Tourism Entropy as a Critical Factor for Destination Sustainability: Assessment of Carrying Capacity Through IoTs and Determination of Tourism Satisfaction

Antonis Theocharous
Cyprus University of Technology, Limassol, Cyprus
Spyridon Avdimiotis
International Hellenic University, Thessaloniki, Greece
Apostolos Karakitsios
Cyprus University of Technology, Limassol, Cyprus

This paper introduces a novel methodology to address the issue of destination carrying capacity assessment and monitoring, while, to develop a new tool to enable effective and optimized policy making in tourism sector. Tourism fits together a chain value, where stakeholders are working as cogwheels targeting the sustainability. An inherent requirement towards the objectives of sustainability is the capacity of policy makers to sustain the available social, economic, and environmental resources. Given the facts that (i) carrying capacity is a complex, dynamic, multidimensional concept, apparently vague, as such it is quite difficult to assess and monitor and (ii) each destination has unique characteristics, varying the extent and intensity of tourism development, the research proposal refines and implements measures and establishes a state-of-the-art data collection network in order to monitor tourism development in terms of sustainability and to establish an integrated carrying capacity estimation framework. Since validity and reliability of input data is a critical issue, different kind of sensors for environmental data (IoTs) are used, as well as web mining methods and field surveys, statistical information and secondary research results. Also, specialized analysis methods are applied and integration of all sources is achieved, defining the characteristics of the tourism destination ecosystem from the perspective of entropy change, in order to provide planning, policy making, and decision support, as well as a set of intelligent tools.

Keywords: carrying capacity, smart tourism, tourism entropy, sustainability, tourism satisfaction, environment protection, policy making, tourism destination, innovation, IoTs

Introduction

In worldwide scale, tourism is acknowledged as a leading dynamic industry, significantly contributing to
economic prosperity, reducing unemployment, strengthening social cohesion, and providing a wide range of benefits to organizations and societies. This is why, at a time when the world is going through the effects of financial recession, social disruption, and institutional crisis, tourism is leading the way out of the doldrums. However, tourism development is most commonly a combination of critical factors that need to be handled with care, in order to avoid disproportional negative environmental, economic, and social cohesion impacts (Avdimiotis, Theocharous, Papadopoulos, & Zopiatis, 2015; Statzu & Strazzera, 2011).

Towards this direction, the research proposal seeks to provide monitoring, policy planning, and implementation tools for sustainable tourism development. In thick lines, the research proposal offers to public authorities, destination management organizations, and tourism development agencies an intelligent decision support infrastructure, to control environmental carrying capacity, comprising wide data collection networks, advanced data analysis, and knowledge engineering technologies.

A critical issue for the development of the Intelligent Tourism Policy Decision Support System is the sufficiency and reliability of input data, which are derived from the measurement of specific carrying capacity, indicators. The research proposal refines measures and establishes a state-of-the-art data collection network to monitor tourism development and set an integrated carrying capacity control framework.

The extent of environmental, social, and economic resilience is linked to the term of carrying capacity which is defined as “the maximum number of persons which could visit a location within a given period, such that local environmental, physical, economic, and socio-cultural characteristics are not compromised, and without reducing tourist satisfaction” (WTO, 1999). Nevertheless, according to existing bibliography, the concept of carrying capacity assessment holds distinctive flaws some of which—in thick lines—are: (1) the complexity of destinations including objective and subjective variables; (2) the extent of destination policy-makers capability to apply measures of action towards sustainability (for instance, the excess of tourist number that can visit the destination without causing permanent environmental damage should entail the possibility to suggest measures and reduce the number of visitors); (3) the wide extent and variation of tourist activities and behavior within the destination; (4) the unique carrying capacity state of each destination, which is not merely depended by the availability of natural and physical resources, but also by the characteristics of the management system, by the type of tourism that characterizes the area, by stakeholders’ perceptions and other local conditions (Guo & Chung, 2016; Ioannides & Billing, 2005).

A significant innovation of the proposal is the development of a state-of-the-art methodology, to address the issue of carrying capacity assessment, by employing IoT techniques (Hou, Zhao, Xiong, Zheng, Chatzimisios, Hossain, & Chen, 2016) in order to select and process environmental data and an Intelligent Decision Support System to monitor and eventually facilitate (public and private) tourism policy bodies. The assessment methodology of the proposed method primarily includes the customization of indicators used, to the attributes of each destination. Therefore, the challenge faced in the current concept seeks to develop decision support systems for tourism destination able to: (1) consider the priorities of the destination under investigation; (2) identify local constraints to tourism development, balancing the demand of new tourist infrastructures and the necessity to protect local environment; (3) select the appropriate set of indicators that can be useful to all tourism policy-makers and stakeholders; (4) define scenarios for the development of the destination.

A critical issue is the sufficiency and reliability of input data, deriving from the measurement of specific carrying capacity environmental indicators.
One of the main aims in developing this methodology is to attempt to define a range of acceptable threshold values for each indicator which compose the framework. Indeed, even if the definition of thresholds is necessarily a choice that implies a certain level of subjectivity, the evaluation of indicators becomes meaningless in absence of reference targets and standards, which are included in a commonly accepted range of (benchmark) values. To address the issue of value range determination, it is required to identify current economic, social, and environmental situation; identify the pressure generated stemming from tourism activities; identify the appropriate set of indicators, customized to the needs and perspectives of the destination; acknowledge benchmark values the activities which are more relevant in the local situation; acknowledge value standards determined by international, European, and national laws.

The innovation of the research proposal system is that a smart integrated system is introduced for the first time, able to collect real time data through a network of sensors, to extract patterns and interpret environmental behavior through data mining methods, and to drive an intelligent decision support system that relates measurements with development policies. In this system input data are carefully selected through a set of carrying capacity indicators adjusted to the preferences, type, and intensity of destination’s tourism product. The system is based on the measurement of carrying capacity; the value of indicators used is customized in acceptable levels, related to environmental indicators, meeting the destination’s tolerance standards. Additionally, knowledge-based models are developed and incorporated in the DSS.

Methodology, Data Sources and Integration

The research proposal methodology includes interdisciplinary work to match effective data collection instruments for each of the three pillars of sustainable development with expertise in the field of sustainable tourism development, information and knowledge modeling methods, as well as intelligent decision support technologies. Public policy making organizations, who are the end users of the developed technologies, are involved in the project from its onset, in order to direct requirements analysis and specific targets. Policy makers in destinations are also involved in the development of the tourism policy model following methodological recommendations (Farsari, Butler, & Szivas, 2011). Moreover, they contribute in the evaluation and feedback collection phase and act as the first adopters of the results. In the research proposal, the technological developments and the data-driven methods are complemented with the necessary content and activated in pilot operation to solve representative real-world problems.

The development of the technological platform is organized in a multi-layer approach, consisting of the following five layers: (1) data layer, (2) intelligence layer, (3) knowledge engineering, (4) inference and decision making, and (5) application layer.
Data Layer

The data layer is concerned with the collection of input data and includes the development of suitable instruments, the establishment of the collection networks, the specification of data sources and data models, as well as the implementation of data management infrastructures (databases). The data layer is designed and implemented giving emphasis to openness, expandability, and scalability.

**Social and cultural data.** Three main sources of social data are foreseen, each one calling for its own specialized instruments. (a) Secondary data and theory from the field of social sciences are coded in order to set the socio-cultural background. Existing indicators and methodologies of assessing the social impact of tourism are reconsidered and further developed. Research results capturing societal problems and their causes, social group profiles and relations between groups and positioning against tourism, factors affecting positive/negative impact, etc., are collected. (b) Primary data collected through social surveys. The instruments are questionnaire-based surveys and are developed tailored to the needs of the research proposal. The surveys are addressed to local inhabitants (capturing multi-item concepts on their current image, response to tourism development and sensitivity) and to tourists (capturing their image on how local society accepts them and possible positive/negative signs on tourism saturation). (c) Social media big data web mining techniques are employed to capture real-time content streams from social media platforms, news media channels, and stakeholder social web sites. The aim is to automatically capture emerging stories, identify opinion leaders, and determine the sentiment expressed towards specific issues by tourists and local inhabitants. Special emphasis is
placed on the social perceptions of sustainable touristic development (environmental conditions and load in resource usage, conflicts between local people and tourists, the economic and cultural impact of tourism to local societies, trends/complaints/wishes of tourists, etc.). The impact of individual sources is measured to guide decisions on strategic and operative levels (policy making, citizen engagement and social innovation, communication strategies, campaign success metrics, etc.). The specific technology platform provides all the necessary infrastructure, including the one for the acquisition, filtering, pre-processing, automated enrichment, storage and indexing of the collected input data.

**Economic data.** Mature and available indicators/measures are employed to collect data on local economy characteristics (e.g., employment, development per sector, resources, etc.). Data sources are existing networks, statistical agencies, and other entities such as chambers of commerce and industry and others. The input data of this category are usually structured in diverse ways depending on the provider, thus management is required by developing the appropriate data model and database.

**Environmental data.** The existing background on environmental indicators are employed and further developed, to define sets of measurable data. Individual characteristics and needs per area/country are considered, in order to adapt a suitable framework for data collection. Novel measurement technologies are used as instruments.

**Collection of data from sensors.** Different approaches to collecting and storing these measurements are used because of the different sensor types.

The following figure shows an approach to the method is implemented:

![Data collection methodology](image)

*Figure 2. Data collection methodology.*

The data collection from the PA-II air quality assemblies’ sensors is carried out in the following steps: (1) Particulates are measured per second. (2) The estimated total mass for PM 1.0, PM 2.5, and PM 10 is averaged by the cloud control panel. (3) The measurements are then uploaded to the cloud infrastructure every 80 seconds, where they are saved for download and view on the interactive map. (4) Then, using the APIs stores all the data on the main server. These sensors use particle laser meters that provide an affordable and accurate
way of measuring smoke, dust, and other particulate pollutants. They have a fan to draw air in front of the laser, causing reflections of any particles in the air. These reflections are used to measure particles in six sizes between $0.3 \, \mu m$ and $10 \, \mu m$. More specifically, this sensor uses the principle of laser scattering, producing scattered light by irradiating the suspended particles in the air. Then, the light scattering is collected to a certain extent and the light scattering curve is eventually plotted over time. Finally, the equivalent particulate diameter and the number of particulates with a different diameter per unit volume can be calculated by a microprocessor based on the MIE scattering theory.

SCP WiFi Smart sensors’ operation is a similar, but without intermediate cloud platform involvement. The sensors are programmed to send their data to main server. More specifically, (1) sensors are programmed to measure particulates per second, (2) the measurements are then uploaded to the main server every 80 seconds approximately, where the estimated total mass for PM 1.0, PM 2.5, and PM 10 averaged and are stored in the database (real-time monitoring of acoustic levels on the streets), air quality, waste management, lighting, etc. The selected sensors are used for real-time monitoring on a destination measuring (1) noise levels, (2) NO$_2$ levels, (3) temperature, humidity and atmospheric pressure, (4) O$_3$ levels.

**Intelligence Layer**

Data collected in the data layer are transformed into structured information, usable in estimating the carrying capacity of touristic destinations. In other words, measurements by environmental sensors, economic parameters, survey data, and web content are transformed into sets of standardized and localized indicators, reference values, rules and qualitative information which can be applied to estimate integrated carrying capacity and thus policy planning and decision making. A solid framework is developed for standardizing and modeling the extracted information and its applicability.

**Social and cultural information.** Three specialized frameworks are applied to different categories of input data: multidimensional factor analysis on questionnaire-based social surveys, web mining/sentiment analysis methods to handle social media sources, and statistical methods (e.g., Bayesian networks and others) to analyze quantitative data. Separate subsets of indicators and corresponding information models are developed for each category, because of the totally different nature of the corresponding domains. Data analysis from the family of multidimensional statistical analysis is employed, in order to extract measurements from complex multivariate data without relying neither on predefined models nor on quantification scales, but making use of the ability offered by these methods to manage large numbers of qualitative variables and to automatically define data-driven scales. Factor analysis method (i.e., multiple correspondence analysis) is used to reveal factors, data structure and to define qualitative scales, while its combination with special clustering methods (i.e., CHA) is used to evaluate behavior, trends, and estimations of complex qualitative indicators (Stalidis, 2013).

Further increasing the accuracy of knowledge extraction algorithms proves increasingly difficult, even when developing advanced hybrid systems that use dependency parsing and combine lexical with machine learning approaches. Linked data resources and semantic knowledge bases that include common and common-sense knowledge are a promising research avenue to increase precision and recall, but often suffer from data quality issues and do not provide the throughput required for real-time web intelligence applications. A fragmentation of the language resources and metadata elements extracted in different projects further complicates the situation.
Addressing these shortcomings and pursuing an integrated approach to knowledge extraction and visualization, the research proposal (i) collects real-time content feeds from news and social media platforms, (ii) extends the existing knowledge extraction portfolio with the ability to process Green content, including factual metadata such as named entities and affective metadata such as sentiment, (iii) aligns the collected unstructured content with the social-cultural, economic, and environmental data, (iv) develops a novel visualization tool that shows evolving opinions in the context of an emerging story—to be integrated into the research proposal dashboard.

**Economic information.** The basis for the economic information framework is the existing work of international, European, and national organizations on standardizing data and issuing guidelines regarding economic indicators for sustainable development. In particular, we mention (a) the Measuring Sustainable Tourism (MST) initiative, operated by UNWTO, through which a statistical framework for sustainable tourism is built and (b) the Global Sustainable Tourism Council Criteria (GSTC), developed in accordance to the formed and widely accepted principles and guidelines, certification criteria and indicators, reflecting sustainable tourism certification standards, indicators, criteria, and best practices from different cultural and geo-political contexts from around the world.

In order to deal with the diversity of existing indicator frameworks, the research proposal contributes by introducing an elaborate information modeling layer on top of data management, in which high level indicators are defined together with their localization and operationalization framework. Furthermore, the research proposal plans to introduce an information extraction process, suitable for exploring data, to discover patterns and relations that lead to additional, more meaningful indicators than the directly measurable ones. For example, Factor and Clustering methods, such as the ones presented in the following subsection, are applied on visitor surveys (e.g., service quality/satisfaction surveys) in order to dynamically define visitor profiles and their typology. Measures such as duration of stay, use of touristic resources, accommodation requirements, etc., can then be estimated per identified profile and associated to specialized policy rules.

The result of the information layer is a specification of economic indicators and a corresponding meta-data model. A mapping is established between high-level indicators and measured data. Theoretical or data-driven estimation and prediction models are incorporated, defining the use of the economic information in the calculation of carrying capacity. A localization framework for the calculation of the indicators is established after the mapping of the current situation, which meets the standards of methodology of smart specialization, as it was implemented in the region of Cyprus (Avdimiotis et al., 2015). To accomplish the mapping task, several analyses are employed: (i) a P.E.S.T. analysis to acknowledge political, social, cultural, technological, and economic environment, (ii) VRIO analysis of competitive advantages, and (iii) finally a SWOT analysis.

**Environmental information.** Data acquisition of environmental data is followed by the definition and customization of indicator sets towards the estimation of tourism carrying capacity. The environmental indicator framework is standardized by developing a corresponding meta-data model. Existing environmental monitoring frameworks are adopted as much as possible, keeping track of international efforts in this field. Further progress is pursued by elaborating on the localization aspects and on progress in the data analysis/mining field. More specifically, leveraging on the rich data sources established by the research proposal (network of sensors) and the site-specific destination analysis, the environmental data streams are mined for patterns and associations with events and actions, leading to prediction models that may explain the
dependencies between tourism development and the environment. In this way, optimal localized indicators and early warning prediction models can be constructed.

Knowledge Engineering Layer

Knowledge is critical and strategic asset, and the key to success in environment, as it facilitates capacities essential for solving problems in highly dynamic environments like ours. For this reason, the platform integrates a layer for appropriately handling the knowledge generated in the lower layers. Therefore, a knowledge engineering Framework is developed. This framework provides with all background and domain knowledge on the impact of tourism on society, economy, and environment. Whereas traditional information systems are used in this context to organize information by indexing, the real advantage of the platform is the capability to organize information by indicating connections between different elements.

It is also important to note that knowledge from several sources is integrated into a unique but modular knowledge base, where each module addresses a separate domain and is tailored to each of the three pillars of sustainable tourism. Technologies that are employed include ontologies (for knowledge representation), rule-based systems and statistical modeling (for inference), and neural networks and Bayesian networks (for prediction).

In addition to the knowledge model, important concepts in the knowledge engineering layer are those devoted to the exploitation, maintenance, and sharing of the knowledge base: (a) Concerning the exploitation, a localization and parametrization framework is developed, which allows the intelligent system to deal with different sets of indicators, calculation methods, carrying capacity limits, priorities, and policies for different areas, considering also input parameters. The resulting complication level is deemed necessary, since the ambition of the research proposal is to establish an extensible, scalable, and generalizable infrastructure for tourism development support at worldwide level. (b) Concerning the maintenance, an efficient knowledge maintenance mechanism is explicitly foreseen and implemented, in order to ensure that the content of the knowledge base is dynamically updated and remains valid on a long-term basis. In this way, the intelligent system functions as a live learning organism. (c) Concerning the knowledge sharing, implementing capabilities for exchanging knowledge with other systems so that each system gets access to more than the knowledge it has been able to build up. The goal here is to reuse knowledge so that unprecedented scenarios can be resolved satisfactorily.

Inference, Decision Making and Policy Modeling

In this layer, intelligent mechanisms are developed that use the knowledge resources in the knowledge base (KB) to provide solutions for policy making and monitoring problems. Human expertise on policy making together with problem solving knowledge extracted from data is introduced in the intelligent engine, in order to solve complicated problems. The core of this layer is policy modeling, i.e., to construct a complete model of sustainable development policies, including rules, limitations, and actions, considering the complex interrelations between policy considerations together with the specific local context. The model built on top of the knowledge engineering layer, so that policies are expressed using the terminology, information definitions, and logic framework developed.

In this stage, extensive consultation and interviews (interviews, focus groups) are conducted with policy makers in the pilot areas. Using hierarchical conceptual models for sustainable tourism policy (Farsari, 2012) as a basis, policy makers in each area elaborate on their own expertise together with the local circumstances.
Using cognitive mapping methods, mental models of policy makers regarding sustainable tourism policies are developed, calibrated, formally documented, and inserted into the system. The aim is to provide policy makers a tool to manage the complexity of sustainable tourism as well as an informed one based on experience (tacit knowledge), academic knowledge, and a powerful monitoring system.

The development of the model takes place in several steps. First, individual models are built for different policy makers in an area. At a second stage individual models are merged to an aggregated one and at a third stage policy makers are invited in group meeting(s) to discuss the merged model and agree on an aggregated version of it. Conceptual models built based on the theory of sustainable tourism are used to further inform the process and ensure that theoretical perspectives in a holistic approach in sustainable tourism are interested into the system. At a fourth stage, the models built for each area are merged further into an overall, hyper-aggregated model of policy-making. At a fifth step, workshops are organized to bring together policy makers from all the pilot areas. During these workshops, besides networking and exchange of know-how, policy makers have the opportunity to discuss the various aggregated models and discuss the hyper-aggregated model to amend it and reach some consensus into an overall model for sustainable tourism development. At a sixth stage, this hype-aggregated model is formalized and documented, inserted into the knowledge, to be made available for testing into the pilot areas. As a seventh stage, feedback is received from users and any necessary amendments on the model are done.

**Application Layer**

The applications to be implemented are “Policy making decision support” and “Monitoring the implementation of policies and the impact of touristic development”. The purpose of this layer is to provide an interface to the users and to implement the necessary functionality for carrying out problem solving tasks. The user environment consists of dashboards where the policy maker is able to select the addressed problem, to set parameters and view the results. The application layer posts queries to the inference layer and receives answers to structured problems.

**Conclusions and Discussion**

Public authorities and DMOs currently responsible for compiling development plans (on national, regional, or local level) are considering spatial, economic, and social parameters; however, planning often relies on obsolete and partial information and is performed using informal methods. Therefore, there is an imperative need for a new approach in policy planning. The research work meets that need and provides a decision support system in two ways: (a) as predefined scenarios, where supports to typical structured problems are provided through parametrizable user-friendly graphical interfaces (e.g., considering current load and projected impact on the environment and local community, to advise on which elements it is proposed to strengthen in specific areas, what is the acceptable level of further development, and which are the suggested best actions), (b) through free queries to the inference engine of the knowledge base in a high-level query language.

Additionally, the research proposal incorporates new technological infrastructure feeding with useful information on the impact of tourism to destinations, in order to (a) provide early warning on alarming situations which may need actions, (b) detect violations of regulations, (c) monitor the progress in the implementation of policies and provide feedback regarding their success.
References


