Forecasting camping tourism demand in America's national parks using a machine learning approach

William L. Rice
The Pennsylvania State University

So Young Park
The Pennsylvania State University

Bing Pan
The Pennsylvania State University

Peter B. Newman
The Pennsylvania State University

Forecasting camping tourism demand in America's national parks using a machine learning approach

Introduction

The campground, as it is found in US national parks, is a uniquely American invention (Meinecke, 1932; Young, 2018). It was forged from a shift in national park use during the early portion of the twentieth century in which organized automobile camping boomed into a major form of tourism (Hogue, 2016). Today camping tourism generates $166 billion annually (The Outdoor Industry Association, 2017). Yet, it remains under researched as a sector of tourism (Triantafillidou & Siomkos, 2013). Markedly, this deficit in research remains even as the demand for camping grows. To date three major questions remain unanswered: 1) what is driving demand, 2) how might the increased demand be met, and 3) how accurately can future demand be predicted. The main focus of this research is the final of these questions. Applying advanced forecasting methodology, including machine learning methods, we seek to understand how the characteristically unique demand for US campsites can be forecasted into the future.

Literature Review

Increasing Campsite Demand

America's national parks hosted over eight million camping overnight stays in 2017 (National Park Service, 2018). From 2013 to 2017, the amount of tent and vehicular campers in national parks increased 24.7% and 24.5%, respectively. Research shows that these campgrounds remain mostly at capacity during the peak summer seasons, with some remaining filled during shoulder seasons (Walls, Wichman, & Ankney, 2018). For some campgrounds in coveted locations, such as Yosemite Valley, visitors arrive in the afternoon prior to their anticipated stay in order to stake their spot in line (Almond, 2017).

Limited Campsite Supply

National parks in the United States are managed by the National Park Service (NPS). The NPS was established under a dual-mandate to both provide for the enjoyment of the people and leave parks unimpaired for future generations (Sax, 1980). Within the context of increasing demand for campsites, the upshot of this ecologically centered mission is a largely prohibitive environment towards the creation of more campsites. These leaves an unmet demand and no means of increasing supply.

Forecasting

Previous attempts to forecast campground demand has been based in largely obsolete methods such as basic linear trend analysis (Marin-Pantelescu, 2015). The importance of accurate demand forecasting for proper management has been emphasized repeatedly in the tourism literature (Hassani, Silva, Antonakakis, Filis, & Gupta, 2017; Li, Song, & Witt, 2005). It has also been shown, however, that there is no one universally superior forecasting method when applied to tourism (Song & Li, 2008). Considering the unique nature of campground demand, being both highly seasonal and linked to an inferior good, we have opted to cast a broad net of forecasting methods—six—in order to assess their effectiveness.

Methodology

Campground reservation data gathered from the Recreation Information Database (RIDB) was used for the analysis. RIDB data is populated from recreation.gov, an online reservation system servicing federally managed public lands in the United States. Historical data is available dating back to its inception in 2007 (Supak, Brothers, Ghahramani, & Van Berkel, 2017). Given the scope of our analysis, only campgrounds operated by the NPS were examined. After removing erroneous information, as in Supak et al. (2017), we were left with a dataset including 3,685,260 reservation records from 108 campgrounds across 32 national parks.

Each of the 108 campgrounds with available data varied widely in their seasonality, available records, administrative rules, and number of years cataloged via recreation.gov. This, combined with computational limitations, led us to choose five campgrounds for an in-depth forecasting analysis: Big Meadows Campground in Shenandoah National Park, Elkmont Campground in Great Smoky Mountains National Park, Mather Campground in Grand Canyon National Park, Moraine Park Campground in Rocky Mountain National Park, and Upper Pines Campground in Yosemite National Park.

Since this is the first study of its kind, we conducted six unique forecasting methods in order to forecast campsite demand. Five forecasting models were selected from the existing tourism literature (Frechtling, 2012; Hassani et al., 2017; Song & Li, 2008) as well as two less explored approaches to both compare their predictive ability and identify the most accurate model: Moving average (MA), ETS with Holt-Winters seasonal method, SARIMA, NNAR, KNN, and the combination of ETS, SARIMA, NNAR and KNN.

Results

Final results are forthcoming, however we have successfully employed each of the above forecasting methods on RIDB data using R Statistical Software (version 3.5.1).

References

Almond, E. (2017, Jun). *Secrets of Yosemite: How to snag a last-second campsite.* The Mercury News. Retrieved from https://www.mercurynews.com/2017/06/07/secrets-of-yosemite-how-to-snag-a-last-second-campsite/

Frechtling, D. (2012). *Forecasting tourism demand*. Routledge.

Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., & Gupta, R. (2017). Forecasting Accuracy evaluation of tourist arrivals. *Annals of Tourism Research, 63*, 112–127.

Hogue, M. (2016). *Thirtyfour Campsites*. Cambridge, MA: The MIT Press.

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 44*(1), 82–99.

Marin-Pantelescu, A., (2015). The perspective of camping tourism in Romania. *Academic Journal of Economic Studies, 1* (4), 45–53.

Meinecke, E. P. M. (1932). *A Camp Ground Policy*. Ogden, UT.

Sax, J. L. (1980). *Mountains without handrails: Reflections on the national parks*. Ann Arbor, MI: University of Michigan Press.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—a review of recent research. *Tourism Management, 29* (2), 203–220.

Supak, S., Brothers, G., Ghahramani, L., & Van Berkel, D. (2017). Geospatial analytics for park & protected land visitor reservation data. In Z. Xiang & D. R. Fesenmaier (Eds.), *Analytics in smart tourism design* (pp. 81–109). Berlin, Germany: Springer.

The Outdoor Industry Association. (2017). *Outdoor recreation economy report*. Boulder, CO: The Outdoor Industry Association.

Triantafillidou, A., & Siomkos, G. (2013). Summer camping: An extraordinary, nostalgic, and interpersonal experience. *Journal of Vacation Marketing, 19* (3), 197–208.

Walls, M., Wichman, C., & Ankney, K. (2018). *Nature-Based Recreation : Understanding Campsite Reservations in National Parks*. Washington, DC.

Young, T. (2018). E.P. Meinecke and the Development of the Modern Auto Campground. *IdeAs*, *12*, 1–22.