Forecasting an Electric Utility's CO_2 Emissions Using SAS/AF® and SAS/STAT® Software: A Linear Analysis

Md. Azharul Islam, The Ohio State University, Columbus, Ohio.
David Wang, The Public Utilities Commission of Ohio, Columbus, Ohio.

ABSTRACT

This paper presents a forecasting model which analyzes and projects CO_2 emissions for an electric utility. The linear spline regression model is used to forecast CO_2 emissions. The model is written on SAS which includes Base SAS®, SAS/AF and SAS/STAT. The forecast model and the SAS/AF application, discussed in this paper, may be one tool that could be used to predict CO_2 emissions and therefore to help take necessary measures for their reduction.

INTRODUCTION

Forecasting an electric utility's CO_2 emissions will play a significant role in estimating the future magnitude and highlighting the temporal growth behavior of its CO_2 emissions. Considering the contribution level (50%) of CO_2 to the global warming and therefore the need for taking necessary measures, it is needless to say the importance of estimating electric utilities' CO_2 emissions for the years to come. The periodic magnitude of CO_2 emissions of an electric utility changes, depending on its periodic electricity generation, the type of fuel used for generation and the environmental activities. The sudden alteration in the growth behavior of a utility's CO_2 emission activities due to energy and environmental changes such as the Clean Air Act Amendment (CAAA) of 1990, the Energy Policy Act of 1992 and so on, may be derived from the available historical time series of its CO_2 emissions. A forecasting model which can measure and estimate the temporal growth behaviors of historical CO_2 emissions and forecast their patterns into the future in a logical manner, is necessary for this purpose. Different researchers have shown empirically that linear spline is one of the useful methods in measuring and projecting the observed characteristics of a time series (Islam, 1995)

This paper presents a forecasting model which analyzes and predicts CO_2 emissions for an electric utility. The concept of assessing time variant linear trends in annual economic data series is considered in the model (Feyzioglu, 1983). The forecasting model is written in SAS. The primary SAS products used in the program were Base SAS, SAS/STAT and SAS/AF. The paper is organized into five sections. The first section presents the theoretical framework of the forecasting model while section two gives a general overview of the SAS program. Section three presents a discussion of the forecast results of the electric utility $"u"$. Section four describes the SAS/AF application and the final section offers concluding remarks.

SECTION 1: FORECASTING MODEL FOR CO_2 EMISSIONS

The mathematical forecasting model of CO_2 emissions provides an empirical estimate of the growth of CO_2 emissions for a particular future horizon, based on historical CO_2 emissions.

The total emissions within the domain of analysis is determined by the initial magnitude of emissions at time plus the magnitudes of a series of emissions. Let C_u be the variable which represents a set of estimated historical time series of CO_2 emissions of an electric utility $"u"$.

Let $H_u = \langle t_1, t_n \rangle = \langle 1979, 1993 \rangle$ be the historical interval of time of size n years (i.e., January 1979 through December 31, 1993) over which $"n"$ annual observations of C_u are available. Let $F_u = \langle t_{n+1,} t_f \rangle = \langle 1994, 2005 \rangle$ be the forecast horizon. Let $D_u = \langle t_{1,} t_f \rangle = \langle 1979, 2005 \rangle$ be the domain of analysis.

The linear spline model for the forecasting time series of CO_2 emissions of the electric utility is outlined as follows:
$C_u t = C_u(t) + e_{u(t)}$ over the estimated historical time series of CO_2 emissions of the utility at H_u (1979 through 1993) and forecast horizon F_u (1994 through 2005)

where:

$C_u t =$ the magnitude of CO_2 emissions by the electric utility u in year t (i.e., historical and forecast horizon of 1979 through 2005).

$C_u(t) =$ the assumed deterministic (predicted) component of the model which is defined on the entire domain of analysis, with:

$$C_u(t) = a_0 + \sum_{i=1}^{k} b_i T_i$$

where:

$a_{0,} b_i = (1, 2, ..., k), 1 \le k < n$ are the parameters to be estimated, T is an appropriate transformation of time and the T's are index numbers representing the years in the domain of analysis that includes both historical and forecast periods.

$k =$ the index number of turning or knot points at time $T_1, T_2 ...., T_k$ in which CO_2 emissions level demonstrated a change relative to the preceding state of emission in the linear spline analysis. Note that number and position of the knot points can differ from one electric utility to another.

The values of the parameters indicate the magnitudes of discrete change in growth over the successive subdomains. That is, the parameter b_i, also called slope of the forecast line, designates the amount of increase (or decrease) in the deterministic component of $C_u(t)$ for every 1-unit increase in t_i. $e_{u(t)} =$ the assumed stochastic or probabilistic component of the model. It is the difference between the observed/historical and the predicted/forecast values of CO_2, with: $e_{u(t)} = C_u(t)$. It encompasses the historical domain of analysis. The model assumes that $e_{u(t)} \sim N[0, \sigma^2]$ and provides with the forecast and 95% lower and upper bounds, which are calculated as $\pm 2\sigma$ (i.e., standard deviation) from the forecast.

It has also been assumed that no external or internal events such as war, oil price increases and recessions will occur until 2005.

SECTION 2: THE SAS PROGRAM

The SAS program uses a number of procedures from Base SAS and SAS/STAT. The PLOT procedure is used to establish the general location of the discontinuities in the time path trajectory and to determine which independent variables are likely candidates to be used in the model. The STEPWISE procedure assists the user in selecting the appropriate model while the REG procedure uses the results from the STEPWISE procedure to generate the weighted least squares estimates of parameters for the linear spline model. Finally, the UNIVARIATE procedure is used to determine the theoretical percent error term (i.e., the standard deviation).

In summary, program was developed to estimate the periodic magnitude of CO_2 emissions and to capture significant events while identifying a model. It also calculates the observed error ratios, the expected error ratios and their comparisons.

SECTION 3: RESULTS

As an example, the forecast results of CO_2 emissions of the electric utility "u" is discussed as follows:

The model parameters estimates, the historical and forecast data, the values of the independent variables of u's forecasting model and the graphical representation of the CO_2 emissions are presented in tables 1 through 3 and in figure 1. The equation of the estimated forecasting model is:

$CO_2 u =$ 39329-2180*T79+4058*T82+3800*T88-4981*T90
(24.4) (-3.0) (4.3) (2.4) (-2.1)

$$R^2 = 0.9505$$

The t values shown in parentheses and the R^2 of the model are significant. T79, T82, T88 and T90 represent the selected independent variables of the model.

It can be seen from the figure 1 that CO_2 emissions, starting from an initial level of 40,964 thousand tons per year in 1979, have dropped to 34,980 thousand tons in 1980. This decline may be attributed to loss of generation of electricity from coal, which led to import of electricity and consequently a drop in CO_2 emissions from 1979 to 1980. It is observed from table 2 and figure 1 that between 1980 and 1983, CO_2 emissions declined slightly. This may be due to variations in the business cycle. Since 1983, CO_2 emissions have followed an increasing trend. In 1993, CO_2 emissions in the electric utility "u" was at 57,823 thousand tons, and it is projected to reach 66,252 in 2000 and 69,863 thousand tons in 2005.

Furthermore, we would expect that the future realizations of CO_2 emissions would be within the $\pm 3.38\%$ (table 3) around the forecast line.

SECTION 4: THE FORECASTING TOOL

A User-friendly menu-driven front end screen was developed using the SAS/AF, SAS Frame and SCL language. The purpose of the menu was to allow users that were unfamiliar with the forecasting methodology or with the SAS language to be able to project CO_2 emissions with a minimum amount of experience.

The primary menu allows the user an opportunity to: edit the data, plot the historical time series, model and forecast the data, and exit out of the application. SAS Frame gives the developer the flexibility of adding additional choices with a minimum amount of work. The first step in any forecasting methodology is to look at the data. The VIEW OF DATA option allows the user to view the data from among selected data sets interactively. The user has the choice of three different data sets and four different graphical presentations. When the user clicks on any of the choices, a plot of data will display immediately. The EDIT option specifies three SAS data files that the user can edit. PROC FSEDIT is used as the mechanism for editing the data. The user has the ability to look at the actual numbers from the data set, make changes and save the changes. The FORECAST option allows the user to choose the following parameters: the company data, the independent variables for the stepwise procedure and regression procedure, the percentage error term which is used to calculate the upper and lower bounds, and the last year of the forecast. The user can generate SAS output on the screen by selecting the RUN Icon. When he or she has completed viewing the results, there is a customized menu which will allow the user to return to the FORECAST option. A simple HELP Icon was included with each screen to briefly explain how to select among the various choices.

The SAS class entitled "Building SCL Applications Using FRAME Entries" was an invaluable source of information. Several SCL programs found in the course notes were modified and used to build certain modules of the SAS/AF application. Also, Anita Hillhouse with SAS Institute was instrumental in facilitating our efforts by answering several of our questions.

CONCLUDING REMARKS

With the aim of estimating current and future comprehensive pictures of an electric utility's CO_2 emissions, it is essential to estimate first its fuel-specific historical emissions data and then to forecast its fuel-specific CO_2 emissions. Forecasting CO_2 emissions will facilitate the decision-making process in selecting the appropriate action to reduce CO_2 emissions. The forecast model and the SAS/AF application, discussed in this paper, may be one tool that could be used to accomplish these important tasks.

REFERENCES

Feyzioglu, G., (1983). "The Analysis and Assessment of Time Variant Linear Trends in Annual Economic Data Series with an Application to Energy Forecasting for the State of Ohio", *Ph.D. Dissertation,* The Ohio State University, Columbus, Ohio.

Islam, M. A., (1995). "Carbon Dioxide Emission Reduction in Ohio: A study of Three Options". *M.C.R.P. Thesis,* Department of City and Regional Planning, The Ohio State University, Columbus, Ohio.

SAS Institute Inc. (1995), *Building SCL Applications Using FRAME Entries*, Cary, NC: SAS Institute Inc.

Contact Author:
Md. Azharul Islam
622 Riverview Drive
Apt. # 1
Columbus, Ohio 43202
(614)-268-4408 (Home)
(614) 644-7691 (Work)
Internet: m.azhar.islam@ohio.gov

TABLE 1: PARAMETER ESTIMATES OF THE FORECASTING MODEL
UNITS = THOUSAND TONS PER YEAR

Model: MODEL1
Dependent Variable: CO_2

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	4	451.64172	112.91043	47.967	0.0001
Error	10	23.53913	2.35391		
C Total	14	475.18085			

Root MSE	1.53425	**R-square**	0.9505
Dep Mean	40761.46990	**Adj R-sq**	0.9306
C.V.	0.00376		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for Ho: Parameter=0	Prob> |T|
INTERCEP	1	39329	1609.2530954	24.439	0.0001
T79	1	-2180.094005	715.98135281	-3.045	0.0124
T82	1	4058.319214	939.82989526	4.318	0.0015
T88	1	3779.967867	1547.8595821	2.442	0.0347
T90	1	-4980.588294	2327.1377394	-2.140	0.0580

TABLE 2: HISTORICAL AND FORECAST OF CO_2 EMISSIONS
UNITS = THOUSAND TONS PER YEAR

Obs	Year	CO_2	Upper Limit	Predicted	Lower Limit
1	1979	40964	41924.43	39328.73	36733.03
2	1980	34980	39600.45	37148.64	34696.83
3	1981	35763	37276.47	34968.54	32660.62
4	1982	33582	34952.49	32788.45	30624.41
5	1983	33887	36954.67	34666.67	32378.67
6	1984	34408	38956.86	36544.90	34132.93
7	1985	40899	40959.05	38423.12	35887.20
8	1986	44267	42961.24	40301.35	37641.46
9	1987	40483	44963.43	42179.57	39395.72
10	1988	43312	46965.61	44057.80	41149.98
11	1989	49164	52997.25	49715.99	46434.74
12	1990	55364	59028.88	55374.19	51719.49
13	1991	57229	59751.21	56051.79	52352.37
14	1992	55597	60473.53	56729.39	52985.25
15	1993	57823	61195.86	57407.00	53618.14
16	1994	.	61918.19	58084.60	54251.02
17	1995	.	62640.51	58762.21	54883.90
18	1996	.	63362.84	59439.81	55516.79
19	1997	.	64085.17	60117.42	56149.67
20	1998	.	64807.49	60795.02	56782.55
21	1999	.	65529.82	61472.63	57415.43
22	2000	.	66252.15	62150.23	58048.32
23	2001	.	66974.48	62827.84	58681.20
24	2002	.	67696.80	63505.44	59314.08
25	2003	.	68419.13	64183.05	59946.97
26	2004	.	69141.46	64860.65	60579.85
27	2005	.	69863.78	65538.26	61212.73

TABLE 3: THEORETICAL PERCENTAGE ERROR RATIO

Univariate Procedure

Variable=TW

Moments				Quantiles (Def=5)		
N	15	**Sum Wgts**	15	100% **Max**	0.0984	99%
Mean	0.033822	**Sum**	0.507326	75% **Q3**	0.058377	95%
Std Dev	0.26403	**Variance**	0.000697	50% **Med**	0.058377	90%
Skewness	1.104542	**Kurtosis**	1.033769	25% **Q1**	0.016928	10%
USS	0.026918	**CSS**	0.009759	0% **Min**	0.000184	5%
CV	78.06419	**Std Mean**	0.006817			1%
T:Mean=0	4.961281	**Pr>|T|**	0.0002	**Range**	0.098216	
Num ^=0	15	**Num > 0**	15	**Q3-Q1**	0.04145	
M(Sign)	7.5	**Pr>=|M|**	0.0001	**Mode**	0.000184	

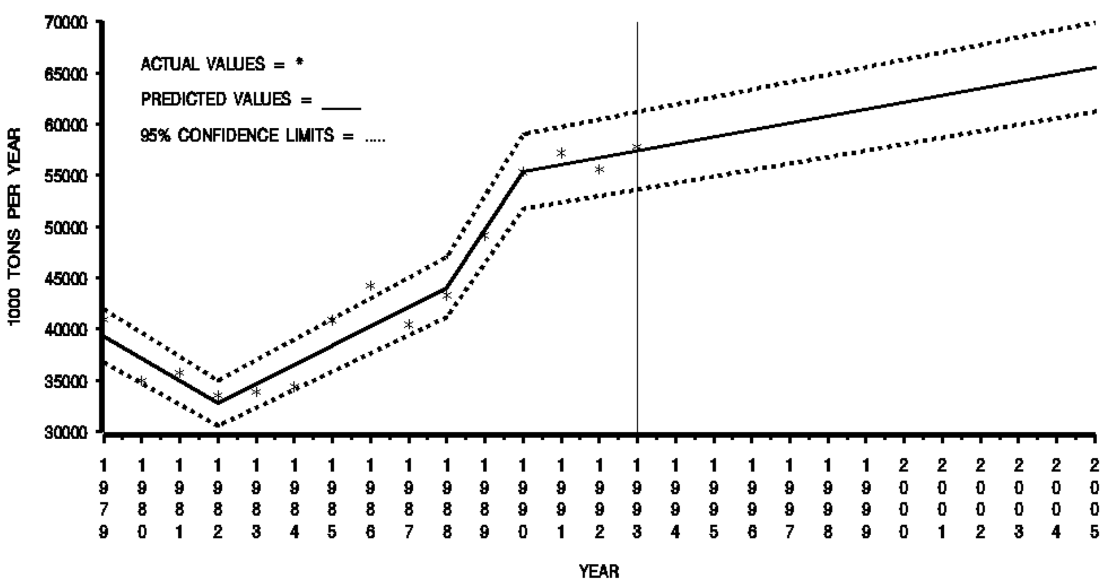
Extremes

Lowest	Obs	Highest	Obs
0.000184(	12)	0.04158(	1)
0.007247(	15)	0.058377(	2)
0.011103(	11)	0.58473(	6)
0.016928(	10)	0.064437(	7)
0.019961(	14)	0.0984(	8)

Missing Value	.
Count	12
% Count/Nobs	44.44

FIGURE 1: TIME TRAJECTORY OF CO2 EMISSIONS

UNITS=THOUSAND TONS OF CO2 EMISSIONS PER YEAR



SOURCE: DIVISION OF FORECASTING, PUCO