Mask cleaning process evaluation and modeling

Pavel Nesladek[a], Steve Osborne[b],

[a] Advanced Mask Technology Center, Raehnitzer Allee 9, 01109 Dresden, Germany, pavel.nesladek@amtc-dresden.com
[b] Sigmameltec Ltd, 3-37-7 Shimoasao, Asao-ku Kawasaki, 215-0022 Japan, sosborne@osbornemgt.com

ABSTRACT

Large error bars in cleaning experiments are commonly accepted in mask making but such errors restrict potential improvements in cleaning and restrict the uniform delivery of megasonic (MS) energy. Hence, large error limits in particle removals have an impact to operational costs based on contamination and breakage. New data handling methods are developed here, which exceed the current capability scatterometric particle measurement methods and which create a better statistical basis for interpretation. These improved data treatment methods employ subdivisions of the mask into regions as small as mm^2. The effective number of runs becomes many thousands of time greater which can compensate for the small number of blanks available for tests due to restricted costs. This new technology is combined with a precise modeling of the MS tracking patterns on a plate and allows better comparisons between theoretical modeling and experimentally observed cleans. The combination of these two methods yields an improved determination of rate kinetics for particle removal. Collectively, these methods provide the basis for better interpretation of the spatial non-uniformities seen in MS spin cleaning methods with obvious consequences to manufacturing costs.

Keywords: Clean process evaluation, Particle removal efficiency, Process modeling, kinetics of cleaning

INTRODUCTION

The main focus of this work is the kinetics of the cleaning process. We compare the theoretical distribution of the direct megasonic beam to the experimentally observed regional cleaning efficiency. To accomplish this, we made a detailed modeling of the trace of the MS head over the mask. We deploy MS power via a sweeping arm over a spinning plate. The distribution model for the megasonic beam was established using the geometrical parameters of the chamber such as arm length, MS beam to mask center distance and kinetics parameters *e.g.* arm swing speed or chuck speed. The modeled path of the MS beam can be matched to traces obtained experimentally using just a few arm swings. These initial steps provide the missing parameters of MS beam width and shape the spatial distribution.

We investigate our megasonic processes only using Si_3N_4 particles. The particle size distribution and formulas used for determining particle removal efficiency (PRE) are discussed in our previous paper [1]. Experimentally obtained spatial PRE distributions exhibit big removal efficiencies close to center of the mask and decrease gradually towards mask edge, usually with a small peak at the radius of reversal point. This matches the MS power distribution model, but differs in detail, as will be discussed later on. For estimation of cleaning kinetics, sequentially cleaned Cr blanks were used. For the analysis of kinetics data, dedicated R script based programs were developed.

As mentioned, we deploy MS power via a sweeping arm over spinning plate. However, there are alternative ways of applying a uniform cleaning force to the mask. *e.g.:* a 'Skirt' shaped nozzle can extend the surface area covered under the MS power head—this avoids a downward directed water stream. As well, one may couple the MS power from the backside of the mask. Other methods include spin cleaning with a stationary megasonic rod, spin cleaning with a clamped mask and an extremely large Piezo surface. High velocity spray is also garnering interest in mask houses as a way to avoid the use of megasonic. Nevertheless, the methods developed here may be interesting and useful to the cleaning community as a whole.

26th European Mask and Lithography Conference, edited by Uwe F.W. Behringer, Wilhelm Maurer,
Proc. of SPIE Vol. 7545, 75450J · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.864335

EXPERIMENTAL

Spin tool model

For simplicity, the movement of MS arm and mask rotation are modeled in polar coordinates and presented as movements of the MS beam center over a standing mask. In this way, the mask rotation can be superimposed to the swing movement of the arm—creating polar diagrams of the MS beam path. Fig.**1a** describes the parameters varied. We assume constant arm speed (true for most tools available today). Fig.**1b** shows the arm angle function. From the arm swing speed ω_a, the center offset $\mathbf{d_c}$ and arm length $\mathbf{r_a}$, the radius **r** can be calculated using Eq.1. This is the radial component of MS beam trace **r** which is shown as a function of time in Fig.**1c**. We assume that the arm stops at the reversal point for a short time $\mathbf{t_d}$, to change arm direction and speed up in the opposite direction. This delay is visible in functions shown in both Fig.**1b** and **1c**.

$$r = \sqrt{r_a^2 + (r_a + r_d)^2 - 2r_a(r_a + r_d)\cos(\varphi)} \qquad [1]$$

The angular position of the beam is a bit more complicated, since all movements are contributing: chuck speed, initial position of the beam, arm movement. The axial position of the beam can be described by the following equation:

$$\gamma = \gamma_{ax} + \gamma_a + \gamma_d + \gamma_0 \qquad [2]$$

Here, γ_{ax} is the angle between $\varphi = 0$ and the positive x axis located at the mask center.
γ_a is correction in angle due to nonlinearity of arm movement,
γ_d is angle origination from chuck rotation which is the main axial component,
γ_0 is the angular position of the beam at process beginning measured from positive x-axis.

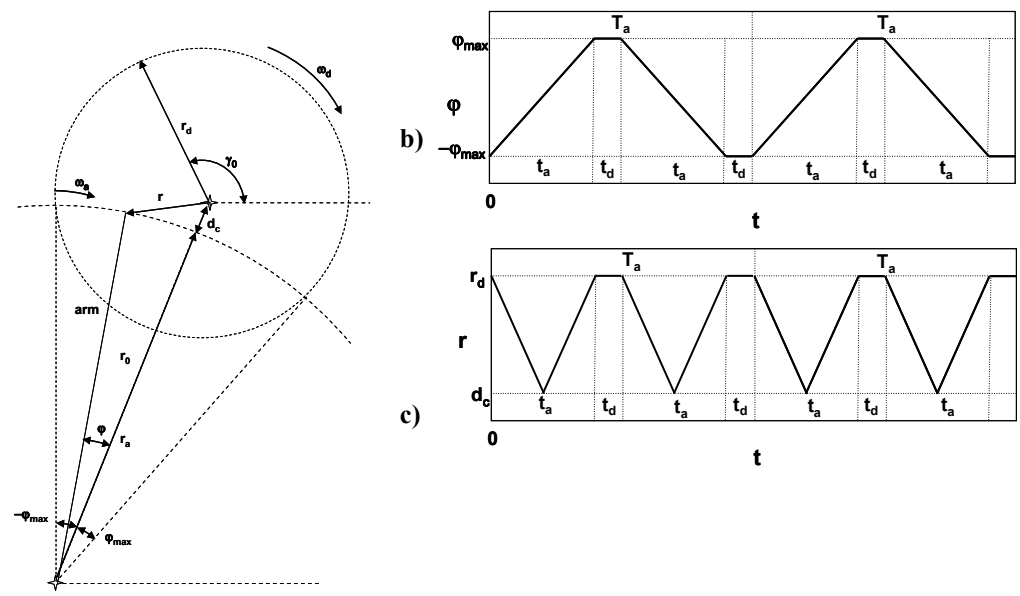


Fig.1a) Parameters influencing the MS beam trace at a spin cleaning tool**:** $\mathbf{r_a}$- arm length (more precisely the distance between the arm axis and the impact point of the media beam), $\mathbf{r_d}$- radius of covered area which may be represented by maximum angle measured from the center of the mask φ_{max} and $-\varphi_{max}$, $\mathbf{d_c}$- center offset, γ_0 - angular starting position, ω_d- mask rotation speed (chuck speed), ω_a- arm swing speed. **1b**) Displays arm angle as function of time. Reference position is the arm pointing towards mask center. The function consists of periods when arm is swinging across the mask within time $\mathbf{t_a}$, as derived from the arm speed ω_a , and reversal point delay $\mathbf{t_d}$ we observed at all tools so far. **1c)** Describes the arm angle radius **r** as function of time and exhibits the delay $\mathbf{t_d}$ at its maximum value

The components of angular movement of the beam are defined by the following set of equations:

$$\gamma_d = \omega_d \cdot t \qquad [3]$$

$$\gamma_a = \arccos\left(\frac{(r_a + d_c)^2 + r^2 - r_a^2}{2\cdot(r_a + d_c)\cdot r}\right)\cdot\sin(\varphi) \qquad [4]$$

$$\gamma_{ax} = -\gamma_{a,(t=0)}\cdot\varphi_{\max} \qquad [5]$$

Using the equations listed, we may predict the trace of the MS beam at varying process conditions. Figure2a & 2b show two spatial distributions which vary in only the arm swing speed. These pictures illustrate the significant variation in spatial pattern when both the chuck speed and arm swing speed ratio equals a small number. The beam follows a Lissajous pattern.

Using broad enough cleaning beam reduces to some degree the non uniformity of the surface coverage, however, for megasonic processes a rather narrow beam of few mm is typical. At such conditions the spatial density of distribution plays substantive role. In our model, we may further assume a Gaussian power distribution and vary the standard deviation of the beam width. We may calculate spatial distribution of the power distribution in this case. Figures 2c and 2d show the spatial distributions for previously mentioned process conditions. This pattern represents the expected distribution of the particle removal efficiencies (PRE), thus slight variations of chuck speed or arm swing speed leads to strong variation in process efficiency. The appropriate combination of both parameters is a crucial point in process optimization. Additionally the reversal point delay, $\mathbf{t_d}$, impacts the path of MS beam over the mask.

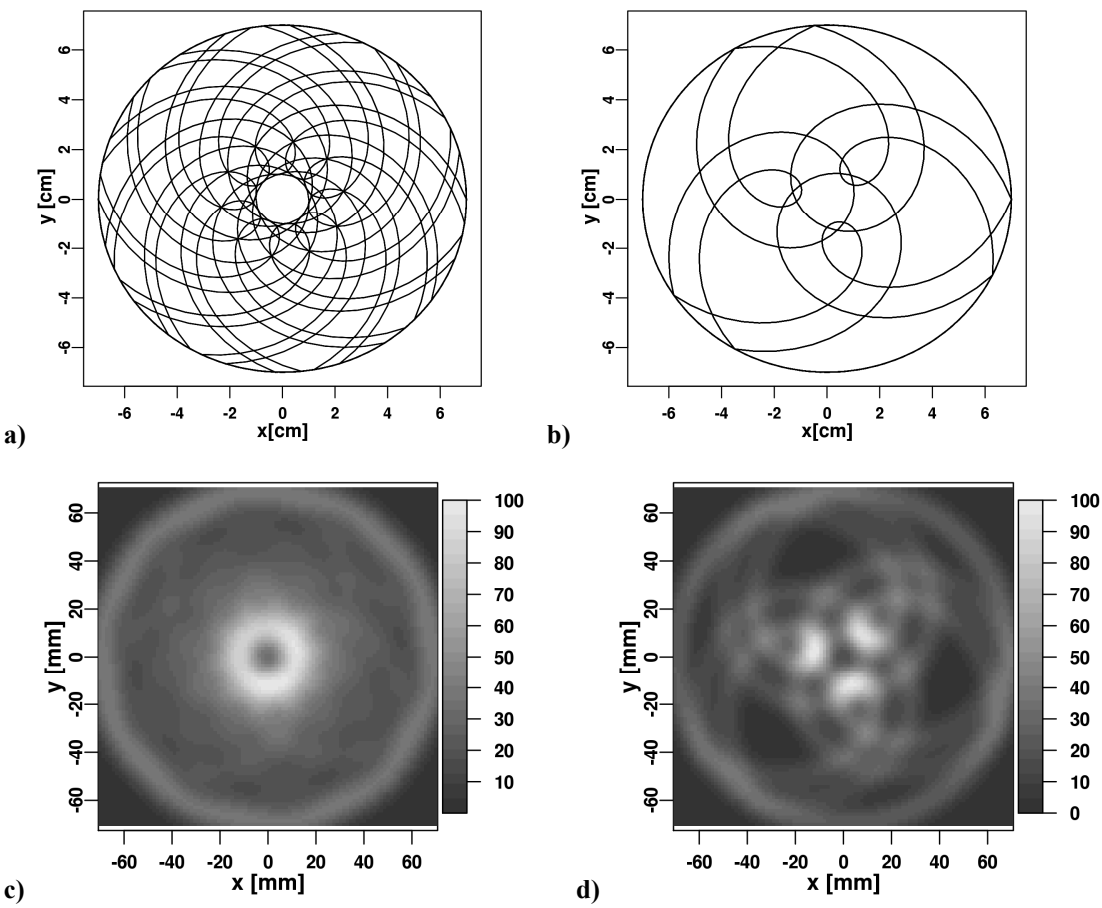


Fig.2) comparison of the spatial MS beam trace at two process conditions 30s process time: chuck period $\mathbf{T_d}$=0.6s ($\boldsymbol{\omega_d}$=100rpm), $\mathbf{d_c}$=10 mm, $\mathbf{r_d}$=70 mm, $\mathbf{r_a}$=400 mm, $\mathbf{t_d}$=0.2s, **t**=30s. **2a**) for arm swing period $\mathbf{T_a}$=1.38s. **2b)** for arm swing period $\mathbf{T_a}$=1.4s. **2c)** megasonic coverage for case **2a**. **2d**) megasonic coverage for case **2b**; Spatial distributions **2c** and **2d** were calculated assuming

Gaussian power distribution within the direct MS beam with standard deviation of 2 mm. Both pictures **2c** and **2d** are normalized on the maximum coverage time.

Impact of the $\mathbf{t_d}$ is as follows. The arm speed is set as period for arm swing across the plate independently from the reversal point setting and the $\mathbf{t_d}$ *i.e.* the real traveling time of the arm is half of the period $\mathbf{T_a}$ minus $\mathbf{t_d}$. In the following Fig.**3** one can see the difference in the MS coverage and judge the intensity of the peak at reversal points in radial MS coverage.

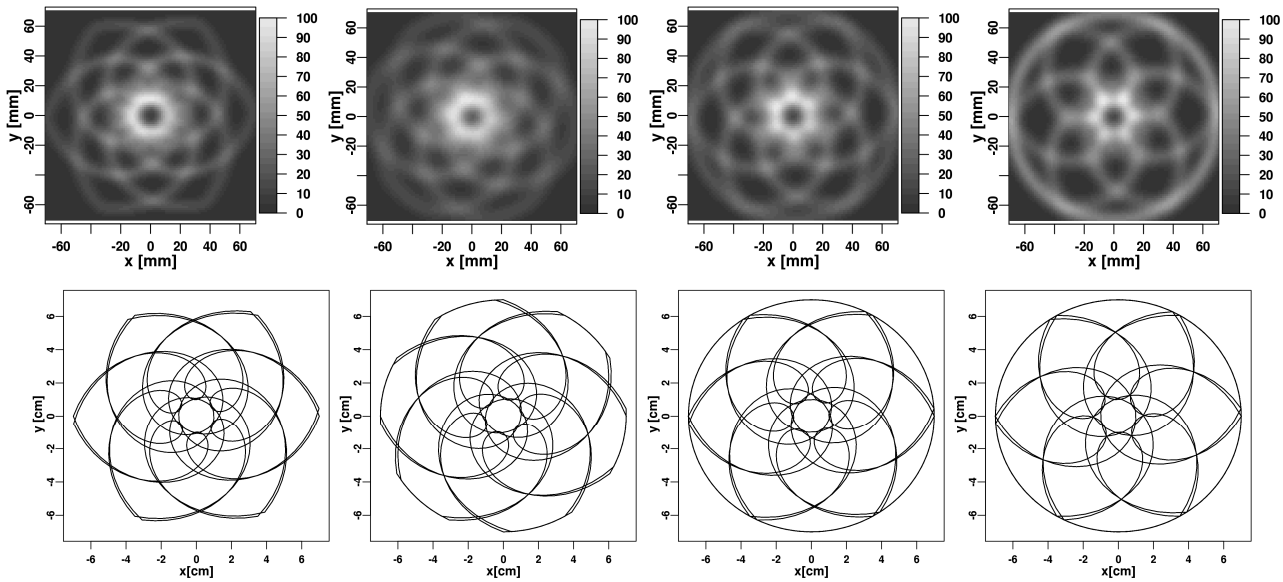


Fig.3) impact of the variation of reversal point delay $\mathbf{t_d}$ at $\mathbf{T_d}$=0.6s, $\mathbf{d_c}$=10 mm, $\mathbf{r_d}$=70 mm, $\mathbf{r_a}$=400 mm, **t**=30s, $\mathbf{T_a}$=1.1s, from left to right the reversal point delay is 0, 0.05, 0.10 and 0.20s

Reflectivity measurements were made with an N&K tool.

RESULTS

Model verification

Using the model described above, the radial distribution of MS 'coverage' can be compared to the experimentally obtained local particle removal efficiencies. By definition, the MS 'coverage' represents the time the center of the MS beam spends at each particular point on the mask. In reality, the MS beam coverage can not equal the MS power 'distribution' which must be influenced by the factors of power damping and local diffusion through the media layer directly beneath the MS beam. In turn, the degree of local diffusion will be influenced by factors of flow rate and mask spin speed. The net effect becomes a beam broadening.

We can estimate the broadened distribution of MS energy around the theoretical beam center by observing the particles cleared by a single circular beam track [1] and evaluate PRE using reflectivity measurements. Fig.**4** shows the co-plot radial function of all parameters experimentally observed: PRE, MS power distribution against the theoretically MS coverage.

The value for reversal point delay $\mathbf{t_d}$ has to be obtained experimentally by fitting of the cleaning trace.

The difference between MS coverage and MS power distribution can be identified in the damping of the power by media layer underneath the MS beam and the media flow. Assuming linear dependence of surface damage to MS energy delivered, one can estimate the distribution of MS energy over the mask using the approach described in [2] and evaluate using reflectivity measurement. Estimation of MS energy distribution was performed using Hoya AR8 Cr blank and measurement of reflectivity distribution at the wavelength of λ=500nm. Fig.**4** shows the co-plot radial function of all three parameters, experimentally estimated PRE, MS power distribution and from model estimated MS coverage. The direct comparison of the three functions identifies weakness of the MS power distribution estimation. The error of that

measurement, as well as the density of points, is about an order of magnitude below the necessary limit. The expected match between the shape of the PRE2 and MS power distribution cannot be confirmed by this measurement. Alternative ways of estimating of the distribution of MS power delivered is under investigation and may clarify the relation between the three parameters plotted in Fig.**4**. Due to this uncertainty in estimation of reflectivity change we have to take alternative way. We decided to make assumption of matching between MS power distribution and PRE2 and use PRE2 for correction of the process time variation across the mask instead of the reflectivity change. The relations between not matching parameters PRE, reflectivity change and MS coverage model have to be clarified by coming experiments.

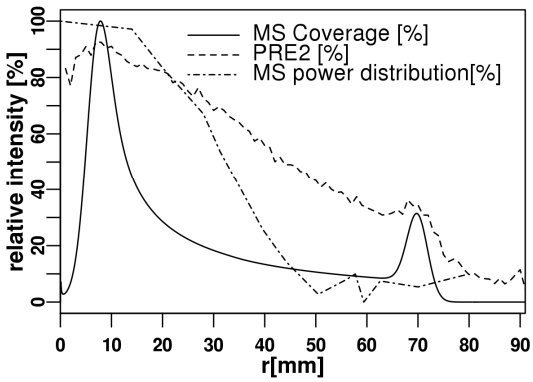


Fig.4) Compares the experimentally PRE as determined on a fully covered particle blank, to the theoretically modeled distributions for MS Power and MS Coverage. A process time of **t**=30s for was used both for the experimental PRE and the theoretical modeled MS Power. The MS coverage measurement was determined using a significantly higher process time of about 40 minutes and very high MS power in order to be able to estimate the reflectivity change.

Kinetics of cleaning

Recent observations make us confident that the kinetics of cleaning can be described using the kinetic rate equation for first order chemical reactions. However, the concentration of reactant must be replaced by particle density per mm^{-2} region for each particle size ($n_0 - n$).

$$\frac{dn}{dt} = k_n (n_0 - n) \qquad [6]$$

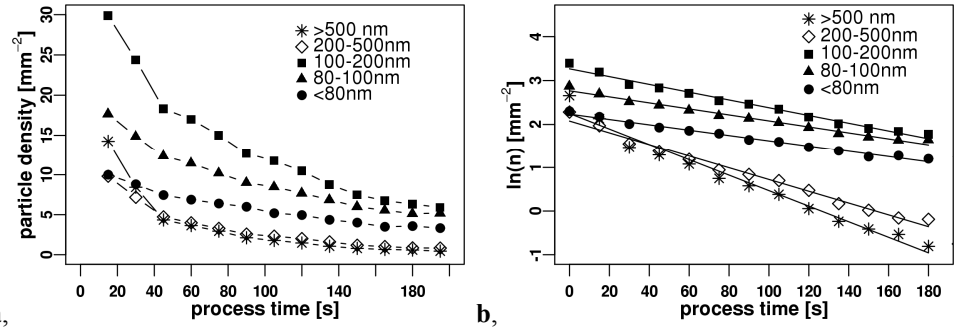
Here **dn/dt** is the change of particle density derived over time. The rate constant $\mathbf{k_n}$ must be determined experimentally for a given set of process conditions, particle sizes and for at least one particular place on the mask.

The description seems to be valid for one place on the mask—more exactly, for places with identical radius from mask center in case the MS energy is distributed uniformly along the circle any radius (as shown in Fig.**2c)**. This effect is caused by the different MS power delivery which equals to different reaction times. More precisely, only the time when MS power affects a particular place on the mask can be counted as process time that affects the kinetic rate.

Equation 7 is derived from Eq. 6 by introducing the function of time

$$\ln(n) = \ln(n_0) - k_n \cdot t \qquad [7]$$

Fig.**5a** plots particle density as a function of time. The method uses a single plate with cleaning is applied in 15s intervals. At each interval, a particle test is performed. By applying Equation 7, we obtain Fig.**5b** where the slopes represent the $\mathbf{k_n}$ for each particle size.



a, b,

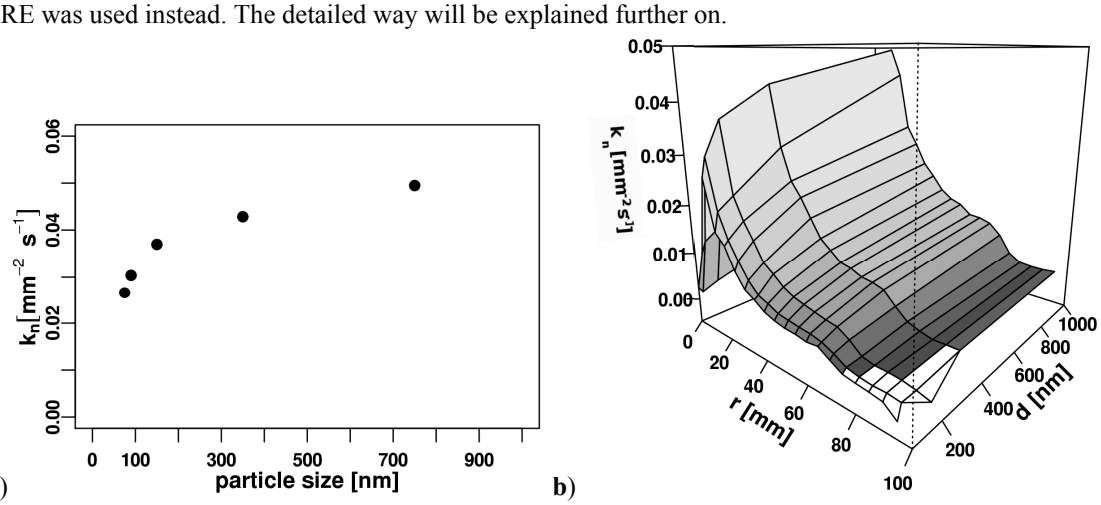
Fig.5a) Displays Particle Density ~~count~~ as a function of Process Time for different particle sizes at **r**=10mm and MS power=13W; **5b)** Redisplays in ln values, each particle size from **5a** using Eq.**7**. The lines corresponds to the rate constant $\mathbf{k_n}$ for each particle category.

The Equation **7** describes the kinetics of particle removal with sufficient cleaning energy to complete particle removal. When cleaning energy is insufficient for a small particle size, the PRE does not reach 95%. In this case, using equation 7 leads to an underestimation of $\mathbf{k_n}$. It implies that, at some process time, the surface can be cleaned completely. Usage of Eq.**8** instead improves the $\mathbf{k_n}$ estimation and provides in addition information about the saturation limit $\mathbf{n_\infty}$.

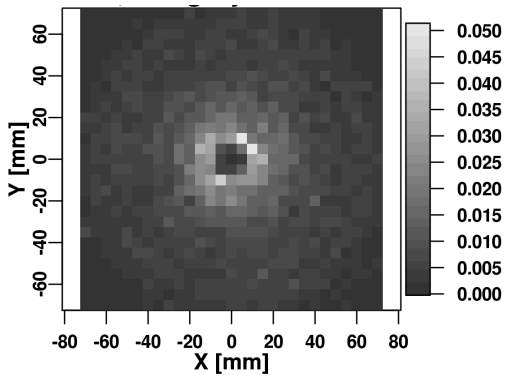
$$\ln(n - n_\infty) = \ln(n_0) - k_n \cdot t \qquad [8]$$

In this way we obtain the rate constants $\mathbf{k_n}$ for each particle size at given radius. Than the rate constant for any place on the mask can be obtained by scaling the $\mathbf{k_n}$ according to the **PRE**. The ratio at **r**=10mm and e.g. **r**=50mm tells us that the removal rate at **r**=50 will be slowed down by factor of 2 compared to radius r=10mm (refer to Fig.**4** PRE2 radial distribution function).

The rate constant $\mathbf{k_n}$ can be estimated using a fully covered mask or can be estimated using only several particle dots on the mask. The uncorrected rate constant $\mathbf{k_n}$ is strongly dependent on radius **r** from mask center, which is assumed to be caused by the difference in MS power distribution across the mask. As previously mentioned the MS coverage is the driving force of the particle removal and that is why we decided to correct the process time for different areas on the mask. Since the MS power distribution measurement is not as precise as expected (Fig.**4**), Particle removal efficiency PRE was used instead. The detailed way will be explained further on.



a) b)



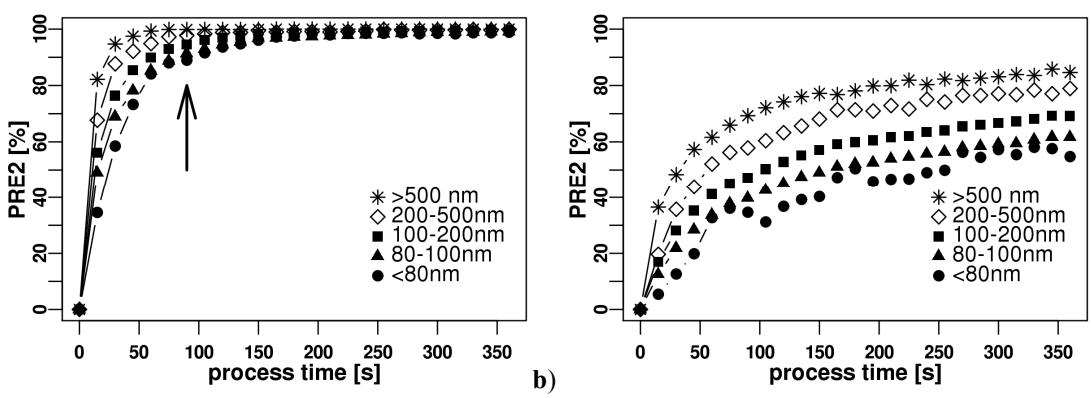
c)

Fig.6a) Displays the kinetic rate constant $\mathbf{k_n}$ as estimated for radius **r**=12mm using particle size categories <80nm, 80-100nm, 100-200nm, 200-500nm and >500nm. **6b**) Estimates the kinetic rate constants $\mathbf{k_n}$ for different particle sizes and radii from mask center. **6c)** Displays an arial plot of the rate constants as estimated across the mask for 25 mm^2 regions using 80-100nm particles. A large variation in $\mathbf{k_n}$ is seen from the mask center towards the edge whereas regions at any particular radius are very uniform.

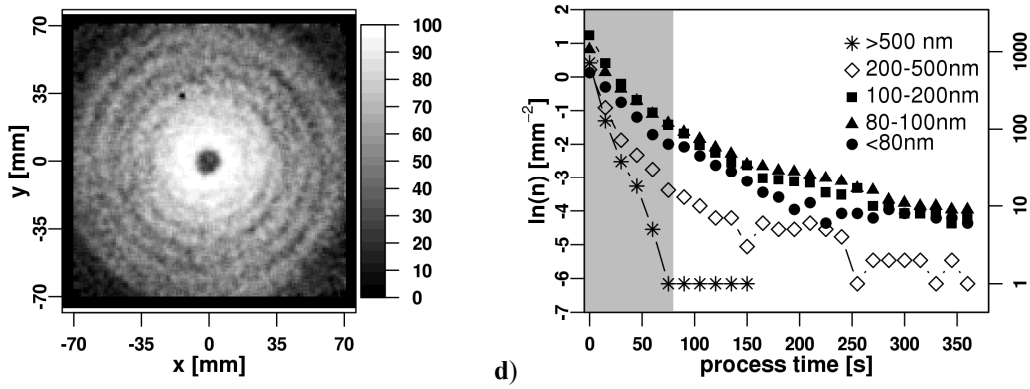
In order to compensate for the radial effect of MS power distribution on the rate constant $\mathbf{k_n}$, the $\mathbf{k_n}$ for each region the mask was divided by a normalized PRE. Figure 7 displays this process graphically. The PRE function was divided by its maximum to obtain the correction function. The rate constant was estimated by linear fit according to Eq.**7**. The time frame for rate estimation was chosen by selection of the process time at which the maximum PRE just exceeds 95%. Figures **7a** and **7b** were obtained by cleaning the same mask in intervals of 15 seconds. Particle measurements were made at each break in the process. The 7th scan is the last we can use for rate constant estimation at least for biggest particles in the central part of the mask. This limits number of the data points we may use for fit according to Fig.**5b** and so increases the scattering especially in the areas with lower particle counts involved in the estimation – center of the mask.

Fig.**7b** shows the PRE as function of process time calculated for the whole mask surface. The function for all particle sizes is more flat than for areas around mask (Fig.**7a**) Use of this function for fit limit estimation leads to using the whole time frame for rate estimation and so underestimates the rate constant in mask center.

Estimation of normalization time period was chosen at which the PRE variation across the mask is well pronounced, but none of the points reached saturation (see Fig.**7**). The compensation based on 5x5 mm^2 leads to increase of $\mathbf{k_n}$ variation due to lower stability when fitting over small particle density. For this reason the $\mathbf{k_n}$ and **PRE** used for compensation of the radius effect were estimated on circular areas with radius difference of maximum 5 mm. The particle counts in the areas selected are between 700 and 16000 particles.



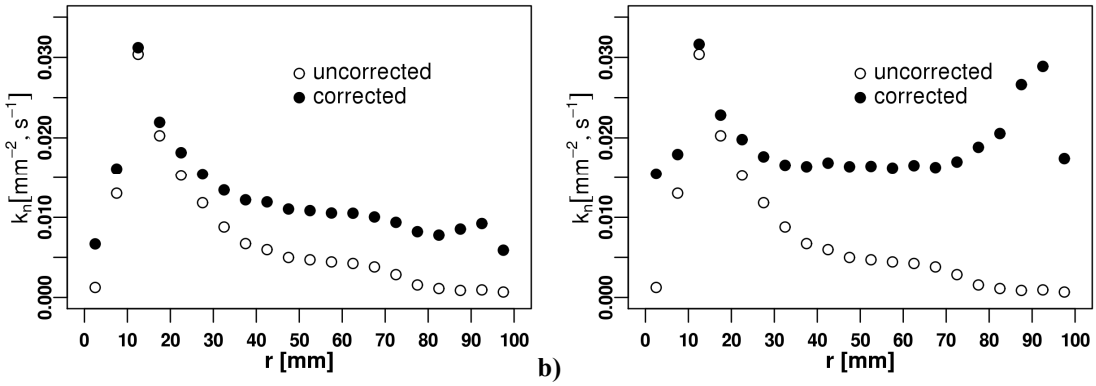
a) b)



c) d)

Fig.7) PRE2 is used to determine the rate constant $\mathbf{k_n}$. **7a)** The plot displays the PRE2 vs. Process Time for different particle sizes as determined over a region equal to the entire mask surface using a radius r=15 mm. The PRE2 was calculated at time t<90s as marked with an arrow. **7b)** Displays the PRE2 determined at the same mask but at the radius r=40mm. At this radius the saturation is not reached and all the timeframe shown can be used for rate estimation **7c**) Displays the arial PRE2 distribution for all particle sizes at 90s process time for a single mask corresponding to the arrow marking 90s of **7a**. This is the maximum process time usable for estimation of the rate constant $\mathbf{k_n}$. **7d)** Displays the logarithm of average particle density per each mm^2 on the mask plotted vs. Process Time **t**. The grey area marks the range in which the data must be fitted in order to avoid distortion of the data for calculating rate constant $\mathbf{k_n}$. This point is clearly given when the large particle density drops below 10^{-6} [mm^{-2}]

Figure **8** shows the comparison of the uncorrected rate constant $\mathbf{k_n}$ obtained by fitting of data for each radius and particle size category according to Eq.**7** and corrected rate constant values. Dividing of the rate constant by the corresponding PRE value shown in Fig.**8a** shows slight improvement, but the corrected data will still show exactly the same trend. Optimum results were obtained by dividing of the rate constant by PRE$^{3/2}$ as shown in Fig.**8b**. The improvement in uniformity of rate constant measured as a ratio between highest value and the smallest value improves by factor of 24 from 2% to 48%. One can observe stronger variation of rate constant at both reversal points in this case at r≈14 and r≈70mm. Above 70 mm where almost no PRE can be observed, the correction leads to overestimation of the rate constant.



a) b)

Fig.8) Displays the variation in rate constant $\mathbf{k_n}$ across the mask for particle size category 80-100nm. **8a**) Compares the uncorrected and corrected data by dividing by PRE. The correction improves the ratio between the highest and smallest value from 2% to 18%. **8b)** Demonstrates the use of a rate constant corrected by PRE$^{-3/2}$. With this correction method, we see the rate constant is completely independent of radius over a broad range of radii not containing the reversal points. This correction improves the ratio from 2% to 48%.

Finally, the kinetics model described in Eq.**8** is demonstrated graphically in Fig.**9**. The reduction of particle density obtained at low cleaning energy is plotted as a function of process time. For comparison, we plot both fitting methods to illustrate their shape match to the raw data. The curvature of the data gives a clear advice to use Eq.**8** instead of Eq.**7**. The saturation limit is about 3.99 and corresponds to PRE value of 79.2 %.

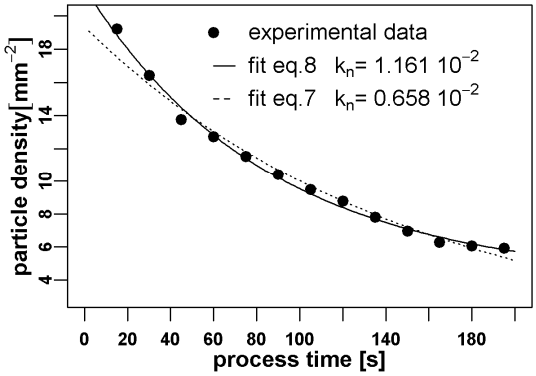
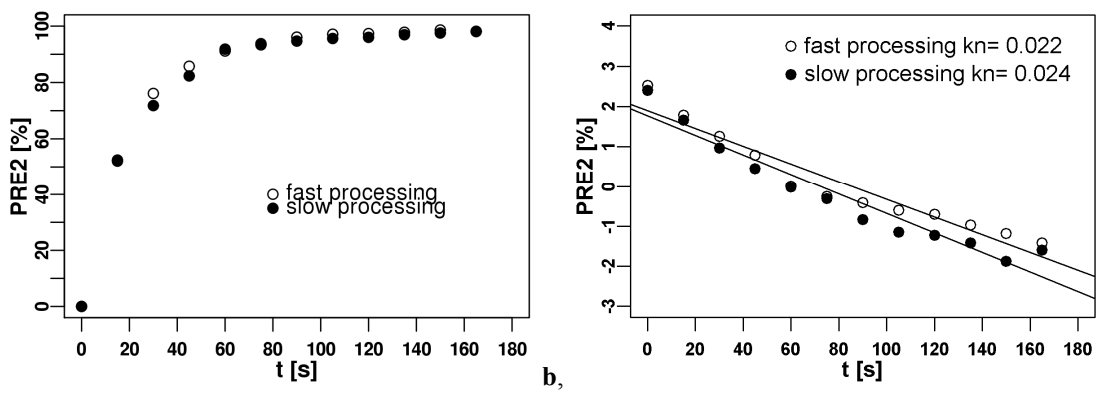


Fig.9) Overlays curves of Particle Density vs. Process Time for the raw experimental data and the same data using Eq.**7** and Eq.**8**. - The data were collected at low cleaning energy to ensure a particle removal efficiency below 95% for particle size 80-100nm.

Impact of waiting time

The waiting time may be a contributing factor to experimental noise. Knowledge of the impact is important for proper evaluation of the experiments. Two samples, one with a delay of 1h between two cleaning steps, the second with a delay of about 24h, indicated no influence on the removal efficiency and cleaning kinetics of Si_3N_4 particles down to 80 nm. Due to the fact that also the particle density is in same range, one can easily compare the cleaning results directly. In Fig.**10** one can see the comparison of both particle removal efficiency and rate constant. Both variables show identical performance within the expected error bars.



a, **b**,

Fig.10a) Comparison of time resolved PRE2 for masks cleaned slowly (1 step in 24h max.) and quickly (1 step per hour). The PRE2 was estimated on r $\in$(10; 15) mm, particle size 80-100nm. Deviation between both data sets is within the range of reproducibility.
10b) The rate constant $\mathbf{k_n}$ also shows no significant difference between samples.

DISCUSSION AND CONCLUSIONS

Our goal has been to target some of the problems experimenters are facing in cleaning and clarify the root cause of the big error bars in experiments. Noise and variation have many contributing factors. The size distribution of the particles applied influences the count statistics. The method used to apply particles influences the uniform distribution of particles on the mask surface. The measurement tools also contribute. Confidence in small particle assay is frustrated by the lower capture rate probability. Different types of particles behave differently. Delay time has been determined to influence removal efficiency for polystyrene latex and latex particles [3]. This occurs by the deformation of the particle on the plate which increases contact surface area and intimacy of contact. On the contrary, for Si_3N_4 ceramic particles, the delay between particle additions and their removals does not influence removal efficiency. The lack of deformability in ceramics may be at cause.

A more general model can be established for particle removal that accommodates variation in particle density, position on the mask, particle size distribution and the impact of the cleaning process itself. Towards this end, a precise model of megasonic head movement over the mask was made utilizing the geometry of the mechanism. MS Coverage is a term

used to define the movement of the MS head over the plate while the term MS Energy Distribution incorporates particle removal data to determine MS beam spread from the parameters including chuck speed, arm swing extents, swing speed and MS arm length. This model provides us information about the exposure time to the direct megasonic beam at each region. It becomes possible to more precisely estimate the impact on cleaning uniformity.

The kinetics of particle removal can be used for direct comparison of different cleaning processes. Correction of the $\mathbf{k_n}$ as function of position on the mask was performed using particle removal efficiency. The PRE does not correlate perfectly with the driving force. However a correction practice of dividing the rate constant by the PRE provides a data set with reduced variation but showing an identical trend. Using $PRE^{3/2}$ for correction instead leads to perfect correction in the area between reversal points covered by the direct MS beam. At areas around reversal points, the $\mathbf{k_n}$ is overcompensated. Careful consideration as to the correction method used and the importance of observing a saturation limit $\mathbf{n_\infty}$ will avoid underestimation of the rate constant $\mathbf{k_n}$.

ACKNOWLEDGEMENT

AMTC is a joint venture of GLOBALFOUNDRIES and Toppan Photomasks and gratefully acknowledges the financial support of the Federal German Ministry of Education and Research (BMBF) under Contract No. 13N9675 (ANTARES).

REFERENCES

1. P. Nesladek, S. Osborne, C. Kohl, "Particle transport and reattachment on a mask surface", EMLC 2009 proceeding
2. S. Osborne, V. Baudiquez, T. Rode, C. Chovino, H. Takahashi, E. Woster, "Damage free megasonic resonance cleaning for the 45nm design rule", 23rd European Mask and Lithography Conference, Proc. of SPIE, **Vol 6533**, 2007.
3. A. Busnaina, *Fundamentals of Adhesion, Wet and Dry Removal of Nano-Particles in Mask Surface Cleaning. Nanoparticle Removal from EUV Masks and Trenches*, Oral Presentation, SEMATECH Mask Cleaning Workshop, Monterey, CA, 2006