

## **AN SUPPORT VECTOR REGRESSION BASED NONLINEAR MODELING METHOD FOR SiC MESFET**

**Y. Xu, Y. Guo, R. Xu, L. Xia, and Y. Wu**

School of Electrical Engineering  
University of Electronic Science and Technology of China (UESTC)  
Chengdu, China

**Abstract**—An approach for the microwave nonlinear device modeling technique based on a combination of the conventional equivalent circuit model and support vector machine (SVM) regression is presented in this paper. The intrinsic nonlinear circuit elements are represented by Taylor series expansions, coefficients of which are predicted by its support vector regression (SVR) model. Example of a SiC MESFET nonlinear model is demonstrated, and good results is achieved.

### **1. INTRODUCTION**

SiC MESFETs are popular devices for power amplifier design in high-power and high-temperature applications because of SiC's superior properties, such as high breakdown voltage, high thermal conductivity, and high saturated electron velocity [1]. The development of SiC devices in wireless applications provides the impetus of researching in the area of nonlinear modeling, which is useful for device performance analysis in designing microwave circuits and characterizing the device technological process.

In general, models for microwave nonlinear devices belong to two categories: physical and empirical. The empirical models, including closed-form equation models (equivalent circuit models) [15, 16] and look-up table models, are by far the most commonly used in nonlinear CAD [2]. Recently, The artificial neural network (ANN) based modeling methods have showed superiorities, while physical-based models are computational cost and lack of accuracy, closed-form models are in defect of robust for different type devices, and look-up table models require a large amount of data [3]. However, although many advantages owned, ANN modeling methods have drawbacks as well. For example, it is difficult to determine the proper ANN

configurations, and on-convex quadric minimization may result in multiple minima [4].

A new microwave active device nonlinear modeling technique based on the combination of the conventional closed-form equation models and support vector machine (SVM) is proposed. Different from ANN, the SVM is based on structural risk minimization (SRM) principle and resolving convex quadratic program (QP), which shows more powerful generalization ability than ANN [5,6]. In this paper, example of SiC MESFET nonlinear modeling utilizing the proposed support vector regression (SVR) based modeling technique are demonstrated. The main frequency independent intrinsic nonlinear elements (source-drain current  $I_{ds}$ , nonlinear capacitances  $C_{gs}$  and  $C_{gd}$ ) are, firstly, modeled by SVR. By using Taylor series expansion of intrinsic nonlinear elements, their coefficients are predicted by related SVR models. With this method, the nonlinear characteristic of SiC MESFET can than be good described in numerical way while preserving the original physical meaning of closed-form equation models.

The organization of this paper is as follows: Section 2 summarizes the theory of support vector regression. Section 3 describes the structure of SiC MESFET used in this paper and the proposed SVR-based nonlinear modeling technique. Section 4 shows the results by using proposed method applied to SiC MESFET.

## 2. SUPPORT VECTOR REGRESSION

SVMs are state-of-the-art tools for linear and nonlinear input-output knowledge discovery [7,8]. Given a training dataset  $(y_i, x_i)$ ,  $i = 1, 2, \dots, n$ ,  $x_i \in R^m$ ,  $n$  is the size of training data. SVR tries to find the mapping function  $f(x)$  between the input variable and the desired output variable. Traditional regression method find the regression function  $f(x)$  by the rule of empirical risk minimization principle, i.e., minimize:

$$R_{emp}[f] = \frac{1}{n} \sum_{i=1}^n L(f(x_i) - y_i) \quad (1)$$

with  $L(x, y, f) = |y - f|_\varepsilon = \max\{0, |y - f| - \varepsilon\}$ .  $L(f(x_i) - y_i)$  represents the error function,  $\varepsilon$  is the insensitive loss function.  $y_i$  is real value,  $f(x_i)$  is the prediction value.

However, the actual risk minimization can not be realized only with the empirical risk minimization. A typical example is the overfitting of ANN. Support vector regression method based on SRM

principle, which minimize the following cost function:

$$\frac{1}{2}\|w\|^2 + C \cdot R_{emp}[f] \quad (2)$$

where  $\frac{1}{2}\|w\|^2$  is the term characterizing the modeling complexity.  $C$  is a regularization which determines the trade off between model complexity and empirical loss function. After some reformulations and introduction of the slack variables:  $\xi_i, \xi_i^*$ . Equation (2) is transformed into primal problem:

minimize:

$$\frac{1}{2}\|w\|^2 + C \cdot \frac{1}{n} \sum (\xi_i + \xi_i^*) \quad (3)$$

subject to:

$$\begin{cases} (w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i \\ y_i - (w \cdot x_i) + b \leq \varepsilon + \xi_i^* \\ \xi_i > 0, \quad \xi_i^* > 0, \quad \varepsilon > 0. \end{cases}$$

According to [5], an improved SVR has been presented, Equation (3) can be changes to minimize:

$$\begin{aligned} \min \phi(w, b) &= \frac{1}{2}\|w\|^2 + C \left( \nu\varepsilon + 1/l \sum_{i=1}^l L_i \right) \\ \text{s.t. } y_i &= \langle w, x_i \rangle + b \geq 1 - L_i, \quad L_i \geq 0, \quad \forall i \end{aligned} \quad (4)$$

where  $C$  (penalty parameter) is a regularization which determines the trade off between model complexity and empirical loss function,  $\varepsilon$  is tolerance of termination criterion, and  $\nu$  ( $0 < \nu < 1$ ) is a constant.

Introducing Lagrange multipliers to solve this problem of convex optimization and making some substitutions, we arrive to the Wolfe dual of the optimization problem:

maximize:

$$W(\alpha, \alpha^*) = \sum (\alpha_i - \alpha_i^*)y_i - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j) \quad (5)$$

subject to:

$$\begin{cases} \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \\ \alpha_i \in \left[0, \frac{C}{n}\right] \\ \alpha_i^* \in \left[0, \frac{C}{n}\right] \\ \sum_{i=1}^n (\alpha_i^* - \alpha_i) \leq C \cdot \nu \end{cases}$$

In order to expand the method to nonlinear decision functions, the input space projects to another higher-dimensional dot product space  $F$ , called feature space, via a nonlinear map  $\varphi: R^m \rightarrow F^d (d \gg m)$ . In this new space the optimal hyperplane is derived. Nevertheless, by using kernel functions which satisfy the Mercer' theorem, it is possible to make all the necessary operations in the input space by using  $\langle \varphi(x_i), \varphi(x_j) \rangle = K(x_i, x_j)$ . The regression estimation function is formulated in terms of these kernels:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (6)$$

where  $a_i$  and  $a_i^*$  are Lagrange multiples,  $K(x_i, x)$  is the kernel function.  $K$  is a symmetric positive definite function, which satisfies Mercer's condition.

It is easy to demonstrate that  $\langle \varphi'(x), \varphi(y) \rangle$  is given the derivative of with  $K(x, y)$  respect to  $x$  [9]. Accordingly, the  $n$ th-order derivatives of  $f(x)$  is

$$g^{(n)}(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K^{(n)}(x_i, x) \quad (7)$$

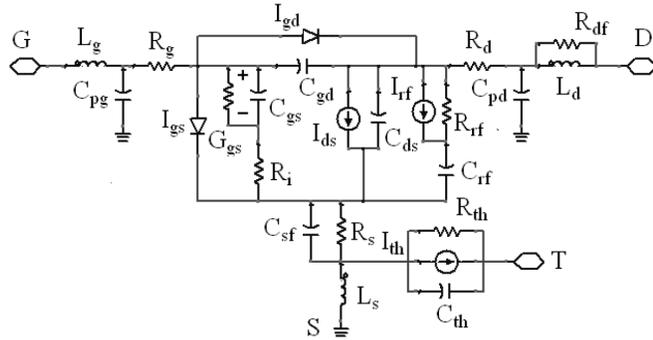
This developed SVR model can then be used to predict outputs for given inputs that were not included in the training data, and the  $n$ th-order derivatives of outputs.

### 3. NONLINEAR MODELING OF SIC MESFETS BASED ON SVR TECHNIQUES

An  $1 \mu\text{m} \times 300 \mu\text{m}$  4H-SiC MESFET is modeled in this paper. This device consists of a  $0.15\text{-}\mu\text{m}$  cap layer ( $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ ),  $0.35\text{-}\mu\text{m}$  channel layer ( $N_d = 1.7 \times 10^{17} \text{ cm}^{-3}$ ), and  $2\text{-}\mu\text{m}$  buffer layer

( $N_d = 1.5 \times 10^{15} \text{ cm}^{-3}$ ) on a semi-insulated 4H-SiC substrate. The buffer layer can prevent damage and deep level impurities in the substrate from the active layer. Due to the lack of p-doping source at the moment, we use an unintentionally weak n-type layer as the buffer in order to minimize the influence of substrate. The caculated cut-off frequency ( $f_T$ ) and maximum frequency of oscillation ( $f_{\text{max}}$ ) are 6.7 GHz and 25 GHz, respectively.

The basic topology of the empirical model for SiC MESFETs is shown in Fig. 1. The Angelov non-linear gate capacitances model ( $C_{gs}$ ,  $C_{gd}$ ) and a modified drain-source current ( $I_{ds}$ ) equation based on Angelov model is used, and detail information about empirical model are showed in [2]. The model has been implemented into ADS as a user-defined model, and its validity has been proved in predicting electrical performance of SiC MESFETs. The simulated  $S$ -parameters of empirical model at different bias will construct the data needed for a SVR model. The simulation has been accomplished by ADS with directly  $50 \Omega$  input-output terms.



**Figure 1.** Equivalent circuit of the empirical model for SiC MESFETs.

Only there main intrinsic nonlinear element ( $C_{gs}$ ,  $C_{gd}$ , and  $I_{ds}$ ) will be considered in our method for simple demonstration. Similar to ANN methods [3], each main intrinsic nonlinear element can be modeled by independent SVR models

$$C_{gs} = f^{SVR}(V_{ds}, V_{gd}) \quad (8)$$

$$C_{gd} = f^{SVR}(V_{ds}, V_{gd}) \quad (9)$$

$$I_{ds} = f^{SVR}(V_{ds}, V_{gs}) \quad (10)$$

Because the SVM is based on SRM principle and resolving convex QP, the SVR have a good ability to compromise model complexity and accuracy, and can get the global minimum results. The nonlinear

model can be carried out by using the similar way of ANN methods [3]. However, the physical mechanism of nonlinear model is still not clearly enough. Here, a more physical nonlinear modeling method is proposed based on above SVR models.

Typically, the transistor is polarized in a bias point  $(V_{ds0}, V_{gs0})$ , and the incremental drain-to-source and gate-to source voltages,  $v_{ds}$  and  $v_{gs}$ , are applied over this DC polarization. With these premises, it is necessary to accurately reproduce the nonlinear elements  $f(V_{ds}, V_{gs}) = F(V_{ds0}, V_{gs0}, v_{ds}, v_{gs})$  dependence and its derivatives with respect to the incremental voltages. In applications of amplifiers and mixers, the usual is to consider up to the third order intermodulation distortion (IMD) [11]. As a result, the nonlinear elements  $Q_g(V_{ds}, V_{gd})$  can be expressed as Taylor series expansion

$$\begin{aligned} Q_g(V_{gs}, V_{gd}) = & Q_g(V_{gs0}, V_{gd0}) + C_{gs1}v_{gs} + C_{gd1}v_{gd} + C_{gs2}v_{gs}^2 \\ & + C_{gs2gd}v_{gs}v_{gd} + C_{gd2}v_{gd}^2 + C_{gs3}v_{gs}^3 + C_{gs2gd}v_{gs}^2v_{gd} \\ & + C_{gs2gd}v_{gs}v_{gd}^2 + C_{gd3}v_{gd}^3 \end{aligned} \quad (11)$$

Accordingly, the  $C_{gs}(V_{ds}, V_{gs})$ ,  $C_{gd}(V_{ds}, V_{gs})$  and  $I_{ds}(V_{ds}, V_{gs})$  expansions are as follows:

$$\begin{aligned} I_{ds}(V_{ds}, V_{gs}) = & I_{ds}(V_{ds0}, V_{gs0}) + G_m v_{gs} + G_d v_{ds} + G_{m2} v_{gs}^2 \\ & + G_{md} v_{gs} v_{ds} + G_{d2} v_{ds}^2 + G_{m3} v_{gs}^3 + G_{m2d} v_{gs}^2 v_{ds} \\ & + G_{md2} v_{gs} v_{ds}^2 + G_{d3} v_{ds}^3 \end{aligned} \quad (12)$$

$$\begin{aligned} C_{gs}(V_{gs}, V_{gd}) = & C_{gs1} + 2C_{gs2}v_{gs} + C_{gs2gd}v_{gd} + 3C_{gs3}v_{gs}^2 \\ & + 2C_{gs2gd}v_{gs}v_{gd} + C_{gs2gd}v_{gd}^2 \end{aligned} \quad (13)$$

$$\begin{aligned} C_{gd}(V_{gs}, V_{gd}) = & C_{gd1} + 2C_{gd2}v_{gd} + C_{gs2gd}v_{gs} + 3C_{gd3}v_{gd}^2 \\ & + 2C_{gs2gd}v_{gs}v_{gd} + C_{gs2gd}v_{gs}^2 \end{aligned} \quad (14)$$

where  $Q_g(V_{ds0}, V_{gs0})$  and  $I_{ds}(V_{ds0}, V_{gs0})$  is the static DC values at bias point, and  $(G_m, \dots, G_{d3}, C_{gs1}, \dots, C_{gd3})$  are coefficients related to the  $n$ th-order derivatives valuated at the bias point.

Different from previous work on  $I_{ds}(V_{ds}, V_{gs})$  [2, 12], the coefficients of each nonlinear elements are directly predicted by SVR  $n$ th-order derivatives models for each nonlinear elements. Benefit to the great generalization ability of SVM, the coefficients of nonlinear elements model can be accurately extracted. For example,

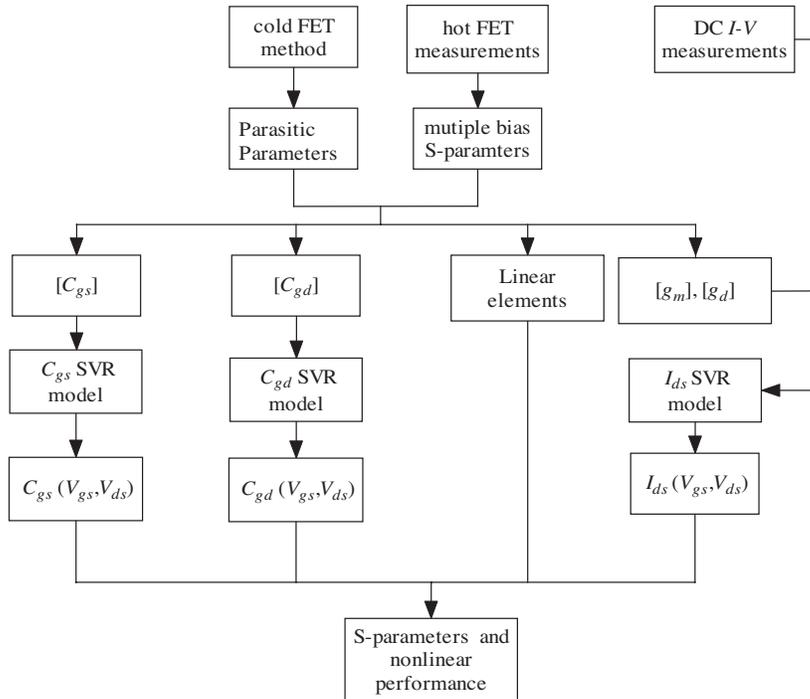
$$G_m = \frac{\partial I_{ds}^{SVR}(V_{gs}, V_{ds})}{\partial V_{gs}} \quad (15)$$

$$C_{gsd} = \frac{\partial C_{gs}^{SVR}(V_{gs}, V_{gd})}{\partial V_{gd}} = \frac{\partial C_{gd}^{SVR}(V_{gs}, V_{gd})}{\partial V_{gd}} \quad (16)$$

$$C_{gd3} = \frac{1}{3} \frac{\partial^2 C_{gd}^{SVR}(V_{gs}, V_{gd})}{\partial V_{gd}^2} \quad (17)$$

$$C_{gs1} = C_{gs}^{SVR}(0, 0)$$

The detailed flow chat is showed in Fig. 2, where  $[C_{gs}]$ ,  $[C_{gd}]$ ,  $[g_m]$  and  $[g_d]$  represents the calculated discrete values at different bias.



**Figure 2.** The proposed SVR based nonlinear modeling flow chat.

The modeling technique can be carried out with the following procedures:

- a) Measurement of the dc  $I$ - $V$  and the multiple biased  $S$ -parameter for a microwave device.
- b) Extraction of the parasitic parameters of microwave device using cold FET method.
- c) Extraction of the intrinsic capacitances,  $g_m$  and  $g_{ds}$  by using the method proposed by Dambrine [13].

- d) Modeling of the intrinsic capacitances and dc current by using SVR technique.
- e) Reconstruction of nonlinear elements expression by calculation of the  $n$ th-order derivatives of nonlinear elements SVR models.
- f) Calculation of the  $S$ -parameters and nonlinear performance by using the nonlinear elements in the nonlinear circuit simulator.

#### 4. MODEL VERIFICATION

LIBSVM-matlab code [14] is used as a basis to implement SVR model.  $v$ -SVR based on radial basis function (RBF) kernel function has been considered in our regression experiments. The parameters ( $\varepsilon$ ,  $\nu$ ,  $C$  and  $\gamma$ ) are extracted by trying with different variable value. The quality of each model is evaluated as its prediction accuracy, measured by mean squared error ( $MSE$ ):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2 \quad (18)$$

$x_i$  is the value of simulated  $S$ -parameters of empirical model,  $y_i$  is the SVR model predicted value and  $N$  is the number of validation data.

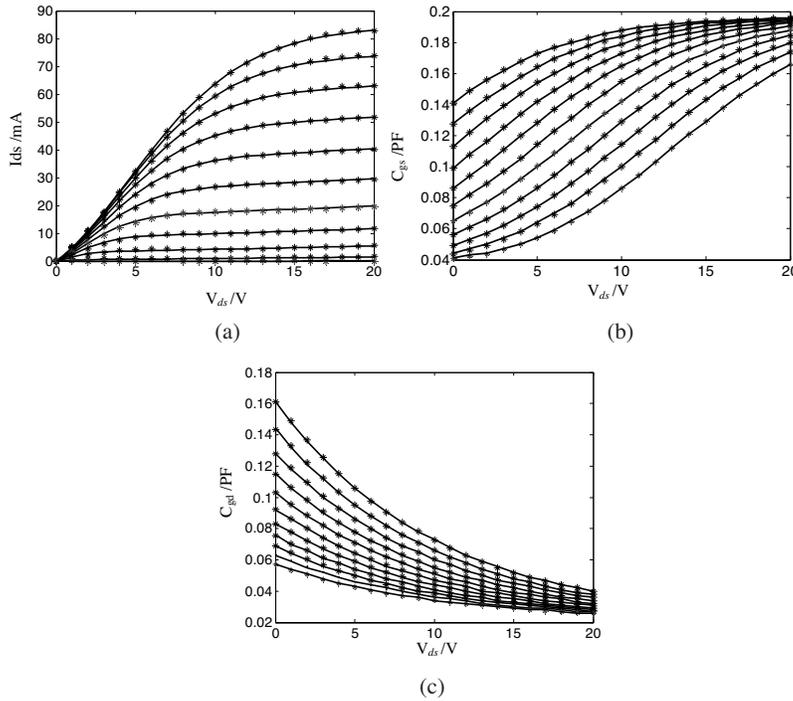
Table 1 showed the SVR variables and samples for training and test data. The same variables (bias points) are selected for each nonlinear element. And the total number of training data is 30 points, which means only 30 sets of  $S$ -parameters and DC I-V data are needed for build a nonlinear model with this method. And the samples are selected with the same steps, which mean the SVR model

**Table 1.** SVR variables and samples selection.

Data	Training Data			Testing Data		
<i>Para.</i>	Min	Max	STEP	Min	Max	Step
$V_{ds}/V$	0	20	4	0	20	1
$V_{gs}/V$	-15	0	-3	-15	0	-1.5

**Table 2.** Results of SVR model.

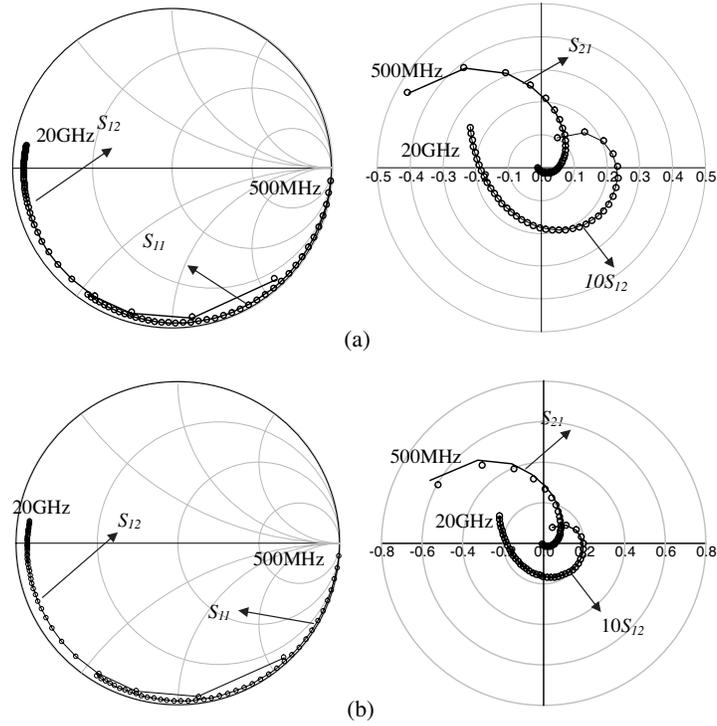
Data	Training Data			Testing Data		
<i>Para./Unit</i>	$I_{ds}/mA$	$C_{gs}/pF$	$C_{gd}/pF$	$I_{ds}/mA$	$C_{gs}/pF$	$C_{gd}/pF$
$MSE$	4.3e-2	4.4e-9	3.2e-9	8.5e-2	1.5e-7	1.7e-7



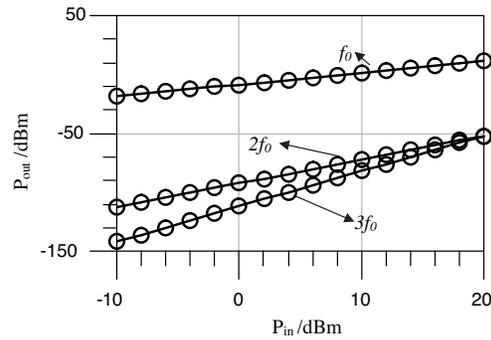
**Figure 3.** The plots of predicted results vs. testing samples. (a) Plot of predicted  $I_{ds}$  results (solid line) vs. testing samples (star), (b) Plot of predicted  $C_{gs}$  results (solid line) vs. testing samples (star), (c) Plot of predicted  $C_{gd}$  results (solid line) vs. testing samples.

is a robust model that with little dependence of the training samples distribution while the results of ANN models are greatly depend on sample selection. Table 2 shows the  $MSE$  results of training and testing data. And Fig. 3 is the plots of predicted results vs. testing samples for each nonlinear element. The results reveal that the SVR model can accurately predict each nonlinear element in less than 1 minute on an Intel Pentium IV 3.0GHz with 1GB of memory and running Windows XP. Besides, it is great convenient that the parameters ( $\varepsilon$ ,  $v$ ,  $C$  and  $\gamma$ ) for each SVR model are the same. It means only four parameters are needed for the nonlinear model, while ANN based method usually require several hundred parameters [3].

Figure 4 and Fig. 5 shows the comparison of  $S$ -parameters and harmonic performance by using proposed method and simulated data using empirical model. The reason of small  $S_{21}$  and output power



**Figure 4.** Comparison of  $S$ -parameters between the empirical model (circles) and the SVR based model (solid line) of SiC MESFET in the frequency range of 500 MHz–20 GHz. bias: (a)  $V_{gs} = -2$  V,  $V_{ds} = 20$  V, (b)  $V_{gs} = -7$  V,  $V_{ds} = 20$  V.



**Figure 5.** Comparison of harmonic performance between the empirical model (circles) and the SVR based model (solid line) for SiC MESFET ( $f_0 = 2$  GHz) at bias  $V_{gs} = -7$  V,  $V_{ds} = 20$  V.

( $P_{out}$ ) are that the simulations were accomplished by directly 50  $\Omega$  input-output terms and the simulation frequency is far over  $f_T$ . The data are selected for demonstration purpose of the proposed method, and the good agreements show the validation of the proposed method.

## 5. CONCLUSION

An approach for the microwave nonlinear device modeling technique based on a combination of the conventional equivalent circuit model and support vector machine (SVM) regression is presented in this paper. The example of SiC MESFET modeling shows that the proposed method can provide fast and accurately modeling, whilst preserve the advantages of closed-form equation model. This technique is very useful for device performance analysis in designing microwave circuits and characterizing the device technological process for relatively new compound semiconductors such as GaN and SiC.

## REFERENCES

1. Charles, E. W., "Comparison of SiC, GaAs, and Si RF MESFET power densities," *IEEE Transaction on Electron Device Letters*, Vol. 16, 451–453, 1995.
2. Xu, Y., "Large-signal modeling of SiC MESFETs," Master dissertation, University of Electronic Science and Technology of China, 2007.
3. Zhang, Q. J., K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design — from theory to practice," *IEEE Transactions Microwave Theory and Techniques*, Vol. 51, 1339–1350, 2003.
4. Xia, L., R. Xu, and B. Yan, "LTCC interconnect modeling by support vector regression," *Progress In Electromagnetics Research*, PIER 69, 67–75, 2007.
5. Bermani, E., A. Boni, A. Kerhet, and A. Massa, "Kernels evaluation of SVM based estimations for inverse scattering problems," *Progress In Electromagnetics Research*, PIER 53, 167–188, 2005.
6. Yang, Z. Q., T. Yang, Y. Liu, and S. H. Han, "MIM capacitor modeling by support vector regression," *Journal of Electromagnetic Waves and Applications*, Vol. 22, 61–67, 2008.
7. Wu, Y. Q., Z. X. Tang, B. Zhang, and Y. H. Xu, "Permeability measurement of ferromagnetic materials in microwave frequency

- range using support vector machine regression,” *Progress In Electromagnetics Research*, PIER 70, 247–256, 2007.
8. Xu, Y., “Modeling of SiC-MESFETs by using support vector machine regression,” *Journal of Electromagnetic Waves and Applications*, Vol. 21, 1489–1498, 2007.
  9. Lázaro, M., F. Pérez-Cruz, and A. Artés-Rodríguez, “Learning a function and its derivative forcing the support vector expansion,” *IEEE Signal Processing Letters*, Vol. 12, 194–197, 2005.
  10. Perez-Cruz, F., M. Lazaro, and A. Artes-Rodriguez, “Mutidimensional SVM to include the samples of the derivatives in the reconstruction of a function,” *2004 European Signal Processing Conference*, 597–600, Vienna, Sep. 2004.
  11. Pedro, J. C. and J. Pérez, “Accurate simulation of GaAs MESFETs intermodulation using a new drain-source current model,” *IEEE Transactions Microwave Theory and Techniques*, Vol. 42, 25–33, 1994.
  12. Santamaría, I., M. Lázaro, C. J. Pantaleón, J. A. García, A. Tazón, and A. Mediavilla, “A nonlinear MESFET model for intermodulation analysis using a generalized radial basis function network,” *Neurocomputing*, Vol. 25, 1–18, 1999.
  13. Dambrine, G., A. Cappy, F. Heliodore, and E. Player, “A new method for determining the FET small-signal equivalent circuit,” *IEEE Transactions Microwave Theory and Techniques*, Vol. 36, 1151–1159, 1988.
  14. Chang, C. C. and C. J. Lin, LIBSVM 2.81(2006): A library for support vector machines. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
  15. Wang, L., R. M. Xu, Y. C. Guo, and B. Yan, “A temperature-dependent current model for pHEMT on GaAs,” *Journal of Electromagnetic Waves and Applications*, Vol. 22, 39–46, 2008.
  16. Shi, Z. G., “Microwave chaotic colpitts oscillator: Design, implementation and application,” *Journal of Electromagnetic Waves and Applications*, Vol. 20, 1335–1349, 2006.