

Prediction of Voter Number using Top Down Trend Linear Analysis Model

Yulyani Dewi

Department of Management
Technology
Institut Teknologi Sepuluh Nopember,
Indonesia

Nani Kurniati

Department of Industrial Engineering
Institut Teknologi Sepuluh Nopember,
Indonesia

Suhartono

Department of Statistics
Institut Teknologi Sepuluh Nopember,
Indonesia

Abstract:-The logistics planning of ballots for General Election (GE) has been estimating 1% increase from previous Voters List (named as Daftar Pemilih Tetap, DPT). This assumed percentage easily adopted from increasing percentage of population as described in Report of Aggregate Population (RAP). In fact, there are several factors different between DPT and RAP. DPT isn't the number of overall population as the RAP states; DPT is the number of potential voters. There are some restriction for who become a voter i.e. more than 17 years old or have married, not be a member of National Army/ Police unless they have retired, must domicile in one area only and may not be registered in two areas, and who are eligible (alive) at the time of GE takes place. DPT becomes the main input for GE Commission, as the basis for ballots logistics planning. The audit from Financial Audit Board (FAB) founded under Reasonable with Exception status, means there isn't optimal absorption budget for logistics procurement at previous GE 2014 which is only 59.44%. One of the reasons may due to over estimation. Therefore, this paper proposed a method for DPT prediction by using Trend Analysis Forecasting Method i.e. Linear Trend and Non Linear Trend. Considering both Bottom Up and Top Down approach for DPT calculation in one province who has several districts, result the Top Down Linear Trend Model is the best method with 3.22% saving in aggregate financial budgeting plan for GE Commission.

Keywords:-Election, DPT, Forecasting, Logistic, Ballots.

I. INTRODUCTION

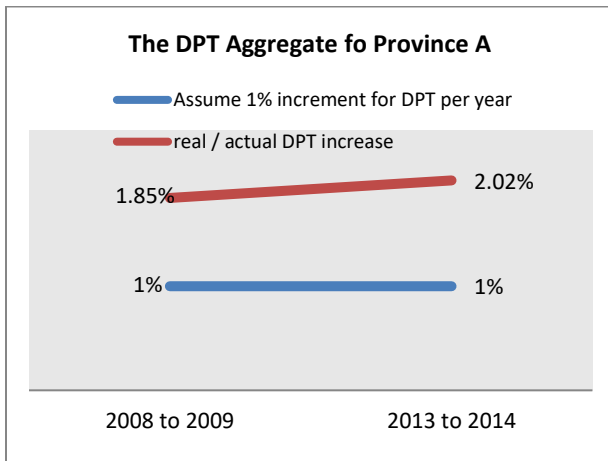
The role and function of the General Election Commission as the General Election Organizer requires accurate data to conduct the GE. GE has characterized as public-free-confidential, honest, and fair conducted in every five years. In the process of organizing the GE, it is required a lot of cost. Currently, the procurement and distribution of ballot logistics has become a measurement of an effective planning as well as the basis for calculating cost of election process. The task of procurement planning division are divided into four tasks and distributed among three level of GE Commission, as follows:

- Procurement of updated Voter List Form, Voting Form, and Vote Count Form are responsibility of the Province GE Commission. Include budget for procurement of ballot, ballot box, voting booth, ink, blind aid, as well as another form needed
- Logistics Distribution costs are budgeted at Districts GE Commission
- While honorarium is planned at all level of GE Commission at National, Province, and District level.

DPT becomes the main input for GE Commission. Problems arise when budget increased year to year; cause inefficient budget as reported in National Budget Report (named as APBN). One of the reasons of the increasing budget due to logistic procurement was planned under assumption of 1% increase from the last DPT. DPT is the basis for ballots logistic planning to determine the budget. Therefore, testing on the above assumption needed. For instance, the following data from the previous two election process; Governor Election and President Election where DPT are generated from the last DPT on 2008-2009 for DPT on 2013-2014. It can be seen from figure 1 that the assumption still questionable. The actual DPT data of Governor Election in 2013 and Presidential Election in 2009, as well as the Governor in 2013 and the President Election in 2014 are different from the predicted 1% increase. Therefore, seems it isn't appropriate if the DPT is predicted under the 1% increase assumption.

The DPT has different characteristics with the population on some aspects. The potential voters require certain consideration as follows:

- who is 17 years old or already married;
- not a member of the Indonesian Army / Police;
- who have retired from members of the Indonesian Army / Police;
- domicile in one district only or may not be registered in two district; and
- be an eligible (alive) residents at the time of the General Election.



Another proof that DPT is predicted less accurately since the budget of IDR 56 million only consume IDR 33.milion or 59.44%, means over estimate. Based on the budget calculation pattern during this DPT has a price, to determine the amount of budget to be planned, for example in 2014 the price of 1 DPT equal to IDR 45thousand from the total budget of Province A 1,4 trillion divided by the total number of DPT in 2013 which is increased by 1% according to the assumption of increase of 30 billion.

Another issue to be considered is that DPT data is not a time series data that is easy to forecast or predicted, because the DPT data is only obtained once every five years when there is a legislative or executive election or the election of regional head and vice regional head. So that the available data is 5 years, as the following table:

Figure 1 Percentage of Aggregate Increase Between DPT and Factual Prediction of Province A

Num	Years	DPT of Governor Election (Pilkada)	DPT of the Presidential and Vice Presidential Election	DPT of regents and vice regents or mayors and deputy mayors	annotation
1	2013	√	-	-	Gub. Election
2	2014	-	√	-	President Election
3	2015	-	-	√	Regent/Mayor Election
4	2016	-	-	-	None Election
5	2017	-	-	√	Regent/Mayor Election
6	2018	√	-	√	Gub. & regent/Mayor Election (simultaneous election)

Table 1. Implementation of General Elections and Governor Election in Province A

2013	2014	2015	2016	2017	2018
30,034,249	30,639,897	18,771,960	-	147,975	30,963,078

Table 2. 2013 (actual), 2018 (adjudgment Plenary General Election ProvJatim) and Pilkada 2010, 2011, 2012, 2015 and 2017 (actual)

Considering that DPT is the source that determines the estimation of budget calculation, it is very important to do a way to improve the prediction pattern on DPT so that the result can be accurate and reliable.

II. METHODS

A. Trend Analysis

Trend analysis according to Zaitun Software Developer Team (2010) is a common trend model for time series data for forecasting. Trend analysis is an analysis used to observe the overall trend of data over an extended period of time.

Trend can be used to predict what condition the data in the future, or can be used to predict the data at a time in a certain time. Some methods that can be used to model the trend, one of which is the Linear Method.

The Linear Method is a method that uses data that randomly fluctuates into a straight line. The linear trend is the tendency of data where the time-based changes are fixed (constant).

Makridakis *et al.* (1993) introduced a new concept in an attempt to make predictions. A prediction will be expressed as a function of the number of factors that determine the outcome of the prediction. Divination does not have to depend on time. Sometimes the astrologer will meet with a measure or unrelated variable (Y) and the independent variable (X) in a simple regression of Y against X. Sometimes also meet the dependent variable (Y) and the independent variable (X₁, X₂, X_k) whose purpose is to find a function to connect Y to all independent variables called multiple regression.

An important difference needs to be established between linear regression models with non-linear regression. All regression models are always written as equations that connect the dependent and independent variables, written in the general form as follows :

$$Y = a + bX \tag{1}$$

First, a and b are coefficients where the equation Y is a linear function of X because if Y we plot (plot) against X will be a straight line.

Second, the equation is linear to the coefficients.

In forecasting using regression, the forecaster must decide how many variables are encountered, which variables should be defined as the dependent variable and which are the independent variables and what functional form to choose. If the data is measured by time, then the functional form is called period series regression, but if there is no time series then the functional form is called cross-sectional regression.

The model of single equation is more limited than the models of multiple equations (simultaneous). Single equations are also

more subtle and may be sufficient for many uses. In the hands of a fortunate forecaster some variations on conventional regression techniques can be used to provide good forecasting results.

In modeling relationships within a set of variables, we need to separate the other variables that interfere with or influence the measurement. This separation process is called 'partialling out of variable'.

B. Simple Y Regression Against Time

The principle of the calculation is directly, if we use Y as the independent variable and X = t as the independent variable, then the goal to be achieved is to obtain a straight line equation :

$$\hat{Y}_t = a + bt$$

Where :

a = intercept

b = inclination

such that for any given t time value, the error is aquaticate :

$$(Y_t - \hat{Y}_t)^2 = e_t^2 \tag{3}$$

If the sum will result in a minimum total. This is a Least Square procedure and an error is expressed as the length of a vertical line and a particular point to the line (a + bt).

For observed values Y is modeled in the form of a pattern and error :

$$Y = \text{pattern} + \text{error}$$

$$Y = a + bt + e$$

$$Y = \hat{Y} + e \tag{4}$$

That if the pattern (data) is expressed as Y, then the pattern translator is declared as the estimator or the value estimator \hat{Y} . The determinants of slope coefficients b for simple linear regression as in (2.1) are as follows :

$$b = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum X^2 - (\sum X)^2} \tag{5}$$

C. Testing Forecasting Model

- ❖ Average Error - AE, or Mean Error - ME.

$$ME = \frac{\sum_{i=1}^n e_i}{n-1} \tag{6}$$

- ❖ Mean Absolute Percentage Error – MAPE

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \tag{7}$$

- ❖ Mean Square Error – MS

$$MSE = \frac{\sum (e_i)^2}{n-1} \tag{8}$$

III. RESULT AND DISCUSSION

A. Forecasting Results with Bottom up Model Trend Linear Analysis Model

The result of forecasting or prediction using Linear Trend Method with Bottom Up model is aggregate prediction model which obtained from counting prediction in each new Regency / City then summed to get total Prediction as Provincial DPT

Prediction data. This linear model can only be applied in 20 districts / cities in Province A that is for Pacitan, Ponorogo, Lumajang, Probolinggo, Pasuruan, Mojokerto, Magetan, Ngawi, Bojonegoro, Lamongan, Pamekasan, Sumenep, Kediri, Blitar, Malang, Pasuruan City, Mojokerto City, Madiun City, Surabaya City. Describe the development of DPT patterned linear trend in the district / city. So from these results DPT can be grouped based on the map as follows:



Figure 2 Bottom up Group for Linear Trend Model

B. Forecasting Results with Top Down Model Trend Linear Analysis Model

Data of Aggregate / Overall DPT Factual of Province A predicted with Top Down Trend Linear Model based on result Prediction. So if grouped into 3 categories will be shown in the map as follows:



Figure 3. The Top Down Group for the Linear Trend Model

From the above grouping results obtained the grouping for Bottom Up and Top Down Model as follows:

- Bottom Up and Top Down Model for Total Prediction of DPT less than 500 thousand equals in 8 districts / cities same as Pacitan, Kediri, Blitar, Probolinggo, Pasuruan, Mojokerto, Madiun, and Batu
- Bottom Up to Total Prediction DPT between 500 thousand - 1 Million is in 17 districts / cities namely Ponorogo, Trenggalek, Tulungagung, Lumajang, Bondowoso, Situbondo, Probolinggo, Mojokerto, Nganjuk. Madiun, Mageetan, Ngawi, Tuban, Gresik, Pamekasan, Sumenep, and Malang, while Top Down Model is in 18 regencies / cities, Ponorogo, Trenggalek, Tulungagung, Lumajang, Bondowoso, Situbondo, Probolinggo, Mojokerto, Nganjuk. Madiun, Mageetan, Ngawi, Tuban, Bangkalan, Sampang, Sumenep, and Malang are different in Gresik and Pamekasan

- Bottom Up for Total Prediction of DPT more than 1 Million is in 13 regencies / cities covering Blitar, Kediri, Malang, Jember, Banyuwangi, Pasuruan, Sidoarjo, Jombang, Bojonegoro, Lamongan, Bangkalan, Sampang and Surabaya in 12 districts / municipalities include Blitar, Kediri, Malang, Jember, Banyuwangi, Pasuruan, Sidoarjo, Jombang, Bojonegoro, Lamongan, Gresik, and Surabaya are not the same in Bangkalan and Sampang

C. *Comparison Between Top Down Models and Bottom Up Models*

After all the results in both models between the Bottom Up Linear Model and the Top Down Linear Model made comparisons, which is the best result, which will be used to predict DPT 2018 and 2019. The comparison yields the best forecast as provided in table 3. The result of the comparison of the precision test is the best data obtained from the Top Down Model.

YEAR	PROVINCE A TOP DOWN	THN	PROVINCE A BOTTOM UP
2013	30,034,249	2013	26,590,818
2014	30,639,897	2014	30,730,259
2015	31,245,545	2015	31,320,045
2016	31,851,193	2016	31,927,425
2017	32,456,841	2017	32,553,474
2018	33,062,489	2018	33,199,347
2019	33,668,137	2019	33,851,429
MSE =	-	MSE =	5,932,692,708,013
RMSE =	-	RMSE =	2,435,712
MAPE =	0	MAPE =	0.0588
	0.00%		5.880%



Table 3: Comparison with accuracy test between Model top Down and Bottom Up

D. *Validation*

The best results are then validated using data from other sources ie BPS to get the validity as provided in table 4. And the result of validity shows that the best DPT predictions are quite valid because the percentage of matches is less than the existing match, and the percentage value is still above 50%.

DATA	CATEGORY					
	Amount < 500.000		Amount 500.000 – 1.000.000		Amount >1.000.000	
BPS 2015	9	District/city	22	District/city	7	District/city
PREDICTION OF DPT 2015	8	District/city	20	District/city	10	District/city
PREDICTION OF DPT 2019	8	District/city	17	District/city	13	District/city
VALIDATION BPS VS PREDICTION OF DPT 2015	89%	Match	91%	Match	57%	Match
	11%	Not Match	9%	Not Match	43%	Not Match
VALIDATION BPS VS PREDICTION OF DPT 2019	89%	Match	77%	Match	57%	Match
	11%	Not Match	23%	Not Match	43%	Not Match

Table 4. Validation of Statistics Indonesia Data with DPT Prediction 2015 and 2019

E. Best Predicted Aggregate Results DPT 2018 and 2019

Using Top Down Linear Trends Analysis The model is known to predict DPT Province Aggregate results in 2018 is 31.197.902 and 2019 is 31.380.624

F. Percentage Budget Savings

Through the best DPT prediction results with Top Down Trend Analysis Model then can be calculated approximately how the budget can be saved in aggregate as around 30 billion people, with ballot price around 45 thousand, we get around 1,386 trillion. means get saving around 46 million or around 3,22 %.

From the calculation can be seen that there is a budget savings from the results of the method used better than just assuming the increase of DPT 1% per year from the Last DPT, and in dollars that can be sparing in aggregate amounted to 46 million.

IV. CONCLUSION AND SUGGESTION

A. Conclusion

1. Whereas the DPT planning with the assumption of 1% increase which has been done by the General Election Commission as the organizer of the National General Election, Provincial General Election, and Regency / City General Election is inappropriate.

2. Considering the final result is no longer the percentage of increase that is used as a benchmark The best prediction, but Prediction model that determines.
3. Different models for the corresponding DPT Prediction in each Regency / City in Province A, can be considered as evidence of plurality of population in Province A, which is spread fairly evenly.
4. improvement of this prediction model of aggregate budget of Province A Provincial General Election can be saved by 3.22%.
5. the best predicted outcome is the Top Down Trend Analysis Linear Model to predict DPT in 2018 and 2019.

B. Suggestion

a). For the Election Commission of Province A

- in the future DPT predictions may be considered to use the model that has been tested in this study.
- Before making predictions in advance the General Election must determine the pattern of data, so that the best model can be adjusted to varying data in each District / City.
- The results of this study will be used as suggestions and suggestions of authors in the plenary meeting of the General Election in March 2018.

b). For Relations and Stake Holder of General Election

- DPT prediction results can also be used to map votes for candidates.

- DPT can be used to map Party Political strength so as to further optimize its socialization / campaign schedule in areas that are expected to be the basis of the vote.

REFERENCES

- [1]. Assauri, S. (1984). Teknik dan Metode Peramalan. Fakultas Ekonomi UI, Jakarta.
- [2]. Buku Pintar Logistik Pemilihan Umum oleh GENERAL ELECTION Republik Indonesia
- [3]. Draper, N. R., Smith, H., & Sumantri, B. (1992). *Analisis regresi terapan*. PT Gramedia Pustaka Utama.
- [4]. Frechtling, D. (2001). *Forecasting Tourism Demand 1st Edition*, Routledge
- [5]. Fauziah Gitri D, Irmaya Fatwa Y.(2011). Laporan Resmi Modul 4 Estimasi Parameter Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember, Surabaya
- [6]. Gilliland, M., et al. (2016). *Business Forecasting: Practical Problems and Solutions*, John Wiley & Sons.
- [7]. Gaspersz, V. (2004). *Production planning and inventory control*. PT Gramedia Pustaka Umum, Jakarta.
- [8]. Hanke, J.E., Wichern, D.W., Reitsch, A.G. (2001). *Business Forecasting, Seventh edition*. Prentice Hall, New York
- [9]. Hanke, John E. (2014) *Guide for : Business Forecasting :9th edition*. Cram101
- [10]. Hill, McGraw. (2009). *Aczel-Sounderpandien : Complete Business Statistics 7th Edition*. Aczel-Sounderpandien
- [11]. Hyndman, R.J. dari Monash University di webnya: <https://robjhyndman.com/talks/Hierarchical.pdf>
- [12]. Kuntoro, H. (2015). *Teori Dan Aplikasi Analisis Seri Waktu*, Zifatama Publisher.
- [13]. Kemenkeu, <http://www.kemenkeu.go.id/sites/default/files/Pajak%20dan%20Pemilu.pdf>.
- [14]. Makridakis, S., et al. (1999). "Metode dan aplikasi peramalan." Binarupa Aksara. Jakarta.
- [15]. Makridakis, S., dan Wheelwright, S. C., McGehee, V. E. (1993). *Metode dan Aplikasi Peramalan. Jilid 1. Edisi Pertama*. Terjemahan Hari Suminto. Binarupa Aksara, Jakarta.
- [16]. Reksohadiprodjo, S. (1988). *Business forecasting*. BPFE. Yogyakarta.
- [17]. Rizal Zaini Ahmad Fathony and Zaitun Time Series Developer Team et al, Zaitun Time Series, retrieved on 14/12/14 from http://zaitunsoftware.com/system/files/zaitun_TS_ENbooklet.pdf
- [18]. RJ Hyndman, RA Ahmed, G Athanasopoulos, and HL Shang (2011). "Optimal combination forecasts for hierarchical time series". *Computational Statistics and Data Analysis* 55(9), 2579–2589
- [19]. RJ Hyndman, RA Ahmed, and HL Shang (2013). [hts: Hierarchical time series](https://cran.r-project.org/package=hts). cran.r-project.org/package=hts.
- [20]. RJ Hyndman and G Athanasopoulos (2013). *Forecasting: principles and practice*. OTexts. [OTexts.org/fpp/](https://otexts.org/fpp/).
- [21]. Subagyo, P. 1986. *Forecasting Konsep dan Aplikasi*. BPFE, Yogyakarta
- [22]. Suhartono. 2013. *Analisa Data Statistik dengan R*. Laboratorium Komputasi ITS. Surabaya.
- [23]. Indah, P. L., Suwignjo, P. dan vanany, I. (2002), "Perancangan sistem pengukuran kinerja dengan Balanced Scorecard pada PT ISN", Seminar Nasional Teknik industry dan Manajemen Produksi II
- [24]. Kaplan, R. S., dan Norton, D. P., (1996), "Balanced Scorecard: Menerapkan Strategi Menjadi Aksi", Erlangga, Jakarta.
- [25]. Mulyadi., (2001), "Akuntansi Manajemen : Konsep, Manfaat dan Rekayasa, Edisi Ketiga", Salemba Empat, Jakarta.
- [26]. Rahman, A. (2012). "Perencanaan Pengukuran Kinerja Di Lembaga Pendidikan Walisongo – Gempol dengan Menggunakan Balanced Scorecard (BSC) dan Analytic Hierarchy Process (AHP)" Tesis: MMT-ITS. Surabaya.
- [27]. Saaty, T. L. (1988). "Multicriteria Decision Making - The Analytic Hierarchy Process", United States America.: Beccles Sufflokk Printed and Bound.
- [28]. Simons, R (1999), "Levers of control: How managers use innovative control systems to drive strategic renewal, 5th edition", Harvard Business School Press, Boston, Massachusetts.
- [29]. Vanany, I. (2009). "Performance Measurement: Model & Aplikasi",: ITS Press, Surabaya.
- [30]. Whittaker, James B., (1993)., "The Government Performance and Result Act of 1993; A mandate for Strategic Planning and Performance Measurement", Educational Service Institute, Virginia, USA.