Airline Emission Forecast Through Empirical Analysis of Its Contributing Factors

J. Nirmalraj and Dr. P. Malliga

***Abstract*—Climate change and environmental degradation are the greatest challenges of this century. Governments and various international organizations are aiming at the reduction and stabilization of emission level that does not endanger the environment. International Civil Aviation Organisation (ICAO) is fully engaged to achieve a global solution to address emission from the civil aviation across the globe. Since the usage of civil aviation is increasing at a rapid rate due to various economical and technological reasons, a clear understanding of the factors affecting the aircraft fuel consumption and its CO_2 emission is essential.**

The primary objective is to estimate the CO_2 emission from the aircrafts, which is done by estimating the fuel consumption as CO_2 emission is directly proportional to the fuel consumption. The data regarding the aircraft fuel consumption and its influencing factors, sector details and trip details are collected from the various databases of an Airline and integrated it. The work has been accomplished with a detailed study of the factors affecting fuel consumption through regression analysis using MINITAB and thereby forecasted the fuel consumption. The validation has been carried out using ANOVA, Durbin-Watson statistics and Mallow's Cp Statistics. The CO_2 emission has been predicted using the forecasted fuel consumption.

The work could be extended to find the optimal values for the factors affecting the fuel consumption so that the CO_2 emission could be minimized. User-friendly software could be developed for the performance engineers to monitor and take managerial decisions based on the fuel consumption and the CO_2 emission.

***Index Terms*— Airline emission, Carbon emission calculator, Empirical analysis.**

I. Introduction

Aviation has recently experienced a phenomenal growth that has created new opportunities and new challenges for airlines, airport operators, aircraft manufacturers, air traffic service providers and other related air transport organizations including Government who make policies in this respect. The development and growth of air transport depends on various factors including economic, trade, fuel price changes, airline productivity gains, airports and airspace capacity. The increase in the personal income affects the level of consumer purchasing power and the tendency to undertake luxury travel. The world economy maintained its growth momentum in 2007 despite higher prices for crude oil and refined products. Developing countries contributed significantly to the emission rate as their average GDP grew by 9.8 percent. The economies of China and India showed remarkable GDP growth at 11.5 percent and 8.9 percent respectively, driven by growing exports, investment and growing demand with the increase in the emission rate. Asia's newly industrialized economies posted a 4.9 percent GDP growth [1]. Worldwide, the total number of annual passengers has grown by 46 percent in the past 10 years, as the number of passengers flying climbed from 1.46 billion to 2.13 billion per year. Freight tonne-kilometre figures show an almost identical rate of increase. The International Air Transport Association (IATA) forecasted that in 2011 the air-transport industries will handle 2.75 billion passengers and 36 million tonnes of freight. International passenger demand is expected to rise from 760 million passengers in 2006 to 980 million in 2011 at an annual average growth rate of 5.1 percent [2].

At the end of 2006 from 1997, the reported number of commercial air transport aircraft in service increased by about 30 percent [1]. It is clear that the world's scheduled airlines experienced material operating profits and overall net profitability despite rising fuel prices. The upcoming private airlines have been one of the greatest transformations in air transport industry along with the increase in emission level.

II. The Objective And Data Collection

There is an ongoing debate about possible taxation of air travel and the inclusion of aviation in an emissions trading scheme, with a view to ensuring that the total external costs of aviation are taken into account. Virtually all airlines with operations, to and fro and within the European Union will come under the scope of the EU's emissions trading scheme from 2012[3]. Airlines are required to submit a monitoring plan by August 2009 and to monitor data from 2010 along with the methodology in which it is done. To meet these criteria the following objectives were framed.

- To analyze the factors affecting the fuel consumption and
- To forecast the fuel consumption thereby calculating the CO_2 emission as CO_2 emission is directly proportional to the fuel consumption [4].

Data collection and integration was a crucial part of this work as three different sources of information was used for analyzing and forecasting fuel consumption. For each and every flight the data such as fuel requirement and fuel efficiency are taken from the Fuel Management System (FMS). The flight performance data was taken from the ARMS server. The eight month duration performance details of various aircraft were collected for all five different types of

Manuscript received May 27, 2010. This work was a part of the academic research activity to support an Airline in the study of carbon emitted from the aircraft during the various trips across the country.

Mr. J. Nirmalraj is a PG student of Quality Engineering and Management, in the Department of Industrial Engineering, Anna University. Mobile: +91-99621- 24723; e-mail: nirmal.jayapal@ gmail.com.

Dr. P. Malliga, Assistant Professor, Department of Industrial Engineering, Anna University e-mail: mallipoo@annauniv.edu.

aircrafts in all the sectors. The corresponding Fuel Management system (FMS) data for all the flights were also collected. The details of the flight such as the seat capacity, engine type, etc were collected from the Manufacturers Manual. The passenger details, the passenger ratio and the cargo details in a particular sector were also collected for all the flights.

III. Empirical Analysis

Multiple regression is a procedure that separates the systematic effects (associated with the explanatory variables) from the random effects (associated with the error term) and also offers a method of assessing the success of the process [5]. Multiple regression is useful [6] in determining whether a particular effect is present; [7] in measuring the magnitude of a particular effect; and [8] in forecasting what a particular effect would be, but for an intervening event. In this work the multiple regression analysis is used to find the various effects of the independent variables such as the time of flight, take of weight, captain of the flight, center of gravity, temperature, etc on fuel consumption which would be the dependent variable to be explained. The independent variables that are more correlated with the dependent variable are segregated first and then they are used to generate the regression equation to predict the fuel consumption. The independent variables are used to generate different regression models with all possible combinations to predict the fuel consumption. The best regression model is arrived using the subset analysis and then the fuel consumption is forecasted and thereby CO_2 emission is predicted from the resulting regression equation.

Multiple regression analysis is well suited to the analysis of data about competing theories in which there are several possible explanations for the relationship among a number of explanatory variables [5]. Multiple regression typically uses a single dependent variable and several explanatory variables to assess the statistical data pertinent to these theories.

A. *Methodological Approach*

The purpose of multiple regression is to establish a quantitative relationship between a group of independent variables (Xs) and a dependent variable, Y [9]. This relationship is useful for understanding which independent variable have the greatest effect, knowing the direction of the effect (i.e., increasing X increases or decreases Y), and using the model to predict future values of the dependent variable, when only the independent variables are currently known. If two variables are correlated, then knowing the value of one variable will allow predicting the value on the other variable. The stronger the correlation, closer the value will fall to the regression line and therefore more accurate prediction. Multiple regression is simply an extension of this principle, where prediction of dependent variable lies with the influence of several other independent variables. Using multiple regression tests, regression models (or theories) are generated and identify precisely which set of variables is influencing the fuel consumption. Figure 1-1 displays the flowchart for empirical analysis.

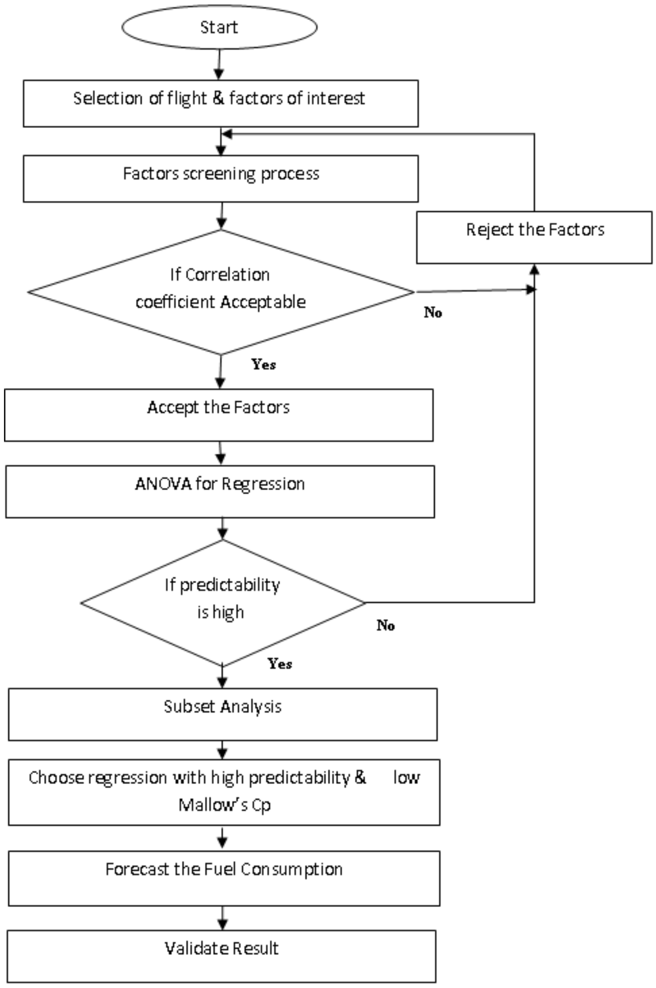

Fig. 1- 1 Flowchart for Empirical Analysis

B. *Factors Screening Process*

Through an extensive literature review various factors affecting the fuel consumption were identified. About 28 factors of interest as shown in Table-1 was considered as the factors affecting the fuel consumption for the flight VT-KFI - understudy. MINITAB was used to find the correlation between the factors of interest.

MINITAB generates an equation to describe the statistical relationship between the independent variables and the dependent variable and to predict new observations. Regression generally uses the ordinary least squares method which derives the equation by minimizing the sum of the squared residuals. Regression results indicate the direction, size, and statistical significance of the relationship between independent variables and dependent variable. Sign of each coefficient indicates the direction of the relationship. Coefficients represent the mean change in the dependent variable for one unit of change in the independent variables while holding other independent variables in the model constant. The null hypothesis is tested for each coefficient and the p-value is found. Therefore, lower the significance level, the stronger the evidence required. Choosing level of significance is an arbitrary task, but for many applications, a level of 5% is chosen [10][11].The regression equation predicts new observations with the given specified independent variables.

The 28 factors of interest were segregated as qualitative and quantitative data. The 17 quantitative factors for the particular tail registration were given as input along with the fuel consumption data in MINITAB. The regression equations were generated and the p-value test for each independent variable was done to screen the uncorrelated variables. The appropriateness of rejecting the null hypothesis in a hypothesis test, p-values range from 0 to 1. According to Sterne JAC and Schervish MJ "the smaller the p-value, the smaller the probability of rejecting the null hypothesis, is a mistake "[12] [13].The commonly used value is 0.05. If the p-value is smaller than or equal to the significant level, often represented by the Greek letter α (alpha), reject the null hypothesis. Because of their indispensable role in hypothesis testing, p-values are used in many areas of statistics including basic statistics, linear models, reliability, and multivariate analysis among many others.

TABLE - 1 FACTORS OF INTEREST

➢ Date	➢ Temperature	➢ Sector
➢ Time of Flight	➢ Captain	➢ Mach
➢ Air Time of Flight	➢ Gravitational force	➢ Center of gravity
Taxi Time of Flight	➢ Fuel flow rate	➢ Altitude
➢ Empty weight	➢ Take of weight	➢ Air Pressure
➢ Weight of Passengers	➢ Departure Fuel	➢ Type of flight
Engine Temperature	➢ Portable water	➢ Time of Delay
➢ Night Flight	➢ Arrival Station	➢ Departure Station

Based on the p-values of each and every independent factor with respect to the dependent variable i.e. fuel consumption, 7 factors were screened and there p-values were analyzed along with the Variance Inflation Factor (VIF) of those variables. The VIF measures the impact of collinearity among the variables in a regression model. The VIF is 1/Tolerance, it is always greater than or equal to 1. There is no formal VIF value for determining the presence of multicollinearity [14]. Values of VIF that exceed 10 are often regarded as indicating multicollinearity, but in weaker models values above 2.5 may be a cause for concern. When the VIF values are high for any of the variables in the model, multicollinearity is probably an issue. The Table-2 shows that the independent variables that were screened are free from the multicollinearity problem. Thus the regression equation arrived from the screened independent variables was found as shown in the Equation 1

$$\text{Fuel consumption} = 306 + 36.7\,X1 + 0.0388X2 + 0.00235X3 - 1344X4 - 0.53\,X5 + 7.39\,X6 - 195\,X7 \qquad (1)$$

The ANOVA for Regression equation was found to test the level of variability within the independent variables in the regression equation and is shown in Table - 3. The p-value in the ANOVA table (0.000) shows that the model estimated by the regression procedure is significant at a α-level 0.05. This indicates that at least one coefficient is different from zero.

TABLE - 2 REGRESSION ANALYSIS TABLE

Variables	**Coef**	**SE Coef**	**P**	**VIF**
Constant	306	3923	0.038	
Time of flight(X1)	36.6864	0.3568	0.000	1.673
Take of Weight(X2)	0.038839	0.003308	0.000	1.127
Altitude(X3)	0.002351	0.002630	0.371	1.524
Mach(X4)	-1343.6	385.4	0.001	1.182
Temperature(X5)	-0.533	1.582	0.236	1.206
Centre of Gravity(X6)	7.393	4.725	0.118	1.025
Gravitational force (X7)	-194.7	396.6	0.124	1.047

TABLE - 3 ANOVA FOR REGRESSION

Source	**DF**	**SS**	**MS**	**F**	**P**
Regression	7	269×10^7	3844×10^5	2733	0
Residual Error	1813	2550×10^5	140702		
Total	1820	2946×10^6			
S = 375.103 R^2 = 91.3% R^2 (adj) = 91.3% R^2 (pred) = 91.01% Durbin-Watson statistic = 1.92491					

The R^2 value indicates that the independent variables explain 91.3% of the variance in fuel consumption. The adjusted R^2 is 91.3%, which accounts for the number of independent variables in the model. Both values indicate that the model fits the data well. The predicted R^2 value is 91.01%. Because the predicted R^2 value is close to the R^2 and adjusted R^2 values, the model does not appear to be over-fit and has adequate predictive ability. The Durbin–Watson statistic (d) is a test statistic used to detect the presence of autocorrelation in the residuals from the regression analysis. The value of'd' always lies between 0 and 4 [15] [16] [17]. If the Durbin–Watson statistic is substantially less than 2, there is an evidence of positive serial correlation. Small values of d indicate successive error terms are, on average, close in value to one another, or positively correlated. If d > 2 successive error terms are, on average, much different in value to one another, i.e., negatively correlated [18][19][20]. In this regression, it implies that the level of statistical insignificance for autocorrelation since the value is close to two this indicates that there is no autocorrelation effect between the variables.

TABLE – 4 SUBSET ANALYSIS TABLE

Variables	R^2	R^2 adj	Mallows Cp	S	X1	X2	X3	X4	X5	X6	X7
1	89.9	89.9	41.7	300	X						
1	41.8	41.7	1716.2	720			X				
2	90.6	90.5	20.4	290	X	X					
2	90.2	90.1	34.7	296	X						X
3	90.9	90.8	12.6	286	X	X	X				
3	90.8	90.7	14.7	287	X	X					X
4	91.1	91.0	7.2	283	X	X	X				X
4	91.0	90.9	10.4	285	X	X	X		X		
5	91.2	91.0	5.4	282	X	X	X		X		X
5	91.1	91.0	7.8	283	X	X	X			X	X
6	91.2	91.0	6.0	282	X	X	X		X	X	X
6	91.2	91.0	7.3	283	X	X	X	X	X		X
7	91.2	91.0	8.0	283	X	X	X	X	X	X	X

A. *Subset Analysis*

Best subset regression identifies the best-fitting regression models that can be constructed with the independent variable screened. Best subset regression is an efficient way to identify models that achieve in predicting the fuel consumption with as few independent variables as possible. Subset models estimate the actual regression coefficients and predict the future responses with smaller variance when compared the full model using all independent variables [21].

MINITAB examines all possible subsets of the independent variables, beginning with all models containing one independent variable, and then all models containing two independent variables, and so on. By default, MINITAB displays the two best models for each number of independent variables as shown in the Table - 4. The table explains that each line of the output represents different regression model. R^2 and adjusted R^2 are converted to percentage value. Independent variables that are present in the model are indicated by X. From the table, it is not clear that which model fits the data. There are 6 models which has the highest adjusted R^2 (91.0%).

Mallows proposed the statistic as a criterion for selecting among many alternative subset regressions [22]. Mallows' Cp compares the precision and bias of the full model to models with the best subsets of independent variables. It helps to strike an important balance with the number of independent variables in the model. A model with too many independent variables can be relatively imprecise while one with too few can produce biased estimates [23]. A Mallows' Cp value that is close to the number of independent variables plus the constant indicates that the model is relatively precise and unbiased in estimating the true regression coefficients and predicting future responses. The simplified results from Best Subset Regression analysis shows that the highest adjusted R^2 (91.0%) and lowest Mallows' Cp value (5.4), and the lowest S value (282) has five-independent variable model with all variables except Mach(X4) and Center of Gravity(X6) .The multiple regression indicates that adding the variable Mach (X4) and Centre of Gravity (X6) does not improve the fit of the model.

B. *Interpreting Regression Result*

The regression equation is generated by giving the historical data of independent variables from the selected best subset model along with historic data of fuel consumption into MINITAB, which produces the regression equation as shown in Equation 2.

$$\text{Fuel consumption} = -2490 + 35.7\,X1 + 0.0411\,X2 + 0.00734\,X3 + 1.78\,X5 - 8\,X7 \quad (2)$$

This regression implies that the level of statistical insignificance for autocorrelation, since the value is close to 2 this indicates that there is no autocorrelation effect between the variables. It also proves that the Durbin-Watson statistic (1.95126) for this regression has increased from the previous regression which was 1.92491. From the equation it is clear that the time of flight has the highest influence since the coefficient is 35.7. The Table - 5 depicts the actual and predicted values of both fuel consumption and CO_2 emission for the available independent variables of VT-KFI during July 2009 to January 2010.

Similar regression equations where generated for all the tail registration which depicted the same result i.e. time of flight, take of weight, altitude, temperature and gravitational force were the primary factors of influence for fuel consumption. The regression equation 2, 3 and 4 of VT-KFI, VT-KFH and VT-KFJ illustrates that the identified independent variables are the major influencing factors of fuel consumption irrespective of the flight. The comparison of the actual data and predicted data for fuel consumption proves that there is a very small residual error in the proposed method of forecasting the fuel consumption for the particular tail registration (VT-KFI). The CO_2 emission for the particular tail registration VT-KFI can be found by multiplying the 3.157 (CO_2Index number) with the value generated from the regression equation with the given parameters. The CO_2 emission per kilometer per hour has been estimated and shown in Table – 5.

Fuel consumption = -1260 + 33.8 X1 + 0.0296 X2 + 0.0161 X3 + 5.14 X5- 12 X7 (3)

Fuel consumption = - 4220 + 37.9 X1 + 0.0411 X2 + 0.00585 X3 + 2.67 X5 - 9 X7 (4)

TABLE - 5 RESULT OF THE REGRESSION

Obs	Fuel Consumption(Actual)	Fuel Consumption (Predicted)	SE Fit	Residual	St Resid	CO2(Actual)	CO2(Predicted)
1	2903	2780	21.8	123	0.34	9164.771	8776.46
2	5893	6041.1	30.8	-148.1	-0.41	18604.201	19071.7527
3	4761	5026.6	36.2	-265.6	-0.74	15030.477	15868.9762
4	5167	5832.7	29.8	-665.7	-1.86	16312.219	18413.8339
5	3049	3306.6	25.8	-257.6	-0.72	9625.693	10438.9362
6	7142	6839.9	68.5	302.1	0.86 X	22547.294	21593.5643
7	4529	4708.8	30.9	-179.8	-0.5	14298.053	14865.6816
8	4328	5191	23.9	-863	-2.41R	13663.496	16387.987
9	2932	3001.9	33.2	-69.9	-0.2	9256.324	9476.9983
10	6154	6312.7	37.1	-158.7	-0.44	19428.178	19929.1939
11	2496	2146.7	32.8	349.3	0.98	7879.872	6777.1319
12	2903	2752	33.5	151	0.42	9164.771	8688.064
13	4499	4435.6	48.4	63.4	0.18	14203.343	14003.1892
14	3135	3404.3	28.6	-269.3	-0.75	9897.195	10747.3751
15	4732	4943.4	24.4	-211.4	-0.59	14938.924	15606.3138
16	5632	5222.8	21.1	409.2	1.14	17780.224	16488.3796
17	4644	4904.4	20.9	-260.4	-0.73	14661.108	15483.1908
18	4470	4398.6	25.7	71.4	0.2	14111.79	13886.3802
19	4412	4742.7	32	-330.7	-0.93	13928.684	14972.7039
20	2555	2519.8	28.3	35.2	0.1	8066.135	7955.0086
21	4703	5126.2	28.1	-423.2	-1.18	14847.371	16183.4134
22	4377	5228.2	24	-851.2	-2.38R	13818.189	16505.4274
23	7460	7230.2	39.1	229.8	0.64	23551.22	22825.7414
24	3107	3187.4	17.1	-80.4	-0.22	9808.799	10062.6218
25	4239	4892.2	41.1	-653.2	-1.83	13382.523	15444.6754
26	4557	4777.6	32.3	-220.6	-0.62	14386.449	15082.8832
27	3019	3708.6	24.4	-689.6	-1.93	9530.983	11708.0502
28	5458	5279.6	33.9	178.4	0.5	17230.906	16667.6972
29	5051	4865.8	28.7	185.2	0.52	15946.007	15361.3306
30	4470	4490	25.4	-20	-0.06	14111.79	14174.93
31	4500	4637	22.1	-137	-0.38	14206.5	14639.009

IV. CONCLUSION

A. *Summary of the Work*

In this work the emission inventory of airline industry has been studied with the impact of aircraft emissions on global warming, the effect of Green house gases and its profile. The Kyoto Protocol and its recommendations were taken into consideration to reduce the effect of carbon-di-oxide on the environment. The fuel consumption factors were statistically analyzed and the best regression model was generated. The validation has been carried out using ANOVA, Durbin-Watson statistics and Mallow's Cp statistics are used to optimize and find the predictability of the model. The regression model was used to find the CO_2 emission for the particular tail registration with various sets of independent variables.

B. *Scope of Work*

The work can be extended in the following ways

- Finding the various set of controllable parameters for the various set of uncontrollable parameters such as temperature and gravitational force for emitting less CO_2 in the atmosphere.
- Software could be developed for the performance engineers to find the various combinations of parameters where they could reduce the fuel consumption using various operational procedures which would be both economical and eco-friendly.
- A bio-fuel analysis could also be done for both ground vehicles and aircrafts to reduce the CO_2 emission from the airline industry.

REFERENCES

[1] N. C. Sekhar, APAC Airports Ramping-up for New Larger Aircraft Regional Report Asia-Pacific, ASIA-PACIFIC – 2008. Available : http://www.bangkok.icao.int/news/icao_reg_apac.pdf

[2] Press Releases No.: 37 Date: 24 October 2007 http://www.iata.org/pressroom/pr/Pages/2007-24-10-01.aspx

[3] Press Release No.: 50 Date: 24 October 2008 IATA Blasts European Union ETS Decision http://www.iata.org/whatwedo/environment/Pages/emissions-europe.aspx

[4] Gareth Horton,November 2006,Forecasts of CO2 emissions from civil aircraft for IPCC,DTI Unique Reference Number 06/2178, QINETIQ/06/02392
Available :http://www.bis.gov.uk/files/file35675.pdf

[5] Daniel l. Rubinfeld Reference Guide on Multiple Regression http://www.fjc.gov/public/pdf.nsf/lookup/sciman03.pdf/$file/sciman03.pdf

[6] Jonathan A. Baker & Daniel L. Rubinfeld, Empirical Methods in Antitrust: Review and Critique, 1 Am. L. & Econ. Rev. 386 (1999).

[7] Thomas J. Campbell, Regression Analysis in Title VII Cases: Minimum Standards,Comparable Worth, and Other Issues Where Law and Statistics Meet, 36 Stan. L.Rev. 1299 (1984).

[8] Arthur P. Dempster, Employment Discrimination and Statistical Science, 3 Stat. Sci.149 (1988).

[9] Hoa Trang Dang,June 8-9, 2006,Gaithersburg, MD 20899,Overview of duc 2006,Document Understanding Workshop,New York Marriott at the Brooklyn Bridge,Brooklyn, New York USA,Presented at HLT-NAACL 2006. Available : http://duc.nist.gov/pubs/2006papers/duc2006.pdf

[10] Stigler S (2008). "Fisher and the 5% level". Chance 21 (4): 12. doi:10.1007/s00144-008-0033-3.

[11] Fisher RA (1925). Statistical Methods for Research Workers (first ed.). Edinburgh: Oliver & Boyd.

[12] Sterne JAC, Smith GD (2001). "Sifting the evidence — what's wrong with significance tests?". BMJ 322 (7280): 226–231. doi:10.1136/bmj.322.7280.226. PMID 11159626.

[13] Schervish MJ (1996). "P Values: What They Are and What They Are Not". The American Statistician 50 (3): 203–206. doi:10.2307/2684655.

[14] ResearchConsultation.com 239-405-7165 (2007) http://www.researchconsultation.com/multicollinearity-regression-spss-collineairity-diagnostics-vif.asp

[15] Gujarati, D.N. (2003) Basic econometrics, 4th ed., Boston, McGraw-Hill,p. 469.

[16] Durbin, J., and Watson, G. S. (1950) "Testing for Serial Correlation in Least Squares Regression, I." Biometrika 37, 409–428.

[17] Durbin, J., and Watson, G. S. (1951) "Testing for Serial Correlation in Least Squares Regression, II." Biometrika 38, 159–179.

[18] Gujarati, Damodar N. (1995): Basic Econometrics, 3. ed., New York et al.: McGraw-Hill, 1995, page 605f.

[19] Verbeek, Marno (2004): A Guide to Modern Econometrics, 2. ed., Chichester: John Wiley & Sons, 2004, Seite 102f.

[20] Bhargava, A., Franzini, L., Narendranathan, W. (1982): Serial Correlation and the Fixed Effects Models, in: Review of Economic Studies, Vol. 49 Iss. 158, 1982, page 533–549.

[21] Spray drying of tomato pulp in dehumidified air: I. The effect on product recovery Athanasia M. Goula and Konstantinos G. Adamopoulos, Department of Chemical Engineering, School of Engineering, Laboratory of Food Process Engineering, Aristotle University of Thessaloniki, 541 24 University Campus, Thessaloniki, Greece Received 14 October 2003; accepted 23 February 2004. Available online 10 April 2004.Avilable: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6T8J-4C4DXVX-1&_user=10&_coverDate=01/31/2005&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=7b679a5c81bdeb3aa8e1c92aaea0ef65

[22] Mallows, C.L. (1973) "Some Comments on Cp" Technometrics, 15, 661-675.

[23] Wikipedia ,the Free Encyclopedia, Mallow's CP Available :http://en.wikipedia.org/wiki/Mallows'_Cp