

# Forecasting the Environment Changes with the ARMA Model for Application to an Oil Tank Fire Hazard

Yuanzhi Ni<sup>1</sup>, Jie Zhang<sup>1</sup>, Ming Lyu<sup>2</sup>, Yuming Bo<sup>1</sup>

<sup>1</sup> School of Automation, Nanjing University of Science and Technology, Nanjing, 210094, China

<sup>2</sup> Information Overall Department, North Information Control Group Co., Ltd., Nanjing, 211153, China

**Abstract**—In oil tank farm, the fire hazard causes a great loss to both people and property. Once there is a fire, to avoid more loss, a reliable, secure and practical method for putting out a fire is needed. And under this circumstance, the accurate forecasting of the changes of the environment, such as temperature or humidity, is of great value. We employed time series analysis method to forecast the temperature and humidity changes when there was a fire accident in the oil tank farm. According to the record, we modelled a suitable ARMA model. And the results show that ARMA is an effective method to predict time varying series in the oil tank farm. Our results in this paper also provide an important way of studying the fire spreading in the oil tank farm.

**Keywords**—time series forecasting; oil tank farm; ARMA; environment change

## I. INTRODUCTION

In the oil tank farm, the fire hazard causes a great loss to both people and property. Once there is a fire, to avoid more loss, a reliable, secure and practical method for putting out a fire is needed. In order to reach this goal, the accurate forecasting of the changes of the environment, such as temperature or humidity, is of great value. Time series forecast focuses on the study combined application and theory. Alysha M. De Liveraa, Rob J. Hyndmana studied the complex seasonal patterns in time series forecasting in [1]. Short-term wind speed forecasting was studied in [2]. The oil market was also object of the study in [3]. There are many approaches for time series forecasting, Kalman filter, Regression analysis, neural network. However, these methods all have certain defects when considering the limitations in the oil tank farm fire hazard. ARMA Time series model is an effective method of forecasting or estimating the trends and changes based on a contiguous finite set of generally complex valued time series observations [4]. This method allows the modelling of the complex systems as “black-boxes” [5].

The ARMA model has been applied into many fields both in theoretical and practical. The problem of time series clustering was studied in [6]. ARMA was also combined with sequential Monte Carlo to estimate the stable process in [7]. Ergin Erdem and Jing Shi used the ARMA based approaches for the forecasting of the wind speed and direction in [8]. Alireza Erfani and Ahmad Jafari Samimi investigated the long memory of Stock Price Index and fitted a fractionally differenced ARMA Model to forecast

out-of-sample data in [9]. Yoshihiro Ohtsuka, Takashi Oga and Kazuhiko Kakamu forecasted electricity demand in Japan with a Bayesian spatial auto-regressive ARMA approach which performs better than traditional approach in [10]. Fong-Lin Chu forecasted tourism demand with ARMA-based methods and the general impression was that the ARMA-based models perform very well in [11]. What’s more, the ARMA, ARIMA, and some other methods were compared together in forecasting the monthly inflow of Dez dam reservoir in [12].

The contributions of this paper are twofold. First, we proposed an ARMA-based method to forecast the environment changes in oil tank farm fire accidents. In the application, we predicted the trend of changes without identifying the complex systems. Second, by introducing this method, we provide an economical way in the study of the fire spreading in the oil tank farm. The large amount of long term experiments can be partly replaced by the results of the forecasts.

## II. METHODOLOGY

In this section, we introduce the typical ARMA model and the evaluation of forecasts.

### A. ARMA model

A typical ARMA model can be described as the following form:

$$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \sum_{j=1}^q \phi_j \varepsilon_{t-j} + \varepsilon_t \quad (1)$$

where  $x_1, x_2, \dots, x_n$  are observed or forecasted values;  $\delta$  stands for constant term of the ARMA model;  $p$  is order of auto-regressive model;  $\varphi_1, \varphi_2, \dots, \varphi_p$  are auto-regressive coefficients;  $\phi_1, \phi_2, \dots, \phi_q$  are coefficients of  $q$ -order moving average model;  $\varepsilon_t$  is white noise with mean value 0, variance  $\sigma^2$ .

The first step of using ARMA model is to test the stationary of the series. The stationary assumption should be checked before identifying the model. The commonly used method is ADF test. For this purpose, the inspection of the run plots and Auto Correlation Function (ACF) plots can be used for deciding on the order of differencing.

In order to determine the order of ARMA model, we need to calculate the statics that describe the sequence

features. The ACF and PACF graphs have been used for deciding on the auto-regressive and moving average terms in the model.

If differencing is used, the model is converted to ARIMA model. At this time, the format of the model evolves into ARIMA ( $p, d, q$ ) model, where  $d$  is the difference order.

### B. Forecasting performance measure

In this study, we consider relative error (RE) as the evaluation of forecasts in (2):

$$RE = \frac{|\hat{x}_t - x_t|}{x_t} \times 100\% \quad (2)$$

Where  $\hat{x}_t$  is predicted value,  $x_t$  is actual measured value.

## III. DATA, MODELLING AND FORECASTING

In this section, we introduce the data source in briefly. Then, we analyze both temperature time varying series and humidity time varying series in the following part. The progress of modeling and estimation of the parameters is demonstrated in detail, and the results are presented in this section as well.

### A. Data source

By employing the practical facilities and sensors, the temperature and humidity data were measured and captured in the fire spreading laboratory in Nanjing University of Science and Technology (NJUST).

We designed and performed the experiments of oil tank fire hazard in small scale. The data were recorded continuously every 3 seconds during the whole fire incident to obtain the environment attributes. In the study, we used the first 100 groups of data to construct the forecasting models and made forecasts in the following 10 groups. The predicted values and real measured values are compared to test the performance of the models.

### B. Model building and parameter estimation

The initial data of temperature is showed in Fig.1.

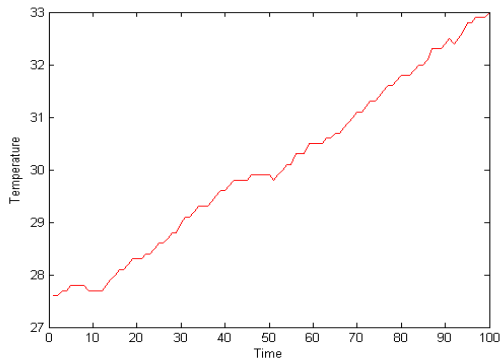


Figure 1. Temperature's curve

The curve shows a rising tendency, so we can give a rough judgment that the series are non-stationary. We give

an ADF unit root test for the temperature sequence, to verify whether it's smooth or not. As shown in Fig.2, it is obvious that the t-statistic for ADF test is -2.534764, larger than the threshold value of test level 1%, 5% and 10%, so we accept the hypothesis that temperature sequence has a unit root, it means that this sequence is not smooth. We can use differential transformation to meet sequence steady condition. After 2 times differential transformation, t-statistic value is -17.64616, much less than the threshold value of test level 1%, and sequence satisfies stationary condition, as shown in Fig. 3 and Fig.4.

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	1.302741	0.9986
Test critical values:		
1% level	-3.491345	
5% level	-2.888157	
10% level	-2.581041	

\*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(TEMPERATURE)  
Method: Least Squares

Figure 2. ADF unit root test for temperature sequence

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-11.43966	0.0000
Test critical values:		
1% level	-3.493747	
5% level	-2.889200	
10% level	-2.581596	

\*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(TEMPERATURE,3)  
Method: Least Squares

Figure 3. ADF unit root test for 2 times differenced temperature sequence

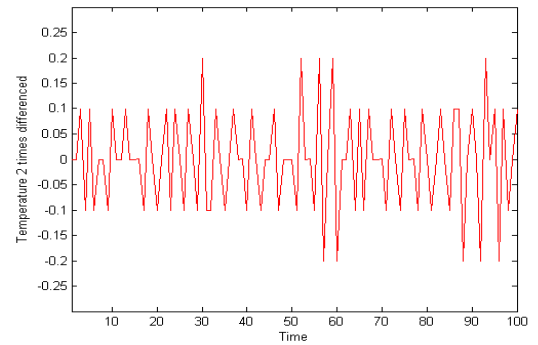


Figure 4. Temperature's curve after 2 times differencing

According to the ACF and PACF of temperature sequence given in Fig.5, we can determine the parameter of the model and q, specifically.

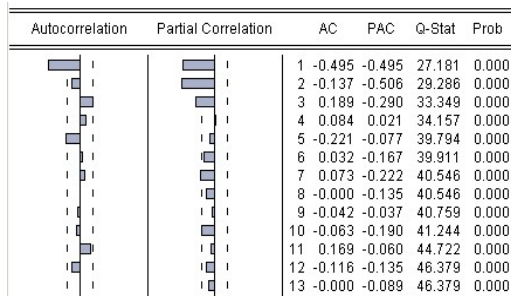


Figure 5. ACF and PACF of Temperature

PACF graph shows column on statistics at lag 2 or probably 3, however, it's not significant at other orders, and therefore the auto-regressive process's order should be 2 or 3. ACF becomes smaller after lag 3, so the moving average process is probably 2 or 3. In this case, we consider 4 forms for this model: ARIMA(2,2,2), ARIMA(2,2,3), ARIMA(3,2,2), ARIMA(3,2,3).

By comparing the coefficients of determination based on AIC and SC guidelines for those four models, as shown in Table 1, ARIMA(2,2,3) has the minimum coefficients, so we believe that ARIMA(2,2,3) is the most suitable model for this situation.

TABLE I. COEFFICIENTS OF DETERMINATION FOR TEMPERATURE MODEL

Model	AIC	SC
ARIMA(2,2,2)	-2.579294	-2.453660
ARIMA(2,2,3)	-2.613204	-2.462444
ARIMA(3,2,2)	-2.583858	-2.432204
ARIMA(3,2,3)	-2.610175	-2.433244

We can use the same method to predict humidity changes. The initial data of humidity is showed in Fig.6.

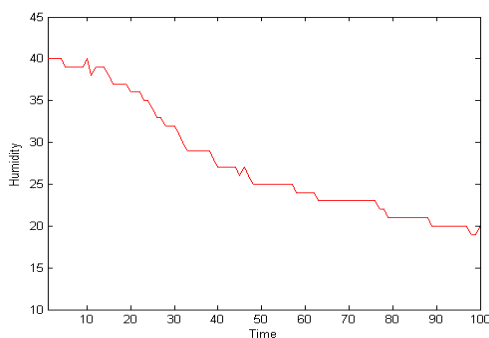


Figure 6. Humidity's curve

The number of the data is 110, and for the purpose of improvement of the forecast accuracy, the first 100 groups of the data are chosen as samples and the rest data are used for test.

The ADF test results tells the series are non-stationary, as shown in Fig.7, the t-statistic for ADF test is -3.372593, larger than the threshold value of test level 1%, so we cannot guarantee this sequence is smooth. Then we use 2 times differential transformation to meet sequence steady condition. The t-statistic value is -8.567498, much less than

the threshold value of test level 1%, and sequence satisfies stationary condition, as shown in Fig. 8 and Fig.9.

Null Hypothesis: HUMIDITY has a unit root  
Exogenous: Constant  
Lag Length: 0 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.867072	0.3467
Test critical values:		
1% level	-3.491345	
5% level	-2.888157	
10% level	-2.581041	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(HUMIDITY)  
Method: Least Squares

Figure 7. ADF unit root test for humidity sequence

Null Hypothesis: D(HUMIDITY,2) has a unit root  
Exogenous: Constant  
Lag Length: 7 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-7.650288	0.0000
Test critical values:		
1% level	-3.497029	
5% level	-2.890623	
10% level	-2.582353	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(HUMIDITY,3)  
Method: Least Squares

Figure 8. ADF unit root test for 2 times differenced humidity sequence

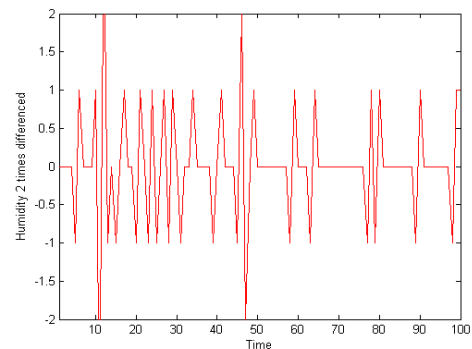


Figure 9. Humidity's curve after 2 times differencing

As the ACF and PACF of humidity sequence are given in Fig.10, we can estimate the parameters of the model generally. According to the graph features, the auto-regressive process's order should be 7, 8 or 9. The moving average process is no more than 2.

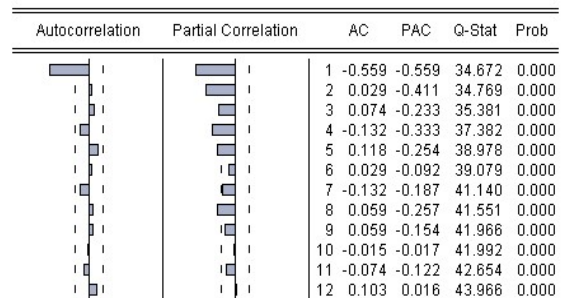


Figure 10. ACF and PACF of humidity

In this case, we consider 6 forms for this model: ARIMA(7,2,1), ARIMA(7,2,2), ARIMA(8,2,1), ARIMA(8,2,2), ARIMA(9,2,1), ARIMA(9,2,2).

TABLE II. COEFFICIENTS OF DETERMINATION FOR HUMIDITY MODEL

Model	AIC	SC
ARIMA(7,2,1)	1.640950	1.873980
ARIMA(7,2,2)	1.476052	1.734975
ARIMA(8,2,1)	1.529167	1.789684
ARIMA(8,2,2)	1.478900	1.765468
ARIMA(9,2,1)	1.451338	1.733685
ARIMA(9,2,2)	1.470481	1.785041

By comparing the coefficients of determination, as shown in Table 2, ARIMA (9, 2, 1) has the minimum coefficients, so we choose ARIMA (9, 2, 1) as the most suitable model.

### C. Forecast results

According to the models we selected, we have the forecasting results compared with real measured value, as shown in Fig.11 and Fig.12.

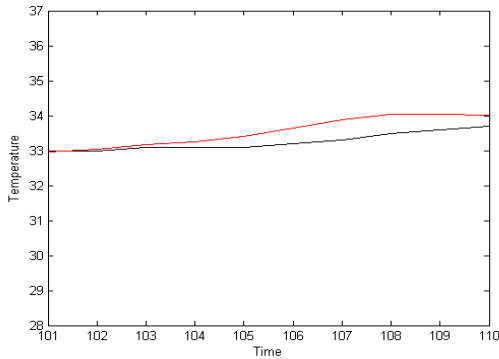


Figure 11. Comparison of the predicted value and the real measured value of temperature (red line is predicted value, black line is measured value)

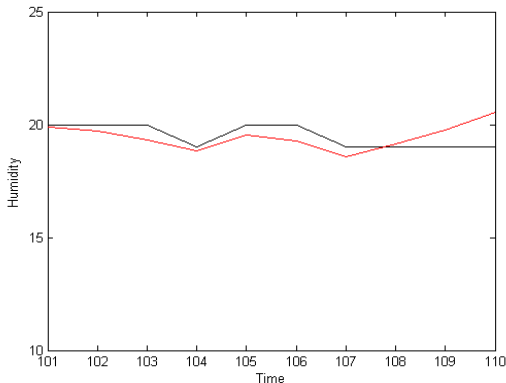


Figure 12. Comparison of the predicted value and the real measured value of humidity (red line is predicted value, black line is measured value)

In order to evaluate the performance of the model, we list the specific value in Table III and Table IV respectively.

TABLE III. PREDICTION FOR TEMPERATURE

No.	Actual value	Predict value	RE
101	33.000	32.968	0.10%
102	33.000	33.029	0.09%
103	33.100	33.171	0.21%
104	33.100	33.259	0.48%
105	33.100	33.410	0.94%
106	33.200	33.658	1.38%
107	33.300	33.874	1.72%
108	33.500	34.027	1.57%
109	33.600	34.027	1.27%
110	33.700	34.015	0.93%

Table III is the comparison between the actual value and the predicted value and the relative error from sequence number 101 to 110. The result shows that the effect of the forecast is decent and acceptable.

But the problem can also be easily indicated. The error of the forecast increases at the end of the series. This kind of phenomenon is not hard to explain. The error accumulation in every step contributes to it.

We can achieve a good result of humidity prediction as well, as shown in Table IV.

TABLE IV. PREDICTION FOR HUMIDITY

No.	Actual value	Predict value	RE
101	20	19.870	0.65%
102	20	19.701	1.49%
103	20	19.336	3.32%
104	19	18.837	0.86%
105	20	19.541	2.30%
106	20	19.258	3.71%
107	19	18.586	2.18%
108	19	19.132	0.69%
109	19	19.756	3.98%
110	19	20.553	8.17%

Table 4 gives the comparison between the actual humidity and the predicted humidity, and the most relative error fluctuates in an acceptable range. The result shows that the effect of the forecast is good as well.

## IV. CONCLUSIONS

In this paper, we employed time series analysis method to forecast the temperature changes when there was a fire accident in the oil tank farm. Without identification of the complex system, we regard the environment as "black-boxes". After recording enough statistics, we constructed some suitable ARMA models. These models could be used to represent the changing progress of the environment attributes in oil tank farm fire hazard. And the results showed that ARMA is an effective method to predict time

varying series in certain situations. Our results in this paper also provided an important way of study the fire spreading in the oil tank farm. To be precise, the forecast results can partly replace the dangerous and time-consuming experiments concerning fire spreading. Also, these models and forecasting results are significant to the further research in the prediction of fire hazard spread trend in oil tank farm.

#### V. ACKNOWLEDGEMENTS

This work has been supported by the National Natural Science Foundation of China(Grant No.61104109),the Natural Science Foundation of Jiangsu Province of China(Grant No.BK2011703),the Support of Science and Technology and Independent Innovation Foundation of Jiangsu Province of China(Grant No.BE2012178),and the Doctoral Fund of Ministry of Education of China(Grant No.20113219110027).

#### REFERENCES

- [1] De Livera A M, Hyndman R J, Snyder R D. Forecasting time series with complex seasonal patterns using exponential smoothing [J]. *Journal of the American Statistical Association*, 106(496): 1513-1527, 2011.
- [2] Lv P, Yue L. Short-term wind speed forecasting based on non-stationary time series analysis and ARCH model[C]//*Multimedia Technology (ICMT)*, 2011 International Conference on. IEEE: 2549-2553, 2011.
- [3] Cheong C W. Modeling and forecasting crude oil markets using ARCH-type models [J]. *Energy policy*, 37(6): 2346-2355, 2009.
- [4] Cadzow J A. ARMA time series modeling: an effective method [J]. *Aerospace and Electronic Systems, IEEE Transactions on*, (1): 49-58, 1983.
- [5] Cortez P, Rocha M, Neves J. Evolving time series forecasting ARMA models [J]. *Journal of Heuristics*, 10(4): 415-429, 2004.
- [6] Xiong Y, Yeung D Y. Time series clustering with ARMA mixtures [J]. *Pattern Recognition*, 37(8): 1675-1689, 2004.
- [7] Huang R, Zheng H, Kuruoglu E E. Time-varying ARMA stable process estimation using sequential Monte Carlo [J]. *Signal, Image and Video Processing*, 7(5): 951-958, 2013.
- [8] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction [J]. *Applied Energy*, 88(4): 1405-1414, 2011.
- [9] Erfani A, Samimi A J. Long Memory Forecasting of Stock Price Index Using a Fractionally Differenced Arma Model [J]. *Journal of Applied Sciences Research*, 5(10): 1721-1731, 2009.
- [10] Ohtsuka Y, Oga T, Kakamu K. Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA approach [J]. *Computational Statistics & Data Analysis*, 54(11): 2721-2735, 2010.
- [11] Chu F L. Forecasting tourism demand with ARMA-based methods [J]. *Tourism Management*, 30(5): 740-751, 2009.
- [12] Valipour M, Banihabib M E, Behbahani S M R. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir[J]. *Journal of Hydrology*, 476: 433-441, 2013