International Journal of Smart Home
Vol. 10, No. 1, (2016), pp. 23-30
http://dx.doi.org/10.14257/ijsh.2016.10.1.03

The Storage Grain and Environment Modeling Based on TS-PLS

Lan Wu, Yanbo Hui and Mingyue Li

College of Electrical Engineering, Henan University of Technology, China
richod@126.com

Abstract

Food security is a strategic cornerstone of national security, and the food safety storage is a recognized problem. Accurate representation of grain storage environment is a fundamental premise for secure storage, which is very difficult because of the complex linkage relation between stored grain and grain storage environment. The characteristics between grain and environment, such as large time lag, two-way coupling, nonlinear, time-varying, make analysis and modeling face more difficulties and challenges. Based on this, combining PLS and TS fuzzy model, a modeling method of dynamic nonlinear multi-input multi-output (MIMO) system is proposed in this paper, which effectively solves the problem of strong nonlinear and correlation, at the same time reduces the computational complexity of data modeling methods. The simulation results show that this modeling could effectively describe the relationship between storage grain and environment.

Keywords*: food security, storage environment, PLS, TS fuzzy model*

1. Introduction

Food security is a strategic cornerstone of national security, and the food safety storage is a recognized problem. According to the research institutes and experts predict that due to the poor facilities, grain handling and transportation spilled omissions, excessive or extensive processing, China's annual grain loss is more than 700 billion kilograms, equivalent to 200 million people a year rations. In this more than 700 billion kilograms of grain lost in vain, the low storage techniques and grain storage environment unsuitable makes food happen rodents, insects and mildew and other phenomena which caused grain loss rate is about 8%, accounting annual household savings loss of about 400 billion kilograms of grain sectors, the storage and transportation business losses more than 150 billion kilograms, the processing enterprise losses more than 150 billion kilograms. These are just food deterioration caused direct loss, if coupled with indirect losses caused by quality deterioration, its serious impact can be imagined. Total national demand will reach 5725 billion kilograms of grain by 2020, and grain production capacity needed to reach 5450 billion kilograms. If only to reduce one percentage point of grain storage loss, to reduce annual losses could reduce about 60 billion pounds, which close to the sum of annual production in Shaanxi Province, one of the big grain output province.

Wheat is a living body. It is not a simple or composite organisms, but extremely complex organism. Wheat contains embryo and endosperm, by cortical bag into a unity. The endosperm is mainly composed of protein and carbohydrates, and the embryo consists of lipids, and a variety of enzymes. The wheat quality changes in the embryo and endosperm simultaneously, which is a complex dynamic multiple-factor change. The change of wheat quality is caused by various factors such as temperature, humidity, gas, insects, mold, and so on. When certain conditions of temperature and humidity, it will produce mold and pests, which could make the temperature, humidity and gas concentration change. There is a great correlation between them, under certain conditions, the stored grain will happen fission trend, occur condensation, vermin, mildew and other phenomena until fever and deterioration. Visible between food quality and storage

ISSN: 1975-4094 IJSH
Copyright ⓒ 2016 SERSC

environment, changing in itself, it is also constantly interact with each other, which is a big time lag, multi-input multi-output (MIMO) nonlinear, variable dynamic processes, these characteristics make mechanism relationship between quality and environment become very complex[1-2]. At the same time, the food quality detection has some problems of can't online real-time and artificial periodic sampling, which makes the predictive control of grain storage environment and quality become very important.

Existing mechanism analysis and modeling between the environment and the quality of stored grain, mostly based on experimental methods of qualitative analysis and quantitative analysis using the mechanism modeling method [3-6]. Based on the experimental methods, using software simulation or experimental equipment set different storage temperature and humidity, to observe food quality changes with time, through the analysis of data collected, fits the change law of food quality in certain environment with time variation. It analyzes the trends between quality and the environment from a qualitative point of view, but with a certain degree of uncertainty and one-sidedness [7]. Mechanism modeling method, with multi-sensor data fusion, to estimate the entire grain storage environment, and research the entire distribution rule, but more focused on the modeling of temperature field, cannot reflect the more detailed local temperature in grain heap. Since the warehouse environment is not only effected by the external environment, but also by the degradation mechanism of grain quality, so it is unscientific to characterize the relationship between the environment and the quality from certain aspect.

Based on this, combination the TS fuzzy model and dynamic PLS model, a modeling method between stored grain and the environment is proposed in this paper, in order to accurately describe the grain storage environment, to ensure the safety of stored grain. This model has the ability to automatically decoupling and matching circuits same as PLS model [8-9], the capacity of deal with nonlinear same as the TS fuzzy model [10], and simplifies and ensures prediction and tracking performance. This method has obvious advantages for strongly nonlinear, MIMO process system.

2. The Modeling for Storage Grain and Environment based on TS-PLS

The system, using the decoupling and dimension reduction characteristics of dynamic PLS, simplifies the MIMO system into a plurality of single-input single-output (SISO) system, resolves data processing complexity and multi-variety correlation. Using fuzzy modeling method resolves effectively the strong nonlinear of crop growth process, establishes a TS-PLS dynamic model for storage grain and environment linkage system, and finally achieves the real-time prediction monitoring of grain quality. Specific method is as follows.

2.1. The Nonlinear PLS Establish External Model

2.1.1. Traditional PLS Model: grain storage process, various factors affect the quality of grain as an argument input data, referred to as $X=(x_{ij})_{l*m}$, grain quality parameters as the dependent variable output data, referred to as $Y=(y_{ij})_{l*n}$, where, *l*, *m*, *n* denote the number of observations systems, input and output variables of dimension. External model can be obtained by calculation, iteration and drawing main element from the input data *X* and output data *Y*:

$$\begin{aligned} X - E^{*} &= TP^{T} = \sum_{r=1}^{R} t_r p_r^T \\ Y - F^{*} &= UQ^{T} = TBQ^{T} = \sum_{r=1}^{R} t_r b_r q_r^T \end{aligned} \tag{1}$$

Where: $R(R<\max(m,n))$ is the number of principal component selection, can be determined by cross-examination or other statistical methods; E^* *and* F^* represents

Copyright ⓒ 2016 SERSC

respectively residual matrix of the input data X and output data Y; t_r *and* p_r denote the r-hidden variables and their input load vector; u_r and q_r represent the r-th output of hidden variables and load vector. Input and output of hidden variables t_r and u_r hidden variable algebraic relations constitutes the internal model of PLS algorithm.

2.1.2. Dynamic PLS Model: Input $x(t)$ after Wx normalized, as the PLS input, obtains the PLS output $y^{PLS}(t)$, together with the ARX gets the model output $y_M(t)$. The process output $y(t)$, as the modeling data, after the lag factor becomes part of the modeling data. In the use of the dynamic predictive PLS model, the model output data participates prediction as part of the model. The whole model can be represented by Equation (2), PLS section matches the variable firstly, and then ARX model is used for modeling the process dynamics property, and the two parts constitute the entire model output.

$$\begin{aligned} y^{PLS}(t) &= PLS\left(x(t)\right) \\ \hat{y}(t+1) &= \sum_{j=1}^{Na} A_j y(t+1-j) + y^{PLS}(t-d) \end{aligned} \tag{2}$$

Where, Na represents the output order, d represents pure time delay. $y^{PLS}(k)$ is calculated by a conventional PLS method.

According to the above model construction, the model building process is as follows:

(1) To create the model, firstly determine the order of ARX model Na, the search term of delay item d of y^{PLS}. While determine the initial value Na and pure delay d of ARX model, and initialize the ARX model parameters A_j.

(2) According to the initial value Na, d and the parameters A_j, transmit the output data into the desired output data of PLS model, and with input data construct PLS model using traditional PLS method.

(3) According to the PLS model, calculating PLS model output Y^{PLS}, new ARX model parameters A_j by the process output y, and re-calculating the new model output $\hat{y}(t+1)$.

(4) Repeating step (2) and (3), until the model converges.

(5) When over the step (4), if the difference between the predicted values and the true value is still higher than the threshold, modify Na and d, until find a better model.

2.2. TS-PLS Model

TS fuzzy model is in each fuzzy rule uses a linear function to represent a local feature, then the membership function of these local model outputs are combined to form an approximate output of the entire model. TS fuzzy model contains a former piece part and the consequent parts, and its identification contains structure identification and parameter identification. The structure identification contains input/output variables selection, fuzzy sets, the type of membership function, the number of fuzzy rules, etc. Parameter identification includes parameter identification and the consequent membership function model. Wherein, after the member model uses ARX models.

In the PLS outer model, MIMO model identification problem is reduced to a plurality of SISO models, here we only need to consider each one separately hidden space SISO dynamic fuzzy inference system. In the *i*-th hidden variable space, using TS fuzzy model to regress score vector t_i and u_i. The following with t and u represent a hidden space dimension of input and output. Definition:

$$q^{-1}y(k) = y(k-1),\ A_l = a_{l0} + a_{l1}q^{-1} + \ldots + a_{\ln}q^{-n},\ B_l = b_{l0} + b_{l1}q^{-1} + \ldots + b_{lm}q^{-m}$$

The *l*-th rule expressed as ARX form:

$$u_l(k) = A_l u(k-1) + B_l t(k-1) \tag{3}$$

Define:

$$\varphi(k) = [u(k-1), u(k-2), \ldots, u(k-1-n), t(k-1), t(k-2), \ldots, t(k-1-m)]$$

$$\theta_l = [a_{l0}, a_{l1}, ..., a_{ln}, b_{l0}, b_{l1}, ..., b_{lm}]$$

Where, n, m is the order of input and output variables. Equation (3) can be expressed as:

$$u_l(k) = A_l u(k-1) + B_l t(k-1) = \varphi(k)\theta_l^T \tag{4}$$

Then, the TS fuzzy model can be expressed as:

$$R_l : if \ \varphi(k) \in \Omega_l$$

$$then \ \ u_l(k) = \varphi(k)\theta_l^T \quad l = 1,2,..., L$$

Where, L defines the number of rules, Ω_l is fuzzy sets, the hidden space output of k time is:

$$u(k) = \frac{\sum_{l=1}^{L} \omega_l(\varphi_k) u_l(k)}{\sum_{l=1}^{L} \omega_l(\varphi_k)} \tag{5}$$

The φk *and* $\varphi(k)$ is equivalent, $\omega_l(\varphi k)$ is the membership functions of $\varphi(k)$ belonging to the l rule. The membership function solved by GK clustering algorithm, equation (5) can be rewritten as:

$$u(k) = \sum_{l=1}^{L} w_l(\varphi_k) u_l(k)$$

$$w_l(\varphi_k) = \frac{\omega_l(\varphi_k)}{\sum_{l=1}^{L} \omega_l(\varphi_k)} \tag{6}$$

Equation (6) is hidden space TS fuzzy model, membership functions' identification and parameter identification of ARX model is to be carried out.

In this paper, using GK clustering algorithm, the membership function is to be solved. Is assumed to be clustering data set, N is the number of samples, the number of clusters is set to c (c=L). GK clustering algorithm to calculate the fuzzy partition matrix $U=(\mu_{ij})_{c*N}$, cluster center vectors $V=[v_1,v_2,...,v_c]$ and covariance matrix $F=[F_1,F_2,...,F_c]$, where, F_i is the i-th cluster of fuzzy covariance matrix, is a positive definite matrix, $\mu_{ij} \in [0,1]$ is membership between the data points z_j and relative cluster centers of v_i. Define the distance square between data point z_j and cluster center v_i as follows:

$$D_{ij}^2 = (z_j - v_i)^T A_i (z_j - v_i), A_i = [\det(F_i)]^{1/(n+m)} \cdot F_i^{-1} \tag{7}$$

The GK clustering algorithm's step is as follows:

Step 1: For a given data set to be cluster Z, select the number of clusters c, fuzzy index r>1, the convergence error ε>0, random initialization fuzzy partition matrix U;

Step 2: Calculate the cluster centers

$$v_i = \frac{\sum_{j=1}^{N} (\mu_{ij})^r z_j}{\sum_{j=1}^{N} (\mu_{ij})^r} \tag{8}$$

Step 3: Calculate the cluster co-variance matrix:

Copyright ⓒ 2016 SERSC

$$F_i = \frac{\sum_{j=1}^{N} (\mu_{ij})^r (z_j - v_i)(z_j - v_i)^T}{\sum_{j=1}^{N} (\mu_{ij})^r} \tag{9}$$

Step 4: Calculate the distance between data points z_j and cluster center v_i :

$$D_{ij}^2 = (z_j - v_i)^T [\det(F_i)]^{1/(n+m)} \cdot F_i^{-1} (z_j - v_i) \tag{10}$$

Step 5: Update fuzzy partition matrix:

$$\mu_{ij} = \left[\sum_{k=1}^{c} (D_{ij} / D_{kj}) \right]^{-2/(r-1)} \tag{11}$$

Repeat step 2 to 5, until the two iterations U is less than convergence error ε. Obtained U, F and V, at time k can be implicit spatial output $u(k)$ is expressed as:

$$\begin{aligned} u(k) &= \sum_{l=1}^{L} w_l(\varphi_k) u_l(k) = \sum_{l=1}^{L} w_l(\varphi_k) \varphi(k) \theta_l^T \\ &= [w_1(\varphi_k)\varphi(k), w_2(\varphi_k)\varphi(k), \ldots, w_L(\varphi_k)\varphi(k)] \cdot [\theta_1, \theta_2, \ldots, \theta_L]^T \end{aligned} \tag{12}$$

Set

$$\begin{aligned} \psi &= [w_1(\varphi_k)\varphi(k), w_2(\varphi_k)\varphi(k), \ldots, w_L(\varphi_k)\varphi(k)] \\ \theta &= [\theta_1, \theta_2, \ldots, \theta_L] \end{aligned}$$

Equation (12) can be expressed as:

$$u(k) = \psi \cdot \theta^T \tag{13}$$

u(k) and ψ *is* known, unknown parameters are determined by the least squares method, the identification of the entire TS-PLS model is completed.

3. Simulation Results and Analysis

In this section, the PLS and TS fuzzy modeling algorithm is mainly validated. In Figure 1, above is comparison the output between identification system and actual system, below is the error curve between identification system and actual system. It is can be seen that the PLS model could fit and predict the actual system, its error limits within ± 0.05 while the signal range is within ± 10. Figure 2 is prediction behavior of the dynamic PLS modeling, from prediction value and true value, we can see that this algorithm is very effective.

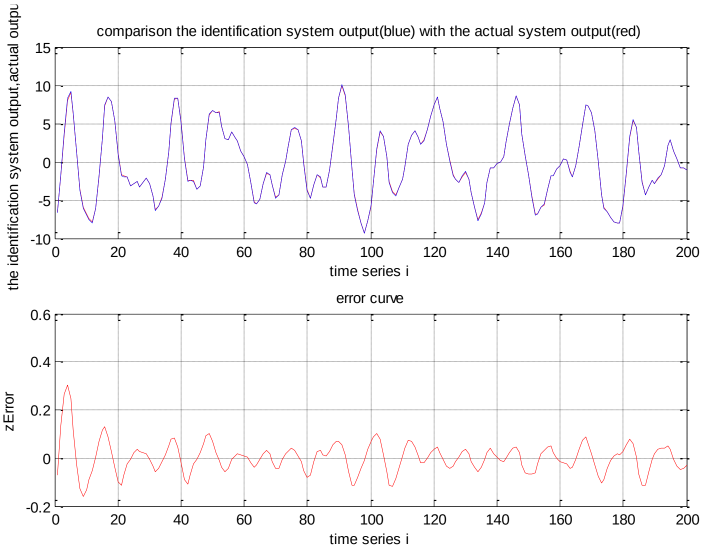

Figure 1. Comparison the Output between Identification System and Actual System

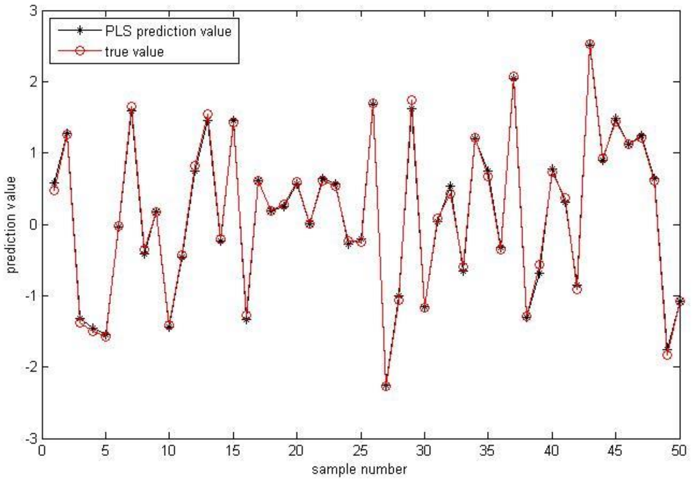

Figure 2. The Dynamic PLS Prediction Behavior

Figure 3 is the algorithm validation. In Figure 3, using 1000 test data to validate the effectiveness of TS model, the TS model output could always follow the original signal very well. Figure 4 is whole system performance. Using 500 data further enlarge following effect. We can conclude that the whole modeling algorithm could follow the trend of actual system change, but exist some errors, which are in allowable range.

Copyright ⓒ 2016 SERSC

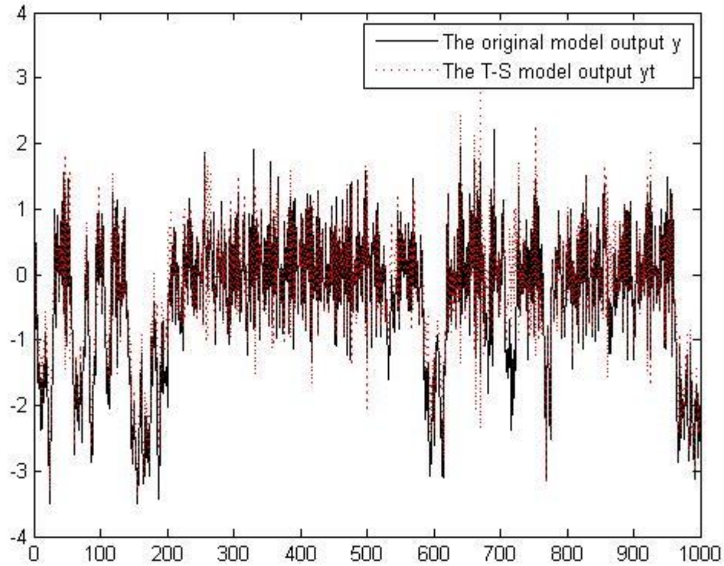

Figure 3. The Comparison between TS Model and Actual System

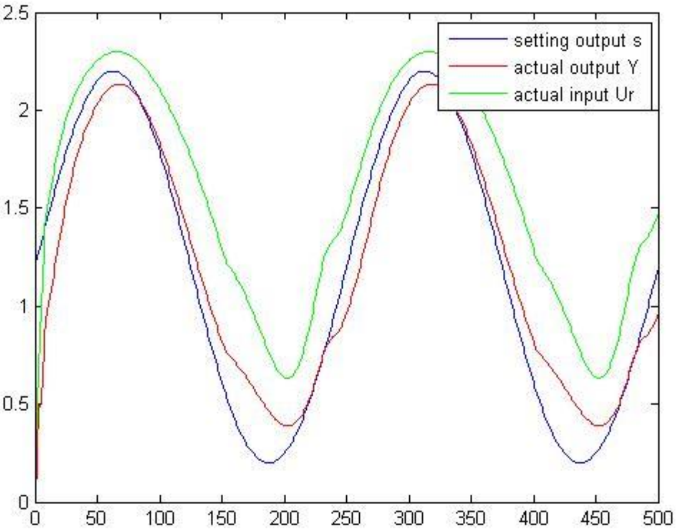

Figure 4. The System Performance

4. Conclusion

Aiming to the characteristics between grain and environment, such as large time lag, two-way coupling, nonlinear, time-varying. combining PLS and TS fuzzy model, a modeling method of dynamic nonlinear multi-input multi-output (MIMO) system is proposed in this paper, which effectively solves the problem of strong nonlinear and correlation, at the same time reduces the computational complexity of data modeling methods. The simulation results show that this dynamic PLS model could effectively predict the actual system, and TS model can more accurate fitting the nonlinear systems.

Reference

[1] D. S. Jayas, "Storing Grains for Food Security and Sustainability", Agric Res, vol. 1, no. 1, **(2012)**, pp. 21-24.
[2] F. Jian and D. S. Jayas, "The Ecosystem Approach to Grain Storage", Agric Res, vol. 1, no. 2, **(2012)**, pp. 148-156.
[3] M. Shrern, B. Simmons, M. Smit and M. Litoiu, "Toward an Ecosystem for Precision Sharing of Segmented Big Data", IEEE Sixth International Conference Cloud Computing, **(2013)**.
[4] G. M. Bender, L. Kot, J. Gehrke and C. Koch, "Fine-grained disclosure control for app ecosystems", Proceedings of the ACM SIGMOD International Conference on Management of data, **(2013)**.
[5] J. Liu, H. Zhou and X. Zhou, "Application and comparison of two mathematical models for simulating grain heat and mass transfer during in-bin drying", International Journal of Digital Content Technology and its Applications, vol. 6, no. 6, **(2012)**, pp. 200-208.
[6] Y. Zhang, X. Xie and B. Huang, "The application of fiber Bragg grating temperature testing system in the barn", Optical Technique, vol. 38, no. 3, **(2012)**, pp. 288-293.
[7] R. L. Wang, X. L. Liu and Y. Zhao, "Effect of Storage Micro-environment on the Variation of Wheat Protein", Modern Food Science and Technology, vol. 30, no. 6, **(2014)**, pp. 47-50.
[8] J. Liang, X. Wang and W. Wang, "Soft-sensor modeling via neural network PLS approach", Journal of Zhejiang University(Engineering Science), vol. 38, no. 6, **(2004)**, pp. 676-681.
[9] X. L. Zhang, L. S. Hu and G. Y. Cao, "Dynamic PLS Modeling Based on Process Data", Journal of System Simulation, vol. 20, no. 10, **(2008)**, pp. 2686-2688.
[10] Z. Liu, Z. Chen and Z. Yuan, "T-S Model Based Fuzzy Generalized Predictive Control", Acta Scientiarum Naturalium Universitatis Nankaiensis, vol. 33, no. 4, **(2000)**, pp. 114-118.

Authors

Lan Wu, research in the food safety information processing.

Yanbo Hui, research in the food safety information processing.

Mingyue Li, research in the food safety information processing.

 Copyright ⓒ 2016 SERSC