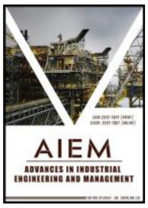
ZIBELINE INTERNATIONAL™ PUBLISHING

ISSN: 2222-7059 (Print)
EISSN: 2222-7067 (Online)

Advances in Industrial Engineering and Management (AIEM)

DOI: *http://doi.org/10.7508/aiem.01.2022.29.31*



CrossMark

REVIEW ARTICLE

FORECASTING LONG-TERM ELECTRIC POWER DEMAND BY LINEAR SEMI-PARAMETRIC REGRESSION

Shengying Zhao[1], Linfeng Zhao[2*]

[1]Newton South High School, Newton, MA 02459, USA
[2]The Physics Department, School of Arts & Sciences Boston University, Boston MA 02215, USA
[*]Corresponding Author Email: zlf975717@gmail.com

ARTICLE DETAILS

Article History:

Received 20 May 2022
Accepted 29 July 2022
Available online 5 August 2022

ABSTRACT

Market economic system put forward high request for electric power demand's forecast. The advantage and disadvantage of ancient forecast models in the world were simply summarized which had been employed in electric demand's forecast, and then semi-parametric model to forecast electric demand was brought forward to enhance the result precision and method rifeness. The semi-parametric model was constructed by the part of linear parameter and non-linear parameter. The portion orderliness knowable was reflected by the part of linear parameter. Instabilization rule was reflected by the part of non-linear parameter. So as to estimate semi-parametric model, the method of PARTIAL RESIDUAL was used. Forecast result was gained from two processes. The first, the part of linear parameter was estimated, then the part of non-linear parameter estimated too. By computing a demonstration, the forecast error of semi-parametric model is not only less but also lower than the outcome of linear regression. Computing result shows that Semi-parametric model to forecast electric power is the arithmetic of high precision, widely utilizable and computed easily.

KEYWORDS

Forecasting, Semi-Parametric, Regression, Core Function, Band

1. Introduction

Electric load predicts the method can roughly be divided into a regular analytic approach, modern analytic approach. For example, regular analytic approach goes back for one yuan, plural linear return, non-linear return, time array predict the method etc. Model of method these parameters estimate technology to be relatively ripe, predict course to be simple, but have, consider factor that load change, precision that predict a short time low, apply to load to predict in middle period (Zhao et al., 2006).

The modern analytic approach for example the expert system predicts the law binds the expert's experience knowledge with statistical method, can provide the conclusion of directionality for the development trend of load, and has overcome the one-sidedness of the single method. But expertise refines difficult, forming of knowledge base difficulty loud, especially suitable for unusually loading the long-term forecast of the mode. The Chaos predict the law utilizing electric load to look for load to change, there are less necessary data materials, can offer the predictable measuring quantitatively of system. Predict precision to be relatively high, looks reconstruction of space course about skewing and imbedding dimension to choose problem take, study further yet (Dechun, 2008).

The above can see these methods each have pluses and minuses to the simple introduction of predicting the method of domestic and international daily load. But most methods predict the precision is relatively low, adaptability is bad. Some methods estimate too the precision is high but calculate the difficulty is great, difficult practical operation, still remain to improve. Besides these methods, it is tall in accuracy whether still can find a kind of prediction, have a good adaptability, the method apt to operate?

Half a parameter abroad has returned to the theory and got sufficient development since the eighties of the 20th century, and has made better application result in predicting actually that analyzed. How is the estimation result that half a parameter model predicts to electric load, how about the result of comparing with other methods, wait a moment for question need answer of providing. These questions will be discussed here, try hard to be solved satisfied.

Congratulations! Your paper has been accepted for journal publication. Please follow the steps outlined below when submitting your final draft. These guidelines include complete descriptions of the fonts, spacing, and related information for producing your journal manuscripts. Please follow them.

2. Specification of Estimation Method for Nonlinear Semiparametric Regression Model

2.1 Specification of Nonlinear Semiparametric Regression Model

Quick Response Code	Access this article online	
	Website www.aiem.com.my	**DOI:** 10.26480/aiem.01.2022.29.31

Cite The Article: Shengying Zhao, Linfeng Zhao (2022). Forecasting Long-term Electric Power Demand by Linear Semi-parametric Regression. *Advances In Industrial Engineering And Management*, 11(1): 29-31.

Nonlinear semiparametric regression model run as in (1):

$$Y_i = f\left(Z_1, B\right) + g\left(X_i\right) + u_i \tag{1}$$

where

$$Z_1 = \left(Z_{1i}, \cdots, Z_{d_1 i}\right), \quad X_i = \left(X_{1i}, \cdots, X_{d_1 i}\right) \tag{2}$$

$f\left(\cdot\right)$ is a function already known, B is an unknown parameter vector, $g\left(\cdot\right)$ is a position function, $u_i\left(i = 1, \cdots, n\right)$ is random error sequence with a mean value of zero (Wang, 2020).

2.2 The Least Square Kernel Estimation Method for Semiparametric Regression Model

The following is an example of least square kernel estimation method for semiparametric model; its estimation principle is:

First step: supposing B is known, estimate $g\left(x\right)$ based on

$$Y_i - f(Z_1, B) = g(X_i) + u_i \tag{3}$$

Choose window width h_n, and work out estimation for $g\left(x\right)$:

$$\hat{g}\left(x, B\right) = \sum_{i=1}^{n} W_{ni}\left(x\right)\left(Y_i - f\left(Z_1, B\right)\right) \tag{4}$$

Where

$$W_{ni}\left(x\right) = K\left(\frac{X_i - x}{h_n}\right) / \sum_{j=1}^{n}\left(\frac{X_j - x}{h_n}\right) \tag{5}$$

Second step: estimate B based on

$$Y_i = f(Z_1, B) + g(X_i) + u_i \tag{6}$$

Work out the least square of B and estimate $\hat{B}$.

Third step: work out the final estimation for $g\left(x\right)$:

$$\hat{g}\left(x\right) = \sum_{i=1}^{n} W_{ni}\left(x\right)\left(Y_i - f\left(Z_1, \hat{B}\right)\right) \tag{7}$$

Fourth step: adjust the window width h_n until satisfactory (Wolwynsik et al., 2004).

3. Example Analysis

(1) Utilize SPSS to count the software and make linear return to analyze to the listed data of Table 1:

The power consumption lives per capitally: Y, Per capital GDP: x_1

$$y = 11.921 + 0.016 x_1 \tag{8}$$

With coefficient correlation Y and x1 is 0.988, the getting closer very in two relations, $r_{0.01} = 0.6226 < r = 0.988$ and $r_{0.05} = 0.4973 < r = 0.988$, it returns to that result is prominent, it is meaningful to return to the equation. 0.95 confidential intervals of returning to coefficient are (0.015, 0.018). So the first estimation of β is 0.016.

(2) Use MATLAB software programming calculation, adopt half a parameter regression model, is explained the per capita relation that lives of the variable between the power consumption Y and per capita GDP x1 is still described with the linear parameter model, the influence right of the per capita Y installed capacity x2 adopts half a parameter form to describe, i.e., $y = \alpha + \beta x_1 + g(x_2) + \varepsilon$. It utilize it is be incomplete to lean towards there aren't difference, each place an estimate at by 0.016; Unless function it β it selects for use to be kernel function K(u)=0.75(1-u2)+, it is if interlocking 0.016, determine there aren't method not to adopt, two times place an estimate at by 0.0253, the square error of fitting is 1.591 (Hall et al., 1992; David et al., 2003).

(3) Utilize SPSS to count the software and make the linear return of two yuan to analyze to the listed data of Table 1:

Table 1: Electricity information statistics

Year	Per capita (yuan/ person)	Per capita installed capacity (KW/ person)	Per capita generation (KW/ person)	power consumption lives per capacity (KW/person)
1986	901.9	0.0873	418.1	21.5
1987	1033.9	0.0941	455.0	24.5
1988	1267.1	0.1040	491.1	29.0
1989	1419.1	0.1124	518.9	33.0
1990	1547.7	0.1206	543.3	40.4
1991	1747.2	0.1308	584.9	45.9
1992	2051.3	0.1421	643.7	54.1
1993	2647.7	0.1543	706.6	62.0
1994	3654.6	0.1668	774.1	73.0
1995	4767.0	0.1793	831.4	82.9
1996	5539.3	0.1933	881.9	93.0
1997	6048.2	0.2057	917.4	101.4
1998	6373.9	0.2222	927.6	111.2
1999	6516.9	0.2373	979.4	116.7
2000	7062.9	0.2523	1081.1	132.1
2001	7516.7	0.2653	1162.7	144.1

y: Per capita life power consumption, x1: Per capita GDP, x3: Per capita generation

$$y = \underset{(-5.191)}{-38.603} + \underset{(2.287)}{0.004} x_1 + \underset{(6.927)}{0.129} x_3 \tag{9}$$

It lay being parameter estimated t value their statistic R2=0.995, $F = 1254.058 \geq F_{0.95}(2,13) = 3.81$ ones that fit error square for 3.0877. So it is prominent to return to the result, it is meaningful to return to the equation. 0.95 confidential intervals of returning to coefficient are respectively (0.000, 0.008) and (0.089, 0.169). Because of 1.592<3.0877, so the average fitting error of half a parameter model is smaller than the average fitting error of the linear regression model of two yuan.

y: Per capita life power consumption, x1: Per capita GDP, x2: Per capita installed capacity

$$y = \underset{(-8.402)}{-33.045} + \underset{(2.606)}{0.003} x_1 + \underset{(11.760)}{567.659} x_2 \tag{10}$$

Lay being parameter estimated value their statistic, R2=0.998 ones that fit error square for 1.9604. $F = 3120.676 \geq F_{0.95}(2,13) = 3.81$. So it is prominent to return to the result, it is meaningful to return to the equation. 0.95 confidential intervals of returning to coefficient are respectively (0.001, 0.005) and (463.377, 671.940). Because of 1.592<1.9604, so the average fitting error of half a parameter model is smaller than the average fitting error of the linear regression model of two yuan.

It lay being parameter estimated value their statistic, ones that fit error square for 3.0877. So it is prominent to return to the result, it is meaningful to return to the equation. 0.95 confidential intervals of returning to coefficient are respectively (0.000, 0.008) and (0.089,

Table 2: Estimate the fitting value in historical data of the per capita power consuming demand

Year	Per capita(X)	βX	G(z)	$\hat{Y}$	Y
1986	901.9	22.8180	-1.4554	21.36	21.5
1987	1033.9	26.1576	-1.7660	24.39	24.5
1988	1267.1	32.0576	-2.7702	29.28	29.0
1989	1419.1	35.9032	-1.8599	34.04	33.0
1990	1547.7	39.1568	0.0918	39.24	40.4
1991	1747.2	44.2041	1.6488	45.85	45.9
1992	2051.3	51.8978	2.1506	54.04	54.1
1993	2647.7	66.9868	-4.9868	62.0	62.0
1994	3654.6	92.4613	-19.4614	72.9	73.0
1995	4767.0	120.6051	-37.7051	82.9	82.9
1996	5539.3	140.1442	-47.1443	92.9	93.0
1997	6048.2	153.0194	-51.6195	101.3	101.4
1998	6373.9	161.2596	-50.0597	111.1	111.2
1999	6516.9	164.8775	-48.1776	116.7	116.7
2000	7062.9	178.6913	-46.5914	132.0	132.1
2001	7516.7	190.1725	-46.0725	144.1	144.1

0.169). Because of 1.592<3.0877, so the average fitting error of semi-parametric model is smaller than the average fitting error of the nonlinear regression model.

4. Conclusion

Relatively calculating indicates: In per capita electric consumption requirement forecasting, half parameter average fitting that regression model estimate error not only very much little, and the average fitting error smaller than the linear regression model of two yuan and estimate. It means that estimates the precision to electric requirement forecasting high, possess the comparative advantage in half a parameter model. Through to plan, total, charge result reveal, fit estimated value close to actual value relatively too number value. This superiority proving half a parameter model and estimating to the requirement forecasting of electric consumption again too.

Half a parameter statistical theory appears later, it is that a lot of places need the discussion to come to talk theoretically. Such as while calculating in modeling, how to choose wide this of best nuclear function and window to be worth studying. Because nuclear function and wide choice of window are not the only one, choose optimum value to possess certain difficulty and subjectivity. I think that it is wide to select the nuclear function and wide basic principle of window for use to choose different nuclear functions and windows many times, according to the result of calculation, adopt the minimal error to confirm best nuclear function and window are wide. Secondly there are many kinds of choices too in estimation methods of half a parameter model, what this selected works are selected is to lean towards incomplete difference laws to estimate, kind of estimating and two minimum laws of estimating of two stages that smooth. Because space restricts, this text has not adopted these two kinds of methods. It estimates how precision is, this is the place for worthy being further studied and calculated. Predict the error is unavoidable, in order to improve the estimation precision of half a parameter regression model, can return to half a parameter the methods and other prediction methods to mix and use, estimate the result may be better like this.

References

Dechun, A., 2008. Forecast and early-warning on coal supply and demand balance in China. China Mining University, (4), Pp. 21-5.

David, R., Wang, M.P., Carroll, R.J., 2003. Semiparametric regression. Cambridge University Press, Pp. 21-26. http://dx.doi.org/10.1017/CBO9780511755453

Hall, P., Marron, J.S., Park, B.V., 1992. Smoothed cross validation. Probability Theory and Related Fields, (90), Pp. 149-173.

Wang, Y., 2007. Prediction of the demand for coke on China's markets in 2010-2020. Metallurgical Information Guide, China, (5), Pp. 39-43.

Wolwynsik, R., Thierry, L. R., Prpnzato, L., 2004. Minimum entropy estimation in semiparametric models. IEEE International Conference, NewYork: IEEE, (ii), Pp. 1045-1048. https://doi.org/10.1016/j.sigpro.2004.11.028

Zhao, L., Han, G., Li, Y., Wang, Y., 2006. Forecasting Electric Power Demand by Semi-parametri model. China Security Science Journal, (8), Pp. 17-21.