The Estimation of Thin Film Properties by Neural Network

Chi-Yen Shen¹, Yu-Ju Chen², Shuming T. Wang¹, Chuo-Yean Chang³, Rey-Chue Hwang¹, *

¹Electrical Engineering Department, I-Shou University, Kaohsiung City, Taiwan
²Information Management Department, Cheng-Shiu University, Kaohsiung City, Taiwan
³Electrical Engineering Department, Cheng-Shiu University, Kaohsiung City, Taiwan

Email address:
cyshen@isu.edu.tw (Chi-Yen Shen), yjchen@isu.edu.tw (Yu-Ju Chen), smwang@isu.edu.tw (S. T. Wang),
cychang@csu.edu.tw (Chuo-Yean Chang), rchwang@isu.edu.tw (Rey-Chue Hwang)
*Corresponding author

To cite this article:
Chi-Yen Shen, Yu-Ju Chen, Shuming T. Wang, Chuo-Yean Chang, Rey-Chue Hwang. The Estimation of Thin Film Properties by Neural Network. Automation, Control and Intelligent Systems. Vol. 4, No. 2, 2016, pp. 15-20. doi: 10.11648/j.acis.20160402.12

Received: February 22, 2016; Accepted: March 21, 2016; Published: March 25, 2016

Abstract: This paper presents a method based on neural network (NN) for estimating the properties of semiconductor thin film. Through the effective learning process, NN is able to catch the relationship between input and output pairs bypassing the complicated statistical steps such as model hypothesis, identification, estimation of model parameters, and verification. Such an estimator then can be developed to be a smart mechanism which can help the technician to set the relevant control parameters in the manufacturing process of thin film. In this research, the thickness and refractive index (RI) of thin film were estimated by the well learned NN model. From the studied results shown, the properties of thin film indeed could be estimated in advance according to the relevant control parameters in the manufacturing process. That also means the estimator we developed could be built and fulfilled its function.

Keywords: Neural Network, Thin Film, Manufacturing Process

1. Introduction

In recent twenty years, the high-tech skills have been developed widely and vitally in various electronic appliances, such as photoelectric, semiconductor and biomedical chips. The electronic industry has become the focus of economic development for many countries. Besides, due to the fast improvement of manufacturing technique, more and more electronic products are requested to be small and exquisite. Their function is requested to be more powerful either.

It is well-known that thin film is an important and indispensable part for many electronic products. Taking the wafer manufacturing as an example, the filming process plays an important and key role in the wafer front-end manufacturing step. Figure 1 shows the flowchart of wafer manufacturing process [1]. Thus, if the relevant manufacturing parameters of filming process could be set quickly and accurately, not only the efficiency of working machine can be greatly improved, but also the time and frequency of machine test can be reduced effectively.

![Figure 1. The manufacturing process of wafer.](image-url)
Generally, Chemical Vapor Deposition (CVD) is the popular method used in the field of thin film manufacturing process. In CVD process, many complicated and nonlinear chemical and physical reactions are hardly analyzed. The phenomena of particle drift and variation are happened very often in the filming process either. Undoubtedly, these uncontrolled factors will affect the quality of filming process very seriously. Thus, how to set the adequate manufacturing parameters for improving the yield rate and reducing the times of machine test has become the most important work in the thin film manufacturing process. Unfortunately, trial-and-error is still the common method taken by the technicians in many companies. As we know, the number of failed manufacturing process and the defective product could be possibly raised, if the manufacturing parameters were determined by the technician based on personal experience only. Thus, several studies about the optimal thin film manufacturing have been proposed [2-5].

Recently, due to the fast development of artificial intelligent techniques, some studies about the optimal control of film manufacturing process were reported. For instance, the genetic algorithm (GA) had been used for searching the optimal parameters of physical vapor deposition manufacturing control process [4]. Hsieh, Tong, Cui and Hwang et al. proposed the film’s control and estimation by using NN techniques [5-14].

Since the powerful learning and modeling capabilities, NN has been widely used in different applications, such as the signal processing and control [15-19]. Basically, through the well learning, NN could generate an efficient mapping between input and output pairs bypassing the complicated statistical steps. The well-trained NN model then can be used for the specific work.

In this research, an artificial intelligent (AI) system based on NN model for the estimation of film properties is studied. Its aim is that according to the estimation information provided, the junior technician with no full experience is able to make a good setting work for the manufacturing parameters in the filming process. Thus, such an AI system can not only help the technician to do the work of film manufacturing very efficiently and easily, but also reduce the rate of defective products and then save the production cost.

2. Neural Network

NN technique is the main tool used for constructing the estimator of film properties. As previous descriptions, the relationship between input and output pairs is expected to be obtained through the well learning of NN.

The NN structure commonly known as multi-layered feed-forward network is used in this study. The supervised NN with error back-propagation (BP) learning algorithm is taken for NN’s training [15-17]. An example of a three-layered feed-forward NN architecture as shown in Figure 2 is the model of selected topology. Each layer is connected to a layer above it in a feed-forward manner, which means no feed-back from the same layer or a layer above. All connections have a multiplying weight associated with them. Training is equivalent to find the proper weights for all connections such that a desired output is generated for a given input set. Once the neural network is well trained, the proper input information could be inferred in accordance with an expected output. In other words, the useful information can be found for helping the technician to do the well control in the manufacturing process.

![Figure 2. A three-layered feed-forward NN architecture.](image)

In this study, the error back-propagation (BP) learning algorithm is used for NN’s training. The major steps of BP learning rule algorithm is briefly summarized as follows [15-17].

1\(^{st}\) step: Initialize all weights \(\omega_{ij}\) to the small random values firstly.

2\(^{nd}\) step: Present an input pattern with the corresponding desired outputs and then calculate the outputs.

3\(^{rd}\) step: Find the error term for all nodes.

4\(^{th}\) step: Adjust weights by

\[
\omega_{ij}(n+1) = \omega_{ij}(n) + \alpha \delta_j X_i + \zeta (\omega_{ij}(n) - \omega_{ij}(n-1))
\]

where \(n+1\), \(n\) and \(n-1\) are the next, present, and previous iteration numbers, respectively. \(\delta_j\) is the error of node \(j\) and \(X_i\) is the \(i^{th}\) input of node \(j\). \(\alpha\) is the learning rate, the step size in the gradient search algorithm. \(\zeta\) is the momentum and its value is between 0 and 1.

5\(^{th}\) step: Present another input pattern and go back to 2\(^{nd}\) step.

3. Experiments

In our study, the thin film data manufactured by using two recipes, LDRXX and TEOSXX, were collected and simulated. Table 1 and Table 2 present the examples of two collected data sets. The numbers of LDRXX and TEOSXX data sets are 102 and 89, respectively.
be estimated by the well-trained NN model. In order to fairly demonstrate the effectiveness of NN model in the estimation of thin film properties, three same size data sets, i.e. LDR-1a, LDR-1b, and LDR-1c, are randomly reorganized from data LDRXX. Similarly, three same size data sets, TEOS-1a, TEOS-1b and TEOS-1c, are randomly reorganized from data TEOSXX. In the simulations of data LDR-1a, LDR-1b and LDR-1c, 70 sets were used for NN’s training and 32 sets were used for testing. For data TEOS-1a, TEOS-1b and TEOS-1c, 59 sets were used for NN’s training and 30 sets were used for testing.

For all data sets, the size of NN in thickness estimation is 4-10-1. The inputs are DT, TEOS factor, HFRF, LFRF. In RI estimation, the size of NN is 5-10-1. The inputs are DT, TEOS factor, HFRF, LFRF and thickness. The mean absolute error (MAE) and mean absolute percentage error (MAPE) are used as the estimated measurements.

\[
\text{MAE} = \frac{\sum_{i=1}^{N} | y_i - \hat{y}_i |}{N}
\]

\[
\text{MAPE} = \frac{\sum_{i=1}^{N} | \frac{y_i - \hat{y}_i}{y_i} |}{N} \times 100\%
\]

Where, \(y_i\) and \(\hat{y}_i\) are actual and estimated values. \(N\) is the total number of estimation data.

In order to observe the distribution behaviors of all training and test data sets, the simple statistical analysis was done. For example, Table 3 lists the information of mean value, variance and standard deviation for the whole LDR-1a data, LDR-1a 70 training data and 32 LDR-1a test data. Similarly, Table 4 and Table 5 list the statistics of data LDR-1b and LDR-1c, respectively. Table 6 lists the estimation errors of LDRXX data series performed by NN. For TEOSXX data, Table 7, Table 8 and Table 9 list the statistics of data TEOS-1a, TEOS-1b and TEOS-1c, respectively. Table 10 presents the estimation errors of TEOSXX data series performed by NN.

### Table 1. The examples of data manufactured by using recipe LDRXX.

| Manufacturing Parameters | DT | cSHDist | TEOS factor | TEOS (mgm) | HE (sccm) | O₂ (sccm) | HFRF (pre) | LFRF (pre) |
|--------------------------|----|---------|-------------|------------|-----------|-----------|------------|------------|
|                          |    |         |             |            |           |           |            |            |
| 25.5                     | 420| 0.9736  | 800         | 1000       | 1000      | 196       | 50.4       |
| 25.5                     | 420| 0.9736  | 800         | 1000       | 1000      | 196       | 51.4       |
| 26.5                     | 420| 0.9736  | 800         | 1000       | 1000      | 196       | 50.4       |
| 26.5                     | 420| 0.9736  | 800         | 1000       | 1000      | 196       | 51.4       |
| 26.5                     | 420| 1.0453  | 800         | 1000       | 1000      | 194       | 50.4       |
| 26.5                     | 420| 1.0453  | 800         | 1000       | 1000      | 194       | 51.2       |
| 25.5                     | 420| 1.0253  | 800         | 1000       | 1000      | 193       | 50.4       |
| 25.5                     | 420| 1.0253  | 800         | 1000       | 1000      | 164       | 51.2       |
| 27.5                     | 420| 0.9936  | 800         | 1000       | 1000      | 196       | 50.4       |
| 27.5                     | 420| 0.9936  | 800         | 1000       | 1000      | 196       | 50.4       |

**Thin Film Properties**

| Thickness | RI |
|-----------|----|
| 423       | 1.4590 |
| 455       | 1.4619 |
| 443       | 1.4627 |
| 457       | 1.4609 |
| 429       | 1.4624 |
| 419       | 1.4816 |
| 416       | 1.4821 |
| 436       | 1.4820 |
| 443       | 1.4804 |

### Table 2. The examples of data manufactured by using recipe TEOSXX.

| Manufacturing Parameters | DT | cSHDist | TEOS factor | TEOS (mgm) | HE (sccm) | O₂ (sccm) | HFRF (pre) | LFRF (pre) |
|--------------------------|----|---------|-------------|------------|-----------|-----------|------------|------------|
|                          |    |         |             |            |           |           |            |            |
| 38.0                     | 310| 1.0386  | 5250        | 4000       | 4200      | 797       | 209        |
| 38.0                     | 310| 1.0386  | 5250        | 4000       | 4200      | 798       | 209        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 793       | 209        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 795       | 210        |
| 38.0                     | 310| 1.0036  | 5250        | 4000       | 4200      | 763       | 199        |
| 38.0                     | 310| 1.0036  | 5250        | 4000       | 4200      | 765       | 200        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 765       | 200        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 795       | 200        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 780       | 200        |
| 38.0                     | 310| 1.0236  | 5250        | 4000       | 4200      | 765       | 200        |

**Thin Film Properties**

| Thickness | RI |
|-----------|----|
| 4983      | 1.4584 |
| 4988      | 1.4587 |
| 4814      | 1.4628 |
| 4830      | 1.4628 |
| 4885      | 1.4634 |
| 4903      | 1.4633 |
| 4832      | 1.4636 |
| 4836      | 1.4629 |
| 4839      | 1.4632 |
| 4682      | 1.4629 |

The values of thickness and RI of thin film are expected to be estimated by the well-trained NN model. In order to fairly demonstrate the effectiveness of NN model in the estimation of thin film properties, three same size data sets, i.e. LDR-1a, LDR-1b, and LDR-1c, are randomly reorganized from data
### Table 4. The distribution behaviors of data LDR-1b.

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 483.39215 | 1.47009  |
| Var.                | 862.65673 | 0.00068761|
| Std.                | 29.37102  | 0.00829221|

#### Training Data (LDR-1b)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 479.58572 | 1.4696674|
| Var.                | 1126.9996 | 0.00074579|
| Std.                | 33.570816 | 0.00863592|

#### Test Data (LDR-1b)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 491.71875 | 1.4710001|
| Var.                | 197.28158 | 0.00056770|
| Std.                | 14.064906 | 0.00753455|

(Var.: Variance, Std.: Standard Deviation)

### Table 5. The distribution behaviors of data LDR-1c.

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 483.39215 | 1.47009  |
| Var.                | 862.65673 | 0.00068761|
| Std.                | 29.37102  | 0.00829221|

#### Training Data (LDR-1c)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 481.62857 | 1.4718357|
| Var.                | 800.9036  | 0.000731506|
| Std.                | 28.300241 | 0.00855281|

#### Test Data (LDR-1c)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 487.25    | 1.4662563|
| Var.                | 1005.5484 | 0.000391548|
| Std.                | 31.710382 | 0.00625738|

(Var.: Variance, Std.: Standard Deviation)

### Table 6. The statistics of estimations for data LDRXX.

| Statistics | LDR-1a | LDR-1b | LDR-1c |
|------------|--------|--------|--------|
| MAE        | 6.3079 | 6.2145 | 6.8258 |
| MAPE       | 1.3182%| 1.3011%| 1.4255%|

#### LDR-1a

| Statistics | Training | Test | Training | Test |
|------------|----------|------|----------|------|
| MAE        | 0.001173 | 0.00001173 | 0.00001173 | 0.00001173 |
| MAPE       | 0.0799%  | 0.0751% | 0.0832%  | 0.2275% |

#### LDR-1b

| Statistics | Training | Test | Training | Test |
|------------|----------|------|----------|------|
| MAE        | 0.001089 | 0.00001089 | 0.001208 | 0.003345 |
| MAPE       | 0.0799%  | 0.0751% | 0.0832%  | 0.2275% |

#### LDR-1c

| Statistics | Training | Test | Training | Test |
|------------|----------|------|----------|------|
| MAE        | 0.001089 | 0.00001089 | 0.001208 | 0.003345 |
| MAPE       | 0.0799%  | 0.0751% | 0.0832%  | 0.2275% |

(Var.: Variance, Std.: Standard Deviation)

### Table 7. The distribution behaviors of data TEOS-1a.

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4967.382  | 1.4596821|
| Var.                | 6752.8295 | 0.00000177311|
| Std.                | 82.175605 | 0.00133158|

#### Training Data (TEOS-1a)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4942.0    | 1.4597865|
| Var.                | 8020.3105 | 0.00000241364|
| Std.                | 89.55618  | 0.00155359|

#### Test Data (TEOS-1a)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 5017.3    | 1.4594768|
| Var.                | 562.2862  | 0.00000048738|
| Std.                | 23.71257  | 0.00698123|

(Var.: Variance, Std.: Standard Deviation)

### Table 8. The distribution behaviors of data TEOS-1b.

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4967.382  | 1.4596821|
| Var.                | 6752.8295 | 0.00000177311|
| Std.                | 82.175605 | 0.00133158|

#### Training Data (TEOS-1b)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4945.085  | 1.459999 |
| Var.                | 7853.424  | 0.00000210681|
| Std.                | 88.619544 | 0.00145149|

#### Test Data (TEOS-1b)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 5011.2333 | 1.4590766|
| Var.                | 1783.7709 | 0.000000594971|
| Std.                | 42.23471  | 0.00077134|

(Var.: Variance, Std.: Standard Deviation)

### Table 9. The distribution behaviors of data TEOS-1c.

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4967.382  | 1.4596821|
| Var.                | 6752.8295 | 0.00000177311|
| Std.                | 82.175605 | 0.00133158|

#### Training Data (TEOS-1c)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4983.356  | 1.4593307|
| Var.                | 4049.8537 | 0.000001109008|
| Std.                | 63.638462 | 0.00105313|

#### Test Data (TEOS-1c)

| Statistics          | Thickness | RI       |
|---------------------|-----------|----------|
| Mean                | 4935.967  | 1.4603734|
| Var.                | 10851.551 | 0.00000241653|
| Std.                | 104.17078 | 0.00155452|

(Var.: Variance, Std.: Standard Deviation)
The examples of superposition plot for thickness estimation are shown in Figure 3 and Figure 4. Figure 3 shows the example of NN’s training result and Figure 4 shows the example of NN’s testing results. In figures, the solid line stands the actual thickness values and the dotted line stands NN’s estimated values.

Similarly, the examples of superposition plot for RI estimation are shown in Figure 5 and Figure 6. Figure 5 shows the example of NN’s training result and Figure 6 shows the example of NN’s testing result. Same as above figures, the solid line stands the actual thickness values and the dotted line stands NN’s estimated values.

4. Results & Discussion

It is known that many unknown factors will affect the properties of thin film in its real manufacturing process. These unknown factors are usually uncontrolled and can be treated as the disturbances. Besides, different manufacturing machines have different physical characteristics. All these conditions might make the films have different properties even they are filmed under the same manufacturing parameters. In our study, we tried to use NN technique to catch the relationships among the film’s properties and the relevant manufacturing parameters so that the film’s properties could be estimated in advance. From the simulation results shown, the relationships between the film’s properties and the manufacturing parameters indeed can be obtained by a well-trained NN model. For both thickness and RI estimations, the plots show that the trends of film’s properties still can be estimated by NN.

5. Conclusion

In this research, the estimation for the properties of semiconductor’s thin film based on NN technique was studied. From the study results shown, we conclude that NN model indeed has the ability to estimate the properties of thin film if NN was well-trained. In other words, these well-trained NN estimators could provide the important information to the technician for setting the proper manufacturing parameters in the filming process. The
A technician is able to make the whole filming process more effective and successful. However, in this research, only a few manufacturing parameters were collected. We do believe that the estimation accuracy for the film’s property could be improved greatly if more relevant manufacturing parameters can be considered and collected.

Acknowledgements

This research was supported by the Ministry of Science and Technology, Taiwan, ROC under the contract No. MOST-103-2221-E-214-050.

References

[1] Q. Michael, & S. Julian, Semiconductor Manufacturing Technology, Prentice Hall, 2000.

[2] C. H. Lai, “Fault diagnosis on CVD equipment via neural network approach,” Master Thesis, National Cheng Kung University, 2007, (In Chinese).

[3] E. Ritter, “Deposition of Oxide film by reactive evaporation”, Journal of Vacuum Science and Technology, vol. 3, issue 4, pp. 225, 1966.

[4] C. H. Li, “Application of genetic algorithm optimization on the physical parameters of the meteorological forecast deposition of thin film semiconductor manufacturing process,” Master Thesis, Southern Taiwan University of Science and Technology, 2009, (In Chinese).

[5] K. L. Hsieh, and L. L. Tong, “Optimization of multiple quality responses involving qualitative and quantitative characteristics in IC manufacturing using neural networks”, Computers in Industry, vol. pp. 46, 1-12, 2001.

[6] W. Z. Cui, C. C. Zhu, and H. P. Zhou, “Prediction of thin film thickness of field emission using wavelet neural networks”, Thin Solid Films, vol. 473, pp. 224-229, 2005.

[7] W. Z. Cui, C. C. Zhu, H. P. Zhao, “Prediction of thin film thickness of field emission using wavelet neural networks,” Thin Solid Films, vol. 473, no. 2, pp. 224-229, 2005.

[8] B. Kim, H. Lee, D. Kim, “Modeling of thin film process data using a genetic algorithm-optimized initial weight of backpropagation neural network,” Applied Artificial Intelligence, vol. 23, no. 2, pp. 168-178, 2009.

[9] C. C. Huang, H. C. Huang, Y. J. Chen, R. C. Hwang, “An AI system for the decision to control parameters of TP film printing”, Expert Systems With Applications, vol. 36, no. 5, pp. 9580-9583, 2009.07.

[10] P. H. Weng, Y. J. Chen, S. M. T. Wang, R. C. Hwang, “The predictions of optoelectronic attributes of LED by neural network”, Expert Systems With Applications, vol. 37, no. 9, pp. 6282–6286, 2010.09.

[11] Y. J. Chen, J. C. Chien, C. Y. Chang, S. M. T. Wang, Y. C. Chiang, R. C. Hwang, “AI transmittance estimator for multi-layer coating TP film”, Applied Mechanics and Materials, vol. 312, pp. 996-1000, 2013.03.

[12] Y. D. Ko, P. Moon, C. E. Kim, M. H. Ham, M. K. Jeong, G. D. Alberto, J. M. Myoung, I. Yun, “Predictive modeling and analysis of HfO2 thin film process based on Bayesian information criterion using PCA-based neural networks,” Surface and Interface Analysis, vol. 45, no. 9, pp. 1334-1339, 2013.

[13] N. M. Sabri, N. D. Md Sin, M. Putech, M. Rusop Mahmood, “Prediction of nanostructured ZnO thin film properties based on neural network,” Advanced Materials Research, vol. 832, pp. 266-269, 2014.

[14] L. Cai, Y. G. Tan, Q. Wei, “On-line thickness measurement of thin film based on neural network,” Applied Mechanics and Materials, vol. 484-485, pp. 307-310, 2014.

[15] A. Khotanzad, R. C. Hwang, A. Abaye, D. Maratukulam, “An adaptive modular artificial neural network: Hourly load forecaster and its implementation at electric utilities,” IEEE Transactions on Power Systems, vol. 10, pp. 1716-1722, 1995.

[16] C. Y. Shen, C. L. Hsu, R. C. Hwang, J. S. Jeng, “The interference of humidity on a shear horizontal surface acoustic wave ammonia sensor”, Sensors & Actuators: B. Chemical, vol. 122, pp. 457-460, 2007.

[17] P. H. Weng, Y. J. Chen, H. C. Huang, R. C. Hwang, “Power load forecasting by neural models”, Engineering Intelligent Systems for Electrical Engineering and Communications, vol. 15, pp. 33-39, 2007.

[18] S. Malinov, W. Sha, J. J. McKeown, “Modelling the correlation between processing parameters and properties in Titanium alloys using artificial neural network,” Computational Materials Science, vol. 21 pp. 375–394, 2001.

[19] M. Toparl, S. Sahin, E. Ozkaya, S. Sasaki, “Residual thermal stress analysis in cylindrical steel bars using finite element method and artificial neural networks”, Computers and Structures, vol. 80, No. 23, pp. 1763-1770, 2002.