Online Supplement for “Monitoring wafer geometric quality using additive gaussian process model”

Linmiao Zhang[†], Kaibo Wang[‡], and Nan Chen^{*†}

[†]Department of Industrial and Systems Engineering, National University of Singapore

[‡]Department of Industrial Engineering, Tsinghua University

March 1, 2015

S.1 Comparing AGP, Exact, and NGP

In Section 4.2.2 we compare AGP with Exact case and NGP model in terms of detecting changes when model parameters are unknown. In the comparison, both AGP and NGP are estimated with different in-control sample sizes:

1. $N_0 = 20, n_0 = 20$,
2. $N_0 = 20, n_0 = 10$,
3. $N_0 = 10, n_0 = 20$,
4. $N_0 = 10, n_0 = 10$.

where N_0 refers to the number of in-control curves, and n_0 refers to the number of measurements on each curve. The comparison results are summarized in Figure [S.1](#).

Figure [S.1a](#) is the one included in the article. Figure [S.1b](#), [S.1c](#) and [S.1d](#) show the comparison results using other sample sizes. We found the results are consistent, and led to the similar findings: overall, AGP performs better than NGP.

*Corresponding author: isecn@nus.edu.sg

Figure S.1: Performance comparison between AGP, Exact and NGP using different sample size

S.2 Justification of Model Assumption

In the AGP model assumption(Eq.1), the residuals ϵ_i are assumed to be independent across different wafers and have identical covariance function form. The justifications of these assumptions are summarized below.

From an engineering perspective, the residuals represent the deviations from the standard profile because of various sources of disturbance. When the production process is stable, the deviations of one wafer may be different from that of another, but the spatial patterns are expected to be similar. In addition, if the disturbance is not persistent and random in nature, it is common to assume that their effects (deviations) are independent. Similar assumptions have been made in related literatures (Plumlee, Jin, Roshan Joseph, and Shi 2013; Jin, Chang, and Shi 2012), and other regression applications when normally distributed *i.i.d* errors are assumed.

On the other hand, different from the regression analysis where ϵ_i is a single number, the residuals here ϵ_i are in fact surfaces. As a consequence, it is difficult to derive rigorous statistical

methods to test their independence and covariance structures. While we would explore this in our future research, here we provide some indirect evidence to validate the assumptions. In the case study, we first estimate the AGP model with measurements from eight IC wafers. Based on the estimated standard profile, the residual surface from each wafer can be obtained. The spatial covariance structures of these eight residual surfaces can be compared through their sample variograms (Cressie 1993). Figure S.2 compares the eight sample variograms and the theoretical

Figure S.2: Sample and model variograms of eight residual profiles

variogram implied by AGP model estimation. It demonstrates that the spatial patterns of the eight residual surfaces are similar, and consistent with the model assumption. The discrepancies, especially when the “distance” is large, are largely caused by limited data in computing sample variogram. In addition, we also plot the residual surfaces in Figure S.3. It further demonstrates that the spatial patterns of each residual surfaces are similar to each other, and no apparent correlations among different surfaces can be observed (indicating independence). In a more quantitative way, we compute the mean squared prediction errors (MSPE) of each residual. They are comparable as shown in Table S.1. These analysis collectively show that the residuals can be

Table S.1: Mean squared prediction errors of predicted residual profiles

Wafer Index	#1	#2	#3	#4	#5	#6	#7	#8
MSE ($\times 10^{-4}$)	7.96	6.89	10.09	8.41	9.27	9.51	8.46	9.56

viewed as independently and identically distributed. In future, we will further study how to test

Figure S.3: Predicted deviation profiles of eight sample wafers

these assumptions in the context of Gaussian process more rigorously.

S.3 Validation of Testing Results

In Section 5 we demonstrate how to apply the proposed method in application. All the data we used in the case study are provided by a real semiconductor fabrication plant. There are only 38 wafers available. In addition, engineers do not have a rigorous classification of the IC samples and OC samples. As a result, it is impossible to evaluate the α and β error based on the real data. In fact, similar problems are encountered in many studies based on the real data. Without necessary replications, comparison results might hold on one set of data but fail on another, simply because of the data randomness instead of performance difference. As a result, we rely heavily on simulations to comprehensively compare the performance of different methods. In this sense, Section 5 focuses more on demonstrating its use in real applications, rather than proving its effectiveness, which has been done through simulation in Section 4. We hope that our proposed method can provide a more reliable quality check for industry and a benchmark for future research development.

Numerically, we also attempt to evaluate our result through comparing with NGP model using real data. Since we cannot evaluate the real α error, we set the test limits based on the χ^2 distribution and the nominal α error. The results are shown in Figure S.4.

Figure S.4 shows that almost every wafer fails GLR test if NGP model is used. T^2 test based on NGP also has more “out-of-control” signals. This result is misleading, and is different from

(a) Control charts based on AGP (our original chart)

(b) Control charts based on NGP

Figure S.4: Monitoring results using different models

empirical analysis (note there is no label on each wafer whether it is IC or OC). To demonstrate this, we randomly pick Wafer #1 and #17, which are rejected by tests based on NGP but not rejected by tests based on AGP. We plot their profiles together with an in-control sample in Figure S.5. Different from those demonstrated in Appendix A-II, Wafer #1 and #17 are not significantly different from in-control sample. Therefore it is likely they are mis-classified as “out-of-control”. This finding, together with simulation results in Section 4, demonstrate that AGP model is more suitable than NGP in the application.

References

- Cressie, N. (1993), *Statistics for Spatial Data, revised edition*, vol. 928, Wiley, New York.
- Jin, R., Chang, C., and Shi, J. (2012), “Sequential measurement strategy for wafer geometric profile estimation,” *IIE Transactions*, 44, 1–12.
- Plumlee, M., Jin, R., Roshan Joseph, V., and Shi, J. (2013), “Gaussian process modeling for engineered surfaces with applications to Si wafer production,” *Stat*, 2, 159–170.

(a) Wafer #1

(b) Wafer #17

(c) In-control sample

Figure S.5: Comparison between two “failed” wafers and an in-control sample