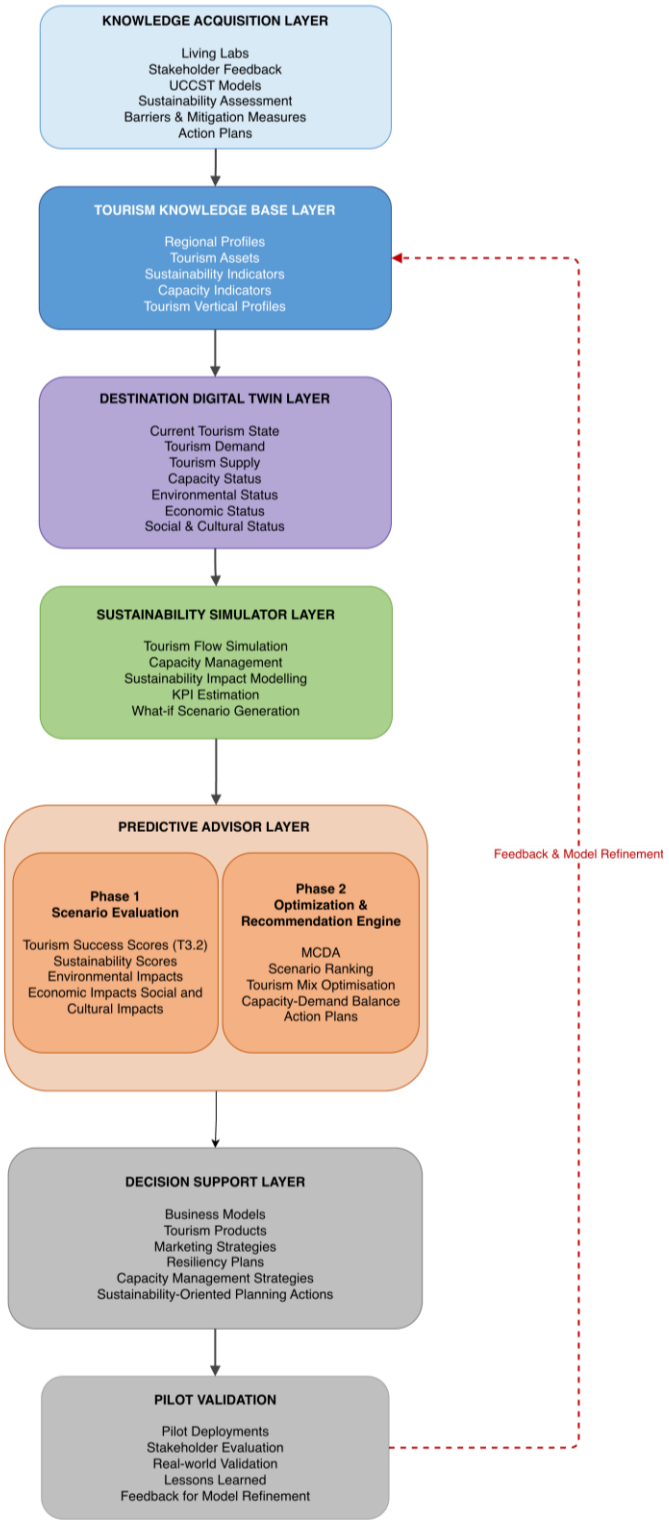

Figure 1: ICT-driven Conceptual Framework of the TOURAL Flow Capacity and Predictive Advisor tools

The conceptual framework presented in Figure 1 illustrates the intended flow of information between knowledge acquisition, destination modelling, simulation, decision support, and pilot validation activities within the TOURAL ecosystem.

The feedback relationship between WP4 Pilot Validation and the Tourism Knowledge Base represents a conceptual evaluation process through which validation outcomes, stakeholder observations, and lessons learned may inform future improvements of the framework.

It should be noted that the current implementation representing the Minimum Viable product (MVP), does not incorporate automated feedback learning, continuous model updating, or self-adaptive optimisation mechanisms. Consequently, the feedback path should be interpreted as a conceptual knowledge refinement process rather than an operational closed-loop learning mechanism.

2.3.1 Knowledge Acquisition Layer

The first layer is responsible for gathering and integrating information generated throughout the TOURAL project.

Knowledge is acquired from multiple sources, primarily originating from WP2, WP3, and WP4 activities.

WP2 contributes policy pathways, stakeholder engagement outcomes, living lab activities, governance mechanisms, and co-design processes that provide contextual information regarding regional priorities, opportunities, and constraints.

WP3 contributes the UCCST models, tourism vertical definitions, sustainability assessments, optimisation outcomes, implementation barriers, mitigation measures, and action plans that establish the conceptual and analytical foundation of the framework.

WP4 contributes business models, tourism products, marketing strategies, resiliency plans, and pilot validation outcomes that provide practical implementation knowledge and operational insights.

Together, these inputs form the knowledge foundation required for simulation and decision-support activities.

2.3.2 Tourism Knowledge Base Layer

The Tourism Knowledge Base serves as the central information repository of the framework, consolidating the knowledge generated throughout the TOURAL project into a unified representation of the tourism ecosystem. The knowledge incorporated within this layer originates primarily from the outcomes of WP2 and WP3, including stakeholder engagement activities, living labs, policy pathways, UCCST models, sustainability assessments, implementation barriers, mitigation measures, and action plans. Additional destination-specific information generated through WP4 activities further enriches the knowledge base whenever available.

Rather than functioning as a conventional data storage component, the Tourism Knowledge Base provides a structured and contextualised description of each destination, integrating tourism assets, regional characteristics, sustainability profiles, infrastructure capacity, tourism verticals, and stakeholder perspectives into a coherent knowledge structure.

This layer acts as the primary knowledge source for the Destination Digital Twin, ensuring that subsequent simulation and decision-support activities are grounded in the conceptual models, sustainability assessment mechanisms, and optimisation principles established in previous work packages. In this sense, the Tourism Knowledge Base represents the operational bridge between the analytical knowledge developed within WP2 and WP3 and the computational tools implemented in WP5.

2.3.3 Destination Digital Twin Layer

At the core of the framework lies the Destination Digital Twin, which serves as a computational representation of the tourism destination and its surrounding ecosystem. Being based on the information stored within the Tourism Knowledge Base, the digital twin integrates tourism

demand and supply characteristics, infrastructure capacity conditions, sustainability-related information, seasonal dynamics, and the tourism vertical profiles established through the UCCST framework into a unified destination model.

The purpose of this layer is to provide a consistent and holistic representation of the current destination state that can support subsequent analytical and decision-support processes. Rather than operating on fragmented datasets or isolated indicators, both the Flow Capacity Tool and the Predictive Advisor interact with this common destination representation, ensuring consistency across simulation outputs, scenario evaluation activities, and recommendation generation processes.

2.3.4 Flow Capacity Tool Layer (T5.1)

The Flow Capacity Tool constitutes the primary analytical engine of the framework and is responsible for translating the conceptual knowledge embedded within the Destination Digital Twin into quantifiable tourism development scenarios. Building upon the destination representation established in the previous layer, the simulator enables the exploration of alternative tourism conditions and assesses their potential implications for destination sustainability and carrying capacity.

The simulator operationalises the sustainability assessment principles, tourism success logic, and indicator structures inherited from the UCCST framework and the optimisation work conducted in WP3. Through the simulation of tourism flows, visitor distribution patterns, capacity conditions, and destination dynamics, it enables the examination of how different tourism development strategies may influence environmental, economic, social, and cultural sustainability dimensions.

Rather than focusing on a single forecast, the Flow Capacity Tool supports the exploration of multiple what-if scenarios representing alternative tourism futures. These scenarios may reflect variations in visitor volumes, tourism profiles, seasonal demand patterns, accessibility conditions, tourism product portfolios, or destination development priorities. For each scenario, the simulator estimates the resulting destination state and evaluates its implications using the sustainability assessment framework established in previous work packages.

The outputs generated by the simulator provide a quantitative representation of the expected consequences associated with alternative tourism development pathways. These outputs include tourism success and sustainability assessments, destination capacity indicators, impact estimates across the four sustainability dimensions, as well as pressure and risk indicators that can support subsequent decision-making activities. As such, the Flow Capacity Tool acts as the analytical bridge between destination modelling and strategic decision support, providing the evidence base upon which the Predictive Advisor operates.

2.3.5 Scenario Evaluation and Predictive Advisor Layer (T5.2)

The Predictive Advisor builds directly upon the outputs generated by the Flow Capacity Tool and constitutes the decision-support intelligence of the framework. Rather than simulating destination behaviour, its primary role is to analyse, compare, and optimise alternative tourism development scenarios in order to support evidence-based planning and decision-making.

Drawing upon the simulation outputs generated by T5.1, the advisor evaluates the performance of alternative tourism strategies from multiple sustainability perspectives. Through the application of multi-criteria decision analysis, optimisation techniques, and explainable reasoning mechanisms, it examines the trade-offs between tourism benefits, sustainability objectives, and destination capacity constraints. This process enables the identification of tourism development pathways that balance economic opportunities with environmental protection, social well-being, and cultural preservation.

The advisor further supports the exploration of alternative tourism mixes, seasonal distribution strategies, and capacity-management approaches, transforming simulation results into actionable insights for destination stakeholders. By systematically evaluating and ranking candidate scenarios, it identifies those that best align with sustainability objectives and regional development priorities.

The outputs generated by the Predictive Advisor are translated into practical recommendations that can support tourism planning and management activities. These recommendations may include preferred tourism development strategies, suggested tourism mixes, capacity-management measures, sustainability interventions, policy directions, and business-oriented actions tailored to the characteristics and needs of each destination.

As such, the Predictive Advisor represents the final analytical stage of the framework, bridging simulation-based assessment with practical decision support and facilitating the development of sustainable, resilient, and evidence-informed tourism strategies.

2.3.6 Decision Support and Validation Layer

The final layer connects the analytical capabilities of the framework with the practical implementation activities carried out within the TOURAL pilot regions. At this stage, the recommendations generated by the Predictive Advisor are utilized to support the design, refinement, and validation of tourism products, business models, marketing approaches, resiliency strategies, and broader destination-management actions.

Rather than representing a purely analytical process, this layer focuses on the translation of simulation results and optimisation outcomes into practical interventions that can be assessed under real-world conditions. The proposed actions and strategies are subsequently examined through pilot activities implemented across the participating destinations, allowing stakeholders to evaluate their applicability, usefulness, and potential contribution to sustainable tourism development.

The pilot activities shall provide valuable insights into the practical deployment of the proposed recommendations, including aspects related to stakeholder acceptance, operational feasibility, sustainability performance, implementation challenges, and overall effectiveness within the destination context. These observations contribute to the overall assessment of the framework and provide evidence regarding the relevance and applicability of the generated recommendations.

As such, this layer represents the operational validation component of the framework, bridging analytical decision support with real-world implementation and ensuring that the proposed strategies are assessed within the complex conditions of actual tourism destinations.

2.4 Framework Workflow

The proposed framework establishes a structured pathway through which knowledge generated across the TOURAL project is transformed into decision-support intelligence for sustainable tourism planning. Beginning with the integration of stakeholder knowledge, sustainability assessments, policy pathways, and UCCST models developed in previous work packages, the framework progressively translates conceptual and analytical knowledge into computational representations, simulation outputs, and evidence-based recommendations.

The Destination Digital Twin provides a unified representation of the tourism ecosystem, enabling the Flow Capacity Tool to assess tourism flows, capacity conditions, and sustainability implications under alternative development scenarios. Building upon these results, the Predictive Advisor evaluates and optimises candidate strategies, transforming analytical outputs into actionable recommendations that can support tourism planning and destination management activities.

The resulting recommendations are subsequently examined within the TOURAL pilot regions, where their applicability and practical value can be assessed under real-world conditions. Through this process, the framework creates a direct connection between the conceptual foundations established in WP2 and WP3, the analytical capabilities developed in WP5, and the implementation and validation activities carried out within WP4. As such, it provides an integrated mechanism for supporting evidence-based, sustainability-oriented tourism planning across participating destinations.

3 Data Framework

The objective of the data framework is to provide a consistent representation of tourism demand, destination capacity, sustainability performance, and tourism behaviour across all TOURAL pilot regions. To support both the Flow Capacity Tool (T5.1) and Predictive Advisor (T5.2), the framework combines destination-specific information, external datasets, and KPI-based sustainability indicators into a unified analytical structure.

3.1 Data Sources

The TOURAL Flow Capacity Tool and Predictive Advisor rely on a combination of regional pilot data**,** external benchmarking references, and a harmonised set of KPIs used to represent sustainability conditions in each destination. At the regional level, the primary data sources are provided by the pilot areas and reflect the current operational, environmental, social, cultural, and economic conditions of each destination. These inputs include tourism demand characteristics, accommodation and accessibility conditions, local infrastructure, destination-management features, selected environmental conditions, and other contextual variables required for the representation of each pilot area. Table 2 provides an overview of the key data categories, their sources, and how they are utilised to support modelling and analysis within the TOURAL framework.

Table 2: Overview of Data Sources and Categories

Category	Source	Used By	Purpose
Tourism flows	Regional statistics	T5.1	Monthly flow forecasting
Accommodation data	Pilot destinations	T5.1/T5.2	Capacity representation
Sustainability indicators	UCCST / WP3	T5.2	KPI Evaluation
Destination characteristics	Pilot datasets	Both	Digital Twin construction
External benchmarks	ETIS, Eurostat, etc	Both	Sustainability benchmarking

Because the availability and granularity of regional data differ across pilots, the project adopted a pragmatic data-collection approach that combines directly available values with expert-informed estimates provided by pilot stakeholders. This allows the system to represent the distinctive characteristics of each pilot while maintaining a common modelling structure across all regions. In this sense, the regional dataset is not limited to purely statistical observations. It also includes structured local knowledge that helps approximate destination conditions in cases where harmonised monitoring data are not yet available.

3.1.1 External datasets and benchmarking frameworks

In addition to pilot-provided data, the TOURAL selected KPIs were informed by a set of established tourism-sustainability benchmarking approaches. These sources were not adopted as a complete indicator system. Instead, they were used as methodological references for choosing indicator domains, shaping the scope of sustainability assessment, and ensuring that the selected KPIs reflect recognised destination-monitoring practice.

The main structural reference was the European Tourism Indicator System (ETIS). ETIS was developed by the European Commission as a management, information, and monitoring tool for

tourism destinations, with the explicit aim of helping destinations assess tourism impacts and improve the sustainable management of tourism activity. ETIS is organised around four broad categories: destination management**,** economic value**,** social and cultural impact**,** and environmental impact**.** This four-part logic strongly influenced the way TOURAL organised its selected KPIs, especially the decision to group them into governance-related, economic, social-cultural, and environmental sections [27].

A second important reference was CO-EVOLVE, which contributed a stronger emphasis on coastal sustainability, ecosystem pressure, and destination diagnosis at local and sub-local scale [4]. Interreg Euro-MED material describing the CO-EVOLVE sustainable tourism indicator toolkit explains that it was developed to monitor tourism and analyse threats and enabling factors for sustainable tourism in Euro-Mediterranean destinations, based on Integrated Coastal Zone Management and Maritime Spatial Planning principles. The same source notes that the toolkit was designed to be adapted to destinations through quantitative, qualitative, and proxy indicators. This perspective was especially relevant for TOURAL because several pilot areas are coastal, island, or environmentally sensitive destinations where tourism pressure must be understood in relation to biodiversity, protected areas, land and water use, and seasonality.

A third contextual reference was the BlueMed initiative[5], which provided a broader Mediterranean sustainability rationale. According to CORDIS, the BlueMed initiative was launched to support blue growth and jobs in the Mediterranean area while also improving social wellbeing, sustainable prosperity, and the good environmental status of the region. Within TOURAL, this orientation supported the inclusion of indicators that connect tourism activity with wider territorial and environmental resilience concerns, particularly in marine, coastal, and island contexts. Rather than offering a fixed tourism indicator catalogue, BlueMed informed the project's strategic understanding that tourism growth should be considered together with ecosystem quality, resource stewardship, and long-term territorial sustainability.

Overall, the role of these benchmarking references can be summarised as follows. ETIS provided the main structural logic for organising the selected indicators. CO-EVOLVE enriched the process with ideas related to environmental pressure, ecosystem sensitivity, and carrying-capacity-oriented diagnosis. BlueMed reinforced the Mediterranean sustainability perspective, especially the need to align tourism growth with ecological resilience and responsible territorial development. The final TOURAL set of KPIs therefore results from an adaptation and synthesis of existing benchmark approaches rather than from the direct adoption of a single framework.

3.1.2 Definition of the TOURAL Set of KPIs

The TOURAL KPIs were created through a three-step process. First, the consortium reviewed established sustainable tourism monitoring approaches, with particular focus on ETIS and related Mediterranean initiatives. Second, candidate indicators were assessed against the operational needs of T5.1 and T5.2, namely tourism-flow simulation, sustainability-impact representation, capacity-pressure detection, and scenario comparison. Third, the selected indicators were transformed into a common pilot template with harmonised names, references, units, and qualitative scales, so that heterogeneous pilot data could be represented in a comparable way.

A central design principle was to preserve the logic of benchmark-based sustainability monitoring while making the indicators usable inside the Flow Capacity Tool and the Predictive Advisor. For this reason, the project defined KPIs rather than adopting a complete external system. Some indicators are fully quantitative, expressed in nights, percentages, kilograms, euros, litres, or

[4] https://co-evolve.interreg-med.eu/

[5] https://cordis.europa.eu/project/id/727453

numbers of infrastructures. Others are qualitative or semi-quantitative and use categories such as low, medium, and high. This hybrid structure was necessary because not all pilots can provide the same level of statistical detail, especially for variables related to perceived safety, biodiversity disturbance, access pressure, or cultural inclusivity.

The final TOURAL KPIs are organised into four main pillars: A. Destination Management, B. Financial/Economic KPIs, C. Social and Cultural Impact, and D. Environmental Impact. This structure follows the core ETIS logic while extending it with additional indicators inspired by coastal sustainability, carrying-capacity concerns, local accessibility, digital inclusion, and environmental pressure.

3.1.3 Description of the selected KPIs

The Destination Management section includes KPIs that describe how sustainably and effectively a destination is organised and managed.

- Hotel numbers with certification for environmental measures / eco-label measures the number of commercial accommodation establishments that hold recognised environmental or sustainability certifications. This indicator was selected because it shows the degree to which tourism enterprises adopt formal sustainability practices.
- Number of beaches with Blue Flag measures how many beaches in the destination hold the Blue Flag distinction. This KPI was selected because it acts as a visible proxy for environmental quality, safety, and organised beach management, especially in coastal and island pilots.
- Percentage of tourists and same-day visitors satisfied with their overall experience in the destination measures the share of visitors who rate their experience very positively. This indicator was selected because visitor satisfaction is a core sign of destination performance and service quality.

Overall, the Destination Management KPIs are mainly aligned with the ETIS destination-management logic, as they assess how effectively the destination is organised, monitored, and managed in relation to sustainable tourism objectives. In particular, the KPI on hotels with environmental or eco-label certification and the KPI on visitor satisfaction are strongly based on ETIS, since they reflect recognised sustainability practices and destination-quality monitoring. The KPI on Blue Flag beaches is also connected to ETIS environmental and quality-management principles, while being particularly relevant to the CO-EVOLVE and BlueMed perspectives because it reflects coastal environmental quality, safety, and organised beach management in Mediterranean destinations. Therefore, this category combines ETIS as the main structural reference with CO-EVOLVE and BlueMed elements for coastal and marine destination contexts.

The Financial/Economic section includes KPIs that describe tourism intensity, enterprise performance, economic benefit, and employment effects.

- Number of tourist nights per month measures the total number of overnight stays in the destination in a given month. This is a basic tourism-flow variable and was selected because it directly supports the simulation of tourism intensity and seasonality.
- Average length of stay of tourists measures how many nights visitors remain in the destination on average. It helps distinguish between short-stay and longer-stay patterns and is useful for understanding demand behaviour.
- Occupancy rate in commercial accommodation establishments in high season measures the proportion of available commercial accommodation occupied during peak season. It was selected because it provides a direct indication of seasonal concentration and tourism pressure on accommodation capacity.
- Tourism revenue per capita measures tourism-generated revenue relative to the resident population. This indicator reflects the economic contribution of tourism at destination level.

- Percentage of tourism spending that leaves the region (Revenue Leakage Index) measures the share of tourism expenditure that is not retained locally. It was selected because strong tourism demand does not necessarily mean strong local benefit when a substantial part of the revenue exits the regional economy.
- Availability of free or low-cost activities and price-range diversity (Economic Inclusivity) measures whether the destination offers a range of affordable tourism options. This indicator was included to reflect inclusiveness and affordability within the tourism economy.
- Direct tourism employment as percentage of total employment in the destination measures the share of local employment directly linked to tourism. It indicates how important tourism is for the local labour market.

Overall, the Financial/Economic KPIs are primarily based on the ETIS economic-value dimension, which focuses on tourism performance, economic contribution, employment, seasonality, and the local value generated by tourism activity. KPIs such as tourist nights per month, average length of stay, occupancy rate, tourism revenue per capita, and direct tourism employment are strongly aligned with ETIS, as they represent core indicators for measuring the economic role of tourism in a destination. The Revenue Leakage Index extends this logic by considering whether tourism income remains within the local economy, while the Economic Inclusivity KPI reflects a broader sustainability concern related to affordability and inclusive access to tourism opportunities. These additional indicators are also consistent with the BlueMed perspective, which links tourism growth with sustainable prosperity and territorial resilience. As a result, this category is mainly ETIS-based, enriched with BlueMed-informed considerations on local economic retention, inclusiveness, and balanced development.

The Social and Cultural Impact section includes KPIs that describe pressure on local communities, accessibility conditions, inclusiveness, heritage sensitivity, and cultural identity.

- Number of tourists per 100 residents measures tourism intensity relative to local population size. It was selected because it is one of the clearest indicators of tourism pressure on everyday life and local carrying conditions.
- Level of safety captures the perceived or observed safety level of the destination. Although often qualitative, it is an important social sustainability variable.
- Percentage of commercial accommodation establishments participating in recognised accessibility information schemes measures how far accommodation providers address accessibility needs and communicate them properly.
- Percentage of public transport that is accessible to people with disabilities and specific access requirements measures how inclusive the local mobility system is.
- Percentage of tourist attractions that are accessible to people with disabilities and/or participating in recognised accessibility information schemes evaluates accessibility in the destination's visitor offer beyond transport and accommodation.
- Number of the destination's events that are not traditional or culturally relevant measures the volume of events not directly linked to local heritage or identity.
- Number of the destination's events focused on traditional/local culture and heritage measures how strongly the destination promotes and preserves local cultural expression through events.
- Heritage Site Stress Index captures visitor density and time-based pressure on heritage spaces. It was selected because it helps represent congestion and stress in sensitive cultural sites.
- Access to Services Index measures how tourism affects local residents' access to transport, health, and other essential services.
- Cultural Inclusivity measures the availability of multilingual information and culturally sensitive tourism materials and services.

- Social Inclusion captures whether tourism-related policies and practices support different user groups, such as families, seniors, and LGBTQ+ travellers.
- Digital Accessibility measures the accessibility of tourism-related digital interfaces, including websites and mobile applications.

Overall, the Social and Cultural Impact KPIs combine the ETIS social and cultural impact logic with additional elements inspired by CO-EVOLVE and BlueMed. KPIs such as tourists per 100 residents, safety level, accessibility of accommodation, accessibility of public transport, and accessibility of attractions are strongly aligned with ETIS, as they describe tourism pressure on communities, quality of life, and inclusive access to destination services. The KPIs related to traditional and non-traditional events, heritage-site stress, access to services, cultural inclusivity, social inclusion, and digital accessibility extend the ETIS framework by addressing cultural identity, heritage sensitivity, resident-service pressure, and inclusive participation. These aspects are also connected to the CO-EVOLVE approach, especially where tourism pressure affects local communities, cultural assets, and carrying-capacity conditions. In addition, the BlueMed perspective supports the inclusion of social cohesion, cultural resilience, and inclusive territorial development. Therefore, this category is based on a combination of ETIS, CO-EVOLVE, and BlueMed, with ETIS providing the main structure and the other frameworks enriching the interpretation of cultural and community-related sustainability.

The Environmental Impact section includes KPIs that describe transport sustainability, pressure on protected areas, waste generation, biodiversity, and environmental resource use.

- Number of green transport means from and to the destination measures the number of environmentally friendlier transport options available for reaching or leaving the destination.
- Types of green transport means from and to the destination identifies the categories of low-emission or sustainable transport options available.
- Number of tourism accommodation and attraction infrastructure located in cultural heritage protected areas measures the extent to which tourism infrastructure is located inside protected cultural zones.
- Number of tourism accommodation and attraction infrastructure located in Natura protected areas measures the extent to which tourism infrastructure is located inside protected natural areas.
- Waste production during tourist periods measures the amount of waste generated during tourism-intensive periods. It was selected because it is a direct expression of tourism-related environmental pressure.
- Percentage of local enterprises in the tourism sector actively supporting protection, conservation, and management of local biodiversity and landscapes measures the degree to which tourism businesses contribute to biodiversity and landscape stewardship.
- Tourism-related CO_2 emissions per capita/day measures carbon emissions associated with tourism activity.
- Water Consumption per Tourist measures how much water is used per visitor.
- Air Quality Index (seasonal) measures seasonal air-quality conditions.
- Ecological Footprint per Tourist captures the use of land and water resources associated with tourism.
- Biodiversity Disturbance Index measures the effect of tourism on habitats and species.
- Light Pollution Change Rate measures changes in artificial light levels in protected or dark-sky areas.
- Noise Pollution Index measures tourism-related noise increase against baseline conditions.

Overall, the Environmental Impact KPIs are based on a combination of ETIS, CO-EVOLVE, and BlueMed. ETIS provides the main environmental-impact logic for indicators related to waste production, water consumption, CO_2 emissions, air quality, green transport, and the environmental behaviour of tourism enterprises. CO-EVOLVE contributes strongly to indicators related to coastal and ecosystem pressure, such as tourism infrastructure in protected areas, infrastructure in Natura areas, biodiversity disturbance, ecological footprint, noise pollution, and light pollution. These KPIs are especially relevant for destinations where tourism activity interacts with sensitive natural, coastal, marine, or protected environments. BlueMed further supports this category by reinforcing the Mediterranean sustainability perspective, especially the need to connect tourism development with ecosystem quality, resource stewardship, blue growth, and long-term environmental resilience. Therefore, this category is not based on a single framework, but represents a synthesis in which ETIS supports general environmental monitoring, CO-EVOLVE strengthens the carrying-capacity and ecosystem-pressure dimension, and BlueMed provides the broader Mediterranean sustainability rationale.

3.2 Data Limitations and Assumptions

The development of the TOURAL Flow Capacity Tool and the Predictive Advisor was affected by important data-related limitations, mainly associated with data availability, granularity, harmonisation, and comparability across pilot areas. Although the project collected pilot-specific information and defined a common set of KPIs for the representation of sustainability conditions, it was not possible to obtain complete and fully homogeneous datasets for all destinations and for all variables of interest. As a result, the current modelling approach combines available measurements, expert-informed estimates, proxy representations, and structured assumptions in order to support a consistent and operational simulation process across the pilots.

Qualitative representation of selected variables

A first limitation concerns the fact that several variables could not be represented through continuous numerical measurements for all pilot areas. This was particularly the case for indicators related to environmental pressure, accessibility conditions, cultural stress, biodiversity disturbance, safety, and other destination-specific sustainability aspects for which harmonised and regularly updated quantitative observations were not available. In order to preserve these dimensions within the modelling process, a number of variables were represented using the qualitative classes low, medium, and high. This choice should not be interpreted as a simplification of the conceptual importance of these indicators, but rather as a necessary methodological adaptation to the data conditions of the pilots. The use of these qualitative levels made it possible to include relevant sustainability aspects in the tool even in cases where precise numeric values were unavailable, incomplete, or not directly comparable across regions.

The use of low**,** medium, and high classes was therefore adopted as a practical representation of relative conditions rather than as an exact measurement scale. In this context, the values assigned to these categories express the approximate level or intensity of a phenomenon within a pilot area. For example, a variable such as biodiversity disturbance, accessibility pressure, or perceived safety may be difficult to quantify directly through a single continuous measure, but it can still be meaningfully represented through an ordered qualitative scale.

Temporal alignment between monthly flows and seasonal KPIs

A second important limitation concerns the temporal alignment between tourism flow predictions and KPI representation. The Flow Capacity Tool operates at a monthly level, generating predictions for tourist flows for specific months. However, the available data for many KPIs did not support a consistent and harmonised monthly representation across all pilot areas. In several

cases, KPI measurements either did not exist at monthly resolution, were only partially available, or were not sufficiently reliable to support a unified modelling approach. For this reason, while tourism flows are predicted on a monthly basis, the KPIs are represented using a seasonal alignment, based on broader tourism-intensity categories, namely low, medium, and high tourism periods. These periods were defined according to the general seasonal patterns of tourism demand in each destination, using available data and expert input regarding the concentration of tourist activity throughout the year.

This approach allows the model to combine fine-grained monthly flow predictions with a more robust and comparable representation of sustainability conditions, avoiding an artificial level of precision for KPIs that cannot be supported by the available data. As a result, while the tourism demand component of the model reflects monthly variability, the KPI values correspond to the broader tourism period in which each month belongs.

Stability of KPI values within each tourism period

An additional assumption is that the selected KPI values are broadly representative of the tourism period to which they correspond. Since the available information did not allow for a fully dynamic update of all indicators on a monthly basis, the model assumes that KPI values remain relatively stable within each low, medium, or high tourism period. In practice, this may not always hold, as certain indicators can vary within the same period due to local events, sudden changes in demand, operational interventions, weather conditions, or short-term pressures on infrastructure and services. These intra-period variations are not explicitly captured in the current version of the tool and therefore constitute a source of modelling uncertainty.

Heterogeneity across pilot datasets

A further limitation is related to the heterogeneity of the pilot datasets. The pilot areas differ significantly in terms of data maturity, local monitoring practices, statistical availability, and the extent to which tourism-related information is systematically recorded. Some pilots were able to provide more structured and quantitative inputs, while others relied more heavily on estimates, qualitative assessments, or partially available evidence. In order to ensure consistency across the project, the modelling process required the adoption of a harmonised data structure and common indicator definitions. However, this harmonisation inevitably involved a degree of abstraction and simplification, as not all locally specific characteristics could be represented with the same level of detail in all pilot areas. Therefore, the current version of the model prioritises methodological consistency and cross-pilot comparability over highly localised detail.

Use of proxy variables and approximations

The project also relied, in some cases, on the use of proxy variables. Proxy representations were necessary when direct measurements were not available for a phenomenon that was nevertheless considered important for the simulation and sustainability assessment. This applies especially to variables related to complex environmental or socio-cultural effects, where direct observation is often difficult, expensive, or not systematically monitored.

Finally, it should be acknowledged that, in several cases, missing or incomplete information had to be approximated through pilot input, seasonal aggregation, or qualitative categorisation. These assumptions were necessary in order to make the tools operational across all pilots within a common methodological structure. Although this introduces uncertainty, it also provides a workable basis for simulation and comparative analysis in a context where complete harmonised datasets are not yet available.

Overall, the main assumptions underlying the current approach can be summarised as follows. First, tourism flows are modelled at a monthly level, while KPIs are represented at a seasonal level through low, medium, and high tourism periods due to data limitations. Second, KPI values are assumed to be broadly representative of the tourism period to which each month belongs. Third, some sustainability variables are represented through low, medium, and high categories because harmonised quantitative measurements were not available for all pilots. Fourth, some variables are approximated through proxy values, expert-informed estimates, seasonal aggregation, or qualitative categorisation when direct observations are missing or incomplete. The framework should be interpreted as a decision-support mechanism rather than an exact predictor of future tourism behaviour.

Future Work / Development Roadmap

While the Flow Capacity Tool and Predictive Advisor are fully operational, several factors currently bound the reliability of their outputs: the availability of pilot data, the reliance on estimated parameters, the limited historical depth of the datasets, the absence of real-time integration, and the broader constraints inherent to a Minimum Viable Product. Future development will address these points through a phased roadmap, beginning in the short term with additional pilot validation and KPI refinement, progressing in the mid term toward real-time data feeds and adaptive learning, and ultimately aiming, in the long term, for a full digital twin with an autonomous recommendation engine.

4 Tourism Flows Impact Analysis and Capacity Control Tool

This section presents the Tourism Flows Impact Analysis and Capacity Control Tool (Flow Capacity tool) developed in the context of Task 5.1. The section describes the functional scope of the tool, the digital-twin-oriented representation of each destination, the input data and modelling variables, the AI-based simulation pipeline, and the way predicted tourism flows are translated into sustainability and capacity-monitoring outputs. The section is organised according to the operational logic of the tool. First, it defines the purpose and outputs of the Flow Capacity Tool. Second, it explains the three core components used to represent each destination: tourism flow, destination capacity, and sustainability performance. Third, it describes the feature space used by the model. Fourth, it presents the simulation pipeline and AI-based monthly flow prediction. Finally, it explains how predicted flows are interpreted through KPI impacts, capacity pressure signals, alerts, and dashboard outputs.

Figure 2 illustrates the front-end interface of the Flow Capacity Tool, developed to provide users with an accessible and comprehensible view of the tool's operational workflow prior to the detailed technical presentation that follows.

Figure 2: Flow Capacity Tool overview dashboard

4.1 Purpose and Functional Scope of the Flow Capacity Tool

4.1.1 Objective of the Tool

The main purpose of the Task 5.1 is to estimate expected tourism flows for a selected destination and future month, and to interpret the implications of these flows in relation to destination capacity and sustainability-related indicators. The tool supports tourism planners, destination managers, regional authorities and project partners in understanding future tourism pressure in a structured and interpretable way. For a selected destination and selected future month, the tool estimates the expected tourist flow and evaluates how this predicted flow may influence the destination's capacity, sustainability conditions and KPI-based performance.

Within the overall TOURAL framework, the Flow Capacity Tool serves as the predictive component that estimates future tourism demand and generates the destination-state information subsequently used by the Predictive Advisor for optimisation and recommendation generation.

4.1.2 Main Outputs of the Simulation

The Flow Capacity Tool produces a structured set of outputs that allow the user to understand the simulated state of the selected destination. The main outputs include:

- predicted tourist flow for the selected destination and month
- qualitative flow class, such as low, medium or high
- overall sustainability score for the simulated state
- category-level scores for environmental, social-cultural and economic dimensions
- detailed KPI values and comparisons
- status labels and pressure signals indicating satisfactory, moderate or critical conditions and
- dashboard-level summaries that support interpretation of the simulated tourism state.

These outputs are designed to be interpretable for non-technical users. The numerical prediction is complemented by qualitative classifications, KPI status labels and explanatory messages, allowing users to understand both the expected level of tourism activity and its potential implications for the destination.

4.2 Digital Representation Model of the Destination

The tool follows a digital-twin-oriented logic, where each pilot destination is represented through a structured combination of tourism demand, contextual characteristics, infrastructure capacity and sustainability-related KPIs. This representation allows the system to simulate the behaviour of the destination under different temporal conditions and to translate predicted flows into meaningful sustainability outputs.

As illustrated in Figure 3, the digital-twin representation of the Flow Capacity Tool is organised around three main components: tourism flow, destination capacity and sustainability performance. These components form the basis through which the tool connects predicted flows with capacity interpretation and KPI-based assessment. The following subsections explain each component and how it contributes to the simulated destination state.

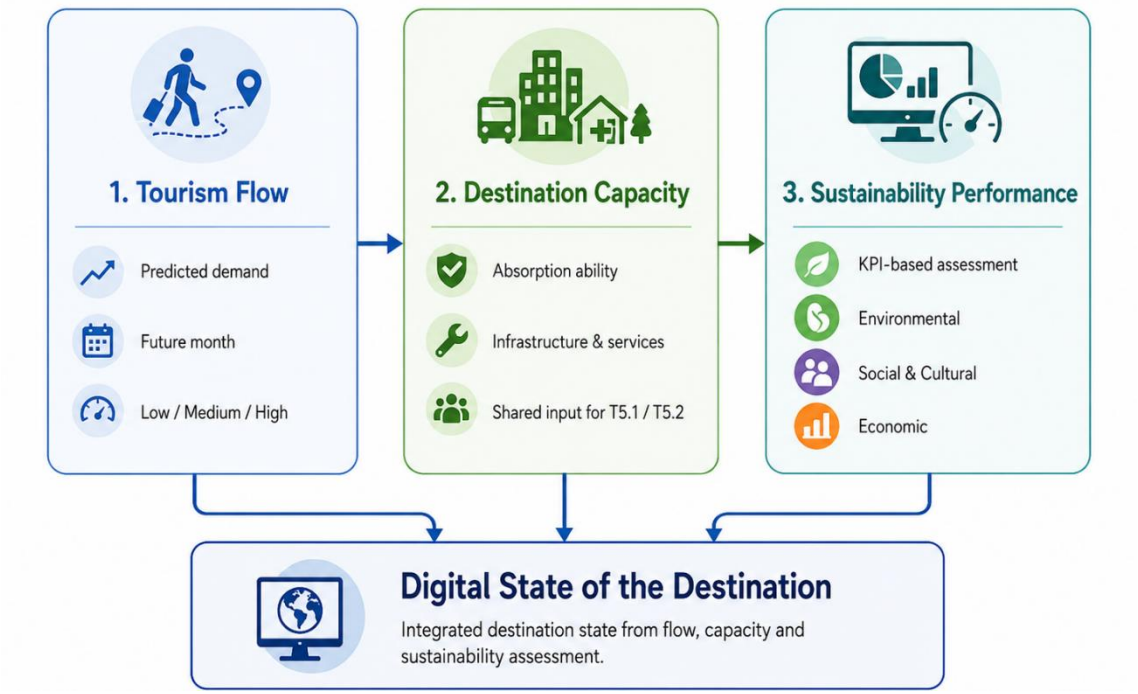

Figure 3: Main system components of the digital-twin representation

4.2.1 Tourism Flow Layer

The tourism flow layer represents the expected level of tourism activity in the selected destination for a specific future month. The central variable of this layer is the predicted tourist flow, which indicates the forecasted number of tourists for the selected month.

The predicted tourist flow is also translated into a qualitative flow class: low, medium or high. This classification supports interpretation by allowing users to understand whether the predicted tourism demand corresponds to a limited, moderate or intense level of activity. The tourism demand layer is the starting point of the simulation process, as it defines the level of tourism pressure that will be assessed by the rest of the system.

4.2.2 Destination Capacity Layer

The destination capacity layer represents the ability of the destination to accommodate the predicted tourism demand. It includes the structural and contextual characteristics that influence how tourism pressure is absorbed by the area. These characteristics include accommodation capacity, resident population, transport accessibility, local services, hospitality infrastructure, health-related infrastructure and other destination-specific factors.

Capacity is interpreted as the destination's ability to absorb incoming tourism flows without creating excessive pressure on infrastructure, services, local communities, environmental resources or cultural assets. This component is informed by data collected through pilot surveys, pilot templates and destination-specific inputs. These inputs describe the current characteristics and limitations of each destination, including accommodation availability, infrastructure readiness, accessibility conditions, local services, environmental sensitivity and other factors that influence how much tourism activity can be sustainably absorbed.

In the Flow Capacity Tool, the capacity layer is not treated as a stand-alone interpretation module. Instead, it is used as an interpretive layer that helps the system determine whether predicted tourism flows are manageable or likely to generate pressure. It therefore supports capacity monitoring and pressure detection within the simulation process.

4.2.3 Sustainability Performance Layer

The sustainability performance layer represents the expected condition of the destination across the main sustainability dimensions. In the Flow Capacity Tool, sustainability performance is organised into three dynamic KPI categories: Environmental Metrics, Social and Cultural Impact and Financial and Economic Performance. These categories allow the system to evaluate how the predicted tourist flow may affect different aspects of the destination.

The Environmental Metrics category captures indicators related to resource use, emissions, waste generation, air quality, water consumption and other environmental pressures. The Social and Cultural Impact category captures indicators related to tourism intensity in relation to residents, accessibility, cultural pressure, service access and social conditions. The Economic Performance category captures indicators related to tourist nights, occupancy, revenue and other economic effects of tourism activity.

Destination Management indicators are not treated as a separate dynamic category in the current monthly simulation logic because many of these indicators are relatively stable over short-term horizons. Indicators such as environmental certifications, Blue Flag beaches or destination-management arrangements usually change through policy decisions, certification processes, infrastructure upgrades or management interventions, rather than as an immediate response to monthly changes in tourist flow. For this reason, they are used as contextual or baseline characteristics where relevant, rather than as dynamic monthly simulation outputs.

4.3 Simulation Inputs and Feature Space

The implemented tool uses a feature space structured according to the role of each variable in the simulation pipeline. Some features are used directly by the tourism-flow prediction model, while others are used after the flow-prediction stage to calculate dynamic KPIs and interpret the simulated destination state.

This distinction is important because the Flow Capacity Tool does not use all variables for the same purpose. The first part of the pipeline estimates the expected number of tourists for a selected destination and future month. The second part uses this predicted flow, together with destination-specific parameters and KPI reference values, to calculate or update flow-sensitive sustainability indicators.

Table 3 summarizes the main feature groups used by the Flow Capacity Tool, their role in the simulation pipeline, indicative variables and the outputs they support. The table distinguishes between variables used directly by the tourism-flow prediction model, variables used after prediction for dynamic KPI calculation, and stable or contextual reference values used for KPI interpretation. The remainder of this section explains these elements in more detail, starting from the scenario inputs provided by the user, continuing with the features used for tourism-flow prediction, and concluding with the dynamic KPI calculation and feature-construction logic.

Table 3: Feature groups by modelling function

Feature group	Role in the pipeline	Indicative features	Main outputs supported
Flow-prediction features	Used directly by the AI-based monthly forecasting model to estimate tourist flow.	Monthly time index, seasonality encodings, historical flows, lagged flows, rolling-flow indicators, flight arrivals; ferry activity, airport/ferry passengers, travel times to ports/airports, road-network density, weather variables, air-quality indicators, pollution indicators, public holidays, bridge days, Google Trends, Wikipedia page views.	Predicted tourist flow, qualitative flow class.

Dynamic-KPI calculation features	Used after flow prediction to calculate or estimate flow-sensitive KPIs and interpret KPI status.	Predicted tourist flow, resident population, bed capacity, accommodation capacity, length-of-stay coefficients, spending-per-night assumptions, CO2 intensity parameters, water-consumption parameters, waste-generation parameters, ETIS-based Excel KPI reference values, pilot inputs, stable KPI reference indicators.	Tourist nights, average length of stay, occupancy rate, tourism revenue per capita, tourists per 100 residents, waste production, CO2 emissions, water consumption, KPI status interpretation.
Stable and contextual reference values	Used as baseline or contextual reference values, optionally selected or interpreted according to the simulated flow class.	Destination-management, accessibility, cultural, green-transport, protected-area, biodiversity, light-pollution and noise-related indicators; ETIS-based Excel KPI reference values; pilot inputs.	KPI status interpretation, complete KPI representation across stable dimensions.

4.3.1 Scenario Inputs

Scenario inputs are the parameters provided by the user through the interface. These inputs define the simulation request and are used by the backend service to retrieve the appropriate destination data, model configuration and KPI definitions. The main scenario inputs are:

- selected region or pilot destination,
- selected year and,
- selected month,

These inputs are passed to the prediction service, which resolves the corresponding destination data, preprocessing logic, model configuration, KPI definitions and scoring rules. This ensures that the simulation is executed consistently for the selected destination and month.

4.3.2 Features for Tourism-Flow Prediction

Flow-prediction features are the variables used by the AI-based monthly forecasting model to estimate the expected tourist flow for a selected destination and future month. These features describe the temporal, historical, mobility-related, environmental, calendar-based and digital-interest conditions that may influence tourism demand.

Temporal and seasonality features describe the position of the target month within the annual tourism cycle. They include the monthly time index, cyclical month encodings, peak-season and shoulder-season indicators, weekend counts and other calendar-related variables. Their role is to help the model capture recurrent seasonal patterns in tourism demand.

Historical-flow features represent previous tourism-flow behaviour through lagged tourist-flow values and rolling-flow indicators. They allow the model to account for demand persistence, recent trends and changes in visitor activity over time. This is important because the expected flow of a future month is strongly influenced by the behaviour observed in previous months.

Transport, mobility and accessibility features describe how accessible the destination is and how visitor movement may evolve. Indicative features include observed flight arrivals, ferry activity, airport passenger counts, ferry passenger counts, travel times to ports and airports, and road-network density where available. These variables provide information on mobility conditions that may affect the expected number of visitors.

Weather and environmental features include mean temperature, precipitation, air-quality indicators and pollution-related indicators. They are used because weather conditions and environmental quality can influence destination attractiveness, travel decisions and seasonal demand patterns.

Calendar-effect features include public holidays, bridge days and related calendar variables. These variables represent short-term demand effects that may increase or reduce tourism activity during specific months.

Digital demand-proxy features capture online interest in the destination through sources such as Wikipedia page views and Google Trends indices. They provide indirect signals of tourism demand and may help identify changes in interest before these changes appear in official tourism statistics.

Together, these feature families form the active input space of the trained tourism-flow prediction model. They are used to estimate the expected tourist flow for the selected destination and month. The predicted flow is then passed to the next stages of the pipeline, where it is classified into a low, medium or high flow class and used by the KPI calculation and interpretation layer.

4.3.3 Dynamic KPI Calculation and Feature Construction

Dynamic KPI calculation features are used after the tourism-flow prediction stage. Their purpose is to translate the predicted tourist flow into sustainability-related indicators that change according to the simulated tourism state. In this part of the pipeline, the predicted tourist flow acts as the main driver connecting the AI-based prediction stage with the downstream calculation and interpretation of flow-sensitive KPIs.

The predicted flow is combined with destination-specific parameters, including resident population, bed capacity, accommodation capacity, average length-of-stay coefficients, spending-per-night assumptions, CO_2 intensity parameters, water-consumption parameters and waste-generation parameters. These parameters support the calculation or estimation of KPIs that are directly affected by changes in tourism intensity, such as tourist nights, occupancy rate, tourism revenue per capita, tourists per 100 residents, waste production, tourism-related CO_2 emissions and water consumption.

The same layer also uses KPI reference values loaded from ETIS-based Excel templates and pilot-provided inputs. These values support comparison, interpretation and status classification. Stable indicators, including destination-management, accessibility, cultural, green-transport, protected-area, biodiversity, light-pollution and noise-related indicators, are not re-predicted every month by the LSTM model. Instead, they are used as baseline or contextual reference values. Where relevant, these indicators are selected or interpreted according to the simulated tourism-flow class, namely low, medium or high. This approach allows the tool to maintain a complete KPI representation even when specific indicators are not directly modelled as monthly dynamic outputs.

The feature space supporting this process is constructed from multiple pilot-provided and external data sources. Pilot inputs and Excel-based KPI templates provide destination-specific parameters, capacity information and KPI reference values. External data services provide additional temporal, mobility, weather, environmental and demand-proxy variables that complement the pilot-provided inputs.

During feature construction, raw data are cleaned, harmonised and aligned to the monthly simulation horizon. Time-dependent variables are aggregated or encoded at monthly level, historical-flow variables are transformed into lagged and rolling indicators, calendar variables are converted into counts or flags, and destination-level parameters are stored as static metadata or used in formula-based KPI calculations. Through this process, heterogeneous input data are transformed into model-ready features that support both tourism-flow prediction and KPI-based sustainability interpretation.

4.4 Simulation Pipeline and Tourism Flow Prediction

The simulation pipeline of the Flow Capacity Tool describes how a user-defined scenario is transformed into predicted tourism-flow and KPI-related outputs. The pipeline connects the user interface, destination data, data-preprocessing steps, the AI prediction model and the sustainability-impact interpretation layer.

The purpose of the pipeline is to provide a structured and repeatable process for estimating future tourist flows and translating them into meaningful destination-level results. The process starts when the user selects a destination and a future month through the interface. The system then retrieves the relevant data, prepares the required model inputs, executes the AI-based prediction process and updates the dashboard with the predicted flow, the flow class, the overall sustainability score, category-level scores and detailed KPI values. This process ensures that each simulation follows the same logic, from scenario selection to dashboard presentation. In this way, the tool provides a consistent workflow for analysing future tourism pressure across different pilot destinations and months. Figure 4 presents the proposed methodology, while Figure 5 illustrates the front-end scenario input interface of the Flow Capacity Tool, where users select the destination and time period that define the simulation request.

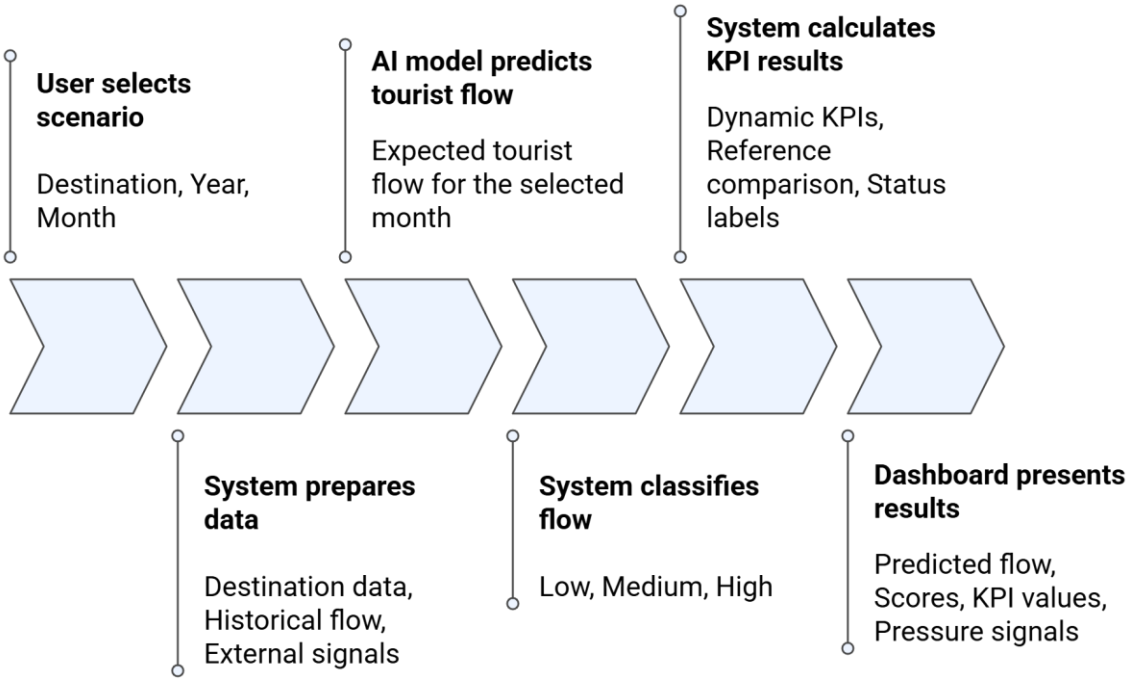

Figure 4: Simulation pipeline and Tourism flow prediction tool

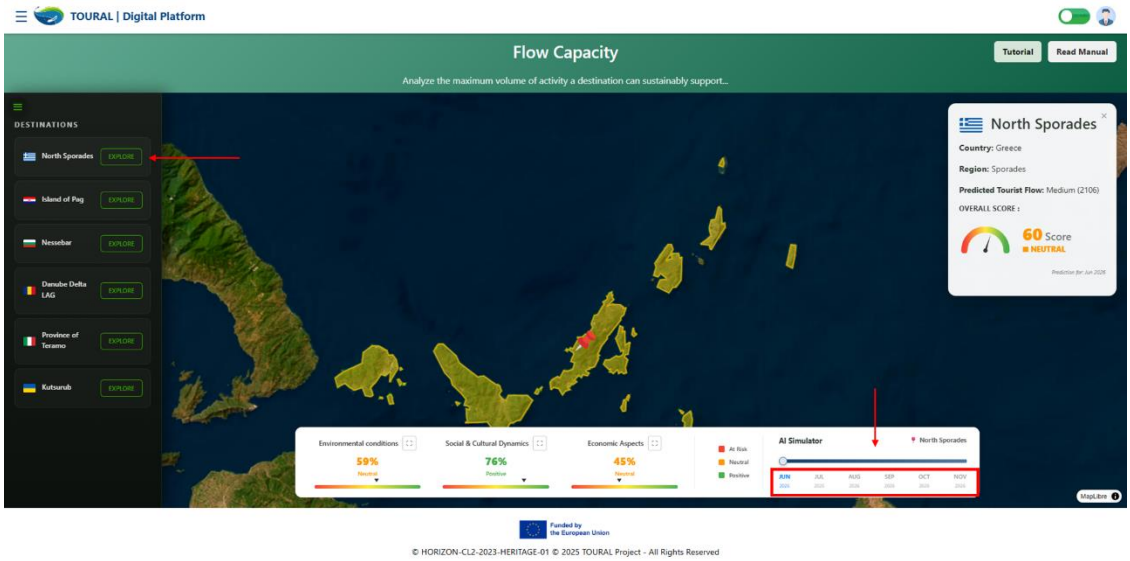

Figure 5: Flow Capacity Tool front-end interface for scenario input selection, including destination and time-period parameters.

4.4.1 Preparing the Scenario Inputs

The first stage of the pipeline transforms the scenario submitted by the user into a structured input for the prediction model. The scenario, namely the selected region or pilot destination, year and month, is defined in Section ‘’Scenario Inputs”. The feature groups used by the model are described in Section ‘’Features for Tourism-Flow Prediction”. This subsection focuses on how the system prepares the required inputs for the selected destination and target month.

When a scenario is submitted, the prediction service identifies the corresponding destination data, preprocessing configuration and model setup. It retrieves the available monthly feature history for the selected destination and prepares the input sequence required by the forecasting model. This sequence contains the recent historical context needed to estimate the tourist flow of the target month. To ensure consistency between model training and model use, the same preprocessing steps are applied during deployment. Raw variables are cleaned, harmonised and aligned to the monthly simulation horizon, following the feature-construction logic described in Section ‘’Dynamic KPI Calculation and Feature Construction”. Time-dependent variables are encoded at monthly level, historical tourism-flow values are represented through lagged and rolling indicators, and external signals are aligned with the selected destination and target period.

The output of this stage is a structured model input that describes the selected destination, the target month, recent demand behaviour, seasonal conditions and relevant external drivers. This input is then passed to the AI-based tourism-flow prediction model.

4.4.2 AI-Based Tourism Flow Prediction

The tourism-flow prediction component estimates the expected number of tourists for the selected destination and target month. This is the main predictive function of the Flow Capacity Tool. The model uses the prepared input data to generate a monthly prediction of tourist flow.

The selected method is based on a hybrid Convolutional Neural Network and Bidirectional Long Short-Term Memory architecture (CNN-BiLSTM). This architecture was selected because tourism demand is influenced by both short-term temporal patterns and sequential dependencies across previous months. The CNN component is used to extract local temporal patterns from the monthly input sequence, while the BiLSTM component captures forward and backward dependencies in the historical data. In this way, the model can learn how tourism activity evolves over time by considering recent demand behaviour, seasonal effects, and contextual drivers.

In the standard configuration, the model uses the previous six months of available data as the input sequence for forecasting. This six-month input window allows the model to capture recent tourism trends, short-term seasonality, and changes in visitor behaviour. Where a complete six-month history is not available, the system can operate with the shorter available history, applying the same preprocessing and prediction logic.

The CNN-BiLSTM model receives monthly features such as historical tourism-flow values, seasonality indicators, transport and mobility signals, weather variables, public-holiday information, and digital demand-proxy signals. These inputs allow the model to estimate tourism flow by considering both the internal structure of the tourism time series and external factors that may influence visitor demand.

Special treatment is applied to observations from the COVID-19 period, because tourism flows during this period were affected by exceptional travel restrictions and abnormal mobility conditions. These observations do not represent normal tourism behaviour and may distort the learning process if treated as standard historical data. For this reason, COVID-19-period observations are assigned lower training weights. This reduces the influence of these abnormal months during model training, while still allowing the model to use the available historical information.

The output of the CNN-BiLSTM model is a monthly tourism-flow prediction for the selected destination and forecast horizon. These predicted values are then passed to the next stages of the pipeline, where they are classified into qualitative flow classes and used by the KPI calculation and capacity-monitoring layer. Therefore, the model does not operate as an isolated forecasting module, but as the predictive core of the broader Tourism Flow Simulation and Capacity Management Tool.

4.4.3 Flow Classification and Dashboard Outputs

After the numerical tourist-flow prediction is generated, the predicted value is translated into a qualitative flow class, such as low, medium or high. This classification helps users understand whether the expected level of tourism activity is limited, moderate or intense.

The flow class is defined using historical quantile thresholds or destination-specific capacity thresholds from the destination configuration. Through this process, the predicted flow is not presented only as a numerical value, but also as an interpretable tourism-pressure level. The flow class is then used by the KPI layer to select and interpret the relevant KPI values for the simulated tourism state.

Once the predicted flow and its corresponding flow class are available, the system performs a mapping between the simulation result and the KPI list. This mapping combines two types of KPI information. The first type includes dynamic KPIs, which are directly affected by changes in tourist flow. These indicators are calculated or updated using the predicted tourist flow together with destination-specific parameters.

The second type includes stable or contextual KPIs, which are not recalculated every month by the prediction model. These indicators are retrieved from the KPI configuration or ETIS-based Excel templates and are selected according to the destination and, where relevant, the simulated flow class. Examples include destination-management indicators, accessibility-related indicators, cultural indicators, green-transport indicators, protected-area indicators, biodiversity-related indicators, light-pollution indicators and noise-related indicators. In this way, the tool maintains a complete KPI representation even when not all indicators are dynamically predicted.

The resulting KPI vector therefore combines flow-sensitive dynamic indicators with stable reference indicators. The combined KPI values are then compared with their reference or target values in order to produce KPI scores, status labels and pressure signals. This step allows the

system to interpret how the predicted tourist flow may affect the sustainability condition of the selected destination.

The outputs of the Flow Capacity Tool are presented through dashboard and reporting elements that translate model results into a form that can be reviewed by project partners, pilot representatives and destination stakeholders. The dashboard presents both numerical and qualitative information, allowing users to understand the predicted tourist-flow level and its expected impact on sustainability indicators.

The main reporting outputs include the predicted tourist flow, flow class, overall score, sustainability-dimension scores, KPI-level values, KPI status labels and pressure signals. These outputs make the simulation results easier to interpret and provide a structured basis for discussing tourism pressure, capacity constraints and sustainability conditions within each pilot destination.

The reporting outputs are interpretive in the sense that they help users identify where attention is required. The Flow Capacity Tool itself remains focused on simulation and monitoring. Its role is to simulate tourism flows, map the predicted flow to the appropriate KPI representation, monitor capacity-related pressure and present interpretable sustainability outputs based on the selected destination and month. Figure 6 illustrates how the Flow Capacity Tool translates the predicted tourism flow into a qualitative flow class and presents the corresponding KPI-based dashboard outputs.

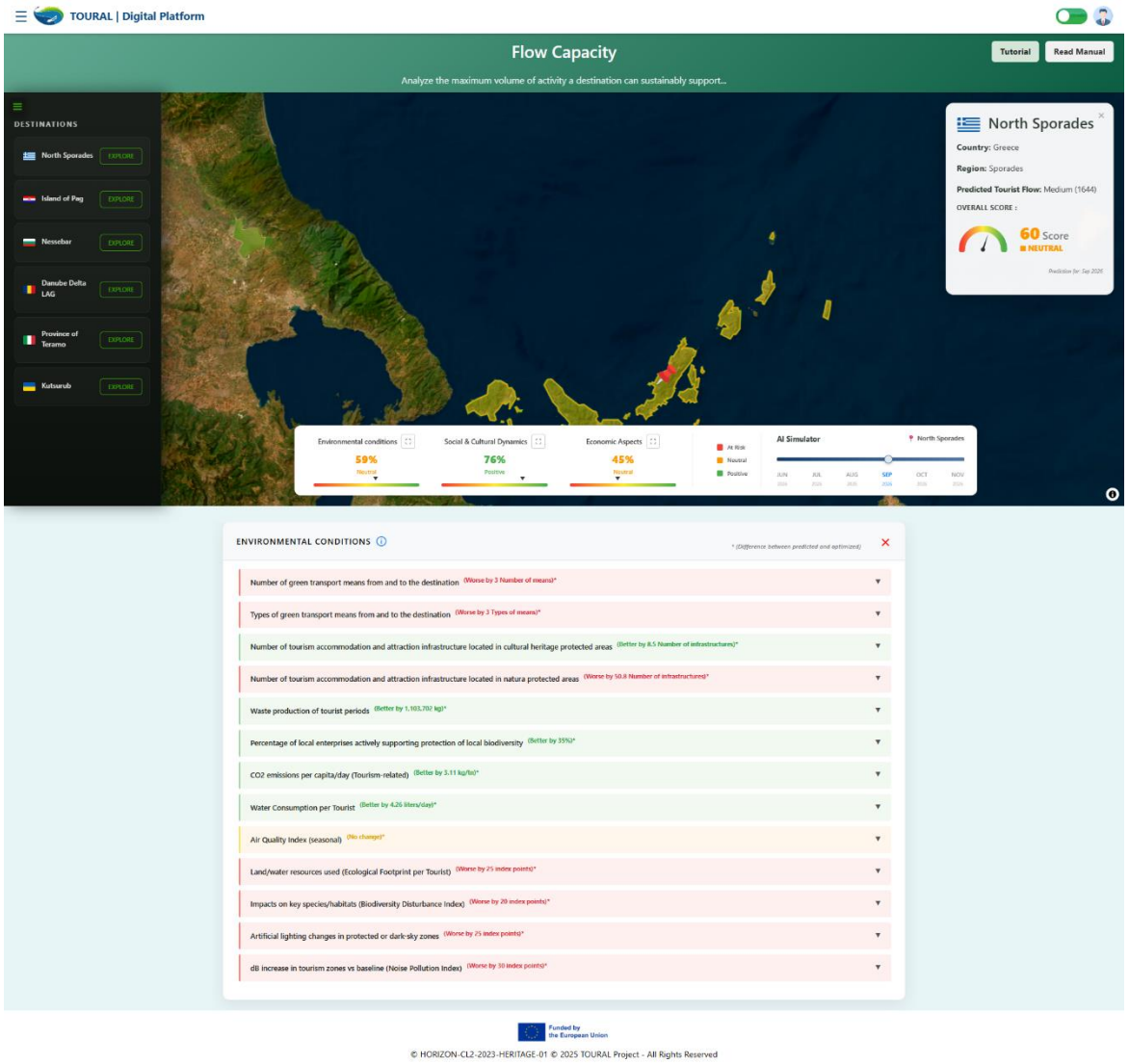

Figure 6: Flow Capacity Tool dashboard showing flow-class summary and indicator-level sustainability outputs.

4.5 Impact Modelling, Capacity Monitoring and Pressure Detection

After the monthly tourist flow is predicted and classified into a low, medium or high flow class, the Flow Capacity Tool translates the simulated tourism state into KPI-based sustainability and capacity-monitoring outputs. This step links the expected number of tourists with destination-specific characteristics, capacity parameters, KPI definitions and reference values. Its purpose is to identify how the predicted flow may affect the destination and where pressure may emerge across environmental, social-cultural and economic dimensions.

The impact interpretation layer does not treat the predicted flow as an isolated forecasting result. Instead, the predicted value is used as the main driver for updating or selecting the relevant KPI values. Dynamic KPIs are calculated or estimated using the predicted tourist flow and destination-specific parameters, while stable or contextual KPIs are retrieved from the KPI configuration and interpreted according to the selected destination and, where relevant, the simulated flow class. The resulting KPI vector provides a complete representation of the simulated destination state.

4.5.1 Sustainability Impact and Capacity Monitoring by Dimension

The simulation framework models the relationship between tourism activity and sustainability outcomes through a set of indicators structured across environmental, social-cultural and economic dimensions. Tourism flows are translated into system impacts by capturing how increased visitor intensity affects resource consumption and environmental pressure, how concentration of activity influences social conditions such as crowding and accessibility, and how tourism demand contributes to economic performance and distribution of benefits.

Each sustainability dimension is evaluated through a set of KPIs, enabling the identification of imbalances and trade-offs across different aspects of the system. This modelling approach allows the system to move beyond simple flow estimation and to provide a holistic view of tourism sustainability.

Capacity monitoring is performed by comparing the predicted demand and the derived KPI values with destination-specific parameters, reference values and scoring rules. For example, predicted tourist flow may be compared with accommodation capacity to assess occupancy pressure, with resident population to assess community pressure, or with environmental coefficients to estimate waste, water consumption and emissions impacts. This allows the tool to identify high-pressure periods, potential over-tourism conditions, infrastructure or service bottlenecks, and sustainability dimensions that may become vulnerable under the simulated flow level.

Table 4 summarises how the predicted tourism flow is interpreted across the three main sustainability dimensions. It presents the impact logic used by the simulation and provides indicative KPI outputs for each dimension.

Table 4: Sustainability impact logic by dimension

Sustainability dimension	Impact logic in the simulation	Indicative KPI outputs
Environmental Conditions	Predicted flows are linked with resource consumption, emissions, waste generation and environmental pressure. Higher tourism intensity may increase pressure on water, waste-management systems, air quality and ecological resources.	Waste production during tourist periods; tourism-related CO2 emissions per capita/day; water consumption per tourist; air-quality indicators; ecological and biodiversity-related reference indicators.
Social & Cultural Dynamics	Predicted flows are interpreted in relation to resident population, accessibility, service availability, cultural assets and community pressure. Higher tourism intensity may increase crowding, pressure on local services and stress on heritage spaces.	Tourists per 100 residents; access to services; heritage site stress; safety level; accessibility indicators; cultural inclusivity and social-inclusion reference indicators.
Economic Aspects	Predicted flows are connected with tourist nights, occupancy, revenue and tourism-related economic activity. The simulation captures both the potential economic contribution and the pressure created by seasonal concentration.	Number of tourist nights per month; average length of stay; occupancy rate; tourism revenue per capita; direct tourism employment reference values; revenue-leakage related indicators.

4.5.2 KPI Scoring, Status Classification and Pressure Signals

Stable KPIs are merged with the dynamic KPI outputs to produce the full KPI vector. Stable values are read from the KPI configuration for the relevant destination, month or flow class. Dynamic values are either model-based predictions or formula-derived estimates. The combined KPI vector is then compared with target or reference values defined in the KPI configuration and scoring logic.

The comparison is direction-aware. Some indicators are better when their value is higher, such as visitor satisfaction or local economic retention. Other indicators are better when their value is lower, such as waste production, CO_2 emissions, water consumption or tourists per 100 residents. Some indicators may also be interpreted against a target value or acceptable range.

The scoring method produces KPI scores, status labels, gap percentages, delta values and human-readable messages, such as “Better by”, “Worse by” or “No change”. These outputs allow the dashboard to show the predicted tourist flow, the overall condition of the destination, category-level scores, detailed KPI cards and pressure signals that support capacity monitoring.

Because the selected KPIs are expressed in different units and scales, raw values are not averaged directly. For example, one KPI may be expressed as a percentage, another as kilograms of waste, another as litres of water, another as euros per capita, and another as the number of tourists per 100 residents. To make these indicators comparable, each raw KPI value is first converted into a normalised score between 0 and 100. This score expresses how close the simulated KPI value is to the optimal or reference value defined for the relevant tourism-flow class: low, medium or high.

For each KPI i, the platform compares the simulated or predicted value V_i with the corresponding optimal or reference value O_i. The calculation then applies the appropriate scoring rule according to the direction and interpretation of the KPI.

$$V_i = \text{simulated or predicted value of KPI i}$$

$$O_i = \text{optimal or reference value of KPI i}$$

For higher-is-better indicators, the KPI score is calculated as:

$$Score_i = 100 \times \frac{V_i}{O_i}$$

For lower-is-better indicators, the KPI score is calculated as:

$$Score_i = 100 \times \frac{O_i}{V_i}$$

For target-based indicators, where the objective is to remain close to a specific reference value, the score is calculated according to the distance from the optimal value:

$$Score_i = 100 - \left|100 \times \left(\frac{(V_i - O_i)}{O_i}\right)\right|$$

In all cases, the resulting score is clipped to the range 0-100:

$$Score_i = min(100, max(0, Score_i))$$

This means that a KPI receives a score of 100 when it reaches the optimal condition or performs better than the reference condition, depending on the KPI direction. Lower scores indicate increasing deviation from the desired benchmark. If the simulated value is better than the optimal value, the score is capped at 100 in order to preserve a common interpretation scale across all indicators.

The platform also calculates the difference between the simulated and reference values. The absolute difference is expressed as a delta value, while the relative difference is expressed as a gap percentage:

$$\Delta_i = V_i - O_i$$

$$Gap_i(\%) = 100 \times \left(\frac{(V_i - O_i)}{O_i}\right)$$

The interpretation of the delta and gap values is also direction-aware. For higher-is-better indicators, a positive difference indicates that the simulated value is better than the reference value, while a negative difference indicates underperformance. For lower-is-better indicators, the interpretation is reversed, because values below the reference condition represent better performance. This logic is used to generate the human-readable messages displayed in the dashboard, such as "Better by 3.6 litres/day", "Worse by 20%" or "No change". The unit used in each message is inherited from the corresponding KPI definition.

After each valid KPI has been converted into a normalised score, the platform calculates the average score for each sustainability section. For a section containing n valid KPIs, the section score is calculated as a weighted average:

$$SectionScore = \frac{(\sum_{i=1}^{n} w_i \times Score_i)}{\sum_{i=1}^{n} w_i}$$

- $Score_i$ is the normalised score of KPI i.
- w_i is the weight assigned to KPI i.
- n is the number of valid KPIs included in the section.

In the current configuration, most KPI weights are equal to 1.0. Therefore, the weighted average behaves as an arithmetic mean:

$$SectionScore = \frac{Score_1 + Score_2 + \cdots + Score_n}{n}$$

The Environmental Conditions score is calculated as:

$$S_env = \frac{(\sum_{i=1}^{n} w_i \times D_i)}{\sum_{i=1}^{n} w_i}$$

where D_i represents the normalised scores of the environmental KPIs.

The Social and Cultural Dynamics score is calculated as:

$$S_soc = \frac{(\sum_{j=1}^{n} w_j \times C_j)}{\sum_{j=1}^{n} w_j}$$

where C_j represents the normalised scores of the social and cultural KPIs.

The Economic Aspects score is calculated as:

$$S_\text{econ} = \frac{(\sum_{k=1}^{n} w_k \times B_k)}{\sum_{k=1}^{n} w_k}$$

where B_k represents the normalised scores of the economic KPIs.

Only valid KPI scores are included in the aggregation. Missing, not applicable or non-comparable values are excluded from the denominator so that they do not artificially increase or reduce the section score. This is particularly important for KPIs that may not be available for all destinations or for all tourism-flow classes.

The final overall sustainability score is calculated as the average of the valid dynamic sustainability-section scores. In the current implementation, the Destination Management section is excluded from the monthly overall score because several of its indicators are relatively stable and do not change directly as a result of monthly tourism-flow variation. Therefore, the overall score is calculated using the Environmental Conditions, Social and Cultural Dynamics, and Economic Aspects sections:

$$OverallScore = \frac{S_env + S_soc + S_econ}{3}$$

To support interpretability, the system includes a qualitative alert mechanism based on KPI status. Indicators are classified into categories reflecting their condition, such as satisfactory performance, moderate deviation, and critical or at-risk conditions. This classification enables users to quickly identify problematic areas and prioritise further examination.

The alert mechanism does not rely on a single metric, but on the combined behaviour of multiple indicators. This provides a more robust view of system performance because tourism pressure may appear differently across environmental, social-cultural and economic dimensions. For example, a simulated month may show positive economic performance while simultaneously creating pressure on water consumption, waste generation or residents' access to services. Table 5 presents the status categories used to communicate KPI conditions and pressure signals in the dashboard.

Table 5: KPI status categories and dashboard interpretation

Status category	Interpretation	Indicative use in the dashboard
Satisfactory	The KPI value is within an acceptable range or shows limited deviation from the reference condition.	Displayed as a positive or stable condition; no immediate pressure signal is required.
Moderate deviation	The KPI value shows a relevant deviation from the reference condition, but the pressure is not yet critical.	Displayed as a warning or monitoring signal; the user can examine the related indicator and period.
Critical / at-risk	The KPI value indicates strong pressure, threshold exceedance or a potentially unsustainable condition.	Displayed as a critical pressure signal; the user can identify the affected sustainability dimension and KPI.

4.6 Impact Modelling per Sustainability Dimension

Beyond the estimation of tourism flows, the Flow Capacity Tool evaluates the potential implications of tourism activity across the sustainability dimensions adopted within the TOURAL framework. Building upon the sustainability assessment logic established in WP3, the simulator translates tourism demand patterns, visitor distribution, and destination utilisation levels into measurable sustainability outcomes.

The modelling process considers how variations in tourism activity influence environmental conditions, economic performance, and social and cultural well-being within the destination. Rather than focusing on a single performance measure, the simulator evaluates multiple indicators simultaneously, enabling the identification of interactions, imbalances, and trade-offs between different sustainability objectives.

This multidimensional assessment approach allows decision-makers to examine the broader consequences of tourism development scenarios and supports a more holistic understanding of destination sustainability. Consequently, the simulator provides not only estimates of tourism activity but also an evaluation of its potential impacts across the environmental, economic, social, and cultural dimensions represented within the UCCST framework.

Table 6: Sustainability dimensions, representative indicators, and example impacts considered by the Flow Capacity Tool

Sustainability Dimension	Representative Indicators	Examples of Assessed Impacts
Environmental	Environmental KPIs, resource utilisation indicators, environmental pressure indicators	Increased resource consumption, pressure on natural ecosystems, environmental degradation risks, seasonality-related environmental stress
Economic	Tourism revenue indicators, occupancy indicators, tourism intensity indicators, employment-related KPIs	Changes in tourism revenue, economic benefits for local businesses, tourism-related employment effects, economic leakage risks
Social	Accessibility indicators, safety indicators, community-related KPIs, service-access indicators	Visitor pressure on local services, accessibility challenges, effects on resident quality of life, social inclusion considerations
Cultural	Heritage-related indicators, cultural-event indicators, cultural inclusivity indicators	Heritage-site stress, impacts on cultural identity, preservation of local traditions, cultural participation and inclusiveness effects

The Table 6 illustrates how tourism activity is translated into sustainability-related consequences within the simulation environment. Rather than evaluating tourism performance exclusively through visitor volumes or economic metrics, the simulator assesses impacts across the environmental, economic, social, and cultural dimensions established in the UCCST framework. This multidimensional perspective enables the identification of trade-offs and supports a balanced evaluation of alternative tourism development scenarios.
The resulting impact estimates can be utilised within T4.2 to assess the sustainability implications of alternative regional business models and support the selection of tourism development strategies that maximise local benefits while minimising environmental and social pressures.

4.7 Tourism Flow Simulation and Destination State Estimation

The primary objective of the Flow Capacity Tool is to estimate how tourism activity influences the overall state of a destination under different operating conditions. Building upon the Destination Digital Twin, the simulator combines destination characteristics, tourism demand patterns, capacity-related information, and sustainability indicators to generate a dynamic representation of destination behaviour.

The simulation process begins with the definition of a destination state, consisting of tourism demand characteristics, tourism supply conditions, infrastructure availability, accessibility

factors, seasonal dynamics, and sustainability-related parameters. Based on this representation, the simulator estimates tourism flows and their distribution across the destination, taking into account both current conditions and alternative scenario configurations.

Rather than focusing solely on visitor numbers, the simulation process evaluates how tourism activity propagates through the destination ecosystem and influences the different components represented within the Digital Twin. This enables the generation of future destination states that reflect potential changes in tourism intensity, spatial distribution patterns, infrastructure utilisation, and sustainability conditions.

The resulting destination state serves as the analytical foundation for subsequent impact modelling, capacity monitoring, pressure detection, and scenario evaluation activities. In this way, tourism flow simulation acts as the core mechanism through which destination dynamics are translated into measurable sustainability outcomes and decision-support information.

4.8 Evaluation & Results

This section presents the evaluation of the Tourism Flow Simulation and Capacity Management Tool. The evaluation is organised into two complementary parts. The first part focuses on the predictive performance of the tourism-flow forecasting component, with emphasis on the comparison between the selected CNN-BiLSTM model and alternative benchmark models. The second part focuses on the impact results presented in the platform, where predicted tourism flows are translated into sustainability scores, KPI-level outputs, and capacity-related pressure signals for each pilot destination.

The evaluation therefore does not assess only the forecasting model in isolation. Instead, it follows the operational logic of the tool: first, the model estimates the expected tourist flow for a selected destination and month. Then, the predicted flow is classified and propagated to the KPI and capacity-monitoring layer. This approach allows the evaluation to capture both the technical accuracy of the prediction model and the practical usefulness of the tool for destination-level sustainability monitoring.

4.8.1 Tourism-Flow Prediction Performance

The tourism-flow prediction component was evaluated using pilot-specific monthly data. In the main body of the deliverable, the evaluation focuses on the North Sporades pilot, with emphasis on Alonissos. The corresponding evaluation results for the Danube Delta pilot are provided in Annex. The purpose of this evaluation was to compare the selected CNN-BiLSTM architecture with alternative forecasting and regression-based approaches and to examine whether the inclusion of the extended feature space improves predictive performance compared with a more limited input configuration. The model produces monthly tourism-flow predictions over a six-month forecasting horizon, allowing the tool to estimate how tourism demand may evolve across the upcoming months rather than providing only a single isolated monthly forecast.

The comparison was carried out using standard regression metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). MAE and RMSE measure the absolute prediction error in the same unit as the target variable, while MAPE expresses the error as a percentage of the observed value. Lower values of MAE, RMSE, and MAPE indicate better predictive performance. R^2 is used as an additional goodness-of-fit indicator, where higher values indicate stronger explanatory performance. Table 7 presents the details related to the evaluation process.

Table 7: Evaluation setup for tourism-flow prediction

Evaluation element	Description
Target variable	Monthly estimated tourist arrivals / tourism flow
Pilot destinations	North Sporades (Alonissos)
Forecasting horizon	One-month-ahead prediction
Main model	Bidirectional Long Short-Term Memory model
Benchmark models	Random Forest, XGBoost, HistGradientBoosting
Input-window configuration	Previous monthly observations used as sequential input
Evaluation metrics	MAE, RMSE, MAPE, R^2
Feature configurations	Single-feature input and full-feature input

The first comparison evaluates the CNN-BiLSTM architecture against the alternative benchmark models. This comparison is important because it shows whether the sequence-based deep learning model provides an advantage over more conventional regression-based approaches. Table 8 presents a comparison of the tourism-flow prediction models.

Table 8: Comparison of tourism-flow prediction models for North Sporades

Destination	Model	MAE	RMSE	MAPE	R^2
North Sporades	Random Forest	161.4697	247.9139	8.744699	0.947053
North Sporades	XGBoost	110.1671	139.0023	13.62742	0.983355
North Sporades	HistGradientBoosting	131.2084	156.0325	15.21311	0.979027
North Sporades	**CNN-BiLSTM**	**90.12965**	**123.7407**	**6.102235**	**0.98681**

For the North Sporades pilot, the CNN-BiLSTM model achieved the best performance across all four-evaluation metrics. It obtained the lowest MAE, RMSE, and MAPE values and the highest R^2 score. Compared with the strongest benchmark model in terms of absolute error, namely XGBoost, the CNN-BiLSTM reduced MAE from 110.1671 to 90.1297 and RMSE from 139.0023 to 123.7407. This corresponds to an approximate reduction of 18.2% in MAE and 11.0% in RMSE. The CNN-BiLSTM also achieved the lowest MAPE value, indicating that it provided the most accurate relative predictions for this pilot.

A second evaluation was designed to examine the contribution of feature-space expansion to the performance of the CNN-BiLSTM model. In this evaluation, the same model architecture was tested under three input configurations, representing progressively richer levels of information. The first configuration, E1, uses only core temporal and historical demand features and serves as the baseline setup. The second configuration, E2, extends the baseline by adding contextual and exogenous variables, such as weather, holidays, search-interest indicators, transport-related variables, tourism nights, and selected local infrastructure features. The third configuration, E3, uses the complete selected feature set retained after preprocessing and missing-value filtering. This experimental design allows the analysis to assess whether additional contextual and engineered variables improve the forecasting performance of the CNN-BiLSTM model, or whether an overly extended feature space introduces noise or overfitting for specific pilot destinations.

- **E1: Core Temporal Features:** In the first experiment, the CNN-BiLSTM model was trained using only the core temporal and historical demand features. These features describe the internal time-series structure of tourism flow, including month, quarter, year, normalised time index, COVID and post-COVID indicators, lagged arrivals, rolling averages, rolling standard deviations, yearly anchors, and year-over-year differences. This setup represents a baseline deep-learning configuration in which the model relies mainly on seasonality, recent demand history, and temporal persistence.
- **E2: Core + Contextual Features**: In the second experiment, additional contextual and exogenous variables were added to the core temporal feature set. These include variables related to weather, precipitation, wind, temperature, holidays, Google Trends or search-interest signals, Wikipedia page views, transport indicators such as ferry or airport activity where available, tourism nights, and selected local infrastructure indicators.
- **E3: All Available Features**: In the third experiment, the CNN-BiLSTM model was trained using the complete selected feature set available after preprocessing and missing-value filtering. This includes all temporal, lagged, rolling, contextual, transport, environmental, search-interest, tourism, and engineered variables retained by the feature-selection step. This experiment represents the maximum-information scenario, where the model receives the richest available representation of each pilot destination.

Table 9 presents a feature-space comparison using the CNN-BiLSTM for the two pilot sites.

Table 9: CNN-BiLSTM feature-space comparison for North Sporades

Destination	Experiment	N_features	MAE	RMSE	MAPE	R^2
North Sporades	E1 core temporal	55	146.7595	173.2707	15.7600	0.9741
North Sporades	E2 core plus context	70	173.2698	208.2514	17.4701	0.9626
North Sporades	E3 all features	93	90.1296	123.7407	6.10223	0.9868

The retained feature set for the North Sporades pilot reflects the available data conditions and the destination-specific characteristics of Alonissos. In addition to core temporal and historical tourism-flow variables, the model incorporates island-specific mobility and tourism indicators, such as ferry-related variables, transport-demand proxies, weather conditions, calendar effects, and digital-interest signals. During preprocessing, only numerical variables with acceptable missing-data coverage were retained, ensuring that the final feature space used for the North Sporades evaluation was consistent, complete, and suitable for model training. The corresponding feature-space configuration and prediction performance for the Danube Delta pilot are reported separately in Annex.

For the North Sporades pilot, the CNN-BiLSTM model achieved the best performance when all available features were used. The E3 all-feature configuration obtained the lowest MAE, RMSE, and MAPE values and the highest R^2 score. Compared with the E1 core temporal configuration, the E3 setup reduced MAE from 146.7595 to 90.1297 and RMSE from 173.2707 to 123.7407. This corresponds to an approximate reduction of 38.6% in MAE and 28.6% in RMSE. The MAPE also decreased substantially, from 15.7600% to 6.1022%, indicating that the full-feature configuration provided more accurate relative predictions. These results suggest that, for North Sporades, the additional contextual, transport, environmental, search-interest, tourism-related, and engineered variables improved the model's ability to capture the dynamics of tourism demand.

4.8.2 KPI-based Impact Results

This subsection presents the KPI-level impact results produced by the Tourism Flows Simulation and Capacity Management Tool for the North Sporades pilot, with emphasis on Alonissos for June 2026. The results illustrate how the predicted tourism-flow condition is translated into environmental, social-cultural and economic KPI outputs. The objective is not only to report individual KPI values, but also to show how these values are compared with reference or optimised conditions and how they are interpreted through direction-aware messages such as "Better by", "Worse by" or "No change". The corresponding KPI-based impact results for the Danube Delta pilot are reported in Annex.

The results are organised into three sustainability dimensions: Environmental Conditions, Social and Cultural Dynamics, and Economic Aspects. For each KPI, the tables report the predicted value, the corresponding optimised or reference value, whether the indicator is stable or dynamic, and the resulting interpretation message. Dynamic indicators are directly affected by the predicted tourism flow and are recalculated through the simulation process. Stable indicators are retrieved from the destination-specific KPI configuration and provide the contextual baseline for interpreting the simulated state. Missing or non-applicable values are marked as N/A and are excluded from the corresponding score aggregation.

For the North Sporades pilot, the tool produced a predicted tourist flow of 2106 visitors for June 2026, corresponding to a medium-flow condition.

Table 10 presents the Environmental Conditions KPI results for the North Sporades pilot. The results show a mixed environmental profile. Some resource-consumption indicators perform positively, especially tourism-related CO_2 emissions per capita/day and water consumption per tourist, both of which remain below or close to the reference values. The Air Quality Index also remains unchanged, indicating a stable air-quality condition for the simulated period. However, other environmental indicators reveal areas of pressure or structural weakness. In particular, the absence of green transport means and green transport types indicates a gap in sustainable mobility options. In addition, the indicators related to tourism infrastructure in Natura protected areas, ecological footprint, biodiversity disturbance, light pollution and noise pollution suggest that environmental sensitivity remains an important issue for the pilot area. These results explain why the Environmental Conditions score reaches a moderate value rather than a clearly positive one.

Table 10: Environmental Conditions KPI Results for the North Sporades (Alonnisos) Pilot - June 2026

Kpi	Predicted	Optimized	Stable/ Dynamic	Message
Number of green transport means	0	3	Stable	Worse by 3 Number of means
Types of green transport means	0	3	Stable	Worse by 3 Types of means
Number of tourism accommodation and attraction infrastructure located in cultural heritage protected areas	0	8.5	Stable	Better by 8.5 Number of infrastructures
Number of tourism accommodation and attraction infrastructure located in natura protected areas	55	4.2	Stable	Worse by 50.8 Number of infrastructures

Percentage of local enterprises actively supporting protection of local biodiversity	85	50	Stable	Better by 35%
CO2 emissions per capita/day	0.15	3.5	Dynamic	Better by 3.35 kg/tn
Water Consumption per Tourist	146.4	150	Dynamic	Better by 3.6 liters/day
Air Quality Index	2	2	Dynamic	No change
Land/water resources used (Ecological Footprint per Tourist)	60	35	Stable	Worse by 25 index points
Impacts on key species/habitats (Biodiversity Disturbance Index)	30 (low)	10 (low)	Stable	Worse by 20 index points
Artificial lighting changes in protected or dark-sky zones	40 (medium)	5 (low)	Stable	Worse by 25 index points
dB increase in tourism zones vs baseline (Noise Pollution Index)	60 (medium)	30 (low)	Stable	Worse by 30 index point

Table 10 presents the Social and Cultural Dynamics KPI results. This dimension achieves the strongest score among the three sustainability dimensions. The results indicate positive performance in several social and accessibility-related indicators, including safety level, accommodation accessibility and public transport accessibility. The number of tourists per 100 residents also remains below the reference pressure level, suggesting that the simulated tourism-flow condition does not generate excessive community pressure for the selected month. At the same time, the results highlight specific social and cultural areas requiring improvement. These include accessibility of tourist attractions, access to services, cultural inclusivity, social inclusion and digital accessibility. Therefore, although the overall social-cultural condition is relatively positive, the KPI-level results show that inclusiveness and accessibility should remain priorities for destination management.

Table 11: Social & Cultural KPI Results for the North Sporades (Alonnisos) Pilot - June 2026

Kpi	Predicted	Optimized	Stable/ Dynamic	Message
Number of tourists per 100 residents	35.86	125.5	Dynamic	Better by 89.64 Number of tourists per 100 residents
Level of safety (low/medium/high)	2 (high)	2 (high)	Stable	No change
Percentage of commercial accommodation establishments participating in recognised accessibility information schemes	90	50	Stable	Better by 40%
Percentage of public transport that is accessible to people with disabilities	100	50	Stable	Better by 50%

Percentage of tourist attractions that are accessible to people with disabilities	8	100	Stable	Worse by 92%
Number of the destination's events that are not traditional/cultural relevant	N/A	N/A	Stable	
Number of the destination's events that are focused on traditional/local culture and heritage	1	0.5	Stable	Better by 0.5 Number of events
Visitor volume per m², time-based load, etc. (Heritage Site Stress Index)	0.1 (low)	0.1 (low)	Stable	No change
Impact of tourism on local access to transport, health, and other infrastructure	60	60	Stable	No change
Signage and guides in multiple languages, cultural sensitivity	60	80	Stable	Worse by 20%
Policies for LGBTQ+ travelers, families, seniors (Social Inclusion)	60 (medium)	100 (high)	Stable	Worse by 40%
Screen-reader compatible websites, mobile app usability (Digital Accessibility)	45	80	Stable	Worse by 35%

Table 12 presents the Economic Aspects KPI results. This dimension records the lowest score among the three categories, indicating that the simulated scenario is weaker from an economic-performance perspective. The number of tourist nights per month and tourism revenue per capita are substantially below the corresponding reference values, suggesting that the predicted tourism-flow condition may not generate the desired level of economic benefit for the destination. The average length of stay performs slightly better than the reference value, which is a positive sign, but this improvement is not sufficient to compensate for the lower values observed in tourism nights and revenue. In addition, the revenue leakage index and economic inclusivity indicator show further weaknesses, suggesting that the economic benefits of tourism may not be fully retained or distributed locally. Direct tourism employment is treated as a stable indicator, but it is not included in the final interpretation due to the absence of a comparable reference value.

Table 12: Economic Aspects KPI Results for the North Sporades (Alonnisos) Pilot - June 2026

KPI	Predicted	Optimized	Stable/ Dynamic	Message
Number of tourist nights per month	6238	71400	Dynamic	Worse by 65162 Number of nights
Average length of stay of tourists (nights)	3.6	3.5	Dynamic	Better by 0.1 Number of nights
Tourism Revenue per Capita (Tourism Revenue)	262.02	2200	Dynamic	Worse by 1,938 EUR/capita

% of tourism spending that leaves the region (Revenue Leakage index)	50	30	Stable	Worse by 20%
Availability of free/low-cost diversity (Economic Inclusivity)	30	70	Stable	Worse by 70%
Direct tourism employment as percentage of total employment	50	N/A	Stable	N/A

The aggregated results show an Environmental Conditions score of 59%, a Social and Cultural Dynamics score of 76%, and an Economic Aspects score of 45%. Based on these section-level scores, the overall sustainability score is 60%, which corresponds to a neutral condition. This means that the simulated tourism state does not indicate a critical sustainability failure, but it also does not represent an optimal or strongly positive condition. The overall result reflects an imbalance between the relatively strong social-cultural performance, the moderate environmental performance and the weaker economic performance.

5 Predictive Advisor on Areas Capacity VS Incoming Tourism Flows

The framework presented in this chapter provides a data-driven methodology for evaluating and improving tourism management strategies at the destination level. The framework integrates a multi-dimensional KPI system spanning four thematic pillars such as sustainable tourism management, tourism volume and employment, socio-cultural pressure and accessibility, and environmental infrastructure, with a scenario-based simulation engine and a metaheuristic optimization layer. For a given destination, flow data describing the current state of tourism activity is ingested alongside static destination parameters and a tourist type mix, from which a comprehensive set of derived metrics is computed, including CO_2 emissions per capita, water consumption, occupancy rate, heritage stress, noise and light pollution, and economic inclusivity. Each metric is evaluated against empirically grounded optimal targets, yielding a utilization score that quantifies proximity to sustainable operating conditions. Optimization of the tourist type composition and, optionally, infrastructure-level variables is performed using the Tree-structured Parzen Estimator (TPE) sampler within the Optuna framework [1], maximizing a composite score defined as the mean deviation of all KPI utilizations from their target values.

The framework is validated across five pilot destinations spanning diverse geographic, demographic, and seasonal contexts: the North Sporades archipelago (Greece), the Island of Pag (Croatia), Nessebar (Bulgaria), the Province of Teramo (Italy), and the Danube Delta (Romania). A sixth pilot site, Kutsurub (Ukraine), is an official TOURAL partner destination located in the Mykolaiv region on the shores of the Dnipro-Buh estuary. However, it could not be integrated into the optimisation tool due to the ongoing conflict, which precluded access to the necessary destination data. These destinations represent the range of destination types targeted by the TOURAL project, encompassing small island and coastal ecosystems, Black Sea heritage sites, inland provincial tourism economies, and transboundary UNESCO wetland territories, and are evaluated across low, mid, and high season scenarios. Results are exported in structured JSON format and may be used to generate evidence-based local action plans, bridging the gap between quantitative optimization outputs and actionable policy recommendations.

5.1 Purpose and Functional Scope

The proposed framework is structured as a three-stage pipeline comprising scenario initialization, KPI evaluation, and optimization-driven scenario refinement, with an optional fourth stage for action plan generation. Figure 7 provides an overview of the architecture.

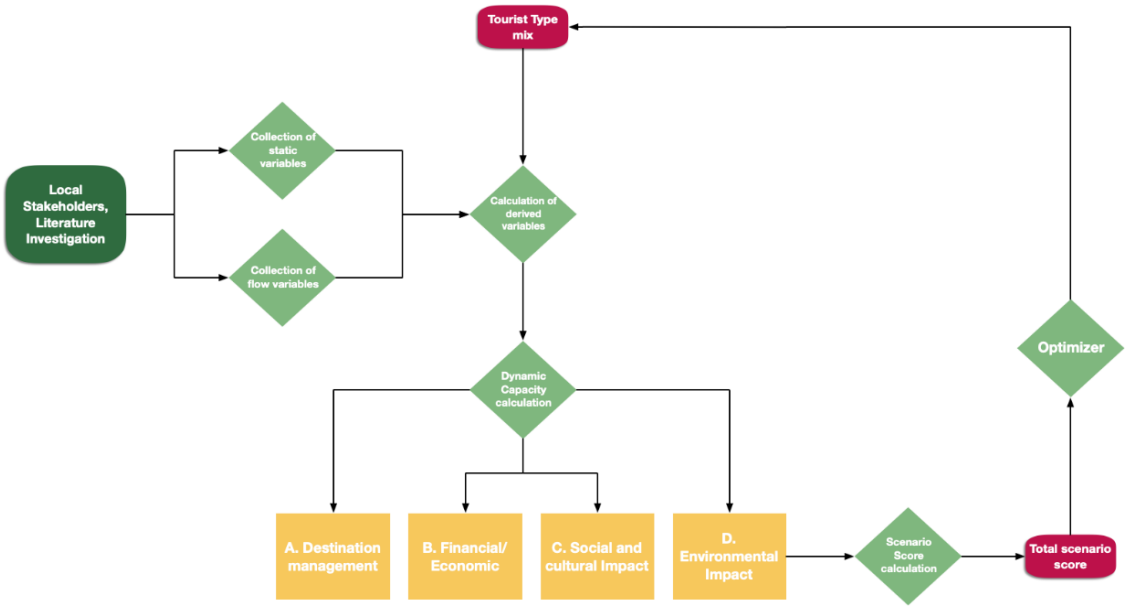

Figure 7: Predictive Advisor Architecture

The first stage, scenario initialization, consolidates all input data into a unified computational object. Two categories of input are distinguished: static destination parameters, which are fixed across simulation runs and include the total number of accommodation establishments, beaches, natural and cultural attractions, the resident population, the geographic area, and the boundaries of the tourist season, and dynamic flow parameters, which represent the observed or assumed state of the destination for a given season and include occupancy rates, waste volumes, green transport availability, accessibility percentages, event counts, revenue leakage, and ecological footprint assessments. A third input stream, the tourist type mix, defines the weighted composition of visitor segments, each characterized by expenditure per night, CO_2 emissions, water consumption, and average length of stay, from which composite averages are derived proportionally. From these inputs, a set of dependent variables is computed through empirical estimation functions, including occupancy rate, CO_2 emissions per capita per day, water consumption per tourist per day, air quality index, biodiversity disturbance level, noise and light pollution indices, heritage stress index, tourist pressure per 100 residents, safety level, social inclusion score, and monthly tourism revenue per capita.

The second stage, KPI evaluation, maps all computed and observed variables onto a structured set of sustainability indicators organized into four thematic pillars: Destination management, Financial and Economic Performance KPIs, Social and Cultural Impact, Environmental Metrics. These indicators are analysed in detail in subchapter 3.1:

For each indicator, a utilization score is computed as the ratio of the observed flow value to an empirically grounded optimal target, expressed as a percentage. Indicators for which lower values are preferable, such as CO_2 emissions, heritage area pressure, and noise pollution, are evaluated using an inverse utilization function that returns 100% when the observed value is at or below the target and decays proportionally as it exceeds it. A composite scalar score aggregating all indicators is defined as 100 minus the mean absolute deviation of all utilization values from 100%, providing a single optimization objective that rewards proximity to target across all dimensions simultaneously.

The third stage, optimization, employs the Tree-structured Parzen Estimator implemented within the Optuna framework to search over a mixed continuous-discrete parameter space. The primary optimization variables are the weights of the tourist type mix, which are sampled as continuous values and normalized to sum to 100% at each trial. In the extended optimization mode, additional

variables are included covering infrastructure-level decisions such as the number of eco-certified hotels, blue flag beach nominations, event programming, and accessibility investments.

At each trial, the updated parameters are propagated through the scenario initialization stage to recompute all derived metrics, and the KPI evaluation stage is then executed to score the resulting configuration. After a configurable number of trials, the optimal parameter set is extracted, the tourist type mix weights are re-normalized, and the optimized scenario is reconstructed for comparison against the baseline.

The final stage, action plan generation, takes as input the baseline and optimized KPI profiles together with the corresponding tourist type mixes, and produces a structured narrative report identifying which tourist segments should be increased, maintained, or reduced, and which infrastructure or policy interventions are most likely to close the remaining gap between observed and target utilization values. This stage is implemented as a prompt-based interface to a locally hosted language model, ensuring that the generated recommendations remain strictly grounded in the quantitative differences between the two scenarios rather than generic sustainability advice.

5.2 Data Models

The framework operates on three categories of structured input data, each stored as a JSON file and loaded at runtime. The separation between static, dynamic, and behavioural data is a deliberate design choice: it allows the same destination to be evaluated under multiple seasonal profiles and tourist compositions without duplicating infrastructure parameters, and it makes the system straightforward to extend to new destinations by providing the appropriate JSON files. This design ensures consistency between the simulation outputs produced by Flow Capacity Tool and the optimisation scenarios evaluated by Predictive Advisor: both tools operate on the same representation of the destination, and any improvement in the underlying data for one tool is automatically reflected in the other.

Static Destination Data

Static data encapsulates the immutable geographic and infrastructure constants of a destination. This file is loaded once per run and its values remain fixed across all scenario variants and optimization trials. The total hotel count represents the full operational capacity at peak season. The number of hotels actually available at a given time is derived dynamically from this figure using a season-level scaling factor (100% for high season, 65% for mid season, 25% for low season). From the available hotel count, the framework derives the total bed stock and the maximum number of bednights per month, which serve as the upper bound for the occupancy-based KPI targets in Pillar B.

Table 13: Static destination parameters: field definitions and data types

Field	Type	Description
Destination	str	Destination name
Total number of hotels	int	Total hotel count (all seasons)
Total number of beaches	int	Total beach count
Total number of attractions	int	Heritage/cultural attractions
Population	int	Permanent resident population
Area (km^2)	float	Destination area in km^2
Rooms per hotel	int	Average rooms per hotel
Beds per room	int	Average beds per room
Tourist period start month	str	Month name (e.g. “May”)
Tourist period end month	str	Month name (e.g. “September”)

Flow Data

Flow data represents a seasonal snapshot of observed or estimated tourism activity at the destination. Unlike static data, flow parameters vary across the three season levels, low, mid, and high, and a separate file is maintained for each. These values reflect real-world measurements or plausible estimates collected from destination management organizations, national statistics, or field surveys, and they constitute the baseline against which the optimization is evaluated. Several flow fields serve as direct inputs to the KPI indicator classes. Others are used by the estimation functions to derive secondary quantities such as CO_2 emissions per capita, water consumption, occupancy rate, and pollution indices before KPI evaluation takes place.

It is important to note that not all fields in this file are passed directly to the indicator classes. Fields such as CO_2 per capita per day, water litres per tourist per day, and tourism revenue per capita represent baseline reference values that are refined by the estimation functions before use. The final values fed to the KPI evaluator are the outputs of those functions, not the raw JSON values. This design allows the framework to model the mitigating effect of eco-certified hotels on emissions and water use, the influence of seasonal occupancy on consumption, and the revenue-amplifying effect of cultural events, rather than treating the observed figures as immutable inputs. Table 14 presents the field definitions and data types of the seasonal flow parameters, which govern how visitor volume and its associated impacts are distributed across the year within the optimisation framework.

Table 14: Seasonal flow parameters: field definitions and data types

Field	Type	Description
Hotel Certified Eco Label	int	Hotels with eco-certification
Number of Blue Flag Beaches	int	Blue Flag certified beaches
Tourist Nights per Month	int	Total tourist bed-nights in month
Tourism Employment Percentage	float	% of workforce in tourism
Accommodation Accessibility Percentage	float	% accessible accommodation
Public Transport Accessibility Percentage	float	% accessible transport
Number of Accessible Attractions	int	Accessible attractions count
Number of Non-Cultural Events	int	Commercial events (concerts, etc.)
Number of Cultural Heritage Events	int	Cultural/heritage events
Number of Green Transport means	int	Number of green transport vehicles/routes
Number of Green Transport Types	int	Distinct types of green transport
Number of accommodations in Heritage Area	int	Hotels sited in heritage zones
Number of accommodation in Natura Area	int	Hotels sited in Natura 2000 zones
Recyclable Waste (kg)	float	Monthly recyclable waste (kg)
Non-Recyclable Waste (kg)	float	Monthly non-recyclable waste (kg)
Local Biodiversity Support Percentage	float	% local biodiversity area protected
Revenue Leakage Percentage	float	% of tourist spending leaving local economy
Digital Accessibility Percentage	float	% digital service compliance
Cultural Inclusivity Percentage	float	% multilingual/inclusive coverage
Economic Inclusivity Index	str	"low", "moderate", or "high"
Ecological Footprint Index	str	"low", "moderate", or "high"

Access to Services	str	One of three ordinal categories
Tourism revenue per capita	float	Baseline EUR/capita (informational only)

Optimal Target Data

Optimal target values define the sustainability thresholds against which all KPI utilization scores are computed. Each target represents the value at which a given indicator is considered to be operating at full sustainability capacity, and it is against this value that the utilization percentage, the core metric of the KPI evaluation stage, is calculated. The optimal target values presented in the table below were defined by the pilot destination representatives in collaboration with the project partners, based on local knowledge, destination-specific constraints, and sustainability ambitions. Targets are destination-specific and expert-calibrated, reflecting the physical constraints, demographic characteristics, and development aspirations of each location rather than universal benchmarks. For destinations with limited empirical data, targets may be informed by European reference frameworks such as ETIS or by comparable destinations in the literature.

Two categories of target require particular attention. First, several targets are not fixed constants but are computed dynamically at runtime from the static destination parameters and the current season level. The optimal number of tourist nights per month, for example, is derived from the season-appropriate occupancy target and the maximum bednights capacity, so it adjusts automatically as the simulation date changes. Similarly, the optimal number of accessible attractions is set to 50% of the total attraction number, and the optimal number of cultural events is set to 50% of all events, both computed from the destination's actual infrastructure rather than from a fixed number. Second, for pressure and pollution indicators, where lower observed values are preferable, the target represents the maximum tolerable threshold rather than a value to be reached.

Utilization for these indicators is computed using an inverse function that returns 100% when the observed value is at or below the target and decays as it exceeds it. The table below lists the optimal targets as configured for all of the integrated pilots and provides the rationale for each value. Table 15 presents the per-archetype fields of the tourist type mix data model, detailing the five parameters used to characterise each archetype within the optimisation framework.

Table 15: Destination-specific optimal KPI targets across the five TOURAL pilot sites

Variable	Unit	Danube Delta	Island of Pug	Nessebar	North Sporades	Teramo
Eco-Certified Hotels	% of available hotels	30	30	30	30	30
Blue Flag Beaches	% of total beaches	30	30	30	30	7
Revenue (Low Season)	EUR/capita/month	900	2200	2200	2200	2200
Revenue (Mid Season)	EUR/capita/month	1500	2200	2200	2200	2200
Revenue (High Season)	EUR/capita/month	2100	2200	2200	2200	2200
Revenue Leakage	% of gross tourism revenue	30	30	30	30	25
Inclusivity	impact index (0–100)	100	100	100	100	100
Occupancy Rate (High Season)	% of maximum bednights	90	90	85	85	55
Occupancy Rate (Mid Season)	% of maximum bednights	60	75	75	70	25
Occupancy Rate (Low Season)	% of maximum bednights	40	60	50	35	15

Average Stay Duration (High Season)	nights	3.2	7.3	7.5	4	6.5
Average Stay Duration (Mid Season)	nights	2.5	6.5	6	3.5	3.5
Average Stay Duration (Low Season)	nights	2.2	6	3.5	3	3.5
Tourism Employment Rate (High Season)	% of total local employment	85	65	80	85	10
Tourism Employment Rate (Mid Season)	% of total local employment	75	40	55	45	7
Tourism Employment Rate (Low Season)	% of total local employment	40	25	25	30	6
Cultural Inclusivity	% multilingual coverage	40	80	80	80	80
Digital Inclusivity	% accessibility compliance	10	80	80	80	30
Safety Level	ordinal (low / moderate / high)	High	High	High	High	High
Accommodations in Heritage Areas	count of infrastructures	0	0	10	10	15
Accommodations in Natura 2000 Areas	count of infrastructures	1	0	5	5	4
Green Transport Options (Low Season)	count of distinct modalities	1	0	25	3	3
Green Transport Variety (Low Season)	count of connections	0	0	3	3	4
Green Transport Options (Mid Season)	count of connections	1	0	85	3	3
Green Transport Variety (Mid Season)	count of distinct modalities	0	0	6	3	4
Green Transport Options (High Season)	count of connections	1	1	210	3	3
Green Transport Variety (High Season)	count of distinct modalities	0	1	8	3	4
Biodiversity Disturbance	impact index (0–100), inverse target	10	10	10	10	10
Accommodation Accessibility	% of establishments or services	50	50	50	50	50
Attraction Accessibility	% of total attractions	50	50	50	50	75
Cultural Events Share	% of all events	50	50	50	50	50
Recyclable Waste Ratio	% of total waste by weight	60	60	60	60	60
Biodiversity Support	% of local tourism enterprises	50	50	50	50	50
CO_2 Emissions	kg CO_2/capita/day, inverse target	3.5	3.5	3.5	3.5	3.5
Water Consumption	litres/tourist/day, inverse target	150	150	150	150	150
Air Quality Index	index 1–6, inverse target	2	2	2	2	2
Ecological Footprint	impact index (0–100), inverse target	35	35	35	35	35

Light Pollution	impact index (0–100), inverse target	5	5	5	5	5
Noise Pollution	impact index (0–100), inverse target	30	30	30	30	30

Tourist Type Mix

The tourist type mix defines the behavioural and environmental profile of the visitor population at a destination. Rather than treating all tourists as a homogeneous group, the framework disaggregates the visitor population into a set of archetypes, each representing a distinct travel motivation with characteristic expenditure patterns, resource consumption, and length of stay. This disaggregation is central to the optimization logic: because the composite KPI score is sensitive to average CO_2 emissions, water consumption, expenditure per night, and length of stay, all of which are derived as weighted averages across the active archetypes, shifting the tourist mix is one of the primary levers available to the optimizer for improving sustainability outcomes without requiring changes to physical infrastructure.

Each archetype is defined by five parameters, listed in the table below. The weight field represents the archetype's current share of the total visitor population, expressed as a percentage, and the set of weights across all archetypes must sum to 100. During optimization, the weights are treated as continuous decision variables, sampled in raw form and normalized to sum to 100% before each trial evaluation, ensuring that the constraint is satisfied throughout the search. The remaining four fields are treated as fixed properties of each archetype and are not subject to optimization. They reflect empirical estimates of the behavioural and environmental characteristics associated with each travel motivation, drawn from tourism impact studies and destination-level surveys. Table 16 presents the per-archetype fields of the tourist type mix data model, detailing the five parameters used to characterise each archetype within the optimisation framework.

Table 16: Per-archetype fields of the tourist type mix data model

Field	Type	Description
Spending per Night	float	Average nightly spend per tourist in EUR. Contributes to the monthly tourism revenue per capita estimate via the weighted average across all active archetypes
CO_2 per Night (kg)	float	CO_2 emitted per tourist per night in kilograms. Used to derive the destination-wide baseline emission rate before adjustment for green transport and eco-certification
Average Stay Duration	float	Mean length of stay in nights. Influences the average stay KPI in Pillar B, the tourist pressure per 100 residents estimate, and indirectly the water consumption adjustment
Water Use per Night (L)	float	Daily water consumption per tourist in litres. Used to derive the baseline water use figure before seasonal, occupancy, and eco-certification adjustments
Weight	float	Share of the tourist mix as a percentage. Must sum to 100 across all archetypes. The primary decision variable during optimization.

The seven archetypes defined for the five pilot cases reflect the destination's existing visitor base and realistic target segments given its natural and cultural offer. They span a wide range of sustainability profiles: nature-oriented and religious tourists tend to have lower CO_2 footprints and moderate expenditure, while diving, honeymoon, and medical tourists generate higher per-

night resource consumption but also higher economic returns. This heterogeneity is what makes the tourist mix a meaningful optimization variable, a destination can improve its composite KPI score by shifting weight toward archetypes whose profile more closely aligns with its sustainability targets, without necessarily reducing total visitor volume. Table 17 presents the behavioural and environmental parameters of the eight tourist archetypes, including average nightly expenditure, CO_2 emissions, length of stay, and water consumption per night.

Table 17: Behavioural and environmental parameters of the eight tourist archetypes

Archetype	Expenditure (EUR/night)	CO_2 (kg/night)	Stay (nights)	Water (L/night)
Nature Tourism	75	7.7	7	125
Diving Tourism	130	10.5	8	380
Adventure Tourism	110	9.2	8	290
Religious Tourism	65	6.5	4	170
Honeymoon Tourism	160	12.8	9	600
Arts Tourism	95	8.8	6	270
Medical Tourism	140	11.5	6	360
Family Leisure	140	11.5	6	360

The behavioural and environmental parameters assigned to each tourist archetype represent composite estimates derived from multiple reference sources rather than single empirically measured values. Their construction follows a three-stage process.

Expenditure per night figures are grounded in Eurostat tourism statistics (dataset tour_occ_nim) and supplemented by European Travel Commission niche market reports covering nature-based, religious, medical, and honeymoon tourism segments (ETC, 2016[6]; ETC, 2021[7]). Where destination-level expenditure surveys were available from national tourism organisations participating in the TOURAL consortium, these were used to validate and adjust the Eurostat baseline figures for the specific economic context of each pilot destination.

CO_2 emission rates per tourist-night are derived from the UNWTO/ITF report on transport-related CO_2 emissions of the tourism sector [2], which provides per-tourist-night emission benchmarks disaggregated by transport mode and accommodation category, and from the EU Best Environmental Management Practice document for the tourism sector [3], which establishes reference emission intensities for Mediterranean accommodation establishments. Higher emission values assigned to Diving, Medical, and Honeymoon archetypes reflect the documented contribution of specialised equipment transport, air travel prevalence, and luxury accommodation energy consumption in those segments.

Water consumption figures are taken directly from the EU BEMP document, which reports empirically measured water use benchmarks for European tourism accommodation in the range of 150–600 litres per tourist per night depending on facility category, season, and service level. The upper bound assigned to Honeymoon Tourism reflects the luxury accommodation benchmark

[6] https://etc-corporate.org/reports/european-tourism-2016-trends-prospects-q4-2016/

[7] https://etc-corporate.org/reports/european-tourism-2021-trends-prospects-q2-2021/

reported in that document, while the lower bound assigned to Nature Tourism reflects the self-catering and low-infrastructure accommodation profile characteristic of that segment.

Average length of stay figures are drawn from Eurostat accommodation statistics and from the European Travel Commission's series of niche tourism market reports[5] , cross-referenced where possible against destination-level arrival and overnight stay data collected during the TOURAL project fieldwork phases. It should be noted that all archetype parameters represent central estimates applicable to the Mediterranean and Central European destination contexts of the TOURAL pilot sites. Their transferability to destination contexts with substantially different visitor profiles or infrastructure conditions should be assessed before application.

It is worth noting that the optimizer does not have visibility into the individual archetype parameters during the search. It only observes the composite KPI score that results from a given weight configuration. The relationship between mix weights and KPI outcomes is therefore implicit and non-linear, which is precisely why a model-free Bayesian search strategy such as TPE is appropriate for this problem rather than a gradient-based method.

5.2.1 Scenario Assembly

This subchapter analyses the scenario assembly process, including the key assumptions adopted to define the input configuration of the Predictive Advisor. In particular, it explains how static destination parameters, seasonal flow variables, and tourist type mix assumptions are integrated into a unified scenario for subsequent KPI evaluation and optimisation.

Season Classification

The simulation date is classified into one of three season levels, high, mid, or low, based on its calendar month relative to the destination's configured tourist season boundaries. The shoulder months immediately before the start and after the end of the peak season are classified as mid season, reflecting the transitional character of those periods in terms of visitor volume and infrastructure utilisation. All months falling within the core tourist window are classified as high season, and all remaining months as low season. This classification drives several downstream quantities, including the number of operational hotels, the season-appropriate KPI targets for occupancy, stay length, employment, and revenue, and the water consumption seasonality multiplier.

Given a datetime timestamp and the destination's tourist_period_start_month and tourist_period_end_month:

$$\text{season} \begin{cases} mid & \text{if } m = m_{\text{start}} - 1 \text{ or } m = m_{\text{end}} + 1 \\ low & \text{if } m < m_{\text{start}} \text{ or } m > m_{\text{end}} \\ high & \text{otherwise} \end{cases} \quad (1)$$

where m is the current calendar month (1–12), m_{start} and m_{end} are the numeric equivalents of the start and end month strings.

Hotel Availability

Not all accommodation establishments operate year-round. The number of hotels available in a given simulation run is derived from the total hotel stock recorded in the static destination data by applying a season-level scaling factor. At full peak season, all hotels are assumed to be operational. At mid season, 65% of the total stock is assumed open, reflecting the partial closure of smaller or seasonal establishments during shoulder periods. At low season, only 25% of the stock remains operational. The result is floored to the nearest integer. This figure is used to compute the total bed stock and the maximum bednights capacity, and it also serves as the denominator for infrastructure pressure KPIs such as heritage and Natura area accommodation counts.

The number of operational hotels is reduced by season:

$$H_{\text{available}} = \begin{cases} H_{\text{total}} & \text{high season} \\ \lfloor 0.65 \cdot H_{\text{total}} \rfloor & \text{mid season} \\ \lfloor 0.25 \cdot H_{\text{total}} \rfloor & \text{low season} \end{cases} \quad (2)$$

where H_{total} is the total number of available hotels in the area.

Maximum Bed-Nights

The maximum bednights figure represents the theoretical upper bound on tourist overnight capacity in a single month, assuming full occupancy across all available accommodation. It is computed as the product of the number of available hotels, the average number of rooms per hotel, the average number of beds per room, and 30 days. This quantity serves as the reference capacity for the occupancy rate estimation function and as the upper bound of the tourist nights search space during optimization. Because it depends on hotel availability, it varies across season levels for the same destination.

$$N_{\text{max}} = H_{\text{available}} \times R_{\text{hotel}} \times B_{\text{room}} \times 30 \quad (3)$$

where R_{hotel} is rooms per hotel and B_{room} is beds per room.

Tourist Type Mix Averaging

The tourist type mix defines a set of K visitor archetypes, each assigned a weight representing its share of the total visitor population. Before any derived metrics are computed, the raw weights are normalized so that they sum to unity, ensuring that the weighted averages are well-defined regardless of the absolute scale of the input weights. Four composite scalars are then derived as normalized weighted averages across all archetypes: average expenditure per night, average CO_2 emissions per tourist per night, average length of stay in nights, and average water consumption per tourist per night. These four scalars replace the individual archetype values for all subsequent computations. They are the quantities passed to the CO_2 estimation function, the water consumption estimation function, the revenue estimation function, and the average stay KPI, meaning that a shift in the tourist type mix weights propagates immediately and consistently through every derived metric and KPI in the evaluation stack. This is the mechanism by which tourist mix optimization translates into KPI score improvement. The tourist type mix JSON supplies a weight w_i for each of the K archetypes. Four weighted averages are derived:

$$\bar{x} = \sum_{i=1}^{K} \frac{w_i}{\sum_j w_j} \cdot x_i \quad (4)$$

Applied to: e (expenditure per night), c (CO_2 per night), s (average stay nights), w (water per night, litres).

These four scalars drive all downstream derived metrics.

5.2.2 Derived Metric Pipeline

After loading static and flow data and computing the tourist mix averages, the following derived metrics are computed in order:

1. Daily population, occupancy rate, tourists per 100 residents
2. CO_2 emissions per capita per day
3. Water consumption per tourist per day
4. Air Quality Index
5. Biodiversity disturbance level
6. Light pollution classification

7. Noise pollution classification
8. Heritage stress classification
9. Safety level
10. Social inclusion category
11. Monthly revenue per capita

All estimation functions are invoked sequentially during the derived metrics pipeline (Section 5.2.2). Their role is to convert raw flow data and tourist mix parameters into the scalar or categorical values consumed by the KPI indicator classes. Where possible, functions are grounded in empirical references from environmental management literature; where precise calibration data are unavailable, coefficients are set conservatively and documented in Annex so that they can be updated as destination-level measurements become available.

5.3 KPI scoring methods

All 33 indicators computed across the four pillar classes are reduced to a common utilization percentage $U \in [0,100]$ that expresses how close the observed value is to its optimal target. This normalization is what makes indicators with different units, scales, and directionalities directly comparable and aggregable into a single composite score. Two scoring methods are defined, selected on a per-indicator basis according to whether a higher or lower observed value represents better performance.

5.3.1 Forward Utilization (positive indicators — higher is better)

Forward utilization is applied to positive indicators, those for which higher observed values are better and the optimal target represents a value to be reached or exceeded. Examples include eco-label certification rate, tourist nights per month, occupancy rate, tourism employment percentage, revenue per capita, biodiversity support percentage, and recycling ratio.

The utilization percentage is defined as the ratio of the observed flow value V to the optimal target V^*, expressed as a percentage:

$$U = \begin{cases} 100.0 & \text{if } V = 0 \text{ and } V^* = 0 \\ 100.0 & \text{if } V^* = 0 \text{ and } V \geq 0 \\ \text{round}\left(\frac{V}{V^*} \times 100, 1\right) & \text{otherwise} \end{cases} \quad (5)$$

where V is the flow (observed) value and V^* is the optimal target.

The first case handles the degenerate situation in which both the observed value and the target are zero, which is interpreted as perfect performance. The destination is not expected to do anything and is not doing anything. The second case handles targets that are set to zero, which would produce a division-by-zero error under the general formula. In this case the utilization is set to 100 regardless of the observed value, on the basis that a zero target represents a condition that is already satisfied. In all other cases, the utilization is the standard ratio, rounded to one decimal place.

Most indicator classes apply an additional cap of 100% to the forward utilization score, preventing over-achieving indicators, where the observed value exceeds the target, from inflating the composite score. This cap ensures that the optimization objective rewards closing gaps rather than accumulating surplus on already-satisfied indicators, and that the composite score remains interpretable as a mean proximity to target rather than an unconstrained average ratio. Indicators where the cap is not applied are those where the target represents a soft ceiling rather than a hard optimum, and exceeding it is genuinely desirable.

Output: utilization percentage $U \in [0,100]$ after capping, or $U \in [0, +\infty)$ where the cap is not applied. A value of 100 indicates that the observed flow meets or exceeds the optimal target; values below 100 indicate underperformance proportional to the gap between V and V^*.

5.3.2 Inverse Utilization (pressure indicators — lower is better)

Inverse utilization is applied to pressure indicators, those for which lower observed values are better and the optimal target represents a maximum tolerable threshold rather than a value to be reached. Examples include CO_2 emissions per capita, water consumption per tourist per day, heritage area accommodation pressure, Natura area accommodation pressure, ecological footprint, biodiversity disturbance, light pollution, and noise pollution.

The scoring logic inverts the standard ratio so that an observed value at or below the target receives a perfect score of 100%, while values exceeding the target are penalized proportionally:

$$U = \begin{cases} 100.0 & \text{if } V^* \leq 0 \text{ and } V \leq 0 \\ 0.0 & \text{if } V^* \leq 0 \text{ and } V > 0 \\ 100.0 & \text{if } V \leq V^* \\ \text{round}\left(\frac{V^*}{V} \times 100, 2\right) & \text{if } V > V^* \end{cases} \quad (6)$$

A flow at or below target is rewarded with a perfect score. Overshoot is penalized proportionally: if $V = 2V^*$, then $U = 50$.

The first two cases handle degenerate targets. When both the observed value and the target are zero or negative, the indicator is interpreted as having no pressure and receives a perfect score. When the target is zero or negative but the observed value is positive, any non-zero pressure on a zero-tolerance indicator is treated as a complete failure, returning a score of 0%.

The third case is the key behavioural distinction from forward utilization: any observed value that meets or is below the target threshold is scored at 100%, regardless of how far below the threshold it falls. There is no reward for being further below the target than necessary, the indicator is simply satisfied. This prevents the optimizer from over-investing in already-compliant pressure indicators at the expense of indicators that are genuinely underperforming.

The fourth case governs overshoot. The penalty is proportional and hyperbolic: a destination emitting twice the optimal CO_2 level scores 50%, one emitting three times the target scores 33%, and so on. Formally, if $V = k \cdot V$ for $k > 1$, then $U = 100/k$. This decay is steeper than a linear penalty for large overshoots. A destination at five times the target scores only 20%, but remains smooth and differentiable, which is consistent with the TPE optimizer's expectation of a well-behaved objective landscape. The result is rounded to two decimal places, one more than the forward utilization, to preserve precision for indicators with small target values where the ratio $V*/V$ varies slowly.

Output: utilization percentage $U \in [0,100]$, where 100 indicates full compliance with or improvement beyond the target threshold, and values below 100 indicate proportional overshoot.

5.3.3 Composite and Category Scores

Once all indicator utilization values have been computed, they are aggregated into summary scores at two levels of granularity: a per-category score that summarizes performance within each of the four KPI pillars, and a global sustainability score that aggregates across all indicators simultaneously. The global score serves as the objective function maximized by the Optuna TPE optimizer.

Category score: For each pillar $k \in \{A, B, C, D\}$, let K_k denote the set of indicators belonging to that category and U_i the utilization percentage of indicator i. The category score is defined as 100 minus the mean absolute deviation of all indicator utilizations from the perfect score of 100:

$$S_k = 100 - \frac{1}{|K_k|} \sum_{i \in K_k} |U_i - 100| \qquad (7)$$

Global sustainability score. The global score applies the same formulation across all N indicators simultaneously, regardless of category membership:

$$S = 100 - \frac{1}{N} \sum_{i=1}^{N} |U_i - 100| \qquad (8)$$

Interpretation: The scoring formulation has several properties that make it well-suited as an optimization objective. A score of 100 is achieved if and only if every indicator is exactly at its optimal target, neither underperforming nor, where the cap is not applied, over-achieving. A score of 0 corresponds to an average deviation of 100 percentage points across all indicators, representing a destination where every KPI is either completely absent or at twice its target on a pressure indicator. The score can theoretically become negative if utilization values deviate by more than 100 points on average from the target, which can occur for pressure indicators evaluated using the inverse utilization function when observed values substantially exceed their thresholds, for example, a destination with CO_2 emissions ten times the optimal target would contribute a utilization of 10%, a deviation of 90 points, to the aggregate.
The use of mean absolute deviation rather than mean squared error is a deliberate design choice. Squared error penalises large deviations disproportionately and would cause the optimizer to focus almost exclusively on the worst-performing indicator at each trial, neglecting incremental improvements across the broader indicator set. Mean absolute deviation treats all deviations linearly, encouraging the optimizer to make balanced progress across all KPIs simultaneously rather than converging to solutions that excel on a small number of indicators while leaving others severely underperforming. This property is particularly important in the context of sustainable tourism, where a destination that scores perfectly on economic indicators but fails on environmental ones cannot be considered genuinely sustainable.

Output: global sustainability score $S \in (-\infty, 100]$, returned as a float rounded to 2 decimal places. In practice, scores below 0 indicate a severely stressed destination and scores above 80 indicate strong overall alignment with sustainability targets across all four pillars.

5.4 KPI Indicator Specifications

Pillar A assesses the adoption of formal sustainability certification across the destination's accommodation stock and coastal assets, with both indicators measuring scheme penetration relative to total available infrastructure. Table 18 presents the KPI specification and scoring method for Pillar A (Destination Management).

Table 18: KPI specification and scoring method - Pillar A: Destination Management

KPI	Description	Input	Output / Scoring method
A1 — Hotel eco-	Measures the share of available hotels holding a recognised eco-label certification relative to	Flow: number of hotels currently holding a recognised eco-label certification. Optimal: the	Forward utilization, capped at 100. Unit: number of hotels.

certification rate	the destination-specific percentage target.	target number of certified hotels, computed as a fixed percentage of the total available hotel stock for the current season, rounded to the nearest integer.	
A2 — Blue Flag beach coverage	Measures the share of beaches holding Blue Flag status relative to the destination-specific percentage target.	Flow: number of beaches currently holding Blue Flag status. Optimal: the target number of Blue Flag beaches, computed as a fixed percentage of the total beach count at the destination, rounded to the nearest integer.	Forward utilization, capped at 100. Unit: number of beaches.

Pillar B assesses destination performance across six economic and volumetric dimensions: visitor throughput, length of stay, accommodation utilisation, employment generation, revenue distribution, and revenue retention. Four indicators use season-dependent optimal targets. Table 19 presents the KPI specification and scoring method for Pillar B (Financial and Economic Performance).

Table 19: KPI specification and scoring method - Pillar B: Financial and Economic Performance

KPI	Description	Input	Output / Scoring method
B1 — Tourist nights per month	Measures actual tourist bednights against the season-appropriate capacity target, derived from maximum bednights and the season-level occupancy target (85% / 70% / 35%).	Flow: number of tourist bednights recorded in the reference month. Optimal: the target number of nights, computed from the total bed capacity multiplied by the season-appropriate occupancy target, floored to the nearest integer.	Forward utilization, capped at 100. Unit: nights/month.
B2 — Average length of stay	Measures the tourist-mix-weighted average stay length against the season-appropriate optimal stay nights target.	Flow: weighted average stay length derived from the tourist type mix at the current weights. Optimal: the season-appropriate target stay length loaded from the destination optimal file.	Forward utilization, capped at 100. Unit: nights.
B3 — Occupancy rate	Measures the fraction of available accommodation capacity in use against the season-appropriate occupancy target.	Flow: occupancy rate estimated from tourist nights divided by maximum bednights capacity, expressed as a percentage. Optimal: the season-appropriate occupancy	Forward utilization, capped at 100. Unit: %.

		target loaded from the destination optimal file.	
B4 — Tourism employment rate	Measures tourism sector employment as a share of total local employment against the season-appropriate employment target.	Flow: tourism employment as a percentage of total local employment, from the flow data. Optimal: the season-appropriate employment target loaded from the destination optimal file.	Forward utilization, capped at 100. Unit: % of total employment.
B5 — Tourism revenue per capita	Measures estimated monthly tourism revenue per resident against the season-appropriate revenue target.	Flow: monthly tourism revenue per capita estimated by the revenue estimation function, incorporating tourist volume, average expenditure, leakage, and event multipliers. Optimal: the season-appropriate revenue target loaded from the destination optimal file.	Forward utilization, capped at 100. Unit: EUR/capita/month.
B6 — Revenue leakage	Measures the share of tourism revenue leaving the local economy. Lower leakage is better. The target is the maximum tolerable leakage rate.	Flow: revenue leakage percentage from the flow data. Optimal: the maximum tolerable leakage rate loaded from the destination optimal file.	Inverse utilization. Unit: % of gross tourism revenue.
B7 — Economic inclusivity	Measures how broadly tourism revenue is distributed across the local economy. Ordinal string mapped to numeric score before evaluation: high=100, moderate=60, low=30.	Flow: economic inclusivity level from the flow data, mapped to a numeric score on a 0–100 scale. Optimal: maximum inclusivity score of 100, corresponding to a high classification.	Forward utilization, capped at 100. Output ∈ {0, 30, 60, 100}. Unit: impact index.

Pillar C assesses social and cultural sustainability across eleven indicators spanning tourist pressure, public safety, accessibility, cultural provision, heritage integrity, and social inclusion. It is the most heterogeneous pillar, combining continuous, ordinal, categorical, and count-based metrics. Tourist pressure and heritage stress are pressure indicators evaluated using the inverse utilisation function; the remaining nine use forward utilisation. Table 20 presents the KPI specification and scoring method for Pillar C (Social and Cultural Impact).

Table 20: KPI specification and scoring method - Pillar C: Social and Cultural Impact

KPI	Description	Input	Output / Scoring method
C1 — Tourist pressure	Measures unique tourists relative to the resident population	Flow: number of unique tourists per 100 residents, estimated from tourist nights, average stay	Inverse utilization.

per 100 residents	against a season-dependent tolerance threshold. Lower pressure is better.	length, and resident population. Optimal: the season-appropriate tolerance threshold, derived from the resident population and a season-level ratio of 1 tourist per 100 residents in low season and 4 in mid and high season.	Unit: tourists per 100 residents.
C2 — Safety level	Measures estimated public safety against a target of high safety. Ordinal string mapped: high=2, moderate=1, low=0.	Flow: safety level estimated at runtime from tourist pressure, occupancy rate, and non-cultural event count, mapped to a 0–2 ordinal score. Optimal: 2, corresponding to a high safety classification.	Forward utilization, not capped (scale bounded at 2). Output ∈ {0, 50, 100}. Unit: ordinal score (0–2).
C3a — Accommodation accessibility	Measures the share of accommodation establishments participating in a recognised accessibility information scheme against the 50% target.	Flow: percentage of accommodation establishments with recognised accessibility participation, from the flow data. Optimal: 50%, consistent with ETIS accessibility participation benchmarks.	Forward utilization, capped at 100. Unit: %.
C3b — Public transport accessibility	Measures the share of public transport services accessible to people with disabilities against the 50% target.	Flow: percentage of public transport services meeting accessibility requirements, from the flow data. Optimal: 50%, consistent with ETIS accessibility benchmarks.	Forward utilization, capped at 100. Unit: %.
C3c — Accessible attractions	Measures the number of accessible tourist attractions against a target of 50% of total attractions and beaches.	Flow: number of tourist attractions that are physically accessible or participating in an accessibility scheme, from the flow data. Optimal: the target count computed as the destination-specific accessibility percentage applied to the total number of attractions and beaches.	Forward utilization, capped at 100. Unit: count of accessible attractions.
C4 — Cultural event ratio	Measures the share of events focused on local culture and heritage against a target of 50% of all events.	Flow: number of cultural heritage events from the flow data. Optimal: the target number of cultural events, computed as 50% of the total event count, making the target dynamic rather than fixed.	Forward utilization, capped at 100. Unit: number of events.

C5 — Heritage stress index	Measures tourist pressure on cultural and natural heritage assets. Lower stress is better. Inverted ordinal scale: low=100, moderate=40, high=0.	Flow: heritage stress classification estimated at runtime from tourist density per attraction and heritage-area accommodation count, mapped to an inverted numeric score: low = 100, moderate = 40, high = 0. Optimal: 100, corresponding to a low stress classification.	Forward utilization, reduces to numeric score. Output ∈ {0, 40, 100}. Unit: impact index.
C6 — Access to services	Measures whether tourism pressure is improving or degrading resident access to essential services. String mapped: positive=100, neutral=60, strained=30, unknown=0.	Flow: access to services classification from the flow data, mapped to a numeric score. Optimal: 100, corresponding to a positive or improved access classification. Note: the current flow data uses abbreviated vocabulary that does not match the expected string keys, causing this indicator to default to 0 across all destinations.	Forward utilization, reduces to numeric score. Output ∈ {0, 30, 60, 100}. Unit: impact index.
C7 — Cultural inclusivity	Measures the share of tourism-facing services available in multiple languages or culturally adapted formats against the destination target.	Flow: cultural inclusivity percentage from the flow data. Optimal: the destination-specific multilingual coverage target loaded from the optimal file.	Forward utilization, capped at 100. Unit: % multilingual coverage.
C8 — Social inclusion	Measures the level of social inclusion associated with tourism activity. Ordinal string mapped: high=100, moderate=60, low=30.	Flow: social inclusion classification estimated at runtime from digital accessibility and economic inclusivity, mapped to a numeric score. Optimal: 100, corresponding to a high social inclusion classification.	Forward utilization, reduces to numeric score. Output ∈ {0, 30, 60, 100}. Unit: impact index.
C9 — Digital accessibility	Measures the share of tourism digital infrastructure compliant with accessibility standards against the destination target.	Flow: digital accessibility percentage from the flow data. Optimal: the destination-specific digital accessibility target loaded from the optimal file.	Forward utilization, capped at 100. Unit: % compliance.

Pillar D assesses environmental performance across eleven indicators spanning green transport, protected area pressure, waste management, biodiversity, emissions, water consumption, and pollution. It is the most technically heterogeneous pillar, combining continuous, count-based, and ordinal metrics. For the four ordinal indicators — ecological footprint, biodiversity disturbance,

light pollution, and noise pollution — a shared numeric impact scale is applied before utilisation is computed:

For ordinal impact levels (ecological footprint, biodiversity disturbance, light pollution, noise pollution), a shared numeric impact scale is applied before evaluation:

Table 21: Impact scaling for Pillar D

String	Impact value
"low"	30
"moderate"	60
"high"	100

This scale is intentionally non-zero at the low end: a value of 30 for low impact reflects the fact that even the best-performing destinations generate some environmental pressure, and a score of zero should be reserved for the theoretical condition of no impact whatsoever. Five of the eleven indicators, heritage area pressure, Natura area pressure, CO_2 emissions, ecological footprint, biodiversity disturbance, light pollution, and noise pollution, are pressure indicators evaluated using inverse utilization. The remainder use forward utilization. Table 22 presents the KPI specification and scoring method for Pillar D (Environmental Metrics).

Table 22: KPI specification and scoring method - Pillar D: Environmental Metrics

KPI	Description	Input	Output / Scoring method
D1 — Green transport means	Measures the number of green transport connections to or from the destination against the season-appropriate target.	Flow: number of green transport connections currently available, from the flow data. Optimal: the season-appropriate target number of connections loaded from the destination optimal file.	Forward utilization, capped at 100. Unit: count of connections.
D2 — Green transport types	Measures the number of distinct green transport modalities against the season-appropriate target.	Flow: number of distinct green transport modalities currently available, from the flow data. Optimal: the season-appropriate target number of modalities loaded from the destination optimal file.	Forward utilization, capped at 100. Unit: count of distinct modalities.
D3 — Heritage area accommodation pressure	Measures accommodation establishments in cultural heritage protected zones against the maximum tolerable	Flow: number of accommodation establishments in heritage protected areas, from the flow data. Optimal: the maximum tolerable count, derived as a fixed percentage of the total	Inverse utilization. Unit: count of infrastructures.

	threshold. Lower is better.	hotel stock and loaded from the destination optimal file.	
D4 — Natura area accommodation pressure	Measures accommodation establishments within Natura 2000 protected areas against the maximum tolerable threshold. Lower is better.	Flow: number of accommodation establishments within Natura 2000 zones, from the flow data. Optimal: the maximum tolerable count, derived as a fixed percentage of the total hotel stock and loaded from the destination optimal file.	Inverse utilization. Unit: count of infrastructures.
D5 — Waste recycling capacity	Measures the share of total tourism waste being recycled against a target of 60% of total waste generated.	Flow: recyclable waste in kilograms, from the flow data. Optimal: the target recyclable volume computed as 60% of total waste, where total waste is the sum of recyclable and non-recyclable fractions reported in the flow data.	Forward utilization, capped at 100. Unit: kg.
D6 — Biodiversity support	Measures the share of local tourism enterprises actively engaged in biodiversity protection against the 50% target.	Flow: percentage of local tourism enterprises actively supporting biodiversity, from the flow data. Optimal: 50% of local tourism enterprises, reflecting the expectation that at least half the sector participates in formal or informal biodiversity stewardship.	Forward utilization, capped at 100. Unit: % of local tourism enterprises.
D7 — CO_2 emissions per capita	Measures adjusted CO_2 emissions per tourist per day against the 3.5 kg/capita/day target. Lower is better.	Flow: adjusted CO_2 emissions per tourist per day, estimated from the baseline emission rate and reduced for green transport availability and eco-certification penetration. Optimal: 3.5 kg CO_2 per capita per day.	Inverse utilization. Unit: kg CO_2/capita/day.
D8 — Water consumption per tourist	Measures adjusted daily water consumption per tourist against the 150 L/day target. Lower is better.	Flow: adjusted water consumption per tourist per day, estimated from the baseline figure and corrected for eco-certification, seasonality, occupancy pressure, and stay length. Optimal: 150 litres per tourist per day, based on BEMP reference values for	Inverse utilization. Unit: litres/tourist/day.

		Mediterranean accommodation.	
D9 — Air quality index	Measures the estimated air quality index against a target of 2 on the 1–6 scale. Lower index is better.	Flow: air quality index estimated from the adjusted CO_2 emission rate on a 1–6 scale. Optimal: 2, corresponding to the good air quality band under standard AQI classifications.	Inverse utilization. Unit: index (1–6).
D10 — Ecological footprint	Measures the overall ecological footprint of tourism against a target impact score of 35. Lower is better. Impact scale: low=30, moderate=60, high=100.	Flow: ecological footprint classification from the flow data, mapped to a numeric impact score via the shared scale. Optimal: 35, set just above the low classification value to enforce a low footprint standard with a small tolerance margin.	Inverse utilization. Unit: impact index (0–100).
D11 — Biodiversity disturbance	Measures biodiversity disturbance from Natura 2000 accommodation pressure against a near-zero target. Lower is better. Impact scale: low=30, moderate=60, high=100.	Flow: biodiversity disturbance classification estimated at runtime from Natura 2000 accommodation count, mapped to a numeric impact score via the shared scale. Optimal: 10, set well below the low classification value to enforce a near-zero disturbance expectation for protected zones.	Inverse utilization. Unit: impact index (0–100).
D12 — Light pollution	Measures estimated artificial light pollution against a near-zero target. Lower is better. Impact scale: low=30, moderate=60, high=100.	Flow: light pollution classification estimated at runtime from hotel count, event count, and population density, mapped to a numeric impact score via the shared scale. Optimal: 5, set well below the low classification value to maintain near-minimal artificial light as a persistent policy target.	Inverse utilization. Unit: impact index (0–100).
D13 — Noise pollution	Measures estimated noise pollution level against a low classification threshold. Lower is better. Impact scale: low=30,	Flow: noise pollution classification estimated at runtime from daily population density and event count, mapped to a numeric impact score via the shared scale. Optimal: 30, equal to the numeric value of a low classification, meaning a low	Inverse utilization. Unit: impact index (0–100).

	moderate=60, high=100.	noise classification is the minimum acceptable condition and scores exactly 100%.	

5.5 Optimization Engine

TOURAL predictive advisor employs the Tree-structured Parzen Estimator (TPE) as implemented in the Optuna framework. TPE is a sequential model-based optimization algorithm that operates by maintaining two probabilistic density models over the parameter space, updated iteratively as new trials are evaluated. Unlike grid search or random search, which treat all regions of the parameter space equally, TPE concentrates sampling effort in regions that have historically produced high objective scores, while retaining sufficient exploration to avoid premature convergence to local optima.

At each iteration, the algorithm partitions the set of previously evaluated trials into two groups according to an adaptive quantile $\hat{y}$: trials whose objective score falls below $\hat{y}$ are considered poor, and those at or above $\hat{y}$ are considered good. Two kernel density estimators are fitted to the parameter values associated with each group:

- $\ell(x) = p(x|\text{score} < \hat{y})$ — density of parameters associated with poor outcomes (9)
- $g(x) = p(x|\text{score} \geq \hat{y})$ — density of parameters associated with good outcomes (10)

New candidate parameter sets are proposed by sampling from $\ell(x)$ and selecting the candidate that maximizes the ratio $\ell(x)/g(x)$. This acquisition function has the intuitive interpretation of favouring parameter regions that are more likely under the good-outcome distribution than under the poor-outcome distribution, simultaneously rewarding proximity to previously successful configurations and penalising proximity to configurations that have already been shown to perform poorly. The ratio formulation is computationally tractable because it requires only the evaluation of the two density models rather than the construction of a full surrogate of the objective function.

The TPE algorithm is particularly well-suited to the predictive advisor optimization problem for three reasons. First, the parameter space is mixed continuous-discrete: tourist type mix weights are continuous variables, while infrastructure counts such as event numbers and accessibility targets are integers. TPE handles mixed spaces natively through the use of appropriate kernel types for each parameter domain, whereas gradient-based optimizers would require relaxation of the discrete variables. Second, the objective function — the composite KPI score — is a black box with no analytical gradient, making gradient-free methods a necessity. Third, the evaluation of the objective at each trial involves running the full derived metrics pipeline and KPI evaluation stack, which is computationally inexpensive per trial but benefits from the sample efficiency of a model-based approach relative to random search, allowing good solutions to be found within a budget of 1,500 trials that is practical on standard hardware.

The Optuna study is initialized with a fixed random seed via the TPESampler to ensure reproducibility of results across runs. The optimization direction is set to maximization of the composite sustainability score S defined in Section 5.5.2.

5.5.1 Search Space

Tourist Mix (always included)

The optimization search space is partitioned into two components: a core component covering tourist type mix weights and tourist volume, which is always included regardless of the optimization mode; and an optional infrastructure component that is activated when

include_infrastructure = True. The two modes allow the framework to be applied in contexts where infrastructure decisions are either within the control of the destination manager (full optimization) or fixed by existing conditions (mix-only optimization).

Tourist type mix weights. For each of the K archetypes $i = 1, \ldots, K$, a raw weight w_i is sampled independently from a continuous uniform distribution:

$$w_i \sim \text{Uniform}(0.01, 1.0) \quad \text{(continuous)} \quad (11)$$

The lower bound of 0.01 prevents any archetype from being completely excluded from the mix, ensuring that the normalized weights remain well-defined and that the optimizer retains the ability to recover excluded archetypes in subsequent trials. After sampling, the raw weights are normalized to sum to 100% and rounded to 4 decimal places:

$$\overline{w_i} \frac{100 \cdot w_i}{\sum_{j=1}^{K} w_j} \quad \text{(normalized to \%, 4 d.p.)} \quad (12)$$

The normalization is applied at every trial before the scenario is updated, ensuring that the sum-to-100 constraint is satisfied throughout the search without requiring it to be imposed as an explicit optimization constraint.

The total number of tourist nights is always included as an optimization variable, regardless of infrastructure mode. It is sampled from a discrete uniform distribution bounded by the destination's maximum bednights capacity:

$$N \sim \text{UniformInt}(0, N_{\max}) \quad (13)$$

This allows the optimizer to discover that reducing tourist volume — rather than merely redistributing it across archetypes — may improve the composite KPI score in scenarios where the destination is operating above its sustainability threshold.

Infrastructure parameters. When include_infrastructure = True, seven additional variables are introduced, each bounded by the destination's physical or operational constraints:

Table 23: Infrastructure parameter search domains for the optional optimization mode

Parameter	Search domain	Unit
hotel_certified_eco_label	$UniformInt(0,\ H_{available})$	number of hotels
blue_flag_beach_count	$UniformInt(0,\ B_{total})$	number of beaches
non_cultural_event_count	$UniformInt(0,\ 10)$	number of events
cultural_heritage_event_count	$UniformInt(0,\ 10)$	number of events
accommodation_accessibility_percentage	$UniformInt(0,\ 100)$	%
public_transport_accessibility_percentage	$UniformInt(0,\ 100)$	%
attraction_accessibility_count	$UniformInt(0,\ A_{total} + B_{total})$	Number of attractions

The upper bounds for hotel eco-certification and beach counts are set to the available stock at the current season level rather than the total baseline stock, preventing the optimizer from proposing configurations that are physically impossible given the season. The event counts are capped at 10

as a pragmatic upper bound reflecting realistic destination event programming capacity. The accessibility percentages are treated as continuous variables since they represent policy targets rather than discrete infrastructure counts. The attraction accessibility count is bounded by the combined total of cultural attractions and beaches, consistent with the definition of accessible attractions used in indicator C3c.

It is important to note that the infrastructure parameters represent aspirational targets or policy levers rather than directly observable quantities — the optimizer is searching over the space of what the destination could invest in or regulate, not over the space of what is currently observed. The practical interpretation of the optimized infrastructure values is therefore as a recommended policy direction rather than a precise operational prescription.

5.5.2 Objective Function

The optimization objective is the global sustainability score $S(\theta)$ defined in Section 5.5, here expressed explicitly as a function of the parameter vector θ:

$$\max_{\theta} S(\theta) = 100 - \frac{1}{N_{\text{ind}}} \sum_{i=1}^{N_{\text{ind}}} |U_i(\theta) - 100| \qquad (14)$$

where θ is the vector of optimization parameters (tourist mix weights, tourist nights, and optionally infrastructure variables), $N_{ind} = 33$ is the total number of KPI indicators across all four pillars, and $U_i(\theta)$ is the utilization percentage of indicator I evaluated under parameter configuration θ. The objective is maximized — a score of 100 represents perfect alignment of all indicators with their targets, and the optimizer seeks the parameter configuration that minimizes the mean absolute deviation from this ideal.

At each trial, the objective is evaluated through a three-step procedure. First, suggest_variables() proposes a candidate parameter vector θ by sampling from the search space distributions defined in Section 5.5.1 and normalizing the tourist mix weights. Second, scenario.update_parameters(θ) applies the proposed parameters to the scenario object and re-executes initialize_tourist_type_mix() and initialize_derived_metrics() in sequence, propagating the new weights through all dependent quantities — composite averages, CO_2 and water estimates, occupancy rate, pollution indices, revenue, and all other derived metrics. Third, a fresh KPIEvaluator instance is instantiated, initialize_indicators() and calculate_kpis() are called to evaluate all 33 indicators, and score() returns the scalar objective value to Optuna.

The re-instantiation of KPIEvaluator at every trial is a deliberate design choice that ensures complete isolation between trials — no state from a previous evaluation can contaminate the current one. This is particularly important for the ordinal indicators whose classifications depend on runtime estimation functions, where any caching of intermediate results could produce inconsistent scores.

5.5.3 Trial Budget and Runtime

The optimization study runs for 1,500 trials, a budget selected empirically to achieve consistent convergence of the TPE surrogate model across the mixed continuous-discrete search space while remaining within a wall time acceptable for interactive use. At an average evaluation time of approximately 50–100 milliseconds per trial on a single CPU core — reflecting the lightweight nature of the Python-based KPI evaluation stack — the total wall time per scenario is approximately 2–3 minutes.

It should be noted that the TPESampler is initialized without a fixed random seed in the current implementation, making optimization results stochastic: two runs on the same scenario with the same input data will in general produce different optimized parameter configurations, though the

objective scores should be similar if the study is well-converged. For reproducibility in reported results, for example, when comparing optimized scenarios across destinations or documenting results in a deliverable, a fixed seed should be passed to the sampler:

$$sampler = TPESampler(seed = 42) \quad (15)$$

This ensures that the sequence of candidate proposals is deterministic and that reported scores and parameter configurations can be exactly reproduced from the input data alone. The current codebase includes this fix in the run_optimization() function in optimization.py and should be verified before any results intended for publication are generated.

Convergence behaviour. TPE requires a warm-up phase of approximately 25–100 random trials before the surrogate model has sufficient observations to guide sampling effectively. Within the 1,500-trial budget, this warm-up represents at most 7% of total evaluations, after which the algorithm transitions to model-guided search. In practice, for the TOURAL search space — which has between 8 and 15 free parameters depending on the optimization mode — convergence of the best observed score typically stabilizes within 500–800 trials, meaning that the remaining trials provide diminishing returns. Future versions of the framework may therefore consider an early stopping criterion based on the improvement rate of the best score over a sliding window of trials, which would reduce wall time without sacrificing solution quality.

5.5.4 Result Extraction

Upon completion of the optimization study, the best parameter configuration is recovered from study.best_params, which contains the raw (un-normalized) weight values as sampled by the TPE algorithm. The extraction procedure mirrors the normalization applied during trial evaluation: raw weights are summed, each weight is divided by the total and multiplied by 100, and the result is rounded to 4 decimal places. The normalized weights are then merged back into the tourist type mix dictionary by updating the weight field of each archetype entry while preserving all other archetype properties unchanged. If infrastructure parameters were included in the search, they are extracted directly from study.best_params without further transformation and assembled into the optimized parameter dictionary alongside the updated tourist type mix. The complete optimized parameter set is then passed to scenario.update_parameters() to reconstruct the optimized scenario object, after which KPIEvaluator is called once more to produce the final optimized KPI profile for comparison and export.

5.6 Action Plan Generation

Delta Computation

Before the action plan prompt is constructed, the differences between the baseline and optimized scenarios are computed and structured into two delta objects: one covering changes in the tourist type mix and one covering changes in KPI utilization values. These deltas serve as the sole factual input to the language model, ensuring that the generated action plan is strictly grounded in the quantitative optimization results rather than generic sustainability advice. Tourist mix deltas (compute_mix_diffs). For each archetype i, the baseline weight w_i^b and optimized weight w_i^o are compared. If the two values differ, a delta record is created containing the archetype name, the baseline and optimized weights, and a direction classification:

$$direction_i \begin{cases} Increase, \ w_i^o > w_i^b \\ Decrease, \ w_i^o < w_i^b \end{cases} \quad (16)$$

Archetypes whose weight is unchanged between baseline and optimized scenarios are excluded from the delta record, reducing the prompt length and focusing the language model's attention on the actionable changes.

KPI results are first flattened from their nested category-indicator dictionary structure into a single-level dictionary via flatten_kpis(). For each indicator, a display value is extracted using the following priority order: utilization_percent is used if available; if not, flow is used; if not, optimal; and finally the raw value. This priority order ensures that the most policy-relevant representation of each indicator is used in the delta computation. Indicators whose display value is identical between baseline and optimized scenarios are excluded from the delta record. For the remaining indicators, the direction of change — closer to or further from the 100% target — is recorded alongside the baseline and optimized utilization values, providing the language model with a structured summary of which KPIs improved, which deteriorated, and by how much.

LLM Integration

Action sentences are generated through a series of individual calls to a locally hosted language model via the Ollama inference server. Two prompt templates are used, one for tourist mix changes and one for KPI changes, each designed to elicit a single concrete, place-level intervention sentence rather than generic policy language. Tourist mix action prompt. For each archetype with a non-zero weight delta, generate_tourist_type_action() constructs the following prompt and submits it to the model:

Table 24: Prompt template for tourist type mix action generation

You are a local tourism policy advisor. Destination: {destination_name}. TASK: The share of this tourist type changes from {b:.1f}% to {o:.1f}% (a {direction}). Propose ONE very concrete local intervention that could realistically cause this change. Use fewer than 25 words. Do NOT mention data, percentages, KPIs, tourists, 'campaign', 'strategy', or 'policy'. Write ONLY the action text, nothing else. Tourist type: {t}

For each KPI with a non-zero display value delta, generate_kpi_action() constructs an analogous prompt supplying the KPI name, category label, unit, and numeric before/after utilization values, with the same length and vocabulary constraints. Inference parameters. Both prompt templates are submitted to the same model instance with identical generation settings, chosen to minimize output variability and encourage factual, actionable responses:

Table 25: Language model inference parameters for action plan generation

Parameter	Value	Rationale
Model	llama3 (configurable)	Locally hosted via Ollama; no external API dependency
Temperature	0.2	Low temperature suppresses creative variation in favour of deterministic, factual outputs
Top-p	0.2	Nucleus sampling restricted to the highest-probability token mass, reinforcing consistency
Max tokens	60	Hard ceiling enforcing the 25-word constraint with a small buffer for tokenization overhead

The combination of low temperature and low top-p is intentional: the action plan is a policy-facing output intended to be read by destination managers, and consistency and conciseness are more valuable properties than creativity in this context. The 25-word constraint forces the model to produce immediately actionable sentences — naming a specific intervention type, venue, or service — rather than abstract recommendations.

The total number of LLM calls issued per run equals the number of significant deltas across both objects:

$$N_{calls} = |\Delta_{mix}| + |\Delta_{KPI}| \qquad (17)$$

In practice, for a full 33-indicator optimization run with K=8 archetypes, this typically yields between 20 and 40 calls per scenario, with a wall time of approximately 1–3 seconds per call depending on model size and hardware. The calls are issued sequentially rather than in parallel to avoid saturating the local inference server.

Output Format

The complete action plan is assembled as a plain markdown string with two clearly delimited sections. The tourist mix section lists one action sentence per archetype with a non-zero weight delta, and the KPI section lists one action sentence per indicator with a non-zero display value delta:

Table 26: Action plan output format for tourist type mix and KPI improvement recommendations

```
=== Tourist Type Mix Actions ===
- {type}: {direction} from {b:.1f}% -> {o:.1f}% -- {LLM action text}
...

=== KPI Improvement Actions ===
- {kpi} (Category {cat}): {b:.2f}{unit} -> {o:.2f}{unit} -- {LLM action text}
...
```

Each line in the tourist mix section identifies the archetype, the direction of change, the baseline and optimized weights in percentage points, and the generated intervention sentence. Each line in the KPI section identifies the indicator name, its pillar category, the unit of measurement, the baseline and optimized display values, and the generated action sentence. The markdown string is written to a plain text file for human review and may optionally be included in the JSON export payload for downstream consumption by a frontend or reporting system.

5.7 Results

North Sporades (Mid Season, Mix-Only Optimization)

Figure 8 - Figure 11 present screenshots from the tool interface for the North Sporades mid-season scenario, showing the baseline KPI profile, the optimized output, and the chart comparison across the four sustainability pillars.

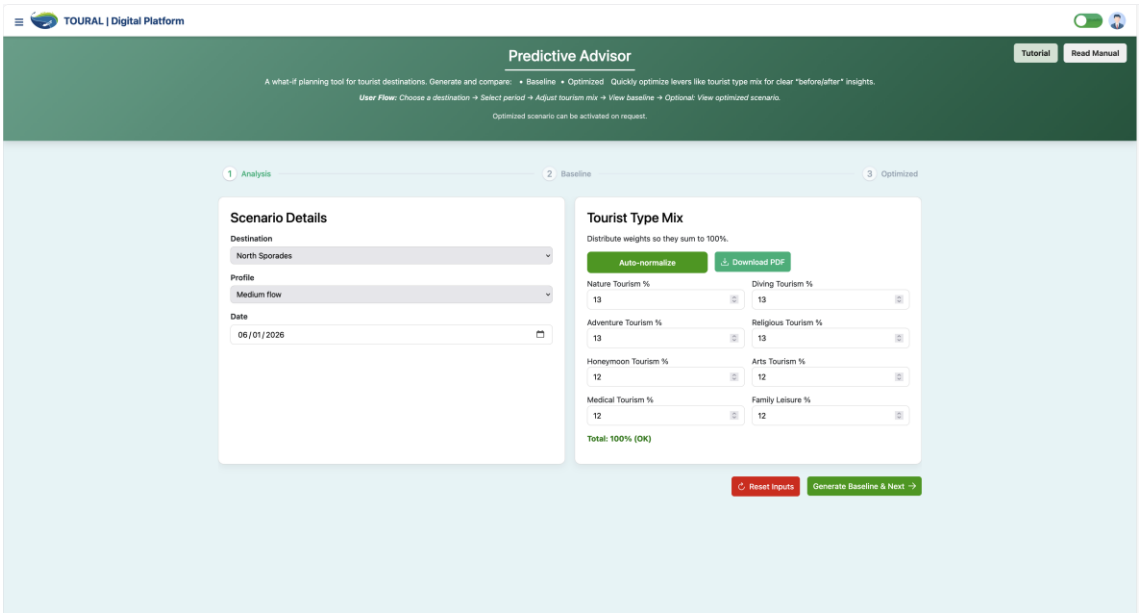

Figure 8: Initial configuration for the North Sporades scenario

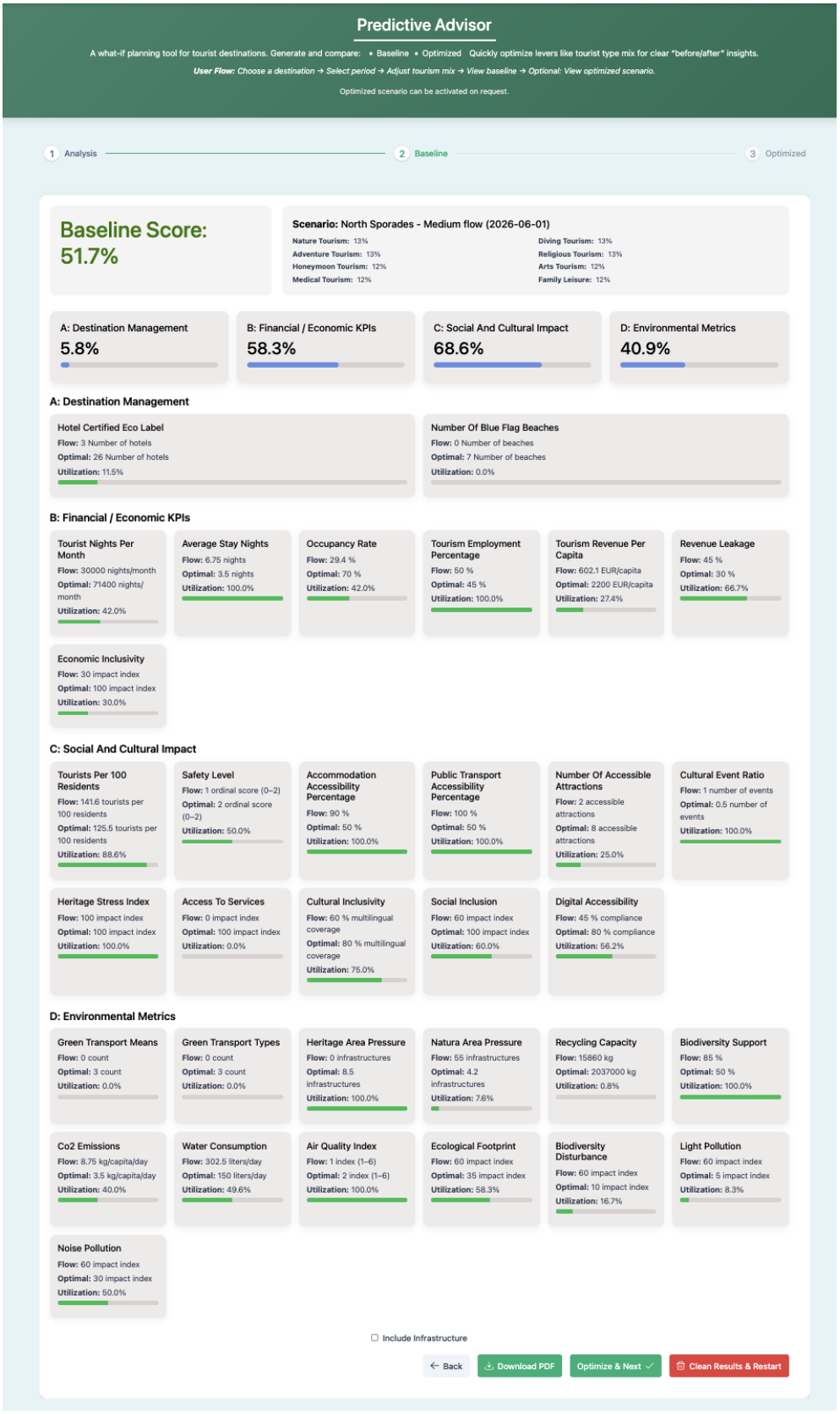

Figure 9: Baseline Results for the North Sporades scenario

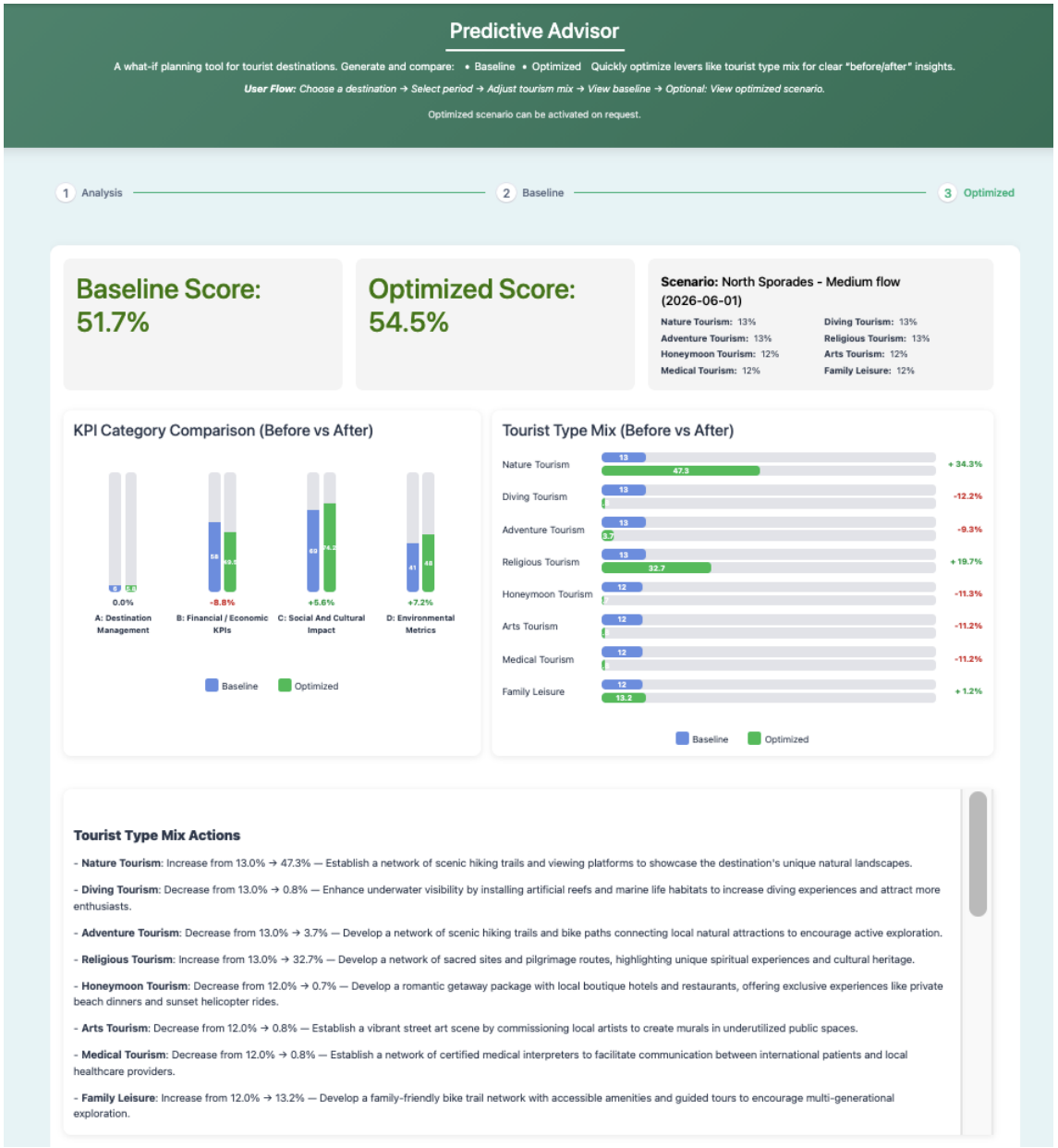

Figure 10: Optimization results for the North Sporades scenario

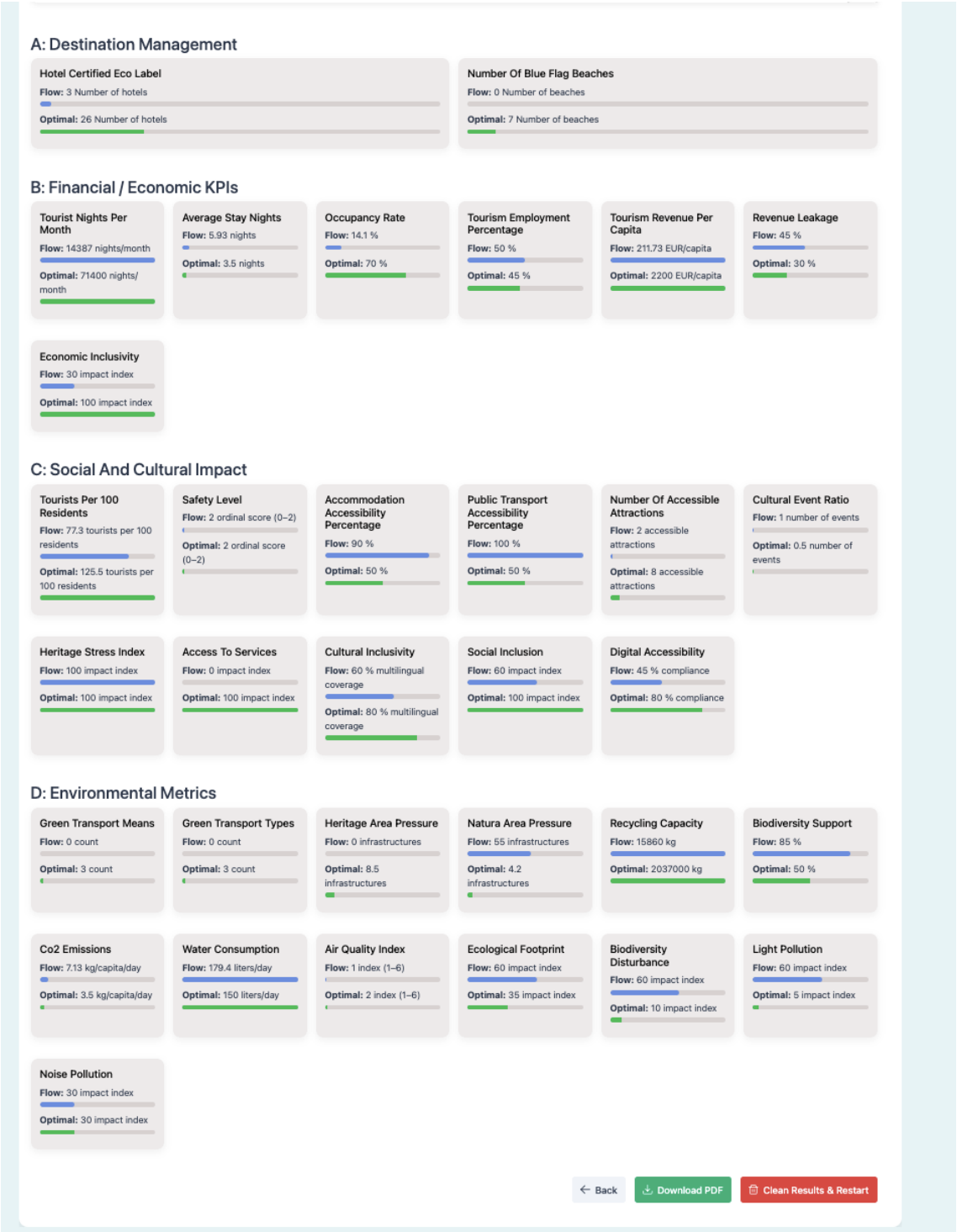

Figure 11: Optimization results for the North Sporades scenario

The North Sporades mid-season scenario is one of the more challenging in the dataset. The baseline composite score of 51.7% reflects a destination caught between two competing pressures: a mid-season volume of 30,000 tourist nights already generating measurable environmental and social stress, while simultaneously falling well short of economic targets. Tourist nights reach only 42% of capacity, occupancy 29.4% against a 70% target, and revenue per capita of 602 EUR just 27.4% of the 2,200 EUR benchmark. Pillar D compounds this picture: ecological footprint, biodiversity disturbance, light

pollution, and noise pollution are all rated moderate, and Natura area pressure is severely exceeded with 55 accommodations against a threshold of 4.2.

The optimized mix concentrates weight on Religious Tourism (rising from 13% to 50.4%) and Nature Tourism (from 13% to 33.5%), the two least resource-intensive archetypes by CO_2 and water consumption. This drives real environmental gains: water consumption falls from 302.5 to 179.8 L/day (improving from 49.6% to 83.4% utilisation), CO_2 from 8.75 to 6.74 kg/day (40.0% to 51.9%), and noise pollution crosses its target entirely (50.0% to 100%). Pillar D improves from 40.9% to 48.2%.

Pillar C also improves from 68.6% to 74.2%, with tourist pressure and safety level both reaching 100% utilisation — a cascading social co-benefit of the environmentally motivated rebalancing.

The critical finding, however, is Pillar B deterioration. Tourist nights fall to 14,727, occupancy to 14.4%, and revenue per capita collapses to 202 EUR (9.2% utilisation). The optimizer, constrained to mix-only search, correctly improves the composite score by reducing volume — but the result is practically counterproductive. No destination manager would accept a 70% revenue reduction for a 2.9-point composite gain.

This exposes a fundamental limitation of the equal-weight formulation for destinations already operating below economic capacity. North Sporades needs more tourists, not fewer — the problem is mix quality and infrastructure, not volume. A weighted objective prioritising Pillar B in mid- and high-season runs, combined with infrastructure optimisation, would likely produce a qualitatively different result. This run should be read as a diagnostic of the trade-off space, not an operational recommendation.

Finally, several indicators are structurally immovable regardless of mix. Natura area pressure remains at 7.6% — 55 establishments in protected zones against a tolerance of 4.2 — and represents the destination's most severe KPI gap and highest regulatory urgency. Pillar A scores 5.75%, reflecting near-total absence of eco-certification and Blue Flag designation. Green transport, recycling capacity, and access to services all score at or near 0%. These structural deficits suppress the composite score floor and can only be addressed through sustained infrastructure investment and regulatory enforcement.

Table 27: KPI utilization scores before and after optimization — North Sporades, mid season

Indicator	Pillar	Baseline (%)	Optimized (%)
eco-label hotels	A	11.5%	11.5%
blue flag beaches	A	0.0%	0.0%
Pillar A average	**A**	**5.8%**	**5.8%**
tourist nights/month	B	42.0%	20.6%
avg stay nights	B	100.0%	100.0%
occupancy rate	B	42.0%	20.6%
employment %	B	100.0%	100.0%
revenue/capita	B	27.4%	9.2%
revenue leakage	B	66.7%	66.7%

economic inclusivity	B	30.0%	30.0%
Pillar B average	**B**	**58.3%**	**49.6%**
tourist pressure	C	88.6%	100.0%
safety level	C	50.0%	100.0%
accomm. accessibility	C	100.0%	100.0%
transport accessibility	C	100.0%	100.0%
accessible attractions	C	25.0%	25.0%
cultural event ratio	C	100.0%	100.0%
heritage stress	C	100.0%	100.0%
access to services	C	0.0%	0.0%
cultural inclusivity	C	75.0%	75.0%
social inclusion	C	60.0%	60.0%
digital accessibility	C	56.2%	56.2%
Pillar C average	**C**	**68.6%**	**74.2%**
green transport means	D	0.0%	0.0%
green transport types	D	0.0%	0.0%
heritage area pressure	D	100.0%	100.0%
natura area pressure	D	7.6%	7.6%
recycling capacity	D	0.8%	0.8%
biodiversity support	D	100.0%	100.0%
CO2 emissions	D	40.0%	51.9%
water consumption	D	49.6%	83.4%
air quality index	D	100.0%	100.0%
ecological footprint	D	58.3%	58.3%
biodiversity disturbance	D	16.7%	16.7%
light pollution	D	8.3%	8.3%
noise pollution	D	50.0%	100.0%
Pillar D average	**D**	**40.9%**	**40.9%**

Table 28: Optimized tourist type mix and LLM-generated policy actions — North Sporades, mid season

Tourist Type	Direction	Baseline (%)	Optimized (%)	Recommended Action
Nature Tourism	Increase	13.0%	33.5%	Establish a network of scenic hiking trails and observation decks offering unique views of the destination's diverse flora and fauna.
Diving Tourism	Decrease	13.0%	4.0%	Establish a marine protected area to safeguard coral reefs and marine life, enhancing diving experiences while promoting sustainable tourism practices.

Adventure Tourism	Decrease	13.0%	2.4%	Develop a network of scenic hiking trails and bike paths connecting local natural attractions to encourage active exploration.
Religious Tourism	Increase	13.0%	50.4%	Develop a sacred site renovation project to restore and enhance the spiritual ambiance of the destination's most revered religious landmarks.
Honeymoon Tourism	Decrease	12.0%	1.4%	Develop a romantic getaway package with local boutique hotels and restaurants, offering exclusive deals for newlyweds.
Arts Tourism	Decrease	12.0%	4.1%	Establish a vibrant street art scene by commissioning local artists to create murals in underutilised public spaces.
Medical Tourism	Decrease	12.0%	2.8%	Establish a network of on-site medical interpreters and cultural liaisons to facilitate seamless communication and care for international patients.
Family Leisure	Decrease	12.0%	1.4%	Develop a network of family-friendly bike trails and parks with amenities like picnic areas, playgrounds, and splash pads.

Table 29: KPI Improvement Actions

KPI	Pillar	Baseline	Optimized	Recommended Action
Tourist nights/month	B	42.0%	20.6%	Implement a 'Shoulder Season' accommodation discount program for visitors arriving between April and June.
Occupancy rate	B	42.0%	20.6%	Implement a 'Green Week' discount program for eco-friendly accommodations, offering 15% off rates during the slowest period.
Revenue/capita	B	27.4%	9.2%	Implement a 'Slow Tourism' certification program for local accommodations and tour operators, promoting low-impact, high-value experiences.
Tourist pressure	C	88.6%	100.0%	Develop a 'Tourist-Friendly' accommodation package incentive for

				local businesses, offering discounts to visitors who stay in locally-owned hotels and B&Bs.
Safety level	C	50.0%	100.0%	Install solar-powered streetlights along high-traffic areas to enhance visibility and reduce risk of accidents.
CO2 emissions	D	40.0%	51.9%	Implement a 'Green Taxi' fleet with electric and hybrid vehicles, replacing 20% of existing taxis within the first year.
Water consumption	D	49.6%	83.4%	Install low-flow showerheads and faucets in all hotels and accommodations.
Noise pollution	D	50.0%	100.0%	Implement a noise-restrictive zoning ordinance for new developments in high-traffic areas, capping decibel levels at 60 dB.

6 Implementation

6.1 TOURAL ICT Suite Overview

The Flow Capacity Tool (T5.1), as presented in Chapter 4 and the Predictive Advisor (T5.2), as presented in Chapter 5, are developed as core components of the TOURAL ICT Suite. Together, they support the TOURAL Digital Platform, which provides a unified environment for accessing, orchestrating, and integrating the project's digital tools.

In Figure 12, the overall TOURAL Digital Platform architecture is presented, showing how the Flow Capacity Tool and the Predictive Advisor are integrated within a common front-end environment and connected to the backend processing services.

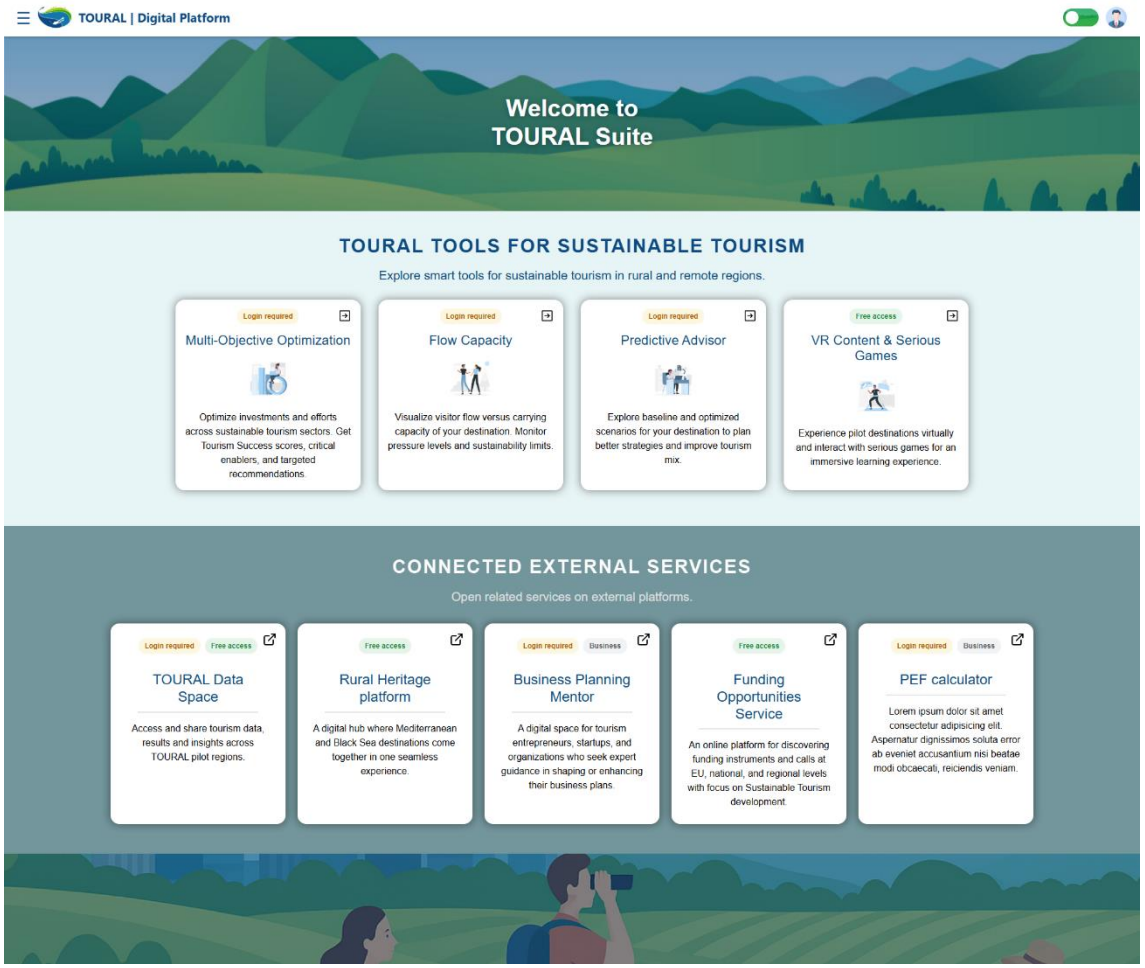

Figure 12: TOURAL Digital Platform

The TOURAL Digital Platform is a modular and interoperable environment that integrates data processing, simulation, and decision-support functions within a unified framework. It supports the full workflow of the system, from data input and modelling to evaluation, and is implemented as a full-stack solution with an Angular frontend and a Node.js backend. The frontend communicates with the backend, which acts as a gateway to the predictive engines. Depending on the computational requirements of each tool, the backend either forwards requests to a local Python microservice managed by PM2 or brokers them to an external processing machine through asynchronous polling and webhook callbacks.

6.2 Technical Architecture and Stack

The TOURAL ICT Suite follows a modular architecture composed of interconnected components responsible for data management, model execution, and user interaction.

The Toural platform is implemented using a modern full-stack architecture, consisting of an Angular frontend and a Node.js backend, integrated with external processing services.

The system is implemented using a combination of backend and frontend technologies, supported by APIs enabling communication between components as it is displayed in Figure 13.

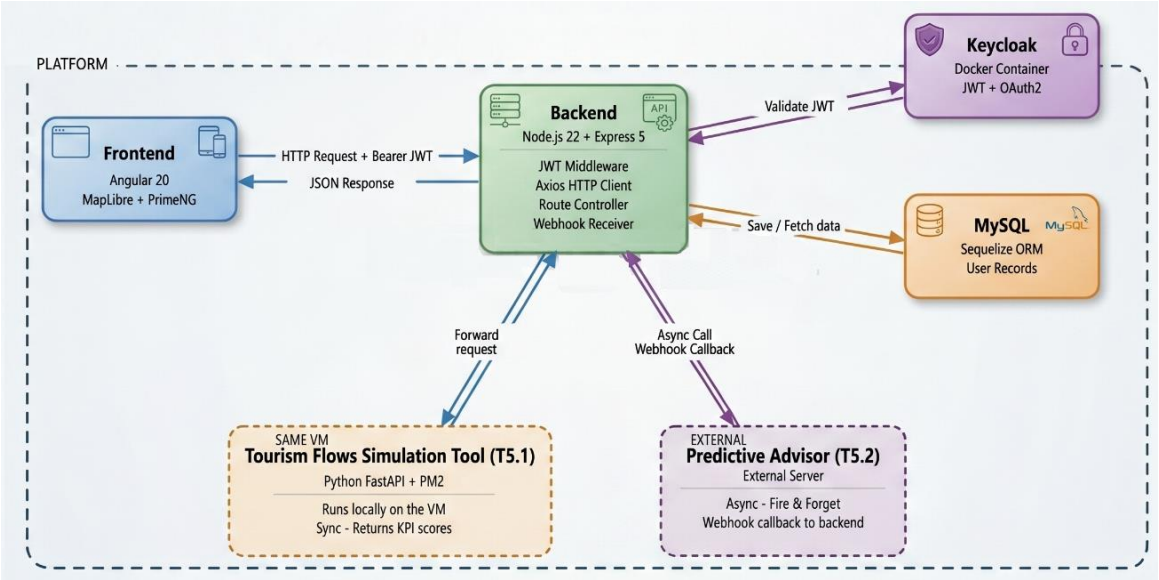

Figure 13: Toural platform architecture

6.2.1 Frontend Presentation

The platform features a modern architecture with a client-side application built on Angular. The frontend leverages robust UI component libraries, including PrimeNG and Bootstrap, to deliver a responsive and accessible user experience, while integrating MapLibre GL for specialized geographic mapping capabilities.

To enhance user experience and platform accessibility, the frontend integrates a context-aware, interactive onboarding system across all core analytical tools (Flow Capacity Tool, Predictive Advisor, and the Multi-Objective Optimization Tool). Powered by the **Driver.js** library, this step-by-step wizard provides dynamic guidance tailored to the user's current workflow phase for every available tool. The implementation leverages DOM-manipulation for automatic scrolling and element highlighting, ensuring users are seamlessly guided through complex data inputs, geographical selections, and analytical dashboards. Furthermore, the wizard utilizes browser session storage for intelligent state management as the user has the option of selection to not see the tutorial again, ensuring that returning users are not disrupted by repetitive tutorials during active analytical sessions. A table of the technologies used for frontend is presented in the Table 30.

Table 30: Frontend technologies used

Category	Technologies Used
Core Framework	Angular 20.x, Angular CLI 20.3.6, TypeScript
UI Components	PrimeNG, Bootstrap
Mapping & Export	MapLibre GL, jsPDF, html2canvas
State & Utilities	LocalStorage (Async persistence), ngx-translate, Driver.js (Onboarding)

6.2.2 Backend Infrastructure

The core backend infrastructure is built on **Node.js** utilizing the **Express.js** framework, acting as the primary API gateway for the platform. It handles cross-origin resource sharing (CORS), environment configurations, and securely routes requests from the frontend client. The backend repository is organized into a modular structure, clearly separating database configurations, controllers, middleware, and API routes.

Depending on the tool's computational needs, the backend routes requests differently:

- **Local Microservices:** Specialized Python microservices (such as Multi-Objective Optimization Tool 3.2 and Flow Capacity Tool 5.1) are in the broader project directory and managed locally.
- **External Services:** Tool 5.2 (Predictive Advisor) is integrated directly within the backend architecture as a broker service. Rather than running locally on the virtual machine, it operates by making secure, external API calls to a separate dedicated machine that processes the heavy predictive data.

A list of technologies used in the backend are displayed in the Table 31.

Table 31: Backend technologies used

Category	Technologies Used
Runtime & API	Node.js (v22.20.0), Express.js (v5.x)
Database & ORM	MySQL, Sequelize
Security & Auth	Keycloak (JWT-based), jsonwebtoken, jwks-rsa, keycloak-connect
Tools & Client	Axios (HTTP Client), dotenv (Environment Config), Nodemon (Development Tooling)

6.2.2.1 *Data Persistence and ORM Integration*

Data management and persistence are handled by a MySQL relational database, which operates as an active system service on the virtual machine.

The Node.js backend interfaces with this database utilizing Sequelize, a robust, promise-based Object-Relational Mapper (ORM). This integration abstracts raw SQL queries and manages the application's data layer seamlessly.

During the backend server startup sequence, Sequelize automatically authenticates the database connection and synchronizes the defined data models (such as user records), ensuring that the underlying database schema is properly created and maintained without requiring manual database migrations.

6.2.2.2 *Identity and Access Management (IAM)*

User security, authentication, and authorization are centralized through Keycloak, which is deployed in a Docker container.

The Express backend securely communicates with Keycloak's REST API using client credentials to manage the user lifecycle, including registrations, password resets, and account deletions.

During login, the backend facilitates the authentication flow by validating credentials against Keycloak and issuing standard JSON Web Tokens (JWT), including both access and refresh tokens.

This architecture ensures that user identities are securely managed outside the core database while enabling robust, token-based API protection across the entire platform.

6.2.3 Deployment & Process Management

For production deployment, the Angular application is compiled and hosted on an **Ubuntu** virtual machine. It is served by an **Nginx** web server configured as a reverse proxy. To ensure secure data transmission over HTTPS, SSL/TLS certificates are provisioned and automatically managed using **Certbot** and Let's Encrypt. To ensure high availability and continuous operation, **PM2** is used as the process manager to daemonize and continuously manage the core Node application and local Python microservices on the production environment.

6.2.4 Architecture & Stack of Flow Capacity Tool (T5.1) & Predictive Advisor (T5.2)

Tool 5.1 operates on a three-tier architecture. The Angular frontend handles interactive spatial and temporal inputs. The Express.js backend acts as a secure reverse proxy and gateway. The actual predictive processing is decoupled into a dedicated Python FastAPI microservice, which is executed locally from a separate directory and managed in the background using PM2 as described above.

- **Frontend Integration:** Relies on MapLibre GL to render interactive map layers and vector polygons. It uses a locking mechanism to prevent race conditions during API calls. Moreover a Driver.js tutorial displayed as soon as the tool opens, helping the users understand the flow of the events.
- **Backend Logic:** Exposed via a protected route (`/predict-flow`) utilizing Keycloak token validation middleware. It acts as an HTTP client (via Axios) to forward payloads to the internal Python FastAPI engine.

Tool 5.2 utilizes a highly decoupled, asynchronous architecture due to the intensive computational requirements of its predictive models. Instead of running locally, the Node.js backend acts as an intermediary broker between the Angular frontend and an external, dedicated processing machine.

- **Frontend Integration:** Uses PrimeNG's p-stepper for a multi-step wizard, complemented by a phase-specific Driver.js tutorial that dynamically adapts to the user's current step (Scenario Input, Baseline Results, or Optimized Plan). It utilizes a polling mechanism, calling the backend at fixed intervals until remote calculations are complete.
- **Backend Logic:** Employs a "fire-and-forget" forwarding mechanism. It registers a pending state in an in-memory Map, forwards the payload to the external processing engine, and waits for a Webhook callback from the external service to update the stored state for the frontend to retrieve.

6.3 API and Interoperability Layer

6.3.1 Flow Capacity Tool (T5.1) API Specification

As described in 6.3.1, the Flow Capacity Tool API exposes a primary endpoint to handle tourism flow prediction requests. The details of this endpoint, including its HTTP method, path, and core functionality, are outlined in Table 32.

Table 32: Core API Endpoint for the Flow Capacity Tool

POST /api/tools/predict-flow	Receives destination and temporal parameters to calculate and return tourism flow predictions and sustainability KPIs.

To successfully execute a prediction, the frontend must provide a specific set of spatial and temporal variables. Table 33 defines the data dictionary and details the required parameters that must be included in the body of the POST request.

Table 33: Request Body Parameters for /predict-flow

Name	Type	Required	Description
year	Integer	Yes	Selected year from the timeline. Must be a 4-digit number, e.g., 2026.
month	Integer	Yes	Month index (1–12) representing January–December.
region	String	Yes	Geographical region name. Example: N. Sporades.
municipality	String	Yes	Municipality name within the selected region.
flow_model_type	String	Yes	Model configuration type. Currently fixed to global.

To illustrate how these parameters are structured in practice,

Table 34 provides a complete example of the JSON request payload sent by the Angular client to the Node.js backend.

Table 34: Example JSON Request Payload

```
{
  "year": 2026,
  "month": 4,
  "region": "sporades",
  "municipality": "alonnisos",
  "flow_model_type": "global"
}
```

Finally, Table 35 demonstrates the structure of a successful JSON response, which is directly used by the frontend to render the interactive dashboards.

Table 35: Example JSON Response Payload


```
{
 "overall_score": 71.19,
 "flow_class": "low",
 "predicted_flow_rounded": 867,
 "section_scores": {
  "section_Environmental": 66.07,
  "section_SocialCultural": 79.12,
  "section_Economic": 68.38
}
```


6.3.2 Predictive Advisor (T5.2) API Specification

Due to the intensive computational requirements of the Predictive Advisor, this tool utilizes an asynchronous, decoupled architecture. The integration is divided into three distinct operational phases: initializing the scenario, receiving the webhook callback from the external engine, and polling for the results.

Understanding Baseline vs. Optimized Scenarios: The Predictive Advisor operates on two primary modes dictated by the `optimize` boolean flag:

- **Baseline Scenario (`optimize: false`):** Calculates current sustainability KPIs based purely on the user's initial tourist distribution mix.
- **Optimized Scenario (`optimize: true`):** Requests the external processing engine to calculate an ideal, recommended tourist mix and generate a strategic action plan to actively improve the destination's sustainability score.

The flow of the endpoints used when the tool is called are presented below, starting with “A” the Scenario Initialization Endpoint , then we have “B” the Webhook Receiver Endpoint , ended up with “C” the Polling Endpoint.

A. Scenario Initialization Endpoint

The primary endpoint receives the destination parameters and tourist distribution mix from the Angular frontend. It forwards this data to the external processing engine and returns an immediate acknowledgment. Table 36 outlines this endpoint.

Table 36: Init Scenario Endpoint

POST	Receives scenario parameters from the frontend, forwards them to the external engine, and returns an immediate acknowledgment.

/api/tools/predictive-advisor/dashboard	

Table 37 details the required structure of the JSON payload sent by the frontend to initiate either a Baseline or Optimized calculation.

Table 37: Request Payload Structure for Init Scenario

Name	Type	Required	Description
meta.userId	String	Yes	The Keycloak ID of the requesting user.
meta.scenario	Object	Yes	Contains the destination name, profile, and selected date.
optimize	Boolean	Yes	Set to false for Baseline calculation, or true for Optimized calculation.
infrastructure	Boolean	Yes	Flag to include infrastructure constraints in calculations.
touristTypes	Array	Yes	Array of objects defining the tourist mix (e.g., Nature Tourism, Diving Tourism) and their respective percentages (value).

To illustrate how these parameters are structured, Table 38 provides a complete example of the JSON request payload, while Table 39 shows the immediate acknowledgment returned by the Node.js backend.

Table 38: Example JSON Request Payload (Init Scenario)

```
{
  "meta": {
    "userId": "*********",
    "scenario": {
      "destination": "North Sporades",
      "profile": "Low flow",
      "date": "2026-04-21"
    },
    "postedAt": "2026-04-21T19:06:54.060Z",
    "tool": "predictive advisor 5.2"
  },
  "optimize": false,
  "infrastructure": false,
  "touristTypes": [
```

```
    { "name": "Nature Tourism", "value": 13 },
    { "name": "Family Leisure", "value": 12 }
  ]
}
```


Table 39: Example JSON Response (Immediate Acknowledgment)


```
{
  "message": "Data forwarded to processing engine.",
  "storedUserId": "******"
}
```


B. Webhook Receiver Endpoint

Once the external processing engine completes its calculations, it pushes the data back to the Node.js server using a dedicated webhook endpoint, detailed in Table 40.

Table 40: Webhook Receiver Endpoint

POST /api/tools/predictive-advisor/results	Acts as a webhook to receive and store the completed predictive calculations pushed by the external processing engine.

The external engine must include the original Keycloak ID to ensure the data is matched to the correct user. Table 41 illustrates the expected webhook payload

Table 41: Example Webhook Request Payload (From External Engine)


```
{
  "meta": {
    "id": "************",
    "destination": "North Sporades",
    "date": "2026-04-21",
    "optimize": false
  },
  "scores": {
    "baseline": 55.6
  },
  "baseline_scenario": {
    "A: Destination management": { "category_average_score": 15 },
    "B: Financial / Economic KPIs": { "category_average_score": 60.1 }
```

```
  }
}
```

C. Polling Endpoint

While the external engine processes the data, the Angular frontend continuously polls the backend to check for completion. This GET endpoint is described in Table 42.

Table 42: Polling Endpoint

GET /api/tools/predictive-advisor/dashboard-results	Returns the final processed data if available, or a "pending" status if the external engine is still computing.

This polling request requires specific query parameters to identify the user and the type of scenario, as defined in Table 43.

Table 43: Query Parameters for Polling Endpoint

Name	Type	Required	Description
userId	String	Yes	The Keycloak user identifier.
type	String	No	Scenario type to retrieve. Accepted values: baseline (default) or optimize

If the external engine is still calculating, the endpoint returns a pending status (Table 44). Once the webhook has successfully updated the backend, the polling endpoint returns the complete dataset containing the sustainability scores and recommendations (

Table 45).

Table 44: Example JSON Response (Pending State)

```
{
 "status": "pending"
}
```

Table 45: Example JSON Response (Completed State)


```
{
 "message": "Data fetched successfully",
 "data": {
  "meta": {
   "id": "**********",
   "destination": "North Sporades",
   "optimize": false
  },
  "scores": {
   "baseline": 55.6
  },
  "baseline_scenario": {
   "A: Destination management": { "category_average_score": 15 },
   "B: Financial / Economic KPIs": { "category_average_score": 60.1 }
  }
 }
}
```

7 Conclusions

The TOURAL system integrates simulation and decision-support functionalities into a unified framework for analysing and improving tourism sustainability. By combining tourism-flow forecasting, sustainability assessment, optimisation, and recommendation generation, the framework supports a transition from descriptive tourism monitoring towards predictive and prescriptive decision-support.

Through the Flow Capacity Tool (T5.1), the system provides a structured representation of tourism dynamics, capturing how visitor flows influence sustainability indicators across environmental, social, and economic dimensions. The evaluation of the forecasting component demonstrated that the proposed modelling approach is capable of accurately estimating tourism flows across diverse pilot destinations while maintaining robustness under varying seasonal conditions. Furthermore, the framework enables the translation of tourism-flow predictions into interpretable sustainability assessments, supporting the identification of emerging pressure conditions and capacity-related challenges.

Building upon these outputs, the Predictive Advisor (T5.2) enables the exploration of alternative tourism development scenarios and the identification of improved sustainability configurations. The optimisation experiments conducted across the pilot destinations demonstrated consistent improvements in overall sustainability performance compared with baseline scenarios. The results further highlighted the importance of balancing environmental, social, and economic objectives, revealing how different tourism profiles and infrastructure configurations can influence destination sustainability outcomes.

The evaluation demonstrates that the TOURAL framework is capable of:

- representing tourism activity in a coherent and interpretable manner;
- forecasting tourism demand and assessing its sustainability implications;
- identifying pressure conditions and system imbalances;
- evaluating alternative tourism development scenarios;
- supporting decision-making through actionable recommendations and optimisation-based insights.

Despite the limitations associated with data availability, indicator completeness, and the use of proxy variables in some pilot regions, the scenario-based approach allows for meaningful analysis and comparison of alternative strategies. The integration of baseline and optimised scenarios provides stakeholders with a transparent mechanism for understanding potential intervention impacts and exploring alternative pathways towards more sustainable tourism development.

Overall, the TOURAL framework demonstrates how forecasting, sustainability assessment, optimisation, and decision-support can be integrated within a single operational environment. The resulting ICT suite contributes to the advancement of evidence-based destination management and provides tourism stakeholders with practical tools for designing more balanced, resilient, and sustainable tourism strategies.

8 References

[1] R. W. Butler, "The concept of a tourist area cycle of evolution: implications for management of resources," *Can. Geogr.,* vol. 24, no. doi: 10.1111/j.1541-0064.1980.tb00970.x., p. 5–12, 1980,.

[2] D. Buhalis and C. Costa, "Tourism Destination Management: Planning, Marketing and Impacts," *Goodfellow Publishers,* 2025.

[3] B. H. Farrell and L. Twining-Ward, "Reconceptualizing tourism," *Annals of tourism research ,* pp. 274-295, 2004.

[4] R. Baggio, "Symptoms of complexity in a tourism system," *Tourism analysis ,* pp. 1-20, 2008.

[5] S. Hartman, , "Destination governance in times of change: a complex adaptive systems perspective to improve tourism destination development.," *Journal of Tourism Futures ,* pp. 267-278, 2023.

[6] E. Kaya and Y. Karakuş, "Managing Tourism Destinations as Complex Adaptive Systems: An MCDM-Based Hybrid Governance Selection Model for Sustainable Regional Development.," *Systems,* vol. 402, 2026.

[7] E. Canestrelli and P. Costa, "Tourist carrying capacity: A fuzzy approach.," *Annals of tourism research ,* pp. 295-311, 1991.

[8] M. Pásková, G. Wall, D. Zejda and Z. Josef , "Tourism carrying capacity reconceptualization: Modelling and management of destinations.," *Journal of Destination Marketing & Management,* 2021.

[9] T. Mai and C. Smith, "Scenario-based planning for tourism development using system dynamic modelling: A case study of Cat Ba Island, Vietnam.," *Tourism management ,* pp. 336-354, 2018.

[10] J. Moreno, J. Parada and D. D. Peña-Miranda, "Agent-Based Model Applied for the Study of Overtourism in an Urban Context.," *Sustainability,* 2025.

[11] M. Keertana and C. Kumar, "Digital Twin Applications in Tourism Destination Management: Predictive Modelling for Tourist Flow and Resource Optimization.," *Asian Review of Social Sciences.,* pp. 43-49, 2025.

[12] E. C. a. P. Costa, "Tourist carrying capacity," *Ann. Tour. Res.,* pp. 295--311, Jan. 1991.

[13] G. W. D. Z. a. J. Z. M. Pásková, "Tourism carrying capacity reconceptualization: Modelling and management of destinations," *J. Destin. Mark. Manag.,* p. 100638, September 2021.

[14] N. C. S. G. a. J. V. D. B. D. Bertocchi, "Venice and Overtourism: Simulating Sustainable Development Scenarios through a Tourism Carrying Capacity Model," *Sustainability,* p. 512, January 2020.

[15] A. T.-D. G. C. J. P. M. J. K. a. G. M. X. Font, "The impact of sustainable tourism indicators on destination competitiveness: the European Tourism Indicator System," *J. Sustain. Tour.,* pp. 1608-1630, July 2023.

[16] M. L. G. a. A. Mariotti, "Sustainable tourism indicators as policy making tools: lessons from ETIS implementation at destination level," *J. Sustain. Tour.,* pp. 1719-1737, July 2023.

[17] M. L.-O. M. G. F. M. G. a. R. C. F. J. Blancas, "How to use sustainability indicators for tourism planning: The case of rural tourism in Andalusia (Spain)," *Sci. Total Environ.,* pp. 28-45, December 2011.

[18] A. P. a. B. P. K. Koens, "Is Overtourism Overused? Understanding the Impact of Tourism in a City Context," *Sustainability,* p. 4384, November 2018.

[19] Y.-T. H. a. C. Tsou, "From carrying capacity to destination resilience: an indicator framework for island tourism," *Tour. Geogr.,* pp. 141-162, February 2026.

[20] L. B. D. D. S. M. A. P. a. S. S. M. Drius, "Tackling challenges for Mediterranean sustainable coastal tourism: An ecosystem service perspective," *Sci. Total Environ,* pp. 1302-1307, February 2019.

[21] W. S. G. V. B. a. D. V. O. P. Vansteenwegen, "Metaheuristics for Tourist Trip Planning," *Metaheuristics in the Service Industry, Lecture Notes in Economics and Mathematical Systems,* pp. 15-31, 2009.

[22] W. S. a. D. V. O. P. Vansteenwegen, "The orienteering problem: A survey," *Eur. J. Oper. Res.,* pp. 1-10, February 2011.

[23] J. D. a. N. M. K. Sylejmani, "Planning the trip itinerary for tourist groups," *Inf. Technol. Tour,* pp. 275-314, September 2017.

[24] X. Y. a. X. J. H. Lyu, "Overtourism to Equilibrium: A System Dynamics & Multi-Objective Model for Sustainable Destinations," *arXiv,* 2025.

[25] J. a. B. R. a. K. B. Bergstra, "Algorithms for Hyper-Parameter Optimization," *Proceedings of the 25th International Conference on Neural Information Processing Systems,* pp. 2545-2554, 2011.

[26] S. S. T. Y. T. O. a. M. K. T. Akiba, "Optuna: A Next-generation Hyperparameter Optimization Framework," *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage AK USA: ACM,* pp. 2623-2631, July 2019.

[27] I. E. a. S. European Commission: Directorate-General for Internal Market, "The European Tourism Indicator System – ETIS toolkit for sustainable destination management," Publications Office, Luxembourg, 2016.

9 Annex – Flow Capacity Tool

Danube Delta Tourism-Flow Prediction Performance

Table 46: Comparison of tourism-flow prediction models for Danube Delta

Destination	Model	MAE	RMSE	MAPE	R^2
Danube Delta	Random Forests	2560.35	3491.65	11.70664	0.909528
Danube Delta	XGBoost	1035.645	1382.458	7.347763	0.985817
Danube Delta	HistGradientBoosting	2331.712	3419.038	10.29427	0.913252
Danube Delta	**CNN-BiLSTM**	**228.4259**	**305.4643**	**13.69227**	**0.919618**

For the Danube Delta pilot, the CNN-BiLSTM achieved the lowest MAE and RMSE values by a substantial margin. In particular, compared with XGBoost, MAE decreased from 1035.6450 to 228.4259 and RMSE decreased from 1382.4580 to 305.4643. This corresponds to an approximate reduction of 77.9% in both MAE and RMSE. These results show that the CNN-BiLSTM model was particularly effective in reducing the magnitude of absolute prediction errors for the Danube Delta dataset. However, XGBoost achieved the best MAPE and R^2 values for this pilot, indicating that it provided stronger relative-error performance and goodness-of-fit according to these two metrics.

Table 47: CNN-BiLSTM feature-space comparison for Danube Delta

Destination	Experiment	N_features	MAE	RMSE	MAPE	R^2
Danube Delta	E1 core temporal	51	612.3840	804.672	9.8420	0.9513
Danube Delta	E2 core plus context	65	356.7420	472.615	16.4380	0.8875
Danube Delta	**E3 all features**	**87**	**228.4259**	**305.4643**	**13.6922**	**0.9196**

For the Danube Delta pilot, the CNN-BiLSTM model also achieved the best performance when the full feature set was used. The E3 all-feature configuration obtained the lowest MAE, RMSE, and MAPE values and the highest R^2 score. Compared with the E1 core temporal configuration, the E3 setup reduced MAE from 612.3840 to 228.4259 and RMSE from 804.6720 to 305.4643. This corresponds to an approximate reduction of 62.7% in MAE and 62.0% in RMSE. MAPE also decreased from 19.8420% to 13.6923%, while R^2 increased from 0.8513 to 0.9196. These results indicate that historical demand and seasonality alone were not sufficient to capture the tourism-flow dynamics of this pilot and that the richer feature space provided important additional predictive information.

The E2 core plus contextual configuration also improved the results compared with the E1 baseline, reducing MAE from 612.3840 to 356.7420 and RMSE from 804.6720 to 472.6150. However, the best performance was achieved only when the complete selected feature set was used. This suggests that, for Danube Delta, the additional transport, environmental, search-interest, tourism-related, and engineered variables included in E3 provided useful complementary information rather than introducing noise. Therefore, for both North Sporades and Danube Delta, the all-feature configuration represents the strongest CNN-BiLSTM setup, supporting the use of the richest available feature space for the operational tourism-flow prediction component.

KPI-based Impact Results

For the Danube Delta pilot, the tool produced a predicted tourist flow of 27767 visitors for June 2026, corresponding to a high-flow condition. The Danube Delta case provides a different

sustainability profile compared with the North Sporades pilot. While the economic score is stronger, several social-cultural and environmental indicators reveal important pressure points related to accessibility, inclusiveness, ecological footprint and biodiversity disturbance.

Table 48 presents the Environmental Conditions KPI results for the Danube Delta pilot. The environmental profile is relatively positive overall, with an Environmental Conditions score of 72%. Several indicators remain stable or close to the reference condition, including air quality, light pollution and noise pollution. The absence of tourism accommodation and attraction infrastructure in cultural heritage protected areas also indicates no additional pressure in this specific category. However, the results also highlight important weaknesses. The lack of green transport means indicates limited sustainable mobility options, while the ecological footprint and biodiversity disturbance indicators remain high compared with the reference values. These results suggest that, although some environmental pressures are controlled, the pilot area remains sensitive to ecosystem-related impacts and requires attention to mobility, biodiversity and resource-use management.

Table 48: Environmental Conditions KPI Results for the Danube Delta Pilot - June 2026

Kpi	Predicted	Optimized	Stable/ Dynamic	Message
Number of green transport means	0	1	Stable	Worse by 1 Number of means
Types of green transport means	N/A	N/A	Stable	-
Number of tourism accommodation and attraction infrastructure located in cultural heritage protected areas	0	0	Stable	No change
Number of tourism accommodation and attraction infrastructure located in natura protected areas	0	1	Stable	Better by 1 Number of infrastructures
Percentage of local enterprises actively supporting protection of local biodiversity	N/A	20	Stable	-
CO2 emissions per capita/day	N/A	N/A	Dynamic	-
Water Consumption per Tourist	N/A	N/A	Dynamic	-
Air Quality Index	1	1	Dynamic	No change
Land/water resources used (Ecological Footprint per Tourist)	90 (high)	60 (medium)	Stable	Worse by 30%
Impacts on key species/habitats (Biodiversity Disturbance Index)	90 (high)	60 (medium)	Stable	Worse by 30%
Artificial lighting changes in protected or dark-sky zones	5 (low)	5 (low)	Stable	No change
dB increase in tourism zones vs baseline (Noise Pollution Index)	30 (low)	30 (low)	Stable	No change

Table 49 presents the Social and Cultural Dynamics KPI results for the Danube Delta pilot. This dimension records a score of 59%, indicating a moderate social-cultural condition. The predicted

number of tourists per 100 residents remains below the reference value, which suggests that community-pressure levels are manageable under the simulated high-flow scenario. Safety, heritage-site stress and access to local services also remain stable. However, the results reveal several areas requiring improvement. Public transport accessibility, attraction accessibility, cultural inclusivity, social inclusion and digital accessibility remain below the desired reference values. The results also indicate limited traditional/local cultural event activity compared with the reference condition. Therefore, the social-cultural profile of the Danube Delta pilot is not critical, but it shows clear gaps in accessibility, inclusive tourism provision and cultural-service readiness.

Table 49: Social & Cultural KPI Results for Danube Delta Pilot - June 2026

Kpi	Predicted	Optimized	Stable/ Dynamic	Message
Number of tourists per 100 residents	300	400	Dynamic	Better by 100 Number of tourists per 100 residents
Level of safety (low/medium/high)	1 (medium)	1 (medium)	Stable	No change
Percentage of commercial accommodation establishments participating in recognised accessibility information schemes	N/A	N/A	Stable	-
Percentage of public transport that is accessible to people with disabilities	10	15	Stable	Worse by 5%
Percentage of tourist attractions that are accessible to people with disabilities	0	10	Stable	Worse by 10%
Number of the destination's events that are not traditional/cultural relevant	1	2	Stable	Worse by 8 Number of events
Number of the destination's events that are focused on traditional/local culture and heritage	1	4	Stable	Worse by 4 Number of events
Visitor volume per m², time-based load, etc. (Heritage Site Stress Index)	60 (medium)	60 (medium)	Stable	No change
Impact of tourism on local access to transport, health, and other infrastructure	60 (medium)	60 (medium)	Stable	No change
Signage and guides in multiple languages, cultural sensitivity	10 (low)	40 (low-medium)	Stable	Worse by 30%
Policies for LGBTQ+ travelers, families, seniors (Social Inclusion)	30 (low)	60 (medium)	Stable	Worse by 30%
Screen-reader compatible websites, mobile app usability (Digital Accessibility)	5	10	Stable	Worse by 5 %

Table 50 presents the Economic Aspects KPI results for the Danube Delta pilot. This dimension achieves the strongest performance for the pilot, with an Economic Aspects score of 80%. The number of tourist nights per month is substantially higher than the reference value, indicating strong tourism activity and high demand. Direct tourism employment also remains relatively high, although it is still below the reference value. However, not all economic indicators perform positively. Average length of stay is lower than the reference value, suggesting that visitors may not remain in the destination for as long as desired. Tourism revenue per capita is also considerably below the reference value, showing that high visitor numbers do not automatically translate into proportional local economic benefit. Economic inclusivity also remains below the reference condition, indicating that the availability of free or low-cost options and price diversity could be improved.

Table 50: Economic Aspects KPI Results for the Danube Delta Pilot - June 2026

Kpi	Predicted	Optimized	Stable/ Dynamic	Message
Number of tourist nights per month	69417	32000	Dynamic	Better by 37418 Number of nights
Average length of stay of tourists (nights)	2.5	3.2	Dynamic	Worse by 0.7 Number of nights
Tourism Revenue per Capita (Tourism Revenue)	24.99	70	Dynamic	Worse by 45.01 EUR/capita
% of tourism spending that leaves the region (Revenue Leakage index)	N/A	N/A	Stable	-
Availability of free/low-cost diversity (Economic Inclusivity)	35	45	Stable	Worse by 10 Price Diversity Index
Direct tourism employment as percentage of total employment	65	75	Stable	Worse by 10%

The aggregated results for the Danube Delta pilot show an Environmental Conditions score of 72%, a Social and Cultural Dynamics score of 59%, and an Economic Aspects score of 80%. Based on these section-level scores, the overall sustainability score is 70%, indicating a generally positive but still improvable condition. The result reflects a stronger economic profile compared with the North Sporades pilot, but also highlights important social-cultural and environmental gaps. In particular, the Danube Delta pilot performs well in terms of tourism nights and overall economic activity, but it requires further attention to accessibility, inclusiveness, cultural-service provision, sustainable mobility and biodiversity-related pressure.

10 Annex – Predictive Advisor Tool

Predictive Advisor - Environmental and Economic Estimation Functions

Table 51 summarizes all fourteen functions. For each function the table lists the inputs, a concise description of the estimation model including key formulas, the output type and precision, and the calibration parameters that may require adjustment when the framework is applied to destination contexts outside the Mediterranean and Central European regions covered by the TOURAL pilot sites.

Table 51: Summary of environmental and economic estimation functions.

Function	Inputs	Model description	Output	Key parameters
Estimate CO_2 emissions	Baseline CO_2 rate, green transport connection count, eco-certified hotel count	Applies two independent exponential decay reductions to the baseline emission rate — one driven by green transport availability and one by eco-label hotel penetration — and blends them with a 60/40 weighting that reflects the greater mitigation potential of transport over accommodation. The combined reduction is capped at 50% to prevent unrealistic results in well-equipped destinations.	Adjusted kg CO_2 per capita per day, rounded to 2 d.p.	Transport decay rate: 0.3; hotel decay rate: 0.1; maximum reduction: 50%; transport weight: 60%
Estimate water consumption	Baseline litres per day, season level, eco-certified hotel count, total hotels, occupancy rate, average stay nights	Adjusts the baseline daily water figure through four independent multiplicative factors: eco-certification (reduces consumption by up to 50% proportional to certified share); seasonality (10% uplift in high season, 10% reduction in low season); occupancy pressure (linear surcharge above 60% occupancy, up to 10%); and stay-length efficiency (15% surcharge for very short stays of two nights or fewer, up to 10% reduction for stays beyond five nights).	Adjusted litres per tourist per day, rounded to 1 d.p.	Season multipliers: 1.10 / 1.00 / 0.90; occupancy threshold: 60%; stay thresholds: 2 and 5 nights
Estimate tourism revenue	Tourist nights, average nightly expenditure, resident population, leakage rate, cultural event count, non-	Computes gross revenue as tourist nights multiplied by average nightly expenditure, then retains the fraction not lost to leakage. Two independent event multipliers — one for cultural heritage events and one for non-cultural events — are applied to the retained revenue	EUR per capita per month, rounded to 2 d.p. Returns 0 if population is zero.	Cultural event uplift cap: +15%; non-cultural cap: +8%; combined maximum uplift:

	cultural event count	to reflect the additional local spending stimulated by events. The result is divided by the resident population to yield a per-capita monthly figure.		approximately +23%
Estimate occupancy rate	Total tourist nights, maximum bednights capacity	Divides total tourist nights by the season-adjusted maximum bednights capacity to express the fraction of accommodation in use. Because the capacity denominator already reflects the seasonal reduction in operational hotel stock, the occupancy rate implicitly accounts for the lower bed availability in shoulder and off-peak periods.	Percentage, rounded to 1 d.p. Returns 0 if capacity is zero.	Capacity derived from season-adjusted hotel count, rooms per hotel, beds per room, and 30-day month
Estimate daily population	Total tourist nights, permanent resident population	Estimates the average number of people present at the destination on any given day by distributing monthly tourist nights evenly across a 30-day reference period and adding the permanent resident population. The result represents the total daily burden on local infrastructure and is used as a density input to the noise and light pollution models.	Persons per day, rounded to 2 d.p.	30-day month assumption; applicable to monthly reference periods only
Estimate tourists per 100 residents	Total tourist nights, average stay nights, permanent resident population	Approximates the number of unique tourist arrivals in the reference period by dividing total tourist nights by the average stay length, then scales this figure relative to every 100 permanent residents. The approximation assumes a homogeneous visitor population with a uniform stay equal to the mix-weighted average, which is reasonable at the destination level when individual arrival records are unavailable.	Unique tourists per 100 residents, rounded to 2 d.p. Returns 0 if stay or population is zero.	Unique arrivals approximated as tourist nights divided by average stay
Estimate safety level	Tourist pressure per 100 residents, occupancy rate, non-cultural event count	Computes a weighted risk score that combines tourist pressure (highest weight, reflecting its role as the primary crowding indicator), occupancy rate (secondary signal for strain on shared public spaces), and non-cultural event count (fixed additive load per event,	Ordinal classification: high, moderate, or low.	Weights: 0.5 (pressure), 0.2 (occupancy), 3.0 per event; thresholds: below 50 = high, 50–100

		capturing the acute crowd pressure of large gatherings). The score is mapped to a three-level safety classification.		= moderate, above 100 = low
Estimate heritage stress	Tourist nights, total attraction count (beaches and cultural sites), heritage-area accommodation count	Combines two pressure sources: a visitation density term that distributes monthly tourist nights across all attractions as a proxy for average footfall per site, and a fixed infrastructure load per accommodation establishment located within a heritage protected zone. The composite score reflects the finding that permanent tourism infrastructure generates disproportionate stress on heritage assets relative to transient visitors.	Ordinal classification: low, moderate, or high.	Heritage accommodation load: 200 units per establishment; thresholds: below 1000 = low, 1000–2999 = moderate, 3000 and above = high
Estimate noise pollution	Average daily population, total event count, destination area in km²	Sums two components: a population density term that applies a logarithmic transformation to reflect the sublinear relationship between density and perceived noise (following the urban acoustic decay model), and an event term that applies a fixed base load per event amplified by 30% to represent the noise spike associated with public gatherings. The composite score is mapped to a three-level classification.	Ordinal classification: low, moderate, or high.	Density coefficient: 20; event base load: 25; event amplifier: 1.3; thresholds: below 100 = low, 100–199 = moderate, 200 and above = high
Estimate air quality index	Adjusted CO_2 emissions per capita per day	Applies a linear scaling to the adjusted CO_2 emission rate to produce an index on a 1–6 scale, where 1 represents the best air quality and 6 the worst, consistent with international AQI band conventions. The divisor is calibrated so that emission rates typical of Mediterranean island tourism (3–15 kg/capita/day) map to the good-to-moderate range of the scale. The result is capped at 6.	Index value on a 1–6 scale, rounded to 1 d.p.	Divisor: 10,000 (calibrated for Mediterranean context); scale bounded at 1 (minimum) and 6 (maximum)
Estimate light pollution	Number of available hotels, total event	Sums three components: a logarithmic population density term (reflecting the sublinear	Ordinal classification: low,	Hotel load: 1.2 per hotel; event load:

	count, average daily population, destination area in km²	relationship between density and sky brightness), a fixed load per hotel (reflecting their continuous nocturnal illumination), and a fixed load per event (capturing the discrete spikes from stage lighting and extended operating hours). The composite score is mapped to a three-level classification.	moderate, or high.	20 per event; density coefficient: 15; thresholds: below 100 = low, 100–199 = moderate, 200 and above = high
Estimate biodiversity disturbance	Number of accommodation establishments within Natura 2000 protected areas	Classifies biodiversity disturbance based solely on the count of accommodation infrastructure within Natura 2000 zones, reflecting the evidence that permanent tourism infrastructure is the primary driver of habitat fragmentation in ecologically sensitive coastal areas, outweighing the effect of transient visitor traffic. Fixed absolute thresholds are applied without further transformation.	Ordinal classification: low, moderate, or high.	Thresholds: below 20 = low, 20–59 = moderate, 60 and above = high
Refine social inclusion	Digital accessibility percentage, economic inclusivity level (low / moderate / high)	Maps the ordinal economic inclusivity level to a numeric score and computes a weighted composite with digital accessibility, assigning greater weight to digital access on the basis that it is a more directly actionable and measurable lever for inclusive tourism at the destination management level. The composite score is mapped to a three-level social inclusion classification.	Ordinal classification: low, moderate, or high.	Digital weight: 0.6; economic weight: 0.4; economic scores: high=100, moderate=60, low=30; thresholds: below 20 = low, 20–59 = moderate, 60 and above = high
Estimate total waste (kg)	Total tourist nights	Estimates total waste generation using a flat coefficient of 2.5 kg per tourist-night, consistent with average solid waste rates for Mediterranean accommodation. This function is not used as a primary KPI input — observed recyclable and non-recyclable waste figures from the flow data take precedence wherever available. It is retained as a fallback for destinations without	Estimated total waste in kg. Informational only — not used in the KPI evaluation pipeline.	2.5 kg per tourist-night; applicable to Mediterranean accommodation contexts

		waste monitoring data and as a sanity check against reported figures.		

Predictive Advisor – Results from the rest of the pilots

Results — Province of Teramo (High Season, Infrastructure and Mix Optimization)

The Province of Teramo produces the strongest optimization outcome in the dataset, with the composite score improving from 57.0% to 72.2% — a gain of 15.1 percentage points, the largest absolute improvement across all destinations and scenarios. This result is particularly significant because it is achieved under high-season conditions, where the destination is already operating at or above capacity on several indicators.

The baseline profile is distinctive. Pillar B is the strongest across all studied destinations at 78.9%, with tourist nights, occupancy, average stay, and employment all at 100% utilisation, though revenue per capita remains severely below target at 8.8% (144 EUR against 2,200 EUR), reflecting the large resident population of a mainland provincial context. Pillar A opens at 50%, with Blue Flag coverage fully met but eco-label certification absent. Pillar C is the weakest at 47.9%, driven by a safety level of 0% — tourist volume and event pressure in high season push the risk score above the low-safety threshold — alongside poor transport accessibility (30%), low cultural event ratio (37.8%), and zero scores on heritage stress and access to services. Pillar D sits at 54.0%, with ecological footprint rated high (35.0%), light pollution high (5.0%), CO_2 above target (49.2%), and water consumption below target (47.7%).

The optimized mix follows the pattern observed across destinations, concentrating on Nature Tourism (rising to 42.5%) and Religious Tourism (to 42.0%), together capturing 84% of visitors. For Teramo, however, infrastructure variables are included in the optimization, which drives the most substantial gains. Pillar C improves from 47.9% to 74.0% — the largest pillar-level gain in the dataset — with safety, accessibility, transport, and cultural event indicators all reaching 100%. This is the clearest illustration in the dataset of the value of infrastructure optimization: indicators that are structurally immovable under mix-only search all reach their targets when infrastructure variables are included.

Pillar D improves from 54.0% to 62.9%, driven by water consumption crossing its target and CO_2 improving from 49.2% to 91.4% — the largest CO_2 gain across all destinations. Structurally persistent gaps in ecological footprint, light pollution, biodiversity disturbance, and recycling capacity remain unchanged. Pillar B shows a negligible decline of 0.3 points, as the shift toward lower-spending archetypes slightly reduces estimated revenue despite higher volume — reinforcing the recurring finding that environmentally optimal mixes carry a revenue cost.

Two aspects of the Teramo result deserve particular attention. First, it is the only destination in the dataset where the baseline safety level scores 0%; the infrastructure optimizer recovers this entirely, confirming that event management is a high-leverage policy lever in high-season continental destinations. Second, the ecological footprint remains rated high in both scenarios despite all other environmental improvements, suggesting that its drivers — land use, transport infrastructure, agricultural intensity — operate independently of tourism activity and that the indicator requires destination-specific calibration for inland provincial contexts. Table 32 presents the full results for the Province of Teramo.

Table 52-Table 54 present the results for province of Teramo.

Table 52: KPI utilization scores before and after optimization — Teramo, High season

Indicator	Sector	Baseline (%)	Optimized (%)
eco-label hotels	A	0.0%	100.0%
blue flag beaches	A	100.0%	100.0%
Sector A average	**A**	**50.0%**	**100.0%**
tourist nights/month	B	100.0%	100.0%
avg stay nights	B	100.0%	100.0%
occupancy rate	B	100.0%	100.0%
employment %	B	100.0%	100.0%
revenue/capita	B	8.8%	6.6%
revenue leakage	B	100.0%	100.0%
economic inclusivity	B	60.0%	60.0%
Sector B average	**B**	**78.9%**	**78.6%**
tourist pressure	C	100.0%	100.0%
safety level	C	0.0%	100.0%
accomm. accessibility	C	60.0%	100.0%
transport accessibility	C	30.0%	100.0%
accessible attractions	C	85.3%	100.0%
cultural event ratio	C	37.8%	100.0%
heritage stress	C	0.0%	0.0%
access to services	C	0.0%	0.0%
cultural inclusivity	C	87.5%	87.5%
social inclusion	C	60.0%	60.0%
digital accessibility	C	66.7%	66.7%
Sector C average	**C**	**47.9%**	**74.0%**
green transport means	D	66.7%	66.7%
green transport types	D	75.0%	75.0%
heritage area pressure	D	100.0%	100.0%
natura area pressure	D	100.0%	100.0%
recycling capacity	D	0.7%	0.7%
biodiversity support	D	60.0%	60.0%
CO2 emissions	D	49.2%	91.4%
water consumption	D	47.7%	100.0%
air quality index	D	100.0%	100.0%
ecological footprint	D	35.0%	35.0%
biodiversity disturbance	D	33.3%	33.3%
light pollution	D	5.0%	5.0%
noise pollution	D	30.0%	50.0%
Sector D average	**D**	**54.0%**	**62.9%**

Table 53: Optimized tourist type mix and LLM-generated policy actions — Teramo, High season

Tourist Type	Direction	Baseline (%)	Optimized (%)	Recommended Action
Nature Tourism	Increase	13.0%	42.5%	Establish a network of scenic hiking trails and observation decks offering unique vantage points for wildlife viewing.
Diving Tourism	Decrease	13.0%	2.4%	Enhance underwater visibility by installing artificial reefs and marine life habitats to increase diving experiences and attract more enthusiasts.
Adventure Tourism	Decrease	13.0%	3.2%	Develop a network of scenic hiking trails with unique viewpoints and interactive wayfinding signs to enhance the sense of adventure and exploration.
Religious Tourism	Increase	13.0%	42.0%	Develop a network of sacred sites and pilgrimage routes, highlighting unique spiritual experiences and cultural heritage.
Honeymoon Tourism	Decrease	12.0%	3.5%	Develop a romantic getaway package with local boutique hotels and restaurants, offering exclusive deals for newlyweds.
Arts Tourism	Decrease	12.0%	3.0%	Establish a vibrant street art scene by commissioning local artists to create murals in underutilised public spaces.
Medical Tourism	Decrease	12.0%	2.8%	Establish a network of on-site medical interpreters and cultural liaisons to facilitate seamless communication and care for international patients.
Family Leisure	Decrease	12.0%	0.6%	Develop a network of family-friendly bike trails and parks with

				amenities like picnic areas, playgrounds, and splash pads.

Table 54: KPI Improvement Actions

KPI	Sector	Baseline (util%)	Optimized (util%)	Recommended Action
eco-label hotels	A	0.0%	100.0%	Establish a 'Green Hotel' certification program with incentives for hotels to adopt eco-friendly practices and display a visible label upon meeting sustainability standards.
safety level	C	0.0%	100.0%	Install solar-powered streetlights along high-traffic areas to enhance visibility and reduce crime.
accomm. accessibility	C	60.0%	100.0%	Install ramps and elevators in existing accommodations to ensure equal access for all visitors.
transport accessibility	C	30.0%	100.0%	Implement a free public transportation system within the destination's city limits, connecting major attractions and accommodations.
accessible attractions	C	85.3%	100.0%	Install wheelchair-accessible ramps and elevators at all attractions within a 6-month timeframe.
cultural event ratio	C	37.8%	100.0%	Establish a Cultural Events Grant Program to incentivise local organisations and artists to host events.
CO2 emissions	D	49.2%	91.4%	Implement a congestion pricing system for private vehicles entering the city centre during peak hours.
water consumption	D	47.7%	100.0%	Install low-flow showerheads and

				faucets in all hotels and accommodations.
noise pollution	D	30.0%	50.0%	Implement a noise-restrictive zone around the city centre, limiting loud music and vehicle traffic to designated areas.

Results — Island of Pag (Medium Season, Mix Optimization only)

The Island of Pag produces a modest optimization outcome, with the composite score improving from 42.58% to 44.26% — a gain of 1.68 percentage points, among the smallest in the dataset. This reflects a structural characteristic of the destination: the majority of underperforming KPIs are governed by infrastructure variables held fixed in this scenario, leaving tourist type mix as the only available lever.

The baseline profile is weak across most pillars. Pillar B opens at 50.5%, with tourist nights and occupancy at roughly half of economic capacity, revenue per capita severely below target at 30.3% (667 EUR against 2,200 EUR), and employment at only 38.5%. Pillar A sits at 50%, with Blue Flag coverage met but eco-label certification entirely absent. Pillar C is the weakest at 36.5%, driven by a tourist pressure ratio of only 22.4% — the destination is significantly under-visited relative to its social carrying capacity — alongside zero scores on transport accessibility, cultural events, and access to services. Pillar D sits at 42.3%, with green transport entirely absent, recycling capacity critically underutilised (0.3%), light pollution severely high (8.3%), and water consumption below target.

The optimized mix concentrates on Nature Tourism (rising to 49.4%) and Adventure Tourism (to 17.3%), selected primarily for low water consumption and CO_2 footprint. This drives a sharp volume reduction: tourist nights fall from 9,283 to 2,106 per month and occupancy from 43.0% to 9.8%. This is a deliberate outcome — tourist pressure improves from 22.4% to 100% and safety level from 50% to 100%, the two most significant KPI-level gains in Pillar C, both driven by the volume reduction. Pillar C consequently improves from 36.5% to 48.1%. However, the indicators that remain at zero — transport accessibility, cultural events, accessible attractions, and access to services — are structurally immovable without infrastructure investment.

Pillar D improves modestly from 42.3% to 44.2%, with water consumption gaining from 44.3% to 63.7% and CO_2 from 35.9% to 40.0%. Structural gaps in green transport, recycling, light pollution, and biodiversity disturbance remain unchanged.

Pillar B declines from 50.5% to 36.8%, driven by the sharp fall in tourist nights, occupancy, and revenue per capita (from 30.3% to 5.8%). This is the most significant economic cost of the optimization, and it is particularly acute for Pag because the baseline is already economically underperforming — the optimization compounds an existing weakness rather than trimming excess from a position of strength.

Two aspects of this result deserve particular attention. First, the no-infrastructure constraint is especially limiting here: the majority of the worst-performing indicators require capital investment to move, and the gap between the achievable optimized score and a plausible infrastructure-inclusive ceiling is likely large. Second, a destination already at 47.8% occupancy falling further to 9.8% risks accommodation business failures and employment losses that would undermine the economic resilience required to sustain any long-term sustainability programme. The score improvement should therefore be interpreted with caution, and any implementation pathway would need to phase the mix shift alongside demand-side stimulation within the targeted

archetypes. Table 49 presents the full results for the Island of Pag. Table 55-Table 57 present the results for province of Island of Pag.

Table 55: KPI utilization scores before and after optimization — Island of Pag , Medium season

Indicator	Sector	Baseline (%)	Optimized (%)
eco-label hotels	A	0.0%	0.0%
blue flag beaches	A	100.0%	100.0%
Sector A average	**A**	**50.0%**	**50.0%**
tourist nights/month	B	47.8%	10.8%
avg stay nights	B	92.5%	95.2%
occupancy rate	B	47.8%	10.9%
employment %	B	38.5%	38.5%
revenue/capita	B	30.3%	5.8%
revenue leakage	B	66.7%	66.7%
economic inclusivity	B	30.0%	30.0%
Sector B average	**B**	**50.5%**	**36.8%**
tourist pressure	C	22.4%	100.0%
safety level	C	50.0%	100.0%
accomm. accessibility	C	10.0%	10.0%
transport accessibility	C	0.0%	0.0%
accessible attractions	C	40.0%	40.0%
cultural event ratio	C	0.0%	0.0%
heritage stress	C	100.0%	100.0%
access to services	C	0.0%	0.0%
cultural inclusivity	C	56.2%	56.2%
social inclusion	C	60.0%	60.0%
digital accessibility	C	62.5%	62.5%
Sector C average	**C**	**36.5%**	**48.1%**
green transport means	D	0.0%	0.0%
green transport types	D	0.0%	0.0%
heritage area pressure	D	100.0%	100.0%
natura area pressure	D	100.0%	100.0%
recycling capacity	D	0.3%	0.3%
biodiversity support	D	20.0%	20.0%
CO2 emissions	D	35.9%	40.0%
water consumption	D	44.3%	63.7%
air quality index	D	100.0%	100.0%
ecological footprint	D	58.3%	58.3%
biodiversity disturbance	D	33.3%	33.3%
light pollution	D	8.3%	8.3%
noise pollution	D	50.0%	50.0%

Sector D average	D	42.3%	44.2%

Table 56: Optimized tourist type mix and LLM-generated policy actions — Island of Pag , Medium season

Tourist Type	Direction	Baseline (%)	Optimized (%)	Recommended Action
Nature Tourism	Increase	13.0%	49.4%	Establish a network of scenic hiking trails and observation decks offering unique vantage points for wildlife viewing.
Diving Tourism	Decrease	13.0%	6.4%	Enhance underwater visibility by installing artificial reefs and marine life habitats to increase diving experiences and attract more enthusiasts.
Adventure Tourism	Increase	13.0%	17.3%	Develop a network of scenic hiking trails with unique viewpoints and interactive wayfinding signs to enhance the sense of adventure.
Religious Tourism	Decrease	13.0%	2.1%	Develop a sacred site renovation project to enhance pilgrims' experiences and improve infrastructure, including accessible pathways and amenities.
Honeymoon Tourism	Decrease	12.0%	0.7%	Develop a romantic getaway package with local boutique hotels and restaurants, offering exclusive experiences like private beach dinners and sunset helicopter rides.
Arts Tourism	Decrease	12.0%	11.6%	Enhance the city's street art scene by commissioning a large-scale mural project in a high-visibility area.
Medical Tourism	Decrease	12.0%	3.2%	Establish a network of certified medical interpreters to facilitate communication between international patients and local healthcare providers.

Family Leisure	Decrease	12.0%	9.4%	Develop a family-friendly bike trail network with accessible amenities and scenic views to encourage more leisurely activities.

Table 57: ***KPI Improvement Actions***

KPI	Sector	Baseline	Optimized	Δ	Recommended action
tourist nights/month	B	47.8%	10.8%	−37.0	Implement a "Shoulder Season" accommodation discount program for visitors arriving between April and June.
avg stay nights	B	92.5%	95.2%	+2.7	Implement a "Stay & Dine" package discount for Category B visitors, offering 10% off at local restaurants and cafes with a minimum 3-night stay.
occupancy rate	B	47.8%	10.9%	−36.9	Implement a "Green Week" initiative, offering discounted accommodations and activities for visitors who arrive by public transportation or carpool.
revenue/capita	B	30.3%	5.8%	−24.5	Implement a "Slow Tourism" certification program for local accommodations and tour operators, promoting low-impact, high-value experiences.
Sector B average	B	50.5%	36.8%	−13.7	Implement a "Clean Beaches" certification program for local businesses, offering discounts and promotions to those that maintain high cleanliness standards.
tourist pressure	C	22.4%	100.0%	+77.6	Establish a "Tourist-Friendly" certification program for local accommodations, offering incentives for eco-

					friendly practices and community engagement.
safety level	C	50.0%	100.0%	+50.0	Install additional streetlights and pedestrian-friendly lighting in high-risk areas to enhance visibility and reduce crime.
Sector C average	C	36.5%	48.1%	+11.6	Implement a "Green Certification" program for local accommodations, offering discounts and promotions to guests who stay at eco-friendly hotels.
CO2 emissions	D	35.9%	40.0%	+4.1	Implement a "Green Taxi" fleet with electric and hybrid vehicles, replacing 20% of existing taxis within the first year.
water consumption	D	44.3%	63.7%	+19.4	Install low-flow showerheads and faucets in all hotels and accommodations.
Sector D average	D	42.3%	44.2%	+1.9	Implement a "Green Certification" program for local accommodations, offering discounts and promotions to guests who stay at eco-friendly hotels.

Results — Nessebar (High Season, Infrastructure and Mix Optimization)

Nessebar produces a strong optimization outcome, with the composite score improving from 62.78% to 70.85% — a gain of 8.07 percentage points, achieved under high-season conditions with infrastructure optimization enabled.

The baseline profile is economically the strongest in the dataset. Pillar B sits at 91.67%, with tourist nights, occupancy, employment, revenue per capita, and economic inclusivity all at 100% utilisation. Revenue leakage is the only financial gap, with 60% of tourism revenue leaving the local economy against a 30% target. Pillar A opens at 51.3%, with Blue Flag coverage met but eco-label certification critically absent (10 of 390 hotels certified). Pillar C sits at 60.45%, with significant internal divergence: tourist pressure, transport accessibility, and digital and cultural inclusivity all reach 100%, while safety level scores 0% — high tourist volume in peak season pushes the risk score above the low-safety threshold — and heritage stress and access to services remain at 0%. Pillar D sits at 50.96%, with water consumption severely below target at 34.81% utilisation (430.9 L/day against a 150 L/day target), CO_2 above target at 67.18%, and structurally high light pollution, biodiversity disturbance, and ecological footprint.

The optimized mix departs from the pattern observed in lower-flow destinations. Nature Tourism is introduced from zero to 37.4% and Adventure Tourism increases from 13.0% to 28.8%, together

accounting for 66.1% of the visitor population. The dominant logic is water and CO_2 reduction: both archetypes consume well below the 360–600 L/night range of the high-weight baseline segments.

The most significant gains are in Pillar D. Water consumption improves from 34.81% to 100% utilisation — the largest single-KPI improvement in this scenario — as the mix shift reduces the weighted average to 149.4 L/day, just below the 150 L/day threshold. CO_2 improves from 67.18% to 75.43%. All other environmental indicators remain unchanged, as they are fixed flow values not adjusted by the optimizer. Pillar A reaches 100%, driven by the infrastructure optimizer setting eco-certified hotels to 1,298 — well above the target of 390, exploiting the absence of an upper bound in the search space.

Pillar C improves from 60.45% to 69.09%, with accommodation accessibility, accessible attractions, and cultural event ratio all reaching 100% through infrastructure optimization. Safety level and access to services remain at 0% in both scenarios, indicating that neither the mix shift nor the available infrastructure variables are sufficient to address these gaps. Safety level in particular persists at 0% despite the volume reduction implicit in the mix shift, suggesting that the risk score at this scale is governed by factors beyond tourist density alone.

Pillar B shows a negligible change of +0.04 points, with a modest decline in revenue per capita (from 3,479 to 2,159 EUR) reflecting the shift toward lower-spending archetypes — reinforcing the consistent cross-destination finding that environmentally optimal mixes carry a revenue penalty regardless of economic starting position.

Two aspects of the Nessebar result deserve particular attention. First, the safety level indicator is the only major KPI that the combined mix and infrastructure optimization fails entirely to address, suggesting that event management and crowd control levers within the current framework are insufficient for a destination of this scale. Second, the water consumption improvement is the largest single-KPI gain in the dataset and illustrates the high leverage of mix shifts in high-consumption coastal destinations: closing a gap of over 280 L/day through demand composition alone is a particularly policy-relevant finding where supply-side interventions are costly or slow to implement. Table 58-Table 62 present the results for Nessebar.

Table 58: KPI utilization scores before and after optimization — Nessebar, High season

Indicator	Sector	Baseline (%)	Optimized (%)
eco-label hotels	A	2.6%	100.0%
blue flag beaches	A	100.0%	100.0%
Sector A average	A	51.3%	100.0%
tourist nights/month	B	100.0%	100.0%
avg stay nights	B	91.7%	93.9%
occupancy rate	B	100.0%	100.0%
employment %	B	100.0%	100.0%
revenue/capita	B	100.0%	98.1%
revenue leakage	B	50.0%	50.0%
economic inclusivity	B	100.0%	100.0%
Sector B average	B	91.7%	91.7%
tourist pressure	C	100.0%	100.0%
safety level	C	0.0%	0.0%
accomm. accessibility	C	60.0%	100.0%

transport accessibility	C	100.0%	100.0%
accessible attractions	C	78.3%	100.0%
cultural event ratio	C	66.7%	100.0%
heritage stress	C	0.0%	0.0%
access to services	C	0.0%	0.0%
cultural inclusivity	C	100.0%	100.0%
social inclusion	C	60.0%	60.0%
digital accessibility	C	100.0%	100.0%
Sector C average	C	60.5%	69.1%
green transport means	D	66.7%	66.7%
green transport types	D	75.0%	75.0%
heritage area pressure	D	87.8%	87.8%
natura area pressure	D	100.0%	100.0%
recycling capacity	D	0.3%	0.3%
biodiversity support	D	44.0%	44.0%
CO2 emissions	D	67.2%	75.4%
water consumption	D	34.8%	100.0%
air quality index	D	100.0%	100.0%
ecological footprint	D	35.0%	35.0%
biodiversity disturbance	D	16.7%	16.7%
light pollution	D	5.0%	5.0%
noise pollution	D	30.0%	30.0%
Sector D average	D	51.0%	56.6%

Table 59: Optimized tourist type mix and LLM-generated policy actions — Nessebar, High season

Tourist Type	Direction	Baseline (%)	Optimized (%)	Recommended Action
Nature Tourism	Increase	0.0%	37.4%	Establish a network of scenic hiking trails and observation decks offering breathtaking views of the surrounding wilderness.
Diving Tourism	Decrease	13.0%	2.4%	Enhance underwater visibility by installing artificial reefs and marine life habitats to increase diving experiences and attract more enthusiasts.
Adventure Tourism	Increase	13.0%	28.8%	Develop a network of scenic hiking trails and bike paths connecting local natural attractions.

Religious Tourism	Decrease	12.0%	2.0%	Develop a sacred site renovation project to enhance pilgrims' spiritual experiences, incorporating modern amenities while preserving historical integrity.
Honeymoon Tourism	Decrease	20.0%	2.8%	Develop a romantic getaway package with local boutique hotels and restaurants, offering exclusive experiences like private beach dinners and sunset helicopter rides.
Arts Tourism	Decrease	12.0%	1.3%	Establish a vibrant street art festival in the city's historic district, featuring local and international artists.
Medical Tourism	Decrease	20.0%	9.5%	Establish a network of certified medical interpreters to facilitate communication between international patients and local healthcare providers.
Family Leisure	Increase	10.0%	15.8%	Develop a network of family-friendly bike trails and rental stations in scenic areas to encourage active exploration.

Table 60: KPI Improvement Actions

KPI	Sector	Baseline	Optimized	Δ	Recommended action
eco-label hotels	A	2.6%	100.0%	+97.4	Establish a "Green Hotel" certification program with clear standards and incentives for hotels to adopt eco-friendly practices.
Sector A average	A	51.3%	100.0%	+48.7	Implement a "Green Certification" program for local accommodations, offering discounts and promotions to guests who stay at eco-friendly hotels.
avg stay nights	B	91.7%	93.9%	+2.2	Implement a "Stay & Dine" package discount for visitors who book

					accommodations and restaurants in the same neighbourhood.
revenue/capita	B	100.0%	98.1%	−1.9	Implement a "Sustainable Accommodation" certification program for local hotels and B&Bs, offering discounts to guests who book eco-friendly rooms.
Sector B average	B	91.7%	91.7%	+0.04	Implement a "Green Certification" program for local accommodations, offering discounts and promotions to guests who stay at eco-friendly hotels.
accomm. accessibility	C	60.0%	100.0%	+40.0	Install ramps and elevators in existing Category C accommodations to ensure equal access for all visitors.
accessible attractions	C	78.3%	100.0%	+21.7	Install wheelchair-accessible ramps and elevators at all Category C attractions within a 2-year timeframe.
cultural event ratio	C	66.7%	100.0%	+33.3	Establish a Cultural Events Grant Program to support local organizations and artists in hosting events, with priority given to Category C initiatives.
Sector C average	C	60.5%	69.1%	+8.6	Implement a "Green Parking" initiative, offering discounted rates for electric and hybrid vehicles at prime locations.
CO2 emissions	D	67.2%	75.4%	+8.2	Implement a bike-sharing program with electric-assist bicycles and incentivize visitors to use them for short trips.

water consumption	D	34.8%	100.0%	+65.2	Install low-flow showerheads and faucets in all hotels and accommodations.
Sector D average	D	51.0%	56.6%	+5.6	Implement a "Green Hotel" certification program, offering incentives for hotels to reduce energy consumption and waste management.

Results — Danube Delta (Low Season, Mix Optimization)

The Danube Delta produces a moderate optimization outcome, with the composite score improving from 47.36% to 54.28% — a gain of 6.92 percentage points, achieved under low-season conditions without infrastructure optimization. The destination has the most underdeveloped baseline in the dataset, reflecting near-total absence of destination management certification, severely underdeveloped economic activity, and a set of social and environmental indicators that are structurally immovable under mix-only optimization.

Pillar A scores 0% in both scenarios — neither eco-label certification nor Blue Flag coverage is met — making this the only destination where Pillar A contributes nothing to the composite score. Pillar B opens at 44.0%, with tourist nights and occupancy severely below target (16.7% and 16.8% respectively) and revenue per capita critically low at 1.9% utilisation (16.79 EUR against a 900 EUR target). Revenue leakage is the strongest financial indicator at 92.3%, a structural feature of a remote destination with limited import dependency. Pillar C opens at 47.73%, with tourist pressure and safety both at 100% — the destination is well within its social carrying capacity — but accommodation accessibility, accessible attractions, and access to services all at 0%, indicating a destination that lacks the service infrastructure to accommodate diverse visitor needs. Pillar D is the strongest baseline pillar at 56.17%, with green transport types, protected area pressure, air quality, ecological footprint, and noise pollution all at or near their targets, though recycling capacity and biodiversity support are entirely absent and CO_2 and water consumption remain below target.

The optimized mix is the most distinctive in the dataset. Religious Tourism rises to 48.4% and Nature Tourism is introduced from zero to 35.6%, together accounting for 84% of the visitor population. Both are low-water, low-CO_2 archetypes well suited to a low-capacity, ecologically sensitive destination.

Pillar B produces the largest gain in this scenario, improving from 44.0% to 68.5% — a 24.5-point increase driven by tourist nights and occupancy both reaching 100% utilisation as the mix shift toward longer-stay archetypes expands total volume. Revenue per capita improves from 1.9% to 7.3% but remains severely below target at 65.71 EUR, a consequence of the low per-night expenditure of the dominant archetypes and the small local population base.

Pillar C shows no change in either scenario, remaining at 47.73% — the clearest illustration in the dataset of the hard ceiling imposed by the no-infrastructure constraint. Every underperforming indicator in Pillar C is governed by physical infrastructure that tourist mix management cannot adjust.

Pillar D improves modestly from 56.17% to 60.50%, with water consumption improving from 46.6% to 90.5% and CO_2 from 33.6% to 46.1%. All other environmental indicators remain unchanged.

Three aspects of this result deserve particular attention. First, the absence of any Pillar A contribution suppresses the composite score ceiling significantly, and the gap between the achievable mix-only score and a plausible infrastructure-inclusive ceiling is likely the largest in the dataset. Second, the dominance of Religious Tourism at 48.4% reflects the specific combination of low water and CO_2 targets, short average stay, and the high volume needed to fill tourist nights from a near-zero baseline — shorter-stay archetypes contribute more arrivals per night, closing the occupancy gap faster. Third, revenue per capita remains critically low even in the optimized scenario, and a mix dominated by the two lowest-spending archetypes in the dataset means that any pathway toward economic viability will require either infrastructure investment to attract higher-spending segments or a fundamentally different approach to local value capture. Table presents the full results for the Danube Delta. Table 61, Table 63 present the results for Danube Delta.

Table 61: KPI utilization scores before and after optimization — Danube Delta, Low season

Indicator	Sector	Baseline (%)	Optimized (%)
eco-label hotels	A	0.0%	0.0%
blue flag beaches	A	0.0%	0.0%
Sector A average	**A**	**0.0%**	**0.0%**
tourist nights/month	B	16.7%	100.0%
avg stay nights	B	100.0%	100.0%
occupancy rate	B	16.8%	100.0%
employment %	B	50.0%	50.0%
revenue/capita	B	1.9%	7.3%
revenue leakage	B	92.3%	92.3%
economic inclusivity	B	30.0%	30.0%
Sector B average	**B**	**44.0%**	**68.5%**
tourist pressure	C	100.0%	100.0%
safety level	C	100.0%	100.0%
accomm. accessibility	C	0.0%	0.0%
transport accessibility	C	20.0%	20.0%
accessible attractions	C	0.0%	0.0%
cultural event ratio	C	100.0%	100.0%
heritage stress	C	100.0%	100.0%
access to services	C	0.0%	0.0%
cultural inclusivity	C	25.0%	25.0%
social inclusion	C	30.0%	30.0%
digital accessibility	C	50.0%	50.0%
Sector C average	**C**	**47.7%**	**47.7%**
green transport means	D	0.0%	0.0%
green transport types	D	100.0%	100.0%
heritage area pressure	D	100.0%	100.0%
natura area pressure	D	100.0%	100.0%
recycling capacity	D	0.0%	0.0%

biodiversity support	D	0.0%	0.0%
CO2 emissions	D	33.6%	46.1%
water consumption	D	46.6%	90.5%
air quality index	D	100.0%	100.0%
ecological footprint	D	100.0%	100.0%
biodiversity disturbance	D	33.3%	33.3%
light pollution	D	16.7%	16.7%
noise pollution	D	100.0%	100.0%
Sector D average	**D**	**56.2%**	**60.5%**

Table 62: Optimized tourist type mix and LLM-generated policy actions — Danube Delta, Low season

Tourist Type	Direction	Baseline (%)	Optimized (%)	Recommended Action
Diving Tourism*	Decrease	13.0%	4.8%	Promote diving tourism opportunities available on the Black Sea Coast (Sulina and Sfântu Gheorghe), as part of the wider Danube Delta pilot region. Enhance underwater visibility by installing artificial reefs and marine life habitats to increase diving experiences and attract more enthusiasts.
Adventure Tourism	Decrease	13.0%	3.7%	Develop a network of scenic hiking trails and bike paths connecting local natural attractions to boost outdoor enthusiasts' experiences.
Religious Tourism	Increase	12.0%	48.4%	Develop a network of sacred sites and pilgrimage routes, highlighting unique spiritual experiences and cultural heritage.
Honeymoon Tourism	Decrease	20.0%	1.5%	Develop a romantic getaway package with local boutique hotels and restaurants, offering exclusive experiences like private beach dinners and sunset helicopter rides.
Arts Tourism	Decrease	12.0%	0.5%	Establish a vibrant street art festival in the city's historic

				district, featuring local and international artists.
Medical Tourism	Decrease	20.0%	4.1%	Establish a network of on-site medical interpreters and cultural liaisons to facilitate seamless communication and care for international patients.
Family Leisure	Decrease	10.0%	1.5%	Develop a network of family-friendly bike trails and parks with amenities like picnic areas, playgrounds, and splash pads.
Diving Tourism*	Decrease	13.0%	4.8%	Enhance underwater visibility by installing artificial reefs and marine life habitats to increase diving experiences and attract more enthusiasts.

*Diving tourism refers to the Black Sea Coast sites of Sulina and Sfântu Gheorghe, not to the Danube Delta waterways. The niche may also include underwater photography focused on natural ecosystems[8].

Table 63*: KPI Improvement Actions*

KPI	Sector	Baseline	Optimized	Δ	Recommended action
occupancy rate	B	16.8%	100.0%	+83.2	Implement a "Free Nightly Parking" program for Category B accommodations during off-peak seasons.
revenue/capita	B	1.9%	7.3%	+5.4	Develop a "Foodie Trail" with 10 local restaurants and cafes offering unique, sustainable dining experiences, promoting high-quality, locally sourced cuisine.
Sector B average	B	44.0%	68.5%	+24.5	Implement a "Green Hotel" certification program, offering incentives for hotels to reduce energy consumption and waste management.

[8] https://rewildingeurope.com/blog/a-week-full-of-surprises-in-the-danube-river-delta/

CO2 emissions	D	33.6%	46.1%	+12.5	Implement a "Green Taxi" fleet with electric and hybrid vehicles, incentivizing drivers to switch from gas-powered cars.
water consumption	D	46.6%	90.5%	+43.9	Install low-flow showerheads and faucets in all hotels and accommodations.
Sector D average	D	56.2%	60.5%	+4.3	Implement a "Green Hotel" certification program, offering incentives for hotels to reduce energy consumption and waste management.
occupancy rate	B	16.8%	100.0%	+83.2	Implement a "Free Nightly Parking" program for Category B accommodations during off-peak seasons.
revenue/capita	B	1.9%	7.3%	+5.4	Develop a "Foodie Trail" with 10 local restaurants and cafes offering unique, sustainable dining experiences, promoting high-quality, locally sourced cuisine.
Sector B average	B	44.0%	68.5%	+24.5	Implement a "Green Hotel" certification program, offering incentives for hotels to reduce energy consumption and waste management.
CO2 emissions	D	33.6%	46.1%	+12.5	Implement a "Green Taxi" fleet with electric and hybrid vehicles, incentivizing drivers to switch from gas-powered cars.
water consumption	D	46.6%	90.5%	+43.9	Install low-flow showerheads and faucets

					in all hotels and accommodations.
Sector D average	D	56.2%	60.5%	+4.3	Implement a "Green Hotel" certification program, offering incentives for hotels to reduce energy consumption and waste management.