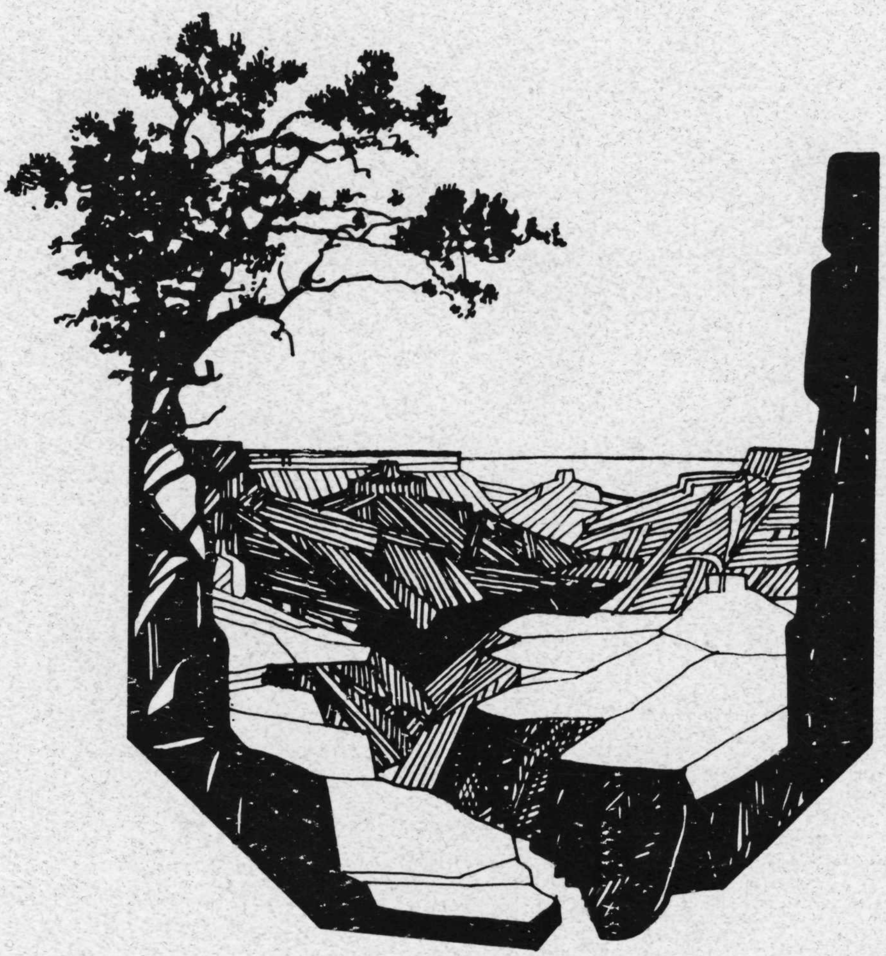
Forecasting Visitation at Grand Canyon National Park Using Time Series Analysis

Peter G. Rowlands

Technical Report NPS/NAUGRCA/NRTR-93/03



National Park Service
U.S. Department of the Interior

Cooperative Park Studies Unit
at Northern Arizona University

National Park Service
Cooperative Park Studies Unit
Northern Arizona University

The National Park Service (NPS) Cooperative Park Studies Unit (CPSU) at Northern Arizona University (NAU) is unique in that it was conceptualized for operation on an ecosystem basis, rather than being restrained by state or NPS boundaries. The CPSU was established to provide research for the 33 NPS units located within the Colorado Plateau, an ecosystem that shares similar resources and their associated management problems. Utilizing the university's physical resources and faculty expertise, the CPSU facilitates multidisciplinary research in NPS units on the Colorado Plateau, which encompasses four states and three NPS regions—Rocky Mountain, Southwest, and Western. The CPSU provides scientific and technical guidance for effective management of natural and cultural resources within those NPS units.

The National Park Service disseminates the results of biological, physical, and social science research through the Colorado Plateau Technical Report Series. Natural resources inventories and monitoring activities, scientific literature reviews, bibliographies, and proceedings of technical workshops and conferences are also disseminated through this series.

Unit Staff

Charles van Riper, III, Unit Leader
Peter G. Rowlands, Research Scientist
Henry E. McCutchen, Research Scientist
Mark K. Sogge, Ecologist
Charles Drost, Zoologist
Elena T. Deshler, Biological Technician
Paul R. Deshler, Technical Information Specialist
Connie C. Cole, Editor
Margaret Rasmussen, Administrative Clerk
Jennifer Henderson, Secretary

National Park Service Review

Documents in this series contain information of a preliminary nature and are prepared primarily for internal use within the National Park Service. This information is not intended for use in open literature. This report has been reviewed and approved for dissemination by the National Park Service. Approval does not signify that the contents necessarily reflect the views and policies of the National Park Service nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Copies of this report are available from the following. To order from the National Park Service, Technical Information Center, use the reference number on the bottom of the report's inside back cover.

National Park Service
Cooperative Park Studies Unit
Northern Arizona University
P. O. Box 5614
Flagstaff, AZ 86011-5614
(602) 556-7466

National Park Service
Technical Information Center
Denver Service Center
P. O. Box 25287
Denver, CO 80225-0287
(303) 969-2130

Forecasting Visitation at Grand Canyon National Park Using Time Series Analysis

Peter G. Rowlands

Cooperative Park Studies Unit
Northern Arizona University
Flagstaff, Arizona 86011

Technical Report NPS/NAUGRCA/NRTR-93/03

October 1993

National Park Service
U.S. Department of the Interior

Cooperative Park Studies Unit
at Northern Arizona University

Table of Contents

List of Figures

List of Tables

Abstract

The purpose of this study was to evaluate visitation trends at Grand Canyon National Park as well as the efficacy of a number of statistical techniques used in forecasting. Visitation data at Grand Canyon National Park were analyzed using time series analysis techniques. Several types of regression models including linear, exponential, polynomial, and transition functions were employed along with exponential smoothing and Autoregressive Integrated Moving Averaging (ARIMA). Projections of visitation were made from 1993 to the year 2005, a 13 year lead time. Results were compared with existing projections based on market analysis.

The most reliable yearly visitation forecasts were those derived using Holt's method of exponential smoothing. Results are comparable to those derived from market analysis, but, unlike the latter more sophisticated method, visitation cannot be broken down into categories such as foreign vs. domestic. Approximately 6 million people are expected to visit Grand Canyon National Park by the year 2005. Visitation at the North Rim, which here includes the North Rim proper, Tuweep, Lees Ferry, and Grand Wash, is predicted to be 740,000. Seasonal analysis revealed that a pattern of seasonal summer spikes and winter troughs of visitation have been the rule for at least a decade; this is expected to continue. There was no evidence of increasing relative importance of "shoulder seasons" as sources of visitation. In instances where market analysis is not feasible due to costs and lack of visitor-use surveys, time series analysis, if used with precaution, can be a valuable tool for projecting potential visitation.

Management recommendations are presented based on the results of this study.

Introduction

Grand Canyon National Park (GRCA) was established in 1919. Since then, the numbers of park visitors have increased from 44,173 during the establishment year to over 4,547,027 in 1992 (J. Sypher, GRCA, personal communication). GRCA as well as many other National Park Service (NPS) units, appears to be receiving ever increasing numbers of visitors with each passing year. In some cases, the "carrying capacity" of the park unit, either in part or in whole, is being overwhelmed. The situations in Yosemite Valley of Yosemite National Park and the South Rim of the Grand Canyon are prime examples of the challenges to park planners and managers of increasing visitor pressure and its effect on park infrastructure as well as natural and cultural resources (National Park Service 1980; General Management Plan—Grand Canyon National Park, in preparation).

Park managers must deal with future visitation trends, whether increasing or decreasing. The limited NPS budget, as well as staff, must be allocated according to both local and regional needs. Management decisions may have to be made on limiting visitation at or below some defined carrying capacity. Such a capacity may be based on the limitations of physical space or infrastructure, resource impacts, e.g., levels of acceptable change (Stankey et al. 1985), data from studies revealing visitor perceptions of crowding, or a combination of these. Accurate forecasts of future park visitation will be helpful in formulating management decisions regarding both the allocation of resources and regulation of visitor use.

Gallipeau (1992) has produced a visitation forecast for the Grand Canyon based on a marketing analysis. Making the tacit assumption that visitation will begin to level off after the year 2000, Galipeau (1992) predicts that by the year 2010, GRCA will experience approximately 6.8 million visitors. Approximately 38% of these would be foreign. However, he did not employ time series analyses (such as ARIMA) on exponential smoothing (Box and Jenkins 1977). Furthermore, forecasts were based on a logistic model based on World Tourism Organization Forecasts, a visitor use study performed at GRCA (Albrecht 1991), and other marketing information such as projected foreign and domestic market shares.

Unfortunately, the resources required to produce reliable, market-analysis based forecasts are not always available due to fiscal constraints or lack of visitor-use and behavioral data. This could be a particular problem for small parks. Forecasting by means of time series analysis should be done with caution (as described below). In some instances where an intensive market analyses is not available, the more straightforward approach of investigating past trends may be a viable alternative.

Forecasting, by no means a straightforward exercise, is one of the more precarious types of statistical inference because it attempts to predict a value outside the region of the sampled data (McClave and Benson 1988). Zar (1984) warns against unwarranted extrapolation using regression equations unless there is good reason to believe that the described function holds for x-values outside the range of those observed, and then only with caution. Wilkinson (1990) adds that forecasting methods and results must be balanced and tempered by extrinsic knowledge, careful examination of residuals, and limited extrapolation beyond the ends of the data. In general, time series forecasting should be confined to the short term. As forecasts are made further into the future, the less certain the accuracy of the forecast. Confidence intervals around the estimate get increasingly broader as the limits of the data are progressively exceeded. Finally, only inferential forecasting models, such as those based on regression, calculate explicit random error components. This allows confidence intervals to be placed around the predictions before the actual value of the time series is observed (McClave and Benson 1988).

This technical report has several objectives:

- Give a brief history of visitation trends at GRCA;
- Compare a number of methods of time series analyses to forecast future visitor numbers based on previous visitation trends;
- Compare the results of time series analyses with forecasts produced by Gallipeau (1992) based on market-analysis.

Methods

Visitation data collected over a period of 73 years came from GRCA files. The park's total visitation data set is the most complete and includes the years 1919 through 1992. Separate North and South Rim data sets are available for the years 1955 through 1992. In the data set, North Rim figures also include visitation at Lees Ferry, Tuweep Ranger Station, and Grand Wash. Specific figures (for only the North Rim of GRCA) were not available. Monthly breakdowns are available beginning in 1980 and continuing through the present time. One of the caveats included with these data was that total park visitation does not always agree with figures derived by adding the North Rim plus the South Rim in any particular year. However, the few discrepancies are very minor and should not affect overall predictions.

A major potential source of error in the data set is due to the park's changing the method of estimating visitor numbers. Daily visitation had always been assessed by counting cars and multiplying by a value based on the average number of visitors per car derived from occasional surveys. This has been the case at least as far back as the early 1940s (J. Sypher, GRCA, personal communication). Multipliers have changed periodically as the park conducted assessments of the average number of visitors per vehicle. Immediately prior to 1991, the multiplier used was 3.4 visitors per car. In 1991, motivated by what seemed to be an overestimation of visitation, the park conducted a systematic survey which revealed that the multiplier varied from 2.2 - 3.4 persons per vehicle depending on the season. Higher values occurred during the summer. Visitation estimates for 1991 and 1992 are based on a more flexible assessment taking into account the seasonal variation in the multiplier. Therefore, park visitation estimates prior to 1991 may be inflated (J. Sypher, GRCA, personal communication).

Several methods of forecasting for the GRCA visitation data set were employed. Data broken down by months between 1980 and 1992 were analyzed both by Winter's seasonal modification of the exponential smoothing method (McClave and Benson 1988) and by Autoregressive Integrated Moving Averaging (ARIMA) developed by Box and Jenkins (1977). Yearly total visitation for the park as a whole as well as separate counts for the North Rim were analyzed by means of Holt's method of exponential smoothing and simple regression (linear and second order polynomial regression) with both untransformed and Log transformed visitation data. Smoothing coëfficients for the level (α) for both the Holt's and Winter's methods were set high, at 0.9, to stress the most recent part of the time series in the calculations. Other weighting coefficients (i.e., for smoothing the estimate of the linear trend [β] and multiplicative seasonality [γ]) were varied by trial and error until the lowest mean error, within series standard error or mean absolute percentage error, depending upon the method, was achieved. Statistical analysis was performed using SYSTAT 5.03 to obtain Holt's exponential smoothing (Wilkinson 1990). Statgraphics 4.0 (Statistical Graphics Corporation 1988) was used to perform regression analysis, Winter's model for seasonal exponential smoothing, and ARIMA since its output is more straightforward than SYSTAT's for these particular tests. TableCurve (Jandel Scientific1992) was employed to fit data to several transition equations (logistic, sigmoid, Gaussian cumulative). Explicit formulae and sample calculations can be found in Neter and Wasserman (1974), McClave and Benson (1988), Wilkinson (1990), and Jandel Scientific (1992).

Results and Discussion

Total Visitation

Visitation at GRCA has been increasing almost steadily and monotonically in a curvilinear fashion since 1919 (Figure 1). The depression era shows up as a minor dip in the secular trend. The first major disruption occurred between 1941 and 1946, coinciding with World War II. Albert Richmond, a historian at Northern Arizona University, revealed that due to gasoline rationing at that time, the park was almost deserted of visitors except for U.S. Army personnel. This was because the U.S. Army had established a formal R&R camp at the park (NPS 1945).

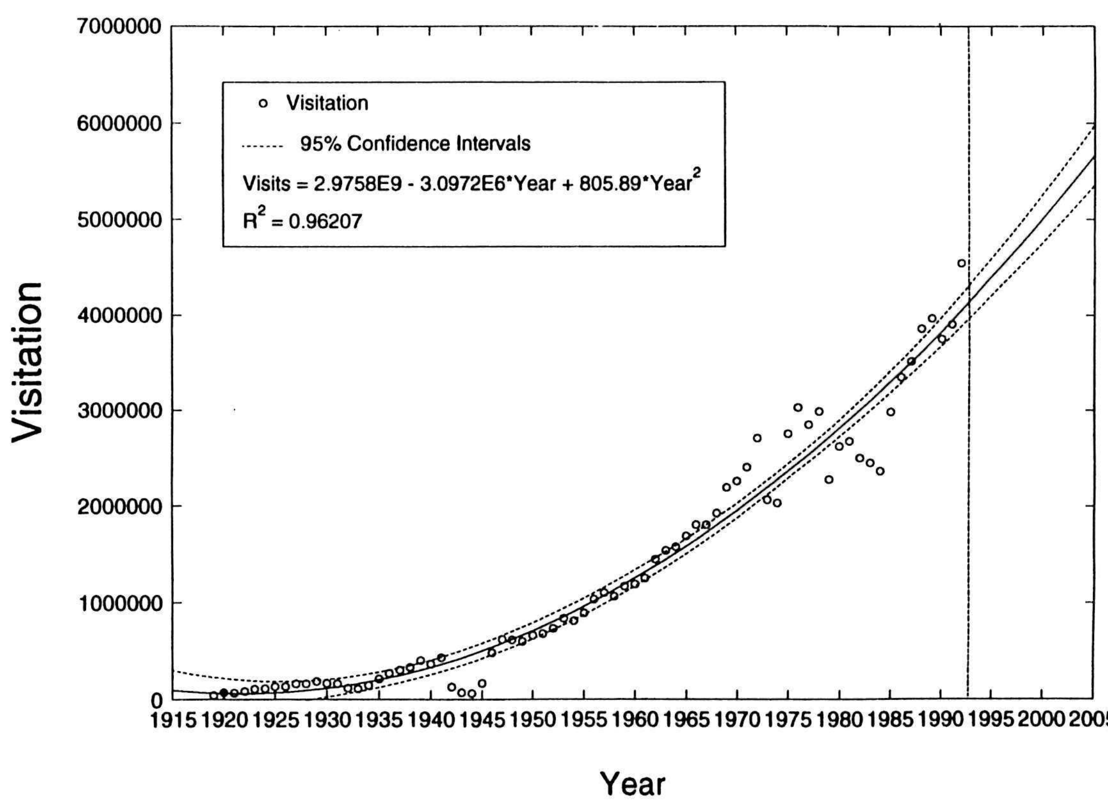


Figure 1. Grand Canyon National Park, total visitation since 1919 with forecasts to 2005. Polynomial regression fit.

park (NPS 1945). The Santa Fe Railway cancelled passenger train trips from Williams to the Grand Canyon during the war years (Richmond 1989). Both the depression era and WW II dips are more pronounced in the Log_e transformed data (Figure 2). The next major disruption coincides with the 1973 Arab oil embargo when visitation dipped from almost 2.8 million to 2 million. There was a quick recovery to just over 3 million visitors followed by an unexplained drop to 2.3 million by 1979 and ending at just under 2.5 million in 1984. This could perhaps be explained by the recession in the U.S. economy during the late Carter administration and the first years of the Reagan presidency. The current rising trend dates from 1985.

Using regression analysis, a second order polynomial (Table 1) was revealed as one of the two best fit curves to these data ($R^2 = 0.962$). Forecasting 13 years into the future to 2005 reveals that estimated visitation ($Y\text{-est}_{2005}$) would approach 5.7 million with 95% confidence intervals of about 5.4 and 6.0 million, respectively (Table 2, Figure 1). Of course, this assumes that all relevant existing social and economic conditions remain unchanged. Transforming the Y data by taking the Log_e of the visitation diminishes R^2 and the closeness of fit (Figures 1 and 2). This exponential model provides the most liberal estimate of $Y\text{-est}_{2005}$, over 13 million visitors by 2005. The linear model is the most conservative at $Y\text{-est}_{2005}$ = 4.1 million but is unrealistic in view of the overall shape of the visitation curve. Grand Canyon had already surpassed 4.5 million visitors by 1992. Along with the second order polynomial, a "sigmoid-curve" like complementary error function model produced the best fit (Figure 3) with an R^2 of 0.961 (Table 1). This model is a Gaussian Cumulative Response curve (GCR) and is part of the same family of equations as the logistic model presented in Gallipeau (1992) presumably based on the same visitation data. However, it uses a longer time series (1919-1992 instead of 1955-1991) and predicts lower visitation by 2005 (4.9 million vs. about 6.3 million) than Gallipeau's model. Although the GCR model fits the greater portion of the data (1919-1982) quite well, it reaches an inflection point at or around 1982 and begins to level off. The park

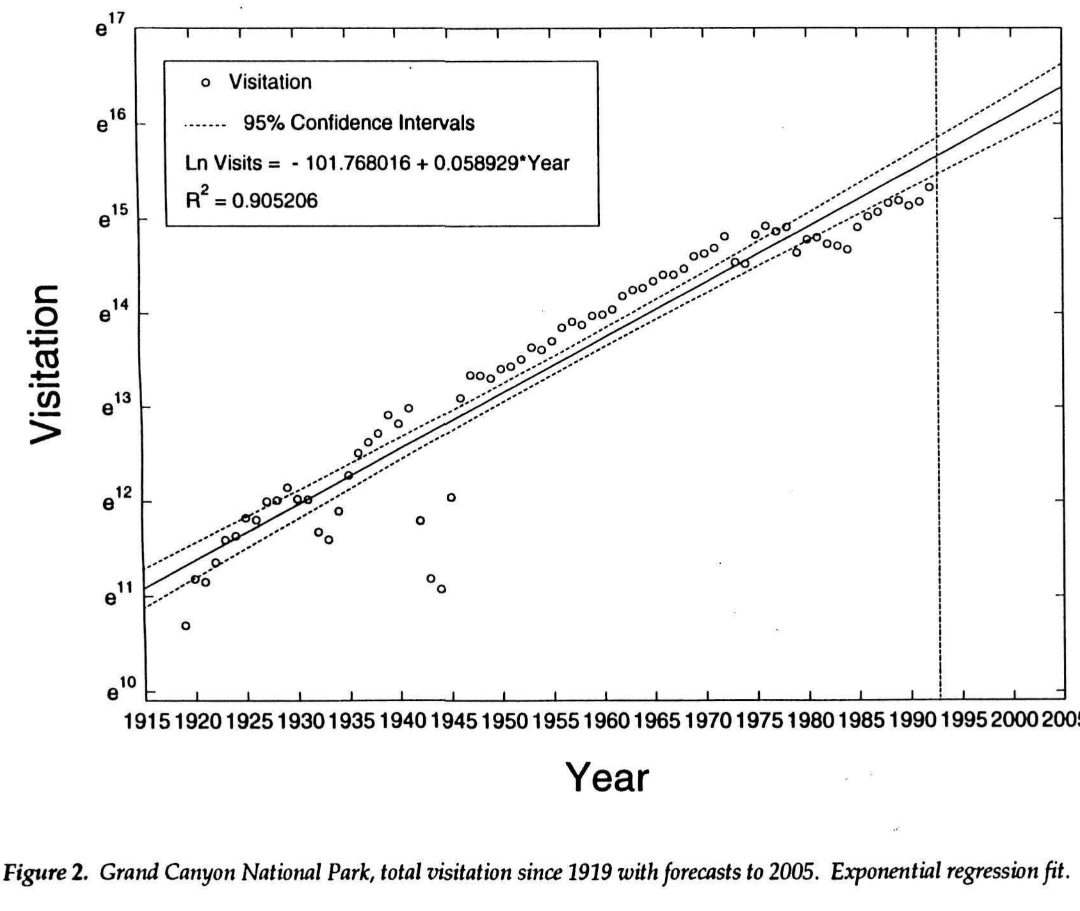


Figure 2. Grand Canyon National Park, total visitation since 1919 with forecasts to 2005. Exponential regression fit.

Table 1. Regression statistics associated with several different models relating changes in Grand Canyon Visitation over time. Total park visitation figures are from 1919 - 1992. North and South Rim figures are from 1955-1992. R^2 is the coefficient of determination for the regression. *F* is Fishers F statistic along with its probability level, *p*. $S.E._{est}$ is the standard error of the estimate. An asterisk after the Durban - Watson Statistic (*d*) indicates that there is a significant level of autocorrelation within the time series.

Data Set - Regression Model	N	Regression Model Equation	R^2	*F*	*p*	$S.E._{est}$	*d*
Total Park - Polynomial, 2 ord.	74	$Y = 2.975816E9 - 3.097175E6*X + 805.8876*X^2$	0.962	900.433	<0.00001	245,841.578	0.717*
Total Park - Gaussian Cumul.	74	$Y = -48635+6261111(1+erf((X-1982.5)/(2^{0.5}*27.5)))$	0.961	577.554	<0.00001	250,516.071	NA
Total Park - Linear	74	$Y = -1.055155E8 + 54561.638401*X$	0.891	590.767	<0.00001	413,154.475	0.262*
Total Park - Exponential	74	$\ln Y = -101.768016 + 0.058929*X$	0.905	687.541	<0.00001	0.42948	0.417*
North Rim - Polynomial, 2 ord.	38	$Y = -6.989322E6 - 583.768591*X + 2.160667*X^2$	0.571	23.309	<0.00001	78,652.995	0.688*
North Rim - Linear	38	$Y = -1.540422E7 + 7944.382865*X$	0.571	47.950	<0.00001	77,553.265	0.688*
North Rim - Exponential	38	$\ln Y = -47.129028 + 0.030184*X$	0.690	80.143	<0.00001	0.22792	0.671*
South Rim - Polynomial, 2 ord.	38	$Y = 1.73158E9 - 1.824682E6*X + 480.51313*X^2$	0.895	148.812	<0.00001	282,242.563	0.772*
South Rim - Linear	38	$Y = -1.398146E8 + 71903.032717*X$	0.891	294.143	<0.00001	283,401.587	0.751*
South Rim - Exponential	38	$\ln Y = -59.447572 + 0.037451*X$	0.899	320.937	<0.00001	0.141317	0.594*

* Data not available

Table 2. Summary of visitation forecasts ($Y\text{-}est_{2005}$ for the Year 2005 and their confidence limits (if applicable) produced by a number of regression models (Table 1) and Holt's exponential smoothing.

Data Set - Forecasting Model	95% Confidence Intervals [1]		
	$Y\text{-}est_{2005}$	Lower	Upper
Total Park - Polynomial, 2 ord.	5,668,008	5,356,727	5,979,288
Total Park - Gaussian Cumulative	4,917,036	4.06E+6 [2]	5.78E+6
Total Park - Linear	4,061,035	3,818,914	4,303,155
Total Park - Exponential	13,054,311	10,248,368	16,628,506
Total Park - Holt's Exponential Smoothing	7,243,467	NA	NA
North Rim - Polynomial, 2 ord.	526,156	304,790	747,522
North Rim - Linear	524,268	446,664	601,872
North Rim - Exponential	653,366	520,124	820,739
North Rim - Holt's Exponential Smoothing	746,983	NA	NA
South Rim - Polynomial, 2 ord.	4,770,019	3,975,375	5,564,662
South Rim - Linear	4,350,981	4,067,394	4,634,567
South Rim - Exponential	6,210,053	5,391,158	7,153,336
South Rim - Holt's Exponential Smoothing	5,866,756	NA	NA

[1] Note that confidence limits as well as the Y-estimate for exponential models have been exponentiated (e.g., $e^{Y\text{-}est}$) to maintain an equivalent scale of reference.

[2] Upper and lower 95% confidence limits estimated from graphical output, actual values were not generated by the program.

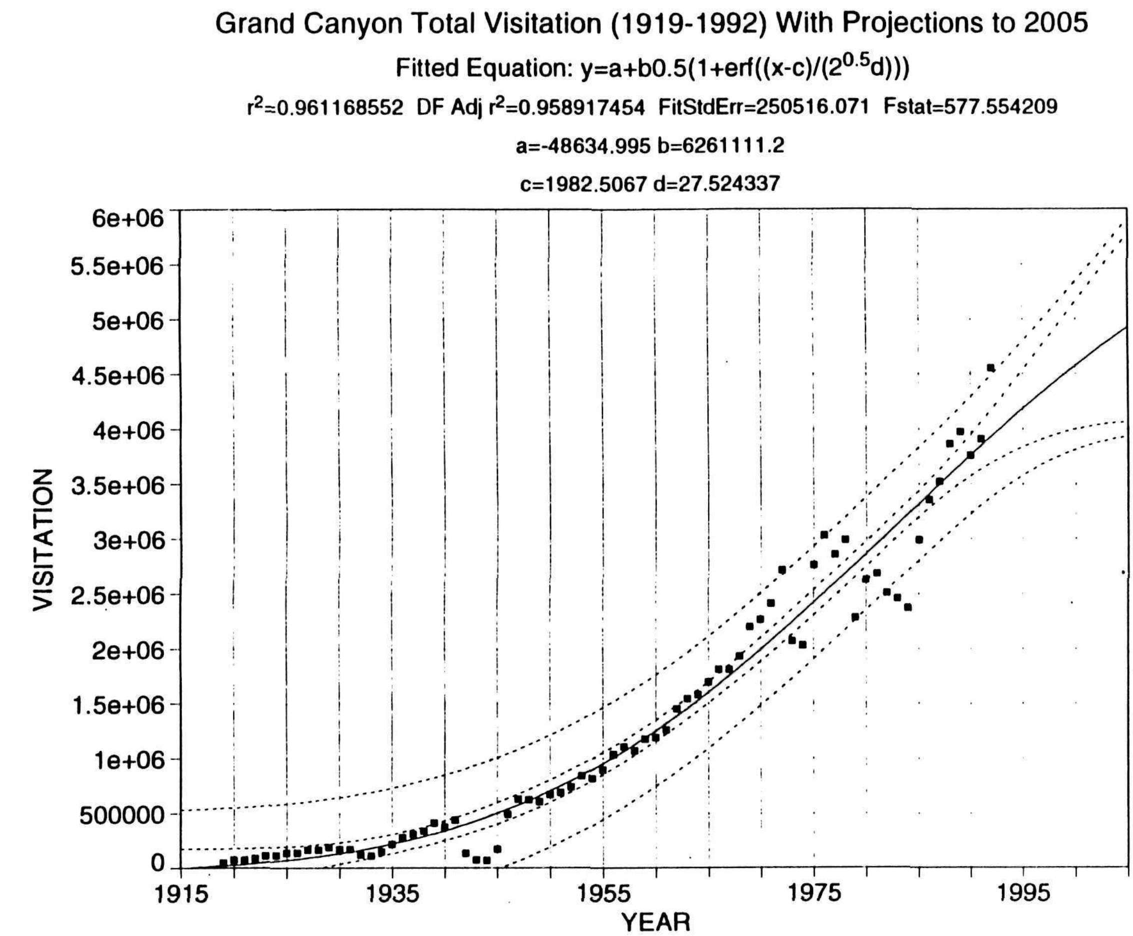


Figure 3. Grand Canyon National Park, total visitation since 1919 with forecasts to 2005. The fitted model is a Gaussian cumulative response curve. The two pairs of dotted lines approximately parellelling the estimate (solid line) are the upper and lower 95% confidence intervals, respectively (Jandel Scientific 1992).

expects to surpass 5 million visitors in 1993, considerably earlier than the model predicts. On the other hand, the upper 95% prediction limit approaches the 6.3 million figure predicted by Gallipeau (1992) who does not indicate his lower confidence or prediction limits (only the upper). On examination of the upper limit given by Gallipeau, it is clear that some overlap between the two models (upper limit of this study vs. lower limit of Gallipeau's) occurs. The significance of the overlap, unfortunately, cannot be ascertained.

Holt's model (Figure 4), which predicts that visitation will exceed 7 million visitors by 2005, provides intermediate forecast values, which are closer to those provided by the polynomial model. Unfortunately, error estimation from Holt's and other models based on exponential smoothing is limited to the original series itself. Confidence limits cannot be assessed for the forecasts derived from exponential smoothing. Nevertheless, a reasonable management strategy for GRCA should take into account the probability that approximately 6-7 million visitors park-wide can be expected by 2005, a 2.5-3% growth rate.

North Rim Visitation

The 1956-1992 North Rim data (including Tuweep, Grand Wash, and Lees Ferry) are more "noisy" (e.g., have greater variations) than the total visitation data set. However, they were treated in a similar manner. The best fit of the regression models is provided by an exponential one ($R^2 = 0.690$) (Figure 5). By the year 2005, 653,000 visitors are predicted. The linear and polynomial models fit the data equally well with R^2 values of 0.571 in both cases (Tables 1 and 2). The latter two models give more conservative forecasts of approximately 524,000 and 526,000 visitors respectively by the year 2005. As might be

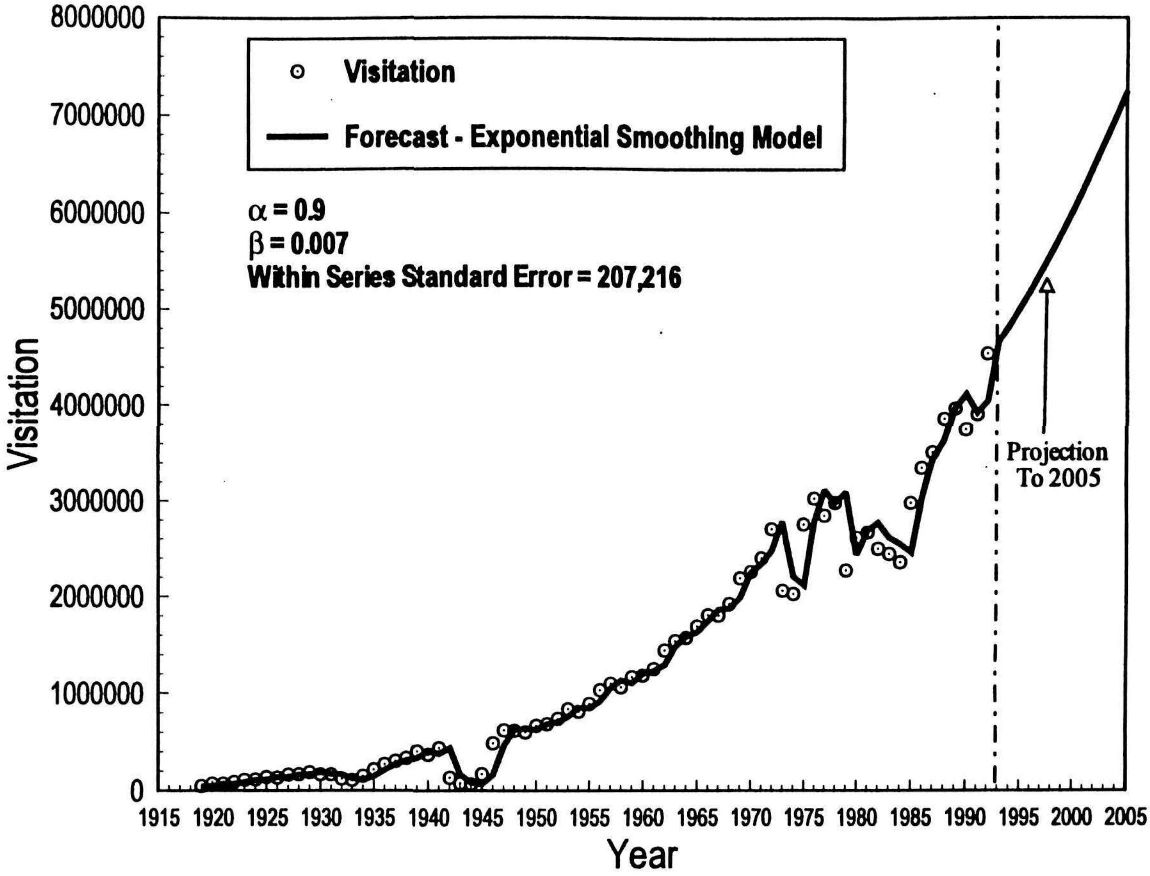


Figure 4. Forecasts of Grand Canyon National Park visitation to 2005. Holts's method of exponential smoothing.

expected, the 95% confidence band is narrower for the linear model than either the polynomial or the exponential one (Table 2). Results from the application of Holt's model were more liberal. Fitting to transition equations did not yield reasonable results. All things remaining equal, approximately 740,000 visitors are predicted to descend on the North Rim of the Grand Canyon by 2005 (Figure 6).

South Rim Visitation

South Rim visitation is increasing at a faster rate than North Rim visitation over the same time period: 1956-1992 (Figure 6). Like the North Rim, South Rim visitation since 1955 is best fit by an exponential model with R^2 = 0.899 (Figure 7). Polynomial and linear models are almost equivalent in terms of goodness of fit as defined by R^2 values of 0.895 and 0.891, respectively. It seems that growth in visitation since 1955 has been closer to exponential even though overall growth since 1919 approaches that defined by a second order polynomial. Note that the linear model for the South Rim data alone predicts a higher level of visitation for the year 2005 than the same model for total park visitation since 1919. Once again, the linear model gives the more conservative estimate of visitation (4.4 million), as well as the narrower confidence bands, for the year 2005 whereas the polynomial and exponential models forecast 4.8 and 6.2 million, respectively. Holt's exponential smoothing model predicts 5.9 million visitors by 2005 (Figure 7), an intermediate value, but one which is closer to the exponential regression model than the others (Table 2). As for the North Rim data set, fitting to transition equations was not fruitful.

Seasonal Multiplicative Effects — Winter's Model and ARIMA

Forecasts for 24 months, resulting from application of Winter's model to 1980-1992 GRCA seasonal data as well as level, trend, and multiplicative seasonal smoothing values (α, β, γ), are presented in Figure 8. Results from ARIMA are shown in Figure 9. Forecasts from Winter's model are also found in Table 3 where they are compared with results from an

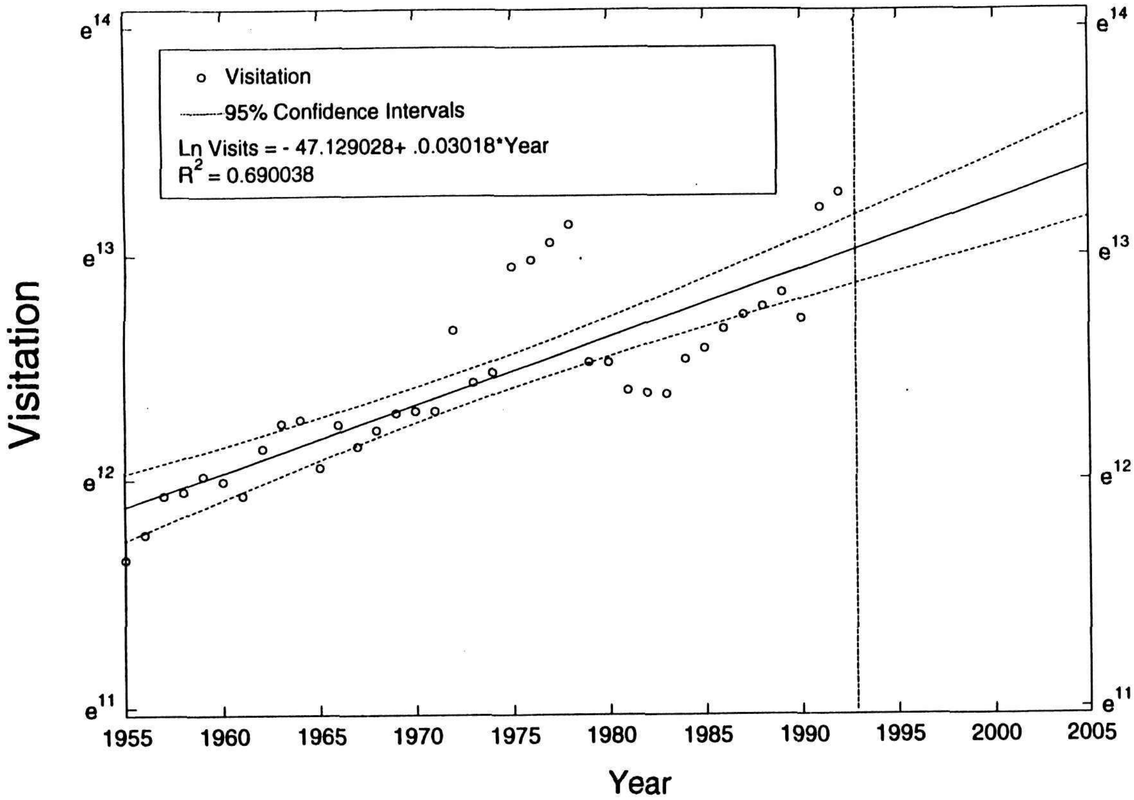


Figure 5. North Rim of Grand Canyon visitation (1955-1992). Exponential regression fit with forecasts to 2005.

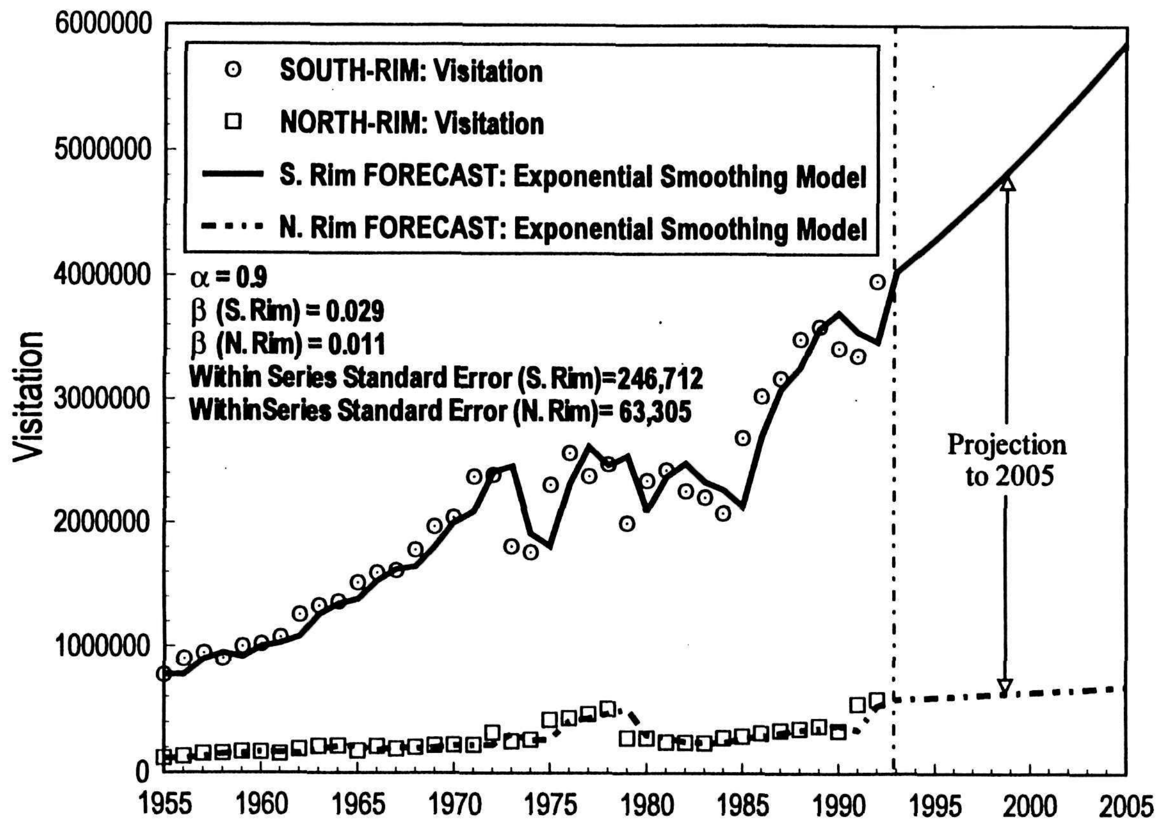


Figure 6. South and North Rim visitation forecasts using Holt's method of exponential smoothing.

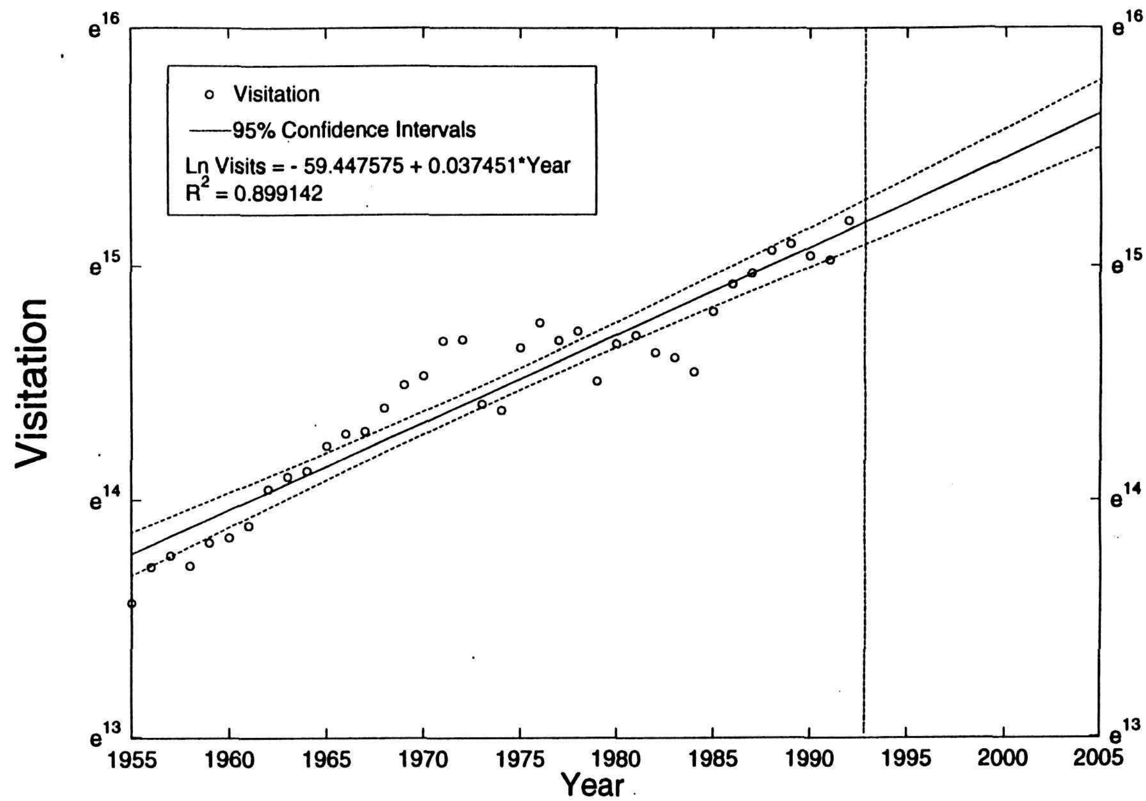


Figure 7. South Rim of Grand Canyon visitation (1955-1992). Exponential regresstion fit with forecasts to 2005.

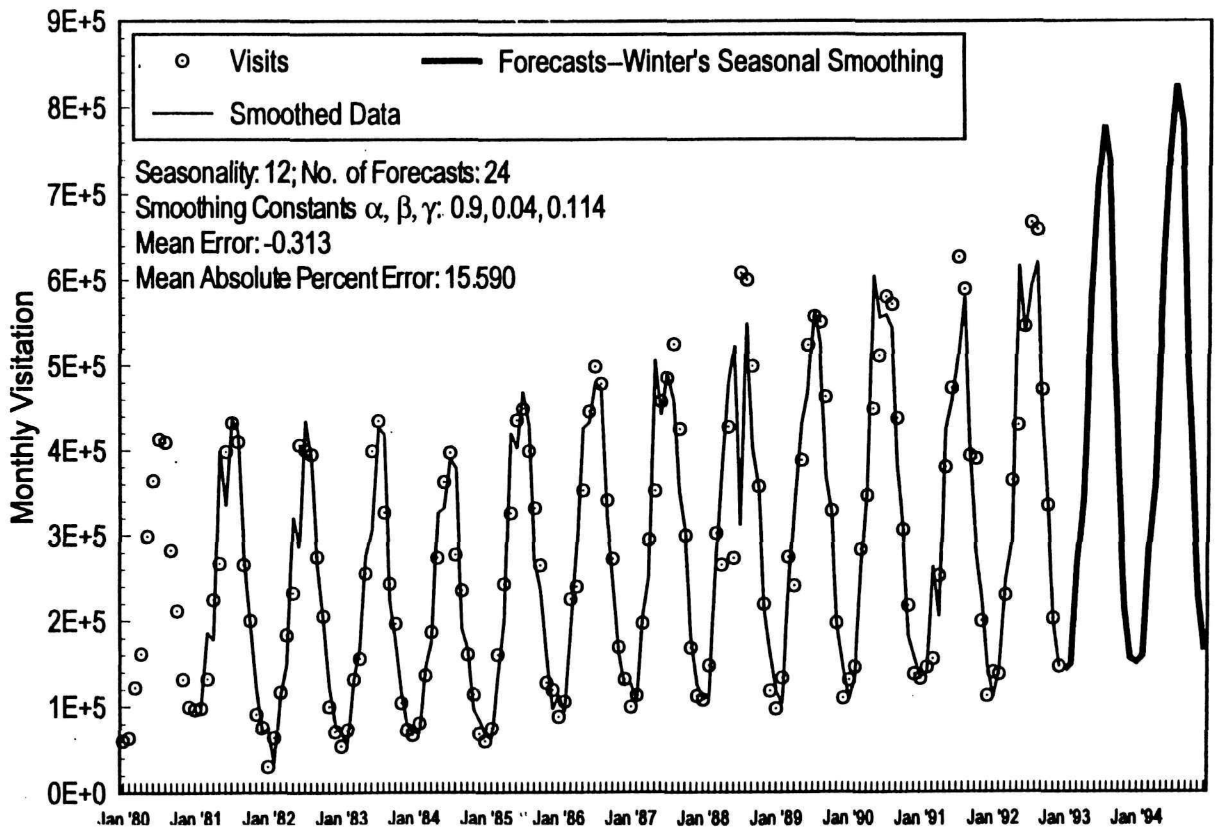


Figure 8. Seasonal variation in total visitation at Grand Canyon National Park smoothed by Winter's method with forecasts through December, 1994.

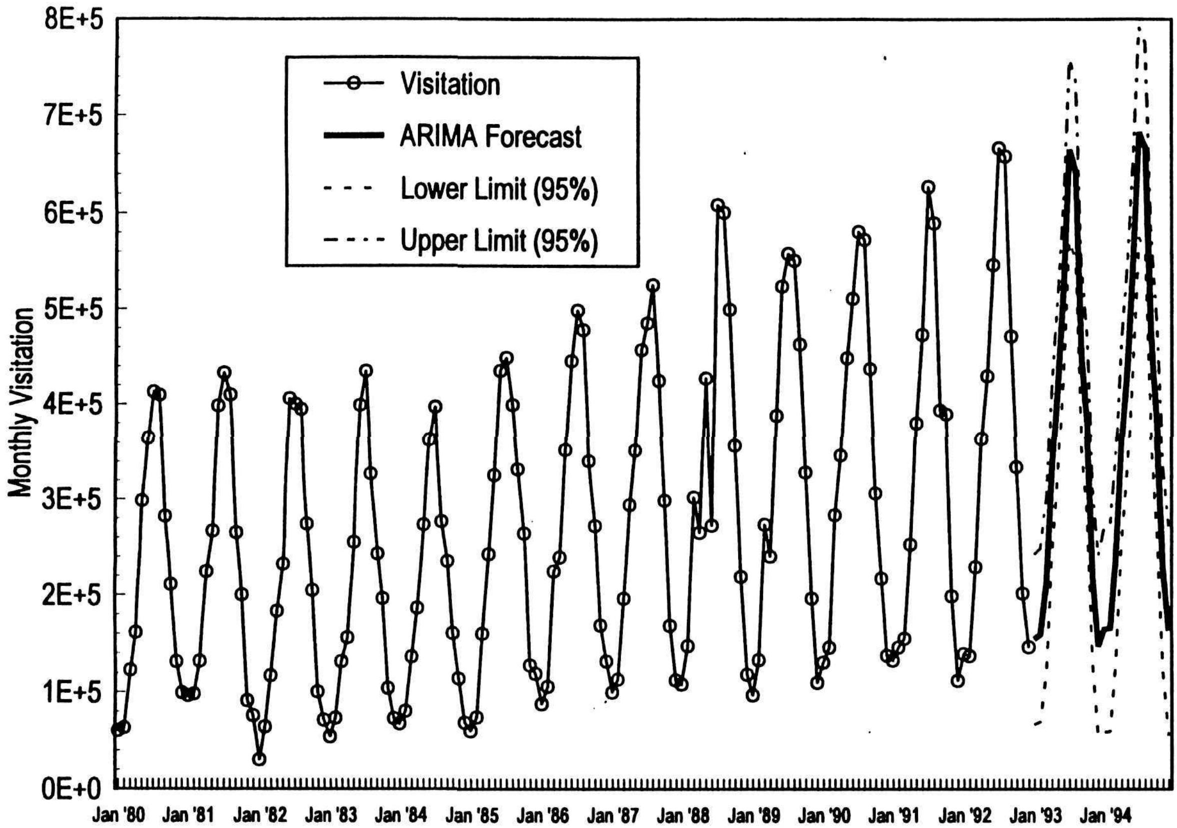


Figure 9. Seasonal variation in total visitation at Grand Canyon National Park smoothed by Autoregressive Integrated Moving Averaging (ARIMA) with forecasts through December, 1994.

ARIMA model whose summary is found in Table 4. All of the ARIMA parameters are significant to at least the $p<0.05$ level. Note that only the ARIMA model is capable of generating confidence intervals since it alone is based on an autoregression. Forecasts from both the Winter's and ARIMA models are similar within an order of magnitude. The former fall within the 95% confidence intervals of the latter except during the late spring and summer months (May-August). Winter's model is more liberal in its forecasts than ARIMA for this period of time.

In comparison with actual visitation obtained for the first 7 months of 1993, both models overestimated January and February visitation, but the actual visitation was within the 95% confidence limits given by the ARIMA. Both June and July actual visitation was outside the upper 95% confidence limits for the ARIMA model but was within an order of magnitude of the forecasts generated by the Winter's model (exponential smoothing does not provide confidence intervals). The latter overestimated June's visitation by 56,000 (an 8.5% difference) and underestimated July's by 32,000 (a -3.9% difference). It would appear that the simpler Winter's model gives more reasonable projections based on my initial comparisons.

There is a very strong seasonality in the past 12 years of GRCA visitation, with a superimposed increasing trend. January and December receive the lowest visitation, July and August the highest. Visitation during all months of the year is generally increasing. Although more visitors are impinging on park resources during non-summer months (off-season), visitation is also increasing during the summer. Overall, the past pattern of strong seasonality has been persistent and is forecasted into the immediate future. There appears to be no development of a "shoulder-season" effect, i.e., an increasing relative density of visitors during the spring and fall months immediately adjacent to the summer season which would tend to broaden out the visitation peaks. The implication for park management is that relative allocation of resources toward visitor management should remain geared toward sharp summer peaks and winter troughs into the immediate future.

Table 3. Forecast results from the application of Winter's Exponential Smoothing Model and ARIMA to 1980-1992 seasonal visitation data for Grand Canyon National Park.

Year and Month		Forecasts: Winter's Model	Forecasts: ARIMA Forecast	95% Confidence Interval Lower	95% Confidence Interval Upper	Actual
1993	January	142,114	154,737	66,463	243,012	99,171
	February	148,836	159,493	69,586	249,399	106,536
	March	267,826	213,037	121,975	304,099	293,719
	April	340,333	330,121	238,237	422,005	353,987
	May	582,543	424,329	331,857	516,800	500,258
	June	708,893	529,838	436,946	622,729	653,211
	July	779,081	665,946	572,753	759,138	810,875
	August	736,283	644,026	550,618	737,434	760,318
	September	499,697	452,574	359,011	546,137	—
	October	359,055	377,681	284,007	471,355	—
	November	213,431	218,054	124,300	311,808	—
	December	155,469	147,344	53,532	241,155	—
1994	January	150,902	164,642	58,419	270,865	—
	February	157,993	166,097	59,259	272,934	—
	March	284,219	237,839	130,561	345,117	—
	April	361,058	363,084	255,490	470,677	—
	May	617,839	443,829	336,009	551,649	—
	June	751,629	554,473	446,491	662,456	—
	July	825,813	683,422	575,323	791,522	—
	August	780,228	667,914	559,730	776,097	—
	September	529,374	478,385	370,141	586,628	—
	October	380,275	374,521	266,234	482,808	—
	November	225,983	227,579	119,261	335,897	—
	December	164,568	163,708	55,367	272,048	—

Table 4. Summary of fitted ARIMA model for Grand Canyon National Park seasonal visitation data (1980-1992).

Parameter[1]	Estimate	Standard Error	t	P
Non-Seasonal Autoregression (1)	0.84810	0.11852	7.15607	0.00000
Seasonal Autoregression (12)	-0.46710	0.07968	-5.86218	0.00000
Non-Seasonal MovingAverage(1)	0.65492	0.16368	4.00132	0.00010
Mean	11465.46244	5517.93005	2.07786	0.03955
Constant	2555.04684			

[1] Model fitted to seasonal differences of order 1 with seasonal length = 12. Estimated white noise variance = 1.99268E9 with 140 degrees of freedom. Estimated white noise standard error = 44639.4. Chi-square test statistic on first 20 residual autocorrelations = 12.9207 with a probability of a larger value given white noise of 0.741477. No backcasting was done.

Comparisons With Marketing Analysis Forecasts

Gallipeau (1992) does not evaluate his data seasonally, so comparisons are not possible in this regard. However, his total annual visitation projection for the year 2005 (period 1955-1992) reveals a visitation of approximately 6.3 million (as read from graphs on pages 31 and 32 of his report). Approximately one-third of the visitors are expected to be foreign. Foreign visitation (not considered in this report) is shown to increase steadily over the period considered and at a rate that exceeds domestic. The 6.3 million figure falls between time series forecasts based on a second order polynomial and Holt's exponential smoothing (Table 2) but is reasonably within an order of magnitude. Gallipeau (1992) expects that the North Rim of the Grand Canyon will attract proportionately fewer foreign visitors than the South Rim (he does not report exactly why except perhaps that the North Rim is less accessible and less publicized). Since foreign visitors are increasing faster than domestic, he concludes that the North Rim proper will have 508,000 visitors in 2010 and lose about 1% of the park's total visitation relative to the South Rim.

Conclusions

There is no doubt that visitation to GRCA is increasing. The important questions are the following: What form does the secular trend take? What is the rate of increase? Are patterns of visitation changing, especially on a seasonal basis? Finally, how should management take advantage of this forecasting tool?

The secular trend in total park visitation since 1919 approaches a second order polynomial according to best fit regression techniques. But since 1955 the curve has been closer to exponential in growth with an unexplained, non-seasonal cyclical variation present in the long-term data. The overall growth rate of visitor numbers has been close to 2.5-3% per year. If economic and social conditions remain relatively unchanged over the next 13 years and the Park Service does not impose restrictions on visitation, reasonable forecasts to 2005 show that GRCA may have to deal with total visitation on the order of 7 million persons per year. Approximately 6.0-6.3 million of these will impinge on the more developed South Rim, the remainder on the more undeveloped North Rim. My estimates for North Rim visitation are more liberal than Gallipeau's, but in my case, the North Rim data include contributions from Lees Ferry, Grand Wash, and Tuweep. Allowing for this difference, time series analysis and market analysis predictions appear to be comparable.

Visitation from 1980 to 1992 has been strongly seasonal with summer peaks, more like spikes, and winter troughs. There is no evidence based on these past patterns that relative changes in seasonal patterns, i.e., the growing relative importance of a shoulder season, is occurring. Problems exist in the data set due to GRCA's changing of the persons per vehicle multiplier. Past visitation records may have been inflated in comparison to the last two years'.

Assuming that resources are available, a market analysis is probably the method of choice in forecasting park visitation. In part, this is because such an analysis can reveal the nature of the visitation in addition to providing projections of numbers of visitors. Such an analysis can provide valuable insights into providing for foreign visitors in terms of multilingual publications, bilingual or multilingual interpretive staff, and other visitor needs, both foreign and domestic. However, in the absence of extensive marketing data, or where breakdown into markets is not relevant, time series analysis, used judiciously, is a viable alternative. In terms of resources preservation and protection, a knowledge of potential markets may be somewhat superfluous; it makes little difference to park resources whether human impacts are due to foreign or domestic visitors. Market analysis results which are used to develop ways of informing different market groups about park purposes and values is an exception. An example is the enhanced development and funding of foreign language brochures which interpret the Park Service's conservation mission in response to predictions of greater proportions of foreign tourists. This kind of management response to a market analysis is more desirable, in terms of addressing conservation of park purposes and values, than using these data to plan additional visitor services infrastructure.

Management Recommendations

(1) The National Park Service should consider setting a "carrying capacity" for visitation to the Grand Canyon, especially the South Rim. Even if park visitation begins to level off just after the year 2000 (Gallipeau 1992), almost 7 million people are expected to enter the park and impact physical, biological, and infrastructural resources by the year 2010 and possibly as early as 2005. Such a carrying capacity should ideally be determined by objective, interdisciplinary research based on the sociology of crowding (e.g., Shelby 1976); the economics and cost/benefit of providing visitor services, as well as infrastructural development and maintenance; and, most importantly, the resilience of the biological, cultural, and physical resources of the Grand Canyon.

(2) The park should decide on a statistically sound and acceptable method of determining visitation multipliers, preferably through the effort of an objective third party, as a means of estimating visitors from vehicle counts. If park management continues to change its estimating methods over a period of time, the long-term integrity and usefulness of the data set as a forecasting and monitoring device will be compromised.

(3) On a park-wide basis, park management should continue, at least in the short term, to allocate resources based on strong seasonality of visitation as no increasing relative importance of "shoulder seasons" visitation can be identified and corroborated.

(4) Due to the increasing relative importance of foreign visitation reported by Gallipeau (1992), the park should plan to allocate additional resources for the preparation and production of information transfer devices. Such devices (e.g., pamphlets, interpretive displays, roadside exhibits) should be available in the principal foreign languages. These materials should also be increasingly oriented toward resources management and protection, as well as explaining the mission of the National Park Service with park purposes and values.

Literature Cited

Albrecht, D. E. 1991. A study of the perceptions, expectations and satisfaction levels of visitors to Grand Canyon National Park. College Station, Texas: Texas A&M University, Department of Rural Sociology and Recreation, Parks and Tourism Sciences.

Box, G. E. P. and G. M. Jenkins. 1977. Time series analysis: forecasting and control, 2nd ed. San Francisco: Holden-Day.

Gallipeau. A. 1992. Grand Canyon National Park market analysis and forecast. Unpublished report to Grand Canyon National Park. National Park Service, Socio-Economic Studies Division, WASO/TNT.

Jandel Scientific. 1992. TableCurve, automated curve-fitting software. San Rafael, California: Jandel Corporation.

McClave, J. T. and P. G. Benson. 1988. Statistics for business and economics. San Francisco: Dellen Publishing Co.

National Park Service. 1945. Superintendent's Monthly Report: August 1, 1945. Grand Canyon National Park Library.

National Park Service. 1980. General Management Plan: Yosemite National Park.

Neter, J. and W. Wasserman. 1974. Applied linear statistical models. Homewood, Illinois: Irwin.

Richmond, A. 1989. Cowboys, miners, presidents and kings: the story of the Grand Canyon Railway. Flagstaff, Arizona: Press of Northland Printing.

Shelby, B. and J. M. Nielsen. 1976. Use levels and crowding in the Grand Canyon. Colorado River Research Technical Report No. 3. Grand Canyon National Park, Arizona.

Stankey, G. H., D. N. Cole, R. C. Lucas, M. E. Peterson, and S. S. Frissell. 1985. The limits of acceptable change (LAC) for wilderness planning. USDA Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report INT-176. 37 pp.

Statistical Graphics Corporation. 1988. Statgraphics, statistical graphics system Version 4.0. Rockville, Maryland: STSC, Inc.

Wilkinson, L. 1990. Systat: The system for statistics. Evanston, Illinois.

Zar, J. H. 1984. Biostatistical analysis, 2nd ed. Engelwood Cliffs, New Jersey: Prentice-Hall.

As the nation's principal conservation agency, the Department of the Interior has responsibility for most of our nationally owned public lands and natural and cultural resources. This includes fostering of our land and water resources, protecting our fish and wildlife, preserving the environmental and cultural values of our national parks and historical places, and providing for enjoyment of life through outdoor recreation. The department assesses our energy and mineral resources and works to ensure that their development is in the best interests of all our people. The department also promotes the goals of the Take Pride in America campaign by encouraging stewardship and citizen responsibility for the public lands and promoting citizen participation in their care. The department also has a major responsibility for American Indian reservation communities and for people who live in island territories under U.S. administration.

NPS D-272

October 1993