Marketing Review St.Gallen

Service Management in the Age of AI and Robotics

Schwerpunkt

Digitalisierung trifft Tradition – Die Jungfraubahnen im neuen Zeitalter

How Many Tourists Next Weekend? – Conceptualizing Al-based Forecasts for Visitor Management in Tourism

Employee Responses to Robots in Hospitality

Acceptance of Social Service Robots in the Hotel Industry

Artificial Intelligence in Family-Owned Tourism Businesses

KI bei Bürgerdienstleistungen von Gemeinden

Chatbots für Dienstleistungen erfolgreich nutzen – Ausgewählte empirische Befunde und Praxisimplikationen

Spektrum

KI-basierte Verhandlungsroboter im B2B-Bereich – Einfluss, Status quo und Zukunftsperspektiven

Mittelstand meets KI-Zukunft!



How Many Tourists Next Weekend?

Conceptualizing Al-based Forecasts for Visitor Management in Tourism

Since the advent of the COVID crisis, crowding and the management of crowds have become increasingly important issues for destination managers. Although researchers have been discussing visitor management for a long time, widespread implementation has only started in 2020. Drawing on the findings of a research project in Germany, the authors conceptualize the role of data sources and forecast mechanisms for smart, data-driven destination management and the implementation of visitor management policies.

Prof. Dr. Robert Keller, Stefan Neubig, Dominik Rebholz, Dr. Dirk Schmücker

In destination management, five main motivations for implementing visitor forecasts can be identified (Figure 1). These five motivations cover various levels of destination management in general and tourism visitor management in particular.

On an operational level, counts and forecasts can be used for visitor management or resource planning (Schmücker et al., 2023). To mitigate overcrowding, counts and forecasts can be communicated to potential or actual visitors with the goal of influencing and "nudging" them to times or spaces that are or will be less overcrowded than others. To facilitate operational planning, utilization management, pricing and marketing activities, forecasting can help businesses match their resources (e.g. personnel) to the expected number of visitors: If a restaurant, hotel, leisure attraction, or tour guide organization anticipates a busy weekend, it might need to revise its shift schedule.

On a more long-term planning level, infrastructure such as parking lots, boat moorings in a marina, or new hiking trails

Management Summary

The paper presents five motivations for implementing visitor counts and forecasts in tourism destinations, with visitor management being the prevalent one. The authors conceptualize different forecasting methods based on available data sources and give concrete recommendations for destination managers.



Prof. Dr. Robert Keller Research Professor, Institute for Innovative and Sustainable Tourism, University of Applied Sciences Kempten, Kempten, Germany

robert.keller@hs-kempten.de

Stefan Neubig

Head of Data Engineering & Al, Outdooractive AG, Immenstadt, Germany stefan.neubig@outdooractive.com

Dominik Rebholz PhD Student,

Institute for Innovative and Sustainable Tourism, University of Applied Sciences Kempten, and Fraunhofer Institute for Applied Information Technology FIT, Business & Information Systems Engineering, Sankt Augustin, Germany dominik.rebholz@hs-kempten.de

Dr. Dirk Schmücker

Scientific Director, NIT–Institut für Tourismus- und Bäderforschung in Nordeuropa GmbH, Kiel, Germany dirk.schmuecker@nit-kiel.de

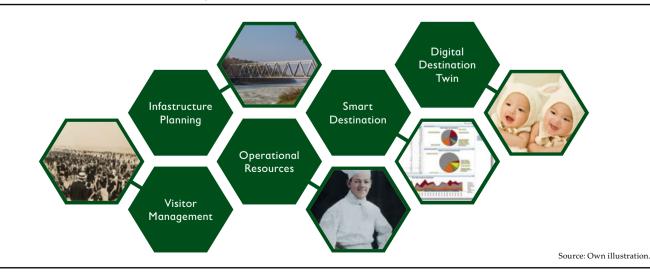


Figure 1: Five Motivations for Using Visitor Counts and Forecasts in Tourism Destinations

can be dimensioned according to counts and forecasts. In all these cases, counts and forecasts need to cover a longer time period than for operational problems.

On the experience level, counts and forecasts can form a building block of a digital destination twin (Samochowiec et al., 2019). A digital destination twin is the digital image of a destination available to actual or potential visitors, businesses, and destination managers. With a digital twin, they have a multitude of different destination data at their fingertips, e.g. water temperatures, available number of rental bikes, open restaurants, delays in public transportation and many more. Counts and forecasts (how crowded a place is now and how crowded it is expected to be in the near future) can be an element of such a digital twin.

On a managerial level, visitor counts and forecasts can form a building block of smart tourism destinations (Gretzel, 2021; Gretzel et al., 2015). The managerial approach of smart destinations (comparable to the better-known concept of smart cities) is that of data-driven and therefore fact-based destination management. It is often materialized in dashboard systems that provide a holistic view of the destination's performance. Counts and forecasts can be fed into such dashboard systems to inform managerial processes.

A Tourism Visitor Management Process

Visitor management is one of the five motivations for using visitor counts and forecasts. Visitor management is the effort to manage "visitor flows, trajectories or corridors" (Beritelli et al., 2020); it can involve different forms of interventions, e.g. "deterrence and enforcement, communicating with visitors and providing education" (Glasson et al., 1995, p. 270) or "information, nudging, pricing, reservation and stopping" (Schmücker et al., 2023).

In order to be able to manage visitor flows through interventions, four consecutive steps need to be taken to form a specific kind of "management circle": generating data, storing and exchanging data, building models and forecasts, and, finally, generating and deploying the results (Figure 2). The circle is closed by measuring the consumers' reaction to the recommendations. This is particularly helpful when this data is used as training input for AI-based models, leading to a refinement of forecasts and recommendations.

Artificial intelligence (AI), specifically machine learning (ML) methods, can be used at various stages of the process. Data can be generated by local sensors, e.g. smart cameras that automatically perform counts and determine the direction of

Main Propositions

- 1 Implementing visitor management systems is more complex than most people think and strongly depends on its overall objective.
- 2 Moderate forecasting inaccuracy as a consequence of simpler models may be okay as concrete visitor numbers are often not needed.
- 3 While machine learning provides the highest accuracy, its data quality and quantity requirements cannot always be met immediately.
- 4 Rule-based anticipation and causal models provide scalable alternatives for cases with less stringent accuracy requirements.

movement. Data can also be generated by global sensors, e.g. background processes in mobile applications for smartphones, which transmit the position of the device. Both methods of generating data can be based on machine learning algorithms.

In this paper, however, we want to focus on the role of ML algorithms in modeling and forecasting demand. The remainder of this paper is devoted to this topic.

Figure 2: The Main Steps of a Visitor Management System

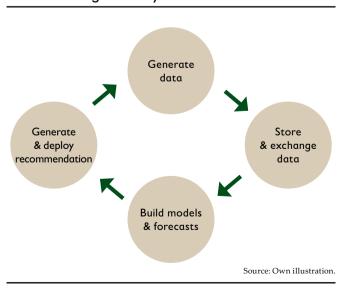
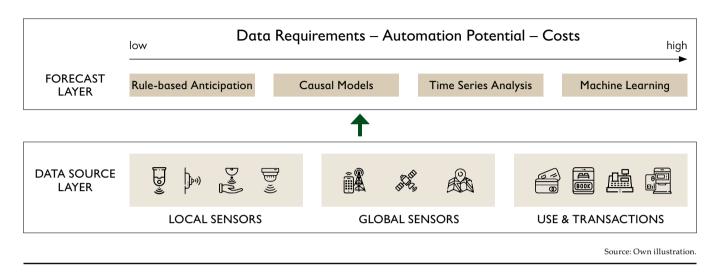


Figure 3: Design Layers of Visitor Forecasts



Design Layers of a Visitor Forecast

To address the challenge of achieving adequate prediction accuracy for a visitor forecast, we need to consider two design layers (Figure 3): first, the data source layer, which encompasses data generation via sensors or IT systems; second, the forecast layer, which can be categorized according to four forecast methods and algorithms. The design of a visitor forecasting strategy depends on the decisions made in the forecast and data source layers, which result in several trade-offs.

In the following sections, we will introduce the data source layer, the forecast layer, and the trade-offs between them. We then present the application of these layers in real use cases from the AIR (AI-based recommender for sustainable tourism) research project.

Data Source Layer

At the data source level, we can identify three categories: local sensors, global sensors, and usage and transaction data (Schmücker & Reif, 2022):

• Local sensors encompass, for example, photoelectric sensors, cameras, and wifi counters. They can be used to measure traffic or visitor movements in cities or rural areas (Muhar et al., 2002; Taczanowska et al., 2018; Zubiaga et al., 2019). The advantages of local sensors are their accuracy and the topicality of the data. The disadvantages are the installation costs, the limited area of coverage, and the high cost of maintenance.

- Global sensors encompass, for example, mobile data, satellite images, or app data. This data is usually aggregated to ensure privacy. The data can include visitor movements in a target area (Ahas et al., 2015; Hardy et al., 2017), but also additional information about the guests, such as their origin (Caceres et al., 2007) or demographics (Monz et al., 2021). The advantages of global sensors are the large target area and the unique types of information. Disadvantages are the limited transparency and the potentially large time lag of the data when combined with sociodemographic information.
- Usage and transaction data encompass, for example, payment data, ticketing data, or subscription data. This includes on-site transactions and online booking data, but also web search data and web page traffic data (Li et al., 2018). Usage and transaction data has a lot of untapped potential with a large amount of unstructured data, but it is often not available because it is mostly owned by companies.

In addition to this data, expert knowledge may be used instead of measured data to generate rules for visitor management systems. Often, local experts and destination management organizations (DMOs) have specialized experience and insight into visitor behavior, including knowledge of crowd dynamics, visitor flows, and visitor occupancy. This data is rich but more subtle.

Forecast Layer

At the forecast level, we discuss four methods: rule-based anticipation, causal models, time series analysis, and machine learning.

- *Rule-based anticipation via expert opinion* is the most basic mechanism. It reflects expert opinion and local knowledge, which is reflected in the hand-crafted rules of the recommender systems (Neubig et al., 2024). Thus, for instance, the system flags high expected crowding or suggests different locations. These systems are easy to implement and configure, but they strongly depend on the knowledge of experts.
- *Causal models* are built using configurational comparative methods (CCMs). CCMs identify the effects of feature combinations and express the model using Boolean algebra (Yakovchenko et al., 2020). These methods have been effective in uncovering causal relationships not previously identified by traditional regression analysis (Whitaker et al., 2020). They can provide a more quantitative approach to rule-based systems and allow for a more fluid process of updating rules. However, they struggle with incomplete or missing data.
- *Time series analysis* is one of the most common tools for analyzing patterns in visitor flows. This method is based on traditional regression statistics and provides a good forecast of visitor numbers (Cellini & Cuccia, 2013). Time series models are very useful for determining long-term trends in large areas but may be insufficient for fine-grained locations with complex external effects.
- *Machine learning models* are a superclass of different models, from regression to decision trees to deep learning (Alzubi et al., 2018). These models can analyze enormous amounts of different data and search for patterns that indicate high or low crowding. Their main drawback is that they require huge amounts of fine-grained, highly accurate data (Bollenbach et al., 2022; Stohr et al., 2024). The procurement of such data can be quite costly, and, in the case of time series data, they cannot be obtained at any cost, but only by waiting for the time series to fill up.

Table 1. Rule based / Rulepation to Guide oking Visitors in 2000 Froundain Ruliges	
Setting	This low mountain area is a popular skiing resort, with a catchment area that includes large parts of northern Germany, the Netherlands, and Denmark. At peak times, it suffers from limited parking capacity and congested roads on the arrival of (especially) day tourists.
Objective	Tourists should be provided with information on nearby alternative skiing areas and f parking availability. Ideally, tourists should already receive this information at home during their planning phase.
Data Source Level	The area has recently installed several sensors that have been collecting data for about one year at its main parking lots.
Application	We tried to use machine learning models to predict future parking space utilization. However, since the data is quite scarce, these models did not perform well. Thus, we went back to rule-based anticipation algorithms. However, we might switch back to other forecast algorithms when more data is available.

Table 1: Rule-based Anticipation to Guide Skiing Visitors in Low Mountain Ranges

Source: Own illustration.

Table 2: Causal Models to Predict Hiking Activities in the Alps

Setting	The Alps are a popular hiking area. Many tourists use the various hiking trails that lead to popular points of interest such as the various lakes, mountains, castles and alpine lodges.
Objective	Since parking lots are quite limited, we aimed for a general data-based prediction of which days will have a high number of visitors.
Data Source Level	Local sensors have been installed along some of the hiking routes to track visitor numbers for an average of approximately one year. However, most hiking routes and POIs do not have sensors installed. In addition, we used holiday and weather data.
Application	Like rule-based anticipation, causal models are more general in their predictions. They can provide an overview of which features are causally connected to tourists going hiking, but they cannot predict a specific number of visitors for a particular day. The causal models can be generally applied to a larger region without providing specific information for particular points of interest. Additionally, with the insights from these data-driven models, we were able to generate some rules that might enrich rule-based systems built from expert opinion.

Source: Own illustration.

These approaches do not necessarily have to be considered in isolation. They can also be used in parallel, for example, by combining different forecasting methods with different confidence values or by combining multiple methods. Another approach is to define one or more standard methods while providing low-threshold intervention options for (trained) human experts (e.g., DMO representatives), such as simple switching between methods ("human in the loop" approach).

The Forecasting Paradox

In visitor management, a forecasting-related paradox can occur: An effective visitor management system will cause visitors to adapt in such a way that the forecast becomes incorrect. This feedback loop can lead to a situation where a recommender system directs visitors to a cold spot because the originally desired point of interest (POI) is predicted to be a hot spot. If visitors follow the recommendation, the cold spot becomes hot and the hot spot becomes cold. Therefore, continuous recalibration is required. The more complex the model, the more computationally intensive and expensive the recalibration.

Illustrative Cases

To illustrate the application of the data source layer and the forecast layer, we construct use cases based on real-world examples. Most of them are based on the German research project AIR, which has been running since 2022. AIR is a joint research project that aims to improve digital visitor management in various German tourist destinations by effectively managing visitor flows. An effective visitor management system requires a visitor forecast. As illustrated in figure 3, there are different methods for visitor forecasting, depending on the data sources

Table 3: Time Series Analysis to Predict Overnight Tourism Demand at the North Sea Coast

Setting	The North Sea coast is a popular holiday destination in Germany. Its utilization is highly seasonal and dependent on national demand.
Objective	We analyzed the demand for overnight stays for a place on the German North Sea coast. The objective was to investigate which parameters would be suitable for forecasting tourism demand.
Data Source Level	Our data included daily overnight demand figures over four years, enriched with school holiday and weather data.
Application	We used random/fixed effects regression models to test the parameters. We found that the mean daytime temperature and the timing of school holidays in Germany were the best fixed-effect variables and that the month of the year, as a proxy for seasonality, was the best random-effect variable.

Source: Own illustration.

Table 4: Machine Learning to Predict Day Visitors at the Baltic Sea Coast

Setting	The area is a beach with several sections. Light barriers have been in place for several years to record the number of people passing through.
Objective	Beach visitors need space for their towels and umbrellas. Accordingly, some parts of the beach can be very crowded on some days. These days should be predicted to be able to point out less crowded beach sections.
Data Source Level	The light barriers have been collecting data for several years. So, there is a good data history. We also used holiday and weather data.
Application	We tested various machine learning algorithms. In this specific case, random forest regression was particularly suitable for short-term forecasts, whereas XGBoost worked best for medium-term forecasts. The more fine-grained the individual sections of the beach are, the harder they are to predict because people often walk along the beach.

Source: Own illustration.

available. The following four use cases are intended to illustrate how data sources and forecasting mechanisms can work together:

- Skiing visitors in low mountain ranges (Table 1): The maximum parking capacity of a popular skiing resort in the area is often exceeded during winter, leading to road congestion. An intelligent visitor management system can automatically predict high visitor numbers and redirect tourists to alternative skiing areas.
- Hiking in the Alps (Table 2): Popular POIs along hiking routes are often crowded, resulting in environmental degradation, congested parking lots, and tourist dissatisfaction. Alternative cold spots can be recommended to balance the number of visitors.
- Overnight tourism on the North Sea coast (Table 3): Visitor numbers at North Sea coastal tourist destinations are subject to strong seasonal fluctuations. Better forecasting of visitor numbers can help stakeholders in the area with their capacity planning.
- Day visitors to the Baltic Sea coast (Table 4): By measuring visitor entries and exits to different beach segments, beach occupancy can be predicted.

Lessons Learned

- 1 *Focus on your goals:* Be clear about what you want to achieve. Only concrete goals allow you to select the most suitable forecasting algorithm.
- 2 Adopt simple models: Use the least complex model that aligns with your objectives. This is likely to save costs and accelerate the market entry of your visitor management system.
- -
- 3 Implement a human-in-the-loop system: Use expert supervision to evaluate automated forecasts and enable manual, low-threshold intervention when necessary. Retain interventions as valuable feedback for further development.
- 4 *Emphasize data quality:* Following the "garbage in, garbage out" paradigm, consider the importance of data quality when applying data-driven methods. If your data is not (yet) sufficient, try simpler solutions.

Key Findings and Learnings

Based on the conceptual considerations and the experiences from use cases and real-world implementations, we discuss our key findings and learnings.

First, implementing a visitor management system is a complex endeavor: A key finding we keep coming across in our research projects is that implementing a visitor management system is more complicated than anticipated. One major challenge is the definition of concrete goals and target variables. A complete specification of the problem and solution space is necessary to determine the next steps for implementing appropriate forecasting systems. Another challenge is overcoming bureaucratic or technocratic procedures when installing local sensors or convincing businesses to give access to their transaction data.

Second, moderate inaccuracy may be sufficient: Depending on the desired visitor management system, not all cases require the highest accuracy. An essential requirement is often to identify expected peak loads rather than actual visitor numbers. In these cases, opting for a less complex and expensive approach may make more economic sense. Moreover, an exact prediction may sometimes even be unrealistic. In some cases, we have found that complex machine learning models quickly reach their limits. At the same time, the devices that collect the necessary data may be limited in their widespread applicability (e.g., for financial or regulatory reasons). In such cases, it may make sense to take a closer look at the use case: What level of precision is required? If it is only a matter of identifying load peaks (e.g., to identify hot and cold spots),

Be clear about what you want to achieve with your visitor management system, then think about technology and algorithms.

simple, rule- and experience-based algorithms may be sufficient. These have the advantage of being independent of the presence of data. By including the option to add irregular events, they may even outperform data-based approaches that do not account for such events in their historical data. Otherwise, rule-based approaches may be less accurate and may be used for scenarios that do not require an accurate estimate but instead focus on peak values.

A third lesson is that it is advisable to keep human knowledge in the loop: In some cases, the trade-off between simplicity and complexity can be difficult to resolve. Even when the data is not sufficient for a full application of machine learning algorithms, machine learning has delivered sound forecasts on many days. A hybrid approach with a "human-in-the-loop", where machine learning approaches make a suggestion and a human optionally switches back to a rule-based baseline in case of suspected inaccuracies, could be a promising solution. Ideally, the machine learning algorithms can even capitalize on this human feedback and incorporate it into future predictions.

Fourth, we learned that causal models (CCMs) can improve rule-based anticipation. Like rule-based anticipation, causal models are more general in their predictions. They can provide an overview of which features are causally related to tourists going hiking, but they cannot predict a specific number of visitors for a particular day. In addition to experience-based occupancy rules, causal models can also be used to derive new, data-driven rules that fit seamlessly into the framework of existing rules. Since the prediction of visitor numbers is not stable and rules need to be constantly updated, we see great potential in combining these two technologies.

Finally, we conclude that machine learning provides the highest accuracy – when done correctly. Despite temporary difficulties, our machine learning-based approaches have produced impressive results. The sticking points in the application of machine learning that need to be overcome is the data, which must be suitable in terms of quality and quantity, and the selection of the right algorithms. Although machine learning algorithms become more complex, the more fine-grained the predictions need to be, we expect them to deliver the best results when the above principles are properly considered.

References

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., Ziemlicki, C., Tiru, M., & Zook, M. (2015). Everyday space-time geographies: Using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn. International Journal of Geographical Information Science, 29(11), 2017–2039. https://doi.org/10.1080/13658816.2015.1063151

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine learning from theory to algorithms: An overview. Journal of Physics: Conference Series, 1142(1). https://doi.org/10.1088/1742-6596/1142/1/012012

Beritelli, P., Reinhold, S., & Laesser, C. (2020). Visitor flows, trajectories and corridors: Planning and designing places from the traveler's point of view. Annals of Tourism Research, 82. https://doi.org/10.1016/j.annals.2020.102936

Bollenbach, J., Neubig, S., Hein, A., Keller, R., & Krcmar, H. (2022). Using machine learning to predict POI occupancy to reduce overcrowding. In D. Demmler, D. Krupka, & H. Federrath (Eds.), Informatik 2022: Informatik in den Naturwissenschaften (pp. 393–408). Gesellschaft für Informatik. https://doi.org/10.18420/inf2022_34

Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin-destination data from a mobile phone network. IET Intelligent Transport Systems, 1(1), 15–26. https://doi.org/10.1049/iet-its:20060020

Cellini, R., & Cuccia, T. (2013). Museum and monument attendance and tourism flow: A time series analysis approach. Applied Economics, 45(24), 3473–3482. https://doi.org/10.1080/00036846.2012.716150

Glasson, J., Godfrey, K., & Goodey, B. (1995). Towards visitor impact management: Visitor impacts, carrying capacity, and management responses in Europe's historic towns and cities. Urban and regional planning and development. Avebury; Ashgate Pub. Co.

Gretzel, U. (2021). Smart tourism development. In P. U. C. Dieke, B. King, & R. Sharpley (Eds.), Tourism in development: Reflective essays (1st ed., pp. 159–168). CAB International. https://doi.org/10.1079/9781789242812.0013A

Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. Electronic Markets, 25(3), 179–188. https://doi.org/10.1007/s12525-015-0196-8

Hardy, A., Hyslop, S., Booth, K., Robards, B., Aryal, J., Gretzel, U., & Eccleston, R. (2017). Tracking tourists' travel with smartphone-based GPS technology: A methodological discussion. Information Technology & Tourism, 17(3), 255–274. https://doi.org/10.1007/s40558-017-0086-3

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. Tourism Management, 68, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

Monz, C., Creany, N., Nesbitt, J., & Mitrovich, M. (2021). Mobile device data analysis to determine the demographics of park visitors. Journal of Park and Recreation Administration, 39(1). https://doi.org/10.18666/10.18666/jpra-2020-10541

Muhar, A., Arnberger, A., & Brandenburg, C. (2002). Methods for visitor monitoring in recreational and protected areas: An overview. In A. Arnberger, C. Brandenburg, & A. Muhar (Eds.), Monitoring and Management of Visitor Flows in Recreational and Protected Areas: Conference Proceedings (pp. 1–6).

Neubig, S., Bečevová, M., Brosda, F., Loges, R., Hein, A., Keller, R., & Krcmar, H. (2024). Beyond sensors: A rule-based approach for cost-effective visitor guidance. In K. Berezina, L. Nixon, & A. Tuomi (Eds.), Information and communication technologies in tourism 2024: ENTER24 International eTourism Conference, Izmir, Türkiye, January 17–19. Springer Nature.

Samochowiec, J., Kwiatkowski, M., & Breit, S. (2019). Unterwegs mit smarten Assistenten: Ein Szenario zum Reisen der Zukunft. Gottlieb Duttweiler Institute. https://doi.org/10.59986/SVTO2855

Schmücker, D., Reif, J., Horster, E., Engelhardt, D., Höftmann, N., Naschert, L., & Radlmayr, C. (2023). The INPReS intervention escalation framework for avoiding overcrowding in tourism destinations. Tourism and Hospitality, 4(2), 282–292. https://doi.org/10.3390/tourhosp4020017

Schmücker, D., & Reif, J. (2022). Digitale Besuchermessung im Tourismus: Ziele, Methoden, Bewertungen. UVK Verlag.

Stohr, A., Ollig, P., Keller, R., & Rieger, A. (2024). Generative mechanisms of AI implementation: A critical realist perspective on predictive maintenance. Information and Organization, 34(2). https://doi.org/10.1016/j.infoandorg.2024.100503

Taczanowska, K., Latosinska, B., Czachs, C., Hibner, J., Muhar, A., Brandenburg, C., & Rothert, M. (2018). Towards standards for quantification of recreational use in forest areas: Indicators and data collection tools applied by the State Forests National Forest Holding, Poland. In XY Monitoring and Management of Visitor Flows in Recreational and Protected Areas: Conference Proceedings (pp. 481–483).

Whitaker, R. G., Sperber, N., Baumgartner, M., Thiem, A., Cragun, D., Damschroder, L., Miech, E. J., Slade, A., & Birken, S. (2020). Coincidence analysis: A new method for causal inference in implementation science. Implementation Science, 15(1). https://doi.org/10.1186/s13012-020-01070-3

Yakovchenko, V., Miech, E. J., Chinman, M. J., Chartier, M., Gonzalez, R., Kirchner, J. E., Morgan, T. R., Park, A., Powell, B. J., Proctor, E. K., Ross, D., Waltz, T. J., & Rogal, S. S. (2020). Strategy configurations directly linked to higher Hepatitis C virus treatment starts: An applied use of configurational comparative methods. Medical Care, 58(5), 31–38. https://doi.org/10.1097/MLR.0000000001319

Zubiaga, M., Izkara, J. L., Gandini, A., Alonso, I., & Saralegui, U. (2019). Towards smarter management of overtourism in historic centres through visitor-flow monitoring. Sustainability, 11(24). https://doi.org/10.3390/su11247254