Mask Reconstruction via Kernel-based Surrogate Modeling for Automated Circulator Design

Marco Atlante, *Student Member, IEEE,* Riccardo Trinchero, *Member, IEEE,* Igor Simone Stievano, *Senior Member, IEEE,* Mihai Telescu, *Member, IEEE,* Vincent Laur, *Senior Member, IEEE,* Noël Tanguy

***Abstract*—This letter introduces a mask-based surrogate modeling framework which generalizes the feature-based paradigm for efficient microwave design exploration using a limited number of full-wave simulations. Unlike feature-based approaches, which extract scalar performance metrics, the proposed method reconstructs the complete S-parameter boundary contour within the operating band. This preserves each design's parametric signature and enables on-demand evaluation of multiple performance specifications without model retraining. The framework employs two Vector-Valued Kernel Ridge Regression models to predict boundary coordinates from geometric parameters. Validation on a three-port circulator demonstrates accurate reconstruction and rapid exploration of thousands of candidate designs.**

***Index Terms*—Surrogate modeling, kernel regression, circulator, machine learning, microwave devices, design optimization.**

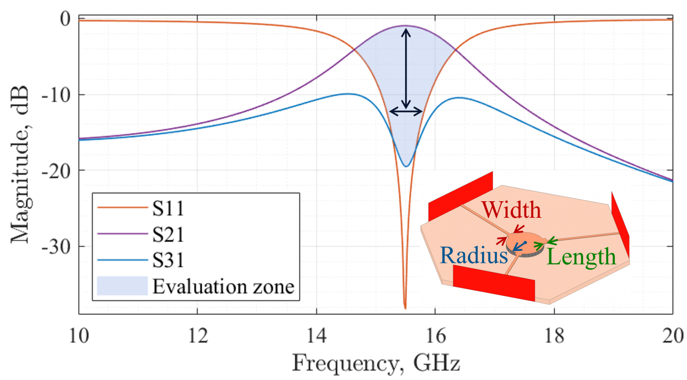

Fig. 1. Simulated S-parameter response for a three-port circulator design (topology shown in inset with geometric parameters indicated). The shaded evaluation zone defines the operating band where performance metrics are extracted from the mask boundaries.

I. Introduction

The design of complex microwave components is typically carried out through repeated computationally expensive electromagnetic simulations, making the exploration of large design spaces difficult and rendering optimization procedures cumbersome. In practical workflows, designers focus on key performance metrics extracted from simulated S-parameters over a target operating band, the region where device specifications must be verified (e.g., the shaded region in Fig. 1 for a circulator design). Surrogate modeling provides an efficient alternative to direct simulation, as it enables learning and reproducing the parametric behavior of quantities of interest from a limited dataset. Once trained, the resulting surrogate model allows for fast and accurate predictions, thus facilitating the design and optimization of microwave devices [1], [2]. This can be achieved using data-driven regression techniques, such as neural or recurrent networks with Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) layers [3], [4]. These models preserve full spectral detail, but also predict frequency regions that are irrelevant to design verification, adding unnecessary information. An alternative approach is represented by feature-based surrogate models [5]–[7], which predict scalar performance figures derived from S-parameters, such as center frequency, bandwidth, or pole–residue parameters. The performance-driven framework in [8]–[10] trains surrogates on designs satisfying performance requirements using feature-based extraction. Fast global sensitivity analysis [11] improves scalability by identifying dominant parameter directions. A recent feature-oriented surrogate for circulator design [12] applies this paradigm to predict device-specific metrics from geometric parameters. Training requires fixing threshold levels (e.g., isolation level) to measure and extract features (e.g., bandwidth) as training targets. However, feature-based approaches fix the target metrics during data preparation. This means that evaluating new performance requirements needs re-extracting features and retraining the model. Moreover, when using threshold-based selection, simulations not satisfying the initial criteria are discarded, wasting computational effort and reducing design-space coverage.

This letter presents a mask-based framework that generalizes the feature-based method. Rather than training separate surrogates for specific scalar metrics, the proposed approach reconstructs the complete S-parameter boundary envelope enclosing the operating region of interest (Fig. 1). The key difference lies in the surrogate's output format: it predicts the mask's frequency-magnitude coordinates directly, preserving each design's parametric signature across the target band. Any metric (bandwidth, center frequency, loss levels) then becomes a derived quantity computed by applying thresholds during evaluation. This generalization yields two practical advantages. First, all simulation data is retained. The framework does not filter geometries based on performance criteria, ensuring that every sampled design contributes to boundary reconstruction without creating gaps in the parameter do-

M. Atlante, R. Trinchero, and I. S. Stievano are with the EMC Group, Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy (e-mail: {marco.atlante, riccardo.trinchero, igor.stievano}@polito.it). (Corresponding author: Marco Atlante.)

M. Telescu, V. Laur and N. Tanguy are with University of Brest, Lab-STICC, CNRS, UMR 6285, F-29200, Brest, France (e-mail: {mihai.telescu, vincent.laur, noel.tanguy}@univ-brest.fr).

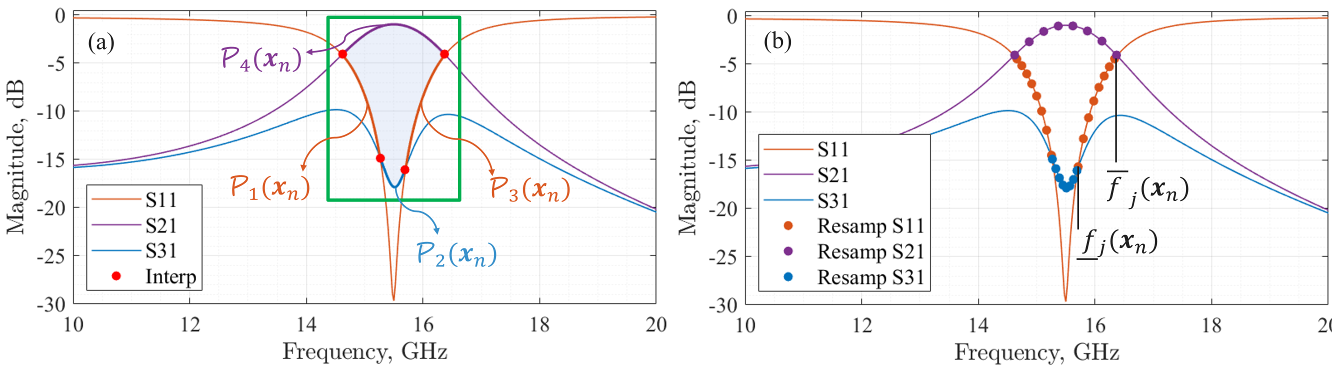

Fig. 2. Mask extraction and regularization for a circulator design. (a) Four-segment boundary structure with intersection points (red circles) delimiting regions dominated by transmission (top), reflection (sides), and isolation (bottom). (b) Uniform resampling: each segment $\mathcal{P}_j(\boldsymbol{x}_n)$ is resampled with N_p equally spaced points ($N_p = 8$ for illustration purposes), yielding fixed-length representations for surrogate training.

main. Second, multiple design scenarios can be evaluated: new specifications require only applying different thresholds to the same predicted masks, eliminating data re-processing and model retraining. Designers can thus explore multiple constraints, vary thresholds, or adapt to evolving requirements using the same trained model. To implement this framework, Vector-Valued Kernel Ridge Regression (VV-KRR) [13]–[17] is adopted to construct the surrogate. Moreover, its general formulation makes it adaptable to other regression techniques and to other devices characterized by frequency-dependent response boundaries, including filters, impedance-matching networks, and various multiport systems.

II. Mask Extraction and Standardization

The framework was validated on a dataset of $N = 150$ full-wave simulations of a three-port Y-junction circulator. Simulations were performed in Ansys HFSS™ over the 10–20 GHz range, encompassing the Ku-band (12–18 GHz) [18], [19]. The device consists of a self-biased hard-ferrite puck embedded in a dielectric substrate [20], [21] and a fixed Y-junction topology with 50 Ω feed lines (shown in the inset of Fig. 1). Three geometrical parameters were varied: resonator radius r ($[990, 1510]$ μm), matching-line length ℓ ($[5, 395]$ μm), and matching-line width w ($[310, 450]$ μm). These parameters were randomly sampled through Latin Hypercube Sampling (LHS) for uniform design-space coverage. For each geometry $\boldsymbol{x}_n = [r_n, \ell_n, w_n]^T \subseteq \mathbb{R}^3$, the corresponding S-parameter responses $|S_{11}(f; \boldsymbol{x}_n)|$, $|S_{21}(f; \boldsymbol{x}_n)|$, and $|S_{31}(f; \boldsymbol{x}_n)|$ describe the reflection, transmission, and isolation behavior of the device. The dataset covers a broad range of responses, including sharp resonances and broadband transitions.

The mask extraction process identifies and samples the boundary envelope of S-parameter responses within the operating region. For the circulator, this envelope captures worst-case isolation and insertion loss. As illustrated in Fig. 2(a), the envelope forms a closed contour composed of $J = 4$ segments $\mathcal{P}_j(\boldsymbol{x}_n)$, each dominated by a different S-parameter: transmission (top), reflection (sides), and isolation (bottom). This four-segment structure is preserved across all geometries in the considered design space. The segments meet at intersection points where S-parameter curves cross (marked as red dots in Fig. 2(a)). These points serve as anchors which define the portions of each curve belong to the envelope. In this work, intersections are located by detecting where the difference between curves changes sign, with sub-sample precision achieved via linear interpolation. A critical challenge arises from the geometry-dependent nature of the S-parameter responses: while the overall mask shape remains qualitatively similar across configurations, each segment exhibits significant differences in frequency span, position, and local density due to resonance shifts induced by changes in $\boldsymbol{x}_n$ (visible in Fig. 3). Consequently, each extracted segment $\mathcal{P}_j(\boldsymbol{x}_n)$ (Fig. 2(a)) contains a variable number of samples, preventing direct concatenation into a unified matrix representation.

To address this, a uniform resampling strategy based on interpolation is employed. Each boundary portion is resampled by defining equally spaced frequency points N_p within its boundaries $[\underline{f}_j(\boldsymbol{x}_n), \overline{f}_j(\boldsymbol{x}_n)]$ and interpolating the corresponding magnitude values (in decibel scale). This process transforms variable-length segments into fixed-size representations, yielding vectors $\boldsymbol{f}_j(\boldsymbol{x}_n) = [f_{j,1}(\boldsymbol{x}_n), \ldots, f_{j,N_p}(\boldsymbol{x}_n)]^T$ and $\boldsymbol{m}_j(\boldsymbol{x}_n) = [m_{j,1}(\boldsymbol{x}_n), \ldots, m_{j,N_p}(\boldsymbol{x}_n)]^T$, as shown by the filled markers in Fig. 2(b). The complete mask is obtained by concatenating all J resampled segments into two aligned vectors, $\boldsymbol{y}_n^{(\text{freq})} = \text{vec}(\boldsymbol{f}_1(\boldsymbol{x}_n), \ldots, \boldsymbol{f}_J(\boldsymbol{x}_n))$ and $\boldsymbol{y}_n^{(\text{mag})} = \text{vec}(\boldsymbol{m}_1(\boldsymbol{x}_n), \ldots, \boldsymbol{m}_J(\boldsymbol{x}_n))$, each containing JN_p elements. To enable effective comparison across different value ranges, both frequency and magnitude vectors are normalized using z-score standardization. Since these coordinates represent distinct physical quantities, they are modeled separately using a dual-surrogate formulation, leading to two point-wise aligned datasets sharing the same geometric inputs:

$$\mathcal{D}_{\text{freq}} = \{(\boldsymbol{x}_n, \boldsymbol{y}_n^{(\text{freq})})\}_{n=1}^N, \ \mathcal{D}_{\text{mag}} = \{(\boldsymbol{x}_n, \boldsymbol{y}_n^{(\text{mag})})\}_{n=1}^N. \quad (1)$$

In the present implementation, $J = 4$ segments and $N_p = 15$ points per segment were used, yielding $D = JN_p = 60$ output coordinates per sample. The choice of $N_p = 15$ was empirically determined to provide sufficient resolution for smooth boundary reconstruction while maintaining manageable model dimensionality. The dataset was partitioned into $L = 100$ training samples and $T = 50$ test samples, each corresponding to a distinct simulated geometry.

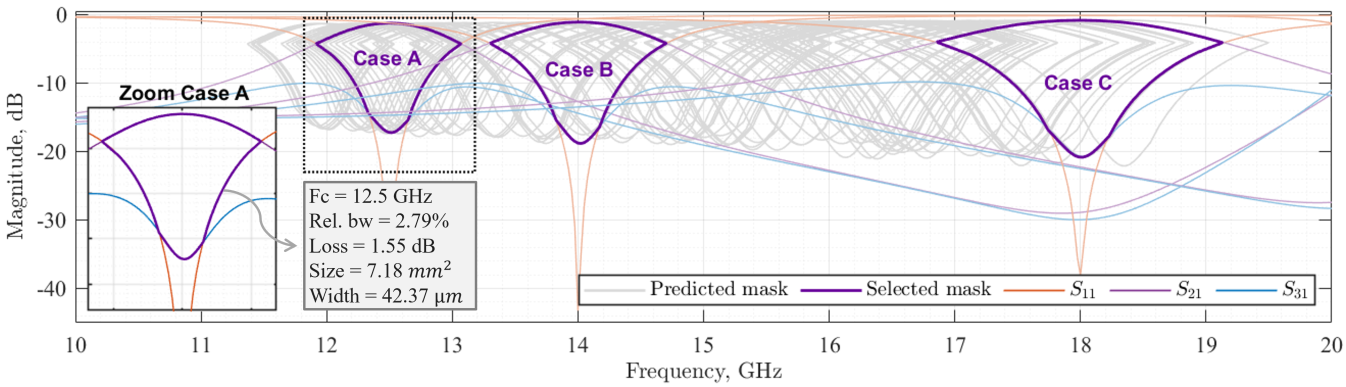

Fig. 3. The gray curves show 100 predicted masks from the surrogate model, sampled from a larger set of 10,000 evaluated geometries. The thick purple curves highlight the three optimal predicted masks (Cases A–C), automatically selected based on distinct performance specifications. Colored curves represent the corresponding HFSS-simulated S-parameters (S_{11} in orange, S_{21} in purple, S_{31} in blue). The inset provides a close-up view of Case A, showing excellent geometric agreement between the predicted mask and the full-wave simulation result, including also the actual values of the design parameters.

III. Surrogate Modeling and Design Optimization

Following the datasets defined in (1), two VV-KRR models [13]–[17] are trained independently for the frequency and magnitude training sets, each with D output dimensions. For each coordinate type $i \in \{\text{freq}, \text{mag}\}$, the model prediction $\hat{\boldsymbol{y}}^{(i)}(\boldsymbol{x})$ for a new geometry $\boldsymbol{x}$, is expressed as:

$$\hat{\boldsymbol{y}}^{(i)}(\boldsymbol{x}) = \sum_{l=1}^{L} k_{\boldsymbol{x}}^{(i)}(\boldsymbol{x}, \boldsymbol{x}_l)\, \mathbf{B}^{(i)}\, \boldsymbol{c}_l^{(i)}, \tag{2}$$

where $\mathbf{B}^{(i)}$ is a $D \times D$ matrix accounting for the correlation in the output dimensions, $k_{\boldsymbol{x}}^{(i)}$ is a scalar kernel accounting for the correlation among the design parameters, and $\boldsymbol{c}_l^{(i)}$ are D-dimensional vectors collecting the model coefficients to be estimated during training. The training of the above VV-KRR model turns out to be equivalent to solving the following Sylvester equation (additional details in [14], [22], [23]):

$$\mathbf{K}_{\boldsymbol{x}}^{(i)}\mathbf{C}^{(i)}\mathbf{B}^{(i)} + \lambda^{(i)}\mathbf{C}^{(i)} = \mathbf{Y}^{(i)}, \tag{3}$$

where $\mathbf{C}^{(i)} = [\boldsymbol{c}_1^{(i)}, \ldots, \boldsymbol{c}_L^{(i)}]^T$ collects the regression coefficients, $\lambda^{(i)}$ is the regularization parameter, $\mathbf{K}_{\boldsymbol{x}}^{(i)}$ is the Gramian matrix of the input kernel computed from the L training samples, and $\mathbf{Y}^{(i)} = [\boldsymbol{y}_1^{(i)}, \ldots, \boldsymbol{y}_L^{(i)}]^T \in \mathbb{R}^{L \times D}$ denotes the training output matrix. This formulation enables computing all D output coefficients jointly within a single matrix operation by leveraging correlations across both input geometries (via $k_{\boldsymbol{x}}^{(i)}$) and output boundary coordinates (via $\mathbf{B}^{(i)}$). This is more efficient than training D independent scalar models, which would require separate coefficient sets and hyperparameter optimization for each output dimension. Equation (3) is efficiently solved via the diagonalization strategy introduced in [23]. For the input kernel, we employ a Matérn 5/2 kernel with bandwidth $\sigma^{(i)}$ [24]. The output kernel is constructed as the empirical correlation matrix $\mathbf{B}^{(i)} = \mathrm{corr}(\mathbf{Y}^{(i)})$, capturing the strong correlations between adjacent boundary points that arise from the continuous nature of S-parameter responses. Modeling these correlations ensures smooth, physically consistent boundary predictions and improves sample efficiency. The hyperparameters $(\lambda^{(i)}, \sigma^{(i)})$ are optimized via 5-fold cross-validation. Training completes in approximately 10 s per model, while each prediction requires only a few milliseconds. To assess accuracy, we compare predicted boundaries against HFSS simulations on the test set. Root-mean-square errors (RMSE) of 0.02 MHz in frequency and 0.18 dB in magnitude demonstrate high accuracy, with relative errors below 0.21% and 0.85%, respectively. Once validated, the surrogates were employed for large-scale design exploration.

In this study, 10,000 new configurations generated using LHS, with the geometric parameters and ranges defined in Sec. II, were evaluated in less than one second. Figure 3 shows a representative subset of the predicted responses (gray curves), illustrating the broad exploration of the design space achieved through geometric parameter variation. To demonstrate the framework's flexibility, three design scenarios emulating typical RF front-end conditions were defined, with target frequencies of 12.5, 14, and 18 GHz and other requirements detailed in Table I.

TABLE I
Target specifications for the representative design scenarios

Specification	**Case A**	**Case B**	**Case C**
Central frequency (Fc)	12.5 GHz	14 GHz	18 GHz
Min relative bandwidth (100·bw/Fc)	2%	3%	3.5%
Isolation/return loss (min. in-band)	14 dB	13 dB	15 dB
Max in-band insertion losses	1.8 dB	1.5 dB	1.2 dB
Maximum size	7.5 mm^2	6.0 mm^2	6.5 mm^2
Minimum line width	20 μm	18 μm	25 μm

All three scenarios were evaluated using the same trained surrogate by filtering the 10,000 surrogate predictions (post-training) against case-specific thresholds. For each candidate, compliance with electrical specifications (isolation, insertion loss, bandwidth) and structural limits (area, line width) was verified. Designs satisfying all constraints were retained, and the optimal solution was selected by maximizing relative bandwidth, yielding the results shown as thick purple curves in Fig. 3. Full-wave HFSS verification of the selected designs confirms close agreement with predicted masks, validating the surrogate accuracy. The inset on the left of Fig. 3 shows the predicted mask overlaid with the corresponding HFSS simulation for Case A, with accurate reconstruction of resonance and transition regions; geometric parameter values are also reported, confirming compliance with design constraints.

The results above demonstrate post-training flexibility in handling multiple specifications, a direct consequence of preserving complete boundaries rather than pre-selecting scalar features. The feature-based approach in [12] uses three metrics: center frequency (Fc), bandwidth (bw) and the frequency span meeting loss requirements (LossBand). We validate the efficiency of the mask-based method quantitatively by comparing prediction accuracy on these three metrics. Consistency is insured by applying the same threshold-based extraction procedure to both simulated S-parameters (feature-based) and predicted boundaries (mask-based). Table II compares relative RMSE values on identical test samples (three feature-based models vs. one mask-based model). For Cases A and B, both approaches achieve errors below 1.1%, confirming that mask-based extraction matches feature-based performance while maintaining specification flexibility.

TABLE II
RELATIVE RMSE (%): FEATURE-BASED [12] VS MASK-BASED. CASES A AND B USE $L = 100$ TRAINING SAMPLES. CASE C (*) USES REDUCED DATASET DUE TO THRESHOLD CONSTRAINTS (SEE TEXT).

Metric	**Case A**		**Case B**		**Case C**	
	Feature	Mask	Feature	Mask	Feature	Mask
F_c	0.0856	0.1003	0.1344	0.0995	0.0947*	0.0802
bw	0.7433	0.9041	0.3031	0.7204	1.0776*	0.7771
LossBand	0.3819	0.8318	0.6437	1.0996	4.8295*	2.9879

Case C, instead, shows better predictions via the proposed approach. The stricter requirements leave only 116 out of 150 simulations meeting the threshold levels, excluding 34 geometries from feature-based model construction and limiting training to $L^* = 80$ samples (errors marked with *). The mask-based method operates independently of target specifications, using all $L = 100$ training samples. This eliminates wasted computational resources and preserves design-space coverage.

IV. CONCLUSIONS

This work demonstrated a mask-based surrogate modeling framework that reconstructs complete S-parameter boundaries, enabling post-hoc evaluation of arbitrary design specifications without retraining. Validation on three Ku-band circulator scenarios achieved accuracy comparable to feature-based methods. The framework requires that the boundary defining the operating region maintains consistent topology across the design space. For the investigated circulator design, all the geometries exhibited identical four-segment closed masks despite varying geometric parameters, enabling fixed-length mask extraction. Devices whose boundary topology changes with design parameters, such as components with parameter-dependent resonance counts or shifts between open/closed boundaries, would require adaptive methods to handle variable, length masks. Extending the framework to such cases and to other passive components represents natural future directions.

REFERENCES

[1] A. Forrester, A. Sobester, and A. Keane, *Engineering Design via Surrogate Modelling*, John Wiley & Sons, 2008.

[2] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker, "Surrogate-based analysis and optimization", *Prog. Aerosp. Sci.*, vol. 41, no. 1, pp. 1–28, 2005.

[3] F. Feng, W. Na, J. Jin, J. Zhang, W. Zhang, and Q. J. Zhang, "Artificial neural networks for microwave computer-aided design: The state of the art", *IEEE Trans. Microw. Theory Tech.*, vol. 70, no. 11, pp. 4597–4620, Nov. 2022.

[4] K. C. Sahu, S. Koziel, and A. Pietrenko-Dabrowska, "Surrogate modeling of passive microwave circuits using recurrent neural networks and domain confinement", *Sci. Rep.*, vol. 15, 13322, 2025.

[5] J. W. Bandler, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Søndergaard, "Space mapping: The state of the art", *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.

[6] J. Jin, et al., "Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters", *IEEE Trans. Microw. Theory Tech.*, vol. 67, no. 10, pp. 4140–4155, Oct. 2019.

[7] B. Liu, H. Yang, and M. J. Lancaster, "Global optimization of microwave filters based on a surrogate model-assisted evolutionary algorithm", *IEEE Trans. Microw. Theory Tech.*, vol. 65, no. 6, pp. 1976–1985, Jun. 2017.

[8] S. Koziel and A. Pietrenko-Dabrowska, "Performance-Driven Surrogate Modeling of High-Frequency Structures", Springer, Cham, Switzerland, 2020.

[9] A. Pietrenko-Dabrowska, S. Koziel, and Q. J. Zhang, "Cost-efficient two-level modeling of microwave passives using feature-based surrogates and domain confinement", *Electronics*, vol. 12, no. 17, 3560, 2023.

[10] A. Pietrenko-Dabrowska and S. Koziel, "Response Feature Technology for High-Frequency Electronics: Optimization, Modeling, and Design Automation", Springer, Cham, Switzerland, 2023.

[11] S. Koziel, A. Pietrenko-Dabrowska, and L. Leifsson, "Improved efficacy behavioral modeling of microwave circuits through dimensionality reduction and fast global sensitivity analysis", *Sci. Rep.*, vol. 14, 19465, 2024.

[12] M. Atlante et al., "Kernel-Based Machine Learning Surrogates for RF Circulator Design", *Asia-Pacific Microwave Conference (APMC)*, Jeju, South Korea, Dec. 2-5, 2025.

[13] J. Shawe-Taylor and N. Cristianini, *Kernel methods: an overview.* Kernel Methods for Pattern Analysis, Cambridge: Cambridge University Press, 2004, pp. 25–46.

[14] N. Soleimani, R. Trinchero and F. Canavero, "Bridging the Gap Between Artificial Neural Networks and Kernel Regressions for Vector-Valued Problems in Microwave Applications," IEEE Transactions on Microwave Theory and Techniques, vol. 71, no. 6, pp. 2319-2332, June.

[15] C. A. Micchelli and M. Pontil, "On Learning Vector-Valued Functions," Neural Computation, vol. 17, no. 1, pp. 177–204, Jan. 2005.

[16] A. Mauricio, L. Rosasco, and N. D. Lawrence, "Kernels for Vector–Valued Functions: A Review," Foundations and Trends in Machine Learning, vol. 4, no. 3, pp. 195–266, 2012.

[17] M. Atlante, R. Trinchero, I. S. Stievano, M. Telescu and N. Tanguy, "IC Modeling via Machine Learning Regressions: A Data-Driven Approach to SPICE Integration," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 15, no. 9, pp. 1814–1822, Sept. 2025.

[18] J. Wang, et al., "Self-biased Y-junction circulator at Ku band,", *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 6, pp. 292–294, Jun. 2011.

[19] Y. Le Noane, et al., "Uncertainty Quantification for a Microstrip Self-Biased Ku-band Circulator", Asia-Pacific Microwave Conference (APMC2023), Dec. 2023, Taipei, Taiwan.

[20] H. Bosma, "On strip line Y-circulation at UHF", *IEEE Trans. Microw. Theory Tech.*, vol. MTT-12, no. 1, pp. 61–72, Jan. 1964.

[21] N. Parker-Soues, et al., "Compact Self-Biased Q-Band Circulators", *IEEE Trans. Magn.*, vol. 57, no. 4, pp. 1–6, Apr. 2021.

[22] F. Dinuzzo, C. S. Ong, P. Gehler, and G. Pillonetto, "Learning Output Kernels with Block Coordinate Descent," Proc. 28th Int. Conf. on Machine Learning (ICML), 2011, pp. 49–56.

[23] N. Soleimani, P. Manfredi and R. Trinchero, "Efficient Implementation of the Vector-Valued Kernel Ridge Regression for the Uncertainty Quantification of the Scattering Parameters of a 2-GHz Low-Noise Amplifier," *2023 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO),* Winnipeg, MB, Canada, 2023, pp. 143-146.

[24] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.