Energy Proceedings
Vol 40, 2024
ISSN 2004-2965

Hybrid Modeling Digital Twin for Natural Gas Station Systems of Long-Distance Pipelines

Weijia Li[1], Jing Gong[1*], Jiuda Zhang[1], Haochen Pan[1], Weihe Huang[2]

1 National Engineering Laboratory for Pipeline Safety/ MOE Key Laboratory of Petroleum Engineering/ Beijing Key Laboratory of Urban Oil and Gas Distribution Technology, China University of Petroleum-Beijing, Fuxue Road No. 18, Changping District, Beijing 102249, P.R. China

2 China National Petroleum Corporation/PetroChina Company Limited, Beijing 100007, P.R. China

(*Corresponding Author: ydgj@cup.edu.cn)

ABSTRACT

The construction of energy system digital twin relies on accurate models. This paper proposes a new modeling method for natural gas station hydraulic systems by integrating physical models and data-driven models to improve the accuracy of models. Taking a natural gas station of long-distance gas pipeline as an example, this paper builds the physical model of a station, including compressors and regulating valves. Then a hydraulic calculation algorithm of the station is developed. A data-driven model Back Propagation Neural Network (BPNN) is introduced for physical model error compensation. Finally, the calculation results show that the hybrid model has better accuracy than the physical model and the energy consumption of key equipment, such as compressors, air coolers in the station is monitored. Moreover, the hybrid model better integrates the advantages of the two types of models, it can serve as a soft sensor to enrich the status monitoring data of station equipment and lay the foundation for further optimization of station energy consumption.

Keywords: energy consumption monitor, natural gas transportation, intelligent model, digital twin

Nonmenclature

Abbreviations	
BPNN	Back Propagation Neural Network
DT	Digital Twin
Symbols	
H	polytropic head
Z	gas compressibility factor
R	specific gas constant
ε	pressure ratio
T_1	suction temperature
T_2	discharge temperature
P_1	suction pressure
P_2	discharge pressure
Q	flow rate
N	speed
C_v	flow coefficient

1. INTRODUCTION

The rise of Digital Twin(DT) technology is due to the rapid development of the cyber-physical systems. Applying DT in the design stage is not only optimizing the design processes but also be beneficial to upcoming events, such as manufacturing planning, product health monitoring.

Physical model is the foundation of digital twin. As for natural gas pipeline transmission, the pressure of the gas is reduced mainly due to friction with the wall of the pipe and heat transfer between the gas and the surroundings. Compressor stations are usually installed to boost the pressure of the gas. In the station, the gas is filtered and metered, and then part of gas is transported to the next station after pressurized through compressors, and the other part is distributed to city gate stations.

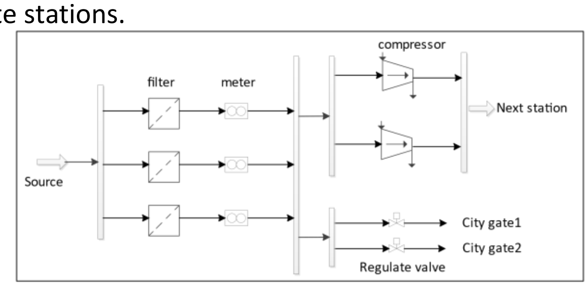

Fig. 1 Schematic diagram of the gas station

This is a paper for 15th International Conference on Applied Energy (ICAE2023), Dec. 3-7, 2023, Doha, Qatar.

Several investigators tried to simulate unsteady condition for pipeline and some of them focused in compressor station modeling.

Botros et al. [1] presented a dynamic compressor station simulation that consists of nonlinear partial differential equations describing the pipe flow together with nonlinear algebraic equations describing the quasi-steady flow through various valves, constrictions, and compressors. Compressor is the key equipment in stations, and optimized control can greatly reduce energy consumption. Bryant [3] simulates The Florida Gas Transmission (FGT) model, almost all possible operational scenarios within each compressor station. Station control has the ability to try and meet multiple setpoints; and, the ability for units to automatically come on-line and off-line. Lei Zhang [4] modeled the natural gas compressing system based on a combined approach that the flow in the pipe was simulated using a finite volume method with a high order upwind scheme considering the real-gas behaviors and the behaviors of elements.

Several works focus on compressor maps. Schultz [5] derived the real-gas equations of polytropic analysis and to show their application centrifugal compressor testing and design.

The dynamic behavior of the gas flow within a compressor station is described by nonlinear partial differential equations. However, the physical model does not take into account unmodeled factors such as noise, equipment aging, and dirt, resulting in a decrease in accuracy over time. In addition, the physical model is slow in computation and cannot meet the real-time requirements of online simulation. Recent advances in computational power and exponential growth of data availability have made data-driven modeling more viable and popular.

On the other hand, data-driven approaches such as machine learning techniques could effectively deal with the modeling challenges when it is hard to construct the mechanism-based model. Black-box data driven modeling can obviously improve the speed of modeling, decrease the computational burden, and contribute to constructing lightweight models to describe the relationship between physical variables. For example, Gaochen Cui[6] proposed a data-driven framework, and used neural network to estimate hyper parameter of pipe, which improves the accuracy of the calculated results effectively. Hamid Asgari[7] developed nonlinear autoregressive exogenous (NARX) models of a heavy-duty single-shaft gas turbine (GT) during start-up operation. Xie H [8]used NARX neural network to predict CO2 compressor vibration performance.

However, the construction of data-driven models highly correlates with the samples of known variables, which scarcely contributes to the understanding of the inner logic of model construction. Combining physics-based and data-driven modeling is a relatively new field of research. Hybrid models have developed in many industries, such as manufacturing, autonomous driving and electricity. However, there is few works in natural gas transmission.

2. METHOD

2.1 Digital twin framework

The digital twin is divided into four layers, based on the device layer, which collects device data through the Internet of Things and sends it to the data center. The second layer is the data layer, where the raw data is classified and stored in different data pools, and the data is integrated and transformed into a unified storage format that is easy to analyze for storage in each data pool. The third layer is the model and algorithm layer, this layer gets the pipe length, pipe material, and equipment parameters of the pipeline from the data layer to establish a model, and uses algorithm simulation to solve the model. The fourth layer is the application layer, which includes application services based on simulation models such as state monitoring, equipment fault diagnosis, and energy consumption optimization. The interface starts the service through calling the service interface, then displays the calculation results, including curves, charts, schemes, to the user.

2.2 Physical model

2.2.1 Quasi-steady models for elements

Compressor

Some parameters are very important for compressor performance, for example isentropic head, isentropic efficiency, rotational speed and power. The equation for head will be :

$$H = \frac{m}{m-1} Z R T_1 (\varepsilon^{\frac{m-1}{m}} - 1) \tag{1}$$

In the above equation, H is the polytropic head and Z is the gas compressibility factor. R is the specific gas constant. ε is pressure ratio of discharge side and suction side.

Using standard polynomial cure-fit procedures for each centrifugal compressor , the relationship between head, speed and flow rate could be found by:

$$H = a_2\left(\frac{n}{n_0}\right)^2 + b_2 Q\left(\frac{n}{n_0}\right) + c_2 Q^2 \quad (2)$$

The power consumption for the compressor driver is currently obtained by:

$$P_e = \frac{MH_{pol}}{\eta_{pol}} + \Delta P_m \quad (3)$$

The gas discharge temperature is obtained by:

$$T_2 = T_1 \varepsilon^{\frac{m-1}{m}} \quad (4)$$

Regulating Valve

The valve characteristics determine the mathematical model of the valve. The equations for the equal percentage valves are:

$$p_2 = \sqrt{{p_1}^2 - \frac{ZGT_1Q^2}{{C_v}^2{N_1}^2}} \quad (5)$$

$$C_v(X) = \frac{C_{vo}}{e^{3.488}-1}e^{3.488x} - \frac{C_{vo}}{e^{3.488}-1} \quad (6)$$

In the equation above, p_1 and p_2 is the pressure where the pipe flow enters and leaves the regulating valve; Z is the average compression factor of the regulating valve; T_1 is the temperature in front of the regulating valve; Q is the volumetric flow rate under standard conditions; C_v is the current flow coefficient of the valve; N_1 is the unit conversion coefficient.

2.2.2 Station system model

The topology of a gas network is defined by a directed graph including N nodes and E edges which represent, respectively, the joints and the pipelines[9].

$A_{ij} = (a_{ij})$ the incidence matrix

$$a_{ij} = \begin{cases} 1 & \text{if node i is the first node of edge j} \\ -1 & \text{if node i is the second node of edge j} \\ 0 & \text{otherwise} \end{cases}$$

where A is the node-branch incidence matrix; where $i = 1...N$ and $j = 1...E$

At each node i $\in$ {1, . . ., N}, flow is exchanged with the exterior of the network with mass flow rate $q_i(t)$ at time t. Of course, $q_i(t)$ can be null as it is for structural nodes. The sign convention is that q_i is positive if the gas is introduced into the network and negative otherwise. Then the mass conservation at the i-th node at time t means

$$\sum_{j=1}^{E} a_{ij}^{+} Q^j(0,t) - \sum_{j=1}^{E} a_{ij}^{-} Q^j(L^j,t) = q_i(t) \quad i = 1...N \quad (7)$$

$$AQ = \sum_{i=1}^{N} q_i \quad (8)$$

2.2.3 Numerical solution

The set of pipe flow, compressor, mass balance, and looping equations can be represented as[10]

$$\begin{cases} F_1(P_1, P_2, ..., P_N) = 0 \\ F_2(P_1, P_2, ..., P_N) = 0 \\ ... \\ F_N(P_1, P_2, ..., P_N) = 0 \end{cases} \quad (9)$$

The multivariable Newton–Raphson iterative procedure for Eq. (10) takes the form

$$P_{k+1} = P_k - \mathbf{J}^{-1}\mathbf{F}(P_k) \quad (10)$$

$$\mathbf{J} = \begin{pmatrix} \frac{\partial F_1}{\partial p_1} & \cdots & \frac{\partial F_1}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial p_1} & \cdots & \frac{\partial F_n}{\partial p_n} \end{pmatrix} \quad (11)$$

2.3 Hybrid model

2.3.1 Data-driven model

Back Propagation Neural Network(BPNN) is a common supervised machine learning algorithm that uses the BP algorithm and the labeled training samples to optimize the randomly initialized network weights to achieve prediction or classification. A three-layer BPNN includes the input layer, the hidden layer and the output layer[11].

In the BPNN, the N-dimensional vector $x = \{x_1, x_2, ..., x_N\}$ represents the input of the network, the L-dimensional vector $h = \{h_1, h_2, ..., h_L\}$ represents the feature representations of the hidden layer, and the affine mapping between the input layer and hidden layer can be established by a nonlinear transform, which can be expressed as follows

$$h_i = f(w_{ij}x_j + a_i) \quad (12)$$

Where W and a represent the weight matrix and bias vector of the network,

W dimension is $N \times L$, m-dimensional vector $y = \{y_1, y_2, ..., y_m\}$ represents output layer of network, respectively $f(\bullet)$ is the sigmoid activation function

$$f(x) = \frac{1}{1+e^{-x}} \quad (13)$$

Output layer is linear:

$$y_k = w_{ki}h_i + b_k \quad (14)$$

W and b represent the weight matrix and bias vector of the network, To optimize parameters, the mean squared error is used as the cost function to optimize the parameters.

$$E = e^T e$$
$$e = y - \hat{y} \qquad (15)$$

Furthermore, to alleviate the network overfitting problem, the L2 norm regularization is introduced, and the cost function is described as

$$[W^*, b^*] = \arg\min \frac{1}{m}\sum\left(\frac{1}{2}\left\|y^{(k)} - \hat{y}^{(k)}\right\|^2\right) + \lambda\|W\|_2^2 \qquad (16)$$

Forward propagation calculates the output of the neural network after it receives an input. Then a backward propagation calculates the gradients for all the parameters of every layer. The parameters are updated based on these gradients to reduce the loss function. In this paper, the neural network was trained with 3000 epochs.

$$w_{ij} := w_{ij} + \triangle w_{ij}$$
$$\triangle w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \qquad (17)$$

<u>2.3.2 Hybrid modeling</u>

In a hybrid model, the physical model is first used to calculate device parameters. However, due to model truncation errors, curve fitting errors, and other reasons, the physical model deviates from the actual measurement results. The error data is fed into the BPNN data model. As the compressor is the core equipment of the station, this article takes the compressor as an example to construct a hybrid model. Utilizing the conventional process parameters and measurement data from SCADA system of the production system to compensate for pressure error of the core equipment compressor in the station. The measurement data in this article is artificially made by adding bias from the model simulation values.

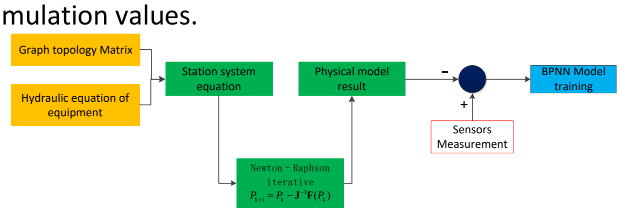

Fig. 2 Schematic diagram of the hybrid model

3. RESULTS AND DISCUSSIONS

3.1 Physical model

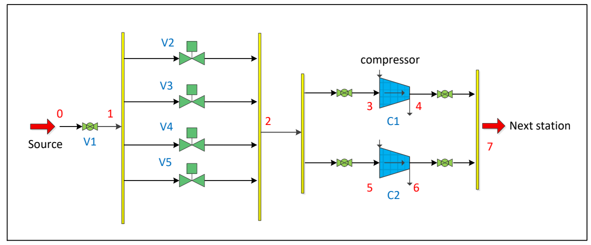

Fig. 3 station system

The topology of station system is in Figure 5. The system contains four parallel regulating valves as resistance elements, such as filters, and two parallel compressors. The source pressure is 92bar and the flow rate to the next station is 789kg/s. The results of physical model includes pressure of nodes(Table1) and flow rate of equipment(Table2).

Table1 Node pressure in station

Node number	Pressure(bar)
1	91.14
2	91.08
3	89.88
4	104.74

Table2 Equipment flow rate in station

Equipment number	Flow rate(kg/s)
V2	197.38
C1	394.79

3.2 Hybrid model

Utilizing the conventional process parameters and production data of the production system to compensate for pressure error of the core equipment compressor in the station. The measurement data in this article is produced by artificially adding a tiny bias from the model simulation result(Figure 4). Specifically, the measurement parameters include compressor inlet temperature T_1, inlet pressure p_1, compressor outlet temperature T_2, outlet pressure p_2, rotational speed N, and inlet flow Q. The output is error of measurement and physical model simulation.

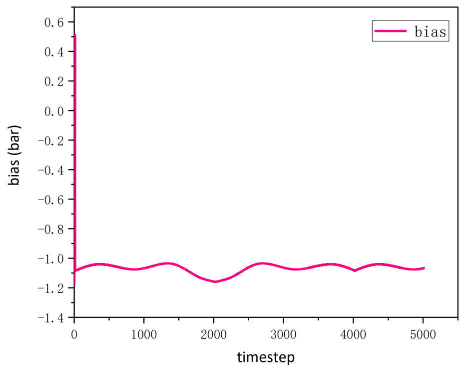

Fig. 4 artificially added bias

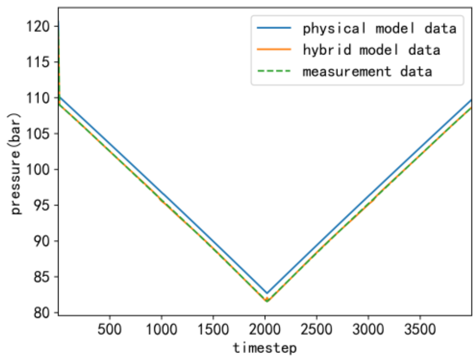

Fig. 5 discharge pressure(training data)

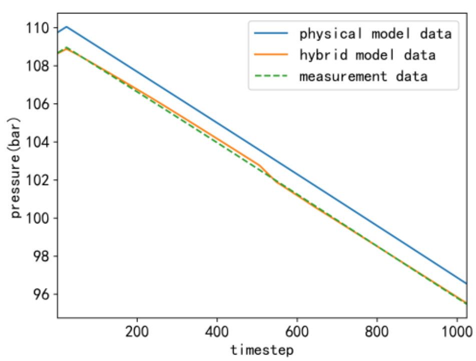

Fig.6 discharge pressure(test data)

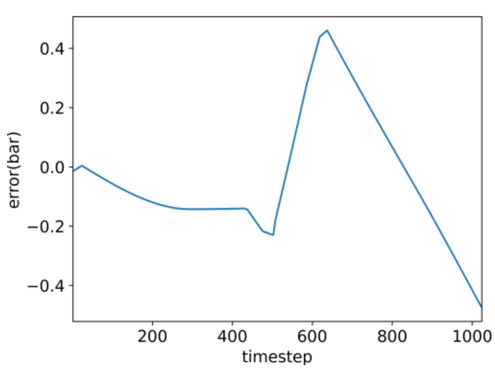

Fig. 7 discharge pressure error(test data)

Figure 5, Figure 6 are results of training set and test set. From Figure 5, Figure 6, it can be seen that the trained model is well validated on the test set. Discharge pressure simulated by pure physical model is compared with measurement and hybrid model. Apparently, the tiny bias is compensated by data-driven model in hybrid model. Figure 7 is the error between hybrid model and measurement. Figure 7 and Figure 5 shows that the error decreases about 50% after applying hybrid model.

4. CONCLUSIONS

In this paper, a novel hybrid modeling method is proposed for hydraulic simulation of stations in long-distance natural gas pipelines. More specifically, the physical model has the capability to simulate station system in quasi-steady state. The error between measurement and physical model are then passed into the BPNN to train it for predicting the unmodeled error. In this way, hybrid model adapts to new measurement information. The data-driven component can be updated online, which aims to further improve the model's generalization ability.

According to the results, the application of the proposed method to a station system demonstrates excellent simulation accuracy. Moreover, the desirable features and performance capabilities of the proposed method will enable scalability of AI methods for error prediction in real-time conditions. The hybrid approach combines the advantages of machine learning and analytic methodologies, which is expected to benefit controller design and simulation of optimization decision.

REFERENCE

[1] Botros K K .Dynamic simulation of compressor [2] Bryant M .Complex Compressor Station Modeling[J].Journal of Petroleum Technology, 1997.
[3] Kralik J , Stiegler P , Vostru Z ,et al.Modeling the dynamics of flow in gas pipelines[J].IEEE Transactions on Systems Man & Cybernetics, 1984, SMC-14(4):586-596.DOI:10.1109/tsmc.1984.6313330.
[4]Schultz, John M .The Polytropic Analysis of Centrifugal Compressors[J].Journal of Engineering for Gas Turbines & Power, 1962, 84(1):69.DOI:10.1115/1.3673381.
[5] Zhang L .Simulation of the transient flow in a natural gas compression system using a high-order upwind scheme considering the real-gas behaviors[J].Journal of Natural Gas Science & Engineering, 2016, 28:479-490.DOI:10.1016/j.jngse.2015.12.012.

[6] Xie H , Tang H , Liao Y H .Time series prediction based on NARX neural networks: An advanced approach[J].Proceedings of International Conference on Machine Learning & Cybernetics[2023-10-13].DOI:10.1109/ICMLC.2009.5212326.
[7] Asgari H , Chen X Q , Morini M ,et al.NARX models for simulation of the start-up operation of a single-shaft gas turbine[J].Applied Thermal Engineering, 2016.DOI:10.1016/j.applthermaleng.2015.09.074.
[8]Cui G , Jia Q S , Guan X ,et al.Data-driven computation of natural gas pipeline network hydraulics[J].Results in Control and Optimization, 2020.DOI:10.1016/j.rico.2020.100004.
[9]Alfredo Bermúdez a b, B M S .Modelling compressors, resistors and valves in finite element simulation of gas transmission networks - ScienceDirect[J].Applied Mathematical Modelling, 2021, 89:1316-1340.DOI:10.1016/j.apm.2020.08.013.
[10] Stoner M A .Steady-State Analysis of Gas Production, Transmission and Distribution Systems[J].Society of Petroleum Engineers, 1969.DOI:10.2118/2554-MS.
[11]Li J , Yao X , Wang X ,et al. Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis[J].Measurement, 2019, 153:107419.DOI:10.1016/j.measurement.2019.107419.