Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/070427

A Fuzzy Collaborative Forecasting Approach for WIP Level Estimation in a Wafer Fabrication Factory

*Toly Chen**, *Chi-Wei Lin, Yi-Chi Wang*

Department of Industrial Engineering and Systems Management, Feng Chia University, Taichung City, 407 Taiwan

Received: 17 Nov. 2012, Revised: 26 Feb. 2013, Accepted: 14 Mar. 2013
Published online: 1 Jul. 2013

Abstract: Forecasting the work-in-process (WIP) level is very important for the control of factory. However, the uncertainty in the WIP level is not easy to deal with. To solve this problem, a fuzzy collaborative forecasting approach is proposed. And also we show the effectiveness of the proposed methodology with a case study.

Keywords: Fuzzy collaborative intelligence, work in process, forecasting.

1. Introduction

Work in process (WIP) indicates the set of unfinished products in the whole factory. Forecasting the WIP level is very important for the control of factory because of several reasons. For example, the accumulation of WIP leads to factory inefficiency [9] In addition, the WIP level is an important parameter that can be properly used to trigger the decision of when to release specific jobs [5]. In the literature, Gupta and Ho [2] proposed the workload balancing algorithm, to minimize the total squared workload deviation on parallel machines. The simulation results showed that the proposed algorithm is a very useful method to determine the standard WIP level efficiently. Qi et al. [9] constructed a simulation model to analyze the effects of arrival distribution, batch size, downtime pattern, and lot release control on cycle time, WIP level, and equipment utilization. In Iriuchijima et al.'s viewpoint [3], appropriate WIP planning can effectively balance production for fixed orders with that for projected demand. Wang et al. [11] established a real-time WIP status monitoring model. Lin and Lee [5] proposed a queuing network-based algorithm to determine the total standard WIP level.

Many factors, such as the release policy, priority combination, the scheduling rule, and the due date assignment policy, will affect the WIP level. As a result, there is considerable uncertainty in the WIP level. However, because of too much human intervention, stochastic methods are not suitable to deal with this uncertainty [4,10]. For this reason, Lin et al. [4] proposed a fuzzy-neural approach, in which some fuzzy linear regression (FLR) equations [7] based on different points of view were used to predict the WIP level, and then a back propagation network (BPN) was constructed to integrate the forecasts to get a single value.

In order to effectively predict the future WIP level, a fuzzy collaborative forecasting approach is proposed. Fuzzy collaborative forecasting is a fairly new field in soft computing, but has considerable potential. In the fuzzy collaborative forecasting approach, a number of experts in this field are invited. These domain experts provide their points of view for the WIP level forecasting. These views are incorporated into the corresponding fuzzy back propagation network (FBPN) to predict the WIP level. Forecasts based on different points of view will be different in nature, so there is space for collaboration. In addition, a representative value has to be concluded from the forecasts. To achieve these goals, a radial basis function (RBF) network is used. Finally, the effectiveness of the fuzzy collaborative forecasting approach is shown with a case study.

2. FBPN for the WIP Level Forecasting

Variables and parameters in the proposed methodology are defined:

(1) a_t: the normalized value of the WIP level at period t.

* Corresponding author e-mail: tolychen@ms37.hinet.net

(2) $\tilde{W}_t$: the WIP level forecast at period t.
(3) $\tilde{o}(t)$: the network output, which is the normalized value of the WIP level forecast at period t.
(4) $\tilde{h}_l(t)$: the output from node l in the hidden layer, $l = 1 \sim L$.
(5) $\tilde{w}^o_l(t)$: the weight of the connection between node l in the hidden layer and the output node.
(6) $\tilde{w}^h_{kl}(t)$: the weight of the connection between input node k and node l in the hidden layer, $k = 1 \sim K$; $l = 1 \sim L$.
(7) $\tilde{\theta}^h_l(t)$: the threshold for blocking off weak signals on node l in the hidden layer.
(8) $\tilde{\theta}^o(t)$: the threshold for blocking off weak signals on the output node.

All fuzzy parameters and variables are given in triangular fuzzy numbers (TFNs), e.g. $\tilde{W}_t = (W_{t1}, W_{t2}, W_{t3})$.

In the fuzzy collaborative forecasting approach, there are G experts, and these experts provide their views on the following issues:

(1) The width of the fuzzy WIP level forecast should be less than $\Psi(g)$.
(2) The membership of the actual value in the fuzzy WIP level forecast should be higher than $s_R(g)$ if the actual value is on the right-hand side of the fuzzy forecast.
(3) The membership of the actual value in the fuzzy WIP level forecast should be higher than $s_L(g)$ if the actual value is on the left-hand side of the fuzzy forecast.

$g = 1 \sim G$.

In the fuzzy collaborative forecasting approach, the view of each expert is incorporated into the corresponding FBPN that is used to predict the WIP level. The configurations of the FBPNs used by the experts are the same:

(1) Number of inputs: K, corresponding to the historical data of the previous K periods after normalization.
(2) Single hidden layer.
(3) Number of nodes in the hidden layer: $2K$.
(4) Network output: the normalized WIP level forecast.
(5) Transformation/Activation functions: For the hidden layer, the hyperbolic tangent sigmoid function is used, while for the others the linear activation function is used.

FBPN training is divided into three steps. The first step is to determine the center value of each parameter. To this end, we consider FBPN as a non-fuzzy BPN, and apply the Levenberg-Marquardt algorithm to solve.

Subsequently, to determine the upper bound of each parameter, such as w^h_{kl3}, θ^h_{l3}, w^o_{l3}, and θ^o_3, the following goal programming (GP) problem can be solved [1]:

Min
$$\sum_{\text{all } t} \pi_t$$
subject to

$$\ln(\frac{1}{W_{t3}} - 1) = \theta^o_3 - \sum_{\text{all } l} w^o_{l3} h_{l3},$$

$$\sum_{\text{all } l} w^o_{l3} h_{l3} - \theta^o_3 = -\ln(1/\pi_t - 1),$$

$$\sum_{\text{all } l} w^o_{l3} h_{l3} - \theta^o_3 \leq -\ln(1/\Psi(g) - 1),$$

$$\sum_{\text{all } l} w^o_{l3} h_{l3} \leq \theta^o_3 - \ln(\frac{1 - s_R(g)}{a_t - s_R(g) W_{t2}} - 1),$$

$$\sum_{\text{all } l} w^o_{l3} h_{l3} \geq \theta^o_3 - \ln(\frac{1}{a_t} - 1),$$

$$\sum_{\text{all } k} w^h_{kl3} x_k - \theta^h_{l3} \geq -\ln(1/h_{l3} - 1),$$

$$\sum_{\text{all } k} w^h_{kl3} x_k - \theta^h_{l3} \leq -\ln(1/h_{l3} - 1),$$

$$k = 1 \sim K; l = 1 \sim L.$$

All actual values will be contained in the corresponding fuzzy forecasts. The objective function is to minimize the sum of the half-ranges (π_t) of the fuzzy WIP level forecasts, calculated according to the second constraint. The third constraint forces the upper bound to meet the expert's requirement $\Psi(g)$. The other constraints limit the changes that should be made to the network parameters (w^h_{il3}, θ^h_{l3}, w^o_{l3}, and θ^o_3) for the same purpose. $s_R(g)$ is referenced in the fourth constraint. If $s_R(g)$ is high, then h_{l3} will be large, leading to a large π_t. After enumerating a number of possible values for them, the goal programming problem is solved many times. In these optimization results, the best one giving the minimum upper bound is chosen.

In a similar way, to determine the lower bound of each parameter (e.g. w^h_{kl1}, θ^h_{l1}, w^o_{l1}, and θ^o_1), the following GP problem is solved:

Min $$\sum_{\text{all } t} \pi_t$$

subject to

$$\ln(\frac{1}{W_{t1}} - 1) = \theta^o_1 - \sum_{\text{all } l} w^o_{l1} h_{l1},$$

$$\sum_{\text{all } l} w^o_{l1} h_{l1} - \theta^o_1 = -\ln(1/\pi_t - 1),$$

$$\sum_{\text{all } l} w^o_{l1} h_{l1} - \theta^o_1 \leq -\ln(1/\Psi(g) - 1),$$

$$\sum_{\text{all } l} w^o_{l1} h_{l1} \leq \theta^o_1 - \ln(\frac{1 - s_L(g)}{a_t - s_L(g) W_{t2}} - 1),$$

$$\sum_{\text{all } l} w^o_{l1} h_{l1} \leq \theta^o_1 - \ln(\frac{1}{a_t} - 1),$$

$$\sum_{\text{all } k} w^h_{kl1} x_k - \theta^h_{l1} \geq -\ln(1/h_{l1} - 1),$$

$$\sum_{\text{all } k} w^h_{kl1} x_k - \theta^h_{l1} \leq -\ln(1/h_{l1} - 1),$$

$$k = 1 \sim K; l = 1 \sim L.$$

$s_L(g)$ is referenced in the fourth constraint. If $s_L(g)$ is high, then h_{l1} will be small, leading to a large π_t.

The forecasting results by these experts can be communicated to each other, so that they can modify their views, and generate more accurate forecasts if all viewpoints are taken into account. To this end, the GP problems are modified to include a collaborative mechanism in the next section.

3. Collaboration Mechanism and Protocol

The view of a domain expert, indicated with $VS_g = \{\Psi(g),\ s_R(g), s_L(g)\}$, is packaged into an information granule encoded using extensible markup language (XML). Subsequently, a software agent is used to convey information granules among the domain experts through a centralized P2P architecture. The communication protocol is as follows:

Input Domain expert E_g, $1 \le g \le G$, provides input data $\tilde{W}_t$ for T periods, where $n \le t \le T+n-1$. In case of computing the FBPN output, the setting vector VS_g is public.

Output Domain expert E_g, $1 \le g \le G$, learns $(D(\tilde{W}_t) - a_t)/a_t$ without anything else, where $D(\tilde{W}_t)$ is computed using the center-of-gravity method:

$$d(\tilde{W}_t) = \frac{W_{t1} + W_{t2} + W_{t3}}{3}.$$

After collaboration, domain expert g refits the corresponding FBPN with two new GP models:

$$\text{Min} \quad \sum_{\text{all } t} \pi_t$$

subject to

$$\ln\left(\frac{1}{W_{t3}} - 1\right) = \theta_3^o - \sum_{\text{all } l} w_{l3}^o h_{l3},$$

$$\sum_{\text{all } l} w_{l3}^o h_{l3} - \theta_3^o = -\ln(1/\pi_t - 1),$$

$$\sum_{\text{all } l} w_{l3}^o h_{l3} - \theta_3^o \le \min_{qq \in t^c(gg)} (-\ln(1/\Psi(qq) - 1)), \quad (1)$$

$$\sum_{\text{all } l} w_{l3}^o h_{l3} - \theta_3^o \le \min_{qq \in t^c(gg)} \left(-\ln\left(\frac{1 - s_R(qq)}{a_t - s_R(qq) W_{t2}} - 1\right)\right), \quad (2)$$

$$\sum_{\text{all } l} w_{l3}^o h_{l3} \ge \theta_3^o - \ln\left(\frac{1}{a_t} - 1\right),$$

$$\sum_{\text{all } k} w_{kl3}^h x_k - \theta_{l3}^h \ge -\ln(1/h_{l3} - 1),$$

$$\sum_{\text{all } k} w_{kl3}^h x_k - \theta_{l3}^h \le -\ln(1/h_{l3} - 1),$$

$$k = 1 \sim K; l = 1 \sim L.$$

and

$$\text{Min} \quad \sum_{\text{all } t} \pi_t$$

subject to

$$\ln\left(\frac{1}{W_{t1}} - 1\right) = \theta_1^o - \sum_{\text{all } l} w_{l1}^o h_{l1},$$

$$\sum_{\text{all } l} w_{l1}^o h_{l1} - \theta_1^o = -\ln(1/\pi_t - 1),$$

$$\sum_{\text{all } l} w_{l1}^o h_{l1} - \theta_1^o \le \min_{qq \in t^c(gg)} (-\ln(1/\Psi(qq) - 1)), \quad (3)$$

$$\sum_{\text{all } l} w_{l1}^o h'_{l1} - \theta_1^o \le \min_{qq \in t^c(gg)} \left(-\ln\left(\frac{1 - s_L(qq)}{a_t - s_L(qq) W_{t2}} - 1\right)\right), \quad (4)$$

$$\sum_{\text{all } l} w_{l1}^o h_{l1} \le \theta_1^o - \ln\left(\frac{1}{a_t} - 1\right),$$

$$\sum_{\text{all } k} w_{kl1}^h x_k - \theta_{l1}^h \ge -\ln(1/h_{l1} - 1),$$

$$\sum_{\text{all } k} w_{kl1}^h x_k - \theta_{l1}^h \le -\ln(1/h_{l1} - 1),$$

$$k = 1 \sim K; l = 1 \sim L.$$

where $VS_g = \{\Psi(qq),\ s_R(qq), s_L(qq)\}$ is the view of domain expert q and so on, according to the nomenclature by Pedrycz [6,8]; $t(qq)$ includes the periods that are satisfactorily forecasted by domain expert q. $t^c(qq)$ is the complement of $t(qq)$, i.e. $t^c(qq) = [1\ T] - t(qq)$; $s_L(qq)$ and $s_R(qq)$ are the satisfaction levels requested by domain expert q. Constraints (1) and (2) force the upper bound of the fuzzy forecast to be less than those by others for a period the domain expert is not satisfied with the forecast. In contrast, in constraints (3) and (4), the lower bound should be greater than those by others for the same period.

4. Aggregation Mechanism

To aggregate the fuzzy WIP level forecasts from G domain experts, a RBF network is used. The RBF network has three layers: input, hidden and output. Inputs to the RBF are the three points of the TFN. For example, if a fuzzy WIP level forecast is (a, b, c), then the corresponding inputs to the RBF are a, 0, b, 1, c, and 0. As there are G FBPNs, the number of inputs to the RBF is $6G$. The reason is straightforward – aggregation results in a convex domain, and each point in it can be expressed with the combination of the corners.

All inputs are normalized and passed directly to the hidden layer without being weighted. The activation function used by the hidden layer is Gaussian:

$$h_i(X_t) = e^{-\sum_{k=1}^{K} (x_{tk} - \hat{x}_{tk})^2 / \sigma_i^2},$$

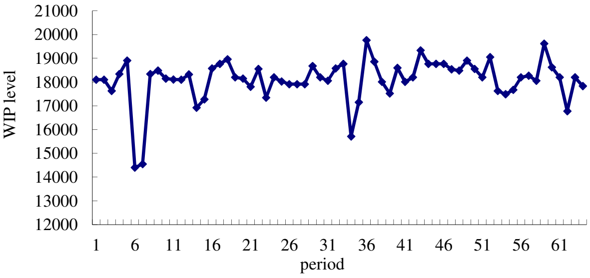

Figure 1: The collected WIP data.

Table 1: The views of the domain experts.

g	$\Psi(g)$	$s_R(g)$	$s_L(g)$
1	3000	0.45	0.25
2	2900	0.45	0.15
3	2500	0.1	0.2

where X_t=[$x_{t1}, \ldots, x_{tK}$] indicates the input vector; $h_i(X_t)$ is the output from the i-th node in the hidden layer, i = 1 ˜ I; $\hat{x}_{tk}$ and σ_i are the center and width of the i-th RBF unit for input variable k, respectively. The activation function used by the output layer is linear:

$$o(X_t) = \sum_{i=1}^{I} w_i h_i(X_t) + w_0.$$

For training the RBF, k-means (KM) is first applied to find out the centers of the RBF units. Then the nearest-neighbour method is used to derive their widths. Finally, the connection weights can be determined by linear regression.

5. A Case Study

The case of forecasting the WIP level in a wafer fabrication factory is used to show the effectiveness of the proposed methodology (see Figure 1). Three domain experts were invited to predict the WIP level in the wafer fabrication factory. However, they had different views for this problem, as shown in Table 1.

For each domain expert, a corresponding FBPN was configured and trained. Every expert used his/her own FBPN to predict the WIP level. The forecasting performances of the three experts before collaboration were shown in Table 2.

Subsequently, they exchanged the views and forecasting results to each other through a centralized P2P network. The central control unit of the centralized P2P network aggregated the fuzzy forecasts using a RBF and then assessed the overall performance (see Table 3).

After receiving such information, the domain experts changed their views and re-train the FBPNs (see Table 4).

Table 2: The forecasting performances of the domain experts before collaboration.

Expert #	RMSE	MAE	MAPE	Average range (Precision)
1	700	573	3.2%	2968
2	674	526	2.9%	2755
3	758	593	3.3%	2464

Table 3: The views of the experts after collaboration.

g	$\Psi(g)$	$s_R(g)$	$s_L(g)$
1	1600	0.45	0.23
2	1800	0.45	0.25
3	1500	0.1	0.2

Table 4: The forecasting performances of the domain experts after collaboration.

Expert #	RMSE	MAE	MAPE	Average range (Precision)
1	669	542	3.0%	2577
2	649	507	2.8%	2526
3	710	557	3.1%	2328

The forecasting performances of the domain experts after collaboration were shown in Table 5.

To compare with the existing approaches, moving average (MA), exponential smoothing (ES), BPN, auto-regressive integrated moving average (ARIMA), and FLR-BPN were also applied to this case. In MA, the number of moving periods was changed from 2 to 10 to determine the best one. The number of inputs to BPN was also set to this value for a fair comparison. A single hidden layer with nodes twice as that of the inputs was configured. In ES, eleven values from 0 to 1 were tried. ARIMA consists of three stages: identification, estimation, and checking. To identify the order in the ARIMA process, the minimum information criterion (MINIC) method was applied. The stationarity in the data was examined by the augmented dickey fuller (ADF) unit root test.

The performances of these approaches were compared in Table 5. The accuracy of WIP level forecasting, measured in terms of MAPE, by the experts, was significantly better than most of the existing approaches by achieving a 43% reduction in MAPE over the comparison basis – MA, which revealed the effectiveness of the FBPN approach. The average advantages over ES, BPN, ARIMA, and FLR-BPN were 43%, 64%, 32%, and 28%, respectively. The accuracy of the WIP level forecasting with respect to MAE or RMSE was also significantly better. Clearly, the FBPN approach provided a good fit for the data collected. In addition, the performance of the WIP level forecasting was indeed improved through expert collaboration.

Table 5: The forecasting performances of the existing approaches.

	RMSE	MAE	MAPE	Average range
MA	963	639	3.7%	5779
ES	934	638	3.7%	5603
BPN	1291	1055	5.9%	7746
ARIMA	853	542	3.1%	5116
FLR-BPN	736	495	2.9%	2714
The proposed methodology	509	365	2.1%	2740

6. Conclusions

The WIP level has the greatest impact for the factory. Data analysis and forecasting in this area is extremely important. There is more and more evidence showing that there is a widespread and long-term trend toward lean production. The WIP level forecasting is considered to be one of the most important tasks to this end. Many evidences also revealed that collaborative intelligence have potential applications in forecasting. In order to effectively predict the WIP level, a fuzzy collaborative forecasting approach is proposed in the present study. In the fuzzy collaborative forecasting approach, a group of experts in the field work collaboratively toward the accurate and precise prediction of the WIP level. The adopted FBPN approach provides the flexibility for these experts to cooperate. Through proper integration, the overall forecasting performance can be maximized. The existing methods lack a formal coordination mechanism. Experts can only modify their views freely, which does not necessarily lead to better prediction performance.

After verifying the effectiveness of the fuzzy collaborative forecasting approach with a case study, the following results were obtained:

(1) The forecasting performances of individual experts were not good enough. Through reference to the forecasts of others, these experts did improve their forecasting performances.
(2) After aggregation, the accuracy and precision of the WIP level forecasting was further improved. It is therefore possible to forecast the WIP level very precisely and accurately using a group of domain experts governed by a centralized P2P network.

More sophisticated learning or collaboration mechanisms can be developed in similar ways in future studies.

References

[1] T. Chen and Y.-C. Wang, Incorporating the FCM-BPN approach with nonlinear programming for internal due date assignment in a wafer fabrication plant, *Robotics and Computer Integrated Manufacturing* **26** (2010), 83-91.

[2] J. N. D. Gupta and J. Ho, Minimizing the total squared workload deviation on parallel machines, *Proceedings of Annual Meeting of the Decision Sciences Institute* (1996).

[3] K. Iriuchijima, H. Sakamoto and M. Fujihara, WIP allocation planning for semiconductor factories, *Proceedings of the IEEE Conference on Decision and Control* (1998), 2716-2721.

[4] Y.-C. Lin, T. Chen and H.-C. Wu, Future WIP level forecasting in a wafer fabrication factory with a hybrid fuzzy and neural approach, *International Journal of Innovative Computing, Information and Control* **7(8)** (2011), 4621-4634.

[5] Y. H. Lin and C.-E. Lee, Total standard WIP estimation method for wafer fabrication, *European Journal of Operational Research* **131 (1)** (2001), 78-94.

[6] V. Loia, W. Pedrycz and S. Senatore, P-FCM: a proximity-based fuzzy clustering for user-centered web applications, *International Journal of Approximate Reasoning* **34** (2003), 121–144.

[7] O. M. Saad, A. H. Amer and E. F. Abdellah, On solving single-objective fuzzy integer linear fractional programs, *Applied Mathematics & Information Sciences* **4 (3)** (2010), 447–457.

[8] W. Pedrycz, Collaborative architectures of fuzzy modeling, *Lecture Notes in Computer Science* **5050** (2008), 117–139.

[9] C. Qi, T. K. Tang and A. I. Sivakumar, Simulation based cause and effect analysis of cycle time and WIP in semiconductor wafer fabrication, *Proceedings of the 2002 Winter Simulation Conference* (2002).

[10] F. Shi and J. Xu, Emotional cellular-based multi-class fuzzy support vector machines on product's KANSEI extraction, *Applied Mathematics & Information Sciences* **6 (1)** (2012), 41-49.

[11] Z. Wang, J. Xu and J. Zhang, Real-time WIP status monitoring model in discrete manufacturing workshop, *Proceedings of Global Congress on Intelligent Systems* (2009), 379-383.

Toly Chen received his Ph.D. in Industrial Engineering from National Tsing Hua University, Hsinchu, Taiwan, R.O.C. in 1996. He is a professor in the Department of Industrial Engineering and Systems Management at Feng Chia University, Taichung, Taiwan, R.O.C. He has been an Industrial Engineer in Vanguard International Semiconductor Corporation for one year. His research interests including fuzzy set modeling, semiconductor manufacturing management, and operations research.

Chi-Wei Lin is an Associate Professor in the Department of Industrial Engineering and Systems Management at Feng Chia University. He received his Ph.D. degree in Industrial Engineering from Purdue University. His research interests are in management of product design and development, optimization of sustainable facility location selection, and optimization of product development project personnel.

Yi-Chi Wang is an Associate Professor in the Department of Industrial Engineering and Systems Management at Feng Chia University. He received his B.S. in Mechanical Engineering from Tatung Institute of Technology of Taiwan, M.S. in Manufacturing Engineering from Syracuse University, New York, and his Ph.D. degree in Industrial Engineering from Mississippi State University in 2003. His research interests are in agent-based manufacturing systems, joint replenishment problems, supply chain system simulation, and optimization of metal cutting conditions. His research has been funded by National Science Council of Taiwan. He has published over 30 articles in refereed journals and conference proceedings. He is a member of SME, IIE, CIIE, and SLEST.