

High season demand forecasting method in logistics warehouse inventory management

陳成[†] 尾白大知[†] 平山淳一[†] 荒宏視[†]
Cheng chen Daichi Ojira Junichi Hirayama Koji Ara

1. Introduction

Demand forecasting is an approach to predict future demand for a product or a service. Demand forecasting techniques are widely used in many different business fields. In electricity field, demand forecasting is important for energy suppliers to control power generation [1]. In marketing field, demand forecasting helps to design suitable production plan and determine appropriate price [2]. In logistics field, which this paper focuses on, demand forecasting can reduce the cost and increase efficiency of supply chain [3].

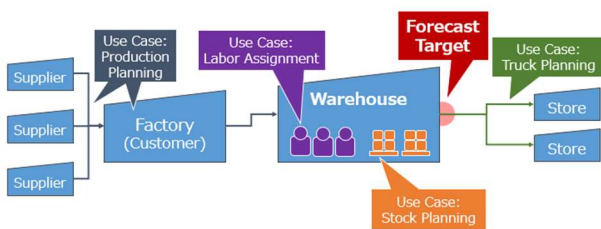


Figure 1 Demand forecasting in supply chain

Demand forecasting plays an important role in different steps of supply chain as shown in figure 1. For supplier, demand forecasting can help suppliers to determine appropriate production plan maximizing profit. In warehouse inventory management, demand forecasting is important in labor management, inventory plan and shipment plan. For labor management, the number of each day's staffs in warehouse is depended on the shipment amount of warehouse. Good demand forecasting can help warehouse manager to decide staff number of warehouse. In inventory plan, each item's inventory number is referred to demand forecasting results. In shipment plan, shipment number to destination have strong relationship to number of trucks used for transport, demand forecasting can help to decide the transport truck planning to reduce the cost of redundant or temporary trucks. In this paper, we are mainly focused on demand forecasting in one part of warehouse inventory management: shipment plan, to reduce cost of temporary transport trucks and labors.

A challenge in logistics demand forecasting is peak (high season) demand forecasting. Peak demand means the point that demand is significant larger than the demand before and after it. Peak demand always occurs around the long vacation such as end of a year, Japan golden week. In these times, to prepare stocks for the long vacation, a huge number of shipment occur. Although peak demand only occurs few times, because the amount of peak demand is significant huge and the price of truck is higher than normal season, the cost of truck and labor in peak period is obvious large than normal season. If the peak demand can be predicted, the shipment of peak period can move forward to normal season where

[†] 日立製作所 Hitachi, Ltd.

the cost of truck and labor is lower than peak period, the cost in peak period would be decreased. Since the peak demand occurs only few time, normal machine learning method would ignore the peak demand in training data to avoid overfitting. Furthermore, the occurrence time of one item's peak demand is always shift in different period. For example, a peak demand occurs in 51th week in 2018 and may shift to 52th week in 2019 because the customers' order request timing is changed based on their sales plan. We call this phenomenon to be peak shift. The peak shift leads the peak demand harder to predict. In previous works, Fragkioudaki et al. [4] proposed a forecasting method which predicts the occurrence places of electricity peak, and builds a special prediction model for predicted peak points. However, their method does not consider the peak shift. If the peak demand cannot predict, the number of labor and truck have a high risk excess or deficiency, the customers' cost increases a lot.

The shift of peak is hard to be judged because of the limitation of training data. For each item, the length of training data often less than 3 years, the all places that peak may occur cannot be detected by the small number of training data. However, the peak occurrence trend is similar for all items in a warehouse, the peak range of each item can be revised by all items' peak occurrence information. Therefore, to decrease the cost of labor and truck in supply chain, in this research, we propose a peak demand forecasting method extracting peak range of each item considering the peak shift by combining the single item and all items' peak probability distribution. We predict the demand in peak range period by a special model different to normal season and validate proposed method on real world 101 items' shipment data.

2. Related works

Demand forecasting technologies are widely applied in many different fields. Taylor et al. [5] proposed a new double seasonal Holt-Winters model outperforms traditional Holt-Winters and ARIMA models in online electricity demand forecasting. Motamedi et al. [6] proposed a forecasting framework for electricity market price and demand forecasting which considering relationship between price and demand. Carbonneau et al. [7] compared the advanced machine learning techniques (e.g., RNN, SVM) for supply chain demand forecasting. They found the machine learning techniques achieve more accurate than simple forecasting techniques (e.g., ARIMA). Recently, many works produce deep learning and probabilistic forecasting in demand forecasting. Salinas et al. [8] proposed a probabilistic forecasting method DeepAR by learning from RNN and similar items' data. Gasthaus et al. [9] proposed SQF-RNN combining the RNNs and quantile function to apply probabilistic forecasting method to large scale real world dataset. However, all these works did not consider predicting peak demand. A few works consider peak demand prediction in their methods. Zhao et al. [10] proposed a forecasting method for electricity market price which considers an occurrence

of the price peak. Fragkioudaki et al. [4] proposed a forecasting method which predicts occurrence places of electricity peaks, and builds a special prediction model for predicted peak points. However, all these works do not consider the peak shift, especially when the peak in test period is different to training period, the occurrence of peak cannot be correctly predicted. Different to these works, our method can predict the peak demand even the peak shift occurs.

3. Proposed method

3.1 Problem definition

The demand forecasting in shipment plan is using past shipment demand of each item to predict future shipment demand. The future demand $d_t, d_{t+1}, d_{t+2}, \dots, d_{t+n}$ after time t can be defined as follows:

$$d_t, d_{t+1}, d_{t+2}, \dots, d_{t+n} = F(d_{t-1}, d_{t-2}, \dots, d_1)$$

where $d_{t-1}, d_{t-2}, \dots, d_1$ are the shipment history demand, F is the prediction method. $d_t, d_{t+1}, d_{t+2}, \dots, d_{t+n}$ are the predicted demand at time $t, t+1, t+2, \dots, t+n$. In other words, future multiply steps demand are predicted based on the shipment demand history data.

In this research, the peak demand p of an item is defined as follows:

$$p = \mu + 2\sigma$$

where μ is the mean shipment demand of the item, σ is the item's demand standard deviation.

3.2 System overview

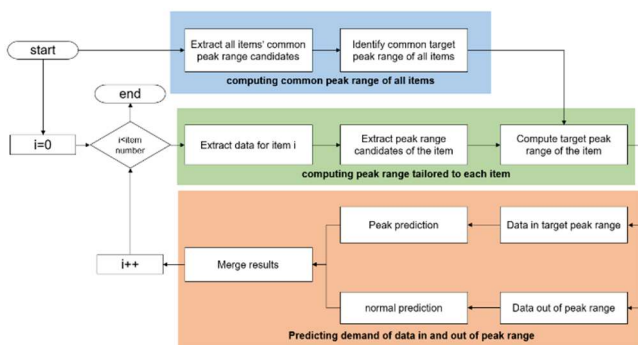


Figure 2 Overview of system

To predict the peak demand, the idea of proposed method is (1) computing the peak range of an item utilizing the all items' peak information. (2) Building 2 prediction models with different strategies for this item, one to predict peak demand (demand of data in peak range), and another one to predict normal demand (demand of data out of peak range). The flowchart of system is shown in figure 2. The proposed method mainly composes with 3 parts: (1) Computing the common peak range of all items. Some random peaks occur because of emergency request. These random peaks are not target in this research, we need first to remove these random peaks and select the range that periodical peak occurs. (2) Computing the peak range tailored to each item. Although we compute the possible peak range for all items, the shipping pattern of each item is little different and not totally same to common peak range. To improve the peak prediction accuracy, we specific the peak range of each item in this step. (3) Predicting demand of data

in and out of peak range. After we know the peak range of each item, using different strategies to predict data in and out of peak range. The detail of each step will be introduced in the rest of chapter.

3.3 Computing the common peak range of all items

The peak demand can simply divide into 2 types. One type is irregular peaks which occur randomly because of emergency request. Another type is periodical peaks which occur periodically because of long vacation. The irregular peaks occur randomly which is hard to predict only leverage shipment history data. In this research, we except to predict irregular peaks and focus on predicting periodical peaks. To predict periodical peaks, the first step is extracting periodical peaks. Because of peak shift, peaks occur not totally same in each period, we should extract the range not the detail position that periodical peak may occur. It is impossible to extract peak range only use one item data, we first compute the common peak range of all items according to all items' peak distribution. It is hard to decide the periodical peak range automatically because the irregular peaks disturb judgement. Our idea is computing the potential candidates of periodical peak range for all items, then utilizing user (warehouse manager) experience to help judge the periodical peak range for all items. There are 3 steps in this part.

3.3.1 User decide the prediction unit and frequency of peak

At first, user has to decide the prediction unit and frequency of peak. The prediction unit (e.g., day, week, month, ...) depends on the business factor such as transport plan. The frequency of peak (e.g., once a year, once a month, ...) is rely on the prediction unit and should be decided based on prediction unit.

3.3.2 Computing all items common peak range candidates based on all items shipping data

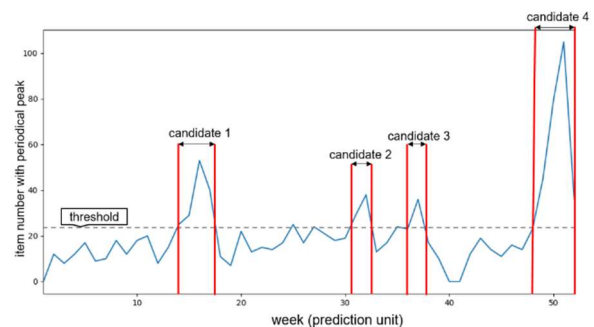


Figure 3 Number of item with periodical peak distribution

After deciding the prediction unit and frequency of peak, we collect shipment data of all items by prediction unit decided in last step and compute the number of items with periodical peak (with multiply peaks at a point in all training period) of each point based on the definition of peak demand in section 3.1. An example of number of item with periodical peak distribution is shown in figure 3. As shown in figure 3, we can find some places number of items with periodical peak is obvious larger than other places since many items have periodical peak in the places, which means these places are possible to be periodical peak for most of items. The places number of items with periodical peak is low means the periodical peak only appears in few items. We set a threshold and extract the continuous range that number of items with periodical peak beyond the threshold. The threshold is defined as follows:

threshold = mean number of items with
periodical peak at each point

We regard the continuous range beyond the threshold as the periodical peak range candidates since a large number of items have periodical peak in this range. Here we set mean number of items with periodical peak of each point as the threshold not a larger number since we want to more potential peak ranges can be contained in the candidates. In figure 3, there are 4 continuous ranges over the threshold, we regard the 4 continuous ranges as the candidates of common peak range.

3.3.3 User select and modify the common peak range from candidates

Not all the peak range candidates are real peak range. Some candidates' ranges are too large or too small, and some peak ranges are not suitable for most of the items. Consider these problems, the target peak range should be chosen by user with warehouse management experience. In this step, we tip the peak range candidates and representative the items of each candidate to user. User could judge each candidate is real peak range or not. Finally, the peak ranges that confirmed by user are set as the target common peak range for all items.

3.4 Computing peak range tailored to each item

Although the possible peak ranges for all items are decided in last step, the common peak range maybe too large or too small for each item. When the common peak range is too large, many normal season data contain in peak range, the accuracy of peak prediction would be worse. When the common peak range is too small, some peak points may not contain in peak range. Considering the problems introduced below, the appropriate peak range tailored to each item should be computed based on common peak range. There are 4 steps to decide target peak range of each item. Figure 4 shows an example of the process.

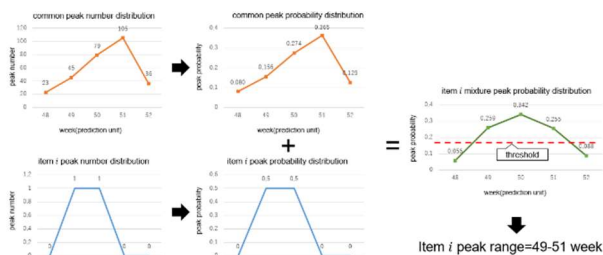


Figure 4 Process of item peak range computation

3.4.1 Computing the item's peak probability distribution

In this step, we compute the item's peak probability distribution. At first, the peaks of an item is extracted based on the peak definition in the section 3.1. Next, the number of peak is grouped by the prediction unit. Finally, the peak probability of an item at position t is computed as follows:

$$\text{position } t \text{ peak probability} = \frac{\text{position } t \text{ total peak number}}{\text{number of peak period}}$$

where total peak number is the sum of whole training data peak number at position t. Peak period is the number of prediction unit period. For example, when the peak occurs once a year and have 3 year data, the number of peak period is 3. The mean of item's peak probability is the probability of each position to be peak according to peak number in the training data.

3.4.2 Computing the common peak probability distribution

In this step, we compute the each common peak range candidate's peak probability distribution. In section 3.3, we get the common peak range candidates that periodical peaks always occur and we know the periodical peak number in each position of common peak range candidate. We regard the periodical peaks only exist in the range of common peak range candidates, and peak probability of position i in a common peak range candidate is computed as follows:

$$\begin{aligned} &\text{position } i \text{ peak probability} \\ &= \frac{\text{peak number of position } i}{\sum_{j=1}^n \text{peak number of position } j} \end{aligned}$$

where n is the length of the common peak range candidate. The mean of common peak probability in this section is the peak occurrence probability in the whole common peak range candidate.

3.4.3 Computing the mixture peak probability distribution

In this step, we mix the item peak probability distribution and common peak probability distribution as the revised peak probability distribution for the item. Limited by the number of each item's peak data, the item peak probability distribution do not reflect the real peak probability distribution of the item, so that when the peak shift occurs, only use item peak probability distribution cannot correctly compute the position of peak. The common peak probability distribution reflects all item's peak distribution, mix the item and common peak probability distribution can revise the item peak probability distribution close to real distribution. Firstly, we judge the peak of an item belong to each common peak range candidate or not. If an item has peak in a common peak range candidate, we mix the two peak probability distribution as follows:

$$\begin{aligned} &\text{mixture peak probability distribution} \\ &= \alpha \times \text{item peak probability distribution} \\ &+ (1 - \alpha) \\ &\times \text{common peak probability distribution} \end{aligned}$$

here α is a hyperparameter used to adjust the weight of two distribution. In our experiment, α is set to 0.3 based on our data. If an item has no peak in a common peak range candidate, we regard no peak in the range of common peak range candidate. For peak of the item not belong to any common peak range candidate, we regard these peaks are irregular peaks and not the target of this research, we ignore these peaks when computing mixture peak probability distribution.

3.4.4 Extracting the peak range of an item

In this step, we extract the peak range of an item based on the mixture peak probability distribution computed in last step. The method is similar to section 3.3, we set a threshold and extract the positions that peak probability over the threshold, the extracted position ranges are set as a peak range of an item. The threshold is set as follows:

$$\text{threshold} = \frac{\max(\text{mixture peak probability})}{2}$$

which means we regard the points over the half of the highest peak probability in the mixture peak probability distribution as peak points because their peak probability is closed to the point with the highest probability to be peak.

3.5 Predicting demand of data inside and outside of

peak range

The final step is using different strategies to predict data in target peak range and out of target peak range. The amount of data in target peak range is small and close to each other, the model to predict peak should be designed more complex so that the peak data would not be ignored. For data not in peak range should use normal prediction method to achieve best accuracy. There are many different strategies can use to make prediction, here we choose to build two decision trees [11] to make prediction. For normal prediction model, we use grid search to find the appropriate parameters for each item. For peak prediction model, the parameter to control minimum number of samples required to be at a leaf node (`min_samples_leaf`) is set to a small value. For normal model, the small `min_samples_leaf` means the model is easy to overfitting. But for peak prediction, since the peak demand and time do not change a lot, the small `min_samples_leaf` can build a more complex decision tree to increase the accuracy of peak demand prediction. Considering the time series of shipment data, in training phase, both of normal model and peak model use all training data to train the model. In test phase, the data belong to peak range use peak model and the data belong to normal season use normal model to make prediction.

4. Experiments

4.1 Experiments setting

To evaluate the effectiveness of the proposed method, we did 2 experiments utilizing real world data.

4.1.1 Experiments data

The experiment data information is summarized in table 1.

Table 1 Experiment data information

Data source	Data range	Prediction unit	Item number
Medical industry	2014-09-29~ 2017-07-31	Week	101

We use a dataset from a real world warehouse serves customers in medical industry from 2014-09-29 to 2017-07-31 for experiments. The target of this research is reducing cost of transport trucks. In logistics work site, the number of necessary trucks are usually decided by week unit, so we set the prediction unit of the experiments to week unit. In other words, we utilize past every week shipment demand to predict future each week shipment demand. Some items' shipment data are sparse and shipment number is low or without peak, we filter these items and select 101 items with large shipment number as the experiment items.

4.1.2 Common peak range decision

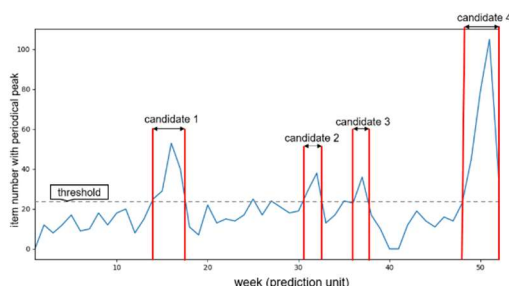


Figure 5 Number of item with periodical peak distribution
The first step of experiment is deciding common peak range of all items, figure 5 is the number of item with periodical peak

distribution of all items in the dataset. Based on the method in section 3.3.2, there are 4 peak range candidates in experiment dataset as shown in table 2. Considering the peak number in each candidate and logistics characteristics, we regard 14-17 week (before Japan golden week) and 48-52 week (end of year) as the common peak ranges of all items.

Table 2 Peak range candidates in experiment dataset

Candidate number	Peak range
1	14 week-17 week
2	31 week-32 week
3	35 week-37 week
4	48 week-52 week

4.1.3 Scenarios

We design 2 scenarios to evaluate proposed method performance for different peak periods:

(1) To evaluate the 48-52 week peak prediction performance, the training period is set to 2014-09-29~2016-09-30, the test period is set to 2016-10-01~2016-12-31 which contains peak season in 48-52 week.

(2) To evaluate the 14-17 week peak prediction performance, the training period is set to 2014-09-29~2017-03-31, the test period is set to 2017-04-01~2017-06-01 which contains peak season in 14-17 week.

We implement experiment to evaluate all 101 items performance under the 2 scenarios.

4.2 Experiment 1

This experiment evaluates performance of each item's peak range that computed in section 3.4.

4.2.1 Compared methods

Baseline: A method usually used in logistics site. The process of method is: (1) Extracting position that demand over the peak demand defined in section 3.1 from training data of the target item as peak position of target item. (2) For each extracted peak position x , the peak range based on the peak position is set as $[x-1, x+1]$. For example, if peak position $x=49$ week, the peak range is set as [48 week, 50 week] according to peak position 49 week.

Proposed: this is the proposed method that computing peak range of the target item by method in section 3.4.

4.2.2 Evaluation method

The evaluation of peak range is considered similar to evaluation of binary classification problem. All positions of data in the predicted peak range are set as positive position and all positions that data out of peak range are set as negative position. Therefore, the performance of predicted peak range can be evaluated by indicators that used for binary classification evaluation. In this experiment, we use precision, recall and f1-score to evaluate performance of predicted peak range.

4.2.3 Results

Table 3 Evaluation of peak range for scenario 1

	precision	recall	F1-score
Baseline	0.397	0.849	0.517
Proposed	0.433	0.849	0.545
Item number	42	42	42

Table 4 Evaluation of peak range for scenario 2

	precision	recall	F1-score
Baseline	0.348	0.826	0.474
Proposed	0.470	0.894	0.582
Item number	22	22	22

Table 3 is scenario 1 evaluation results, item number in table 3 means the number of item with peak demand data in 48-52 week of training data. The precision, recall and F1-score in table 3 are average results of 42 items with peak demand data in 48-52 week. Table 4 is scenario 2 evaluation results, item number in table 4 means the number of item with peak demand data in 14-17 week of training data. The precision, recall and F1-score in table 4 are average results of 22 items with peak demand data in 14-17 week.

From the results in table 3 and table 4, proposed method performance is better than baseline method for all 3 evaluation methods. The precision in this experiment indicates predicted peak position is real peak position or not. The precision of proposed method is higher since peak range computed by proposed method is dynamic and tailored to the target item considering the other items' peak distribution, so that predicted peak range is close to true peak range and not contains many positions without peak demand. However, precision of baseline method is low since the peak range computed by baseline method is static for all items, for some items' peak position not shift a lot, peak range for these items is too large and many predicted peak positions are not real peak positions. The recall in this experiment indicates real peak position is predicted or not. Both 2 methods' recall results are higher than 80% which means peak range computed by both 2 methods contains most of real peak positions. Not all true peak positions contains in peak range since some peak demand randomly occur influencing the results.

4.3 Experiment 2

This experiment evaluates predicted demand accuracy comparing to 2 existing methods.

4.3.1 Compare methods

ARIMA [12]: ARIMA model is a generalization of an autoregressive moving average (ARMA) model.

Peak auto prediction: a peak prediction method proposed by [4]. This method consists in training three basic sub-models: (1) A classification tree to predict the occurrence of peak. (2) A regression tree to predict the value of demand at a point in case a peak has been predicted by the classification tree. (3) Another regression tree to predict the value of the price at a point in case a normal point has been forecasted by the classification tree.

Proposed: This is the proposed method.

4.3.2 Evaluation method

In experiment, we use MAE (mean absolute error) as evaluation method because MAE directly indicates error amount demand between predicted value and real shipment demand. MAE can link to economic indicator (e.g., shipment cost) and widely use in real world shipment planning. The definition of MAE is:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where y_i is the ground truth demand at time i , \hat{y}_i is the predicted value at time i . Based on the definition, the smaller the MAE, the better the prediction results.

4.4 Results

Table 5 101 item average prediction MAE results

	Proposed	ARIMA	Peak auto prediction
Scenario 1	827.262	900.546	891.415
Scenario 2	841.651	918.665	951.668

Table 6 Number of items with the lowest MAE

	Proposed	ARIMA	Peak auto prediction
Scenario 1	49	30	22
Scenario 2	46	29	26

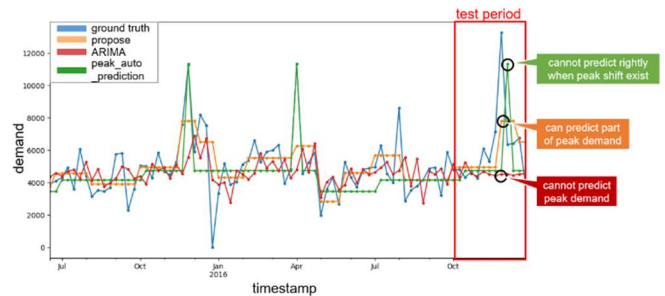


Figure 6 Scenario 1 one item prediction results

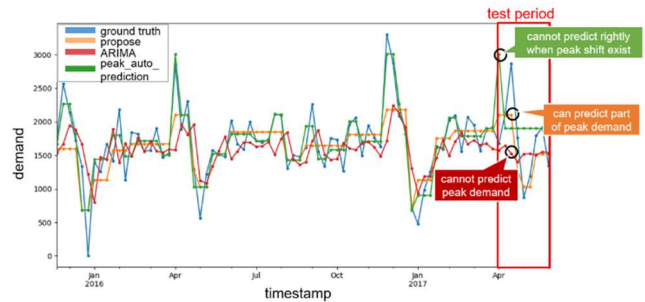


Figure 7 Scenario 2 one item prediction results

Table 5 presents the 2 scenarios 101 items average MAE results, table 6 presents 2 scenarios each method number of item MAE lower than other 2 methods results. Figure 6 shows one item prediction results under scenario 1, figure 7 shows one item prediction results under scenario 2. From the results we can find proposed method MAE results are average 7.8% better than existing method with the lowest MAE for each scenario. Besides, for proposed method, the number of item with the lowest MAE is larger than other 2 compared methods. This is because proposed method can predict the peak demand even the peak periods are different in training and test phase. The peak auto prediction method performance is the worst because peak auto prediction method cannot correctly predict the position that peak occurs. As shown in figure 6 and figure 7, when the peak shift occurs, peak auto prediction method may predict a wrong point not the peak to be peak point and predict huge demand to the wrong point increasing the error amount to both real peak point and wrong predicted peak point. Different to the peak auto prediction method, the proposed method predicts the peak range not the peak point. When the peak shift occurs, the proposed method still can predict part of peak demand to decrease the error amount. ARIMA cannot predict the peak demand so its performance is worse than proposed method. In scenario 2, ARIMA performance is better than peak

auto prediction method since peak auto prediction method predicts wrong places to be peak increase the error amount. For both 2 scenarios proposed method is better than the existing methods, which means proposed method is robust for different peak periods.

5. Conclusion

This research is focused on high season (peak) demand forecasting in logistics field. High season demand is obvious larger than the period before and after it and always occur around long vacation. If the high season demand cannot be predicted, temporary trucks and labors increase the transport cost. However, the number of high season data is significant less than normal season and the occurrence time of high season is always shift in different periods which lead high season demand hard to predict. To solve the issue, we propose a new high season demand forecasting method, the proposed method is comprised as follows:

1. Computing the common peak range of all items that high season may occur.
2. Computing the appropriate peak range tailored to each item based on the common peak range.
3. Using different strategies to predict the demand in and out of peak range, the high season demand can be predicted even the high season occurrence time is shift.

Based on the experiments using real world data, our proposed method decreases test period (including high season) demand forecasting error amount 7.8% comparing to existing demand forecasting method.

Reference

- [1] L. Hernandez et al., "A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings", *IEEE Communication Surveys & Tutorials*, Vol.16, pp. 1460-1495 (2014)
- [2] Kris Johnson Ferreira, Bin Hong Alex Lee, David Simchi-Levi, "Analytics for an Online Retailer: Demand Forecasting and Price Optimization", *Manufacturing & Service Operations Management*, Vol.18, No.1 (2015)
- [3] Luis Aburto, Richard Weber, "Improved supply chain management based on hybrid demand forecasts", Vol.7, pp. 136-144 (2007)
- [4] A. Fragkioudaki, A. Marinakis, R. Cherkaoui, "Forecasting price spikes in European day-ahead electricity markets using decision trees" 12th International Conference on the European Energy Market (EEM) pp. 1-5 2015.
- [5] J W Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing", *Journal of the Operational Research Society*, Vol.54, pp. 799-805 (2003).
- [6] Amir Motamedi, Hamidreza Zareipour, William D. Rosehart, "Electricity Price and Demand Forecasting in Smart Grids", *IEEE Transactions on Smart Grid*, Vol. 3, pp. 664-674 (2012).
- [7] Real Carbonneau, Kevin Laframboise, Rustam Vahidov, "Application of machine learning techniques for supply chain demand forecasting", *European Journal of Operational Research*, Vol.184, pp. 1140-1154 (2008).
- [8] David Salinas, Valentin Flunkert, Jan Gasthaus, "DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks", <https://arxiv.org/abs/1704.04110>
- [9] Jan Gasthaus, Konstantinos Benidis, Yuyang Wang, Syama Sundar Rangapuram, David Salinas, Valentin Flunkert, Tim Januschowski, "Probabilistic Forecasting with Spline Quantile Function RNNs", *Proceedings of Machine Learning Research*, Vol.89, pp. 1901-1910 (2019).
- [10] Jun Hua Zhao, Zhao Yang Dong, Xue Li, Kit Po Wong, "A Framework for Electricity Price Spike Analysis With Advanced Data Mining Methods", *IEEE Transactions on Power Systems*, Vol.22, pp. 376-385 (2007)
- [11] L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classification and Regression Trees", Wadsworth, Belmont, CA, 1984.
- [12] Brockwell, P.J., and Davis, R. A., "Time series: theory and methods". Springer Science & Business Media, 2009.