Korean J. Chem. Eng., 35(4), 1033-1044 (2018)
DOI: 10.1007/s11814-017-0235-6

pISSN: 0256-1115
eISSN: 1975-7220

A novel system dynamics model for forecasting naphtha price

Byeonggil Lyu, Hweeung Kwon, and Il Moon†

Department of Chemical and Biomolecular Engineering, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 30722, Korea

(*Received 12 July 2017 • accepted 21 August 2017*)

Abstract−Fluctuations in naphtha price are directly related to the profit of petrochemical companies. Thus, forecasting of naphtha price is becoming increasingly important. To respond to this need, a naphtha crack (the price gap between naphtha and crude oil) forecasting model is developed herein. The objective of this study was to design a reasonable forecasting model that is immediately available and can be used to develop various naphtha supply strategies. However, it is very difficult to forecast a price value with a high accuracy. Therefore, the proposed model focuses not on the price value but on the direction of the crack. These considerations are vital to a company's decision-making process. In addition, a system dynamics model that considers causal relations is proposed. It was developed based on heuristics, statistical analysis, seasonal effects, and relationships between factors that affect naphtha price, and it exhibits an accuracy rate of 84%-95% in forecasting of the naphtha crack three months in advance.

Keywords: Naphtha Crack, Causal Loop, Seasonal Effect, System Dynamics, Forecasting

INTRODUCTION

Naphtha, a crucial product of crude oil distillation, is a primary raw material for petrochemical plants. Thus, the profit of a petrochemical company is highly affected by the price of naphtha. Conversely, condensate is used as a resource in the petrochemical industry. Thus, companies are faced with a decision-making problem of whether to use naphtha or condensate. As the dealing price of condensate is similar to the price of crude oil, "crack," which is the price gap between naphtha and crude oil, serves as vital information for petrochemical companies. When the naphtha crack is elevated, the use of condensate as a resource is increased, whereas the use of naphtha is reduced. Additionally, owing to transit times, the price of raw materials at the point-of-use and the point-of-purchase are different. Thus, the production decision-making process involves a thorough examination of both potential revenues and costs, with price being central to revenue-generation possibilities [1].

A well-developed decision-making support model helps a company obtain more increased profit, save energy and cost. To help the decision-making process, several decision support models have been suggested. Previously, an integrated capacity planning framework was employed in the semiconductor manufacturing process with the mixed integer linear programming (MILP) model [2]. The MILP model is also used for strategizing biomass utilization. In addition, a developed optimization model helps policymakers to solve problems related to energy supply [3]. Another MILP optimization-based decision-making support model was used in a wind-powered hydrogen supply system. The model and case studies were useful to design an operate hydrogen supply system [4]. Without the prediction of decision variables, these decision support models are a good help for decision makers. However, if decision support models include forecasting models with high accuracy and reasonable explanation, they will be more helpful to the decision-making process. Therefore, the development of a price forecasting model with high accuracy and reliability is required for petrochemical companies to increase profit. With this motivation, a model that focuses on the forecasting of the "naphtha crack" is developed herein. Previous research on forecasting methodology will be introduced in this section.

In the past, numerous forecasting methods have been suggested based on statistical time-series methods. In particular, a hybrid model that combines the autoregressive integrated moving average (ARIMA) and support vector machines (SVMs) has been presented [5]. The prediction performance of this model is higher than those of models using ARIMA or SVMs alone. An artificial neural network (ANN) model has been used in various studies involving activated sludge filamentous bulking [6], fixed-bed reactors [7], phase equilibria [8], and operational optimization of crude oil distillation systems [9]. In terms of price forecasting, system-marginal-price-based short-term forecasting methods involving ANNs [10], electricity price [11], and crude oil price [12] have been proposed. Moreover, some researchers have combined this statistical model with the ANN model to obtain increased forecasting accuracy. The ARIMA model was used with a wavelet transform to forecast electricity price [13]. In addition, a model using the adaptive neuro-fuzzy inference system (ANFIS) has been constructed [14]. The proposed model is considered to be more powerful than the statistical and exponential smoothing methods, especially when the formulated model is nonlinear. In addition, a hybrid model using a least-squares support vector machine (LSSVM) and autoregressive moving average with external input (ARMAX) has been proposed [15]. Based on statistical theory, a variety of models have been suggested for the forecasting of the price of gas components,

†To whom correspondence should be addressed.
E-mail: ilmoon@yonsei.ac.kr

natural gas, naphtha, and crude oil [16]. Forecasting models have been developed for the forecasting of natural-gas spot price using gamma test analysis, nonlinear models, local linear regression (LLR), dynamic local linear regression (DLLR), and ANNs [17]. These methods have promoted the growth of forecasting fields and have motivated the advancement of other novel forecasting methods. Some other types of forecasting models have also been proposed in previous studies. The kriging model is used to forecast corrosion in the crude oil overhead [18]. This model shows better forecasting accuracy than ANNs and the multiple linear regression (MLR) model. Correlogram, a diagram that reports the correlation index of two time series X and Y as a function of the time shift introduced between them, is used to model commodity fluctuations and forecast commodities [19]. A price/cost model for a refinery for short-term and medium-term prediction has also been developed using correlograms [20]. An interesting methodology for forecasting the system marginal price has been suggested by Yeo's group [21,22]. They proposed a forecasting model based on the supply and demand of power, and the model shows better forecasting results than the wavelet and neural network models.

These methodologies are well known and widely used in forecasting. However, no prior studies on forecasting naphtha crack have been conducted. In our previous research, a multiple regression model has been applied to forecast naphtha crack by taking more than 20 factors into account. This model exhibited high accuracy and focused on determining whether naphtha crack spread is positive or negative [23]. However, this model is limited because it does not explain the origin of the selected factors and the relations between them. Therefore, we started developing a forecasting model that can explain the forecasting result. First, forecasting of the naphtha supply and demand based on time serial data analysis and causal relation analysis is conducted [24]. Furthermore, we developed a naphtha purchase optimization model from our forecasting model [25]. Our forecasting model is based on the supply and demand causal relation system dynamics model; recently, we developed a decision support model for hedge trading from the purchase optimization model [26]. The optimization model would be of good help for decision-making in the petrochemical industry. However, the forecasting model, which is the basis of the optimization model, also lacks the explanation about the forecasting result.

Previous studies have suggested various models for forecasting, and these models have contributed to future decision-making. However, as these models were only developed on a theoretical basis, a reasonable explanation of the results was difficult to establish. Hence, during the decision-making process in actual situations, it is hard to answer the question "why is the price increasing or decreasing?" Accordingly, a reliable model that can explain and support decisions is necessary for field engineers. A model that considers causal relationships and reasonable equations between factors can be very helpful to field engineers. Thus, by using a combination of the existing mathematical theory and logical descriptions, a more updated forecasting model compared to our previous model is described in this paper. This research explains how a causal model can be constructed, describes the method used herein for forecasting, and validates the accuracy of the results.

The main objectives of this paper are i) to develop a new causal loop of the naphtha supply/demand chain based on heuristic methods and statistical theory, and ii) to forecast the increase or decrease direction of the naphtha crack using the causal loop and regression method. To achieve these goals, this paper is organized as follows. First, two forecasting models are used: an MLR model and a system dynamics (SD) model. The MLR model is a widely used conventional model. However, it has limitations in the explanation of causal relationships. To solve this limitation, a new strategy is suggested: an SD model. Second, a detailed SD model with a causal loop diagram is described. Finally, the forecasting results are analyzed.

FORECASTING METHODOLOGY

The MLR model is a widely used statistical method owing to its simplicity. However, this model cannot explain the causal relationships between forecasting results. Therefore, a new strategy that combines causal relationships with linear regression is proposed in this paper. Furthermore, the forecasting performance is also compared with an MLR forecasting model.

1. Multiple Linear Regression

First, a simple MLR model is used in this research. The MLR model, represented by Eq. (1), is widely used in forecasting. Here, y is the forecasting objective factor (the naphtha crack in this paper); x_i's are the major affecting factors, such as the price of products, supply, and demand; and β_i's are the coefficients of the selected factors. This model was developed using the Statistical Package for the Social Sciences(SPSS) software.

$$y=\Sigma_{i=1}^{n}\beta_i x_i+\varepsilon. \tag{1}$$

2. System Dynamics

The MLR model is widely used in forecasting because it is simple to use. However, this method has two major limitations in the forecasting of the naphtha crack. First, it is hard to explain the relationship between factors, and second, same-time factors cannot be used because we do not know the value of each factor. Therefore, the MLR model can only use previous data and not same-time data. To solve these problems, a new strategy for the forecasting of the naphtha crack is suggested in this paper.

SD is a problem-solving theory based on the feedback control theory. SD defines a system using variables related to the presented problem and models the relationships among the variables. It is based on operations research and the synthesis of numerous subjects such as systems theory, control theory, information feedback theory, decision-making theory, mechanical systems, and computer science [27]. SD utilizes various control factors such as feedback loops and time delays to observe how the system reacts and behaves in response to trends. Additionally, under complex and dynamic system behaviors, policy and decision makers can be assisted by SD to make decisions [22]. This theory is used to analyze supply chains [29], perform modeling of electricity generation capacity systems [30], and plan wastewater management strategies [31]. In this research, an SD model is applied to develop a naphtha crack forecasting model. The forecasting procedure using an SD model is described in Fig. 1. The naphtha crack forecasting model is suggested based on a causal loop diagram, a quantitative

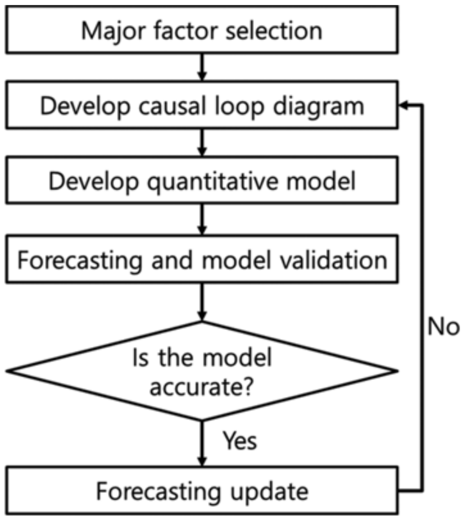

Fig. 1. SD model development procedure.

model, and a previous dataset.

SYSTEM DYNAMICS MODEL DESCRIPTION

The SD modeling procedure (Fig. 1) is described in this section. The major factor selection procedure using a field engineer's heuristic and statistical analysis is first conducted. Based on the selected factors, a causal loop diagram and SD model are developed. Then, quantitative modeling and naphtha crack forecasting are conducted. Further details on the model are explained below.

1. Major Factor Selection

Among the procedures underlying naphtha crack forecasting, the selection of affecting factors and development of a causal loop diagram are the most important procedures. To select appropriate factors for the naphtha crack, a selection procedure is conducted using Pearson product moment correlation coefficient (PPMCC) analysis and the heuristic of field engineers. The forecasting model can be reliable when it reflects an engineer's experience. Therefore, a field engineer's heuristic is the most important factor to be considered. PPMCC is also used to select major factors, and it is the most commonly used estimator in measuring such correlations [32]. PPMCC is a measure of the linear correlation between two variables and gives a value between −1 and 1. A value of 1 indicates complete positive correlation, 0 indicates no correlation, and −1 indicates complete negative correlation. The coefficient is expressed in Eq. (2). The covariance of the two variables is divided by the product of their standard deviations.

$$\mathrm{p}_{X,Y}=\frac{\mathrm{cov}(X, Y)}{\sigma_X\sigma_Y}. \tag{2}$$

Among the various factors that affect the naphtha crack, specific factors were chosen based on their PPMCC values for the naphtha crack and a trading expert's heuristics. This research is focused on the forecasting of Asia's naphtha crack; therefore, it mainly considers factors in the Asian region. The supply and demand of naphtha, export from the Middle East and India, and Europe's naphtha crack are major factors affecting Asia's naphtha crack.

2. Causal Loop Diagram

After factor selection, a causal loop diagram is developed using

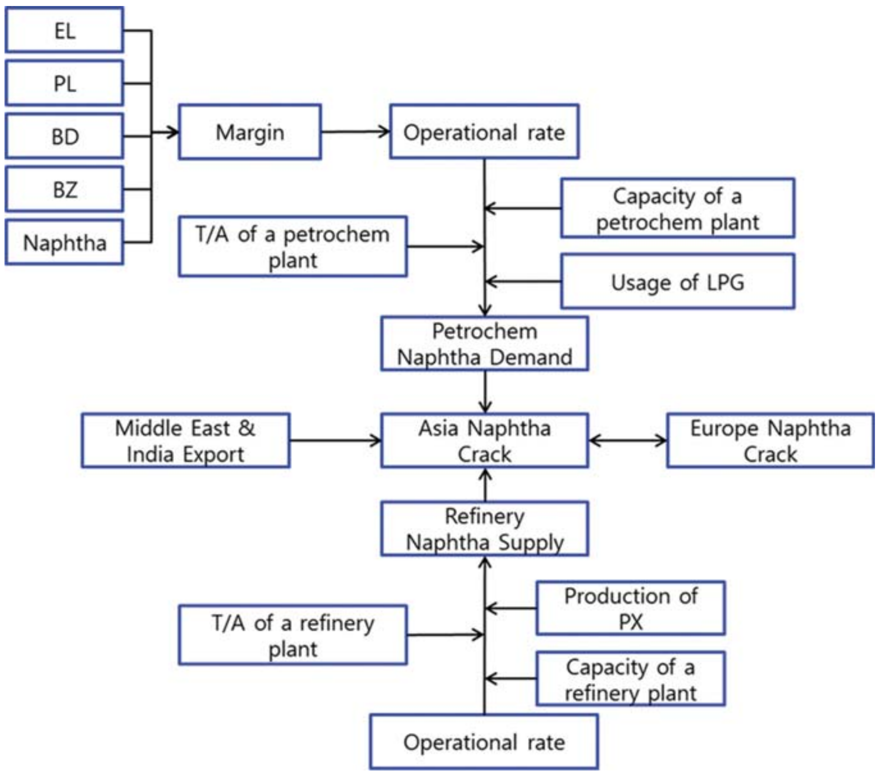

Fig. 2. Causal loop diagram of the Asian naphtha crack.

the selected factors. The world naphtha market is divided into two areas: East of the Suez Canal (EOS) and West of the Suez Canal (WOS). EOS consists of the East and the Middle East, and WOS consists of the European Union and the United States. Countries in East Asia, such as South Korea, Japan, and Taiwan, are naphtha consumers, and Middle Eastern countries and India are naphtha providers. As a result, the EOS naphtha network is developed by the naphtha supply and demand chain of these countries. For WOS, gasoline and propane are the important factors that affect the supply and demand of naphtha. Surplus naphtha in Europe moves to North America or Asia. Among the various factors regarding EOS and WOS, certain key factors were chosen to produce the causal loop of the naphtha crack. Therefore, the causal loop diagram is based on EOS and WOS. The developed loop consists of two main loops: the causal loop of Asia (EOS) for the Asian naphtha crack, and the causal loop of Europe (WOS) for the European naphtha crack. Both naphtha cracks affect each other.

2-1. Causal Loop of Asia

The causal loop for Asia (shown in Fig. 2) consists of the supply and demand of naphtha. Among the various factors affecting the supply and demand of naphtha, the key factors for building a causal model are identified as follows: margin, capacity of a plant, operational rate of a plant, turnaround (T/A) of a plant, and naphtha substitute. The demand of a petrochemical company for naphtha is almost the same as the demand for raw material, which is affected by its operational rate and plant capacity. Furthermore, the operational rate is affected by the operating margin. Thus, the capacity of a petrochemical plant, its rate of operation, and the operating margin are considered to be key factors affecting its demand.

The operating margin is defined by the price gaps of ethylene (EL), propylene (PL), butadiene (BD), benzene (BZ), and naphtha. The capacity of a petrochemical plant is based on Asian petrochemical plants, and the rate of operation is affected by the economic situation and T/A. The larger the capacity of a petrochemical plant, the higher the amount of naphtha can that can be cracked at a time. Therefore, a petrochemical plant with a larger capacity has a higher demand for naphtha. However, as the operational rate of a plant is not always 100%, it is the most important factor to be considered. The factors affecting the operational rate can be classified as economic factors (e.g., the margin on a product) and non-economic factors (e.g., T/A of a plant). Petrochemical companies repair their plants according to their T/A plans. During T/A, the cracking unit must stop its operation. Thus, the operational rate is reduced by performing T/A. Besides the effect of T/A, the operational rate of a plant depends on the market conditions, which is an economic factor. If the margin of the petrochemical industry becomes high, the inventory will decrease, which will lead to an increase in the operational rate, and vice versa. As a result, naphtha demand can be calculated using the margin and the operational rate of a plant. The margin is defined as the difference between the prices of the main products of petrochemical plants and naphtha. Using the equation between the margin and the operational rate developed with the linear regression method, the predicted operational rate can then be calculated.

In addition, the use of liquefied petroleum gas (LPG), which is an alternative to naphtha, affects the demand for naphtha. The products of naphtha cracking, such as ethylene and propylene, can be obtained through the cracking of LPG. Therefore, LPG can be used as an alternative to naphtha, and the use of LPG causes a reduction in the demand for naphtha. The price of LPG increases in winter as its demand as a heating fuel increases; conversely, it decreases during in summer. In fact, petrochemical companies use LPG when the price of LPG is lower than the price of naphtha. Therefore, the demand for naphtha can be determined based on the operational rate of a plant and the use of LPG.

The demand for petrochemical naphtha is calculated as follows. First, the price gap between the product and naphtha defines the margin. Next, the operational rate is calculated using the regression equation. Finally, the demand for naphtha is calculated while considering the capacity of a plant, the T/A of a plant, and the use of LPG.

The supply of naphtha can be calculated from the production of an oil company. The naphtha produced is used to obtain petrochemicals, e.g., para-xylene (PX), which is also an important factor in the supply of naphtha. T/A affects an oil plant's rate of operation as a non-economic factor. The operational rate and T/A enable the calculation of the supply of naphtha. However, it is difficult to calculate the operational rate using only its margin because it is affected by several factors. Therefore, an average operational rate with seasonal effects was used to forecast its supply. The production of PX is taken into account because it yields a higher margin than naphtha. As a result, oil companies are less likely to supply naphtha to petrochemical companies; rather, they are more inclined to produce it for themselves. Therefore, the amount of naphtha used to produce PX should be excluded from the supply of naphtha, which can thus be calculated using the relationship between the operational rate of oil plants and the amount of PX production.

The supply of naphtha from a refinery is calculated in a way similar to the calculation of the demand for naphtha. Based on the operational rate of a refinery, the supply of naphtha is calculated while considering the T/A of a plant, the capacity of a plant, and the production quantity of PX.

2-2. Causal Loop of Europe

Since European countries have a shorter shipping time compared with Asian countries, it is difficult to analyze the relationship between a European petrochemical company's margin and operational rate. This challenge led to the development of a model based on propane and gasoline price data. Gasoline is mainly used as a car fuel, whereas propane is used as a heating fuel and a raw material for the petrochemical industry. When gasoline consumption increases, the gasoline and naphtha demand also increases, accordingly leading to a decrease in the surplus naphtha amount that moves to Asia. Increased use of propane as a heating fuel decreases the surplus of naphtha. Therefore, based on the prices of gasoline and propane, the European loop is proposed as shown in Fig. 3.

Propane M and gasoline M represent the first month's future prices of propane and gasoline, respectively. Propane M+1 and gasoline M+1 denote the second month's future prices of propane and gasoline, respectively. The spreads of the first month's future price and second month's future price affect the propane and gaso-

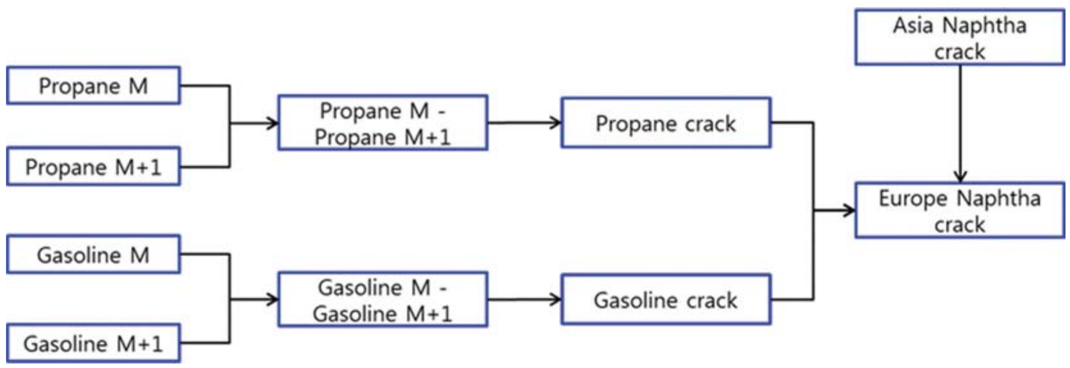

Fig. 3. Causal loop diagram for the European naphtha crack.

line crack. Crack indicates the price gap of propane, gasoline, and crude oil, leading to the calculation of Europe's naphtha crack.

3. System Dynamics Modeling

An SD model was developed based on the causal loop diagram shown in 2. Vensim, a widely used SD software, was used for modeling. This software supports drawing the causal loop diagram using the editing interaction formula and optimization of parameters. Our model is calculated by Intel Core i7-6700K CPU and it shows very short computational time less than 1 minute. Fig. 4 shows the flowsheet used for forecasting the naphtha crack. It comprises submodels of the Asian naphtha supply and demand and the European naphtha crack loop. Every arrow represents an interaction formula in the form of a linear expression between the factors.

3-1. Asian Naphtha Supply and Demand

As mentioned in Section 2.1, the Asian naphtha demand model consists of the margin, capacity, T/A, operational rate, and use of LPG. The margin is calculated from the prices of EL, PL, BD, BZ, and naphtha by seasonality. The operational rate is calculated from the margin by using a regression equation. By multiplying the operational rate with the capacity of an Asian company's plant considering the T/A, the Asian petrochemical company's raw material demand can then be determined. A small amount of LPG is used instead of naphtha when the use of LPG as a resource produces greater profit compared with the use of naphtha. The analysis shows that the use of LPG is affected by the margin ratio of

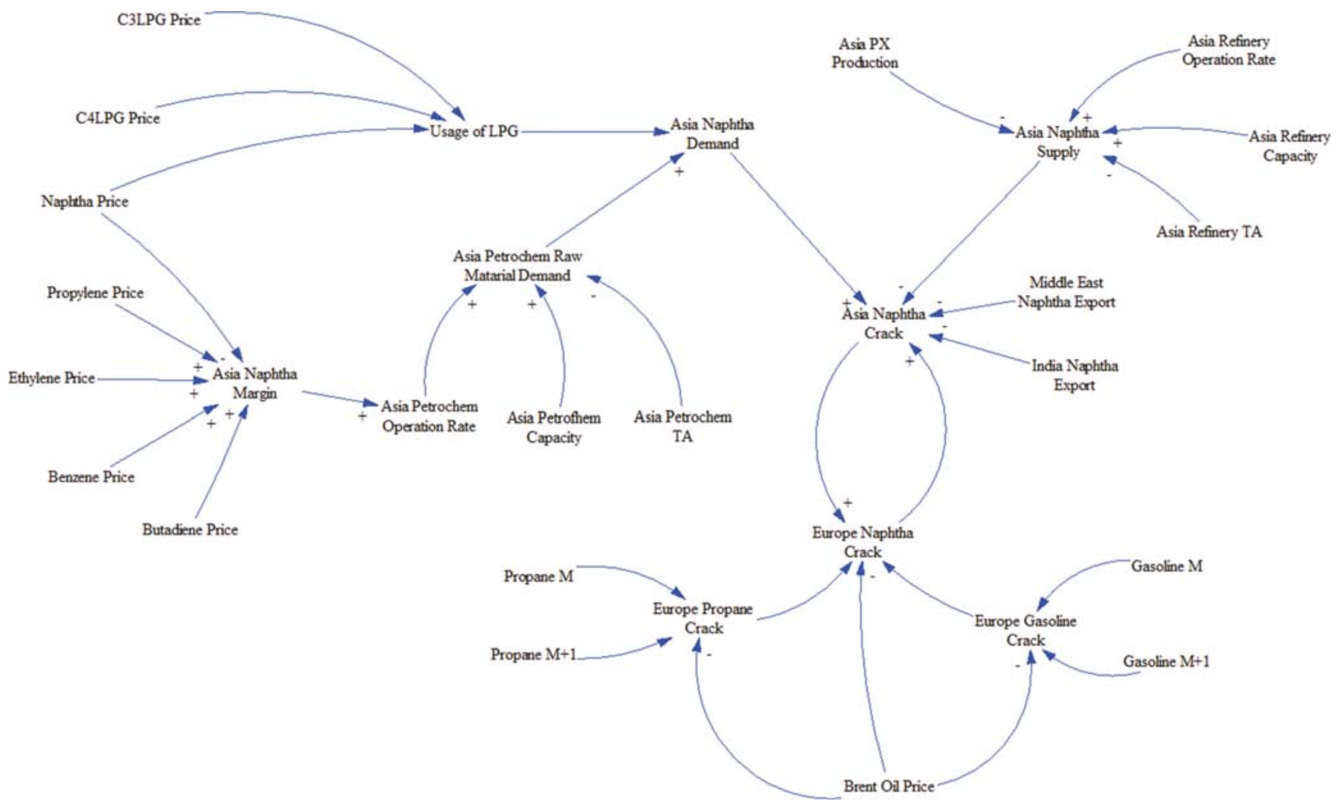

Fig. 4. SD model of the naphtha crack.

LPG and naphtha. Accordingly, the use of LPG is calculated using the regression equation from the margin ratio of LPG and naphtha. As a result, the petrochemical company's naphtha demand is calculated by subtracting the use of LPG from the Asian petrochemical company's raw material demand. The Asian naphtha supply model is similar to the demand model. It considers capacity, T/A, operational rate, and production of PX. As opposed to the demand model, the operational rate is calculated from the average of previous data and seasonality. The T/A and the production of PX value are based on the seasonality calculation. By multiplying the operational rate with the capacity of the Asian refinery's plant while considering the T/A and production of PX, the Asian refinery's naphtha supply can be calculated.

3-2. European Naphtha Crack

The European loop is considered using monthly future prices. The first month's future price and the second month's future price gaps for gasoline and propane are calculated based on their respective seasonal effect values. Using these two values and the crude oil price (Brent oil price) values, the gasoline crack and propane crack can then be calculated. Finally, the propane crack and the gasoline crack can be used to calculate the European naphtha crack. Since Europe's naphtha crack and Asia's naphtha crack affect each other, Europe's naphtha crack can be used to forecast Asia's naphtha crack.

4. Quantitative Model

The causal loop diagram requires a quantitative model that contains mathematical equations and data. Each relationship has a mathematical interaction formula. However, finding the exact equation is very hard work. Therefore, a linear expression, which is the most general form of equations, is used to express the equations.

4-1. Interaction Formula

The interaction formulas between the factors are described below in Eqs. (3)-(12). The parameters p_1-p_{20} are calculated using the linear regression method with the dataset of the previous 30 months. The Asian naphtha crack, which is the final objective variable, is calculated by intermediate variables and the intermediate variables are calculated from the primary variables. Eq. (3) represents the Asian naphtha margin calculation. It consists of the prices of the products produced from naphtha and the naphtha price.

$$\begin{aligned}&\text{Asia Naphtha Margin}=\text{Propylene Price}+\text{Ethylene Price}\\&\quad+\text{Benzene Price}+\text{Butadiene Price}-\text{Naphtha Price}\end{aligned}\tag{3}$$

The operational rate of a petrochemical plant is affected by the margin. The calculated margin from Eq. (3) is used for the calculation of an Asian petrochemical plant's operational rate as follows:

$$\text{Asia Petrochem Operation Rate}=\text{Asia Naphtha Margin}\times p_{19}+p_{20}\tag{4}$$

The calculated operational rate is used for the calculation of the Asian petrochemical plant's raw material demand. The Asian petrochemical plant's raw material demand in Eq. (5) consists of Korean, Japanese, and Taiwanese petrochemical plant capacity data from the industry, their T/A capacity, and calculated operational rates.

$$\begin{aligned}&\text{Asia Petrochem raw Material Demand}\\&\quad=\text{Asia Petrochem Operation rate}\\&\quad\times\text{Asia Petrochem Capacity}-\text{Asia Petrochem TA}\end{aligned}\tag{5}$$

Petrochemical plants use naphtha and LPG as their raw materials. Thus, the Asian naphtha demand is calculated using Eq. (6).

$$\begin{aligned}&\text{Asia Naphtha Demand}\\&\quad=\text{Asia Petrochem Raw Material Demand}-\text{Usage of LPF}\end{aligned}\tag{6}$$

The use of LPG in Eq. (7) is derived from seasonal price data. It consists of the seasonal price gap of naphtha-C3 LPG and naphtha-C4 LPG. The seasonal price is explained further in Section 4.2.

$$\begin{aligned}&\text{Usage of LPG}=(\text{Naphtha Price}-\text{C3LPG Price})\times p_{16}\\&\quad+(\text{Naphtha Price}-\text{C4LPG Price})\times p_{17}+p_{18}\end{aligned}\tag{7}$$

Asian naphtha mainly comes from refineries. Similar to Eq. (5), the naphtha supply consists of the operational rate, capacity, and T/A. Some of the produced naphtha is used as raw material for producing PX in the refinery. Therefore, Eq. (8) considers the amount of PX production.

$$\begin{aligned}&\text{Asia Naphtha Supply}=\text{Asia Refinery Operation Rate}\\&\quad\times\text{Asia Refinery Capacity}-\text{Asia Refinery TA}\\&\quad-\text{Asia PX Production}\end{aligned}\tag{8}$$

The European naphtha loop considers European propane, gasoline, and Brent oil. The propane and gasoline crack is calculated using Eqs. (9) and (10).

$$\begin{aligned}&\text{Europe Propane Crack}=(\text{Propane M}-\text{Propane M}+1)\\&\quad\times p_{12}-\text{Brent Oil Price}+p_{13}\end{aligned}\tag{9}$$

$$\begin{aligned}&\text{Europe Gasoline Crack}=(\text{Gasoline M}-\text{Gasoline M}+1)\\&\quad\times p_{14}-\text{Brent Oil Price}+p_{15}\end{aligned}\tag{10}$$

The European naphtha crack is calculated using the Asian naphtha crack, European propane crack, gasoline crack, and Brent oil price. Each price data comes from seasonal price data.

$$\begin{aligned}&\text{Europe Naphtha Crack}=\text{Asia Naphtha Crack}\\&\quad\times p_7+\text{Europe Propane Crack}\times p_8\\&\quad+\text{Europe Gasoline Crack}\times p_9+\text{Brent Oil Price}\times p_{10}+p_{11}\end{aligned}\tag{11}$$

Asian naphtha comes from refineries in the Middle East and India. Therefore, the final Asian naphtha crack calculation is shown in Eq. (12), and it considers the European naphtha crack, Asian naphtha demand, Asian naphtha supply from refineries, Middle Eastern naphtha export, and Indian naphtha export.

$$\begin{aligned}&\text{Asia Naphtha Crack}=\text{Europe Naphtha Crack}\times p_1\\&\quad+\text{Asia naphtha Demand}\times p_2+\text{Asia Naphtha Supply}\times p_3\\&\quad+\text{Middle East Naphtha Export}\times p_4\\&\quad+\text{India Naphtha Export}\times p_5+p_6\end{aligned}\tag{12}$$

4-2. Seasonal Effect

The affecting factors and affected factors in the causal loop diagram occur at the same time and change simultaneously. Therefore, the "time problem of forecasting" occurs. In other words, when we want to forecast the value of a factor "A" using the value of another factor "B," we cannot forecast the value of the A because we do not know the value of B and A at the same time. Accordingly, it is difficult to derive the equation for affecting factors and affected factors using linear regression. However, through the introduction of seasonality analysis, the time problem of forecasting can be solved. Three types of coefficients are determined from the seasonality analysis of actual data: the initial price coefficient, price increase coefficient, and seasonal effect coefficient. The coefficients are determined through model calibration using information obtained from previous data analysis and are applied in Eqs. (13), (14), and (15).

$$M = M_m * F_s. \tag{13}$$

$$M_m = M_0 + \alpha \sum_{i=1}^{m-1} M_i. \tag{14}$$

$$F_s = [f_1, f_2, f_3 \ldots\ldots f_{12}]. \tag{15}$$

Here, M indicates a time-series variable with a seasonal effect, such as the prices of EL, PL, BZ, BD, naphtha, and crude oil. It is composed of the product of the basic price M_m and monthly seasonal effect coefficients F_s, of which there are 12 monthly constant coefficients. M_m is calculated from the sum of the initial value M_0 and the product of past values and the price increase coefficient α. The analysis is conducted using Vensim, an SD software program. As an example, the seasonal effect calculation model and results of EL, PL, BD, BZ are shown in Figs. 5-8.

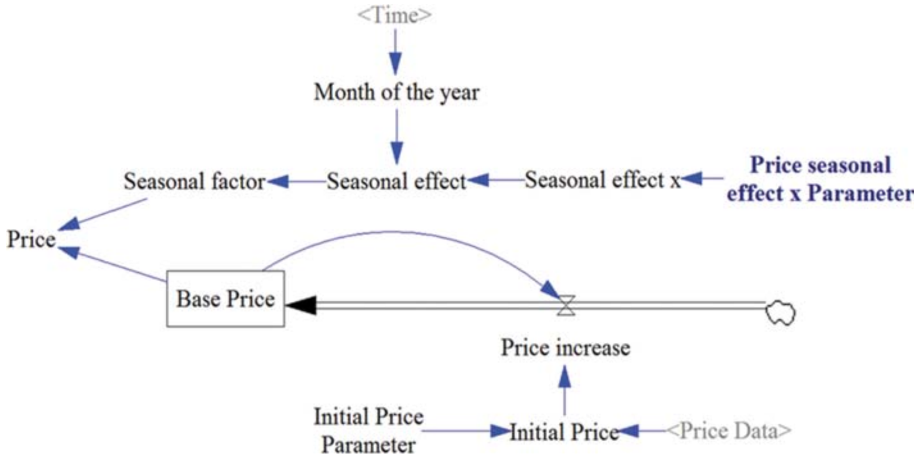

Fig. 5. Method of calculating seasonal factors.

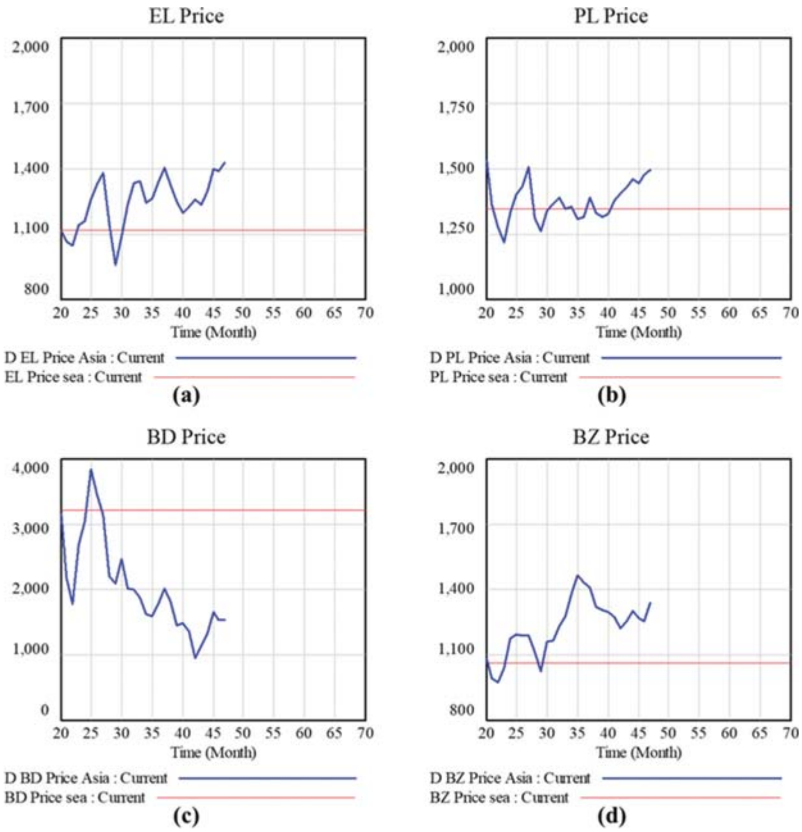

Fig. 6. Calculation procedure 1 for seasonal factor. (a) Ethylene price. (b) Propylene price. (c) Butadiene price. (d) Benzene price.

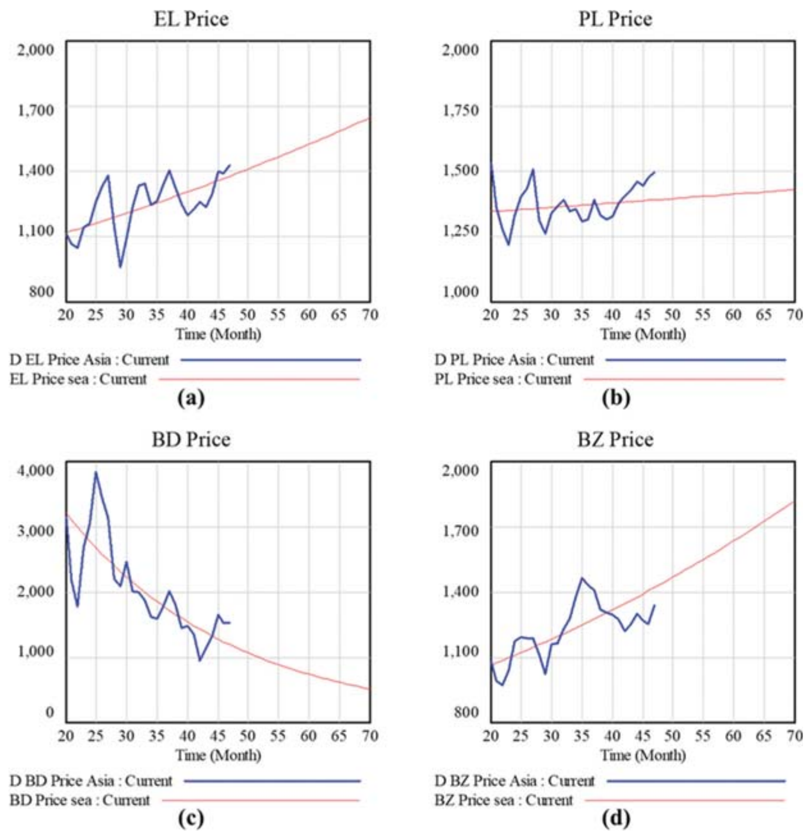

Fig. 7. Calculation procedure 2 for seasonal factors. (a) Ethylene price. (b) Propylene price. (c) Butadiene price. (d) Benzene price.

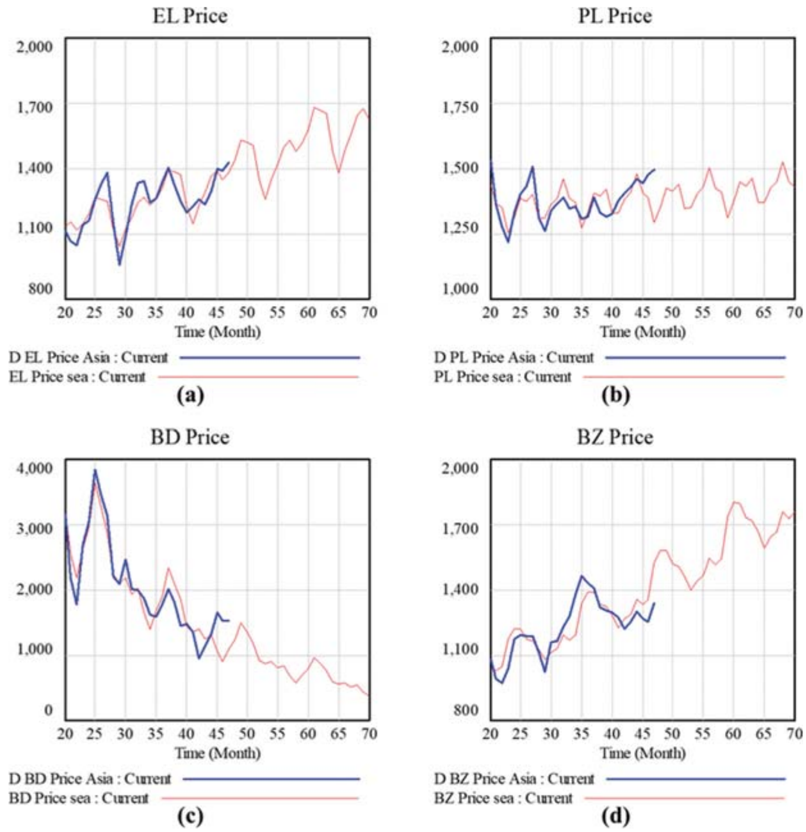

Fig. 8. Calculation procedure 3 for seasonal factors. (a) Ethylene price, (b) propylene price, (c) butadiene price, and (d) benzene price.

Table 1. List of factors obtained from seasonality analysis

Ethylene price	Asian refinery T/A
Propylene price	Asian refinery operational rate
Butadiene price	Asian PX production
Benzene price	Middle Eastern naphtha export
Naphtha price	Indian naphtha export
Brent oil price	Propane M
C3LPG price	Propane M+1
C4LPG price	Gasoline M
Asia petrochem T/A	Gasoline M+1

The development of the seasonal effect model involves the following three steps. The first step, as shown in Fig. 6, is to define the initial value of the factor. Generally, this value is the same as the initial value of the real data; however, when an actual value does not show an increasing or decreasing trend, it is set to the middle value between the maximum and the minimum value. The second step, as shown in Fig. 7, is to define the slope of the line. An exponential smooth curve is used as a basic curve, and its slope variation is determined by the price increase coefficient. Fig. 8 shows the final step; each value is multiplied by a monthly seasonal factor. In this way, the seasonal effect model is completed. Table 1 shows a list of every factor whose value is based on seasonality analysis. These factors have the same-time value according to the seasonality. Additionally, regression equations among the factors were applied to the forecasting model.

4-3. Data Preparation and Parameter Calculation

Before the forecasting procedure, input data preparation and calculation of the parameter values in Eqs. (3)-(12) are required. The input data are fetched from actual field and seasonality analyses; the parameter values are calculated from the training region. After the preparation and calculation, the data and parameter values are applied to the model.

As shown in Fig. 9, the forecasting models work on a cycle that includes a 30-month training and three-month forecasting period. The overall cycle is repeated, and after each cycle, the schedule period is shifted a month later. The overall forecasting range is demonstrated from April 2012 to December 2013.

RESULTS

The objective of forecasting is to obtain the direction of the naphtha crack (increasing or decreasing) not the value because this model is expected to be used in actual decision-making processes in the field. To make more profit, companies want to know the best timing to purchase their resources. Thus, it is more important to determine whether the crack is increasing or decreasing rather than obtaining an accurate value. Of course, a model that produces highly accurate values would be extremely helpful in the decision-making process; however, it is very hard to construct a regression model with high accuracy. Thus, this research focuses on the directional accuracy of the forecasted crack rather than the accuracy of the obtained value. If the forecasting model shows similar or better directional accuracy, it can assist in industrial deci-

Fig. 9. Time window movement for the forecasting of the naphtha crack.

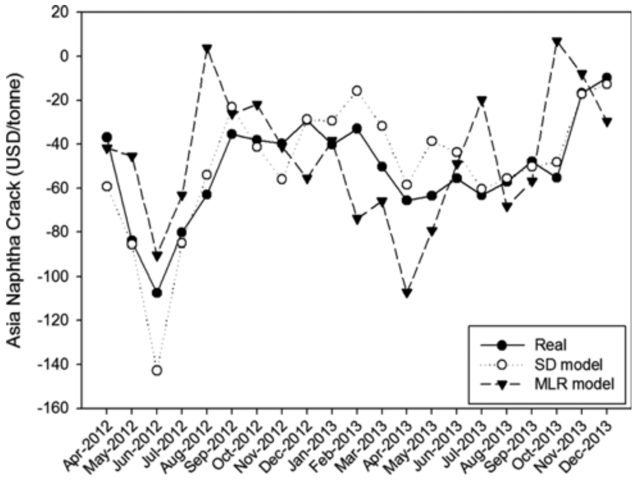

Fig. 10. One-month forecasting results for the naphtha crack.

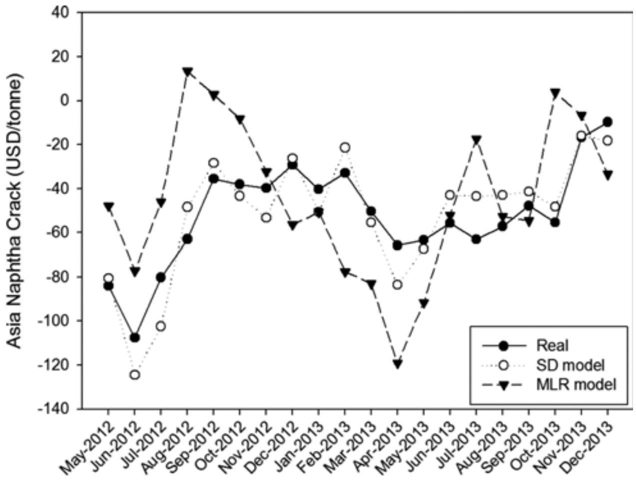

Fig. 11. Two-month forecasting results for the naphtha crack.

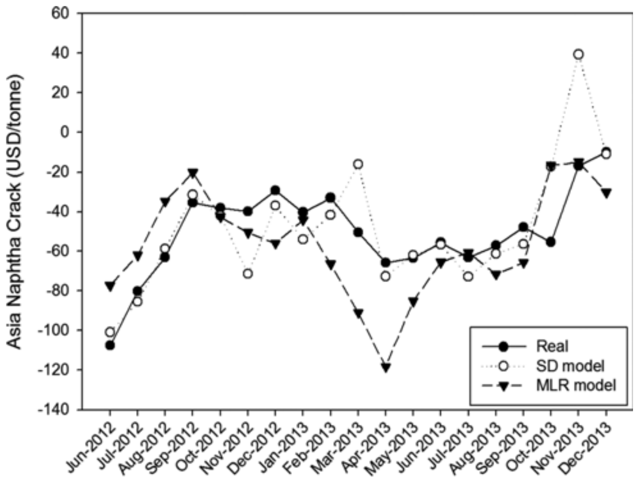

Fig. 12. Three-month forecasting results for the naphtha crack.

sion-making processes. The forecasting results for the naphtha crack are shown in Figs. 10-12 and Tables 2 and 3.

The forecasting in this study focused on Asia's naphtha crack and was conducted using two models: a simple MLR model and an SD model based on the causal loop. The overall forecasting period was from April 2012 to December 2013 (21 months). As shown in Table 6, the SD model has better directional accuracy than the MLR model. The one-month forecasting result shows the highest directional accuracy, i.e., 95.24%, with a forecasting period of 21 months. The two-month forecasting result exhibits a directional accuracy of 85.00% with a forecasting period of 20 months. Finally, the three-month forecasting result shows a directional accuracy of 84.21% with a forecasting period of 19 months. However, the MLR model has a lower directional accuracy, i.e., 61.90% for the one-month period, 55.00% for the two-month period, and 52.63% for the three-month period. Regarding short-term directional agreement, the forecasting of the SD model is highly accurate. Compared with the MLR model, which does not consider same-time variables, the proposed model shows much better directional accuracy in terms of actual data. As shown in Figs. 10-12, the accuracy of the actual value of the SD model is not significantly higher than that of the MLR model. However, the objective of this model is to forecast the directional accuracy, not the actual value accuracy because increasing or decreasing of crack is key point for decision makers. Therefore, on the basis of the accuracy of direction, the SD model shows higher accuracy compared with the MLR model. As the SD model shows better directional accuracy, it is expected that the SD model will be more helpful in the decision-making process. The best quality of this model is that it can explain why the crack is increasing or decreasing using the causal loop. Thus, this model produces good results with a reasonable description.

CONCLUSIONS

Naphtha is the most important resource in the petrochemical industry because a petrochemical company's profit is primarily affected by the price of naphtha. Since there are many factors and cases that must be considered, forecasting the price of naphtha is very difficult but of great interest to petrochemical companies. The objective of this research is to define the price gaps (cracks) of naphtha and crude oil and forecast their fluctuations. The naphtha crack is crucial in the determination of whether naphtha needs to be purchased, which is the reason behind forecasting the crack rather than the price. Forecasting using the SD model is based on the causal loop of the affecting factors. Herein, the calculations were performed using the linear regression method. To increase the forecasting accuracy, the seasonal effect of each factor was also considered. Using the causal loop, seasonal effect, and regression equations, an SD model was developed. The one-, two-, and three-month model forecasts for naphtha crack exhibited directional accuracies in the range of 84%-95%. This model has three main features: i) it focuses on the crack rather than the price, ii) it explains the causal relationships, and iii) it aims to accurately forecast not the value but the direction. The motivation and objective of this study were to realize practical applications of the proposed model. Therefore, the developed model has the features that are described herein and is being used by a petrochemical company

Table 2. Forecasting results for the naphtha crack (USD/ton)

	Real naphtha crack		1-Month forecasting naphtha crack SD model			1-Month forecasting naphtha crack regression model			2-Month forecasting naphtha crack SD model			2-Month forecasting naphtha crack regression model			3-Month forecasting naphtha crack SD model			3-Month forecasting naphtha crack regression model		
	Value	Vari-ation	Value	Vari-ation	Matching status	Value	Vari-ation	Matching status	Value	Vari-ation	Matching status	Value	Vari-ation	Matching status	Value	Vari-ation	Matching status	Value	Vari-ation	Matching status
Apr-12	−36.92	−	−59.34	−	○	−41.72	−	○												
May-12	−84.05	−	−85.59	−	○	−45.42	−	○	−80.63	−	○	−47.88	−	○						
Jun-12	−107.59	−	−142.85	−	○	−90.25	−	○	−124.60	−	○	−77.26	+	×	−100.93	−	○	−77.34	+	×
Jul-12	−80.29	+	−85.04	+	○	−63.12	+	○	−102.44	+	○	−45.97	+	○	−85.55	+	○	−61.98	+	○
Aug-12	−63.03	+	−54.11	+	○	3.78	+	○	−48.50	+	○	13.38	+	○	−58.81	+	○	−34.77	+	○
Sep-12	−35.55	+	−23.15	+	○	−26.10	+	○	−28.46	+	○	2.77	+	○	−31.52	+	○	−20.20	+	○
Oct-12	−38.12	−	−41.39	−	○	−21.82	+	×	−43.42	−	○	−8.28	+	×	−41.92	−	○	−42.55	−	○
Nov-12	−39.82	−	−56.01	−	○	−41.08	−	○	−53.29	−	○	−32.30	+	×	−71.45	−	○	−50.49	−	○
Dec-12	−29.21	+	−28.76	+	○	−55.51	−	×	−26.45	+	○	−56.44	−	×	−36.93	+	○	−56.02	−	×
Jan-13	−40.32	−	−29.45	−	○	−38.40	−	○	−50.65	−	○	−50.72	−	○	−54.04	−	○	−44.07	−	○
Feb-13	−32.95	+	−15.76	+	○	−73.88	−	×	−21.43	+	○	−77.58	−	×	−41.56	−	×	−66.43	−	×
Mar-13	−50.37	−	−31.73	+	×	−65.91	−	○	−55.38	−	○	−82.99	−	○	−16.14	+	×	−90.97	−	○
Apr-13	−65.74	−	−58.50	−	○	−107.17	−	○	−83.49	−	○	−119.16	−	○	−72.67	−	○	−118.22	−	○
May-13	−63.49	+	−38.71	+	○	−79.30	−	×	−67.42	−	×	−91.63	−	×	−61.94	+	○	−85.22	−	×
Jun-13	−55.65	+	−43.71	+	○	−48.94	+	○	−42.95	+	○	−52.07	+	○	−56.77	+	○	−65.55	−	×
Jul-13	−63.19	−	−60.48	−	○	−19.92	+	×	−43.51	+	×	−17.44	+	×	−72.92	−	○	−60.62	−	○
Aug-13	−57.15	+	−55.54	+	○	−68.10	−	×	−42.95	+	○	−52.68	+	○	−61.26	+	○	−71.62	−	×
Sep-13	−47.86	+	−50.19	+	○	−56.77	+	○	−41.39	+	○	−54.67	+	○	−56.33	+	○	−65.69	−	×
Oct-13	−55.43	−	−48.11	−	○	7.02	+	×	−48.31	−	○	3.79	+	×	−17.27	+	×	−16.67	+	×
Nov-13	−16.81	+	−17.25	+	○	−8.04	+	○	−15.99	+	○	−6.61	+	○	39.34	+	○	−14.92	+	○
Dec-13	−9.82	+	−12.70	+	○	−29.56	−	×	−18.26	−	×	−33.48	−	×	−10.93	+	○	−30.15	−	×

Table 3. Comparison of directional accuracy

	1 Month	2 Months	3 Months
MLR model	61.90%	55.00%	52.63%
SD model	95.24%	85.00%	84.21%

for decision-making. However, there are many factors that affect the naphtha crack. The proposed model can be made more reliable and the accuracy of its output can be increased by considering more factors, such as the effect of shale gas and new shipping courses.

ACKNOWLEDGEMENTS

Financial support from Hanwha Total Co., Ltd. is gratefully acknowledged.

LIST OF ABBREVIATIONS

MILP : mixed integer linear programming
ARIMA : auto regressive integrated moving average
SVMs : support vector machines
ANN : artificial neural networks
ANFIS : adaptive neuro-fuzzy inference system
LSSVM : least-squares support vector machine
ARMAX : autoregressive moving average with external input
LLR : local linear regression
DLLR : dynamic local linear regression
MLR : multiple linear regression
SD : system dynamics
PPMCC : pearson product moment correlation coefficient
EOS : east of Suez Canal
WOS : west of Suez Canal
T/A : turnaround
EL : ethylene
PL : propylene
BD : butadiene
BZ : benzene
LPG : liquefied petroleum gas
PX : para-xylene

REFERENCES

1. G. Dooley and H. Lenihan, *Resour. Policy*, **30**, 208 (2005).
2. J.-H. Ryu, *Korean J. Chem. Eng.*, **30**, 27 (2013).
3. M. Lee and J. Kim, *Korean J. Chem. Eng.*, **34**, 1604 (2017).
4. M. Kim and J. Kim, *Int. J. Hydrog. Energy*, **42**, 3899 (2017).
5. P.-F. Pai and C.-S. Lin, *Omega*, **33**, 497 (2005).
6. J.-M. Brault, R. Labib, M. Perrier and P. Stuart, *Can. J. Chem. Eng.*, **89**, 901 (2011).
7. X.-G. Zhou, L.-H. Liu, W.-K. Yuan and J. L. Hudson, *Can. J. Chem. Eng.*, **74**, 638 (1996).
8. S. Mandal and A. K. Jana, *Int. J. Hydrog. Energy*, **38**, 1244 (2013).
9. L. M. Ochoa-Estopier, M. Jobson and R. Smith, *Comput. Chem. Eng.*, **59**, 178 (2013).
10. B. Szkuta, L. Sanabria and T. Dillon, *IEEE Trans. Power Syst.*, **14**, 851 (1999).
11. R. Gareta, L. M. Romeo and A. Gil, *Energy Conv. Manag.*, **47**, 1770 (2006).
12. R. Jammazi and C. Aloui, *Energy Econ.*, **34**, 828 (2012).
13. A. J. Conejo, M. A. Plazas, R. Espinola and A. B. Molina, *IEEE Trans. Power Syst.*, **20**, 1035 (2005).
14. P. Visetsripong, P. Sooraksa, P. Luenam and W. Chaimongkol, *SICE Annual Conference*, 659-663 (2008).
15. X. Yan and N. A. Chowdhury, *Int. J. Electr. Power Energy Syst.*, **53**, 20 (2013).
16. J. Myklebust, A. Tomasgard and S. Westgaard, *OPEC Energy Review*, **34**, 82 (2010).
17. N. Salehnia, M. A. Falahi, A. Seifi and M. H. M. Adeli, *J. Nat. Gas Sci. Eng.*, **14**, 238 (2013).
18. K. Tak, J. Kim, H. Kwon, J. H. Cho and I. Moon, *Korean J. Chem. Eng.*, **33**, 1999 (2016).
19. D. Manca, *Comput. Chem. Eng.*, **57**, 3 (2013).
20. R. Rasello and D. Manca, *Comput. Aided Chem. Eng.*, 433 (2014).
21. K. J. Lee, T. H. Lee, L.-H. Kim and Y. K. Yeo, *Korean J. Chem. Eng.*, **28**, 1505 (2011).
22. T. H. Lee, K. J. Lee, B. W. Jo, L. H. Kim and Y. K. Yeo, *Korean J. Chem. Eng.*, **28**, 1331 (2011).
23. C. Sung, H. Kwon, J. Lee, H. Yoon and I. Moon, *Comput. Aided Chem. Eng.*, 145 (2012).
24. B. Lyu, H. Kwon, J. Lee, H. Yoon, J. Jin and I. Moon, *Comput. Aided Chem. Eng.*, 829 (2014).
25. H. Kwon, B. Lyu, K. Tak, J. Lee, J. H. Cho and I. Moon, *Comput. Chem. Eng.*, **84**, 226 (2016).
26. H. Kwon, K. Tak, J. H. Cho, J. Kim and I. Moon, *Ind. Eng. Chem. Res.*, **56**, 1267 (2017).
27. X. Tang, B. Zhang, M. Höök and L. Feng, *Energy*, **35**, 3097 (2010).
28. A. Aslani, P. Helo and M. Naaranoja, *Appl. Energy*, **113**, 758 (2014).
29. M. A. Rendon-Sagardi, C. Sanchez-Ramirez, G. Cortes-Robles, G. Alor-Hernandez and M. G. Cedillo-Campos, *Appl. Energy*, **123**, 358 (2014).
30. H. Qudrat-Ullah, *Energy*, **59**, 285 (2013).
31. R. Rehan, M. A. Knight, A. J. A. Unger and C. T. Haas, *Tunn Undergr Sp Tech*, **39**, 116 (2014).
32. O. Erdem, E. Ceyhan and Y. Varli, *Physica A: Statistical Mechanics and its Applications*, **414**, 274 (2014).