# Smoothing Techniques Versus Regression Models Applied to Daily Hospital Patient Forecasting

**Nader Vargha**
Associate Professor of Economics
Department of Business and Economics
Missouri Western State College
St. Joseph, MO 64507
Tel: 816-271-4339
Fax: 816-271-5871
E-mail: vargha@mwsc.edu

Paper presented at the 38th Annual
Missouri Valley Economic Association
February 28 - March 2, 2002
Memphis, Tennessee
“Smoothing Techniques Versus
Regression Models Applied to Daily
Hospital Patient Forecasting"

Multiple regression models have become the prime method of choice in the area of forecasting. Economists, especially in the academic arena, in many instances give no consideration to any other method but regression analysis in order to examine cause and effect relationships and to use such relationships to forecast future outcomes. Regression models appear to have become the ‘Cadillac’ of forecasting methods.

Undoubtedly, regression analysis is a very powerful tool of forecasting. However, there is another question of interest, which forms the basis for this study: could other methods, such as smoothing techniques, be used with equal effectiveness in place of a regression analysis? Could a smoothing technique produce a forecast of greater or, at least equal, accuracy? This question takes on added importance in cases where the effectiveness of regression analysis is questionable due to unavailability of data for certain appropriate explanatory variables, or where time appears to be the most important independent variable.

One such case involves forecasting daily patient census for a hospital, which on the surface seems to be an excellent candidate for an econometric modeling and forecasting by means of multiple regression. In the regression model for determining patient census, explanatory variables, such as a hospital ward, day of the week and day of the month, are chosen and used to determine patient census. Thus, it is assumed that patient census is a function of these and other quantitative variables. However, there are other possible explanatory variables that are qualitative in nature and cannot be effectively included in a regression equation, not even via the use of dummy variables. One such variable is the training background of the physician who admits patients to a hospital. Some medical schools presumably teach a more conservative approach to patient care which results in longer hospital stay for the patient. Some other medical schools may train their students to rely on a minimum hospital stay for their patients. Such differences in training, for example, have a direct and significant effect on daily patient census, yet there is no effective method of quantifying such a variable for inclusion in a multiple regression model.

The objective of this study is to produce forecasts of daily hospital patient census using a smoothing technique, and to compare these forecasts to forecasts generated by means of multiple regression model, which was estimated in a previous study (Vargha, 2000), to determine the viability of a smoothing technique relative to regression analysis for such a case. For this purpose, the total patient census is first disaggregated into individual wards and analyzed for presence of trend and seasonal variations. Next, each data set containing seasonal variations is deseasonalized to obtain data with possibly trend variations only. Subsequently, two smoothing techniques, both capable of dealing with trend variations in the data, are employed to produce short-term forecasts of daily patient census. The results of these forecasts are then compared to the forecasts obtained via multiple regression analysis using various error measurements to determine the viability of such forecasting methods in producing accurate forecasts for this particular event.

## Literature Review

In recent times, economic forces have created increasingly tough and pervasive incentive for hospitals to reduce their costs. Nursing staff compensation constitutes, by most estimates, over 50 percent of total operating costs for hospitals. Thus, any cost reduction schemes merit an examination of nursing staff. Generally, hospitals would seem to have two choices: (1) cut back the nursing staff, or (2) improve the efficiency of scheduling the nursing staff. A method that allows the hospital administrator to project, with a high degree of accuracy, the number of patients in a ward at a given time will result in a more efficient scheduling of the nursing staff and thus help lower costs. Similarly, in the long run, cost saving can result from planning based on more accurate projections.

Empirical work in this area is rather scarce. Helmer, Oppermann, and Suver (1980) developed a model for forecasting the number of patients by level of care. For a 220-bed short-term acute-care hospital, they developed six different intensity-of-care level categories. Next, they used regression analysis to develop a model capable of predicting the number of patients by care level for each hospital ward by month, day, and shift.

The number of patients for a given care level was regressed with different wards, months of the year, days of the week, shifts of the day, and day of the year. After running numerous regressions using over 9,600 observations, Helmer et al. came up with seven models, four for individual care levels and three for combinations of care levels. They report very satisfactory results at a significance level of 0.05 ($\alpha = 0.05$).

In a related study, Wilson and Schuiling (1992) attempted to identify variables that influence hospital laboratory workload in order to develop workload forecasts for such laboratories. They identified four independent variables—inpatient admissions, acuity days, length of stay, and discharge days—as being important determinants of billable procedures in a laboratory. Then they regressed billable procedures (the dependent variable) on the four independent variables. To overcome potential effects of serial correlation, the Cochrane-Orcutt procedure is employed.

Wilson and Schuiling tested the resulting regression model for a one-year period and found the annual percent forecast error to be 0.7 percent. They also attempted to forecast laboratory workload using an exponential smoothing technique called Winter's Exponential Smoothing model. Tests showed this model was not quite as accurate as the regression model in predicting. For the same one-year time period laboratory workload was forecasted using a subjective forecasting method (a method based on experience and history). The annual percent forecast error for the subjective forecast was 11.6 percent, much higher than the error for the regression model.

In a recent study, the author of this study (Vargha, 2000), developed a regression model for a medium-sized Mid-western hospital. Daily patient census data was collected from seven different wards for a period of 364 days. Assuming daily patient census is a function of day of the week, day of the month, day of the year, and the ward of the hospital, a regression model was constructed and estimated. The Durbin-Watson test of this model indicated existence of serial correlation, which is not uncommon in business and economic data. Consequently, the remedy of taking the lag of the dependent variable and adding it as another explanatory variable was chosen, and the model was re-estimated.

An examination of the results for the revised regression showed the following: the F-test again indicated there is a significant relationship between total daily census and the ward, time, day, and month variables. The two models were tested for relative accuracy by using each regression equation to generated forecasted values for each of the observed census values. Finally the MSE (mean squared error) and RMSE (root mean squared error) were determined for each model. The revised model, adjusted for serial correlation proved to generate a more accurate forecast.

The accuracy of forecasts generated with regression models, such as the one described above, could potentially be enhanced, if additional explanatory variables can be identified and included in the estimation of the model. This is not always possible. There are possible explanatory variables that are qualitative in nature and often cannot be effectively included in a regression equation because of unavailability of necessary data.

One such variable is the training background of the physician who admits patients to a hospital. Some medical schools presumably teach a more conservative approach to patient care which results in a longer hospital stay for the patient. Other medical schools may train their students to rely on a minimum hospital stay for their patients. Such differences in training, for example, have a direct and significant effect on daily patient census. This factor could theoretically be included in a regression model by means of dummy variables. However, lack of data and the wide range of different training backgrounds make it very difficult, if not impossible, to include such explanatory variables in a multiple regression model.

In the absence of such variables, and under the assumption that daily patient census is a function of day of the week, day of the month, day of the year, and the ward of the hospital only, the following interesting question emerges: would the accuracy of the forecast improve if one assumes patient census is a function of time and wards, and thus used other techniques, such as smoothing techniques to forecast daily patient census by each ward separately? This study attempts to provide some answers to this question.

## Methodology

This study was done in three steps. In step one, the revised regression model was tested for accuracy. In step two, smoothing techniques were used to forecast daily patient census for each ward. In the final step, the results of steps one and two were compared to conclude whether or not smoothing techniques are, under certain conditions, a viable alternative to a multiple regression models.

### *1. Testing The Multiple Regression Model*

The testing method employed here is a method called an Ex Post forecast. In order to generate an ex post forecast, the time period for which observed values are available is divided into two periods: the estimation period and the ex post forecast period. The values for the estimation period are used to re-estimate an estimated model. The model is then used to forecast values for the ex post forecast period, for which observed values are available. The results thus obtained are compared to observed values to determine the accuracy of the model.

In this instance, the model adjusted for serial correlation was used twice to re-estimate another regression equation.

**TABLE I**
**Regression Analysis**
1 - WEEK EX POST FORECASTING PERIOD

| Day | NO2 - 1 week | | SO2 - 1 week | | SO3 - 1 week | | NS2 -1 week | | PCU - 1 week | | SICU - 1 week | | SSDU - 1 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-23 | 24 | 27 | 27 | 28 | 26 | 26 | 10 | 14 | 8 | 14 | 13 | 16 | 12 | 14 |
| 11-24 | 17 | 25 | 23 | 28 | 24 | 26 | 12 | 12 | 8 | 12 | 16 | 14 | 12 | 13 |
| 11-25 | 12 | 21 | 24 | 25 | 21 | 24 | 10 | 12 | 13 | 11 | 16 | 14 | 11 | 12 |
| 11-26 | 12 | 18 | 28 | 25 | 25 | 22 | 8 | 10 | 15 | 12 | 12 | 13 | 11 | 11 |
| 11-27 | 18 | 21 | 30 | 29 | 29 | 27 | 14 | 12 | 11 | 16 | 16 | 14 | 15 | 13 |
| 11-28 | 27 | 23 | 27 | 30 | 29 | 28 | 14 | 15 | 16 | 14 | 20 | 16 | 16 | 15 |
| 11-29 | 31 | 28 | 31 | 29 | 31 | 28 | 15 | 15 | 15 | 16 | 20 | 18 | 16 | 16 |
| RMSE | | 5.66 | | 2.67 | | 2.27 | | 2.04 | | 3.68 | | 2.45 | | 1.25 |
| MAPE | | 0.32 | | 0.09 | | 0.08 | | 0.15 | | 0.32 | | 0.14 | | 0.08 |

**TABLE II**
**Regression Analysis**
2 - WEEK EX POST FORECASTING PERIOD

| Day | NO2 - 2 week | | SO2 - 2 week | | SO3 - 2 week | | NS2 -2 week | | PCU - 2 week | | SICU - 2 week | | SSDU - 2 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-16 | 32 | 30 | 32 | 31 | 31 | 31 | 15 | 15 | 16 | 16 | 20 | 16 | 15 | 15 |
| 11-17 | 31 | 29 | 31 | 31 | 31 | 30 | 15 | 14 | 16 | 16 | 21 | 17 | 15 | 14 |
| 11-18 | 21 | 27 | 23 | 29 | 27 | 28 | 13 | 13 | 16 | 14 | 13 | 16 | 16 | 13 |
| 11-19 | 18 | 22 | 26 | 25 | 27 | 26 | 13 | 11 | 14 | 14 | 17 | 12 | 14 | 13 |
| 11-20 | 23 | 23 | 29 | 29 | 31 | 29 | 11 | 14 | 14 | 15 | 16 | 16 | 15 | 15 |
| 11-21 | 29 | 26 | 31 | 30 | 29 | 30 | 15 | 13 | 15 | 15 | 16 | 16 | 16 | 15 |
| 11-22 | 27 | 28 | 26 | 31 | 24 | 30 | 12 | 15 | 11 | 16 | 18 | 16 | 13 | 16 |
| 11-23 | 24 | 27 | 27 | 29 | 26 | 27 | 10 | 14 | 8 | 14 | 13 | 17 | 12 | 14 |
| 11-24 | 17 | 26 | 23 | 28 | 24 | 28 | 12 | 12 | 8 | 12 | 16 | 14 | 12 | 13 |
| 11-25 | 12 | 21 | 24 | 25 | 21 | 25 | 10 | 12 | 13 | 11 | 16 | 14 | 11 | 12 |
| 11-26 | 12 | 18 | 28 | 25 | 25 | 24 | 8 | 10 | 15 | 12 | 12 | 13 | 11 | 11 |
| 11-27 | 18 | 21 | 30 | 30 | 29 | 28 | 14 | 12 | 11 | 16 | 16 | 14 | 15 | 13 |
| 11-28 | 27 | 24 | 27 | 30 | 29 | 30 | 14 | 15 | 16 | 14 | 20 | 16 | 16 | 15 |
| 11-29 | 31 | 28 | 31 | 29 | 31 | 30 | 15 | 15 | 15 | 16 | 20 | 18 | 16 | 16 |
| RMSE | | 4.66 | | 2.88 | | 2.41 | | 2.00 | | 2.99 | | 2.92 | | 1.51 |
| MAPE | | 0.22 | | 0.08 | | 0.07 | | 0.14 | | 0.21 | | 0.15 | | 0.08 |

First, the data consisting of the first 358 of the 365 observations was used to estimate a new regression equation leaving the last seven observations for the ex post forecast period. The new equation thus estimated was then used to forecast the daily patient census for each ward for periods 359 through 365. Next, the observed values for each ward and for periods 359 to 365 were compared against the forecasted values and the mean squared errors and the mean average percentage errors were calculated. The results are reported in Table I.

**The wards were as follows:**

| | | |
|---|---|---|
| 2 North | – | Orthopedics (NO2) |
| 2 South | – | General Surgery (SO2) |
| 3 South | – | Oncology and Neurology (SO3) |
| 2 North Sub | – | Urology and Gynecology (NS2) |
| PCU | – | Progressive Care Unit (PCU) |
| SICU | – | Intensive Care Unit (SICU) |
| SSDU | – | Surgical Step-down Unit (SSDU) |

Next, the data consisting of the first 351 of the 365 observations was used to estimate a new regression equation leaving the last 14 observations for the ex post forecast period. The new equation thus estimated was then used to forecast the daily patient census for each ward for periods 352 through 365. Next, the observed values for each ward and for periods 352 to 365 were compared against the forecasted values and the mean squared errors and the mean average percentage errors were calculated. The results are reported in Table II.

*2. Forecasting Using Smoothing Techniques*

In this step, the same data was used to forecast daily patient census via two different smoothing technique methods. The smoothing techniques chosen were Brown’s Linear Exponential Smoothing technique (sometimes referred to as a Double Exponential Smoothing) and Holt’s Two Parameter Linear Exponential Smoothing. Both of these methods are capable of forecasting more than one period into the future. Both of these techniques are also capable of handling data sets that contain trend variations. Conceptually, these two techniques are similar, except that in Holt’s technique a second smoothing constant is added to deal with trend in the data. The reason for using two smoothing methods was simply the attempt to improve the accuracy achieved with the smoothing techniques.

Neither of these two methods is capable of dealing with seasonal variation in the data. Thus, an autocorrelation analysis was done on the actual data for each of the seven wards to determine presence of trend and seasonal variations in the data. The autocorrelation coefficients were tested at 95 percent confidence level to determine if they were significantly different from zero.

For each of the seven wards, the autocorrelation analysis and test indicated some trend variation. Seasonal variation was detected for wards 2 North, 2 South, 2 North Sub, Intensive Care Unit, and Surgical Step-down Unit. Consequently, it was decided to deseasonalize the data for these five wards. The ratio-to-moving-average method was employed for this process using seven periods for the moving average since the autocorrelation coefficients for the 365 days showed a ‘spike’ for the seventh, fourteenth, twenty-first, etc. periods.

The ratio of observed values to the moving averages were used to construct a seasonal index chart for seven days and fifty-two weeks. An adjustment factor was calculated and used to adjust the mean value for each of the seven days. The resulting values were divided by 100 to obtain an adjusted seasonal index for each day of the week. Finally, the actual data for each day of the week was divided by the appropriate seasonal index to obtain a series of deseasonalized data for all 365 days.

As in the estimation earlier using regression analysis, in this part the total period, for which data was available, was divided into an estimation period and an ex post forecast period. First, the observed values for the first 358 time

**TABLE III**
**Brown’s Linear Exponential Smoothing**
1 - WEEK EX POST FORECASTING PERIOD

| Day | DSNO2 - 1 week | | DSSO2 - 1 week | | SO3 - 1 week | | DSNS2 -1 week | | PCU - 1 week | | DSSICU - 1 week | | DSSSDU - 1 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-23 | 22.0 | 25 | 28 | 29 | 26 | 29 | 9 | 13 | 8 | 15 | 12 | 16 | 11 | 14 |
| 11-24 | 16.0 | 25 | 23 | 29 | 24 | 29 | 11 | 13 | 8 | 15 | 15 | 16 | 12 | 14 |
| 11-25 | 14.0 | 24 | 23 | 29 | 21 | 29 | 11 | 13 | 13 | 15 | 19 | 16 | 12 | 14 |
| 11-26 | 14.0 | 24 | 26 | 29 | 25 | 29 | 10 | 13 | 15 | 15 | 15 | 16 | 12 | 14 |
| 11-27 | 19.0 | 24 | 31 | 29 | 29 | 29 | 14 | 13 | 11 | 15 | 16 | 16 | 15 | 14 |
| 11-28 | 26.0 | 24 | 28 | 29 | 29 | 29 | 13 | 13 | 16 | 15 | 18 | 16 | 15 | 14 |
| 11-29 | 27.0 | 23 | 32 | 29 | 31 | 29 | 14 | 13 | 15 | 15 | 18 | 16 | 15 | 14 |
| RMSE | | 6.92 | | 3.70 | | 4.11 | | 2.24 | | 4.12 | | 2.24 | | 1.85 |
| MAPE | | 0.37 | | 0.13 | | 0.13 | | 0.19 | | 0.33 | | 0.13 | | 0.15 |

periods were used to generate forecasts for the last seven periods (ex post forecast period). This was done first using the Brown's method and then Holt's method. The observed values were used for wards 3 South - Oncology and Neurology (SO3) and PCU - Progressive Care Unit (PCU), since autocorrelation analysis did not indicate seasonal variations for these wards.

For the other five wards, 2 North - Orthopedics (NO2), 2 South - General Surgery (SO2), 2 North Sub - Urology and Gynecology (NS2), SICU - Intensive Care Unit, and SSDU - Surgical Step-

**TABLE IV**
**Holt's Two Parameter Linear Exponential Smoothing**
1 - WEEK EX POST FORECASTING PERIOD

| Day | DSNO2 - 1 week | | DSSO2 - 1 week | | SO3 - 1 week | | DSNS2 -1 week | | PCU - 1 week | | DSSICU - 1 week | | DSSSDU - 1 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-23 | 22.0 | 25 | 28 | 29 | 26 | 28 | 9 | 12 | 8 | 14 | 12 | 16 | 11 | 14 |
| 11-24 | 16.0 | 25 | 23 | 29 | 24 | 28 | 11 | 11 | 8 | 14 | 15 | 16 | 12 | 13 |
| 11-25 | 14.0 | 25 | 23 | 29 | 21 | 28 | 11 | 11 | 13 | 14 | 19 | 16 | 12 | 13 |
| 11-26 | 14.0 | 25 | 26 | 29 | 25 | 28 | 10 | 11 | 15 | 14 | 15 | 16 | 12 | 13 |
| 11-27 | 19.0 | 25 | 31 | 29 | 29 | 28 | 14 | 11 | 11 | 14 | 16 | 16 | 15 | 13 |
| 11-28 | 26.0 | 25 | 28 | 29 | 29 | 28 | 13 | 10 | 16 | 14 | 18 | 16 | 15 | 13 |
| 11-29 | 27.0 | 25 | 32 | 29 | 31 | 28 | 14 | 10 | 15 | 14 | 18 | 16 | 15 | 13 |
| RMSE | | 7.30 | | 3.70 | | 3.57 | | 2.51 | | 3.55 | | 2.24 | | 1.85 |
| MAPE | | 0.39 | | 0.12 | | 0.12 | | 0.16 | | 0.30 | | 0.12 | | 0.14 |

**TABLE - V**
**Brown's Linear Exponential Smoothing**
2 - WEEK EX POST FORECASTING PERIOD

| Day | DSNO2 - 2 week | | DSSO2 - 2 week | | SO3 - 2 week | | DSNS2 -2 week | | PCU - 2 week | | DSSICU - 2 week | | DSSSDU - 2 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-16 | 29.0 | 29 | 33 | 30 | 31 | 29 | 14 | 15 | 16 | 15 | 19 | 18 | 14 | 15 |
| 11-17 | 30.0 | 29 | 31 | 30 | 31 | 29 | 14 | 15 | 16 | 15 | 20 | 18 | 15 | 15 |
| 11-18 | 24.0 | 29 | 22 | 30 | 27 | 29 | 14 | 15 | 16 | 15 | 15 | 18 | 17 | 16 |
| 11-19 | 22.0 | 29 | 24 | 30 | 27 | 29 | 16 | 15 | 14 | 15 | 21 | 19 | 16 | 16 |
| 11-20 | 24.0 | 30 | 30 | 30 | 31 | 29 | 11 | 15 | 14 | 15 | 16 | 19 | 15 | 16 |
| 11-21 | 28.0 | 30 | 32 | 30 | 29 | 29 | 14 | 16 | 15 | 15 | 14 | 19 | 15 | 16 |
| 11-22 | 24.0 | 30 | 27 | 30 | 24 | 29 | 11 | 16 | 11 | 15 | 16 | 20 | 12 | 16 |
| 11-23 | 22.0 | 31 | 28 | 30 | 26 | 29 | 9 | 16 | 8 | 15 | 12 | 20 | 11 | 16 |
| 11-24 | 16.0 | 31 | 23 | 30 | 24 | 29 | 11 | 16 | 8 | 15 | 15 | 20 | 12 | 16 |
| 11-25 | 14.0 | 31 | 23 | 30 | 21 | 29 | 11 | 16 | 13 | 15 | 19 | 21 | 12 | 16 |
| 11-26 | 14.0 | 32 | 26 | 30 | 25 | 29 | 10 | 17 | 15 | 15 | 15 | 21 | 12 | 16 |
| 11-27 | 19.0 | 32 | 31 | 30 | 29 | 29 | 14 | 17 | 11 | 15 | 16 | 21 | 15 | 16 |
| 11-28 | 26.0 | 32 | 28 | 30 | 29 | 29 | 13 | 17 | 16 | 15 | 18 | 22 | 15 | 17 |
| 11-29 | 27.0 | 33 | 32 | 30 | 31 | 29 | 14 | 17 | 15 | 15 | 18 | 22 | 15 | 17 |
| RMSE | | 9.68 | | 4.23 | | 3.41 | | 4.06 | | 3.16 | | 4.26 | | 2.70 |
| MAPE | | 0.44 | | 0.13 | | 0.11 | | 0.31 | | 0.23 | | 0.25 | | 0.18 |

**TABLE VI**
**Holt's Two Parameter Exponential Smoothing**
2 - WEEK EX POST FORECASTING PERIOD

| Day | DSNO2 - 2 week | | DSSO2 - 2 week | | SO3 - 2 week | | DSNS2 - 2 week | | PCU - 2 week | | DSSICU - 2 week | | DSSSDU - 2 week | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST | CENS | FCST |
| 11-16 | 29.0 | 28 | 33 | 30 | 31 | 30 | 14 | 15 | 16 | 16 | 19 | 17 | 14 | 15 |
| 11-17 | 30.0 | 28 | 31 | 30 | 31 | 30 | 14 | 16 | 16 | 16 | 20 | 17 | 15 | 15 |
| 11-18 | 24.0 | 28 | 22 | 30 | 27 | 30 | 14 | 16 | 16 | 16 | 15 | 17 | 17 | 15 |
| 11-19 | 22.0 | 28 | 24 | 30 | 27 | 30 | 16 | 16 | 14 | 16 | 21 | 17 | 16 | 16 |
| 11-20 | 24.0 | 28 | 30 | 30 | 31 | 30 | 11 | 16 | 14 | 16 | 16 | 17 | 15 | 16 |
| 11-21 | 28.0 | 28 | 32 | 30 | 29 | 31 | 14 | 17 | 15 | 16 | 14 | 17 | 15 | 16 |
| 11-22 | 24.0 | 28 | 27 | 30 | 24 | 31 | 11 | 17 | 11 | 16 | 16 | 17 | 12 | 16 |
| 11-23 | 22.0 | 28 | 28 | 30 | 26 | 31 | 9 | 17 | 8 | 16 | 12 | 17 | 11 | 16 |
| 11-24 | 16.0 | 28 | 23 | 30 | 24 | 31 | 11 | 17 | 8 | 16 | 15 | 17 | 12 | 16 |
| 11-25 | 14.0 | 28 | 23 | 30 | 21 | 31 | 11 | 18 | 13 | 16 | 19 | 17 | 12 | 16 |
| 11-26 | 14.0 | 28 | 26 | 30 | 25 | 31 | 10 | 18 | 15 | 16 | 15 | 17 | 12 | 16 |
| 11-27 | 19.0 | 28 | 31 | 30 | 29 | 31 | 14 | 18 | 11 | 16 | 16 | 17 | 15 | 16 |
| 11-28 | 26.0 | 28 | 28 | 30 | 29 | 31 | 13 | 18 | 16 | 16 | 18 | 17 | 15 | 16 |
| 11-29 | 27.0 | 28 | 32 | 30 | 31 | 31 | 14 | 19 | 15 | 16 | 18 | 17 | 15 | 16 |
| RMSE | | 7.30 | | 4.23 | | 4.57 | | 5.06 | | 3.76 | | 2.45 | | 2.66 |
| MAPE | | 0.32 | | 0.14 | | 0.14 | | 0.39 | | 0.26 | | 0.13 | | 0.17 |

down Unit (SSDU) deseasonalized data was used. The prefix 'DS' indicates deseasonalized data. The same process was repeated using Holt's method. The results of these forecasts are reported in tables III and IV.

Next, the same process was repeated using the first 351 periods for estimation and the last 14 periods for ex post forecasting. Once again, the two smoothing techniques, Brown's Linear and Holt's Two Parameter Linear Smoothing, were used to generate the ex post forecast. As in the case of 7-period forecasts, the original observed values were used for 3 South - Oncology and Neurology (SO3) and PCU - Progressive Care Unit (PCU), as no seasonal variations were determined for these two wards. For all the other five wards deseasonalized data were used. The results are presented in tables V and VI.

*3. Comparing The Two Approaches*

For each type of forecast two error measures, the root mean squared error (RMSE) and the mean absolute percentage error (MAPE) were calculated; they are summarized in tables VII and VIII. In comparing these measures to determine the accuracy of smoothing techniques relative to a regression model, the following conclusions were arrived at for the 1-week ex post forecasts (table VII):

a. Brown's method versus regression comparison showed that regression results are best in all cases except for the SICU.

b. Holt's method versus regression comparison showed that regression results are best, using both the RMSE and MAPE, in all cases except for the PCU and SICU.

**TABLE VII**
**Error Measure Comparisons**
1 - WEEK EX POST FORECASTS - DESEASONALIZED DATA

| Ward | Regression RMSE | Brown's RMSE | Holt's RMSE | Regression MAPE | Brown's MAPE | Holt's MAPE |
|---|---|---|---|---|---|---|
| NO2 | 5.66 | 6.92 | 7.30 | 0.32 | 0.37 | 0.39 |
| SO2 | 2.67 | 3.70 | 3.70 | 0.09 | 0.13 | 0.12 |
| SO3 | 2.27 | 4.11 | 3.57 | 0.08 | 0.13 | 0.12 |
| NS2 | 2.04 | 2.24 | 2.51 | 0.15 | 0.19 | 0.16 |
| PCU | 3.68 | 4.12 | 3.55 | 0.32 | 0.33 | 0.30 |
| SICU | 2.45 | 2.24 | 2.24 | 0.14 | 0.13 | 0.12 |
| SSDU | 1.25 | 1.85 | 1.85 | 0.08 | 0.15 | 0.14 |

## TABLE VIII
### Error Measure Comparisons
2 - WEEK EX POST FORECASTS - DESEASONALIZED DATA

| | Regression | Brown's | Holt's | Regression | Brown's | Holt's |
|---|---|---|---|---|---|---|
| Ward | RMSE | RMSE | RMSE | MAPE | MAPE | MAPE |
| NO2 | 4.66 | 9.68 | 7.30 | 0.22 | 0.44 | 0.32 |
| SO2 | 2.88 | 4.23 | 4.23 | 0.08 | 0.13 | 0.14 |
| SO3 | 2.41 | 3.41 | 4.57 | 0.07 | 0.11 | 0.14 |
| NS2 | 2.00 | 4.06 | 5.06 | 0.14 | 0.31 | 0.39 |
| PCU | 2.99 | 3.16 | 3.76 | 0.21 | 0.23 | 0.26 |
| SICU | 2.92 | 4.26 | 2.45 | 0.15 | 0.25 | 0.13 |
| SSDU | 1.51 | 2.70 | 2.66 | 0.08 | 0.18 | 0.17 |

c. Brown's versus Holt's method comparison showed that Holt's method, as should be expected, does better.

For the 2-week ex post forecasts (table VIII), comparisons yielded the following conclusions:

a. Brown's method versus regression comparison showed that regression results were best on a consistent basis.

b. Holt's method versus regression comparison showed that regression results were best in all cases except for the SICU.

c. Brown's versus Holt's method comparison showed that Brown's method produced slightly better results, but that there was no real difference between the two methods.

Inspection of the forecasts generated via the smoothing techniques and

## TABLE - IX
### Brown's Linear Exponential Smoothing
1 - WEEK EX POST FORECASTING PERIOD - SEASONAL VALUES

| Day | NO2 | | | SO2 | | | NS2 | | | SICU | | | SSDU | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST |
| 11-23 | 22.0 | 25 | 27 | 28 | 29 | 30 | 9 | 13 | 14 | 12 | 16 | 17 | 11 | 14 | 15 |
| 11-24 | 16.0 | 25 | 26 | 23 | 29 | 29 | 11 | 13 | 14 | 15 | 16 | 17 | 12 | 14 | 15 |
| 11-25 | 14.0 | 24 | 21 | 23 | 29 | 27 | 11 | 13 | 12 | 19 | 16 | 14 | 12 | 14 | 13 |
| 11-26 | 14.0 | 24 | 20 | 26 | 29 | 27 | 10 | 13 | 10 | 15 | 16 | 13 | 12 | 14 | 13 |
| 11-27 | 19.0 | 24 | 23 | 31 | 29 | 30 | 14 | 13 | 13 | 16 | 16 | 16 | 15 | 14 | 14 |
| 11-28 | 26.0 | 24 | 25 | 28 | 29 | 30 | 13 | 13 | 14 | 18 | 16 | 18 | 15 | 14 | 15 |
| 11-29 | 27.0 | 23 | 27 | 32 | 29 | 30 | 14 | 13 | 14 | 18 | 16 | 18 | 15 | 14 | 15 |

## TABLE - X
### Holt's Linear Exponential Smoothing
1 - WEEK EX POST FORECASTING PERIOD - SEASONAL VALUES

| Day | NO2 | | | SO2 | | | NS2 | | | SICU | | | SSDU | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST |
| 11-23 | 22.0 | 25 | 28 | 28 | 29 | 30 | 9 | 12 | 12 | 12 | 16 | 17 | 11 | 14 | 14 |
| 11-24 | 16.0 | 25 | 26 | 23 | 29 | 29 | 11 | 11 | 12 | 15 | 16 | 17 | 12 | 13 | 14 |
| 11-25 | 14.0 | 25 | 22 | 23 | 29 | 27 | 11 | 11 | 10 | 19 | 16 | 13 | 12 | 13 | 12 |
| 11-26 | 14.0 | 25 | 21 | 26 | 29 | 26 | 10 | 11 | 9 | 15 | 16 | 13 | 12 | 13 | 12 |
| 11-27 | 19.0 | 25 | 24 | 31 | 29 | 29 | 14 | 11 | 11 | 16 | 16 | 16 | 15 | 13 | 13 |
| 11-28 | 26.0 | 25 | 26 | 28 | 29 | 30 | 13 | 10 | 11 | 18 | 16 | 18 | 15 | 13 | 14 |
| 11-29 | 27.0 | 25 | 29 | 32 | 29 | 30 | 14 | 10 | 11 | 18 | 16 | 18 | 15 | 13 | 14 |

comparison to actual daily patient census data revealed another problem that casts serious doubt on the results of the forecasts: in some cases the forecasted values showed the same identical number of patients for the entire test period. For example, for SO2 both Brown and Holt methods produced a forecast of 29 patients for each of the seven day ex post period. Similarly, for the 2-week ex post period the forecasted values for SO2 were 30 patients. This problem was detected also for SO3, PCU, and SICU. The actual data clearly does not support the forecasted results.

One possible explanation for this problem could lie in the fact that there is definite seasonal variation present in the data for some of the wards, such as SO2. Therefore, using deseasonalized data for these wards to generate a forecast may result in the problem of obtaining the identical forecast for each day of the test period. No plausible explanation could be arrived at for the occurrence of the problem explained above for wards SO3 and PCU. Autocorrelation analysis showed no seasonal variations and thus the original data was used to generate the smoothing technique forecasts.

To see if this problem could be corrected, the forecasts generated from deseasonalized data for NO2, SO2, NS2, SICU, and SSDU were converted back to seasonalized figures by multiplying each day's census by that day's seasonal index. The resulting seasonal patients census forecasts were then used to recalculate the two error measures. The results are presented in tables IX through XIV, where DSFCST values represent forecasts based on deseasonalized data and SFCST values represent forecasted values, which have been converted to seasonal values.

For both Brown's and Holt's Linear Exponential Smoothing techniques and for both the 1 week and 2 week ex post forecasting periods, the seasonal forecast values for the five wards (NO2, SO2, NS2, SICU, and SSDU) showed changes from the constant value throughout the period. The forecast accuracy results, as measured by error measures, were rather mixed.

Using Brown's method and a 1-week ex post forecast period, when forecasted values were converted back into seasonal values, the error measures improved significantly for ward NO2. For ward SO2 there was a slight improvement. The error measurements deteriorated for wards NS2, SICU, and SSDU. For a 2-week ex post forecast period, error measures improved significantly for NO2 and SO2, deteriorated some for NS2 and SSDU, and deteriorated significantly for SICU.

Using Holt's method and a 1-week ex post forecast period, conversion of forecasted values into seasonal values resulted in a significant deterioration in the error measurements for ward NO2, a slight deterioration for SICU, and some improvement for wards SO2, NS2, and SSDU. For a 2-week ex post forecast period, error measures improved significantly for NO2 and SO2, deteriorated some for NS2 and SSDU, and deteriorated significantly for SICU.

In none of the cases described above, where there was an improvement in the error measures, did these error measures improve sufficiently to make the forecast superior to those generated via the regression analysis. The closest error measure was for ward NO2 where

**TABLE - XI**

**Brown's Linear Exponential Smoothing**

2 - WEEK EX POST FORECASTING PERIOD - SEASONAL VALUES

| Day | NO2 | | | SO2 | | | NS2 | | | SICU | | | SSDU | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST |
| 11-16 | 29 | 29 | 31 | 33 | 30 | 31 | 14 | 15 | 16 | 19 | 18 | 19 | 14 | 15 | 16 |
| 11-17 | 30 | 29 | 29 | 31 | 30 | 30 | 14 | 15 | 16 | 20 | 18 | 19 | 15 | 15 | 16 |
| 11-18 | 24 | 29 | 24 | 22 | 30 | 28 | 14 | 15 | 14 | 15 | 18 | 15 | 17 | 16 | 14 |
| 11-19 | 22 | 29 | 23 | 24 | 30 | 27 | 16 | 15 | 12 | 21 | 19 | 15 | 16 | 16 | 14 |
| 11-20 | 24 | 30 | 27 | 30 | 30 | 30 | 11 | 15 | 15 | 16 | 19 | 19 | 15 | 16 | 15 |
| 11-21 | 28 | 30 | 29 | 32 | 30 | 31 | 14 | 16 | 17 | 14 | 19 | 21 | 15 | 16 | 17 |
| 11-22 | 24 | 30 | 32 | 27 | 30 | 31 | 11 | 16 | 17 | 16 | 20 | 22 | 12 | 16 | 17 |
| 11-23 | 22 | 31 | 31 | 28 | 30 | 31 | 9 | 16 | 17 | 12 | 20 | 21 | 11 | 16 | 17 |
| 11-24 | 16 | 31 | 29 | 23 | 30 | 30 | 11 | 16 | 17 | 15 | 20 | 21 | 12 | 16 | 17 |
| 11-25 | 14 | 31 | 24 | 23 | 30 | 28 | 11 | 16 | 15 | 19 | 21 | 17 | 12 | 16 | 15 |
| 11-26 | 14 | 32 | 23 | 26 | 30 | 27 | 10 | 17 | 13 | 15 | 21 | 17 | 12 | 16 | 14 |
| 11-27 | 19 | 32 | 27 | 31 | 30 | 30 | 14 | 17 | 17 | 16 | 21 | 21 | 15 | 16 | 16 |
| 11-28 | 26 | 32 | 29 | 28 | 30 | 31 | 13 | 17 | 18 | 18 | 22 | 24 | 15 | 17 | 17 |
| 11-29 | 27 | 33 | 32 | 32 | 30 | 31 | 14 | 17 | 19 | 18 | 22 | 24 | 15 | 17 | 18 |

## TABLE - XII
## Holt's Linear Exponential Smoothing
2 - WEEK EX POST FORECASTING PERIOD - SEASONAL VALUES

| Day | NO2 | | | SO2 | | | NS2 | | | SICU | | | SSDU | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST | CENS | DSFCST | SFCST |
| 11-16 | 29 | 28 | 31 | 33 | 30 | 31 | 14 | 15 | 16 | 19 | 17 | 18 | 14 | 15 | 16 |
| 11-17 | 30 | 28 | 29 | 31 | 30 | 30 | 14 | 16 | 16 | 20 | 17 | 18 | 15 | 15 | 16 |
| 11-18 | 24 | 28 | 24 | 22 | 30 | 28 | 14 | 16 | 14 | 15 | 17 | 14 | 17 | 15 | 14 |
| 11-19 | 22 | 28 | 23 | 24 | 30 | 27 | 16 | 16 | 13 | 21 | 17 | 14 | 16 | 16 | 14 |
| 11-20 | 24 | 28 | 27 | 30 | 30 | 30 | 11 | 16 | 16 | 16 | 17 | 17 | 15 | 16 | 15 |
| 11-21 | 28 | 28 | 29 | 32 | 30 | 31 | 14 | 17 | 18 | 14 | 17 | 19 | 15 | 16 | 16 |
| 11-22 | 24 | 28 | 32 | 27 | 30 | 31 | 11 | 17 | 19 | 16 | 17 | 19 | 12 | 16 | 17 |
| 11-23 | 22 | 28 | 31 | 28 | 30 | 31 | 9 | 17 | 18 | 12 | 17 | 18 | 11 | 16 | 17 |
| 11-24 | 16 | 28 | 229 | 23 | 30 | 30 | 11 | 17 | 18 | 15 | 17 | 18 | 12 | 16 | 16 |
| 11-25 | 14 | 28 | 24 | 23 | 30 | 28 | 11 | 18 | 16 | 19 | 17 | 14 | 12 | 16 | 15 |
| 11-26 | 14 | 28 | 23 | 26 | 30 | 27 | 10 | 18 | 14 | 15 | 17 | 14 | 12 | 16 | 14 |
| 11-27 | 19 | 28 | 27 | 31 | 30 | 30 | 14 | 18 | 18 | 16 | 17 | 17 | 15 | 16 | 16 |
| 11-28 | 26 | 28 | 29 | 28 | 30 | 31 | 13 | 18 | 20 | 18 | 17 | 19 | 15 | 16 | 17 |
| 11-29 | 27 | 28 | 32 | 32 | 30 | 31 | 14 | 19 | 21 | 18 | 17 | 19 | 15 | 16 | 18 |

the RMSE for the regression forecast was 5.66 while the RMSE for Brown's forecast was 5.69. In all the other cases, the difference was much more substantial. The Brown 2-week ex post forecast for ward PCU remained the only forecast with error measures lower than those of the regression forecast.

## Conclusion

In comparing the regression model and its results to those obtained via the smoothing techniques using strictly error measures, such as the RMSE and MAPE used in this case, one is led to conclude that the regression modeling of real-world event still seems to be a more powerful tool of forecasting, ceteris paribus. In addition to error measure comparisons, a careful inspection of the forecast results presented above yields the following conclusion: even in cases where the error measures are fairly close, or even lower for the smoothing techniques, the forecast casts some doubt on the ability of the method to accurately predict. Conversion of forecast values based on deseasonalized data to seasonal values results in variations in the forecasted values, but often at the expense of forecast accuracy, as measured by error measures RMSE and MAPE.

Thus, recognizing that one study is not sufficient to generalize conclusions, this author believes one can still argue strongly in favor of regression models as a method of forecasting.

## TABLE VIII
## Error Measure Comparisons
1 - WEEK EX POST FORECASTS - SEASONALIZED DATA

| Ward | Regression RMSE | Brown's RMSE | Holt's RMSE | Regression MAPE | Brown's MAPE | Holt's MAPE |
|---|---|---|---|---|---|---|
| NO2 | 5.66 | 5.69 | 6.30 | 0.32 | 0.29 | 0.33 |
| SO2 | 2.67 | 3.07 | 3.12 | 0.09 | 0.10 | 0.10 |
| SO3 | 2.27 | 4.11 | 3.57 | 0.08 | 0.13 | 0.12 |
| NS2 | 2.04 | 2.30 | 2.20 | 0.15 | 0.15 | 0.17 |
| PCU | 3.68 | 4.12 | 3.55 | 0.32 | 0.33 | 0.30 |
| SICU | 2.45 | 2.88 | 3.14 | 0.14 | 0.14 | 0.14 |
| SSDU | 1.25 | 2.00 | 1.65 | 0.08 | 0.12 | 0.10 |

## TABLE XIV
### Error Measure Comparisons
1 - WEEK EX POST FORECASTS - SEASONALIZED DATA

| Ward | Regression RMSE | Brown's RMSE | Holt's RMSE | Regression MAPE | Brown's MAPE | Holt's MAPE |
|---|---|---|---|---|---|---|
| NO2 | 5.66 | 6.60 | 6.60 | 0.32 | 0.28 | 0.28 |
| SO2 | 2.67 | 3.40 | 3.40 | 0.09 | 0.11 | 0.11 |
| SO3 | 2.27 | 3.41 | 4.57 | 0.08 | 0.11 | 0.14 |
| NS2 | 2.04 | 4.38 | 5.39 | 0.15 | 0.34 | 0.41 |
| PCU | 3.68 | 3.16 | 3.76 | 0.32 | 0.23 | 0.26 |
| SICU | 2.45 | 5.02 | 3.42 | 0.14 | 0.27 | 0.17 |
| SSDU | 1.25 | 3.11 | 2.96 | 0.08 | 0.20 | 0.19 |

## References

Helmer, F. Theodore, Oppermann, Edward B., Suver, James D., "Forecasting Nursing Staffing Requirements By Intensity-of-Care Level," *Interfaces*, Vol. 10, No. 3, June 1980, pp 50 - 55.

Taylor, H. W., "A Study of Factors Affecting Laboratory Workload," *Clinical Biochemistry*, 4: 179 - 182.

Wilson, J. Holton, Schuiling, Steven J., "Forecasting Hospital Laboratory Procedures," *Journal of Medical Systems*, Vol. 16, No. 6, 1992, pp 269 - 279.

Bails, Dale G., Peppers, Larry C., *Business Fluctuations - Forecasting Techniques and Applications*, 2nd ed. Prentice Hall, Englewood Cliffs, New Jersey, 1993.

Vargha, Nader, *Forecasting Average Daily Census for a Medium-sized Hospital*, presented at the Southwest Business Symposium, Edmond Oklahoma, 2001.