# A STUDY ON LOAD CURRENT FORECASTING USING STATISTICAL ANALYSIS

Toshihiro Nagumo
TEPCO Holdings, Inc. – Japan
nagumo.toshihiro@tepco.co.jp

Hirokazu Ito
TEPCO Holdings, Inc. – Japan
ito.hirokazu@tepco.co.jp

Tsuneyo Sano
TEPCO Holdings, Inc. – Japan
sano.tsuneyo@tepco.co.jp

## *ABSTRACT*

*Building facilities for distribution systems and operation of them need to be considered based on the amount of load current of the consumers receiving the supply from the distribution system. On the other hand, due to the rapid extension of interconnection to distribution systems with distributed power sources such as photovoltaics(PV) in recent years, both load current and power generation current coexist, making it difficult to gather precise information about consumer load current. In this context, the authors are considering a method to individually forecast load current using statistical analysis.*

*At present, sensor-built-in switch called ITSW, which is aimed at improving reliability of supply and maintaining/improving quality of electric power, has gradually been introduced to distribution systems in Japan, allowing us to obtain current values not only at conventionally measured threshold points in distribution systems but also at any points where the ITSW is installed (at the midpoint, near the terminal, etc.). Therefore, in this paper we examined the accuracy of load current forecast in a specified zone in a distribution system using the load current forecast method we are proposing and ITSW measurement data.*

## INTRODUCTION

Tokyo Electric Power Co. has measured the distribution line current value at a distribution substation and used the measurement history information to obtain load current in each distribution system in the past. However, interconnection to distribution systems with distributed power sources such as photovoltaics in recent years is increasing at a rapid rate, and load current and power generation current coexist in the current measured in interconnected distribution systems as shown in Fig. 1, making it difficult to obtain measurements separately for each. In actual practice, the current values forecast based on empirical rules of the operators are used in many cases. However, at the time of temporary power interchange between distribution systems using a switch installed in them, or for new commercial-scale consumers or interconnection to distributed power sources, a method for theoretically forecasting load current is important.

When building facilities for power distribution systems, expansions of the facilities are planned based on the annual maximum load current forecast for each system, but if we cannot obtain accurate load current, it may lead to overloading the facilities or excessive capital investment.

In addition, for the operation of a distribution system it is important to obtain the load current in the unit of "section" divided by automatic switches as shown in Fig. 2. A multi-sectionalized, multi-connection method composed of multiple sections is commonly adopted in distribution systems in Japan. By opening and closing automatic switches by remote control, system configuration can be changed for the purpose of maintaining the quality of electric power or power can be sent from a nearby system at the time of a power outage or accident until the system is restored. Therefore, it is important to obtain accurate load current not only per distribution system but also per section. A switch called ITSW equipped with CT or VT has gradually been introduced since FY 2012 in order to obtain detailed current/voltage information within a system, but the current value calculated by ITSW is again the total of the load current and power generation current, so the above-mentioned problem may occur. Therefore, we will need to have a method to forecast load current in a distribution system for the future, too.

## METHOD FOR FORECASTING LOAD CURRENT

First, characteristics of daily value of load current measured in a distribution system and the proposed forecasting method are introduced in greater detail. As you can see in Fig. 3, the load current of consumers receiving the supply from a distribution system changes in a characteristic way, depending on how electricity is used by each consumer, etc. [1]. The proposed method makes forecasts by statistical analysis using information about "type of contract" and "contract capacity", which are determined when the power vendor and consumers sign the power supply contract.

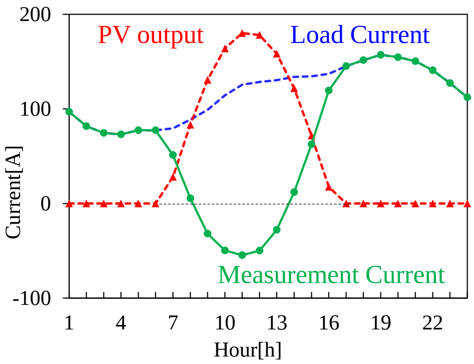



Fig.1 Example of measured current at distribution system that PV is interconnected

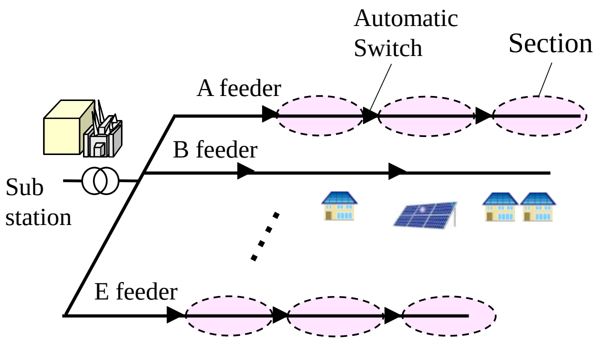



Fig.2 Image of distribution system

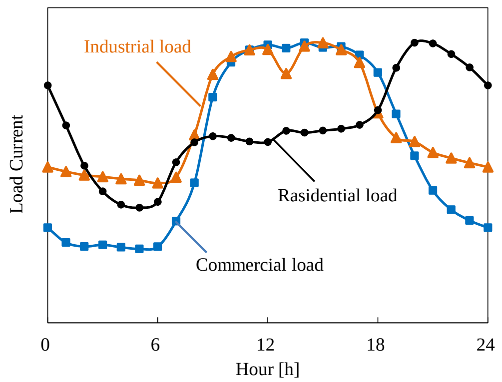



Fig.3 Example of load current

The contract type is believed to reflect characteristics of the form of demand (residential load, industrial load, etc.). Our method classifies distribution systems into different categories by cluster analysis, which is a statistical analysis technique, based on contract capacity ratios of each contract type shown in Table 1. Then, a forecast model is made for each classified pattern. In this method, the load current every on-the-hour is forecast separately on weekdays and holidays for each month during the year. There are different forecast models for weekdays and holidays, because it is known from statistical analysis that they show different load curves [2]. This paper creates the forecast models and evaluates the forecast accuracy using measurement data from distribution systems with less interconnection to distributed power sources.

## Classification of contract capacity

An example of the creation of a forecast model for average weekday current is shown below. First, to make the forecast model, we extracted 184 distribution system lines (568 sections) with less PV interconnection (100 kW or less) in TEPCO's service area. Next, in order to obtain the load profile of each section, four composition ratios of contract capacity are calculated. Although the distribution systems have different demand characteristics such as housing zone, commercial zone, and industrial zone depending on the composition ratio by the contract classification, the sorting method and threshold values have not been revealed. However, in creating the forecast

Table.1 Contract type

| Supply voltage | Contract type |
|---|---|
| Low voltage | Lighting service |
| | Low voltage power service |
| High voltage | Commercial power service |
| | High voltage power service |

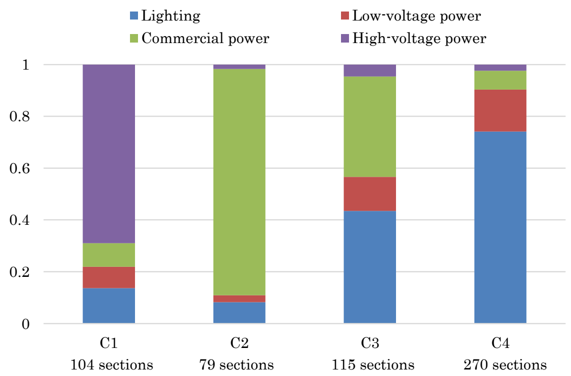



Fig. 4 Category composition ratio of each cluster.

models, the sorting methods and threshold values need to be clarified. Then, in this method, using these composition ratios of contract capacity as parameters, the load patterns are clustered by a non-hierarchical clustering technique called the k-means method. However, the k-means method is dependent on the initial value, so four clusters, which are determined from a dendrogram created by a hierarchical clustering technique called Ward's method with normalized monthly average current values of the target distribution lines, and the composition ratios of contract capacity averaged out for each cluster are used as the initial number of clusters and the initial value of cluster center to enter, respectively.

In Fig.4, the final values of cluster center for each cluster classified by the k-means method are shown. Cluster 1 (C1) is high-voltage power, Cluster 2 (C2) is power for industrial use, Cluster 4 (C4) is classified into a section with many contracts for electric light. Cluster 3 (C3) is classified into the section with the contract for electric light and for power for industrial use.

## Multiple regression analysis

Next, create the forecast models by obtaining partial regression coefficients for Formula (1) by multiple regression analysis, using the contract capacity in the distribution system clustered for each section as the explanatory variable and the monthly average value of system load current to be forecast as the response variable. The correlation coefficients of 16 partial regression coefficients (4 contract types multiplied by 4 clusters) obtained for each hour are 0.97 or greater at all time slots, indicating that these correlation coefficients can fully explain the average value of distribution line current, which is the response variable [3].

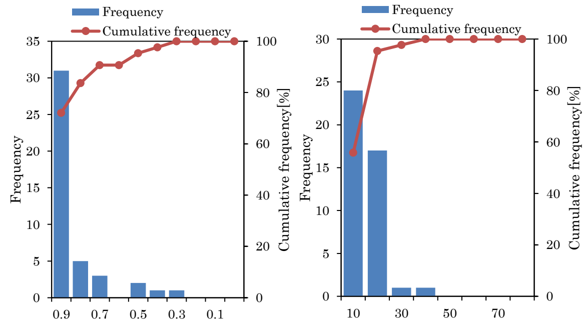
$$Y^{(t)} = \sum_{c=1}^{4}\sum_{w=1}^{4} a_{c,w}{}^{(t)} X_{c,w}{}^{(t)} \qquad (1)$$

$Y$ : Average value of distribution line current
$a$ : Partial regression coefficient
$X$ : Contract capacity
$c$ : Cluster (1 to 4)
$w$ : Contract type (electric light, power, high-voltage power, power for industrial use)
$t$ : Every on-the-hour (1 to 24).

## Results of distribution line current forecast

We forecast distribution line current of the distribution system and evaluated accuracy by extracting 49 lines different from the 186 distribution lines used to generate the forecast models. The mean error of absolute value was 24.0A and absolute error ratio was 18.6% for all time slots. We calculated the correlation coefficient between the actual measurements and forecast values for each distribution system and the standard deviation of the forecast error (24 hours) to evaluate reproducibility of the shapes of load curves. Fig. 5 shows their histogram. In the same Figure (a), the correlation coefficient is 0.8 or greater in 84% of the distribution systems, and in the same Figure (b), the standard deviation is 20A or less in 95% of the distribution systems. We can see that the variation in the forecast error is also small for each time slot. In another study, we investigated the plus / minus sign of the forecasting error for each hour and confirmed that the plus or minus sign is biased in the 78% distribution system [3]. Therefore, on the whole, the measured value and the forecasted value have errors in the same direction for the whole time, and the error value also changes over time with almost the same value, that is, the load curve shape is almost reproduced. Therefore, the shape of the load curve is reproduced, but for the distribution systems with deviations between the actual measurements and forecast values, we may be able to improve forecast accuracy by correction with a uniform value for all time slots. Accordingly, taking into account forecasting actual load current of PV interconnection distribution lines, we performed the corrections by uniformly adding the night-time mean error value between 19:00 ~ 5:00, which is outside the time slot for PV power generation, to the forecast values in this method.

We forecast the average weekday current in June 2014 in the 49 distribution system lines for the evaluation. Fig. 6 shows the histogram of the forecast the error of absolute values after the corrections. From the same Figure, the forecast accuracy was with the mean error of absolute values of 50A or less in 88% of the distribution systems before the corrections, but it was with the mean error of absolute values of 20 A or less in 88%, the mean error of absolute values for all time slots was 11.6 A and the error ratio of absolute values was 9.7% after the corrections. We found that the forecast accuracy improves with the




(a) Correlation coefficient of measured and forecast current (b) Standard deviation of forecast error[A]

Fig. 5. Histogram of correlation coefficient and standard deviation.

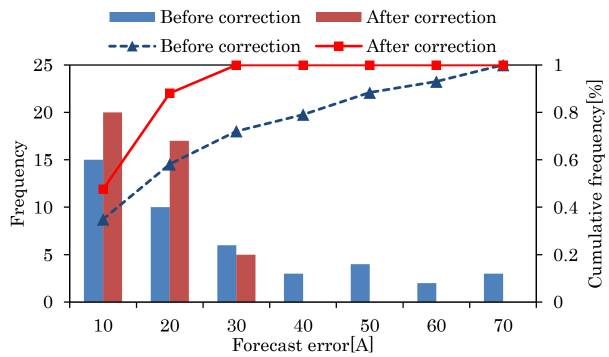



Fig. 6. Histogram of absolute error of before and after correction.

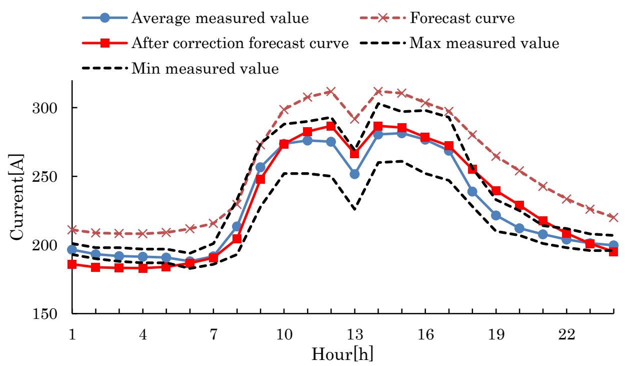



Fig. 7. Example of load forecast by forecast model in weekday of June.

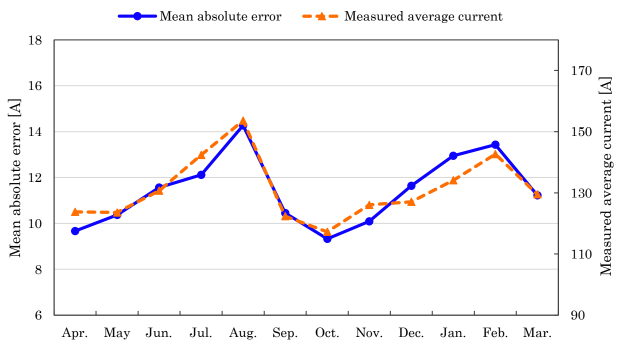



Fig. 8. Monthly results of forecast error.

corrections using the night-time mean error.

Fig. 7 shows an example of the forecast of the load curve for weekdays in June with the above-mentioned corrections. The forecast values before the corrections generally managed to reproduce the load curve, but on the whole, they exceed the maximum actual measurements for June. On the other hand, we can see that the forecast values after the correction almost overlap with the actual measurements, indicating that the forecast is accurate.

We created the monthly forecast models for the 184 distribution system lines between April 2014 and March2015 to do the same study, and we conducted annual evaluation of the forecast models with separate 49 distribution system lines for evaluation.

Fig. 8 shows the forecast errors of absolute values for average weekday current and mean current values for each month of the year. According to the Figure, the largest mean error of the year is 14.4A in August. We can also see that the forecast error is greater in August and February. The trend in the shift of the forecast error during the year shows larger errors during the seasons with large current, so we compared the error with the monthly average current in the distribution systems for evaluation. The correlation coefficient was 0.92. Therefore, we can see that forecast errors of absolute values are strongly correlated with mean current measurements. We asked actual operators for opinions about the forecast accuracy of our method. They answered that the accuracy of the method was good enough for practical use. Since the current values of the distribution systems are often several hundred amperes, it is presumed that if the forecast error of about several tens of amperes of this method is considered, the possibility that the operation quality of the operator will not be significantly lowered is small. So we were able to confirm the effectiveness of our method.

## LOAD CURRENT FORECAST IN SECTION BY USING ITSW

This chapter examines the forecast accuracy when forecasting current per section instead of for the entire distribution system. If we can forecast load current per section, it will be advantageous in operating distribution systems, such as changing by operating a switch or load interchange. The forecast model is the same as the previous Chapter, but here we forecast load current in specific sections by using measurements by the ITSW installed in each section as correction values. As validation data, in distribution systems including 86 sections in which the amount of PV interconnection is 100 kW or less and ITSW is installed as before, 2016 when the load becomes heavy in winter. We forecast the average current for each section for weekdays in January of the year.

Fig. 9 shows the histogram of forecast errors of absolute values when the 86 sections for verification are forecast. The absolute error was 15 A or less in 92% of the forecast points, and the mean absolute error was 5.9 A. Therefore, we were able to be confirm that this method is effective in

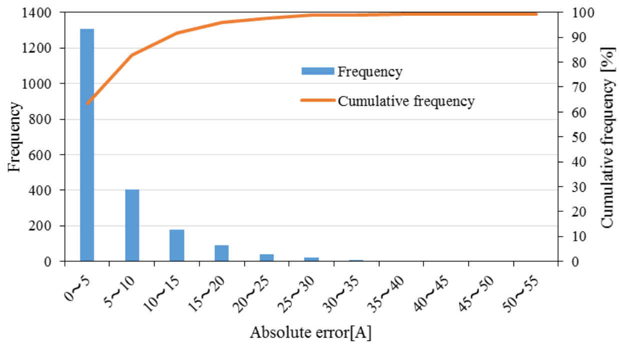
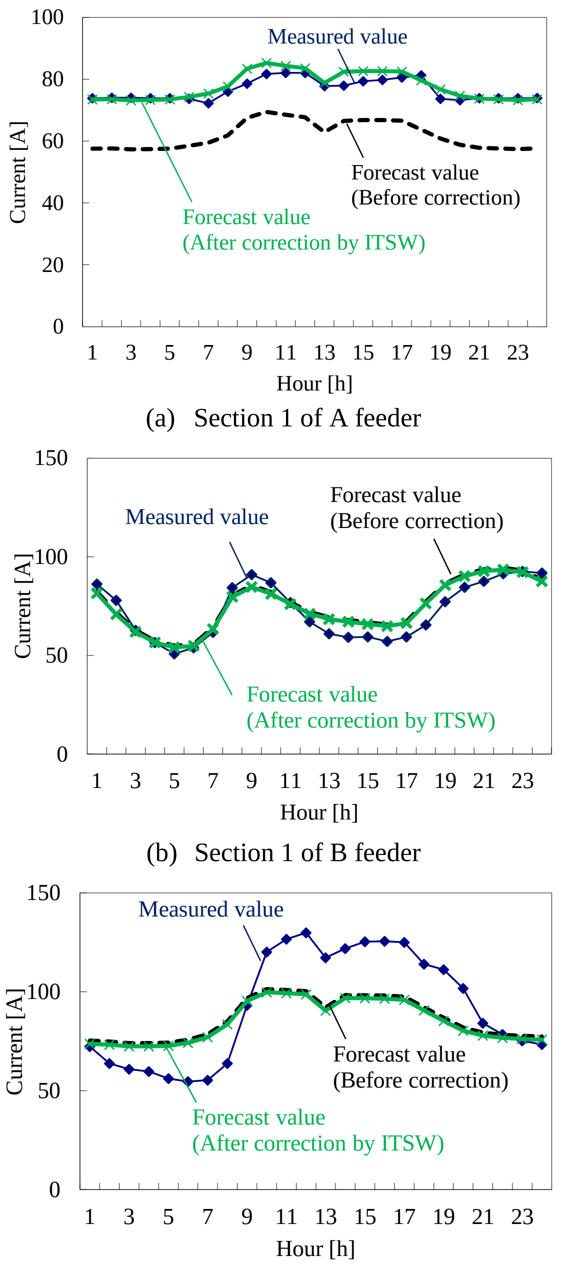



Fig.9. Histogram of absolute error




(a) Section 1 of A feeder

(b) Section 1 of B feeder

(c) Section 3 of C feeder

Fig. 10. Example of load forecast correcting by ITSW

forecasting section current by using the measurements by ITSW for corrections. Fig. 10 shows an example of forecast results. The Figure (a) shows the results of forecast for one section in Distribution Line A, and the Figure (b) shows the results of Distribution Line B. We can see that the forecast for the current values in the sections is close to the actual measurements of the current with the corrections by the measurements by ITSW. On the other hand, (c) shows the forecast results for three sections in Distribution Line B. The load curve is different from the actual values, and the forecast values deviate from the actual values even after the corrections. This may be because one customer in this section accounts for a large portion of the total contract capacity (high-voltage power 1,690 kW) and the system current is dominated by the use pattern of this consumer, so it cannot be calculated by the statistical forecast formula. It is known from the results of previous analysis by the authors that forecast accuracy is poor in a system like this [4], so we need to take different measures and it remains to be solved.

Since this forecasting method is a method to statistically calculate representative values from the entire learning data due to the nature of the method called multiple regression analysis, the load curve also expresses the average shape. So there is a problem that the distribution system having a peculiar load curve can not express the load curve shape. There are many cases where specific customers occupy the majority of the supply amount in the systems with low forecasting accuracy. In particular, when the customer is a special type of business, the possibility that the forecasting accuracy is lowered has been confirmed [5]. In the future, it is needed to consider a load management method using a method of searching for a case that is difficult to forecast based on the customer contract capacity or industry type information, and reading automatic meter values of dominant customers.

## CONCLUSION

Actual load current will become difficult to know with the rapid increase of the amount of PV interconnection to distribution systems. This paper investigated a method of forecasting it using a statistical technique. The forecast model created by this method can mostly reproduce the shapes of load curves. When we forecast the weekday average values of distribution line current in June, we were able to forecast it with great accuracy after the corrections, with the mean error of absolute values of 11.0 A in all time slots and the error ratio of absolute values of 8.8%. Furthermore, when we conducted an annual evaluation with this method, the mean error for each time slot was 24 A at a maximum at 9 o'clock on weekdays in August and February. Comments from operators also suggest that the method is good enough for practical use.

We forecast current for each section using the measurements by ITSW to correct the values, and the results showed that the method could forecast it very accurately with the mean absolute error of 5.9 A. We were also able to confirm the possibility of forecasting load current per section. We are going to forecast the section current in spring and summer and analyze characteristics of systems with poor forecast accuracy in order to improve forecast accuracy further in the future.

## REFERENCES

[1] T.Sano, I.Tezuka, Y.Fukuda, 2004 : “The Load Forecasting Technology in the Electric Power Distribution System”, *IEEJ Trans. Power Syst.*, Vol.124, No.6 p.824-834 (in Japanese)

[2] T.Sano, E.Kamiya, T.Nagumo, 2014 : “A Basic Study for the Load Forecasting in the Electric Power Distribution System using Statistical Classification”, Proc. of 2014 Annual Conference of Power & Energy Society, IEE Japan, No.181 (in Japanese)

[3] E.kamiya, Y.Ogihara, T.Sano, 2015 : “A Study of load forecasting systems throughout the year”, Proc. of 2015 Annual Conference of Power & Energy Society, IEE Japan, No.208 (in Japanese)

[4] T.Sano, H.Ito, H.Takigasaki, 2016: “Factor Analysis of the Error in the Load Forecasting of the Electric Power Distribution system”, 2016 National Convention Record IEE Japan, No.6-152 (in Japanese)

[5] T.Nagumo, H.Ito, H.Takigasaki, T.Sano 2017: “Trend analysis of electric power distribution system with low forecast accuracy in load forecasting”, 2017 National Convention Record IEE Japan, (in Japanese)