Bull. Hiroshima Inst. Tech. Research Vol.48 (2014) 87-91

Article

Revisit to Modelling Global Annual Average Temperature - A Parametric Approach

Satyabrata PAL*, Arunava GHOSH**, Subhabaha PAL*** and Sanpei KAGEYAMA****

(Received August 22, 2013)

Abstract

Global annual average temperature is regarded as a precise indicator of the warming of the globe over centuries and if we track the available record on it, it is observed that the available period of 133 years (1980-2012) can be partitioned into three distinct phases, 1880-1936, 1937-1980 and 1981-2012, based on temperature variation pattern (Pal *et al*., 2013). The content of the said paper deals with the development of models, parametric and nonparametric, which produce high precision levels (as measured by value of R^2 coefficient) as high as 0.84. In this paper, a class of parametric models has been developed which is a combination of three mathematical functions, namely, cubic, exponential and trigonometric, respectively. An extensive search of models belonging to the above-mentioned class identifies four different models whose R^2 values lie within the range, 0.826 to 0.860 as a distinct improvement in the value of R^2 vis-a-vis the precision level, over that obtained in Pal *et al*. (2013).

Key Words: Global warming, Parametric models

1. Introduction

The aspect of global warming is the most dreadful environmental phenomenon the impact of which is looming large as time advances. The concern is that human activities are the principal factors triggering this alarming state to which the future of the earth is almost bound to reach unless appropriate measures are devised to annihilate the underlying effects of global warming.

Recalling the partitions mentioned in Pal *et al*. (2013), it is observed that the third phase (1981 - 2012) is the most critical period. The present study is referred to this phase, and efforts have been made to develop models which are superior (in terms of having greater precision levels, as measured by R^2 values) to the models available in the literature so far surveyed. After wide search in the quest for superior models it is found that a combination of three models (all distinct in nature), namely, cubic, exponential and trigonometric, produces higher values of R^2 coefficient as high as 0.86. The identified class (which is, indeed, a combination) contains four models denoted by M_1, M_2, M_3 and M_4, respectively. The model equations, predicted values under different models, standard errors of estimates and 95%confidence limits are given in Tables 1, 2 and 3, respectively. Table 4 contains the values of different precision criteria in respect of the above four models. The graph plots of the four models along with the data set are

* Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Faculty of Agriculture, Mohanpur, Nadia, West Bengal, 741252, India
e-mail: satyabrata.pal@hotmail.com
** Uttar Banga Krishi Viswavidyalaya, Department of Agricultural Statistics, Pundibari, Coochbehar, West Bengal, 736165, India
e-mail: arunava_ubkv@yahoo.co.in
*** SAP Development Lead, Manipal Global Education Services, Bangalore, Karnataka, 560002, India e-mail:subhabaha@msn.com
**** Department of Environmental Design, Hiroshima Institute of Technology, Hiroshima, 731-5193, Japan
e-mail:s.kageyama.4b@it-hiroshima.ac.jp

given in Annexure.

Section 2 is devoted to a description of the source of data and the method employed in the paper. Section 3 presents the results and subsequent discussions on the findings evolved in the paper. The last section, Annexure, displays the four graph-plots (along with plots of distribution of residuals) in respect of the four models respectively.

Over all, it can be mentioned that the two models, M_1 and M_2, have been identified and these two models are superior to the models obtained in Pal *et al*. (2013), diagnostic check results (on independence and normality of the estimated model errors) being satisfied in case of these two models.

2. Materials and Methods

The data sources are:

IPCC Report 2007 - http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml#1

Temperature Data Source - http://data.giss.nasa.gov/gistemp/tabledata_v3/GLB.Ts+dSST.txt

This paper has taken into consideration the parametric and nonparametric models considered in Pal *et al*. (2013). Different structures of models created as combinations of mathematical functions of various kinds are here explored.

The structure of the combination which is found as the best is of the type: Cubic + Exponential + Trigonometric, and the model expression is given below.

$$y = a + bt + ct^2 + dt^3 + he^t + q\sin^u(wt) - r\cos^v(zt) + \varepsilon,$$

where ε's are errors and follow NID $(0, \sigma^2)$, the values of u and v are fixed at 5 and 3 in case of the models, M_1, M_2 and M_3, and the values of u and v are fixed at 3 and 4 in case of the model M_4 respectively. Also, t means years.

Remark: In case of each model, various values for the pair, (u, v), are tried but the optimum value for the pair (u, v), as given above, maximizes R^2 value obtained by fitting the respective model. For each such model, the remaining coefficients are estimated till the global convergence with respect to each such is reached after successive iterations.

3. Results and Discussions

This section contains the results obtained after fitting the models to the data. These are given in four tables, Table 1 to Table 4 respectively. Also given are the results on diagnostic checking (Draper and Smith, 1998) on the errors obtained after fitting the models (vide Table 5). It is evident from the p-values (Shapiro Wilk test and Kolmogorov Smirnov test) and z-values (Run Test) listed in Table 5 that the errors (after fitting the model in each case) are independent and also are normally distributed, implying that the results on diagnostic checks are satisfied. In fact, it can be seen from Table 5 that the probability values/levels are greater than .05/.025 as the case may be, and hence the null hypotheses concerning independence and normality can't be rejected, or in other words the model errors are independent and are normally distributed.

Table 1: Estimated model equations

Model	Equations
M_1: (u=5, v=3)	$y = 14.20 - 0.01t + 0.002t^2 - 0.00004t^3 - 103\text{E-}17e^t + 0.057\sin^u(887.7t) + 0.057\cos^v(117.7t) + \varepsilon$
M_2: (u=5, v=3)	$y = 14.24 - 0.02t + 0.003t^2 - 0.00007t^3 + 1.05\text{E-}15e^t - 0.076\sin^u(-54.9t) - 0.046\cos^v(-0.957t) + \varepsilon$
M_3: (u=5, v=3)	$y = 14.19 - 0.009t + 0.002t^2 - 0.00004t^3 - 147\text{E-}17e^t - 0.04\sin^u(-45.47t) + 0.06\cos^v(0.774t) + \varepsilon$
M_4: (u=3, v=4)	$y = 14.16 - 0.006t + 0.002t^2 - 0.00004t^3 - 142\text{E-}17e^t - 0.06\sin^u(-0.83t) + 0.04\cos^v(-1.24t) + \varepsilon$

For example, $103\text{E-}17e^t = 103.10^{-17}e^t$

Table 2: Predicted values under different models

Year	Observed values	Predicted values			
		M_1	M_2	M_3	M_4
1981	14.28	14.236	14.135	14.257	14.187
1982	14.09	14.135	14.206	14.187	14.239
1983	14.27	14.178	14.304	14.129	14.196
1984	14.12	14.246	14.212	14.126	14.167
1985	14.08	14.182	14.136	14.199	14.176
1986	14.15	14.134	14.169	14.207	14.131
1987	14.28	14.245	14.213	14.216	14.206
1988	14.35	14.270	14.216	14.292	14.243
1989	14.24	14.195	14.223	14.289	14.285
1990	14.39	14.258	14.306	14.274	14.334
1991	14.38	14.371	14.295	14.266	14.274
1992	14.18	14.305	14.277	14.228	14.290
1993	14.20	14.221	14.271	14.296	14.281
1994	14.28	14.352	14.364	14.376	14.306
1995	14.43	14.440	14.378	14.385	14.396
1996	14.32	14.379	14.394	14.423	14.409
1997	14.45	14.394	14.455	14.458	14.477
1998	14.61	14.510	14.512	14.486	14.481
1999	14.39	14.479	14.456	14.457	14.453
2000	14.40	14.405	14.392	14.387	14.472
2001	14.52	14.511	14.522	14.470	14.443
2002	14.60	14.599	14.627	14.571	14.512
2003	14.60	14.533	14.604	14.558	14.573
2004	14.52	14.510	14.505	14.591	14.601
2005	14.65	14.632	14.579	14.626	14.643
2006	14.59	14.643	14.612	14.621	14.590
2007	14.62	14.561	14.569	14.603	14.583
2008	14.49	14.579	14.532	14.545	14.575
2009	14.59	14.633	14.611	14.562	14.562
2010	14.66	14.590	14.645	14.611	14.613
2011	14.55	14.546	14.568	14.596	14.592
2012	14.56	14.567	14.554	14.550	14.551
2013	NA	14.417	14.669	14.321	14.346

NA-Not available

Table 3: Standard errors of estimates and 95% confidence limits

Model coefficients	M_1			M_2			M_3			M_4		
	SE	LCL	UCL	SE	LCL	UCL	SE	LCL	UCL	SE	LCL	UCL
a	0.07	14.06	14.34	0.07	14.1	14.3	0.07	14.0	14.3	0.08	14.01	14.32
b	0.02	−0.05	0.03	0.02	0	0	0.02	0	0	0.02	−0.05	0.04
c	0.00	−0.00	0.01	0.00	0	0	0.00	0	0	0.00	0.00	0.01
d	0.00	−0.00	0.00	0.00	0	0	0.00	0	0	0.00	0.00	0.00
h	0.00	0.00	0.00	0.00	0	0	0.00	0	0	0.00	0.00	0.00
q	0.04	−0.02	0.13	0.03	0	0	0.03	0	0	0.03	−0.13	−0.00
r	0.03	−0.12	0.01	0.03	0	0	0.03	0	0	0.05	−0.15	0.07
w	0.02	887.7	887.8	0.01	−54	−54	0.02	−45	−45	0.02	−0.88	−0.79
z	0.02	117.6	117.7	0.02	−1	0	0.02	0.74	0.81	0.03	−1.29	−1.18

SE = standard error; LCL = lower confidence limit; UCL = upper confidence limit

Table 4: Values of different precision criteria for four models

Models (*e.d.f.* = 23)	MSE (mean square error)	R^2	MAE (mean absolute error)	F_o	Pr. (F greater than F_o)
M_1: (u=5, v=3)	0.00599	0.860	0.0528	17.71	> 0.0001
M_2: (u=5, v=3)	0.00610	0.858	0.0530	17.35	> 0.0001
M_3: (u=5, v=3)	0.00684	0.840	0.0593	15.14	> 0.0001
M_4: (u=3, v=4)	0.00742	0.826	0.0632	13.73	> 0.0001

e.d.f. = error degrees of freedom

Table 5: Diagnostic checking for four models

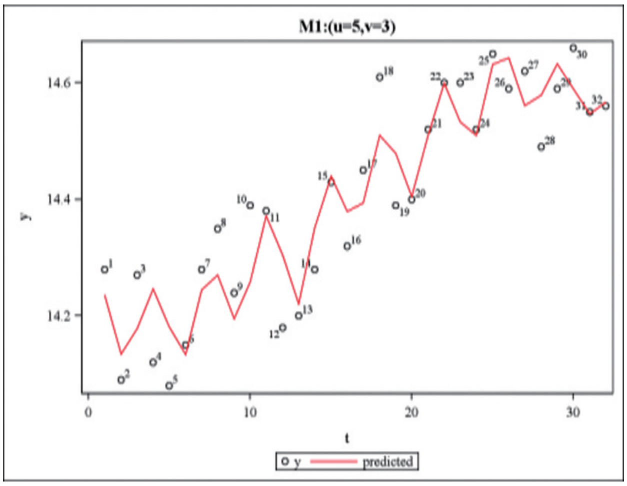
Models	Test for Normality				Test for Independence	
	Shapiro-Wilk test		Kolmogorov-Smirnov test		Run test	
	Statistic (W)	p-value	Statistic (D)	p-value	z-value	p-value
M_1: (u=5, v=3)	0.979	0.76	0.098	> 0.15	0.822	0.411
M_2: (u=5, v=3)	0.971	0.543	0.121	> 0.15	0.00	1.00
M_3: (u=5, v=3)	0.961	0.30	0.116	> 0.15	1.286	0.199
M_4: (u=3, v=4)	0.966	0.396	0.091	> 0.15	0.639	0.523

W: Value of Shapiro-Wilk test statistic. D: Value of Kolmogorov-Smirnov test statistic

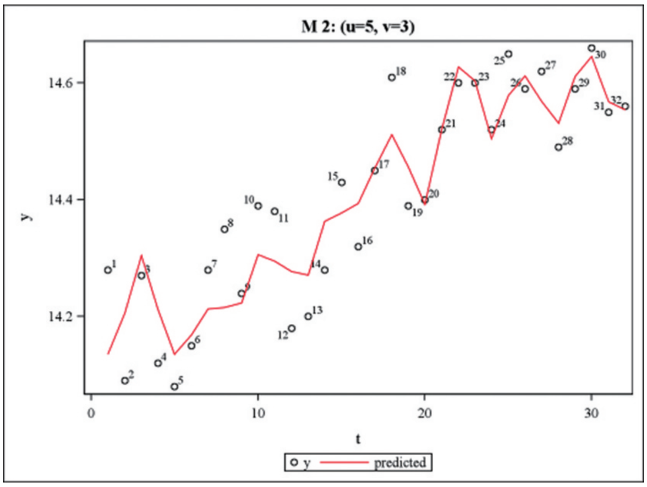
ANNEXURE

Here, the subscript, 't' denotes year, e.g., t = 1, 2, 3, ⋯ mean first, second, third, ⋯ years, respectively.

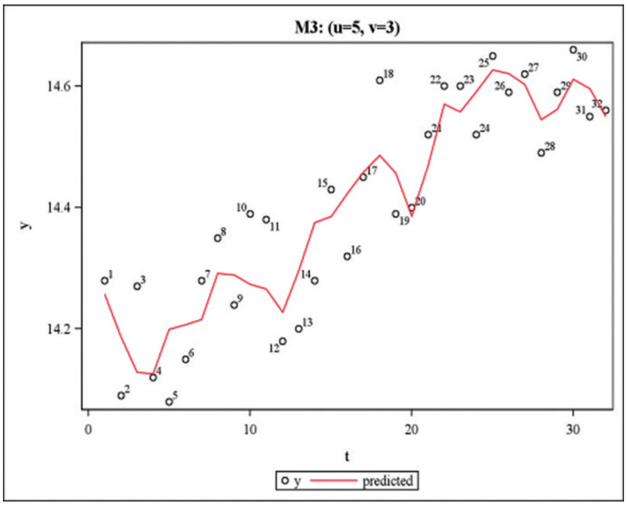
O_i represents the observed value of temperature at the i-th point of time, i.e., y_i.



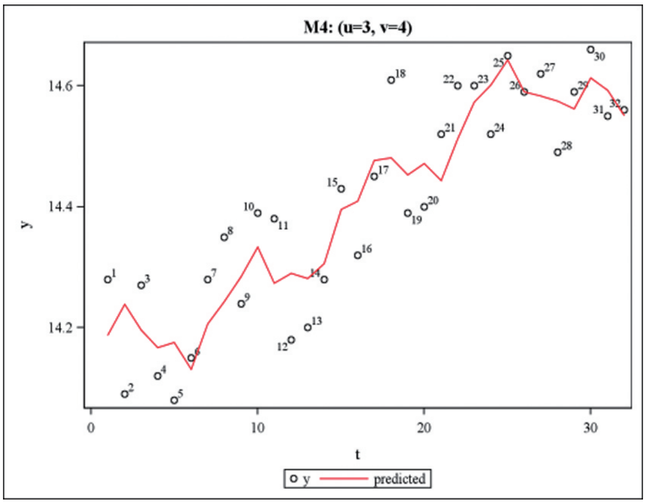
M_1



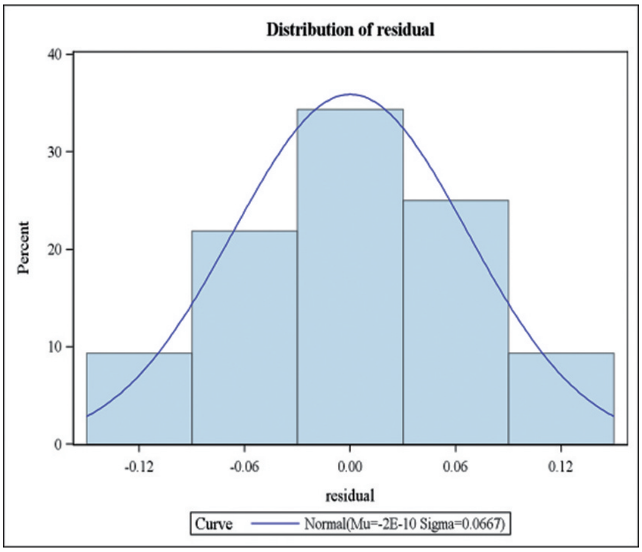
M_2



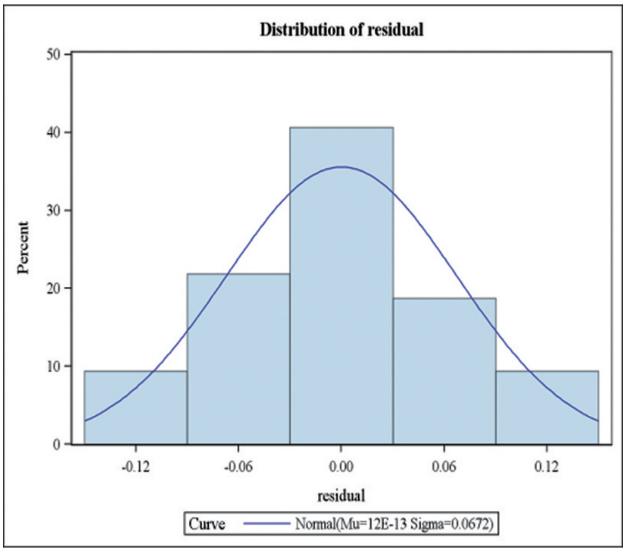
M_3



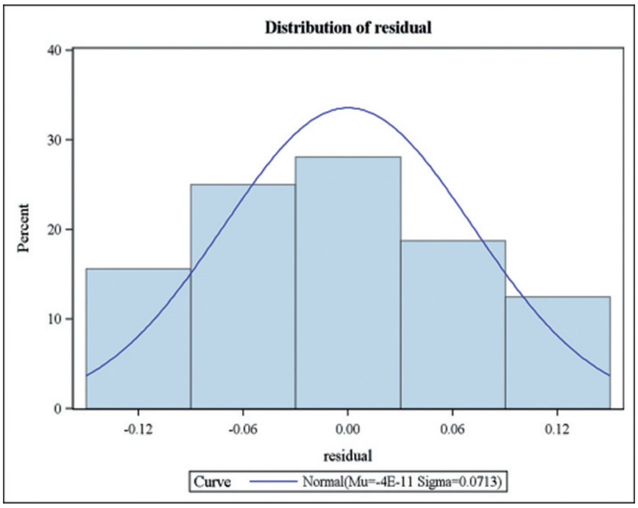
M_4



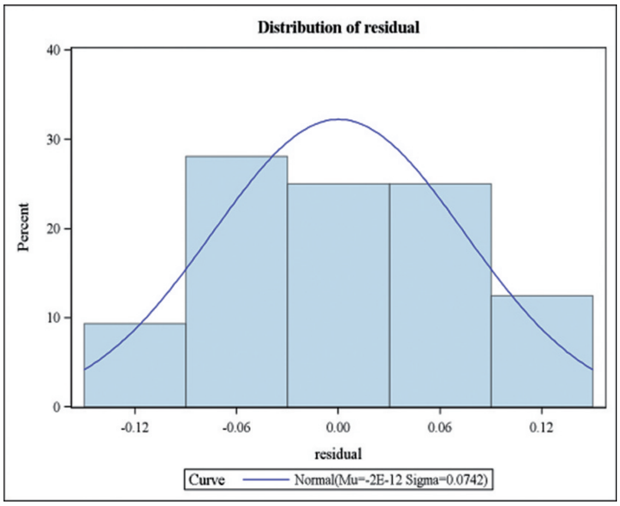
M_1 : (u=5, v=3)



M_2 : (u=5, v=3)



M_3 : (u=5, v=3)



M_4 : (u=3, v=4)

References

Draper, N. R. and Smith, H. (1998). *Applied Regression Analysis,* John Wiley and Sons.

Pal, S., Pal, S. and Kageyama, S. (2013). Modelling the Global Mean Temperature. *Bull. Hiroshima Inst. Tech. Research* 47, 149-152.