Machine-Learning Optimization of Energy Efficiency in Microwave Applicators with Plasma

Williams, C.E. a,b, Murphy, E.K. c, Yakovlev, V.V. a

*Center for Industrial Mathematics and Statistics, Department of Mathematical Sciences, WorcesterPolytechnic Institute, Worcester, MA, USA
bDepartment of Physics, WorcesterPolytechnic Institute, Worcester, MA, USA
cThayer School of Engineering, Dartmouth College, Hanover, NH, USA

Abstract

With a widening range of applications of microwave plasma in processing technologies, there is a demand of computer-aided design of efficient and controllable applicators containing microwave plasma. The use of modern computational tools here is limited due to a principal uncertainty in data on the plasma medium for electromagnetic simulations. In this paper, based on simplified characterization of microwave plasma, we propose to embed a model of an applicator with plasma in a machine-learning optimization procedure aiming to maximize energy efficiency of the system. To demonstrate functionality of the procedure featuring the dynamic training of the decomposed radial-basis-function network, we use a model of a conventional cylindrical applicator with a coaxial quartz vessel containing plasma. The model is built in the finite-difference time-domain simulator QuickWave. For a neutral gas chosen, the procedure finds the geometry of the applicator and the plasma frequency satisfying the optimization goal. The network is readily reconfigurable, so the procedure can be easily reformulated for estimating geometrical parameters of other systems or physical parameters of the plasma and thus be practical in the design of microwave systems in scenarios lacking adequate data on plasma medium.

Introduction

Microwave (MW) plasma has shown a significant contribution to applications in processing technology [1] with demonstrable impact on the efficiency and quality of various processes. Examples include production of synthetic diamonds [1-3], surface processing for semiconductor manufacturing (deposition, etching, cleaning) [1], [4], plasma-based decomposition of various substances (including CO₂ [5] and water [6]), etc. However, advancement of these applications is constrained by the challenges associated with development of efficient and controllable MW applicators for industrial use [1].

While computer simulations aiding in the design of such applicators have been reported [2], [4, 7-12], the use of advanced electromagnetic (EM) modeling remains limited due to the absence of adequate input data for plasma in the system models. The direct measurement of complex permittivity of plasma is challenging [13, 14], and its properties are specific to the MW system in which it is ignited and maintained. Hence, the complex permittivity of plasma is mostly estimated by theoretical calculation [13, 15]. Recently, a simple physics-driven approach to characterization of MW plasma for EM modeling using the finite-difference time-domain (FDTD) technique was proposed in [16, 17]. In this approach, plasma is represented by the electric conductivity, which is determined by the plasma frequency \( f_p \) and the collision frequency \( \gamma \). Using this approach, in which \( f_p \) and \( \gamma \) are estimated based on typical values of the underlying physical parameters of plasma, modeling of a conventional MW applicator affirms well-known behavior of the electric field in presence of plasma [16, 17].

Further developing this work and embracing an apparent uncertainty in input data on plasma medium in EM models, in this paper we introduce a machine-learning (ML) approach to computational study of MW applicators with plasma. In the first implementation of the concept, an FDTD model is embedded in an ML optimization procedure, which allows one to find, for a particular gas, the geometry of the applicator and plasma frequency ensuring optimal performance. Optimality here means energy efficiency, and, accordingly, for the considered one-port applicator, the goal of optimization is a value of the reflection coefficient \( |S_{11}| \) below some threshold...

Keywords: Electric conductivity, Collision frequency, Neutral gas, Optimization, Plasma frequency, RBF network.
Plasma Model and an ML Procedure

Representation of MW plasma medium in an EM model requires knowledge of its electric conductivity \( \sigma \) [1, 21]:

\[
\sigma = \varepsilon'' \varepsilon_0 \omega E_0,  \tag{1}
\]

where

\[
\varepsilon'' = \frac{\omega^2 f_p}{\omega (\omega^2 + \omega_c^2)}.  \tag{2}
\]

\( \omega_p = 2\pi f_p \), \( \omega = 2\pi f \), \( f \) is the frequency of the EM field, and \( \varepsilon_0 \) is permittivity of vacuum. From (1), (2), it is seen that characterization of MW plasma may be reduced to determination of the plasma frequency and the collision frequency [16, 17]. Except the two fundamental constants, the mass of an electron (\( m \)) and the charge of an electron (\( q \)), the former is conditioned by electron density \( N \) [1, 22]:

\[
\omega_p = \sqrt{\frac{N q^2}{m \varepsilon_0}}.  \tag{3}
\]

While \( N \) is a function of physical system parameters such as power and pressure, the exact dependences are system specific and typically not known [16, 17]. Determining electron density is not a trivial task. To this end, researchers employ special techniques (such as versions of perturbation method [15, 23] and power measurement [24]) to directly find it or, when possible and sufficient, use some estimations (including based on corresponding modelling data [4]).

The collision frequency depends on the neutral gas density \( n_\text{e} \), the average electron velocity \( \varepsilon_\text{e} \), and the cross section of electron-neutral particle collision \( \sigma_{\text{eN}} \) [25-27]:

\[
\gamma = n_\text{e} \varepsilon_\text{e} \sigma_{\text{eN}}.  \tag{4}
\]

While \( n_\text{e} \) also depends on physical parameters of a MW system [16, 17], in [4] it is calculated using the ideal gas equation and, for low pressure plasma, is determined to be \( 1.94 \times 10^{22} \) (1/m³). \( \varepsilon_\text{e} \) is found in [4].

This work is licensed under a Creative Commons Attribution 4.0 Licence.

https://creativecommons.org/licenses/by/4.0/
\[
X = \{G_1, G_2, \ldots, G_N, f_p\}, \quad \text{where} \quad G_j (j = 1, \ldots, N) \quad \text{are geometrical parameters of the modeled applicator, and} \quad i = 1, \ldots, P, \quad \text{where} \quad P \quad \text{is the number of points (input-output pairs) of modeling data. The network output vectors} \quad Y_j = \{S_1|1\}, \ldots, S_1|K]\] obtained by taking \( K \) equally spaced values of frequency characteristics of the reflection coefficient \( |S_1| \) over a specified frequency range. The network uses linear RBFs, i.e.,

\[
\Phi_{i,j} = \|X_i - C_j\|, \quad (6)
\]

where \( \Phi_{i,j} \) is the system matrix and \( C_j \) is the \( j \)th of \( \eta \) centers. The RBF centers are chosen to match the sampled points, which results in zero training error \( (\eta = P) \), as the training matrix is full rank. The network is coupled with a linear model, and the weights are constructed by solving the corresponding linear system. Choosing RBF centers equal to the training points is a convenient programming solution, known to be practical for dimensional input domains that are not excessively high [29]. To reduce concerns for overfitting, similarly to [18], a regularization term, \( \lambda = 10^{-3} \), is included when solving the linear system. The optimal geometrical parameters of the system \( G_i \) and the plasma frequency \( f_p \) are, therefore, determined for the chosen gas.

In the optimization procedure implemented in a MATLAB code, training/testing data for the network \( (X \text{ and } Y) \) are generated by QuickWave. A MW plasma applicator to be optimized is represented by a fully parameterized model. Following [18], the procedure is realized in two separate modules responsible for database (DB) generation and network operations. At the first stage, the initial DB is created. The user selects the number of FDTD time-steps necessary to converge to steady state and the stopping condition limiting optimization operations at the second stage (i.e., either maximum CPU time, or maximum DB size). The user also enters the intervals of variation for all the design variables and the number of subdivisions.

At the second phase, the user sets optimization constraints – the frequency of interest and the upper tolerance for the value of \(|S_1|\) at that frequency. Next, the user launches the iterative procedure of dynamic training/optimization [18]. Immediately after the initial database is constructed, the weights of the RBF network are computed, and the resulting RBF network is minimized using standard minimization techniques using a nonlinear constraint and a parameter \( \beta \) \( (0 \leq \beta \leq 1) \). Choosing various values of \( \beta \) ensures the procedure does not get trapped within local minima. For each \( \beta \) value, a minimum \( (i) \) is determined, \( (ii) \) simulated using QuickWave, and \( (iii) \) added to an updated database. The procedure essentially continues in this fashion; iteratively building a larger and larger database and a more and more accurate RBF network to approximate the true mapping. The procedure outputs the results as they are produced.

One particular advantage of using neural networks in this context is also worth mentioning. In [19], an RBF network efficiently worked within a general global optimization algorithm designed to solve expensive black-box functions. The crucial aspect of the approach to microwave optimization [18] that employed the technique [19] was in the use of a global optimization to achieve a similar goal – to limit the number of expensive full-wave FDTD analyses of microwave systems. In contrast to neural networks which have recently become common in a variety of deep learning and AI algorithms, the approach [18] and its replication in the present paper for modeling microwave systems with plasma are much simpler and, in both implementations, they appear to be convenient practical tools for controlling an FDTD solver, processing the data, and finding an optimum with small FDTD data sets.

### Optimization Results

In the illustrative optimization described in this section the applicator in Fig. 2 is characterized by constant parameters \( t = 1.5 \text{ mm}, l = 220 \text{ mm}, \) and \( L = 240 \text{ mm} \) and is fed by WR284 oriented either cross-sectionally, or longitudinally. The CAD goal is to find an inert gas, the values of plasma frequency and of two geometrical parameters \( (G_1 = D \) and \( G_2 = d) \) that provide a minimum value of \(|S_1| \leq 0.1 \) at \( f_0 = 2.45 \text{ GHz} \). In the plasma model [16, 17], an inert gas defines the collision frequency, so in this illustration we consider eight typical gases with their values of \( \gamma_1 \) as specified in Table 1. This reduces the analysis to eight optimization problems with three design variables, \( D, d, \) and \( f_p \), for which the following specifications are applied:

\[
90 \leq D \leq 120 \text{ mm}, \quad (7)
\]

\[
40 \leq d \leq 70 \text{ mm}, \quad (8)
\]

\[
0.4 \leq f_p \leq 8.0 \text{ GHz}. \quad (9)
\]

The bounds for \( f_p \) in (9) correspond to the values of plasma density of the order of \( 10^{16} - 10^{18} \text{ m}^{-3} \) and, as such, are consistent with the value of \( n_e \) accepted in (4) as a parameter of a low-pressure plasma.

The procedure starts with an initial database of 27, by choosing 3 subdivisions for each variable, and the stopping criterion is set for the database size reaching 150 points. The FDTD model features a fine uniform mesh with 1.5 mm cells. When run on an advanced Windows workstation, the steady state is reached in less than 2 min.
Optimized frequency responses of $|S_{11}|$ are shown in Figs. 3-4 for the systems with the longitudinal and cross-sectional orientations of WR284, respectively.

Each figure contains eight curves for the considered inert gases as well as examples of three non-optimized characteristics for comparison. In 7 runs out of 16, the optimal solutions satisfying the goal $|S_{11}| < 0.1$ were found: generation of the DB points and training of the network were over then before reaching the stopping criterion. For the other 9 cases, with 150-point DB’s, the goal was not reached, but the best solutions were close to it ($|S_{11}| = 0.1 \pm 0.22$). Moreover, for cases 2, 3, 4 in the system in Fig. 2(a) and for cases 1, 2, 3 in the system in Fig. 2(b), the optimal values of the design variables were found at the upper limits of those intervals. The results imply that it might be possible to reach “better” values of $|S_{11}|$ within some wider intervals for the design variables. The fact that the desirable values of $|S_{11}| < 0.1$ were not reached in those 9 cases is an indicator that in the chosen solution space those desirable values cannot be achieved.

![Diagram](image)

**Fig. 2.** 2D (a) and 3D (b), (c) views of the modelled cylindrical cavity with a coaxial quartz vessel (containing plasma) and an input rectangular waveguide oriented longitudinally (b) and cross-sectionally (c) oriented rectangular waveguide.

The shapes of the non-optimized curves (obtained from randomly chosen values of the design variables) are fully consistent with the results in [17], where these characteristics are exemplified and appear to be strongly dependent on $\gamma$ and $f_p$. By including these curves in Figs. 3-4, we demonstrate that, for the considered microwave system, the optimized geometry results in significant improvement of the theorized energy efficiency. In accordance with Fig. 4, improvement can be estimated as going from 35-50% to 95-99%.

Computationally, the developed procedure demonstrates excellent performance by quickly converging to the “best solutions”. This is...
apparantly ensured by the underlying ML algorithm [18] known for quick finding local optima (satisfying the constraints) with small FDTD modeling data sets. Table 4 shows that, when dealing with three design variables of the MW applicator in Fig. 2, the procedure needs, depending on the gas, 48 to 151 DB points to find characteristics of $S_{11}$ corresponding to energy efficiency at least 95%.

**Discussion and Conclusions**
The ML procedure employing the CORS sampling in the dynamic training of the decomposed RBF network has been introduced for determining parameters of MW applicators with plasma that lead to a minimum reflection (or maximum energy efficiency). The procedure demonstrates its robustness and computational effectiveness when working with an FDTD model of a simple conventional system but is expected to successfully optimize other complex applicators characterized by more geometrical parameters as well. The FDTD model behind the data used for optimization, however, relies on the simplified characterization of conductivity of plasma. Consequently, each optimized frequency characteristic of the reflection coefficient (such as in Figs. 3-4) is found for a concrete gas and certain geometrical parameters of the applicator and plasma frequency.

An RBF network used in the core of the optimization here provides certain benefits (e.g., a zero-training error discussed above), yet its overall performance is outside of the scope of this study as somewhat secondary factor to the overall CORS RBF algorithm. The main advantage of the latter is in efficient handling of expensive black-box functions [19] and the ability to sharply reduce the number of full-wave FDTD analyses [18].

The results in the previous section emphasize the benefits of using neural networks in the key pursuits of this study. The optimization technique with a simple RBF network embedded into it [18] excels here in reducing the number of full-wave FDTD analyses of microwave systems with plasma. In response to the principal uncertainty of data on the plasma medium, these results can be seen as a favorably argument in the use of a ML algorithm. It is apparent that the approach used in the presented illustrative optimization could potentially aid in the design of microwave applicators with plasma.

More specifically, the reported optimization results also suggest the following. When using the proposed optimization technique in CAD of microwave applicators with plasma, one can either be satisfied with the obtained local minimum, i.e., the geometry corresponding to the value of $S_{11}$ achieved with the considered solution space, or expand it aiming to find another optimum providing alternative geometry for a lower value of the reflection coefficient. The exploration for better results can also include introduction of additional design variables $G_j$; e.g., in the systems considered here this is encompassed by $L$ and $l$.
While, in practice, the geometry may be directly adopted for a prospective physical prototype with a chosen gas, realization of optimal value of $f_p$ depends on special factors (such as the ability to measure (and ultimately control) electron density) and remains outside the scope of this study. Here, we are concerned with the concept of an ML approach to computational study of MW applicators with plasma given the intrinsic uncertainty in characterization of plasma in EM models. Our optimization procedure is based on representation of plasma conductivity via $\gamma$ and $f_p$ and, since $\gamma$ is inherent to the gas (and hence can be considered as given), we are only concerned with an optimal value of $f_p$. The RBF network can be easily adapted to alternative more sophisticated physics-driven plasma models or to the models based on data collected from measurements (and processed by other ML procedures). Overall, with the uncertainties input data for the EM models, our optimization procedure can be instructive in determining the operational bounds of applicators with MW plasma and therefore assist in their CAD.

The principal uncertainty of input data on the plasma medium also implies that simulation results produced by the plasma model [16, 17] are, by definition, more or less inadequate for their traditional use. However, they are not intended for conventional direct comparison with measurements. Rather, they are more appropriate for estimating characteristics of microwave applicators and computational studies of their possible behaviour in the initial stages of applicator design (presuming that experimental tuning of their performance will follow).

In these circumstances, the optimization technique outlined in this paper is to help handle this data uncertainty. We have described a procedure seeking best geometrical parameters and plasma frequency, but the problem can be easily reformulated (and the network straightforwardly reconstructed) to focus on other parameters of interest. With data on some plasma or system parameters more defined than others, the ML procedure could be used to approximate such parameters or find them belonging to some plausible ranges. As an element of applicable CAD, this could our help shorten the timeframe between design and efficiently operating physical prototype.

Acknowledgments
The work of C.E.W. was partially supported by the 2022 WPI Summer Training in Arts & Sciences Research (STAR) Fellowship.

Conflicts of Interest
The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References
10. Zhang, W.; Tao, J.; Huang, K.; Wu, L. Numerical investigation of the surface wave