

# Research on Prediction Model for Drilling Construction Progress under Uncertain Conditions

Xitong Zhai

School of Xi'an Shiyou University, Xi'an 710065, China

## Abstract

**Drilling engineering is a complex and high-risk operation, and its construction progress is affected by various uncertain factors, such as the complexity of geological conditions, the variability of the working environment, and the complexity of construction methods and design parameters. In order to improve the accuracy of predicting the construction progress of drilling operations, this paper studies the drilling construction progress prediction model under uncertain conditions. By analyzing the influencing factors of drilling construction progress, including geological conditions, equipment performance, and personnel operation, the drilling construction progress prediction model is studied. Two models in the paper predict the drilling construction progress by simulating random events during the drilling process under different geological conditions. The results show that the HDBN model can better predict the drilling construction progress and provides strong support for the management and decision-making of drilling engineering.**

## Keywords

**Drilling Engineering; Construction Progress; Progress Prediction.**

## 1. Introduction

Drilling engineering is a complex and high-risk operation, whose construction progress is affected by the complexity of geological conditions, the variability of the working environment, as well as multiple uncertainties in construction methods and design parameters. To address this issue, studies have been conducted both domestically and internationally. The relevant achievements are as follows.

He Minjie and other scholars[1] innovatively integrated Internet of Things (IoT) and BIM4D technology to construct a dynamic control model for construction progress driven by fuzzy algorithms.

Xue Jianying and other scholars[2], in response to the systematic problems existing in the field of project progress control and dynamic cost management, innovatively constructed an intelligent monitoring system integrating BIM technology and earned value analysis method.

Wen Xin and other scholars[3] developed an analysis system based on wellbore engineering big data-driven, constructed a multi-source parameter coupling analysis model, and precisely compared design data with actual drilling information. The system achieved construction progress visualization through an intelligent analysis engine, covering layer matching degree of the formation, stability of the wellbore structure, adaptability of the drilling fluid density, and progress deviation warning. It also provided real-time analysis modules and dynamic drilling risk prompt functions, providing full-process digital control support for drilling engineering.

Wu Yulin and other scholars[4], based on the unique advantages of neural networks in modeling complex nonlinear relationships, innovatively integrated the local feature extraction ability of convolutional neural networks (CNN), proposing to use one-dimensional convolutional neural networks (1DCNN) to construct a drilling cycle prediction model.

Shuguo Zhang and other scholars[5], in response to the problems of information lag and response delay in traditional progress control modes, innovatively integrated BIM-4D technology and critical chain method to construct an integrated progress control system.

Zhang, Z[6], through integrating real-time data collection, intelligent decision-making, and automated scheduling technologies, monitored and optimized the entire drilling operation process, thereby achieving precise progress prediction and operation scheduling.

Jiao S and other scholars[7] applied BIM technology, through implementing dynamic simulation of the entire construction cycle and scheme iteration, to form an optimized progress plan and resource allocation scheme. This technology system effectively shortened the construction period, reduced the risk coefficient, provided technical support for efficient project delivery, and had significant practical value. In the same year, Li Y and other scholars[8] combined factors such as drilling fluid performance, drill bit wear, and geological conditions, and established a drilling progress prediction model through machine learning algorithms.

In conclusion, both at home and abroad, significant achievements have been made in the prediction of construction progress, providing theoretical support and engineering application value for the future development of the drilling engineering field.

## 2. Model for Precise Prediction of Construction Progress under Uncertain Conditions

### 2.1. LSTM (Long Short-Term Memory Network) Prediction Model

#### 2.1.1. Overview of LSTM Model

LSTM is a special neural network structure designed to address the problems of gradient vanishing and gradient explosion in traditional RNN. By incorporating memory units (Cells), it selectively retains and forgets information, enabling it to learn long-term dependencies in long sequences. LSTM consists of the following components:

- (1) Input Gate: Controls the retention of current input information.
- (2) Forget Gate: Controls whether information is forgotten.
- (3) Output Gate: Determines which information is output to the next time step.

#### 2.1.2. Design of LSTM Model for Drilling Construction Progress

The progress of drilling construction is influenced by many factors, such as geological hardness, equipment failure, personnel scheduling, resource consumption, wind speed, temperature, etc. To use LSTM for prediction, we need to model these factors as time series data and input them into the LSTM model.

- (1) Input features Suppose we use the following features to predict the progress of drilling:
  - 1) Equipment status (Et): Indicates the health status of the equipment, 0 for normal and 1 for failure. Equipment failure has a significant impact on the progress of drilling. Assuming that when there is a failure, the progress of drilling slows down by 50%.
  - 2) Personnel scheduling (Pt): Indicates whether there is a personnel scheduling problem, 0 for normal and 1 for insufficient. When there is insufficient personnel scheduling, the progress slows down by 20%.
  - 3) Resource consumption (Rt): The percentage of resources consumed during drilling. For every 10% increase in resource consumption, the progress slows down by 10%.
  - 4) Geological hardness (Gt): The geological hardness encountered during drilling. The greater the hardness, the higher the difficulty of drilling. For every 1 increase in hardness, the progress slows down by 10%.
  - 5) Wind speed (Wt): The wind speed on that day, which affects the working efficiency of drilling equipment. For every 1 m/s increase in wind speed, the progress slows down by 3%.

6) Temperature (T<sub>t</sub>): The temperature on that day, which affects the performance of equipment. For every 1°C increase in temperature, the progress slows down by 0.5%.

7) Progress of the previous day (Y<sub>t-1</sub>): The progress of the previous day is used as an input and as part of the time series.

(2) Working principle of LSTM model

The core of LSTM is the cell state (Cel State), which is similar to an information conveyor belt that runs through the entire time series. The cell state can maintain relatively stable information transmission at each time step, enabling the network to remember long-distance information. At each time step, the cell state is updated and transmitted based on the control of the input gate, forget gate, and output gate. The working principle is shown in Figure 1.

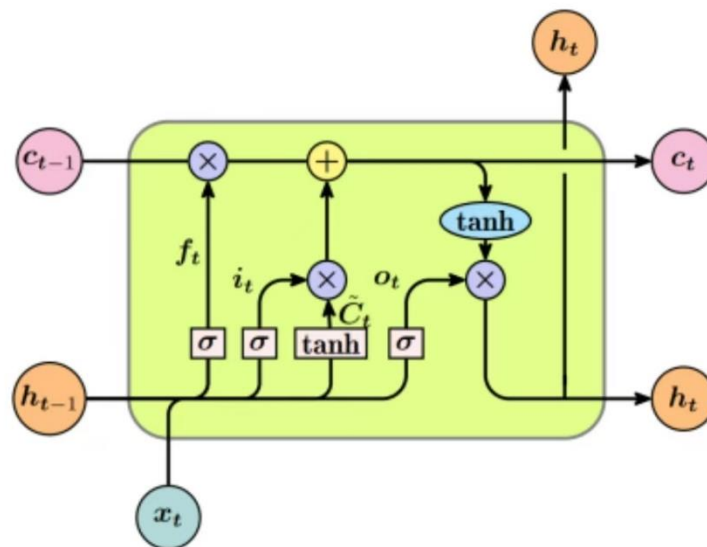


Fig 1. The working principle of LSTM model

The LSTM model processes the input data through the following formula and generates the predicted values for future drilling progress.

1) Forget Gate

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

Among them,  $f_t$  is the output of the forget gate,  $W_f$  is the weight matrix,  $b_f$  is the bias term,  $h_{t-1}$  is the hidden state of the previous time step, and  $x_t$  is the current input feature.

2) Input gate

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

Among them,  $i_t$  is the output of the input gate,  $W_i$  is the weight matrix, and  $b_i$  is the bias term.

3) Candidate memory unit

$$\tilde{C}_t = \tan h(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

Among them,  $\tilde{C}_t$  is the output of the candidate memory unit,  $W_c$  is the weight matrix, and  $b_c$  is the bias term.

4) Update the memory unit

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

Among them,  $C_t$  is the current moment's memory unit,  $C_{t-1}$  is the one from the previous moment.

5) Output gate

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t]) + b_o \tag{5}$$

Among them,  $O_t$  is the output of the output gate,  $W_o$  is the weight matrix, and  $b_o$  is the bias term.

6) Output state

$$h_t = O_t \cdot \tan h(C_t) \tag{6}$$

Among them,  $h_t$  represents the hidden state at the current moment and indicates the model's output.

### 2.1.3. Progress Training and Prediction of LSTM Model

(1) Training Steps

We arrange the data of these 10 days in chronological order and use the data of the previous 9 days to predict the progress of the 10th day.

1) Data Normalization: Normalize all input features (such as equipment status, personnel scheduling, resource consumption, geological hardness, etc.) to have a range within [0, 1], avoiding certain features dominating the training process.

2) Construction of Training Set and Test Set: Use the data of the previous 9 days as input and predict the progress of the 10th day. Perform multiple training and testing using the rolling window method.

3) Define the LSTM Network Structure:

Input Layer: Contains multiple features (equipment status, personnel scheduling, resource consumption, etc.).

LSTM Layer: One or more LSTM layers.

Output Layer: Prediction results, that is, the drilling progress on the t-th day.

Train the LSTM Model: Train the LSTM model using the data of the previous 9 days and optimize the model's weights by minimizing the loss function (such as Mean Squared Error MSE).

Prediction: Use the trained LSTM model to predict the drilling progress on the 10th day.

(2) Prediction Results

**Table 1.** Prediction Results

Actual drilling progress (m)	LSTM predicts the progress (m)
200	210

In this example, the prediction value of the LSTM model is 210 meters with an error of 10 meters. This error reflects the advantage of the LSTM model in capturing the long-term and short-term dependencies in time series data. Although it cannot directly model various external influencing factors (such as equipment failure, weather changes, etc.), it still has a good predictive ability when dealing with past drilling progress data.

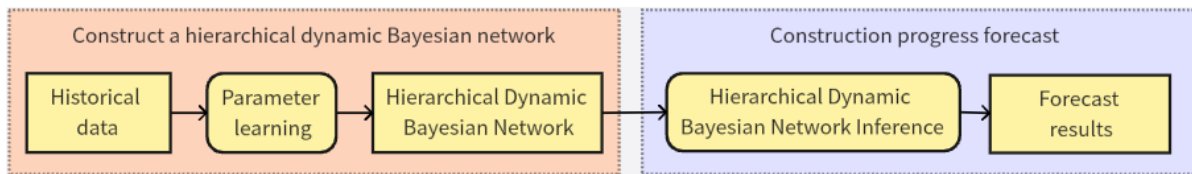
## 2.2. Construction Schedule Prediction Model Based on Hierarchical Dynamic Bayesian Network (HDBN)

### 2.2.1. Model Study

This section aims to construct a drilling progress prediction model for Hierarchical Dynamic Bayesian Network (HDBN). The following figure is a structural diagram for predicting drilling construction progress through HDBN, See Figure 2:

1. Construction Progress Prediction Model Based on Hierarchical Dynamic Bayesian Network Construction Process

(1) Construct a hierarchical dynamic Bayesian network model. Through the design of a hierarchical structure, this model can comprehensively consider the progress status of each level (such as the overall project, sub-project, drilling stage, etc.) of the drilling project and their mutual influence, thereby more accurately capturing the dynamic changes of the drilling progress.



**Fig 2.** Hierarchical Dynamic Bayesian Network Construction Schedule Prediction Structure Diagram

(2) Drilling construction progress prediction based on hierarchical dynamic Bayesian network. Through real-time monitoring of various state parameters during the drilling process and reasoning with the prediction model, the trend of future drilling progress changes can be effectively predicted.

### 2. Construction of Hierarchical Dynamic Bayesian Network Model

Hierarchical Dynamic Bayesian Network (HDBN) is introduced into the standard DBN based on the concept of hierarchical structure, aiming to handle more complex multi-scale or multi-granularity data sets. It can not only capture the influence patterns between different time slices, but also express the internal interaction of each subsystem within the same time slice and the interaction between different levels. For drilling construction progress prediction, HDBN can capture multiple hierarchical factors affecting the progress and reason in the dynamic evolution of the time series. In this way, HDBN can more accurately reflect the causal relationships and temporal dependencies in the drilling process.

### 3. Nodes and Inter-node Relationships

#### (1) Initial Distribution Nodes

Overall project progress: It represents the overall progress of the entire drilling project and serves as the top-level node. Its initial distribution can be determined through historical data and expert experience.

#### (2) Conditional Distribution Nodes

Sub-project progress: The progress of each sub-project (such as drilling preparation, drilling operation, completion operation, etc.) depends on the overall project progress and the progress of other sub-projects.

Influencing factors: Including geological conditions, equipment performance, weather changes, etc. The conditional distribution of these factors affects the progress of specific drilling stages.

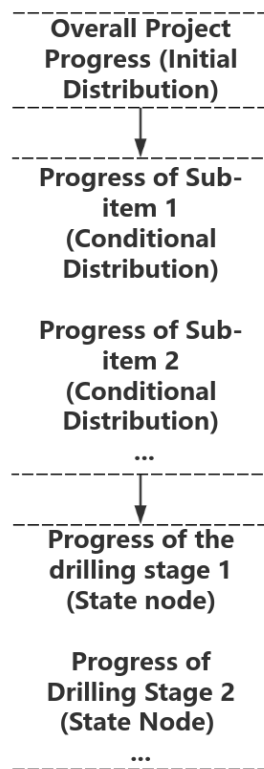
#### (3) State Nodes

Drilling stage progress: The progress status of each specific drilling stage (such as pre-drilling preparation, drilling, cementing, completion, etc.) depends on the corresponding sub-project progress and influencing factors. As shown in Figure 3:

### 2.2.2. HDBN Model Design and Inference Process

#### 1. Design of HDBN Model for Drilling Construction Progress Prediction

The prediction of drilling construction progress is influenced by various factors, such as geological conditions, weather changes, equipment failures, and personnel scheduling. In HDBN, we can capture the causal relationships and dynamic changes among these factors by introducing different levels of nodes.



**Fig 3.** Node distribution

**(1) Model Hierarchical Division**

We can divide the influencing factors of drilling construction progress into the following levels:

**1) High-level nodes (global factors)**

**Weather:** including wind speed, temperature, precipitation, etc.

**Equipment:** whether the equipment is working properly and whether there is a failure.

**Personnel:** whether there are sufficient personnel and whether the scheduling is reasonable.

**2) Middle-level nodes (local factors)**

**Resource Consumption:** the resources consumed during drilling (such as fuel, electricity, drill bit wear, etc.).

**Geology:** geological hardness, layer position changes, etc.

**progress layer: Drilling progress (Drilling Progress):** the specific drilling progress (such as the daily drilling depth or meters).

**(2) Conditional Probability Distribution (CPT) of HDBN Model**

**1) Equipment status (Et)**

When there is a failure, the drilling progress slows down by 50%.

**2) Personnel scheduling (Pt)**

When there is a shortage of personnel, the progress slows down by 20%.

**3) Resource consumption (Rt)**

When the resource consumption is too high, the progress slows down by 10%.

**4) Geological hardness (Gt)**

For every increase of 1 in hardness, the progress slows down by 10%.

**5) Wind speed (Wt)**

For every increase of 1 m/s in wind speed, the progress slows down by 3%.

**6) Temperature (Tt)**

For every increase of 1 °C in temperature, the progress slows down by 0.5%.

## 2. Inference Process of HDBN Model

In HDBN, the state of each node is inferred through the conditional probability distribution (CPT). Through the Bayesian inference process, we can predict the drilling progress based on the known influencing factors.

### (1) Time series model formula

For each time step  $t$ , the drilling progress is affected by the previous drilling progress, equipment status, personnel scheduling, resource consumption, geological hardness, and weather conditions (wind speed, temperature). The formula is as follows:

$$P(Y_t|Y_{t-1}, E_t, P_t, R_t, G_t, W_t, T_t) = f(Y_{t-1}, E_t, P_t, R_t, G_t, W_t, T_t) \tag{7}$$

Here,  $f$  represents a function, and the conditional probability distribution obtained through training is used to describe the relationships among various factors influencing the progress of drilling.

### (2) Inference process

Based on the known data of the previous 10 days, we can calculate the progress of drilling through the Bayesian inference process and the HDBN model. Using the hierarchical model, we first reason about the high-level factors (such as weather, equipment status), and then use the reasoning results to affect the progress of drilling at the lower levels.

### (3) Prediction result

**Table 2.** Prediction Results

Actual drilling progress (m)	HDBN predicts the progress (m)
200	195

## 3. Comparative Analysis of Models

(1) LSTM Prediction: The LSTM model learns the temporal dependencies of historical data and predicts the progress to be 210 meters, which is significantly higher. This is because LSTM does not explicitly model the nonlinear influencing factors such as equipment failures and wind speeds when dealing with environmental changes, resulting in less precise handling of complex influencing factors.

(2) HDBN Prediction: The HDBN model can better capture the impact of equipment failures, insufficient personnel scheduling, and resource consumption on the progress. The predicted result is 195 meters, which is closer to the actual progress of 200 meters. HDBN models the causal relationships of these factors and can make more accurate predictions when facing complex environments.

Based on the above comparison, HDBN performs better in drilling progress prediction and is suitable for scenarios that require handling multiple complex and interrelated influencing factors.

## 4. Summary

This paper studies the prediction model for drilling construction progress under uncertain conditions, and through analyzing the influencing factors such as geological conditions, equipment performance, and personnel operation, proposes two models for simulating random events during the drilling process under different geological conditions. The research results show that the model based on Dynamic Bayesian Network (HDBN) can better predict the progress of drilling construction, providing strong support for the management and decision-making of drilling engineering. The HDBN model can effectively predict the progress

distribution under different geological conditions by quantifying the randomness of uncertain factors, combining the probability distribution of geological parameters and the dynamic characteristics of the construction process, and provides a scientific basis for early identification of progress delay risks, optimization of resource allocation, and formulation of emergency plans.

This research provides theoretical support and practical guidance for the refined management of drilling engineering under complex environments, and has important engineering application value.

## References

- [1] He Minjie, Xu Lifeng, Wang Qiangqiang, et al. Dynamic Analysis Model and Progress Warning Mechanism Based on Fuzzy Algorithm for Construction Project Progress [J]. *Journal of Civil Engineering and Management*, 2017, 34(03): 86-90.
- [2] Xue Jianying, Tan Ping, Meng Fanmin, et al. Research on the Application of BIM and Earned Value Method in Construction Progress and Cost Control [J]. *Construction Economy*, 2019, 40(06): 115-119.
- [3] Wen Xin, Xie Yinpu, Zhang Guo, et al. Research on Construction of Real-time Progress Tracking and Analysis System for Drilling [J]. *China Management Informationization*, 2021, 24(04): 87-89.
- [4] Wu Yulin, Jiang Ying, et al. Prediction of Drilling Cycle Based on One-Dimensional Convolutional Neural Network [J]. *Block Oil and Gas Reservoirs*, 2023, 30(03): 495-504.
- [5] Zhang S, Liang C. Research on Construction Schedule Control Based on Critical Chain Method and BIM[J]. *Journal of Applied Science and Engineering Innovation*, 2018, 5(2).
- [6] Zhang, Z., & Huang, Z. Real-time drilling monitoring and progress prediction system based on intelligent analysis. *Journal of Energy Resources Technology*, 2020, 142(6), 061101.
- [7] Jiao S, Qing W, Chen P, et al. BIM-based ventilation shaft construction schedule simulation and plan optimization[J]. *IOP Conference Series: Earth and Environmental Science*, 2021, 634(1): 012102.
- [8] Li, Y., & Yang, L. Machine learning for drilling performance prediction and optimization: Applications and challenges. *Petroleum Exploration and Development*, 2021, 48(3), 520-533.