

Control-Oriented Modelling of an Experimental $Ar-O_2$ Plasma Process

Petar Iordanov

Department of Electronic and Computer Engineering
University of Limerick
Limerick, Ireland
Email: petar.iordanov@gmail.com

John Ringwood

Department of Electronic Engineering
National University of Ireland
Maynooth, Co. Kildare, Ireland
Email: john.ringwood@eeng.nuim.ie

Abstract—Development of a control-oriented model of an experimental plasma reactor is presented in this paper. The model structure is suitably partitioned in order to facilitate any subsequent control design. In the model, the linear dynamical part is conveniently separated from a static nonlinearity, which in turn allows identification to be performed for both parts independently. Validation results indicate that the model gives a reasonable representation of the studied plasma process.

I. INTRODUCTION

Etch conditions in a *reactive ion etching* (RIE) system depend strongly on many process variables, such as pressure, gas flows, radio-frequency (RF) power, etc. Various etch recipes are typically used by manufacturers, depending on the configuration of the etching chamber (reactor) and the material to be etched. These recipes include a combination of process variables that can be manipulated in order to achieve etch profile with prescribed properties. While this may seem a reasonable approach, a recipe designated for a specific RIE chamber may not necessarily provide satisfactory results on another chamber of the same brand and specifications. Moreover, the etching process is quite sensitive to the variation of plasma variables, such as charged particle flux, electron density, neutral densities, etc., [1]. Plasma variables may vary significantly from run to run under the same operating conditions. Variables can be disturbed by various factors such as wall depositions and inaccurate set point adjustment in actuators. In other words, the “recipe”-type approach for RIE controls plasma and etch parameters in an open-loop manner, and does not take into account any possible process perturbations that may occur during operation. In an open-loop configuration, etching performance may vary significantly from wafer to wafer, reducing achievable wafer tolerances.

Real-time closed loop control of plasma-assisted semiconductor manufacturing processes such as RIE could yield greatly improved performance. A strategy to reduce the effect of disturbances, which has received some attention in recent years, has been to control plasma variables such as the electron density rather than to attempt to implement feedback control of a variable such as etch rate directly, [2], [3]. Successful implementation of this control strategy would enable etch recipes to be specified in terms of plasma variables as opposed to manipulated variables such as RF power and gas flow rates.

The efficiency of a closed loop control strategy is, to a large extent, determined by the process model. A good process model must adequately describe the static and dynamic behaviour of the process, and yet be simple enough to facilitate the design of a feedback system. Linear dynamical data-based models are easy to derive and are suitable for most control strategies known to date. However, such models provide limited process insight and are usually specific to an operating point of a particular reactor. Possibly, a more reasonable approach to model plasma processes is to exploit the information that is available about the physical and chemical interactions occurring in the process. Models based on this methodology are usually referred to as *first-principles based models*. Although, first-principles based models take advantage of variables that describe plasma chemical kinetics, chamber geometry, delivery of RF power to the plasma, actuators, sensors, etc., such models tend to rapidly become quite complex as the number of chemical species in the plasma increases. In such cases, this type of process models would become virtually unusable for control design. First-principle based models help the understanding of physical and chemical interactions present in a plasma discharge, however their complexity does not allow them to be employed alone. In addition, plasma science has not yet evolved to the extent that all plasma interactions are completely understood and quantified.

A control-oriented model of an experimental plasma process is presented in this paper. The available knowledge about the plasma process components allow construction of a grey-box model in which linear dynamics are separated from the static nonlinearity, thus facilitating any subsequent control design.

The paper is structured as follows: Section 2 gives a description of the plasma process under consideration and the setup for computer-based data acquisition. Section 3 is dedicated to the development of a control oriented model of the plasma process. This section presents a Wiener-type model in which the linear dynamic part and the static nonlinear part are separated. For the static nonlinear part, various model structures are studied and assessed in section 4. Some validation results are provided in section 5, while conclusions are detailed in section 6.

II. PROCESS DESCRIPTION AND CONTROL OBJECTIVES

A. Process Description

BARIS (**BA**sic **RA**diofrequency **I**nductive **S**ystem) is an inductively coupled plasma chamber with an internal antenna, [4]. The discharge chamber, shown in Figure 1, consists of a cylindrical, stainless steel vacuum vessel of internal diameter 20cm and length 90cm. A water cooled copper antenna with 11 turns, positioned along the axis of a 30cm long quartz tube, is inserted into a 5cm diameter port at one end of the chamber. The antenna is 10cm long, 4cm in diameter is surrounded in the quartz tube by air at atmospheric pressure and has no direct contact with the plasma. Three 70mm vacuum ports for diagnostic purposes are situated above the antenna. One port is used by a pressure gauge fitted with an interface for computer-based monitoring. Another port is used for an RF compensated Langmuir probe, which allows measurement of ion flux. The chamber is also equipped with another side port with a quartz glass for optical diagnostics. A fibre optic for an optical spectrometer is rigidly attached to that port. Argon and oxygen flows into the chamber are regulated by two mass flow controllers. The chamber is evacuated through a pumping port by a turbomolecular pump backed by a rotary pump and pressure is regulated by a system of a gate valve with a dedicated controller. Power is generated by an RF generator at frequency of 13.56MHz. To ensure an efficient power transfer into the plasma and minimum generation of reflected waves, the plasma system is supplied with a matching network consisting of a controller and a standard “L”-configuration matching unit. The matching unit comprises of two variable capacitors manipulated by DC servomotors. The matching controller tunes the capacitors so that the source impedance of 50Ω is matched with the plasma discharge impedance.

Actuation and data collection are centralised in an Intel Pentium 4-based PC containing an analog output board, an analog input board, and an RS232 serial interface. A schematic of the chamber and experimental set up is shown in Figure 2.

The chamber geometry is not suitable for an industrial application such as etching and the plasma chemistries have not been chosen with any particular application in mind. However, the simplicity of the configuration lends itself to experimentation and rapid adaptation.

B. Control Objectives and Controlled Variables

The main objective of this work is to derive a multivariable model for control of ion flux and atomic Oxygen concentration in an Argon/Oxygen plasma using RF power, gate valve position, and the flow rates of Argon and Oxygen as actuators. Due to certain restrictions on the chamber, the flow of O₂ is limited to 5sccm, while the flow of Ar is fixed at 100sccm. RF power can be manipulated in the range [120W, 500W]. The gate valve position is expressed in terms of residence time, since the latter has a truly physical meaning and can be easily linked to first-principle basics.

The RF-compensated Langmuir probe tip is positioned in the middle of the chamber in both the radial and axial

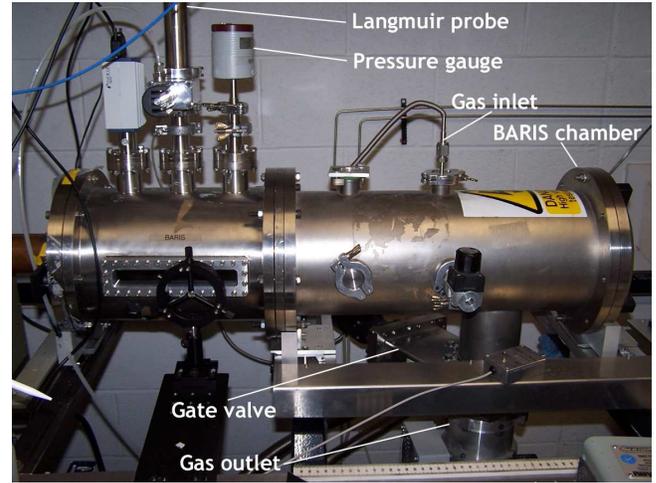


Fig. 1. BARIS Chamber

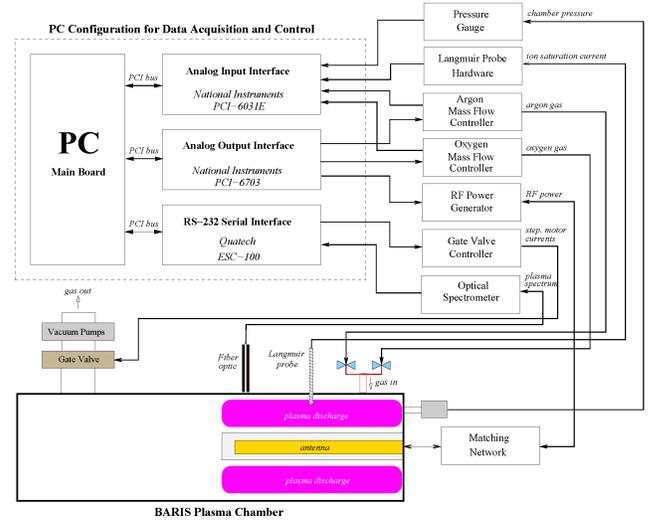


Fig. 2. Configuration for Data Acquisition and Control

directions. The probe is biased at -50V so that only positive ion saturation current is collected. Here, a factor of importance is the effective surface area of the probe tip, which in turn depends on the bias voltage. Since this factor is a constant, no calculations of the effective surface area of the probe tip were performed, thus ion flux was measured in mA rather than mA/cm².

Atomic Oxygen concentration was inferred using actinometry. Actinometry is a technique that gives a relative measurement of the concentration of atomic plasma species using only optical emission spectral (OES) data. Here, relative atomic Oxygen concentration is estimated by comparing the emission line intensity of an inert gas (in this case, Argon) with that of atomic Oxygen. In an Argon-Oxygen plasma, the relative atomic Oxygen concentration, [O], can be written

$$[O] = \frac{I(\lambda_{O1} = 777nm)}{I(\lambda_{Ar} = 750nm)} \quad (1)$$

where, $I(\lambda)$ is the emission intensity at wavelength λ . The Argon and Oxygen emission lines, $750nm$ ($3s^23p^54s - 3s^23p^54p$) and $777nm$ ($2s^22p^33s - 2s^22p^33p$), respectively, are well known actinometry lines, [5]. The commercial Ocean Optics USB2000 spectrometer was employed as a sensor in this work. Its optic fiber was positioned at one of the chamber view-ports where the light intensity reaches its maximum. An integration time of $40ms$ was selected as a trade-off between measurement noise and measurement frequency. A Windows-based software was specially developed for USB2000 and integrated into the existing data-acquisition configuration.

III. DEVELOPMENT OF A CONTROL-ORIENTED MODEL OF THE PROCESS

The selection of a model structure is an important step that precedes identification of a process model. A thorough preliminary analysis has been done on the studied plasma reactor that takes into account the plasma chemistry, RF power delivery, as well as the responses of actuators and sensors. A model structure has been devised based on the all available information about the process configuration. In order to reduce the complexity of the model, while preserving its accuracy within an acceptable level, several simplifications were assumed. First, the RF power generator has a fast response time, and therefore its dynamics can be neglected. Secondly, for the considered operating space, the plasma impedance variation is small and tuning of the matching network can be achieved fairly quickly regardless of the operating point. Power measurements have indicated that the reflected power is negligible compared to the forward RF power, i.e., maximum RF power is delivered to the plasma virtually at all times. Therefore, the dynamics of the matching network was not taken into account in the model. In addition, the Langmuir probe system dynamics and the optical spectrometer response time are negligible compared to the dynamics of the actuators, and can be reasonably ignored.

A. Modelling of the Actuators and Residence Time

There are two main actuators that determine the dynamical behaviour of the plasma process: The O_2 mass flow controller and the system gate valve - controller.

The input to the mass flow controller is an analog signal which corresponds to the desired flow rate, i.e. a setpoint signal. The mass flow controller is fitted with an analog output, which can be used for monitoring of the actual flow rate. A model of the mass flow controller was derived from a sequence of step responses from the analog input to the analog output using different step amplitudes and initial conditions. The responses indicated a consistent dynamical behaviour, which allows the use of a linear dynamical model. Linear models are normally expressed as *transfer functions*. A transfer function is the linear mapping of the Laplace transform of the model input(s) to the model output(s). The following model was derived for the mass flow controller:

$$M(s) = \frac{e^{-\tau_{dm}s}}{\tau_{mfc}s + 1} \quad (2)$$

In eqn. (2), τ_{mfc} is the time constant of the mass flow controller, while τ_{dm} is the actuator time delay. Both model parameters were identified using standard least squares identification routines, [6], giving $\tau_{mfc} = 0.069s$ and $\tau_{dm} = 0.123s$.

The chamber exhaust flow is controlled by a gate valve, which is fitted with an interface for computer control. The gate valve is driven by a stepper motor that is connected to a dedicated controller, which can be remotely programmed by a PC using the standard RS232 serial interface. The stepper motor works in *open-loop*, so no indication of the actual position is available. However, an estimate of the current valve position can be returned by the controller. Model identification of the gate valve system was performed in the same fashion as for the mass flow controller. A linear model that gives an adequate representation of the system was identified. The model has the same structure as the mass flow controller model, and can be expressed as:

$$G(s) = \frac{e^{-\tau_{dv}s}}{\tau_{gvc}s + 1} \quad (3)$$

where τ_{gvc} is the time constant of the gate valve system and τ_{dv} is the system time delay. The model parameters returned from the identification are $\tau_{gvc} = 0.091s$ and $\tau_{dv} = 0.27s$.

Residence time is a physical parameter of the plasma chamber, and is the average time a gas species spends in the chamber. Residence time is a function of the volumetric pumping rate and the area of the chamber exhaust, which in turn depends only on the gate valve position. For a particular gate valve position, the residence time, τ , can be calculated from the following equality:

$$\frac{d\hat{p}}{dt} = \frac{\hat{F}k_B T_N}{V} - \frac{\hat{p}}{\tau} \quad (4)$$

where k_B is Boltzmann's constant, \hat{p} is the chamber pressure, V is the volume of the chamber, T_N is the gas temperature and \hat{F} is the total throughput mass flow rate. The residence time can be determined by setting the left side of eqn.(4) to zero. A more accurate estimate can be provided by looking at the transient changes in pressure when the flow rate is stepped. The relationship $\hat{F} \rightarrow \hat{p}$ can be described by a first-order transfer function, such as

$$R(s) = \frac{\hat{p}(s)}{\hat{F}(s)} = \frac{k}{\tau s + 1} \quad (5)$$

In this notation, \hat{F} is the actual flow rate into the chamber achieved by the mass flow controllers. Dynamic estimation of the residence time was performed by stepping up and down the flow rate at different valve positions, from $\phi = 400$ to $\phi = 600$. A flow rate step size of $10sccm$ was used in this study, with an initial flow rate of $100sccm$.

B. Structure of the Model

The model structure proposed in this work is illustrated by the block diagram in Figure 3. In the diagram, P , F_{O_2} and τ are the set points for RF power, O_2 flow rate and residence time, respectively. The actual O_2 flow rate and

residence time are represented by \hat{F}_{O_2} and $\hat{\tau}$, respectively. Each of the controlled variables, ion flux, Γ_n , and atomic Oxygen concentration, $[O]$, can be expressed as functions of the manipulated variables:

$$\begin{aligned}\Gamma_n &= P f_{11}(\cdot) + F_{O_2} f_{12}(\cdot) + \tau f_{13}(\cdot) \\ [O] &= P f_{21}(\cdot) + F_{O_2} f_{22}(\cdot) + \tau f_{23}(\cdot)\end{aligned}$$

In eqn.(6)-(7), f_{ij} are generally nonlinear functions of the input variables. Each of the functions, however, can be separated into linear and nonlinear parts. The following separation can be made:

$$\begin{aligned}f_{11}(\cdot) &= k_{11}(\tau, F_{O_2}) \\ f_{12}(\cdot) &= k_{12}(P, \tau)M(s)R(s) \\ f_{13}(\cdot) &= k_{13}(P, F_{O_2})G(s)R(s) \\ f_{21}(\cdot) &= k_{21}(\tau, F_{O_2}) \\ f_{22}(\cdot) &= k_{22}(P, \tau)M(s)R(s) \\ f_{23}(\cdot) &= k_{23}(P, F_{O_2})G(s)R(s)\end{aligned}\quad (11)$$

Note that, in eqns.(8)-(13), k_{ij} are static nonlinear functions, while $M(s)$, $G(s)$ and $R(s)$ are linear dynamic functions. From eqn.(6)-(7) and (8)-(13), the model can be rewritten as

$$\underbrace{\begin{bmatrix} \Gamma_n \\ [O] \end{bmatrix}}_y = \underbrace{\begin{bmatrix} k_{11}(\cdot) & k_{12}(\cdot) & k_{13}(\cdot) \\ k_{21}(\cdot) & k_{22}(\cdot) & k_{23}(\cdot) \end{bmatrix}}_{\text{static nonlinearity}} \times \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{e^{-\tau_{dm}s}}{(\tau_{mf}e^s+1)(\hat{\tau}s+1)} & 0 \\ 0 & 0 & \frac{e^{-\tau_{dv}s}}{(\tau_{gvc}e^s+1)(\hat{\tau}s+1)} \end{bmatrix}}_{\text{linear dynamics}} \underbrace{\begin{bmatrix} P \\ F_{O_2} \\ \tau \end{bmatrix}}_u \quad (14)$$

The model expressed by eqn.(14) is referred to as a *Wiener model* [7] and can also be written in the following form:

$$y = N(u)L(s)u \quad (15)$$

where $N(u)$ represents the static nonlinearity, $L(s)$ is the transfer matrix, while y and u are the model outputs and model inputs, respectively. Note that $L(s)$ is a parameter-dependent function, since the residence time $\hat{\tau}$ is, in fact, a varying parameter. Also note that $N(u)$ and $L(s)$ do not commute.

IV. MODELLING OF THE STATIC NONLINEARITY

A series of measurements were performed in order to model the static nonlinearity, $N(u)$. Ion flux and atomic Oxygen concentration were measured as the RF power, O_2 flow rate and gate valve position were stepped using a grid of discrete levels. After each step, sufficient time was allowed for measured variables to stabilise, before they were recorded. The number of levels for each manipulated variable and the level spacing were selected by following general *design of experiment* (DOE) rules, [8], [9]. A $5 \times 5 \times 5$ equally-spaced grid was initially used to identify the degree of nonlinearity and to provide an appropriate distribution of evaluation points.

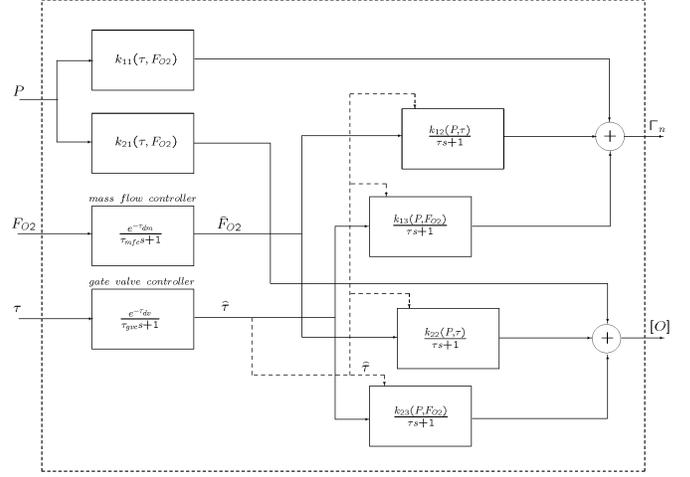


Fig. 3. Structure of the Plasma Model

A final $9 \times 9 \times 9$ grid spanning the entire operating space was selected:

$$\begin{aligned}P &: [120, 140, 160, 180, 200, 250, 300, 400, 500] \\ F_{O_2} &: [1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 5] \\ \tau &: [0.87, 0.74, 0.62, 0.49, 0.39, 0.33, 0.28, 0.25, 0.22]\end{aligned}$$

Various model structures were studied in order to find a parsimonious representation of the process. Here, complexity is defined with regard to the number of parameters that describe the relationship between the three manipulated variables (P , F_{O_2} , τ) and the two controlled variables (Γ_n , $[O]$). In this investigation, the number of parameters was arbitrarily limited to 30. The number of appropriate model structures that reasonably approximate the steady-state input-output relationship was narrowed down to three: an exponential function, a polynomial function and a neural network model. For the first two model representations, each process output was modelled independently. Fitting was performed in two steps: First, the Nelder-Mead simplex method, [10], was employed followed by the nonlinear least-squares method, [11], both using a *mean square error* (MSE) cost function. The non-gradient based Nelder-Mead method is well-known for handling non-smooth functions and generally returns a good suboptimal solution, reasonably close enough to the true optimum. In the second step, a nonlinear least-squares method was employed using the solution returned by the first method as an initial condition, thus producing an improved fit. Various artificial neural network configurations were considered, including radial basis function and *multilayer perceptron* (MLP) networks, [12]. MLP network configurations with a single hidden layer

and two hidden layers as well as “tansig” and “logsig” neurons were examined. The neural net structure that demonstrates a good compromise between model complexity and accuracy was selected. The best choice was an MLP network with two hidden layers featuring five “tansig” neurons in each hidden layer and a linear output layer.

The fitting performance of each model structure was assessed by looking at the MSE over the entire grid of measurements. The average percentage errors, between model and measurement in all three cases, are displayed in Table I, and the results clearly underline the performance superiority of the neural network model over the other two models. In addition, the neural network model can be easily inverted if that is required by a control design strategy, thus allowing the use of open-loop linearisation.

| Model type | Γ_n error [%] | [O] error [%] | Overall error [%] |
|-------------|----------------------|---------------|-------------------|
| Exponential | 2.28% | 3.91% | 6.19% |
| Polynomial | 1.55% | 2.32% | 3.87% |
| MLP (5+5+2) | 0.48% | 0.99% | 1.47% |

TABLE I
COMPARISON BETWEEN THE THREE MODEL STRUCTURES

V. MODEL VALIDATION

The validity of the complete Wiener model is now examined. Five sets of transient responses were recorded, each with a fixed duration of 300s, using a sample period of 10ms. The first 30s of the recorded data, containing the initial transient, was removed. In this analysis, the three manipulated variables were allowed to vary independently from one another, while the step sizes and step durations changed randomly within prescribed limits. The same input sequences were also passed through the model and the output sequences were recorded. A small 30s-section from one of the sets, best illustrating a large portion of the operating space, is shown in Figure 4. The top three graphs illustrate step changes in actuator set-points, while the bottom two show the measured and modelled ion flux and atomic Oxygen concentration. For illustration purposes, the time on the x-axis was adjusted to start from 0s. For the presented 30s time-frame, percentage MSE error was calculated between the measured controlled variables and their modelled counterparts. The errors for ion flux and atomic Oxygen concentration were 4.1% and 2.8%, respectively. Although both errors are larger than the errors indicated in Table I, the discrepancy can be explained with the variation of the measured variables over time. The recorded data in the remaining four sets of measurements was also analysed and the errors for ion flux and atomic Oxygen concentration were found to lie within $\pm 5.3\%$ and $\pm 5.6\%$, respectively. These variations are within acceptable limits and represent unmodelled effects. The graphs in Figure 4 demonstrate a good agreement between model and measurements.

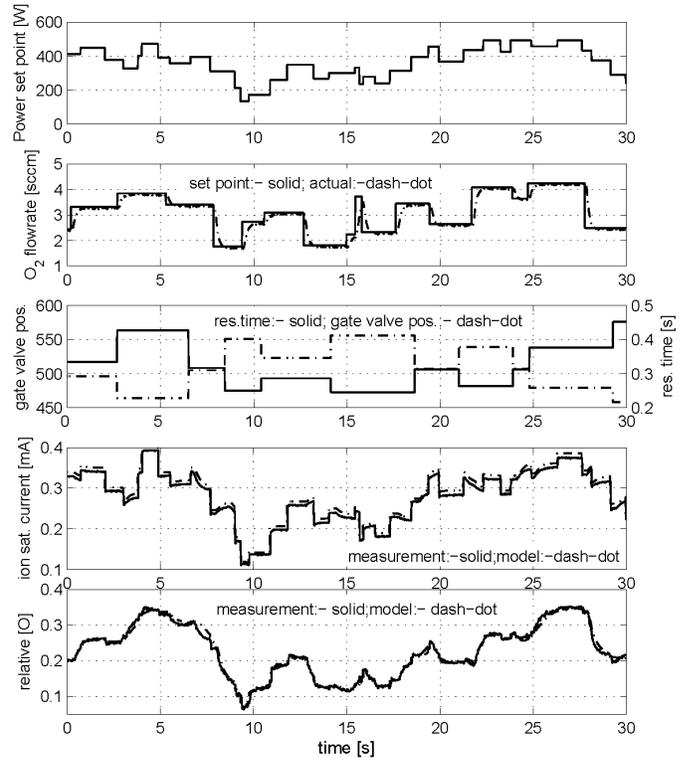


Fig. 4. Comparison between measurements and the MLP-based model

VI. CONCLUSIONS

In this work, a control-oriented model of an experimental plasma process has been derived. An efficient modelling approach is presented, which allow first-principles knowledge about the process dynamics to be fully exploited and modelled separately from the static nonlinearity. The valid separation of the model into a linear dynamic part and static nonlinearity can facilitate future control design, in particular, allowing open-loop linearisation to be employed. Various nonlinear structures were studied for the modelling of the static nonlinearity, and an MLP-based neural network model has achieved better approximation than exponential and polynomial functions, which are described by far more parameters. Analysis of the validation data has proven good agreement between the process and its model, with variations lying within acceptable limits.

ACKNOWLEDGMENT

This work has been undertaken within the PlasMAC project, a joint collaboration between National University of Ireland at Maynooth and Dublin City University (DCU), and supported by Science Foundation Ireland (SFI). The authors would like to thank Ronan Faulkner (NCPST, DCU) for his continuous help during the experimental work.

REFERENCES

- [1] T.A. Badgwell, T. Breeduk, S.G. Bushman, S.W. Butler, S. Chatterjee, T.F. Edgar, A.J. Toprac and I. Trachtenberg, Modelling and control of microelectronics material processing, *Computers Chemical Engineering*, Vol. 19, pp 1-41, 1995.
- [2] K.J. McLaughlin, T.F. Edgar and I. Trachtenberg, Real-time monitoring and control in plasma etching, *Proceedings of American Control Conference*, San Diego, California. pp 310, 1999.
- [3] B.A. Rashap, M.E. Elta, H. Eternad, J.P. Fournier, J.S. Freudenberg, M.D. Giles, J.W. Grizzle, P.T. Kabamba, P.P. Khargonekar, S. Lafortune, J.R. Moyne, D. Teneketzis and F.L. Terry Jr, Control of semiconductor manufacturing equipment: Real-time feedback control of a reactive ion etcher, *IEEE Transactions on Semiconductor Manufacturing*, Vol. 8(3), pp 286297, 1995.
- [4] D. Doherty and P. Iordanov, Preliminary Experimental Study of an Argon Plasma Process, Internal Report: EE/JVR/02/2005, National University of Ireland, Maynooth, Ireland, 2005
- [5] R.E. Walkup, K.L. Saenger and G.S. Sewyn, Studies of Atomic Oxygen in $O_2 + CF_4$ RF Discharges by Two-photon Laser-induced Fluorescence and Optical Emission Spectroscopy, *Journal of Chemical Physics*, Vol. 84, pp 2668-2674, 1986.
- [6] L. Ljung, *System Identification Toolbox*, MathWorks, 2004.
- [7] N. Wiener, *Nonlinear problems in random theory*, John Wiley, New York, 1958.
- [8] R.O. Kuehl, *Design of Experiments: Statistical Principles of Research Design and Analysis*, Brooks/Cole, 2000.
- [9] M.J Anderson and S.L. Kraber, Keys to Successful Designed Experiments, *American Society for Quality Conference*, Vol. 11, pp 1-6, 1999.
- [10] J.A. Nelder and R. Mead, A simplex method for function minimization, *The Computer Journal*, Vol. 7, pp 308-313, 1964.
- [11] D. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *SIAM Journal of Applied Math*, Vol. 11, pp. 431-441, 1963.
- [12] M. Nørgaard, O. Ravn, N.K. Poulsen and L.K. Hansen, *Neural networks for modelling and control of dynamic systems : a practitioners handbook*, Springer, 2000