RF Sensing for Real-Time Monitoring of Plasma Processing

Craig Garvin and Jessy W. Grizzle

University of Michigan Electronics Manufacturing Laboratory, 3300 Plymouth Road, Ann Arbor, MI 48105-2551
garv@eecs.umich.edu, grizzle@eecs.umich.edu

A novel sensing system based on the microwave resonance probe is compared to standard RF metrology. The system uses an antenna in the glow discharge to excite the bulk plasma at a frequency range of $30MHz$ to $1GHz$. Standard RF metrology is implemented by measuring the fundamental and five harmonics of the RF power signal. An experiment varying power, pressure, CF_4 and O_2 is constructed. Using a subset of the data to regress a model, standard RF sensing reconstructs the experimental variables with a best average R^2 of 0.586 at a high model coefficient variance (σ_b^2), whereas the novel sensing system results in a best average R^2 of 0.804 and an order of magnitude lower σ_b^2.

INTRODUCTION

Ever shrinking geometries and larger substrates are mandating improvements in sensor systems for control and diagnostics of plasma processing. Experts in industry and academia have recognized the RF signal ($13.56MHz$) and its harmonics as a potential source of process information. A number of researchers have used RF metrology as a tool for plasma diagnostics. Maynard *et al.* (1) have used RF metrology for end-pointing of an industrial etch process. Spanos *et al.* (2,3) have extensively used RF metrology in their plasma diagnostic and control work. A number of researchers (4–7) have used ion flux information obtained via RF metrology to characterize etching. Researchers at the Adolph Slaby Institute (8,9) have developed a diagnostic system that uses plasma physics models to infer process - relevant information from the fundamental frequency exciting the discharge and multiple harmonics resulting from plasma non linearities. Attempts to model the relationship between measured RF parameters and plasma physics have been presented in (10–12). This list is by no means complete. The common element is the attempt to relate measurement of the $13.56MHz$ signal to plasma, wafer, and tool conditions.

The ionospheric plasma community has followed a radically different approach to extracting information from plasmas, namely the use of resonance probes (13). A resonance probe is like a Langmuir probe in that it is an active probe. However, unlike the Langmuir probe, the resonance probe operates at a frequency far above the standard RF signal. In its standard mode of operation, an antenna is inserted into the plasma and driven over a frequency range centered about the plasma frequency in order to infer electron concentration.

This article investigates the use of a resonance probe in a micro-electronics processing plasma. The primary goal is to evaluate the relative observability of the state of the plasma under standard RF metrology and resonance probe techniques. More precisely, the aim of most plasma diagnostics is to *detect* a change in the plasma state and to *isolate* the source of this change. Accordingly, in this first study, a very simple experiment is conducted to compare the abilities of the two measurement techniques to detect and isolate plasma changes due to the variation of generator power, chamber pressure, and gas chemistry.

In contrast to the ionospheric plasma community's use of only plasma resonance frequency information, analysis of the resonance probe data presented in this article employs all frequency information. To underscore the frequency range of the modified resonance probe monitoring, it is referred to as *broad band sensing* whereas methods relying on the plasma's RF signal and its harmonics are referred to as *narrow band sensing*. It could be anticipated that a signal containing a broader range of data may carry with it more information about the plasma's condition. Initial results indicate that broad band sensing exhibits a sensitivity to plasma parameters that is significantly stronger than that achieved through conventional narrow band sensing.

EXPERIMENTAL SETUP

An experiment was designed in order to simultaneously collect broad band and narrow band data as a function of widely varying plasma conditions, as shown in Figure 1. Experiments were performed on a GEC research reactor, described in Ref. (14). Power is delivered using an ENI generator and matching network, and generator power is measured using the built in ENI power meter. Pressure is measured using an MKS barratron, and gas flow rates are measured with MKS flow meters with MKS gas correction factors.

Narrow band sensing is implemented using a Werlatone C1373 $1.5MHz$ to $80MHz$ directional coupler rated at 750 Watts power with a nominal $-30dB$ coupling between main line and sensor ports received by a Tektronix TDS 420 digital storage oscilloscope. Groups of 1000 points of both forward and reverse waves are sampled at 200 Ms/sec, resulting in a total of about 65 periods of each wave. Each waveform is then digitally filtered using separate 6^{th} order discrete Chebyshev bandpass filters centered at the fundamental frequency and first five harmonic components respectively. The magnitude and relative phase of forward and reverse wave at each frequency component is then calculated.

CP449, *Characterization and Metrology for ULSI Technology: 1998 International Conference*
edited by D. G. Seiler, A. C. Diebold, W. M. Bullis, T. J. Shaffner, R. McDonald, and E. J. Walters

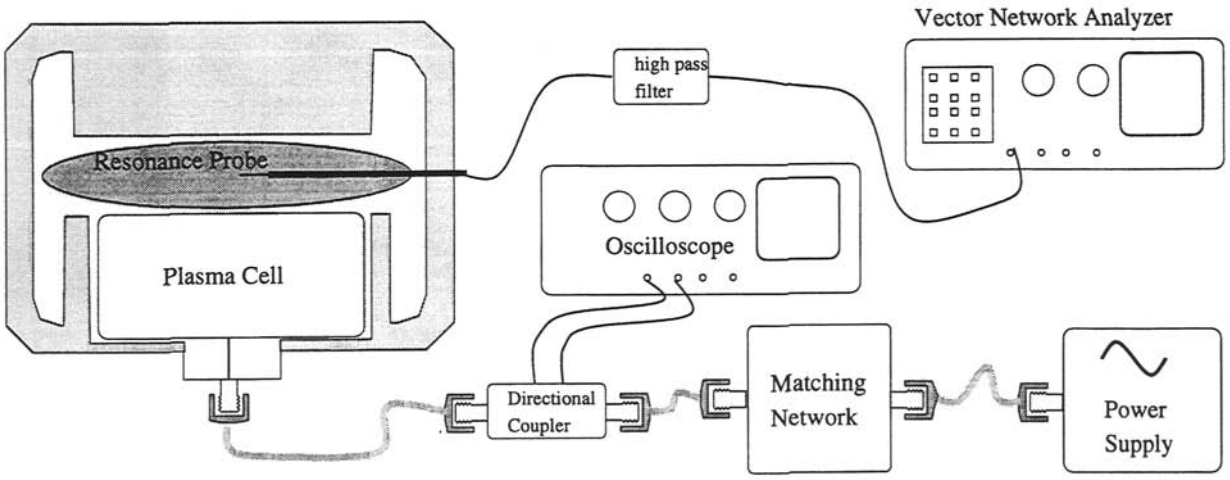

FIGURE 1. Experimental Setup for Comparing Broad Band and Narrow Band Sensing

TABLE I. Experimental Variables and Levels

Variable	Level 1	Level 2	Level 3
Power	$60W$	$70W$	$80W$
Pressure	$40mTorr$	$50mTorr$	$60mTorr$
CF_4 flow	$40sccm$	$50sccm$	-
O_2 flow	$0sccm$	$2sccm$	-

Broad band sensing is achieved by a resonance probe constructed of a length of RG402u stainless steel rigid coaxial cable. Approximately one inch of center conductor is exposed to the plasma to act as an antenna. The probe is inserted in the bulk plasma using an O-ring compression sealed vacuum port. A Hewlett Packard 8712B vector network analyzer drives the resonance probe over a range of $30MHZ$ to $1GHz$ at a power level of 0 dBM. A Mini Circuits $25MHz$ high pass filter is used to isolate the vector network analyzer from the discharge. After calibration, the complex reflection coefficient (Γ) is recorded at 801 frequency points linearly uniformly spaced between $30MHz$ and $1GHz$. To maximize accuracy, the data is collected using the analyzer's lowest IF bandwidth. The set-points of power, pressure and flow rate for the GEC as well as data acquisition are controlled with a Macintosh Quadra 950 running LabVIEW data acquisition and control software. All data is logged and written to file automatically.

EXPERIMENT AND INITIAL RESULTS

The goal of the experiment is to evaluate the ability of broad band and narrow band sensing to isolate basic plasma perturbations due to changes in power [1], pressure and chemistry. Accordingly, a full factorial experiment is performed, as summarized in Table I. Standard statistical methods are then used to construct a linear regression model to predict the variables listed in Table I based on measured broad band and narrow band data, respectively.

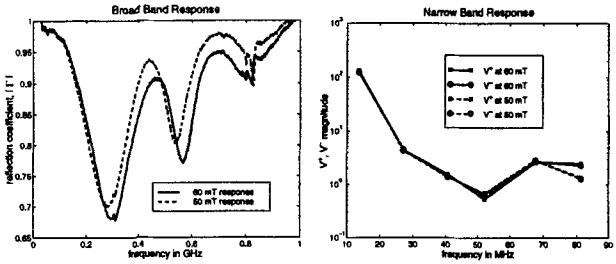

FIGURE 2. Broad Band and Narrow Band Response for 50 and 60 mTorr Plasma Pressures

It is informative to consider the response of the two sensing systems to a variation in pressure. Figure 2 shows both the broad band and narrow band magnitude response to a change in pressure from 50 $mTorr$ to 60 $mTorr$ with power at a constant 60 W, and a chemistry of pure CF_4 at 50 $sccm$ flow rate. In the broad band signal, there is a clear and distinct trend differentiating the two pressure conditions. Although both plasmas show absorptive peaks near 300, 600 and $800MHz$, the precise magnitude and location of these

[1] Power level refers to a generator set point and not a calculated delivered power to the plasma.

peaks changes visibly with pressure. Though it remains to be seen whether this qualitative observation will be elicited in the statistical analysis, the structure of the response using the broad band sensor suggests that substantial information about fundamental plasma physics may be embedded in the sensor data. The narrow band signal is also reported in Figure 2. A log scale is used due to the wide dynamic range of the signal[2]. It is more difficult to discern a pattern in the narrow band response, but this by no means indicates that structure is lacking.

EXPERIMENTAL ANALYSIS

The experiment described in Table I results in 36 experimental points, each consisting of a specific level of the independent variables. For each experiment, 12 complex points of narrow band and 801 complex points of broad band data are collected. The goal of the experimental analysis is to generate a linear regression model from each of the sensor systems to the plasma state as represented by power, pressure and chemistry. It is assumed that if the same regression methods are used in both cases, differences in fit can be attributed to fundamental differences in observability between the two systems. To present a direct comparison of methods, no additional transformations (such as using impedance or standing wave ratio representations) are used. The narrow band data is considered as $\mathbf{V}^+$, $\mathbf{V}^-$, and phase. The broad band data is considered as $\mid \Gamma \mid$, and phase.

Two statistical parameters are used in the regression. The first is the standard R^2 evaluation of residuals as a function of output, defined as

$$R^2 = 1 - \frac{(y-\widehat{y})^T(y-\widehat{y})}{y^T y}, \tag{1}$$

where y is the zero mean variable being estimated and $\widehat{y}$ is the estimate. The second is σ_b^2, the variance in the model parameters, which is estimated [3] by

$$\sigma_b^2 \approx s^2 \cdot \|(x^T x)^{-1}\|, \tag{2}$$

$$s^2 = \sum_{i=1}^{n} \frac{(y_i - \widehat{y}_i)^2}{n}. \tag{3}$$

In the above, the induced 2-norm ($\|\cdot\|$) is used to reduce the covariance matrix of model coefficients to a single parameter.

The model variance, σ_b^2, essentially quantifies the relationship of the conditioning of the pseudo inverse ($(x^T x)^{-1}$) to the residual ($(y_i - \widehat{y}_i)^2$). It is clear that as more sensor data is used to estimate y, the fit will improve. However, more data means poorer conditioning of $(x^T x)^{-1}$. If the negative effects of poor conditioning outweigh improved fit, model variance increases. Physically, this suggests a model with poor robustness. It can be expected that a model with a high σ_b^2 would fit a new data set very poorly due to the extreme sensitivity of the model to sensor noise.

TABLE II. Experimental Variables and Levels Used for Modeling Data Set

Variable	Level 1	Level 2
Power	$60W$	$80W$
Pressure	$40mTorr$	$60mTorr$
CF_4 flow	$40sccm$	$50sccm$
O_2 flow	$0sccm$	$2sccm$

The regression model is developed using a subset of the experimental data set, referred to as the *modeling set*, and tested on the full data set. The modeling set is composed of the 16 extremal experiments, shown in Table II. This modeling set is used for evaluating the relationship between principle components and σ_b^2. The fit (R^2) is then evaluated on the entire 36 experiment data set. This approach to evaluating model performance confounds two factors. By testing the model with new data, robustness is evaluated. Additionally, because the new data is at different experimental levels (the midpoints), the linearity of the underlying physics plays a major factor in fit performance. Because such a small data set is used for developing the model, it is expected that the fit to the entire 36 experiment data set may be relatively poor. Since the goal of this initial investigation is to compare the two measurement systems under similar conditions, only relative performance is of interest.

In order to study σ_b^2, principle component regression as described in (15–19) is used. Subsets of the original measurement data are formed using an increasing number of principle components (PC's). A plot of σ_b^2 as a function of PC's guides the selection of an optimum number of PC's. Principal component regression can be directly implemented on the narrow band data. It is straight forward to take the singular value decomposition of the 16×18 matrix composed of rows of experimental levels and columns of centered and scaled $\mathbf{V}^+$, $\mathbf{V}^-$, and phase. The same direct approach is not computationally practical for the broad band data where singular values of a 16×801 complex matrix are required. Accordingly, a preliminary reduction by FFT is performed on the data. A discrete Fourier transform of the data is taken, and the frequency components sorted by magnitude. The data is then reconstructed with increasing numbers of components until a compromise between data reduction and fit is found. For our data set, using the 300 most significant frequency components results in an R^2 *of the FFT reduced broad band sensor data to the unprocessed sensor data of* 0.999, and yields a tractable data set upon which to perform a principle component regression.

[2]The 6^{th} harmonic is above the directional coupler frequency response of $80MHz$, slightly affecting the coupling.

[3]Precise conditions for the convergence of Eq. (2) to actual variance can be found in Sen (15) or other statistics texts.

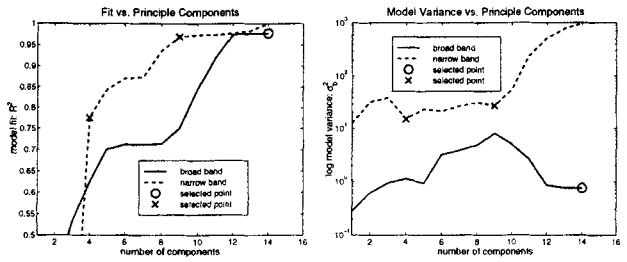

FIGURE 3. R^2 and σ_b^2 as a Function of Number of Principle Components on the 16 Experiment Modeling Data Set

Regression models fitting the four parameters of power, pressure, CF_4 flow and O_2 flow are generated using the PC's of the sensor data corresponding to the n strongest singular values. The procedure is performed for both broad band and narrow band data *on the 16 experiment modeling set.* Figure 3 shows both R^2 and σ_b^2 as a function of the number of PC's used in the regression model. It should be noted that Eq.(2) becomes less accurate as the number of measured data points approaches the number of experiments. A sharp drop in σ_b^2 serves as an indicator that the variance estimate is no longer accurate. In both the broad and narrow band cases, σ_b^2 is no longer accurate at 15 PC's and accordingly, data is only plotted to 14 PC's.

Choosing an optimal number of PC's for the broadband data is relatively straightforward. The variance estimate, σ_b^2, decreases monotonically above 9 PC's. Clearly, a 14 principle component model is best for the broad band data. Determining the optimal number of PC's in the narrow band case is more difficult. The variance estimate, σ_b^2, dips to a clear local minimum at 4 PC's. Another slightly higher local minimum is seen at 9 PC's. The model fit at 4 PC's is much poorer than that at 9 PC's. This may offset the better conditioning suggested by low σ_b^2. It makes sense to to test both these models against the complete 36 experiment data set.

TABLE III. R^2 and σ_b^2 of Broad and Narrow Band Sensor vs Experimental Variables: Entire 36 Experiment Data Set

	Broad Band		Narrow Band			
	14 PC's		9 PC's		4 PC's	
Variable	R^2	σ_b^2	R^2	σ_b^2	R^2	σ_b^2
Power	0.807	3.85	0.836	88.6	0.791	8.67
Pressure	0.885	2.46	0.667	192	0.416	25.9
CF_4	0.501	3.78	0.303	143	0.0570	14.9
O_2	0.997	0.0010	0.539	4.05	0.602	0.269

Table III summarizes R^2 and σ_b^2 data regression models using both the broad and narrow band sensor data to estimate experimental variables *in the complete 36 experiment data set.* With the exception of power, the fit with the broad band data is substantially better than that of the narrow band data. Even the fit to the power levels reflects favorably on the broad band sensor. Using only impedance information, the broad band sensor is able to determine power levels almost as accurately as the narrow band sensor, which has access to both forward and reverse voltage magnitudes. Of particular interest is the lower model variance achieved by broad band sensing. The low σ_b^2 promises a more robust parameter estimation under realistic process conditions, and will require additional experiments to confirm.

The R^2 and σ_b^2 data for 4 and 9 PC regression models using narrow band sensor show the trade offs that appear to be present narrow band sensing. Results of more than doubling the model size for the narrow band sensor are mixed. Average R^2 does improve from 0.466 to 0.586 but the variance estimate degrades by more than an order of magnitude. Additionally, R^2 performance is mixed. The CF_4 flow estimate improves dramatically with more PC's, but the O_2 flow estimate is actually worse. Clearly, there is a tradeoff in the narrow band data between improved predictive ability and model robustness.

FUTURE WORK

The promise shown by this initial experiment motivates several areas of future work. The range of narrow band sensing must be extended to provide an effective comparison. In this experiment, the coupler bandwidth limited data collection to 6 harmonics. Work at the Adolph Slaby Institute indicates that valuable information may be found at far higher harmonics. It appears that the main difference between broad and narrow band sensors is in model robustness. Accordingly, experiments should be designed to draw out these differences.

Although results using a resonance probe are promising, the potential of an intrusive diagnostic is clearly limited. In order to be practical, a non-intrusive application of broad band sensing is needed, as proposed in Figure 4. A combination of broad band combiner and coupler are used to add the broad band signal to the standard $13.56MHz$ power wave. The combiner has very strong directivity. As a result, power from the $13.56MHz$ power supply does not damage the analyzer. The broad band coupler is used to make a reflection measurement using the standard 'external test kit' mode of the network analyzer.

Finally, the plasma frequency response shows very interesting characteristics. An attempt to relate these to physical parameters through more sophisticated plasma models might be informative.

CONCLUSION

Power, pressure, CF_4 and O_2 levels relevant to microelectronics processing were varied using a full factorial experiment performed on a GEC reference cell. Standard RF sensing (referred to as 'narrow band sensing') was compared

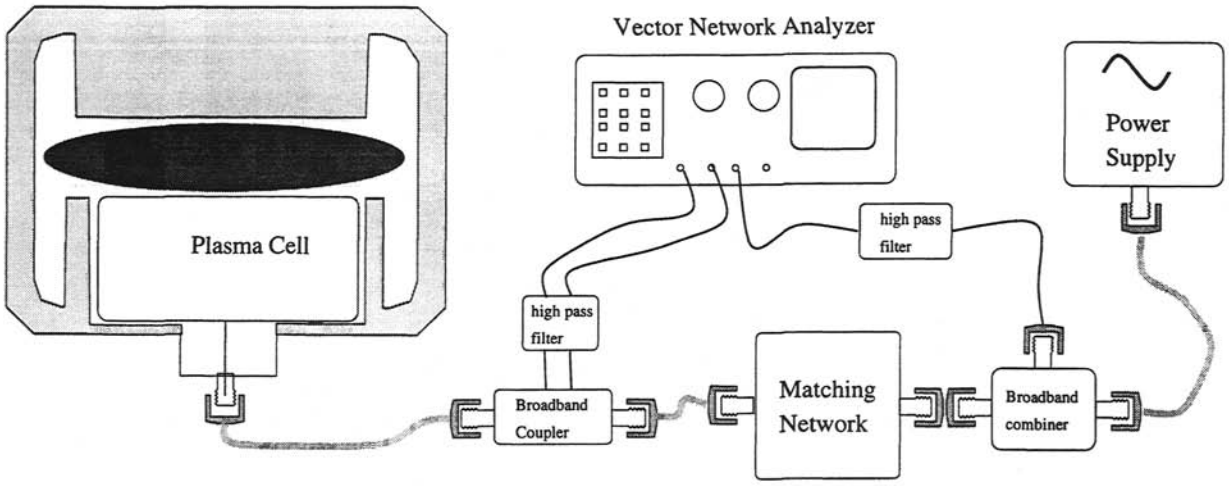

FIGURE 4. Potential Non-intrusive Setup for Broad Band Sensing

to a novel sensing technique based on resonance probes used in ionospheric research (referred to as 'broad band sensing'). Standard statistical techniques were used to regress a linear model against narrow band and broad band data respectively. A much better fit to the data was obtained using broad band sensing. Suggestions for further work include using more harmonics for the narrow band data and methods for designing a non intrusive broad band system.

ACKNOWLEDGMENTS

The authors sincerely thank Steven C. Shannon, and Professor Mary L. Brake for use of and assistance with the GEC cell. Sven G. Bilén constructed the resonance probe used in this research. Hyun-Mog Park designed and wrote the automated data collection algorithm. Dr. Dennis S. Grimard, Dr. Helen Maynard and Oliver Patterson provided many useful discussions.

This work was supported in part by AFOSR/ARPA MURI Center under grant # F49620-95-1-0524 and The Semiconductor Research Company under contract #97-FC-085.

[1] H. L. Maynard, E. A. Reitman, J. T. C. Lee, and D. E. Ibbotson, J. Echem. Soc. **143**, 2029 (1996).

[2] A. J. Miranda and C. J. Spanos, J. Vac. Sci. Technol. A **13**, 1888 (1996).

[3] A. Ison, W. Li, and C. J. Spanos, in *1997 IEEE International Symposium on Semiconductor Manufacturing* (The Institute for Electrical and Electronics Engineers, San Francisco, California, 1997), Vol. 1997.

[4] A. J. van Roosmalen, W. G. M. van den Hoek, and H. Kalter, J. Appl. Phys. **58**, 653 (1985).

[5] N. Hershkowitz, J. Ding, and J. Jenk, in *Proceedings of the IEEE International Conference on Plasma Science* (The Institute for Electrical and Electronics Engineers, Santa Fe, New Mexico, 1994), Vol. 1994, p. 110.

[6] S. Bushman, T. F. Edgar, I. Trachtenberg, and N. Williams, in *Proceedings of SPIE* (International Society for Optical Engineering, Bellingham, Washington, 1994), Vol. 2336.

[7] M. A. Sobolewski, IEEE Transactions on Plasma Sciences **23**, 1006 (1995).

[8] M. Klick, J. Appl. Phys. **79**, 3445 (1996).

[9] M. Klick, W. Rehak, and M. Kammeyer, Jpn. J. Appl. Phys. **36**, 4625 (1997).

[10] P. A. Miller, L. A. Romero, and P. D. Pochan, Phys. Rev. Letts. **71**, 863 (1993).

[11] V. A. Godyak, R. B. Piejak, and B. M. Alexandrovich, J. Appl. Phys. **69**, 3455 (1991).

[12] N. Hershkowitz, IEEE Transactions on Plasma Sciences **22**, 11 (1994).

[13] K. Takayama, H. Ikegami, and S. Miyasake, Phys. Rev. Letters **5**, 238 (1960).

[14] J. Hargis, P. J. *et al.*, Rev. Sci. Inst. **65**, 140 (1994).

[15] A. Sen and M. Srivastava, *Regression Analysis: Theory, Methods, and Applications* (Springer-Verlag, New York, New York, 1990).

[16] A. Ison and C. J. Spanos, in *1996 IEEE International Symposium on Semiconductor Manufacturing* (The Institute for Electrical and Electronics Engineers, Tokyo, Japan, 1996), Vol. 1996.

[17] R. Chen and C. J. Spanos, IEEE Transactions on Semiconductor Manufacturing **10**, 307 (1997).

[18] R. Chen, H. Huang, C. J. Spanos, and M. Gatto, J. Vac. Sci. Technol. A **14**, 901 (1996).

[19] S. Leang and C. J. Spanos, IEEE Transactions on Semiconductor Manufacturing **9**, 101 (1996).