PHM DATA CHALLENGE 2016

The PHM Data Challenge is a competition open to all potential conference attendees. This year the challenge is focused on tracking the health state of components within a wafer chemical-mechanical planarization (polishing) system. Participants will be scored based on their ability to predict average removal rate of material during polishing at particular tool settings and as the performance of the tool degrades over time.

This is a fully open competition in which collaboration is encouraged. The teams may be composed of any combination of students, researchers, and industry professionals. The results will be evaluated by the Data Challenge Committee and all teams will be ranked. The top three scoring teams will be invited to present at a special session of the conference and will be recognized at the Conference banquet event.

Data Challenge Chairs
Nicholas Propes, Seagate Technology, nicholas.c.propes@seagate.com
Justinian Rosca, Siemens Corporate Technology, justinian.rosca@siemens.com

Teams
Collaboration is encouraged and teams may be comprised of one of more students and professionals. The team judged to have the first, second, and third best scores will be awarded prizes of $600, $400, and $200 respectively contingent upon:

- Having at least one member of the team attend the PHM 2016 Conference
- Presenting the analysis results and technique employed at a special session within the Conference program
- Submitting a peer-reviewed conference paper (Submission of the challenge special session papers is outside the regular paper submission process and follows its own modified schedule.)
- The top entries will also be encouraged to submit a journal-quality paper to the International Journal of Prognostics and health Management (ijPHM).
- The organizers of the competition reserve the right to both modify these rules and disqualify any team for any practices it deems inconsistent with fair and open practices.

Registration
Teams may register by contacting the Competition organizers (justinian.rosca@siemens.com and nicholas.c.propes@seagate.com) with their name(s) and a team alias under which the scores would be posted. Please note: In the spirit of fair competition, we allow only one account per team. Please do not register multiple times under different user names, under fictitious names, or using anonymous accounts. Competition organizers reserve the right to delete multiple entries from the same person (or team) and/or to disqualify those who are trying to "game" the system or using fictitious identities.

Annual Conference of the Prognostics and Health Management Society – PHM Data Challenge 2016

Key Conference Dates

Competition Open   May 15, 2016
Final Validation Set Posted   Aug 22, 2016
Competition Closed   Sept. 8, noon PST, 2016
Preliminary Winners Announced   Sept. 12, 2016
Winners Announced   Sept. 29, 2016
Winning Papers Due   Sept. 30, 2016
PHM Conference Dates   Oct. 2-6, 2016

System and Data Description

This year's challenge is focused on the combination of physics-based modeling and statistical approaches for prediction. It is not required that the solution you select use a physics-based modeling approach. However, additional points will be given to those approaches that provide some physical connection to the data such as health states of various components, relationship between data and model parameters / states, etc.

The system under investigation is a wafer Chemical-Mechanical Planarization (CMP) tool that removes material from the surface of the wafer through a polishing process. Figure 1 depicts the CMP process components and operation.   The CMP tool is composed of the following components:

- a rotating table used to hold a polishing pad
- a replaceable polishing pad which is attached to the table
- a translating and rotating wafer carrier used to hold the wafer
- a slurry dispenser
- a translating and rotating dresser used to condition a polishing pad.

A wafer is placed on the underside of a wafer carrier in the CMP tool, the CMP tool recipe is set (e.g. set-points for speeds, forces, polish time, etc.), and the polishing process is started.   During the polishing process, the wafer is pressed against a polishing pad and both the wafer / wafer carrier and polishing pad / table are rotated in the same direction. A slurry composed of abrasive materials and chemicals are dispensed onto the pad during the polishing process.   After polishing is completed, the polishing pad may be conditioned to improve its polishing properties by using a dresser.   The dresser is typically composed of a hard material such as diamond that is pressed across the pad to roughen the pad's surface to prepare it for future polishing operations.
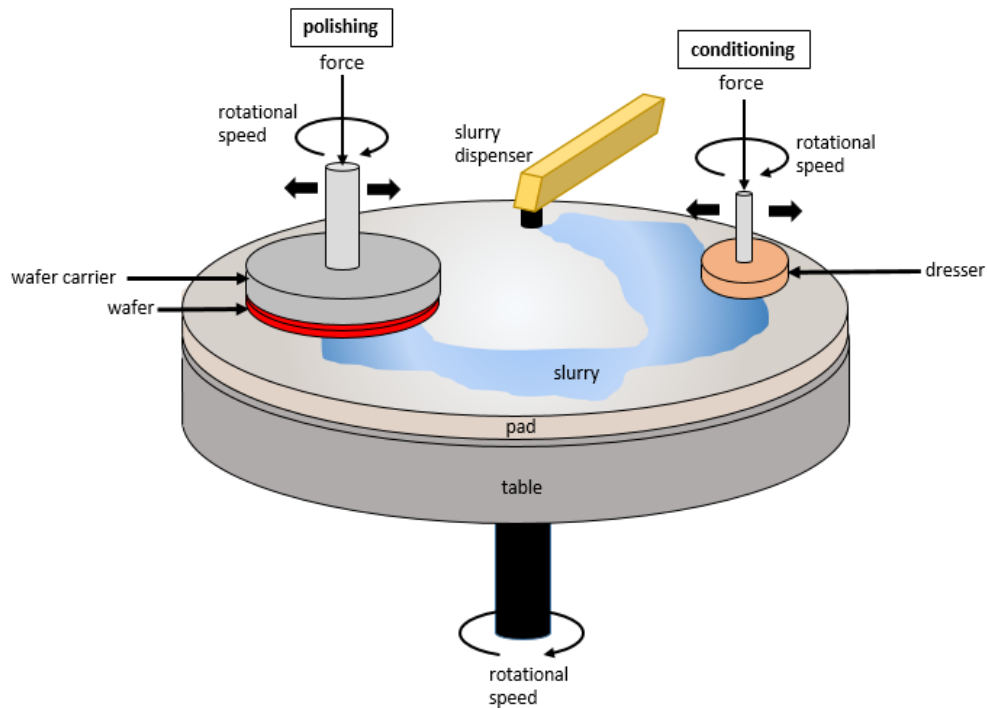
Figure 1. Chemical Mechanical Planarization (Polishing) of wafer. This process removes material from wafer surface.

During the polishing process, the polishing pad's ability to remove material is diminished. Over time, the polishing pad has to be replaced with a new pad.   Similarly, the dresser's capability to roughen the polishing pads is also reduced after successive conditioning operations and after a while the dresser must be replaced.

Objective
The primary objective of this challenge is to predict polishing removal rate of material from a wafer using physics-based modeling methods and the data provided.   The condition of the polishing pad and dresser change over time as they are being used. If these states can be estimated, then polishing time estimates can possibly be improved.

Data Description
Training and test data sets are provided to you to establish your methods. The training data represents data collected during various runs of the CMP tool for specified wafers over time. Data is given in the Table 1 format described below. Each row of the data represents an instance of all measurement variables at any given time. An average rate of material removal from a wafer is given separately in Table 2, which has a corresponding wafer identification number and stage. The average rate of removal was determined from measurements of the thickness of the material before and after CMP polishing.

Table 1: Time Series Data Description

Column Symbol	Column Name	Description
x1	MACHINE_ID	Numeric ID of machine
x2	MACHINE_DATA	Numeric ID of wafer ring location in machine
x3	TIMESTAMP	Seconds
x4	WAFER_ID	Number representing ID of wafer
x5	STAGE	A or B representing a different type of processing stage
x6	CHAMBER	Chamber in machine for wafer processing
x7	USAGE_OF_BACKING_FILM	A usage measure of polish-pad backing film
x8	USAGE_OF_DRESSER	A usage measure of dresser
x9	USAGE_OF_POLISHING_TABLE	A usage measure of polishing table
x10	USAGE_OF_DRESSER_TABLE	A usage measure of dresser table
x11	PRESSURIZED_CHAMBER_PRESSURE	Chamber pressure
x12	MAIN_OUTER_AIR_BAG_PRESSURE	Pressure related to wafer placement
x13	CENTER_AIR_BAG_PRESSURE	Pressure related to wafer placement
x14	RETAINER_RING_PRESSURE	Pressure related to wafer placement
x15	RIPPLE_AIR_BAG_PRESSURE	Pressure related to wafer placement
x16	USAGE_OF_MEMBRANE	A usage measure of polishing membrane
x17	USAGE_OF_PRESSURIZED_SHEET	A usage measure of wafer carrier flexible sheet
x18	SLURRY_FLOW_LINE_A	Flow rate of slurry type A
x19	SLURRY_FLOW_LINE_B	Flow rate of slurry type B
x20	SLURRY_FLOW_LINE_C	Flow rate of slurry type C
x21	WAFER_ROTATION	Rotation rate of wafer
x22	STAGE_ROTATION	Rotation rate of stage
x23	HEAD_ROTATION	Rotation rate of head
x24	DRESSING_WATER_STATUS	Status of dressing water
x25	EDGE_AIR_BAG_PRESSURE	Pressure of bag on edge of wafer

Table 2: Average Material Removal Rates

Column Symbol	Column Name	Description
x4	WAFER_ID	Number representing ID of wafer
x5	STAGE	A or B representing a different type of processing stage
y	AVG_REMOVAL_RATE	The average rate of material removal (the measure to predict)

Annual Conference of the Prognostics and Health Management Society – PHM Data Challenge 2016

Note: Each column of data in Table 1 has been scaled with hidden values to protect proprietary information.

Training data is given in a collection of files *"CMP-training-ddd.csv"* representing instances for all 25 columns (*x1,…,x25*) described in Table 1, plus a removal rate file *"CMP-training-removalrate.csv"* described in Table 2. Test data is given in a collection of files *"CMP-test-ddd.csv"* representing the 25 variables (*x1,…,x25*). Participants need to predict the missing values of the *AVG_REMOVAL_RATE* (*y*) for each wafer identifier and stage. A correct submission will be given by a zip archive [.zip] containing the predicted removal rate file, in the same two column format that was given for the training data, with the *WAFER_ID,* the *STAGE,* and the *AVG_REMOVAL_RATE* (*x4, x5,* and *y*) representing the prediction of average *AVG_REMOVAL_RATE* (*y*) for each *WAFER_ID* (*x4*) and *STAGE* (*x5*) in the test data. The submission file name should be the team alias, e.g. "team_alias.zip" and it should contain *"CMP-test-removalrate.csv"*.

The competition training data and the test data is available at:
https://www.dropbox.com/s/1ihr7wkmrj7nfyw/2016%20PHM%20DATA%20CHALLENGE%20CMP%20DATA%20SET.zip?dl=0

Questions about the data challenge should be emailed directly to the organizers (nicholas.propes@seagate.com and justinian.rosca@siemens.com). A summary of the question asked and the answer will be posted for all participants on the PHM society Data Challenge general discussion web page: https://www.phmsociety.org/forum/5182

Acknowledgement: We would like to thank Steve Mossey from Savigent for support for extracting the initially very large data files.

Scoring
During active competition, scoring will be calculated using mean squared error (MSE) accuracy. Only one submission will be accepted per team per week (on Monday by Noon PST deadline). Please send all submissions to Nicholas Propes (nicholas.c.propes@seagate.com).

The final submission is to include a 1-page description of the physics-based modeling method utilized (if any) and the final predicted average removal rates on a validation dataset to be posted several weeks before the competition closes.

After the competition is closed, the final score will be calculated for all submissions based on the MSE accuracy (90% weight) and the physics-based modeling approach (10% weight).   The physics-based modeling approach will be judged on the following criteria:
- Estimation of dresser condition and effect on removal rate of polishing pad (3%)
- Estimation of polishing pad condition and effect on removal rate of polishing pad (3%)
- Effect of other parameters on the removal rate of polishing pad (4%)

References
1. https://en.wikipedia.org/wiki/Chemical-mechanical_planarization
2. https://www.crystec.com/alpovere.htm

3. http://www-mtl.mit.edu/researchgroups/Metrology/PAPERS/Thesis_Xiaolin_final.pdf
4. http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1189&context=etd
5. http://escholarship.org/uc/item/4ct2n4jh
6. http://homepages.rpi.edu/~schwed/Research/sundararajan_jecs_cmp.pdf
7. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.8.6597&rep=rep1&type=pdf