Behavioral Modeling of Power Amplifiers With Modern Machine Learning Techniques

Sener Dikmese[1,2], Lauri Anttila[1], Pablo Pascual Campo[1], Mikko Valkama[1], and Markku Renfors[1]

[1]Electrical Engineering, Faculty of Information and Communication Sciences, Tampere University, Tampere, Finland
[2] Ericsson AB, Stockholm, Sweden
e-mail: {sener.dikmese; lauri.anttila; pablo.pascualcampo; mikko.valkama; markku.renfors} @tuni.fi

***Abstract*—In this study, modern machine learning (ML) methods are proposed to predict the dynamic non-linear behavior of wideband RF power amplifiers (PAs). Neural networks, k-nearest neighbor, and several tree-based ML algorithms are first adapted to handle complex-valued signals and then applied to the PA modeling problem. Their modeling performance is evaluated with measured data from two basestation PAs. Gradient boosting is seen to outperform the other ML approaches and to give comparable performance to the generalized memory polynomial (GMP) reference model in terms of both the normalized mean squared error (NMSE) and adjacent channel error power ratio (ACEPR). This is the first study in the open literature to consider modern ML approaches, besides neural networks, for PA behavioral modeling.**

***Index Terms*—Behavioral modeling, power amplifiers, generalized memory polynomial, machine learning, neural networks, gradient boosting, tree based approaches, decision tree.**

I. Introduction

Modern wireless communication systems such as 4G/5G and 802.11 based systems utilize nonconstant-envelope I/Q-modulated signals such as orthogonal frequency division multiplexing (OFDM). Due to the high peak-to-average-power-ratio (PAPR), such signals will naturally excite the nonlinearities of the TX chain, especially of the power amplifier. In order to compensate for this nonlinear behavior, different techniques can be applied in the transmitter digital domain, with digital predistortion (DPD) being the most popular. In this context, behavioral modeling techniques predicting the nonlinearity of the power amplifier (PA) are of particular interest [1], [2].

Memory polynomial (MP) based models, such as the generalized MP (GMP), have been widely used for wideband PA modeling [3]. While GMP typically provides excellent modeling performance, the complexity of GMP and similar polynomial based models is high when modeling saturated PAs. In addition, a given polynomial based model is only valid within a narrow power range, therefore limiting their applicability. From these perspectives, machine learning (ML) based modeling techniques may offer increased performance with reasonable complexity. Neural network (NN) based structures have earlier been applied to PA modeling and digital predistortion (DPD) in [4], [5].

To the best of the authors' knowledge, besides NN based ML approaches, other modern/powerful ML methods such as decision tree (DT), random forest (RaF), gradient boosting (GB) and k-nearest neighbor (k-NN), have not been previously studied in the context of PA behavioral modeling. Thus, in this study, we propose to use these modern ML algorithms for PA behavioral modeling. The ML algorithms are designed to work with real-valued signals only, therefore they first need to be adapted to handle complex-valued signals. We then apply the ML techniques to data measured from two RF PAs, and compare their performance to conventional MP based techniques in terms of normalized mean squared error (NMSE) and adjacent channel error power ratio (ACEPR), considering wideband 5G new radio (NR) waveforms.

The remainder of the paper is organized as follows. The ML methods applied in this study are briefly introduced in Section II. The experimental results and the measurement setup are given in Section III. Finally, closing remarks are provided in Section IV.

II. Machine Learning Algorithms

Tree, k-NN and NN based approaches have been widely considered in ML applications covering both classification and regression [6]. The most common tree based models are DT, RaF, and GB based on the decision process.

DT regression is drawn upside down with its root at the top. Recursive binary splitting, which is also known as greedy approach, is a common technique to obtain good performance in DT based ML approaches, in which all the features are taken into account and different split points are tested using a cost function [6].

RaF regression is a flexible learning technique which provides good performance without hyper-parametric tuning. This method creates a forest as a collection of DTs and hence, creates additional randomness in the model to improve the performance and also alleviate the overfitting problem [6].

GB regression method is usually the best tree based ML technique. GB provides a prediction model in the form of an ensemble of weak prediction models, which are commonly DTs. Boosting, which can be interpreted as an optimization algorithm on a suitable cost function, is used to increase the performance of traditional tree based approaches [6].

The k-NN method is one of the simplest ML algorithms. It builds on the assumption that similar behaviors are near to each other. In this method, it is not required to build a model

This work was supported in part by the Finnish Cultural Foundation and the Academy of Finland (under grant #301820).

and impose additional assumptions as in the tree based methods. As a drawback, when the number of examples, predictors and independent variables increases, the computational time of the method increases significantly [7].

The NN regression technique is considered as a framework for several ML applications in different fields [8]. The NNs are composed of highly interconnected nodes, inspired by the structure of the human brain. Each connection of the NN is associated with a weight value which determines the importance of this relationship in the neuron [8]. In this work, we utilize supervised learning with the backpropagation algorithm [4].

Most ML algorithms do not naturally support complex notations. Therefore, we adopt a matrix formulation with the real and imaginary parts of the complex input and output signals written separately. While the available k-NN and NN algorithm implementations support both multiple inputs and outputs with matrix notations, the tree based algorithms do not support multiple outputs. Hence, for the tree-based methods separate real and imaginary training networks are applied, while the k-NN and NN require only a single training network in the fitting process.

While there are several studies on PA modeling and DPD with NN based ML in the literature, we are currently not aware of any studies that consider other modern ML methods such as the tree based approaches. In the following, we apply tree, k-NN and NN based ML algorithms to model PAs with memory effects.

III. Measurement Results

The PA modeling procedure with the different ML algorithms is illustrated in Fig. 1. As seen in the block diagram, MATLAB and PYTHON interfaces are used together for the purposes of signal processing and ML, respectively.

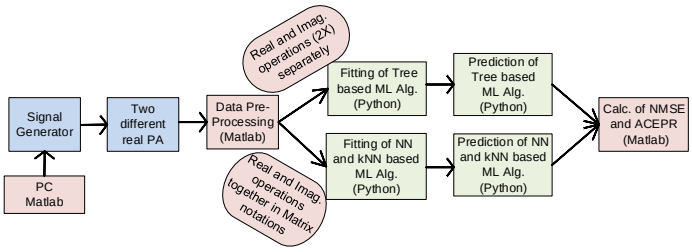

Fig. 1. Scheme of PA modeling using machine learning algorithms.

The test setup used for the experiments consists of a vector signal transceiver (VST) with MATLAB as the host-processor and two different PAs: a Skyworks SKY66293-21 medium-power PA designed for the n78 NR operating band, and a laterally diffused metal oxide semiconductor (LDMOS) 100-Watt peak-power base station (BS) PA designed for the n2 and n3 NR operating bands. Firstly, the baseband I/Q component samples are generated in the VST, and fed to the PA with the corresponding power and carrier frequency. The PA output RF signal is then input to the VST, where the received I/Q data samples are extracted for further digital processing.

All the signals used in this work have been generated according to the NR BS radio transmission and reception standards. Two different bandwidths, 100 and 60 MHz, are considered in the experiments. Additionally, to decrease the noise variance in the measured signals, statistical averaging is implemented to the received I/Q samples, such that the final data is the result of 10 averaged measurements.

To evaluate the performance of the different behavioral modeling techniques, we use the normalized mean square error (NMSE) and the adjacent channel error power ratio (ACEPR). The NMSE evaluates the full-band modeling accuracy of the PA behavioral model, and can be defined as

$$\mathrm{NMSE_{dB}} = 10\log_{10}\frac{\sum_{n=1}^{N}|e_{\mathrm{model}}[n]|^2}{\sum_{n=1}^{N}|y_{\mathrm{meas}}[n]|^2}, \tag{1}$$

where $e_{\mathrm{model}}[n] = y_{\mathrm{meas}}[n] - y_{\mathrm{model}}[n]$ is the error signal between the measured signal and the predicted signal. On the other hand, the ACEPR evaluates only the out-of-band modeling performance, computing a ratio between the error signal power over the adjacent channel and the desired channel power of the measured signal ($P_{\mathrm{error}}^{\mathrm{adj}}$ and $P_{\mathrm{meas}}^{\mathrm{ch}}$), as

$$\mathrm{ACEPR_{dB}} = 10\log_{10}\frac{P_{\mathrm{error}}^{\mathrm{adj}}}{P_{\mathrm{meas}}^{\mathrm{ch}}}. \tag{2}$$

Training and prediction stages are considered separately in this study, as shown in Fig. 1. The training networks of all used ML methods are applied once and saved for the future test signals. In the data processing stage, the training and test data are split as 80% and 20% of the total number of samples $N = 300000$, respectively. While 20 hidden layers and nearest neighbors are applied in NN and k-NN, respectively, maximum depth of 10 and 400 trees are considered in both RaF and GB regression methods. The maximum depth and minimum sample splits are 10 and 2 in the DT regression. For the reference MP model, we use polynomial order $P = 9$ (with odd orders only considered) and memory order $M = 4$, while for the GMP model, aligned signal term $K_a = 7$, envelope term $L_a = 4$, lagging envelope terms $K_b = 5$, $L_b = 3$, $G_b = 2$ and leading envelope terms $K_c = L_c = G_c = 0$ are considered (please refer to [3] for more details). Similarly, we use $M = 4$ in all the ML algorithms. Once the training stage is completed, the different test signals are applied in the prediction stage using the saved fitting network. Finally, the NMSE and ACEPR are calculated after forming the complex signals from the real-valued predictor outputs.

The power spectral densities (PSDs) of the PA input, and the actual and predicted PA output using MP, GMP and ML algorithms, considering 60 MHz NR signal and the 100-Watt peak-power BS PA are shown in Fig. 2. This PA has 40-dB gain and +50-dBm 1-dB compression point. The corresponding NMSE and ACEPR results of the same algorithms are calculated as in Table I. In terms of NMSE, GB based ML outperforms the other ML techniques and the MP model, and has quite similar performance compared to the GMP model.

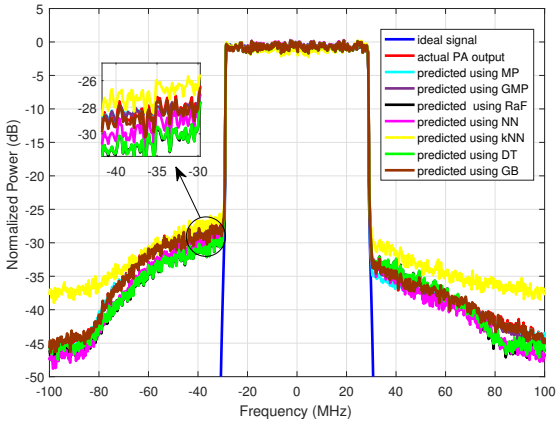

Fig. 2. Normalized PSDs of the ideal, measured, and predicted PA output signals. 100-Watt peak-power PA operating at 48-dBm output power under 60 MHz NR signal.

TABLE I
NMSE AND ACEPR OF ML AND POLYNOMIAL BASED ALGORITHMS FOR THE 100-WATT PEAK-POWER BS PA.

Machine Learning and Polynomial based Algorithms	NMSE	ACEPR
Memory Polynomial	-33.9138	-44.4245
Generalized Memory Polynomial	-38.1549	-46.4166
Neural Network	-31.5821	-39.9753
k-nearest neighbor	-22.7671	-34.6219
Random Forest	-31.3270	-38.3730
Decision Tree	-31.2149	-38.2673
Gradient Boosting	-38.0011	-43.4606

Similarly, GB has the best ACEPR performance out of all the ML techniques and reaches a similar performance compared to MP based modeling. NN is the second best ML algorithm in terms of both NMSE and ACEPR, but is still several decibels behind GB.

Similarly, the PSDs and NMSE/ACEPR of the same models considering a 100 MHz NR signal on a Skyworks SKY66293-21 medium-power PA are shown in Fig. 3 and Table II, respectively. This PA has 33-dB gain and +31.5-dBm 1-dB compression point. GB based ML outperforms not only the other ML approaches, but also the reference MP based techniques, in terms of both NMSE and ACEPR. It is also noteworthy, that all the other ML methods perform clearly worse in modeling this particular PA.

IV. CONCLUSIONS AND FUTURE WORK

This paper proposed a novel behavioral modeling scheme for RF PAs employing modern machine learning techniques. The performance of decision tree, random forest, gradient boosting, k-nearest neighbor and neural network were evaluated with measured data from two commercial basestation PAs excited with wideband 5G NR signals. Gradient boosting was shown to perform the best, yielding similar or even better performance compared to the memory and generalized memory polynomial reference models.

In this work, a single power level near to the saturation level of each PA was considered to evaluate the performance of the learning techniques. In future studies, a single predictor will be studied to model a PA at different output power levels. Instead

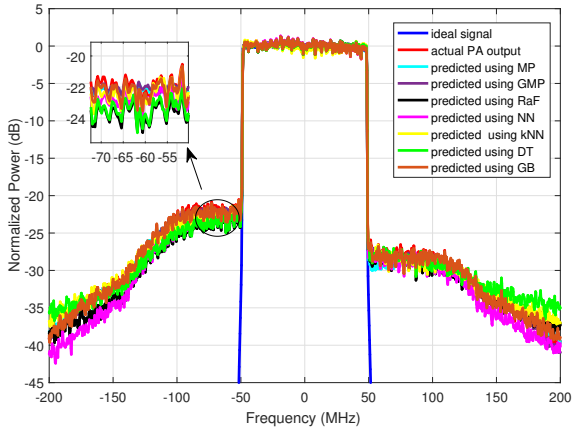

Fig. 3. Normalized PSDs of the ideal, measured, and predicted PA output signals. Skyworks PA operating at 29-dBm output power under 100 MHz NR signal.

TABLE II
NMSE AND ACEPR OF ML AND POLYNOMIAL BASED ALGORITHMS FOR THE SKYWORKS SKY66293-21 MEDIUM-POWER PA.

Machine Learning and Polynomial based Algorithms	NMSE	ACEPR
Memory Polynomial	-27.6417	-35.9327
Generalized Memory Polynomial	-31.6969	-37.5741
Neural Network	-25.6709	-32.8280
k-nearest neighbor	-22.7605	-33.4338
Random Forest	-25.5790	-32.3666
Decision Tree	-24.3239	-31.3957
Gradient Boosting	-32.8320	-41.5048

of fitting and saving separate networks for each power level in the training stage, this can be achieved simply by tagging each measured data set with a power level dependent tag. This will decrease the predictor computational complexity significantly. Moreover, due to the page limit, this study does not cover the detailed computational complexity analysis, which remains a topic of future work. Applying similar machine learning techniques for digital predistortion linearization is an obvious future extension of this work.

REFERENCES

[1] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microwave magazine*, vol. 10, no. 7, pp. 52–64, 2009.
[2] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 6, pp. 1510–1520, 2010.
[3] D. R. Morgan *et al.*, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, 2006.
[4] F. Mkadem and S. Boumaiza, "Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 4, pp. 913–923, 2011.
[5] M. Bhuyan and K. K. Sarma, "Learning aided behavioral modeling and adaptive digital predistortion design for nonlinear power amplifier," *IEEE Sensors Journal*, vol. 16, no. 16, pp. 6167–6174, Aug 2016.
[6] C. Sheppard, *Tree-based machine learning algorithms: Decision trees, random forests, and boosting*. CreateSpace Ind. Publish. Platform, 2017.
[7] J. Song, J. Zhao, F. Dong, J. Zhao, Z. Qian, and Q. Zhang, "A novel regression modeling method for PMSLM structural design optimization using a distance-weighted KNN algorithm," *IEEE Transactions on Industry Applications*, vol. 54, no. 5, pp. 4198–4206, Sep. 2018.
[8] M. A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2018.