Advanced Engineering Informatics 51 (2022) 101443

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Data science and reinforcement learning for price forecasting and raw material procurement in petrochemical industry

Chia-Yen Lee [a,*], Bai-Jian Chou [b], Chen-Feng Huang [c]

[a] Department of Information Management, National Taiwan University, Taipei City 106, Taiwan
[b] Institute of Manufacturing Information and Systems, National Cheng Kung University, Tainan City 701, Taiwan
[c] LCY Chemical Corp, Taipei City 105, Taiwan

ARTICLE INFO

Keywords:
Raw material procurement
Price forecasting
Deep learning
Reinforcement learning
Digital transformation

ABSTRACT

Petrochemical industry is one of the major sectors contributing to the world-wide economy and the digital transformation is urgent to enhance core competence. In general, ethylene, propylene and butadiene, which are associated with synthetic chemicals, are the main raw materials of this industry with around 70–80% cost structure. In particular, butadiene is one of the key materials for producing synthetic rubber and used for several daily commodities. However, the price of butadiene fluctuates along with the demand–supply mismatch or by the international economy and political events. This study proposes two-stage data science framework to predict the weekly price of butadiene and optimize the procurement decision. The first stage suggests several the price prediction models with a comprehensive information including contract price, supply rate, demand rate, and upstream and downstream information. The second stage applies the analytic hierarchy process and reinforcement learning technique to derive an optimal policy of procurement decision and reduce the total procurement cost. An empirical study is conducted to validate the proposed framework, and the results improve the accuracy of price forecasts and the procurement cost reduction of the raw materials.

1. Introduction

The petrochemical industry refers to the industry that manufactures chemicals using petroleum or natural gas as raw materials, and its finished products are called petrochemical products. The scope of the petrochemical industry does not include raw materials such as oil, natural gas, gasoline, and naphtha, but refers to upstream industries that cover the raw materials such as ethylene, propylene, butadiene, benzene, toluene, and xylene, and these chemical reactions on polymerization, esterification, alkylation are made into the synthetic resins and plastics, synthetic rubber, human fiber raw materials and petrochemicals in the mid-stream industries.

In general, ethylene, propylene and butadiene, which are associated with synthetic chemicals, are the main raw materials of this industry with around 70–80% cost of finished goods. In particular, butadiene (BD) is one of the key materials for synthetic rubber and used for several daily commodities. Butadiene is generated from naphtha by catalytic cracking and is the main raw material for producing synthetic rubber such as polybutadiene rubber (PBR), styrene-butadiene rubber (SBR), acrylonitrile–butadienestyrene (ABS), adiponitrile (ADN), specialty polymers, etc. The final products of butadiene are like tires, shoe materials, various kinds of people's livelihood and industrial supplies, etc., which are closely related to the daily life supplies of the public. However, the price of butadiene fluctuates along with the demand–supply mismatch, the international economic fluctuations, and world-wide political events.

There are three factors which significantly affect the price fluctuation of BD. First, the market effect caused by the gap between supply and demand: the supply side is limited since BD is a by-product of ethylene and cracking plant is built for ethylene demand. More ethylene demand, more ethylene supply, and thus more BD obtained. The demand side is affected by the booming cycle of the world-wide environment. For example, the growth of China's auto market in recent years has driven the demand for auto tires, which has affected the synthetic rubber industry and the development of other downstream products of BD. Second, the change of production cost caused by the fluctuation of crude oil price: the main raw material for butadiene is Crude C4 from naphtha cracking plant, and the upstream of Crude C4 is crude oil, which means that the price of crude oil directly affects the cost of BD. Third, the psychological reaction of market sentiment caused by the fluctuation of

* Corresponding author.
E-mail address: chiayenlee@ntu.edu.tw (C.-Y. Lee).

https://doi.org/10.1016/j.aei.2021.101443
Received 17 January 2021; Received in revised form 29 September 2021; Accepted 18 October 2021
Available online 3 November 2021

crude oil price. Since raw material purchasers attempt to buy with lower price and build the inventory level, and thus they can earn the spread if the future price increase. That is, the fluctuation of crude oil price leads the market sentiments of BD price. This phenomenon can be seen everywhere and critical, but it is difficult to measure and control. Fig. 1 illustrates the factors affecting BD price except market sentiments.

Recently, affected by the crude oil production reduction agreement made by the Organization of Petroleum Exporting Countries (OPEC) and the geopolitical tensions in the Middle East, the international crude oil prices have shown a dramatic fluctuation and uncertainty to the world economy. This impact influences the production capacity and price of the downstream markets. In order to respond to the fluctuation of oil prices, we need to predict the prices of raw materials and clarify their future trends. That is, providing the relevant information on the future market development to management can effectively support the procurement decision and operational strategy.

Thus, from digital transformation aspect, this study defines the scope and focus of the transformation on the BD procurement decision, and addresses this issue through technology adoption using the improved business tools and emerging technologies for cost-effectiveness improvement [1,2,30]. Particularly, the digital data-driven technology (i.e. big data analytics) and the artificial intelligence (AI)-based technology (i.e. reinforcement learning) are two digital transformation enablers developed for BD price forecasting and procurement decision optimization. In addition, experience and knowledge of the procurement staff is urgent to stored due to the talent gap and the declining birth rate. This study suggests using AI method to structurize the knowledge of the practical procurement decision. The digital enabled automating and supporting knowledge intensive tasks is critical for knowledge management and bringing digital transformation benefits to enhance core competence [1,5,17].

In practice, price prediction and procurement decision are two activities and usually separated into different divisions. This causes a disconnection between these two and undermines the integration of more comprehensive factors across divisions for making better decision. This study fills the gap between these two and proposes a two-stage data science framework to support price prediction and procurement decision of the raw materials. The first stage “Price Forecasting” collects the historical prices, contract prices, capacity operating rate (i.e. rate at which a factory utilizes its capacity), and downstream substitutes in the whole supply chain, and then build a price prediction model by recurrent neural network. The second stage “Procurement Decision” collects historical inventory, demand and procurement data, and price forecast results from the previous stage, and then builds a procurement decision model by RL. Finally, based on the results of price prediction, the procurement strategy is generated to minimize the total procurement cost and supports proactive decision-making. An empirical study of a petrochemical manufacturer is conducted and we collect data of the butadiene procurement from 2009 to 2018 to validate the proposed two-stage data science framework. To the best of our knowledge, there is no previous study proposing an integrated framework embedded with RL used in procurement decision of purchasing raw materials.

The remainder of this paper is organized as follows. Section 2 reviews the related work in literature. Section 3 introduces the fundamentals and methodologies used in this study. Section 4 presents the proposed two-stage data science framework including the BD price prediction model and procurement decision model. Section 5 describes an empirical study of a petrochemical manufacturer and validates the proposed two-stage framework. Section 6 concludes.

2. Related work in literature

There are several studies predicting the raw material price (eg. crude oil price) in literature. Ref. [3] suggested a forecasting model by ordinary least squares (OLS) to predict the 1-, 3-, 6-, 9-, and 12-month ahead crude oil prices for 1991 to 2001. They tested the unbiasedness of the OLS model and found that the predictive accuracy is outperformed by the 1- and 12-month forecasts. Ref. [4] built the Autoregressive Integrated Moving Average (ARIMA), backpropagation neural network (BPNN), and support vector machine (SVM) to forecast the crude oil price in 2000–2003. The results showed that SVM performed well in time series dataset. Ref. [6] suggested a divide-and-conquer strategy and proposed a hybrid method which decomposed the oil price into several independent components called intrinsic mode functions (IMFs) and the residuals by the empirical mode decomposition (EMD). They applied SVM with the particle swarm optimization (PSO-SVM) method and the generalized autoregressive conditional heteroskedasticity (GARCH) model to forecast the nonlinear and time-varying components in 2013, respectively, and then summed all the forecasted components as the final forecasted values of crude oil prices. Ref. [7] developed two deep learning techniques including stacked denoising autoencoders (SDAE) and bootstrap aggregation to forecast the oil prices for 1986–2016.

For the procurement decision, there are several studies focusing on the optimal production planning or contract selection. Given demand scenarios, [26] considered a petroleum supply chain structure including crude sector, refining sector, petrochemicals sector, and downstream

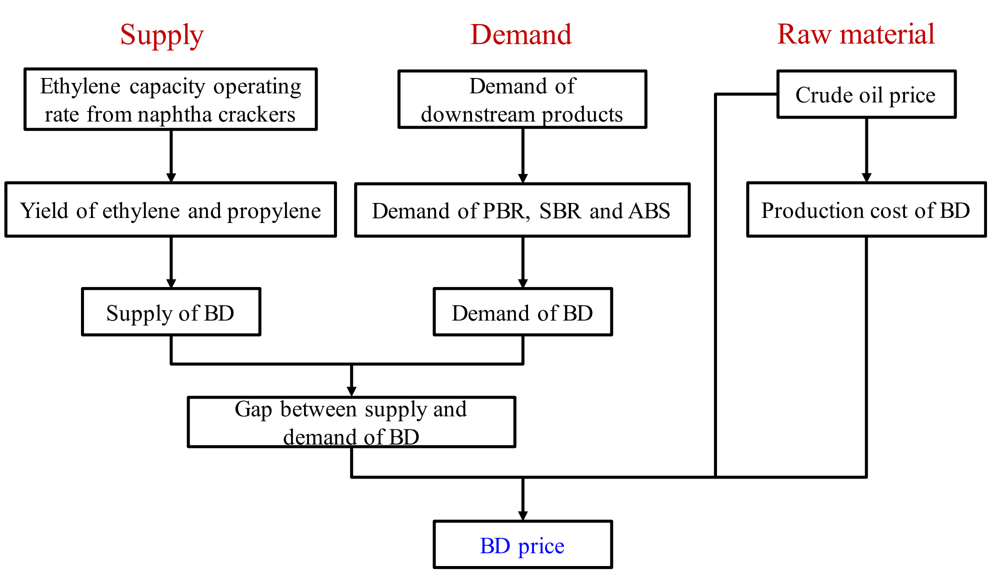

Fig. 1. Factors affecting BD price.

chemicals sector, and proposed multi-period stochastic linear programming model to simulate the effect of uncertainties in market prices and demands. Ref. [27] used optimal control theory to consider JIT procurement, backlogging, and warehouse capacity for minimize the net present value of the total cost. Ref. [28] considered the contract types including discount, bulk discount, and fixed duration, and proposed a mixed-integer nonlinear two-stage stochastic programming to select optimal contract under supply and demand uncertainty. The approach was manufacturer-centric to sign or not contracts with suppliers in order to hedge against spot market supply uncertainty, however, the sales contract related to customer's decision was not considered. Ref. [29] proposed the geometric Brownian motion and built a multi-period stochastic dynamic programming (SDP) to optimize the refinery operation. The solution approach (SDP) was developed through an orthogonal array of fractional factorial design for sampling method and a multivariate adaptive regression splines for value function approximations.

However, previous studies focused on developing prediction models and optimal production planning separately, and few studies connected the forecasted result to the practical procurement decision. In addition, the mathematical programming builds the decision over time using variables with subscript timestamp (eg. daily, weekly, monthly) and then lots of the decision variables generated lead to computational burden (particularly with integer/binary variables). For addressing uncertainty, stochastic programming or robust optimization are suggested, however, with more variables or constraints. Reinforcement learning (RL) is successfully applied to some uncertain and dynamic environment. It uses dynamic programming to formulate a sequential decision with different stages over time, and adapts to rapidly-changing environment (eg. demand fluctuation and price volatility) via constant model training from data. RL not only optimizes the procurement decision in a stochastic environment, but also generates an optimal policy which provides the best action corresponding to the state (i.e. environment) we face. Thus, this study develops time-rolling price forecast with the practical factors and suggests RL to offset the procurement risks when the real price fluctuates.

Generally, there are three types of procurement methods of raw materials in the petrochemical industry: (1) The downstream of the petrochemical industry signs a contract directly with the naphtha cracking plant and receives a stable supply with a fixed contract price formula each month; (2) the intermediate traders buy raw materials from the upstream plants, and then sell them to the downstream at the contract or spot price of the market; (3) the surplus goods caused by over-production of cracking plants or production shutdown/decrease of the downstream customers are usually sold to a spot market. Since a stable supply from the contract with a fixed contract price formula [28], the scope of procurement discussed in this study focuses on the spot market, which shows a significant fluctuation of the BD price. In practice, in order to cope with the uncertain and urgent demand in the downstream, a firm may need to buy the raw materials from the spot market one or two times per month. Thus, this study shows how to procure in the spot market.

3. Fundamentals

This section introduces some fundamental techniques of data science and RL. Least absolute shrinkage and selection operator (Lasso) and random forest (RF) are used for feature selection (FS) to identify important variables affecting BD price. Long short-term memory (LSTM) and RL are used for price forecasting and optimal procurement decision, respectively. Analytic hierarchy process (AHP) is used to extract the preference structure from decision maker and form the rewards in RL. These techniques used for data science framework optimizing the BD procurement decision in Section 4 are described as follows.

3.1. Random forest (RF)

One of the variable selection techniques used in this study is RF, which can select important factors significantly affecting BD price. Tree structure forms a nonlinear model and partially considers the interaction effect among multiple variables. RF is an ensemble learning which combines multiple weak classifiers (eg. classification and regression tree (CART) or C4.5) to form a strong classifier by randomly resampling the subsets of the dataset (i.e. bagging) [8,9]. RF divides the raw dataset into training data and out-of-bad (OOB) data. Resample subsets from the training data by bootstrapping procedure and create several CARTs. Integrate several trees by unweighted majority vote for predicting response variable, and evaluate the results by OOB dataset. OOB can also provide the importance of each factors.

3.2. Least absolute shrinkage and selection operator (Lasso)

Lasso [10] identifies the relevant independent variables that significantly affect the response variable by suggesting the L1 regularization term in the ridge regression. Lasso shrinks the coefficients of the independent variable to zero by adjusting the penalty parameter. Let β_0 and β_p be the intercept and coefficients of independent variables in the regression. The Lagrangian form of Lasso can be formulated as:

$$\widehat{\beta}^{Lasso} = \underset{\beta}{argmin}\left\{\frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}x_{ij}\beta_j\right)^2 + \eta\sum_{j=1}^{p}|\beta_j|\right\} \tag{1}$$

where $\eta\sum_{j=1}^{p}|\beta_j|$ is a LASSO penalty term and $\eta \geq 0$ is a predetermined penalty parameter controlling the shrinkage. If η is small enough, the penalty term is close to zero and the estimator becomes the ordinary least squares (OLS); however, given a large enough η, the estimates of β_j shrink toward zero. Thus, η parameter decides the selected variables in the regression model. In fact, increasing the parameter η increases the bias and decrease the variance of β_j. Lasso can be used in small data with a larger number of variables. Lasso has two limitations [11]. Lasso is sensitive to outliers and a bias estimate of β_j may cause that some relevant factors cannot be identified.

3.3. Long short-term memory (LSTM)

The traditional neural networks are developed with feedforward or backpropagation mechanism, and each input is independent and there is no design of memory function in the network structure. It is difficult to deal with a complicated sequential problem which human brain can handle, such as linguistic materials, self-correlated signals, or time-series dataset. Recurrent neural network (RNN) with memory function was developed to address the issue by using the memory cell [12]. Each input can be temporarily memorized with a storage unit and is available input of training next neuron with new observation.

LSTM is one type of RNNs commonly used for time-series dataset. The structure of LSTM with the input gate, the forget gate, and the output gate as shown in Fig. 2 [13]. When the input gate is turned on, the last output can be written to the cell, otherwise the cell ignores the written information. Forget gate is used to control whether the memory cells clear the previously-stored content. Output gate is used to control the availability of the information from the memory cells. The "ON" or "OFF" of these gates are determined by training and learning process.

To simplify the LSTM structure, the gated recurrent units (GRUs) was developed with only two gates: reset gate which determines how to integrate the new input with the previous memory, and update gate which defines how much of the previous memory to keep around [14,15]. GRU has several variants and we introduce a typical one. Let s_t describe the output state vector, x_t be the input vector, z_t be the update gate vector, r_t be the reset gate vector. Let W_z, U_z, and b_z, be the parametric matrices and intercept vector of update gate. Let W_r, U_r, and b_r,

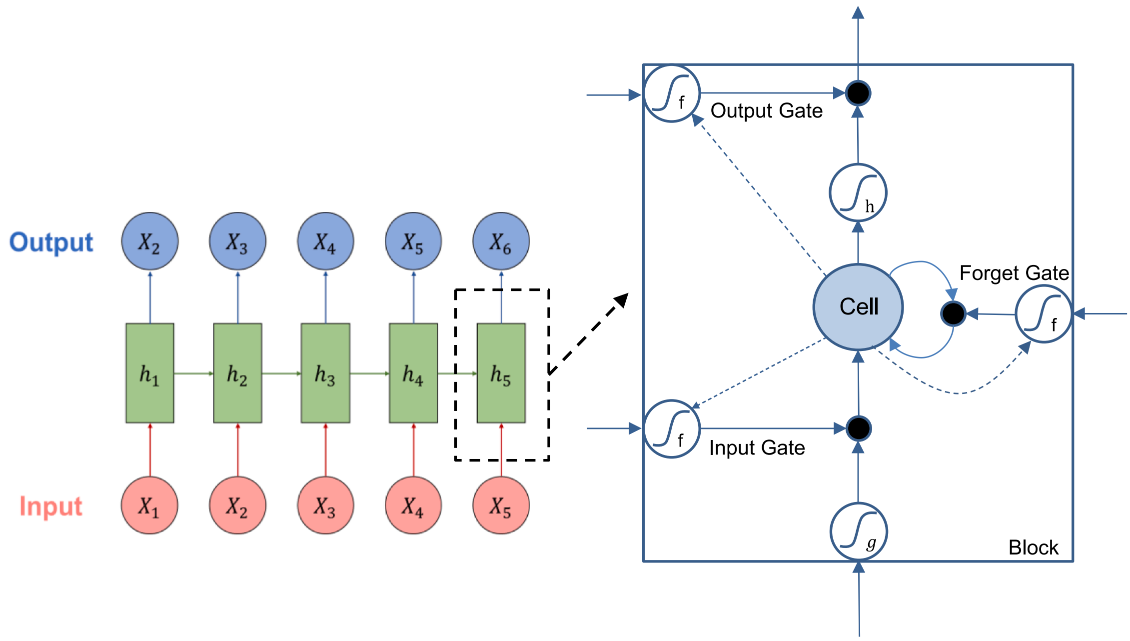

Fig. 2. LSTM [13].

be the parametric matrices and intercept vector of reset gate. Let W_s, U_s, and b_s, be the parametric matrices and intercept vector for calculating the state candidate. The notation φ represents a sigmoid function. The operator $^\circ$ denotes the Hadamard product. s_t is a hyperbolic tangent function representing the state candidate. The GRU formulates the following equations.

$$z_t = \varphi(W_z x_t + U_z s_{t-1} + b_z) \tag{2}$$

$$r_t = \varphi(W_r x_t + U_r s_{t-1} + b_r) \tag{3}$$

$$s_t = \tanh(W_s x_t + U_s(r_t{}^\circ s_{t-1}) + b_s) \tag{4}$$

$$s_t = z_t{}^\circ s_{t-1} + (1 - z_t)^\circ s_t \tag{5}$$

Fig. 3 illustrates the GRU. Update gate is formulated by the Eq. (2) and forgets some information from previous state, where z_t determines how much of the previous memory to retain. Reset gate integrates the new input with the previous memory by Eq. (3) and generates the state candidate by Eq. (4). Finally, the new output state s_t is calculated by Eq. (5).

3.4. Reinforcement learning (RL)

In RL, Markov decision process (MDP) is a mathematical framework that describes an agent interacting with an uncertain environment and supports the stochastic and sequential decision making [16]. For each time step, the agent takes an action a in a certain state s. After the action is performed, the agent receives a reward r and the environment transits from current state s to a next state s' based on the transition probability $P^a_{s,s'}$. The state transitions of MDP satisfies the Markov memoryless property; that is, the new state s' only depends on current state s and is conditionally independent to all previous states and actions.

The elements of MDP are defined as follows. Let $\mathscr{A}$ be a set of actions, $\mathscr{S}$ be a set of states, $\mathscr{P}$ is a transition probability function from state s to state $s^{'}$ after taking action a, $\mathscr{R}$ is a reward function after taking action a from state s to state $s^{'}$, $\mathscr{T}$ is the set of decision time step, which is finite in this study, and π is a policy which is a mapping from perceived states of the environment to actions to be taken when in those states s. The goal of a MDP is to find an optimal policy π^* to minimize the expected total reward from state-value function $V_\pi(s)$, which is defined as follows [18]:

$$V_\pi(s) = \mathbb{E}_\pi\left[\sum_{t=0}^{T} \gamma^t R\left(s_t^{'} | s_t\right)\right] \tag{6}$$

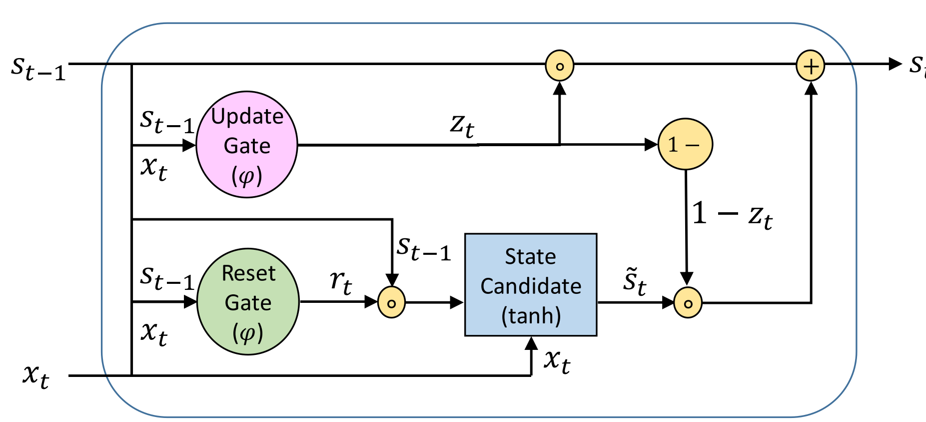

Fig. 3. Gated Recurrent Unit (GRU).

where $\mathbb{E}_\pi[\bullet]$ denotes the expected value of a random variable and γ is the discounting factor. Similarly, we define another value of taking action a in state s under a policy π, denoted action-value function $Q_\pi(s,a)$, which is defined as follows,

$$Q_\pi(s,a) = \mathbb{E}_\pi\left[\sum_{t=0}^{T}\gamma^t R(s_t^{'}|s_t,a_t)\right] \tag{7}$$

The optimal policy π^* is to minimize $V_\pi(s)$ in (2). Denote $V^*(s)$ as the minimum achievable reward at state s, then we can find $V^*(s)$ at every state recursively by solving the following Bellman's optimal equations [19,20].

$$V^*(s) = \min_a \sum_{s^{'}} Q^*(s,a) = \min_a \mathbb{E}_{\pi^*}\left[\sum_{t=0}^{T}\gamma^t R(s_t^{'}|s_t)\right] = \min_a \mathbb{E}_{\pi^*}\left[R_t + \gamma\sum_{t=0}^{T}\gamma^t R(s_{t+1}^{'}|s_{t+1})\right] = \min_a \mathbb{E}\left[R_t + \gamma V_{t+1}^*(s)\right]$$

$$= \min_a \sum_{s^{'}} p(s'|s,a)[r+\gamma V^*(s')] = \min_a \sum_{s^{'}} P(s,a,s^{'})[R(s^{'}|s,a)+\gamma V^*(s')] \tag{8}$$

The Bellman optimality equation for $Q^*(s,a)$ is

$$Q^*(s,a) = \mathbb{E}\left[R_t + \gamma \min_a Q_{t+1}^*(s,a)\right]$$

$$= \sum_{s^{'}} p(s^{'}|s,a)\left[r + \gamma \min_{a^{'}} Q^*(s^{'},a^{'})\right] = \sum_{s^{'}} P(s^{'}|s,a)[R(s^{'}|s,a)+\gamma V^*(s')] \tag{9}$$

The above equations show backwards induction, which is a dynamic programming approach and efficient method for solving the Bellman's equations. The optimal action and the optimal value function can first be determined at the end of episode, then the optimal path can be obtained recursively from the last to the first period based on the Bellman optimal equations.

3.5. Analytic hierarchy process (AHP)

The analytic hierarchy process (AHP), a multi-criteria decision-making approach, arranges factors in a hierarchical structure [21] by using pairwise comparison to investigate the weights (i.e. importance) of all elements (factors or alternatives) from the decision maker. Pairwise comparison investigates the importance to the objective one pair a time with respect to all of the elements (factors and alternatives) in the same group. Given p elements, i.e., $A_1,\ldots,A_p$, the decision maker assigns relative importance a_{kl} representing the importance of A_k relative to the importance of A_l to extract the weights based on the decision maker's preference structure. Let the weights $w_1,\ldots,w_p$ corresponding to each element and $a_{kl} = w_k/w_l$. Formulate the matrix forms A (also called reciprocal matrix) of the pairwise ratios:

$$\mathbf{A} = \begin{bmatrix} w_1/w_1 & \cdots & w_1/w_p \\ \vdots & \ddots & \vdots \\ w_p/w_1 & \cdots & w_p/w_p \end{bmatrix} \tag{10}$$

Calculate the eigenvalues $\lambda_1,\cdots,\lambda_p$ of matrix A to obtain the eigenvectors (i.e. weights) of each element in the hierarchical structure. Finally, to validate the weight consistency of the decision maker, conduct a consistency test which calculates the two indices: the consistency index (CI) and the random consistency index (RI) [22]. The CI is defined as $\mathrm{CI} = (\lambda_{max}-p)/(p-1)$, where λ_{max} is the maximal eigenvalue of matrix A. The RI is the mean value of CI obtained from matrix A with the random number $\frac{1}{9},\frac{1}{8},\cdots,1,\cdots,8,9$ generated by the simulation with sample size 500 matrices according to different p. The definition of the consistency ratio (CR) for the typical acceptance criterion is $\mathrm{CR} = \mathrm{CI}/\mathrm{RI} \le 0.1$. Thus, AHP extracts the weight of each variable through a systematic approach. Since the performance of RL significantly depends on reward values, this study suggests AHP to assess the rewards; particularly, from three aspects of price, inventory, and procurement action.

4. Data science framework for BD procurement decision

This section proposes a data science framework for BD procurement decision in Fig. 4, which includes price forecasting (1st stage) and procurement decision (2nd stage). The main analysis processes are data preprocessing, FS, data integration, predictive model, decision model, and model evaluation.

The data collection includes the historical price, supply & demand, upstream/downstream material price and capacity, and other industrial data. We identify an appropriate time period for training prediction model since the historical data of BD price showed different patterns; in particular, before and after economics crisis in 2008. Thus, this study suggests using data after 2008. After data clean and merge, the dataset is almost complete with a few missing values related to contract price or capacity data, and we impute them by domain experts or using K-nearest neighbor (k-NN) [23].

FS can select the important variables which significantly affect the BD price. It can improve the prediction accuracy, avoid the overfitting, and reduce the computational running time. This study suggests using one nonlinear model (i.e., RF) and one linear model (i.e., Lasso), respectively, since there is no idea about the true geometric causal relationship between predictors and the response variable in the collected dataset [24]. RF can consider the nonlinear interaction effects among variables and form tree structure for illustration. The Lasso is a shrinkage method which effectively addresses the high dimension issue (i.e. $p > n$) which may occur in our case study, where p is the number of variables and n is the number of observations, and the issue undermines the prediction accuracy [11]. Finally, the cross validation (CV) and voting mechanism combining two selection techniques are applied to enhance the robustness of FS techniques [24].

After FS, we transform the time series data into a new format for supervised learning, and the sliding window (SW) and data normalization are applied. Fig. 5 illustrates an example. Given the window size is 4, we use data in the past two periods to forecast the price in the future two periods. We slide the raw dataset and reorganize them into one row, and also merge the other two predictors (X) into the new format. For normalization, the z-score is used to re-scale the values.

For the prediction models, we suggest three methods— time series decomposition (TSD), deep neural network (DNN), and LSTM since they are applicable to time series and nonlinear prediction. For TSD, we suggest "divide-and-conquer" strategy and decompose the raw time series (complex) data into several independent and additive (simple) components, such as trend, seasonality, and randomness. We model each component respectively, and then sum each estimated component to approximate the original pattern. Generally, modelling these simple components is easy and helpful to clarify the causal relationship and improve the interpretability [25]. In TSD, this study uses SVM for modelling the trend pattern, remains seasonality effect, and suggests

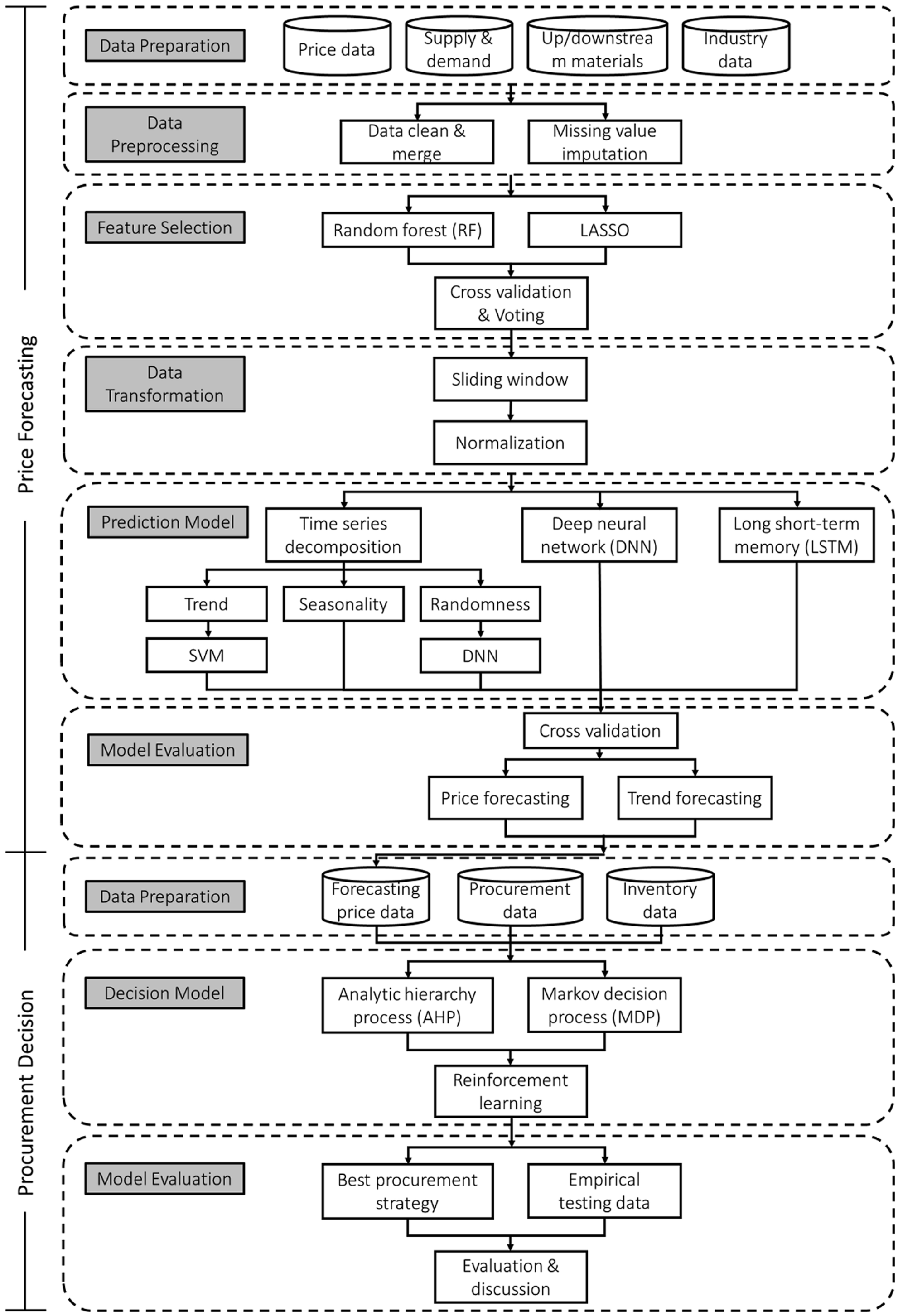

Fig. 4. Two-stage data science framework for procurement decision.

DNN for modelling the randomness. In addition, we also suggest DNN and LSTM for price prediction, respectively. Finally, we provide the performance evaluation of three methods.

In the procurement decision stage, we collect the forecasting price from previous stage, the historical procurement decision (including time and quantity), the historical inventory data, etc. In order to build the reward function used in MDP, the AHP is applied to form a hierarchical structure of each pair (state, action) and extract the preference (i.e. reward) from the decision maker by pairwise comparison through the questionnaire. In this study, the state is a two-dimensional (2D) composite state $s \in S = \{(p, i)\}$ described by the forecasting price (p) and inventory level (i), and the action $a \in A$ is the procurement decision. For MDP, we discretize the state space and action space, and then the reward can be described explicitly by AHP. Finally, the MDP is used to derive the optimal policy by learning from the observed samples of outcomes in the environment. The objective of MDP is to maximize the expected reward. That is, the proposed data science framework aims to minimize the expected total procurement cost of BD.

Proposition 1. *Assume state space S is finite and countable, and action space A is finite for each $s \in S$. Based on the backward value iteration algorithm defined by the Bellman's equation (8), there exists a deterministic*

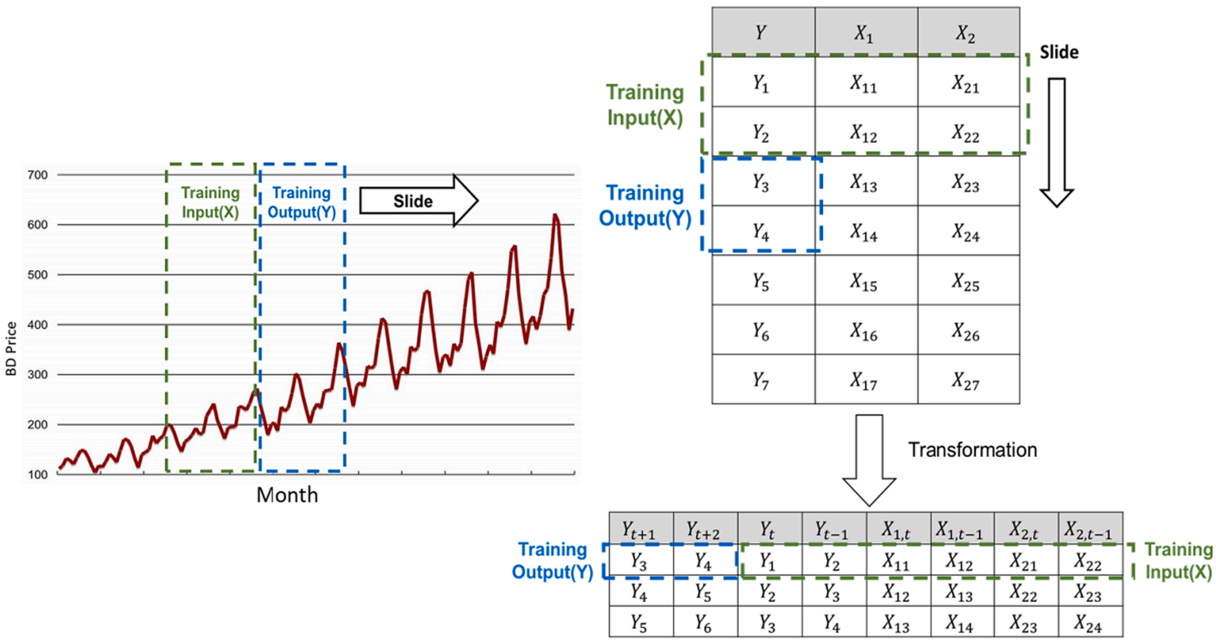

Fig. 5. Data transformation by sliding window.

optimal policy of the procurement decision that minimizes the expected total cost over the planning horizon.

Proof. According to the MDP model defined above, the set of all possible states and the set of all possible actions are both finite and countable. Therefore, based on the Proposition 4.4.3 of [16], there exists a deterministic Markovian policy that is optimal. ■

An empirical study is conducted to verify the effectiveness and robustness of the proposed two-stage data science framework in the next section.

5. Empirical study

This section conducts an empirical study of a leading petrochemical manufacturer in Taiwan to validate the two-stage proposed data science framework. Without losing generality, all data is transformed for proprietary information protection.

5.1. Price forecasting

The data is mainly collected from different sources in the case company, market surveys, and industrial reports in 2009–2018. After

Crude Oil

	Variable	r.c.	c.c.
	X01	17.9591	0.5195
★	X02	18.3465	0.5229

Upstream

	Variable	r.c.	c.c.
★	X03	2.2597	0.8511
★	X04	0.8613	0.1990
	X05	1.6192	0.5878
	X06	-0.3626	-0.1536
	X07	2.0611	0.5830
★	X08	-0.4406	-0.4160
	X09	0.0002	0.4336

Butadiene

	Variable	r.c.	c.c.
	Y		
	X10	0.9936	0.9997
	X11	1.0051	0.9997
	X12	0.8651	0.8741
	X13	0.9153	0.8584
★	X14	0.8382	0.9401
	X15	0.8153	0.9406
	X16	0.8615	0.9390
	X17	0.8751	0.8425
	X18	0.8802	0.9512
	X19	0.2148	0.0719
	X20	0.3655	0.1411
	X21	0.0012	0.4902

Downstream

	Variable	r.c.	c.c.
★	X22	5.6786	0.7958
	X23	0.7643	0.9258
	X24	0.8698	0.9424
	X25	0.9213	0.7151
	X26	0.8003	0.9272
	X27	0.8897	0.9313
	X28	0.0008	0.5036
	X29	0.0005	0.4319
★	X30	1.0112	0.9810
	X31	1.3185	0.9753
	X32	0.1570	0.1345
★	X33	0.7595	0.3052
	X34	1.4520	0.6067

Legend	
(pink)	Butadiene(BD)
(yellow)	r.c. > 1.5
(blue)	r.c. < 0
★	Important variables identified by feature selection

Fig. 6. Upstream and downstream variables affecting BD price.

business understanding, we collect 35 variables including one response variable (i.e. weekly BD price) with 459 samples. To clarify the interaction effects and causal relation among multiple variables, we need to understand the industrial domain knowledge and the upstream/downstream structure of the BD raw material. Fig. 6 illustrates the relationship among all variables, and shows the regression coefficient (r.c.) and correlation coefficient (c.c.) of each variable (X01-X34) with respect to one response variable (Y).

Fig. 6 is also helpful to connect the domain knowledge and improves the interpretability of the BD price forecasting. For example, lots of yellow marks (r.c. $>$ 1.5) or blue marks (r.c. $<$ 0) are presented in the upstream and it implies that the price and supply of upstream materials significantly affect the BD price. In particular, two blue marks (X06 and X08) showing the negative effects of BD price represent the marginal profit of naphtha cracking. The increase of marginal profit increases the supply of upstream materials related to BD, and thus the increasing demand of BD leads to the decease of BD price. In addition, one yellow mark (X22) in the downstream presents the price of natural rubber, and it is a substitute of the synthetic rubber which is one of main products produced by the BD. Thus, when price of natural rubber increases, the demand of synthetic rubber increases and then the price of BD increases.

In data preprocessing, we need to shift the values in predictors one or several periods corresponding to the periods shown in the response variable because the "lag" effect causal relation between upstream and downstream. After data preprocessing (i.e. data merge and missing value imputation), the FS is applied; in particular, RF and Lasso. In RF, we calculate the mean decrease Gini (MDG) to evaluate the importance of the variables, and use the out-of-bag (OOB) error to decide the best number of selected variables. In Lasso, we calculate mean squared error (MSE) and draw the shrinkage of coefficients diagram shown as Fig. 7 to decide the best penalty parameter $\lambda = 0.653$ for variable selection. 10-fold cross validation is used in this study for the robustness of FS. Finally, Table 1 shows the 8 selected variables by the voting. We also calculate the variance inflating factor (VIF) which provides a metric that measures the variance of an estimated regression coefficient, and the variance is increased if the multicollinearity occurs among these selected variables. The result shows that all VIFs are less than 3.616 which are smaller than 10 (a typical threshold used in literature), and it implies that these 8 selected variables are not highly correlated.

This study forecasts the BD prices of the future 13 weeks due to reserving the variability of the lead time (including the shipping time) for procurement operations. For price forecasting, we apply TSD, DNN, and LSTM (including GRU). In TSD method, the "divide-and-conquer" strategy is used to decompose the original pattern into three sub-patterns as Fig. 8. SVM is used for modelling the "trend", and DNN is used for the "randomness", respectively. For seasonality component in Fig. 8, we can see that BD price gradually increases until the year end due to the Christmas and New Year festival, and also downstream

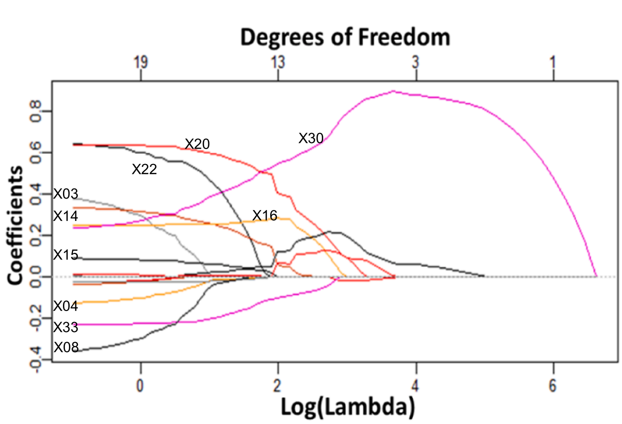

Fig. 7. Lasso shrinkage of coefficients.

procurement in the year end increases the BD price because the rubber production in Southeast Asia interrupts by northeastern monsoon from January to April. The seasonal factor is an adjustment and can be calculated by simple statistics. Finally, sum of three components presents the original BD price prediction. The TSD not only enhances the prediction accuracy but also improves the interpretability.

Besides to TSD, we also suggest DNN and LSTM for BD price forecasting as mentioned in Fig. 4. In DNN, we set MSE as loss function and Adam algorithm as optimizer. The DNN structure (eg. the number of neurons in each layer) and its parameters (eg. dropout rate) are identified by using grid search. LSTM reorganizes the dataset with 435 samples, 8 time steps (i.e. 8 week), and 72 features equal to 9 variables (1 response variable plus 8 selected predictors) multiply 8 time steps. LSTM uses a similar setting as DNN to identify the best configuration. We forecast the future 13 weeks of BD price. The 10-fold cross validation is applied to evaluate the model robustness; in particular, with two indices for prediction evaluation: (1) mean absolute percentage error (MAPE); and (2) trend accuracy, where both of them are easy-understanding and improve management's interpretation. To be specific, trend accuracy categorizes the actual prices from period t to t + 1 into three categories: up, down, and flat, which imply the "trend". The trend accuracy is to calculate the "hit rate" which the trend of predictive prices matches the trend of actual prices.

Tables 2 and 3 show the MAPE and trend accuracy of BD price forecasting among different models, respectively. The period of the testing dataset is from March 2017 to March 2018. Based on the request from the company, we provide the 3th-week forecast, 8th-week forecast, and 13th-week forecast. The current method (i.e. As-Is) in the company is mainly based on human experience with partial information collected information semi-automatically. On one side, the IT division collects relevant data including the historical receiving orders, the marketing report, the third-party report, etc., and builds the dashboard to summarize and visualize the information. On the other side, the procurement division manually collects the unstructured data including upstream and downstream news, world-wide news, quotation, regional conflict, political issue, etc., and also refers to IT dashboard to make a price forecast. This study compares TSD, DNN, and LSTM. The result shows that the proposed three models are generally better than current method, and the LSTM provides the best performance. Fig. 9 shows the BD price forecast by LSTM; in particular, the right-hand-side is the zoom-in diagram of the testing dataset. The "furry" diagram shows price forecasting for the future 13 weeks and each red line shows the time-rolling mechanism. Thus, this study suggests using LSTM for BD price forecasting.

5.2. Procurement decision

Based on the BD price forecasting, this section builds a policy for procurement decision through the RL technique. First, we collect the monthly procurement data related to the unit price and amount purchased from January 2015 to February 2018. The procurement decision is mainly affected by the BD consumption, inventory, safety stock, and future price. The lead time is one month due to the overseas shipping. In addition, the month-start and month-end inventory data are collected and thus we can derive the monthly BD consumption. To obtain the optimal policy, this study used Python 3.6 and applied the OpenAI Gym platform on PC with Intel® Core™ i7-4770 CPU @ 3.40 GHz processor and 8.0 GB of memory.

To develop RL technique, we define State (S) as a composite state with price state (PS) and inventory state (IS), i.e. a vector $S = (PS, IS)$. The $PS = \{1, 2, 3\}$ indicates {down, flat, up} regarding the future trend of BD price by comparing the actual price in one specific month and the forecast price in next month. The $IS = \{1, 2, 3, 4\}$ indicates {low, median, high, very high} regarding the inventory level including safety stock. We define the Action $A = \{1, 2, 3, 4\}$ indicates {do nothing, low, median, high} regarding the amount purchased restricted by the size of

Table 1
FS by voting through RF and Lasso.

Var. (X)	02	22	14	08	03	33	30	04	05	21	23	19	12	06	28	18	31	24	15	26	17	27
RF	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Lasso	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Voting	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1

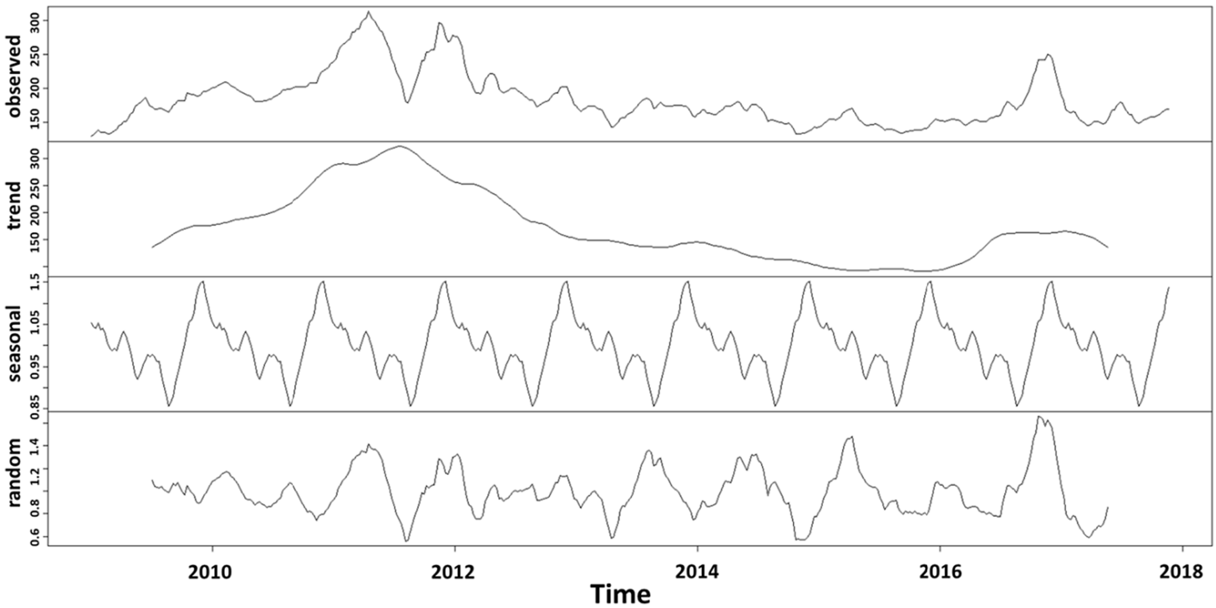

Fig. 8. Time series decomposition.

Table 2
A comparison of MAPE.

MAPE	As-Is	To-Be (data science model)		
Foreseen period	Current method	TSD	DNN	LSTM
Week 3	0.126	0.128	0.115	0.097
Week 8	0.246	0.184	0.167	0.158
Week 13	0.272	0.237	0.216	0.191

* MAPE as smaller as better.

Table 3
A comparison of trend accuracy.

Trend Accuracy	As-Is	To-Be (data science model)		
Foreseen period	Current method	TSD	DNN	LSTM
Week 3	0.594	0.723	0.702	0.726
Week 8	0.529	0.717	0.744	0.754
Week 13	0.588	0.683	0.737	0.773

* Trend accuracy as larger as better

cargo ship. Thus, the procurement decision can be investigated, that is, we can observe a "path" from the current state s to next state s' by taking some action. Based on these paths in the historical dataset of procurement and inventory, the Transition Probability $T(s, a, s')$ can be calculated.

For the Reward Function $R(s^{'}|s, a)$, we suggest the reward value by using AHP to extract the "weights" from the decision maker. We design the questionnaire and deliver to the procurement staffs and management. Fig. 10 illustrates an example of a partial hierarchy regarding scenario $PS = 2$ (similar hierarchy is used for the other two scenarios $PS = 1$ and $PS = 3$). A question example is a statement like "given the price state is flat ($PS = 2$) and inventory state is low ($IS = 1$), what is your preference when you compare A = 1 with A = 2?" The pairwise comparison in this case ($PS = 2, IS = 1$) also includes the comparison between A = 1 and A = 3, between A = 1 and A = 4, between A = 2 and A = 3, between A = 2 and A = 4, between A = 3 and A = 4. We formulate a score matrix as Eq. (10) and calculate the eigenvectors to obtain the weights 0.0288, 0.0220, 0.0334, 0.0425 with respect to A = 1, 2, 3, 4, respectively. Through a pairwise comparison and consistency test, AHP provides the weights as reward values in Table 4. Table 4 shows that, given a lower IS, the weight increases over the amount purchased to avoid the risk of out-of-stock. In fact, the company tends to purchase BD when the stock space is available. However, in a higher IS, purchasing too much is not allowed to avoid the risk of excess inventory; in particular, the weight of IS = 3 and A = 4 drops to 0.0767 due to a limited stock space. Thus, the procurement decision can be evaluated, that is, we obtain a reward through a "transition" from the current state to next state by taking some action.

After explicitly defining the state, action, and reward, the MDP can be applied to derive the "policy". This policy provides a mapping from state to action in order to maximize the expected reward in the long run. If all the relations among state, action, and reward are well-defined, then the MDP is suggested; otherwise Q-learning algorithm is suggested for partial information or complicated environment. Given the learning rate $\alpha = 0.9$ and the discounted factor $\gamma = 0.8$, Fig. 11 illustrates that the maximal reward increases and converges quickly at a high reward value. This result implies that the agent has learned to maximize its total reward earned in an episode by behaving optimally at every state, and the optimal policy is formulated as Fig. 12. First, there are two blank states (PS = 2, IS = 2) and (PS = 2, IS = 4) in Fig. 12 because there is no record found about these two states from the historical dataset, and thus the Q-value (i.e. $Q(s, a)$) cannot be obtained. Second, given PS = 3 (i.e. the forecast price goes up), the amount purchased increases along with the decrease of inventory state. Third, given PS = 1 (i.e. the forecast price goes down), the low amount purchased is fixed and independent of

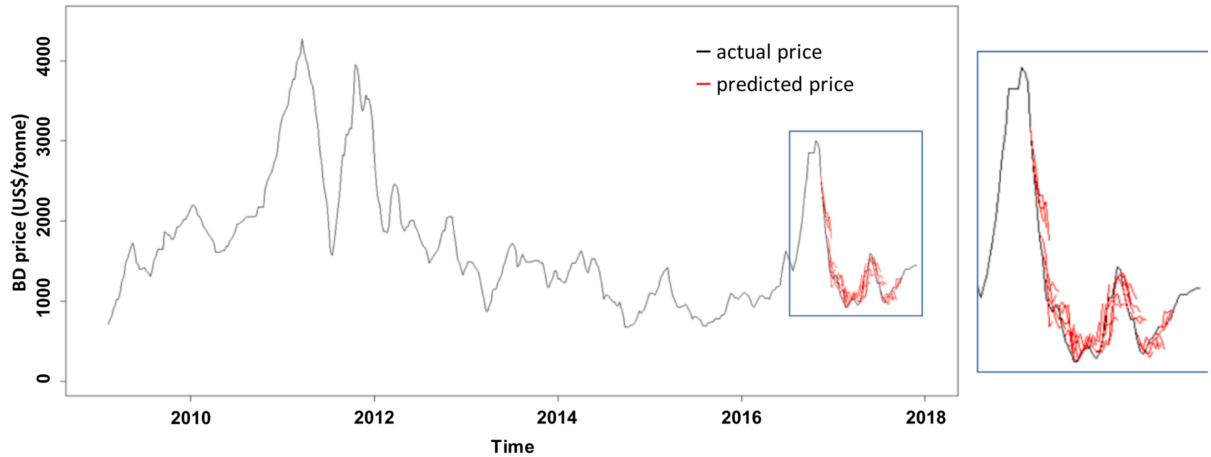

Fig. 9. 13th-week BD price forecast by LSTM.

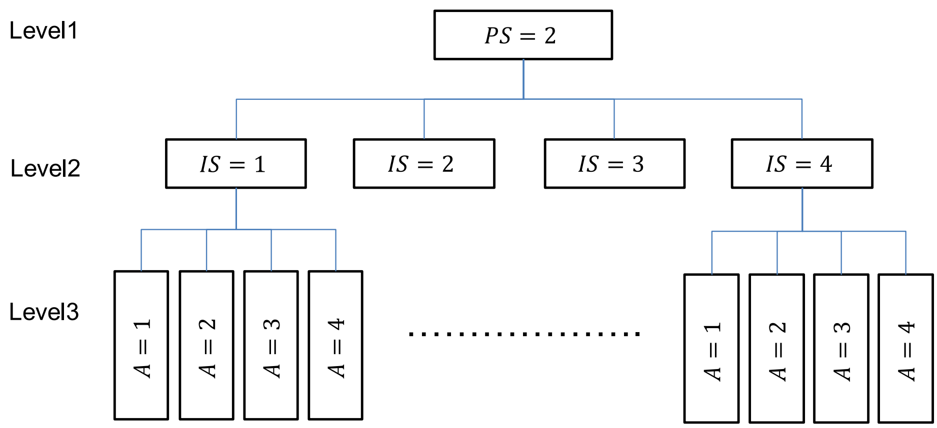

Fig. 10. AHP hierarchy regarding scenario.$PS = 2$

Table 4
Weights extracted by AHP regarding $PS = 2$.

PS = 2	Weights (the decision maker's preference)
IS = 1 and A = 1	0.0288
IS = 1 and A = 2	0.0220
IS = 1 and A = 3	0.0334
IS = 1 and A = 4	0.0425
IS = 2 and A = 1	0.0356
IS = 2 and A = 2	0.0355
IS = 2 and A = 3	0.0447
IS = 2 and A = 4	0.0703
IS = 3 and A = 1	0.0579
IS = 3 and A = 2	0.0800
IS = 3 and A = 3	0.1174
IS = 3 and A = 4	0.0767
IS = 4 and A = 1	0.2482
IS = 4 and A = 2	0.0843
IS = 4 and A = 3	0.0136
IS = 4 and A = 4	0.0091

the inventory state. This conservative action avoids "buy high and chase low", reduces the variability of amount purchased, and ensures an expected reward maximization in the long run. Finally, in fact, Table 4 and Fig. 12 also illustrate that the practical knowledge and experience of the procurement staff can be partially extracted and stored for knowledge management.

5.3. Evaluation

This study compares the current policy of the company employs, the order-up-to-level (s,S) procurement policy which orders the quantity to the target level S when the inventory level decreases to the reorder point s, and the optimal policy derived from RL to validate the proposed two-stage data science framework. To be specific, the company's current procurement process mainly depends on the on-hand inventory. According to the forecasted demand, if the inventory level is below the safety stock, the current policy prefers to purchase raw materials because of an unacceptable shortage cost including capacity loss and reputation damage. While if the inventory is sufficient, the current policy refers to the market quotation and the human-judged price prediction (based on the third-party report) to make procurement quantity. The company purchases at least once a month on average. This study calculates the total procurement cost and average inventory in 2017 for the evaluation of the three methods. Though the 13-week ahead price forecasting is developed, we consider 1-month ahead forecast for the cost evaluation due to the procurement lead time (i.e. overseas shipping) equal to 1-month. The total procurement cost is the sum of the monthly procurement cost which is calculated by multiplying the monthly procurement quantity by the actual unit price. Fig. 13, Fig. 14 and Fig. 15 show the procurement decision by current policy, (s,S) policy, and RL optimal policy, respectively. The gray bar chart indicates month-start inventory level, the blue curve shows actual price, the orange curve shows forecast price, and action {A1, A2, A3, A4} indicates the amount purchased from low to high. Note that the price and inventory are rescale without changing the pattern, and the scale is removed due to

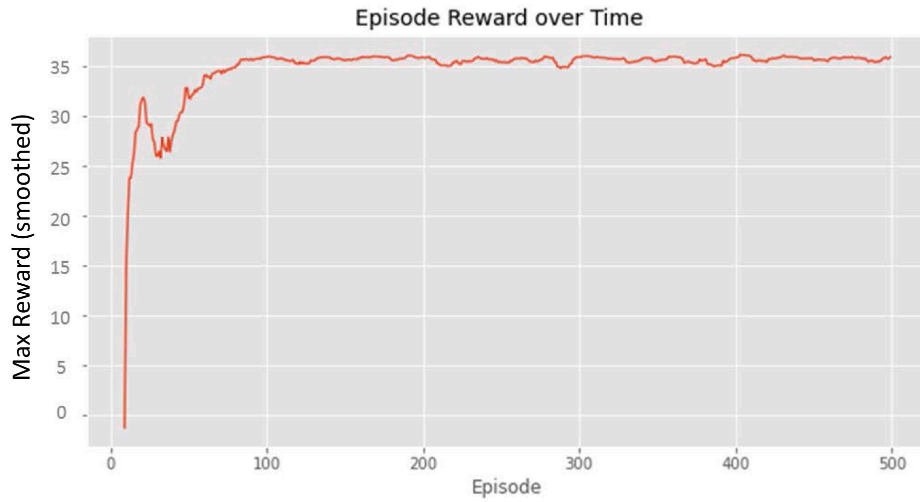

Fig. 11. Reward Convergence in RL.

IS=4	A=2		A=1
IS=3	A=2	A=1	A=2
IS=2	A=2		A=3
IS=1	A=2	A=2	A=4
	PS=1	PS=2	PS=3

Fig. 12. Optimal procurement policy by RL.

the confidential issue.

Fig. 13 illustrates that the company tends to purchase large amount when the actual price is relatively low. In spite of a higher price, the company must purchase a proper amount when the inventory is low. The results are typically generated from human experience, and thus the procurement decision (i.e. action) may show a larger fluctuation (eg. we can see 5 times for action A1 and 3 times for action A4). Fig. 14 shows that when the inventory level is below the reorder point, then the order-up-to-level (s,S) is initiated. This (s,S) policy only considers the inventory level without the forecasted information related to demand or price. The result shows that the procurement quantity of each purchase is very close (eg. 7 times for action A2 and 5 times for action A3), and the purchasing behavior could occur even at a moment that the price is higher. However, in Fig. 15, the RL optimal policy is derived based on the price forecasting, and thus the procurement decision is more robust (we can see there is no action A4 adopted and RL generally suggests action A2 or A3). To validate the optimal policy, there is an example. The actual price in March 2017 was rapidly falling, but the case company may purchase the highest amount (action A4) because the inventory is not enough. If the price forecast is available, it might suggest to postpone a larger amount purchased since the future price decreases; that is, the RL optimal policy suggests action A2.

Table 5 describes the performance evaluation of current policy, (s,S) policy, and optimal policy in 2017. The result shows around 12.3% reduction of the annual total cost when comparing RL optimal policy with current policy. In particular, the amount purchased by optimal

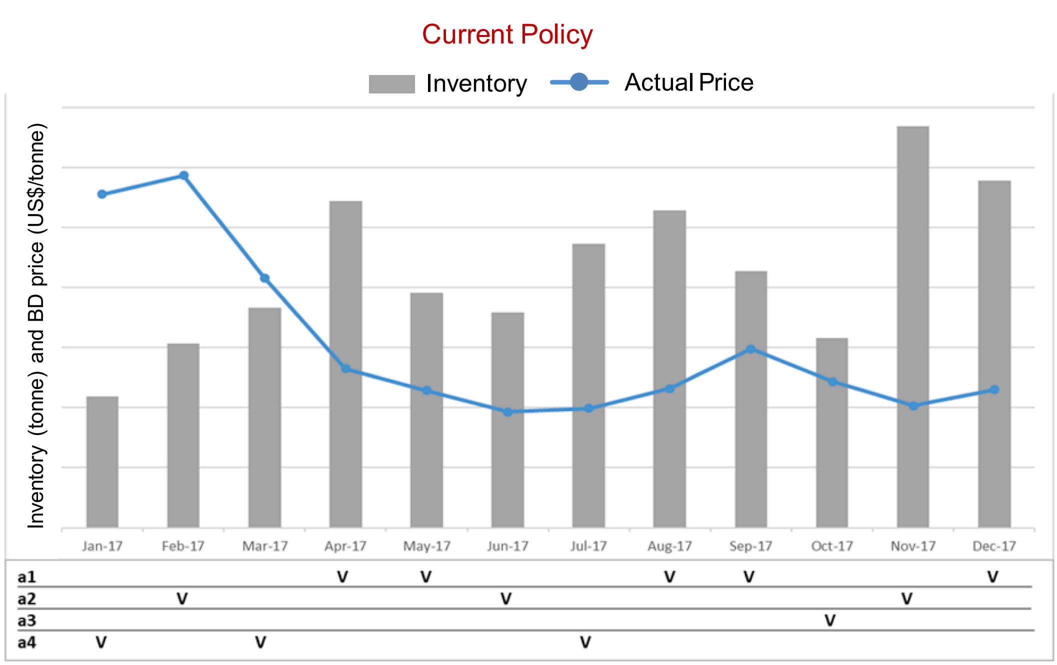

Fig. 13. Procurement decision by current policy.

Fig. 14. Procurement decision by order-up-to-level (s,S) policy.

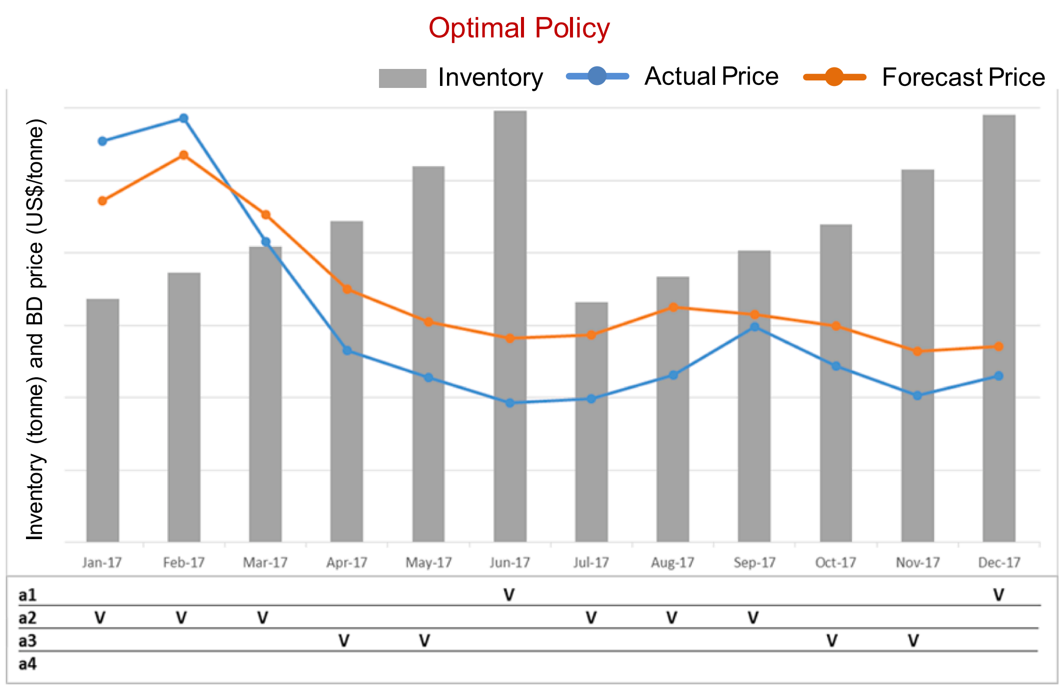

Fig. 15. Procurement decision by RL optimal policy.

Table 5
Performance evaluation of three policies in 2017.

	Current policy	(s,S) policy	Optimal policy
Average inventory (tonne)	3112.921	1812.788	3197.935
Standard deviation of inventory (tonne)	743.6546	302.1546	489.6778
Amount purchased (tonne)	25,301	35,430	36,835
Total cost (US$)	44596113.4	42324694.5	39091618.5

policy is higher but with lower cost since it considers the relationship between inventory level and forecasted price. In addition, the optimal policy also improves the standard deviation of inventory level (the smallest one is (s,S) policy) because a robust procurement decision is illustrated in Fig. 15 and more A2 actions are presented in the bottom-left triangle of Fig. 12.

6. Conclusion

For the petrochemical industry, it is capital-intensive and the procurement cost of raw materials accounts for about 70% to 80% of the production cost structure of downstream products. If the material cost can be reduced, it can significantly improve the profitability. This study proposes a two-stage data science framework to build the linkage between price prediction and procurement decision of the raw materials, and an optimal policy which minimizes the procurement cost can be derived by RL. The result shows that the price prediction performance of this study is better than the current method about 8.1% and the optimal procurement decision can reduce the total procurement cost about 12.4%. This validates the robustness and effectiveness of the proposed data science framework.

This study illustrated a successful digital transformation study related to petrochemical procurement decision. By using the digital data-driven technology and the AI-based technology for BD price forecasting and procurement decision optimization, the total cost reduction of BD procurement is significant and the practical knowledge of the procurement policy can be optimally formulated by RL as Table 4 and Fig. 12. The proposed data science framework provides a comprehensive analysis integrating factors related to BD price and procurement, and thus plays a critical digital transformation enabler to support the business process automation and enhance core competence [2,5,17].

For the future research, we suggest the following items for potential improvement. First, the multiple objectives including the total procurement cost, due date, on-time delivery, variability of inventory level could be considered in RL. Second, the slope information of price ups and downs should be counted for improving the amount purchased. Finally, the price discount of raw materials could be considered, that is, the discount is larger when the amount purchased is larger.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by LCY Chemical Corporation and Ministry of Science and Technology (MOST106-2218-E-031-001), Taiwan.

References

[1] D. Schallmo, C.A. Williams, L. Boardman, Digital transformation of business models—best practice, enablers, and roadmap, Int. J. Innovat. Manage. 21 (8) (2017) 1740014.
[2] C.-H. Lee, C.-H. Chen, A.J.C. Trappey, A structural service innovation approach for designing smart product service systems: Case study of smart beauty service, Adv. Eng. Inf. 40 (2019) 154–167.
[3] S. Abosedra, H. Baghestani, On the predictive accuracy of crude oil futures prices, Energy Policy 32 (12) (2004) 1389–1393.
[4] W. Xie, L. Yu, S. Xu, S. Wang, A New Method for Crude Oil Price Forecasting Based on Support Vector Machines, in: v.A.G.D. Alexandrov V.N., P.M.A. Sloot, J. Dongarra (Eds.), Computational Science – ICCS 2006. ICCS 2006. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2006, pp. 444–451.
[5] Y.-H. Wang, C.-H. Lee, A.J.C. Trappey, Modularized design-oriented systematic inventive thinking approach supporting collaborative service innovations, Adv. Eng. Inf. 33 (2017) 300–313.
[6] J.-L. Zhang, Y.-J. Zhang, L.u. Zhang, A novel hybrid method for crude oil price forecasting, Energy Econ. 49 (2015) 649–659.
[7] Y. Zhao, J. Li, L. Yu, A deep learning ensemble approach for crude oil price forecasting, Energy Econ. 66 (2017) 9–16.
[8] L. Breiman, Random Forests, Mach. Learn. 45 (1) (2001) 5–32.
[9] A. Liaw, M. Wiener, Classification and regression by randomforest, R News 2 (3) (2002) 18–22.
[10] R. Tibshirani, Regression shrinkage and selection via the lasso, J. Roy. Stat. Soc. B 58 (1) (1996) 267–288.
[11] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2 ed.:, Springer, 2008.
[12] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, Neural Comput. 9 (8) (1997) 1735–1780.
[13] A. Graves, Supervised Sequence Labelling with Recurrent Neural Networks, Springer, 2012.
[14] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, *arXiv:1406.1078*, 2014.
[15] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, *arXiv:1412.3555*, 2014.
[16] M.L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, second ed., John Wiley & Sons Inc, Hoboken, New Jersey, 2005.
[17] W.J.C. Verhagen, P. Bermell-Garcia, R.E.C. van Dijk, R. Curran, A critical review of Knowledge Based Engineering: An identification of research challenges, Adv. Eng. Inf. 26 (1) (2012) 5–15.
[18] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, 2 ed.:, The MIT Press, 2018.
[19] R.E. Bellman, *Dynamic Programming*, Princeton, Princeton University Press, NJ, 1957.
[20] C.-F. Chien, C.-H. Wu, Y.-S. Chiang, Coordinated capacity migration and expansion planning for semiconductor manufacturing under demand uncertainties, Int. J. Prod. Econ. 135 (2) (2012) 860–869.
[21] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.
[22] W. Ho, Integrated analytic hierarchy process and its applications– A literature review, Eur. J. Oper. Res. 186 (1) (2008) 211–228.
[23] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27.
[24] C.-Y. Lee, T.-L. Tsai, Data science framework for variable selection, metrology prediction, and process control in TFT-LCD manufacturing, Robot. Comput. Integr. Manuf. 55 (2019) 76–87.
[25] C.-Y. Lee, C.-F. Chien, Pitfalls and protocols of data science in manufacturing practice, J. Intell. Manuf. (2020).
[26] W.B.E. Al-Othman, H.M.S. Lababidi, I.M. Alatiqi, K. Al-Shayji, Supply chain optimization of petroleum organization under uncertainty in market demands and prices, Eur. J. Oper. Res. 189 (3) (2008) 822–840.
[27] J. Arnold, S. Minner, B. Eidam, Raw material procurement with fluctuating prices, Int. J. Prod. Econ. 121 (2) (2009) 353–364.
[28] B.A. Calfa, I.E. Grossmann, Optimal procurement contract selection with price optimization under uncertainty for process networks, Comput. Chem. Eng. 82 (2015) 330–343.
[29] R. Chen, T. Deng, S. Huang, R. Qin, Optimal crude oil procurement under fluctuating price in an oil refinery, Eur. J. Oper. Res. 245 (2) (2015) 438–445.
[30] G. Vial, Understanding digital transformation: A review and a research agenda, J. Strateg. Inf. Syst. 28 (2) (2019) 118–144.