Machine learning enhanced predictions of ICRF heating: overcoming numerical limitations via data curation

A. Sanchez-Villar,[1] Z. Bai,[2] N. Bertelli,[1] E. W. Bethel,[3] J. Hillairet,[4] T. Perciano,[2] S. Shiraiwa,[1] G. M. Wallace,[5] and J. C. Wright[5]

[1)] *Princeton Plasma Physics Laboratory, Princeton, NJ 08540, USA*
[2)] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[3)] *San Francisco State University, San Francisco, CA 94132, USA*
[4)] *CEA-IRFM, F-13108 Saint-Paul-Lez-Durance, France*
[5)] *MIT Plasma Science and Fusion Center, Cambridge, MA 02139, USA*

(*Electronic mail: asvillar@pppl.gov)

(Dated: 28 February 2025)

In this work we present the development of robust surrogate models for Ion Cyclotron Range of Frequencies (ICRF) and High-Harmonic Fast Wave (HHFW) heating predictions in fusion plasmas. Building upon our previous efforts to achieve real-time capable models, we identify the cause of the outliers found using TORIC in certain HHFW heating scenarios. The outliers are observed to be spurious ion Bernstein wave (IBW)-like modes caused by a wavelength control algorithm designed to address challenging scenarios with high perpendicular wavenumbers. The effect arises from the modulation in the perpendicular susceptibility, which can induce sign reversal and IBW-like propagation for scenarios featuring normalized ion Larmor radius $\lambda_i \gg 1$. We use TORIC with this algorithm disabled to generate a novel HHFW-NSTX database that is free of outliers. Surrogate models trained on this database, including Random Forest Regressors (RFR), Multi-Layer Perceptrons (MLP), and Gaussian Process Regressors (GPR), demonstrate the ability to accurately predict HHFW heating profiles, with regression scores of $R^2 \in [0.93-0.99]$. Additionally we demonstrate that it is possible to generalize predictions beyond training data by the use of both RFR and GPR models, enabling the prediction of scenarios previously limited to the original model. GPR models also provide uncertainty quantification, offering insights into model confidence. This work introduces a comprehensive Verification, Validation, and Uncertainty Quantification (VVUQ) methodology for surrogate modeling, applicable not only to ICRF heating but also to other RF heating challenges and fusion physics problems. Beyond accelerated inference, these models show performant extrapolation capabilities, providing an alternative for addressing numerical challenges.

I. INTRODUCTION

The demonstration of robust, reliable and efficient Radio-Frequency (RF) actuator systems is critical for the realization of effective magnetic confinement fusion reactors. Among RF systems, Ion Cyclotron Range of Frequencies (ICRF) heating[1–3] plays a central role in heating plasma through resonant interactions between RF waves and charged particles. This mechanism is highly flexible, allowing energy transfer to either ions (i) or electrons (e) depending on the chosen frequency (ω), ion species, and resonance conditions. By depositing power into the plasma, ICRF systems contribute to achieving the temperature conditions necessary for sustaining fusion reactions and optimizing plasma performance.

The design, operation and performance of ICRF-actuator relies on the achievement of high-fidelity ICRF modeling codes[4–12] which are generally built on physics-based computational methods. These computational codes feature high computational costs that make their application challenging for specific scenario optimization, limited for inter-shot predictive modeling, and unfeasible for real-time control.

Artificial intelligence and machine learning (AI/ML) frameworks appear as a disruptive technology that can solve some of the challenges of fusion energy sciences[13] such as disruption prediction and control[14], experimental analysis[15], scenario planning and experimental design[16], enhanced diagnostics[17], and model extraction or reduction via surrogate model development. In particular, surrogate models have been found useful to accelerate inference while preserving accuracy in multiple plasma physics areas such as disruption mitigation[18,19], transport[20,21], turbulence[22], plasma-surface interaction[23], kinetic equilibrium reconstruction[24], and auxiliary heating[25–27], amongst others.

In the context of RF heating, fast and accurate surrogate models were achieved[26] for the lower hybrid current drive using coupled GENRAY-CQL3D[28,29] simulations as the reference model. More recently, ML-based surrogate models have been trained to successfully achieve real-time capable predictions of ICRF heating[27] for the high harmonic fast wave[30] (HHFW) at NSTX[31], and the ion cyclotron (IC) minority heating[32] at WEST[33]. The models[27] were trained on a database generated with TORIC[5–7], an ICRF semi-spectral full-wave solver for toroidally axisymmetric plasmas. The ICRF heating surrogate models achieved[27] were based on the Random Forest Regressor[34,35] (RFR) architecture, and the Multi-Layer Perceptron[36,37] (MLP). These models showed an acceleration in the predictions of multi-species heating

profiles by six-to-seven orders of magnitude compared to that of the reference model TORIC. Such result demonstrated the feasibility of training effective scenario-specific surrogate models of ICRF heating in different fusion devices featuring diverse heating schemes and plasma conditions via ML approaches.

In certain scenarios, numerical approximations within simulation codes can fail to reproduce physical results. In the context of ICRF simulations, a well-known issue is the presence of numerical pollution in cyclotron resonances for low local parallel wavenumber $k_\parallel$, so that $(\omega - n\Omega_{ci})/k_\parallel v_{th} \gg 1$, where n is the harmonic resonance number, Ω_{ci} the ion cyclotron frequency, and v_{th} the thermal velocity. This effect has been reported in AORSA whenever the numerical solution presents slow waves with high perpendicular wavenumber[38]. A similar effect was reported in TORIC[5] regarding the appearance of spurious shear Alfvén modes for scenarios with low toroidal mode number (N_φ) and non-zero poloidal field ($B_{\rm pol}$).

More recently we reported the appearance of spurious modes in specific scenarios of the HHFW heating when using the quasi-local approximation in TORIC, where these scenarios were treated as outliers[27]. The outlier scenarios were identified automatically by a Gaussian-filtering process, and represented roughly a 23% of the cases explored in the study. These cases were reported to manifest for cases with low toroidal mode number and non-zero poloidal field. Sensitivity studies showed that the outliers could be mitigated or even eliminated by reducing the poloidal resolution, which could be attributed to limiting the lowest values of $k_\parallel$ in the poloidal spectrum explored that occur in the most negative values of the poloidal mode number (m). The outliers exhibit spike(s)–both in the electric field and power deposition–at arbitrarily located poloidal angles (θ) but consistently after specific harmonic resonances.

Thus, the HHFW heating profiles in outlier scenarios presented a portion characterized by a smooth power absorption observed in standard HHFW heating scenarios, and an outlier portion featuring a set of one or more spikes in the heating profile. In the absence of an operational physics-based model for these scenarios, which were already known to be challenging for the modeling point of view, we proposed an approach[27] based on an alternative application of surrogate models, commonly-used for accelerated inference. The surrogates were trained uniquely on physics-conforming data, a dataset curated by adequate outlier identification and filtering. Then we employed the trained models to infer profiles in outlier scenarios. An analysis of the predictions showed that for minor outliers, where the outlier portion of the signal was smaller than the magnitude of the physical absorption, inferred predictions were able to correct the profiles while maintaining the standard heating portion of the profile. On the contrary, with increased importance of the outliers, inferred predictions were also anticipating a higher heating in the standard heating region, especially in scenarios dominated by the numerical instability.

In this work we show the origin of the outliers found in the TORIC-HHFW model[27]–which had previously prevented us from modeling these challenging scenarios– and describe the correction applied in TORIC to generate a novel database for HHFW scenarios (see §II). The novel database is employed in §III (i) to verify the ML-based predictions in outlier scenarios presented in our previous work[27], showcasing applicability of surrogates to overcome certain numerical artifacts, and (ii) to generate a new set of models for the outlier-free HHFW database for NSTX. A discussion on the surrogates' applicability and the results obtained is detailed in §IV, followed by concluding remarks in §V.

II. METHODOLOGY AND DATASET

A. The origin of outliers in TORIC-HHFW

Wave propagation in hot magnetized plasmas with a Maxwellian distribution function is described by the hot plasma dielectric susceptibility tensor $\boldsymbol{\chi}_s$[39,40]. The plasma susceptibilities are a function of λ_s, which depends on the Larmor radius of species s (ρ_s) normalized with the perpendicular wavenumber ($k_\perp$) as

$$\lambda_s = \frac{1}{2}k_\perp^2\rho_s^2 = \frac{1}{2}n_\perp^2\frac{\omega^2}{\Omega_s^2}\frac{v_{ths}^2}{c^2}\,, \tag{1}$$

where c is the speed of light in vacuum, $n_\perp$ is the perpendicular refractive index, and Ω_{cs} and v_{ths} are the cyclotron frequency and thermal velocity of species s, respectively. The validity of the FLR approximation[40,41] used in TORIC is limited to $\lambda_s \ll 1$. Dedicated attention to monitoring scenarios with local high λ_s is necessary, particularly challenging for ion interactions with very short wavelength modes in ICRF plasmas, such as the ion Bernstein wave (IBW).

One of the features of TORIC–version integrated in TRANSP[7,42]–is the wavelength control (WLC) algorithm which addresses this IBW modeling challenge. In the IC minority heating regime it is well-known that fast wave (FW) modes feature linear mode conversion to IBW after the ion-ion hybrid resonance[40], which follows the dispersion relation[41,43]

$$n_\perp^2 = -\frac{n_\parallel^2 - S}{\sigma_{\rm FLR}}\,, \tag{2}$$

$$\sigma_{\rm FLR} = -\frac{1}{2}\sum_{s=i}\left(\frac{\omega_{ps}v_{ths}}{\Omega_{cs}c}\right)^2\zeta_{0s}^2 Z(\zeta_{2s})\,, \tag{3}$$

$$\zeta_{ns} = \frac{\omega - n\Omega_{cs}}{k_\parallel v_{ths}}\,, \tag{4}$$

$$Z(\zeta) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{e^{-u^2}}{u-\zeta}du + i\gamma\sqrt{\pi}e^{-\zeta^2}\,, \tag{5}$$

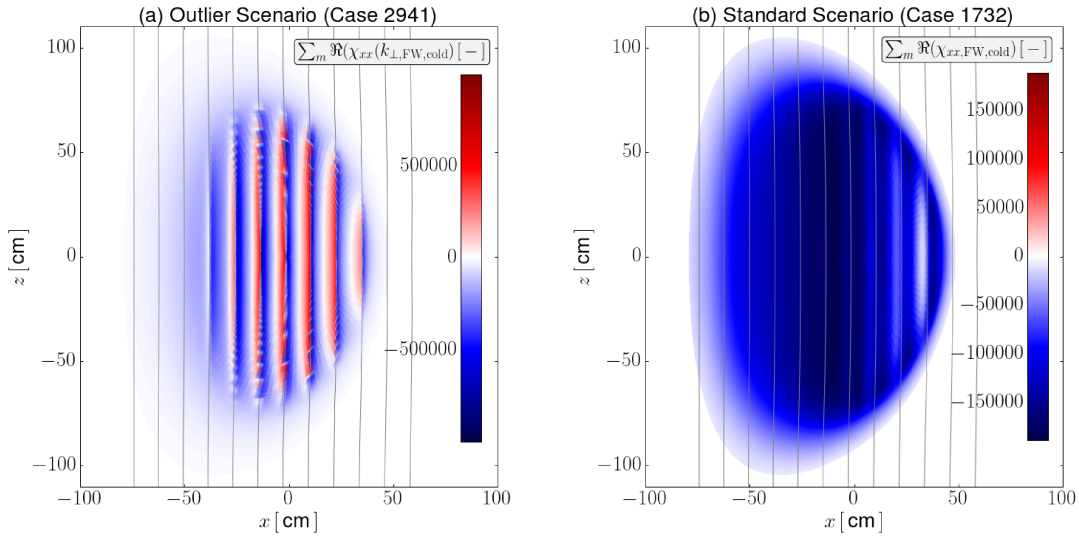

FIG. 1: Comparison of the sum of the real part of the hot dielectric tensor susceptibilities χ_{xx} over all poloidal mode numbers and evaluated using the perpendicular wavenumber of the FW cold-root for (a) an outlier scenario and (b) a standard scenario. The properties are obtained directly from TORIC. The figures' colorbars are forced to be symmetric to highlight sign reversal (from blue to red). The harmonic resonances are represented in grey.

with $n_\parallel$ the parallel refractive index, S the cold plasma tensor component in Stix notation[39], $Z(\zeta)$ the plasma dispersion function with $\gamma = 0$ for $\mathrm{Im}(x) > 0$, $\gamma = 1$ for $\mathrm{Im}(x) = 0$, and $\gamma = 2$ for $\mathrm{Im}(x) < 0$. In specific IC minority heating scenarios the IBW wavelength decreases rapidly such that ion $\lambda_{s=i} > 1$, and the WLC algorithm in TORIC can be used to ensure that the IBW wavelength is always greater than the radial mesh step. The algorithm increases the real part of σ_{FLR} by

$$\sigma^* = \sigma_{\mathrm{FLR}} + \delta\sigma(S_{\mathrm{FLR}}, d\psi)\,, \tag{6}$$

where $\delta\sigma \propto -S_{\mathrm{FLR}} d\psi^2$, with $d\psi$ is the flux coordinate step determining the radial resolution, and S_{FLR} the zero-order term of the FLR approximation contributing to ϵ_{xx}[40,41]. As $S_{\mathrm{FLR}} \sim S$, by replacing σ_{FLR} with σ^* in (2), the WLC algorithm reduces the $n_\perp$ of the IBW (i.e. wavelength increase).

In the HHFW regime higher order IBWs can also propagate after each harmonic resonance. The nth IBW (with $n \geq 2$) features a resonance at the nth harmonic resonance and a cutoff right after (i.e. higher magnetic field) the $(n+1)$ harmonic resonance (i.e. at $\omega = (n+1)\Omega_{ci}$)[40]. In this region a narrow frequency gap whose width decreases with n allows for the confluence of the FW and the nth IBW. Given the narrow frequency gap it is expected that the amount of power transferred to IBW waves is negligible in practically most HHFW scenarios[44]. To force IBW evanescence throughout the simulation domain, the WLC algorithm in TORIC forced a similar correction to that described in (6)

$$\begin{aligned}\sigma^* &= \delta\sigma\left(\chi_{xx}(k_{\perp,\mathrm{FW,cold}}), d\psi, m\right) \\ &\propto -|m|\chi_{xx}(k_{\perp,\mathrm{FW,cold}}) d\psi^2\,,\end{aligned} \tag{7}$$

where χ_{xx} is the xx component of the quasi-local conductivity kernel[6], or in other words, the hot dielectric tensor for Maxwellian plasmas evaluated using the local FW root obtained from the cold plasma dispersion relation[40]. Here the xx component refers to that defined in the Stix frame, where the coordinate system is oriented with the z-axis aligned with the background magnetic field and x, y lying in the perpendicular plane, and the wave vector lies in the x-z plane[39]. In this case then the coefficient $\delta\sigma$ is scaled by $|m|$, to force the evanescence of the IBW mode even for highly negative m, and more importantly, now proportional to the quasi-local conductivity kernel component.

In our exercise of generating a database for NSTX[27] we explored a variety of NSTX scenarios with parametric combinations unexplored prior to this work. Big-data analysis showed that these outlier scenarios were frequent and driven by plasma, magnetic, and excitation parameters. For these scenarios we find that the parametric combination is such that $\chi_{xx}(k_{\perp,\mathrm{FW,cold}}) \neq S$, especially in a narrow region right after harmonic resonances with high λ_i, where $\chi_{xx}(k_{\perp,\mathrm{FW,cold}}) > 0$. The effect can be understood as a modulation on top of a smooth FLR coefficient $S^{(0)}$ which increases for higher λ_i and lower $k_\parallel$, and that appears after the harmonic resonances. Figure 1 shows the net effect (i.e. summed over all poloidal modes) of this modulation with the properties obtained from TORIC, where the modulation can be observed to be present in both (a) an outlier scenario and (b) a standard scenario. However, the amplitude of the modulation is much larger than the negative value of $S^{(0)}$ inducing sign reversal in the quasi-local dielectric constant. As a result, in these scenarios, we find that the correc-

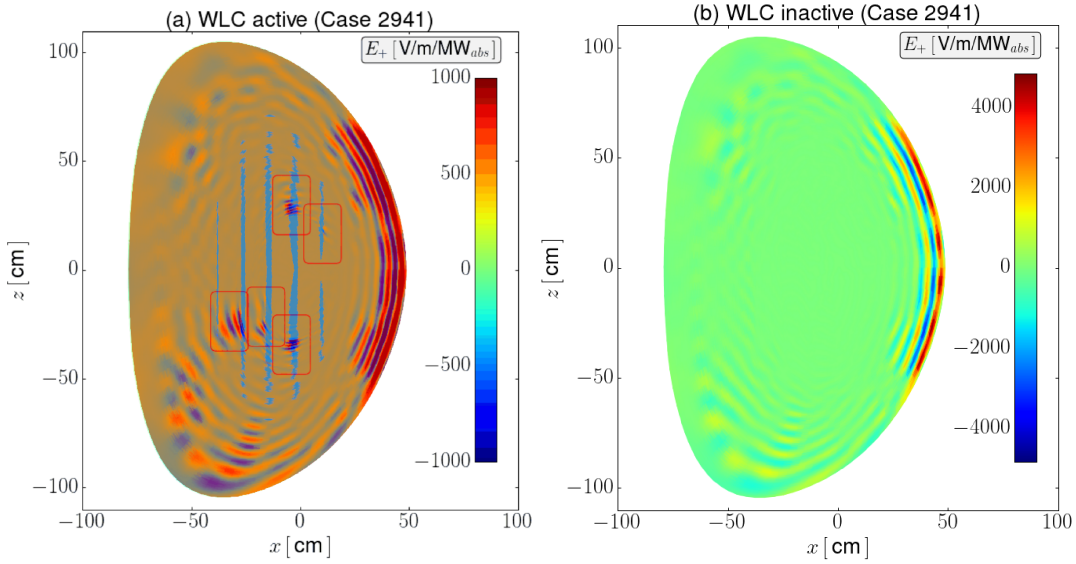

FIG. 2: Comparison of the effect of WLC algorithm on the left-hand polarized wave electric field result from TORIC. Results for WLC (a) active and (b) inactive are shown for an outlier scenario (case 2941: $N_\varphi = 5$, $n_{e0} = 1.1 \times 10^{20}$ m^{-3}, $T_{e0} = 4.9$ keV and $\alpha = 3.8$). Results with WLC active are shown from [-1000, 1000] V/m/MW$_{\text{abs}}$ to highlight the IBW-like modes characteristic of an outlier scenario (red rectangles). A shadowed blue region represents the region of $\sigma^*(m = -31) < 0$, correlated with the IBW-like modes appearance. Actual range of the E_+ field in (a) is [-2095, 1960] V/m/MW$_{\text{abs}}$.

tion $\delta\sigma$ changes sign and so does σ^*, thus allowing linear mode conversion to an IBW-like mode where $\sigma^*(m) > 0$. Figure 2 shows a comparison of the left-hand polarized component of the electric field solution from TORIC, for an outlier scenario where the wavelength control algorithm is (a) active and (b) inactive. With WLC active, TORIC applies a correction to the real part of the wave equations coefficients that results in the IBW-like spurious modes shown in red rectangles in Fig. 2(a). We also highlight in cyan the shadow where the sign of $\delta\sigma$ of the lowest poloidal mode (i.e. $m = -31$ for our simulations) is negative, and thus the IBW-like mode can appear. On the contrary, when WLC is inactive, the results exhibit an outlier-free electric field solution with a FW propagation from the low field side towards the core. We term these outlier modes as IBW-like since they cannot be adequately described by the algorithm, and accurate representation of the mode conversion to nth IBWs would require a numerical algorithm handling higher-order derivatives. The WLC routine–originally designed to force IBW evanescence in HHFW scenarios–is operating inadequately in these scenarios and causing linear mode conversion of FW to spurious IBW-like modes. For future users, we recommend deactivating WLC for HHFW scenarios, and use it only in IC minority heating scenarios featuring under-resolved IBW modes, which can occur particularly in the high-field side.

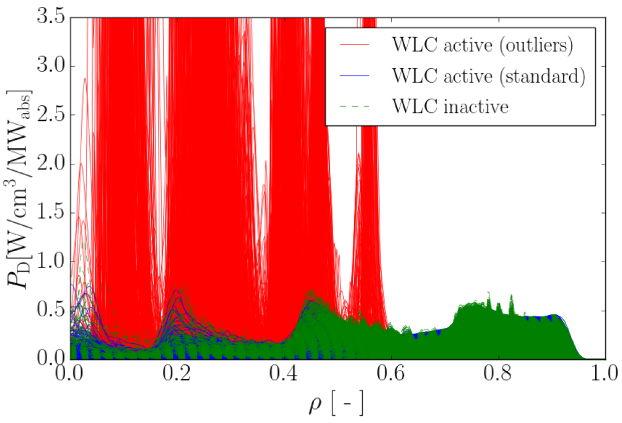

FIG. 3: Deuterium power absorption profiles in the HHFW database generated with WLC active, showing outliers (red solid) and standard (blue solid) cases, compared to those obtained with WLC inactive (green dashed). Note the vertical axis has been truncated for visualization purposes; the actual maximum value of the outlier profiles reaches 47.8 W/cm^3/MW$_{\text{abs}}$.

B. An outlier-free HHFW dataset for NSTX

We repeat the database shown in our previous work for the HHFW in NSTX[27], in the exact same parametric space combinations, to verify that no outliers are present in any of the numerical solutions from TORIC when the WLC algorithm is inactive. For simplicity, as a first ap-

proximation we assume that the plasma property (e.g. electron density n_e) profiles with respect to normalized radius ρ follow the relation

$$n_e = n_{e1} + (n_{e0} - n_{e1})(1 - \rho^{\alpha})^{\beta}\,, \tag{8}$$

where subscripts 0 and 1 refer to core and edge values, and α and β are the inner and outer profile exponents, respectively. The input parameters scanned across the database are: (i) the toroidal mode number ($N_\varphi \in [5-21]$), (ii) the core plasma density ($n_{e0} \in [0.5-2]\times 10^{20}\mathrm{m}^{-3}$), (iii) the core electron temperature ($T_{e0} \in [1-5]$ keV) and (iv) the inner profile exponent ($\alpha \in [2-10]$). Another assumption is that the plasma equilibrium is fixed[45]. A Latin hypercube sampling method is used to generate the database representing the selected parametric space.

For each parametric combination we use TORIC to compute the outputs of this direct problem, which are the 1D electron and deuterium HHFW heating profiles. The novel database is comprised of 12240 simulations. Figure 3 shows a comparison of the deuterium power absorption profiles obtained with WLC active which are outliers (red solid) or standard scenarios (blue solid), against the TORIC database results with WLC inactive (green dashed). All deuterium power absorption profiles comprised within the new database (green) show a standard scenario behavior (blue), and therefore no trace of outlier features in the database with WLC inactive. Note that the figure has been truncated for comparison purposes and that the actual maximum value of the outlier deuterium power deposition profiles is 47.8 $\mathrm{W/cm^3/MW_{abs}}$. An equivalent effect is observed in the electron heating profiles and it is not shown here for the sake of brevity.

C. Machine learning methodology

With a novel dataset free of numerical artifacts we implement surrogate models for two main applications: (i) overcoming numerical artifacts providing physical predictions, and (ii) a final set of surrogates. The models trained here are based on the RFR, MLP and also, the Gaussian Processes Regressor[46] (GPR) methodology based on TensorFlow and GPFlow[47] implementation. The RFR is an ensemble-based learning method that averages the predictions of multiple decision trees, providing robust and interpretable regression outputs. The MLP is a feedforward artificial neural network that uses layers of interconnected nodes that can learn complex nonlinear relationships between input and output data. For further details on these methods, we refer the reader to our previous publication[27] and relevant literature[34–37]. The GPR is a ML method for regression tasks, which defines a distribution over a functional space $\mathcal{GP}\{\mu(\mathbf{x}), \mathcal{K}(\mathbf{x},\mathbf{x}')\}$, characterized by a mean function $\mu(\mathbf{x})$ and a kernel $\mathcal{K}(\mathbf{x},\mathbf{x}')$. Given N training points $\mathbf{x} = \{\mathbf{x}_i \in \mathbb{R}^d\}_{i=1}^N$ and their corresponding observations $\mathbf{y} = \{y_i = g(\mathbf{x}_i) \in \mathbb{R}\}_{i=1}^N$, the model infers a latent function $g : \mathbb{R}^d \to \mathbb{R}$. This function is trained to infer and capture the relationship between the input and output variables in the form of a Gaussian probability distribution. Standard ML regression methods, such as RFRs and standard MLPs, are deterministic, meaning that they provide a single-point prediction for a given input. In contrast, GPR models are probabilistic models that estimate a probability distribution for the outputs. Specifically, GPR provides a Gaussian probability distribution characterized by a mean value (μ) and a standard deviation (σ), allowing for both point predictions and uncertainty quantification (UQ)[48].

TABLE I: Optimal hyperparameters, the variance (σ^2) and the lengthscales (l), for the final GPR models.

Target	$\mathcal{K}_{\mathrm{linear}}\,(\sigma^2)$	$\mathcal{K}_{\mathrm{Matern32}}\,(\sigma^2)$	$\mathcal{K}_{\mathrm{Matern32}}\,(l)$
P_e	1317.9	16917.7	[81.0, 46.5, 49.3, 54.6]
P_{D}	0	43977.5	[98.2, 48.5, 47.7, 44.9]

To train these models we choose a kernel (covariance) function $\mathcal{K}$ as

$$\mathcal{K} = \mathcal{K}_{\mathrm{linear}} + \mathcal{K}_{\mathrm{Matern32}}, \tag{9}$$

where $\mathcal{K}_{\mathrm{linear}}$ and $\mathcal{K}_{\mathrm{Matern32}}$ are a linear and a Matern kernel[49] with $\nu = 3/2$, based on the results shown in Wallace[26]. Interestingly, as we will show in the next section, the hyperparameter tuning of RFR and MLP models for both electron and deuterium absorption made for the filtered dataset generated with WLC active also result in successful training for the surrogates trained on the new dataset. For the GPR model we tune the other hyperparameters (length-scale and variance) using the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with box constraints (L-BFGS-B), as implemented in the `scipy.optimize.minimize` function from SciPy[50]. Overall, the length-scale addresses the smoothness and the variance the scale of the values of the learnt functions. Table I shows the results of the hyperparameter tuning of the final models.

In order to estimate the accuracy of the surrogate models trained for regression, we here use two standard regression accuracy metrics such as the mean squared error (MSE) and the coefficient of determination (R^2) in scikit-learn[51], defined as

$$\mathrm{MSE} = \frac{1}{N_y}\sum_{j=1}^{N_y}\frac{1}{N}\sum_{i=1}^{N}(y_{ij} - \hat{y}_{ij})^2\,, \tag{10}$$

$$\mathrm{R}^2 = \frac{1}{N_y}\sum_{j=1}^{N_y}\left(1 - \frac{\sum_{i=1}^{N}(y_{ij} - \hat{y}_{ij})^2}{\sum_{i=1}^{N}(y_{ij} - \bar{y}_j)^2}\right), \tag{11}$$

where y_{ij} and $\hat{y}_{ij}$ represent the ground truth and predicted value of output variable j for case i, $\bar{y}_j$ is the mean (amongst all cases) of the ground truth values for output variable j, N_y is the number of output variables, and N the number of cases.

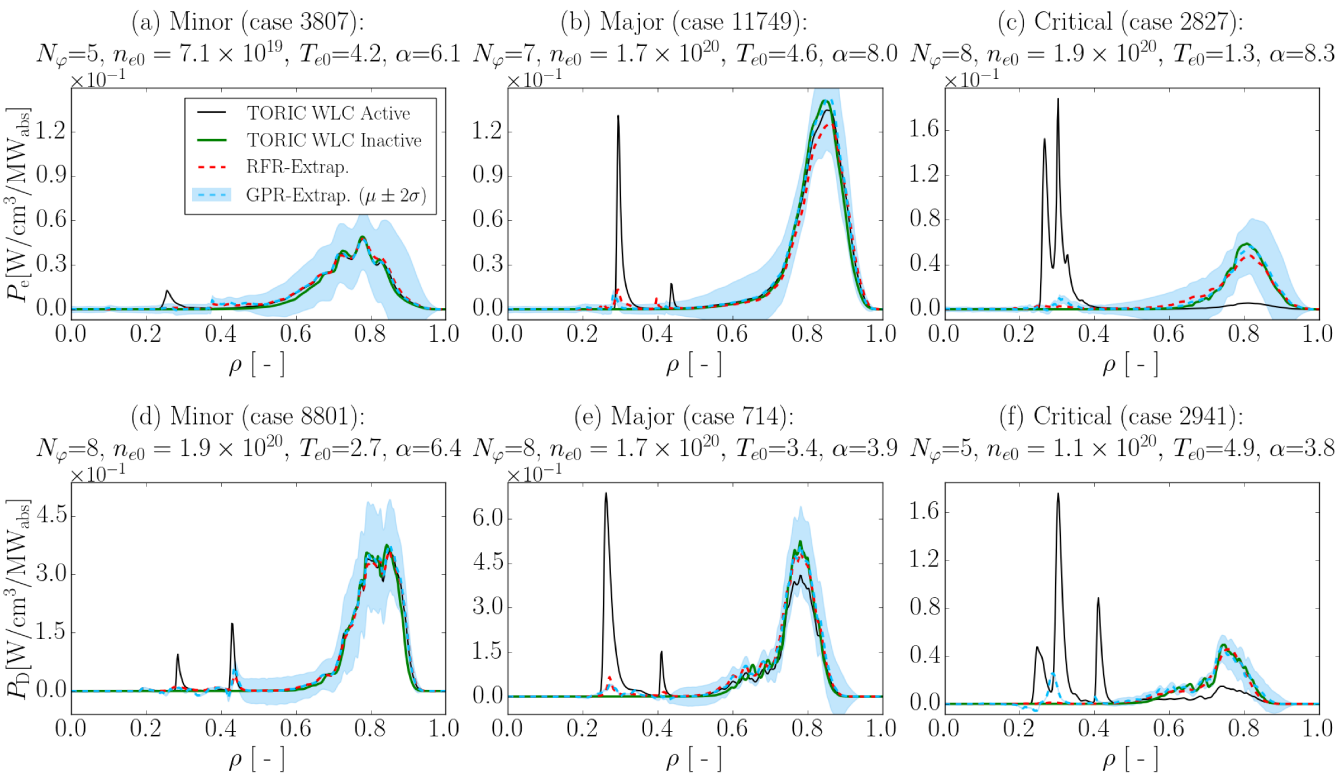

FIG. 4: TORIC-ML RFR (red dashed) and GPR (light-blue dashed) predictions of electron (top) and deuterium (bottom) power deposition density in (a,d) minor, (b,e) major, and (c,f) critical outlier scenarios compared to TORIC solutions with WLC active (black solid) and inactive (green solid). The 95% confidence level in the GPR prediction is added on top of the GPR mean predicted value. Key input parameters (n_{e0}[m^{-3}] and T_{e0}[keV]) are shown in the title for each case.

III. ML-BASED ICRF HEATING MODELS FOR NSTX

The novel outlier-free HHFW-NSTX database generated using TORIC with WLC inactive allows for verification of the inferred heating profiles obtained by the surrogate models trained with TORIC WLC active filtered-data. This filtering, reported in Ref.[27], was carried out using an outlier identification metric based on comparison of Gaussian-filtered profiles with raw profiles. Here then, we employ the results from TORIC with WLC inactive to verify the capabilities of these surrogates to predict physical profiles even in the presence of numerical artifacts in the simulation code used for the training.

Figure 4 shows a comparison of the RFR (red dashed) and GPR (light-blue dash-shaded) predictions compared to both the estimations of TORIC with WLC active (black solid) and inactive (green solid). The results are shown for both electron (top row) and deuterium (bottom row) power deposition predictions for, from left to right, minor (Fig. 4a and d), major (Fig. 4b and e) and critical (Fig. 4c and f) non-conformance outliers. The outlier non-conformance refers to the relative importance or maximum value of the outlier portion of the profile compared to the physical one. For reference, we show the comparison for the same scenarios as depicted in our previous results[27] (i.e. cases 3807, 11749, 2827, 8801, 714, and 2941). Note that Figure 4(f) shows scenario 2941; its corresponding left-hand polarized electric wave field component E_+ is shown in Fig. 2, obtained with (a) active and (b) inactive WLC. In all scenarios, TORIC's results with WLC inactive (green solid) show no outlier (sharp) features in both electron and deuterium heating profiles, coherently to what was shown in Fig. 3. However, more importantly, the extrapolated inferred predictions show the capability to predict the heating profiles with WLC inactive, even without being trained on them. Thus, the physics of these models captures the physics-conforming version of the model (to certain degree of accuracy) even in these challenging scenarios. We demonstrate here that unlike MLP models[27], and in addition to RFR models, GPR models can also be employed to extrapolate accurately. As the importance of these numerical artifacts increases we observe that the heating predicted by the surrogates differed from that of the original model. We here verify that the surrogates are not only able to eliminate the outlier features from the inferred heating profiles, thus improving their robustness, but also to estimate correctly the difference in heating caused by the numerical artifact. This effect can be ob-

TABLE II: Regression accuracy of extrapolated inference of surrogates trained with filtered data from TORIC with WLC active, against TORIC values without outliers (i.e. WLC inactive). Regression accuracy metrics shown are for mean values (μ) of the R^2 and MSE.

Target	ML-model	μ(MSE) $[(W/cm^3/MW_{abs})^2]$[a]	$\mu(R^2)$ -
P_e	RFR	4.2×10^{-5}	0.58[b]
P_e	GPR	6.8×10^{-6}	0.89
P_D	RFR	4.4×10^{-4}	0.94
P_D	GPR	3.6×10^{-4}	0.96

[a] TORIC's power deposition units are in terms of MW of absorbed power.

[b] 83.9% cases feature higher R^2 than the mean value $\mu(R^2)$. The median value for the RFR is R^2=0.948.

TABLE III: TORIC-ML surrogate scoring metrics for each target species, dataset, and ML-method used. The test R^2 and MSE are shown, including training coefficient of determination (R^2_{tr}) for reference. Surrogate's training times (t_{tr}) are shown for one single thread.

Target	Dataset	Method	R^2_{tr}	t_{tr} [s]	R^2	MSE
P_e	NSTX	RFR	0.99	3	0.96	1.7×10^{-5}
P_e	NSTX	MLP	0.97	4	0.97	2.2×10^{-5}
P_e	NSTX	GPR	0.99	2810	0.99	2.1×10^{-6}
P_D	NSTX	RFR	0.99	3	0.93	1.2×10^{-4}
P_D	NSTX	MLP	0.98	21	0.98	6.6×10^{-5}
P_D	NSTX	GPR	0.99	4015	0.99	8.6×10^{-6}

served in the difference between the TORIC WLC Active curve (black) and WLC Inactive curve (green) in the standard (Gaussian-like) portion of the heating profiles of electrons and deuterium, to a minor degree in major outliers (see Fig. 4(b,e)), and more evidently in critical outliers (see Fig. 4(c,f)).

The GPR model also provides uncertainty quantification via the standard deviation in the profile inferred by the GPs, which is shown as a 95% confidence interval in Fig. 4 (blue shaded area). The uncertainty in GPR predictions tends to be smaller towards the core, compared to for instance at an intermediate radius or at the edge, where it is often higher. This denotes that the variance in the dataset is higher in the edge than in the core, and so is the sensitivity to the selected input parameters.

In order to better characterize the extrapolation accuracy of these surrogates to anticipate heating in outlier scenarios we compute the mean of R^2 and the MSE, where the ground truth used is the result of TORIC with WLC inactive, and the prediction is that of the RFR and GPR models trained with the TORIC WLC active dataset filtered of outliers. Table II summarizes their corresponding regression accuracies. For electrons, the RFR and GPR models are both able to reproduce the profiles with a fair to excellent accuracy. Amongst the two models, the GPR model seems to capture more accurately the profile, showing an excellent regression score of $\mu(R^2)$ = 0.89. On the contrary, the RFR model shows a hindered performance $\mu(R^2)$ of 0.58. This effect can also be noticed in the mean MSE of the profiles (μ(MSE)) for electrons, which is one order of magnitude larger for the RFR, 4.2×10^{-5}, than for the GPR, 6.8×10^{-6}. Analyzing the statistics of RFR P_e accuracy we observe that 83.9% of the RFR extrapolated P_e heating profiles feature higher R^2 than the mean $\mu(R^2)$, and 80% a value above R^2 = 0.7, with the median being R^2=0.94. The regression accuracy of extrapolated inference is observed to be higher for deuterium absorption, where both RFR and GPR surrogate models are capable to capture the physics of ion heating accurately while being trained with a filtered WLC algorithm data. RFR and GPR feature a $\mu(R^2)$ = 0.94 and 0.96, respectively. This comparable accuracy can also be observed in terms of MSE values, where the RFR model accuracy is 4.4×10^{-4} compared to that of the GPR, 3.6×10^{-4}. This difference in MSE values between electron and deuterium heating is primarily due to the different scales of the targets. While R^2 provides a normalized, unitless measure of relative fit, MSE captures absolute error in the units of the predicted quantity squared. As deuterium heating profiles have higher magnitudes than electron heating, they naturally yield higher MSE values, even when the relative prediction quality is comparable. Note that although GPR predictions denote increased sensitivity to very minor non-conformance outliers still included in the training dataset (see Figs. 4(d) or (f) close to the location of the spikes), these are still able to capture HHFW heating in these scenarios and more accurately than the RFR models.

Beyond the above-mentioned capabilities, with the new dataset we train a set of surrogate models based on RFR, MLP, and now also GPR, to provide real-time capable predictions of both electron and deuterium HHFW heating in NSTX. Table III shows the summary of the performances for these models. As can be noticed from the R^2 and MSE values achieved, the surrogates show excellent regression accuracies for all models and in both predicting electron and deuterium heating. For electron heating predictions the accuracies are R^2=0.96, 0.97, and 0.99 for the RFR, MLP and GPR models, respectively. Similarly for the deuterium ion heating predictions the accuracies are R^2=0.93, 0.98, and 0.99 for the RFR, MLP and GPR models, respectively. Therefore, for both species, the GPR models feature the highest accuracy, followed by the MLP models, and lastly the RFR models. Thus, these models show no evidence of underfitting, given their high test regression scorings. Moreover, by comparing both the training and test set coefficient of determination (R^2_{tr} and R^2), we can conclude that no model presented here feature any level of overfitting.

TABLE IV: Cross validation of the multiple models using MSE. MSE mean (μ) and standard deviation (σ) of the 5-fold for each model are compared to the corresponding MSE of the final model developed (Test).

Target	Data	Variable	RFR	MLP	GPR
P_e (NSTX)	5-fold	μ(MSE)	2.1×10^{-5}	2.3×10^{-5}	1.4×10^{-6}
		σ(MSE)	2.9×10^{-6}	3.8×10^{-6}	2.0×10^{-7}
	Test	μ(MSE)	1.7×10^{-5}	2.2×10^{-5}	2.1×10^{-6}
P_D (NSTX)	5-fold	μ(MSE)	1.2×10^{-4}	6.9×10^{-5}	1.3×10^{-5}
		σ(MSE)	8.8×10^{-6}	1.0×10^{-5}	1.2×10^{-6}
	Test	μ(MSE)	1.2×10^{-4}	6.6×10^{-5}	8.6×10^{-6}

Another important performance metric of the surrogates is their training time. As can be noticed, the training process of RFR and MLP models takes 3-21 s while for GPR models it takes 2810 and 4015 s for electron and deuterium predictions, respectively. That is a significant difference, featuring 3 orders of magnitude increased training time for the GPR models (e.g. $\mathcal{O}$(h)). Regarding the inference time, t_I, we note that the exact values may vary depending on the hardware and system load at runtime. MLPs are the fastest ($\sim 10\mu$s), RFRs are roughly 5-10 times slower ($\sim 100\mu$s), and GPR models are the slowest of all architectures ($\sim$ 1ms). While GPR models can be 50-100 times slower than MLPs in average inference time, they remain within real-time feasible limits. These estimates are consistent with timings reported in our previous work[27]. In our observations, inference times tend to show more variability than training times, likely due to their sensitivity to system load and operating system scheduling effects at short time scales.

Table IV summarizes the results of the cross-validation process carried out for all the surrogates via a five-fold. This process is ensures that the randomization in the data splitting did not affect surrogate performance of the final models. For that we take a 5-fold split of the training dataset and permute the validation set (1 out the 5) and use the rest for training. The results of this cross-validation process is shown for the RFR, MLP, and GPR models compared to the mean value of the final surrogates. The comparison is made taking the MSE of each model and computing its mean μ(MSE) and standard deviation σ(MSE) of the MSE. All cross-validations show that each model's training has been made independently of the data splitting. This is observed first by the fact that condition $\sigma(\text{MSE}) \ll \mu(\text{MSE})$ holds for all models. Additionally, the mean error of the 5-folds is consistently comparable to that of the test set MSE in the definitive models.

Figure 5 shows the comparison of the electron (top row) and deuterium (bottom row) HHFW heating estimations of TORIC with WLC inactive (green solid) with the RFR (red dashed), MLP (blue dashed), and GPR (lightblue dashed and shaded) surrogate predictions. From left to right, we show scenarios with increasing MSE of the RFR model, the one featuring the lowest regression accuracy of the three algorithms. Figures 5 (a,d) show the best scenarios, where all the models are able to capture accurately the physics of HHFW heating, which are strong single-pass absorption scenarios featuring electron Landau damping for the electrons and high harmonic absorption for the ions. In these scenarios the HHFW power is absorbed in the first pass through the plasma column, avoiding multiple reflections. Figures 5 (b,e) show the comparison in the typical heating scenarios, that is in the median of the MSE of the RFR model. Here we observe the same type of scenario, where no significant differences can be observed in terms of surrogates' accuracy. This implies that all the models are generally able to accurately reproduce the physics of HHFW heating described by TORIC in the parametric range considered. Figures 5 (c,f) shows the worst type of prediction, which is characterized by a non-negligible heating across the plasma column and strong variations, even in the vicinity of the plasma core. These scenarios denote multi-pass absorption, where the wave propagation front crosses the plasma column and is reflected and bounced back, performing multiple passes across the plasma. In these more intricate scenarios we can observe that RFR models are able to capture the overall trends of the HHFW heating but often misestimate in the amplitude of the heating. This effect is more subtle for the MLP regressors, where the accuracies are improved. However, the most accurate regressions are featured by the GPR models, which are able to accurately reproduce the heating profiles even in the worst cases. These models also provide a quantification of the uncertainty, represented by the 95% confidence interval (light blue shaded area) which can be observed to be smaller for the electron predictions than for the ions, and are significantly larger in deuterium scenarios closer to the plasma edge. However, this metric is case-dependent and further investigation is still required.

IV. DISCUSSION

The capability to accurately reproduce the one-dimensional ICRF heating profiles from the surrogates developed can be observed in Fig. 5 as well as it is reflected in the surrogates' performances reported in Table III. For instance, the new models reported show an improvement in regression accuracy in comparison with our previous work[27], achieving $R^2 \in [0.93 - 0.99]$. Thus, this work establishes a simplified yet effective methodology for implementing surrogate models to reproduce one-dimensional ICRF heating predictions. The RFR, MLP, and GPR models showcase different algorithms capable of delivering accelerated inference of ICRF heating profiles, reducing prediction times from minutes–required by TORIC or other ICRF heating codes– to microseconds or tens of microseconds. This reduction in computational time allows the use of the physics embedded within these computational codes to streamline the optimization of

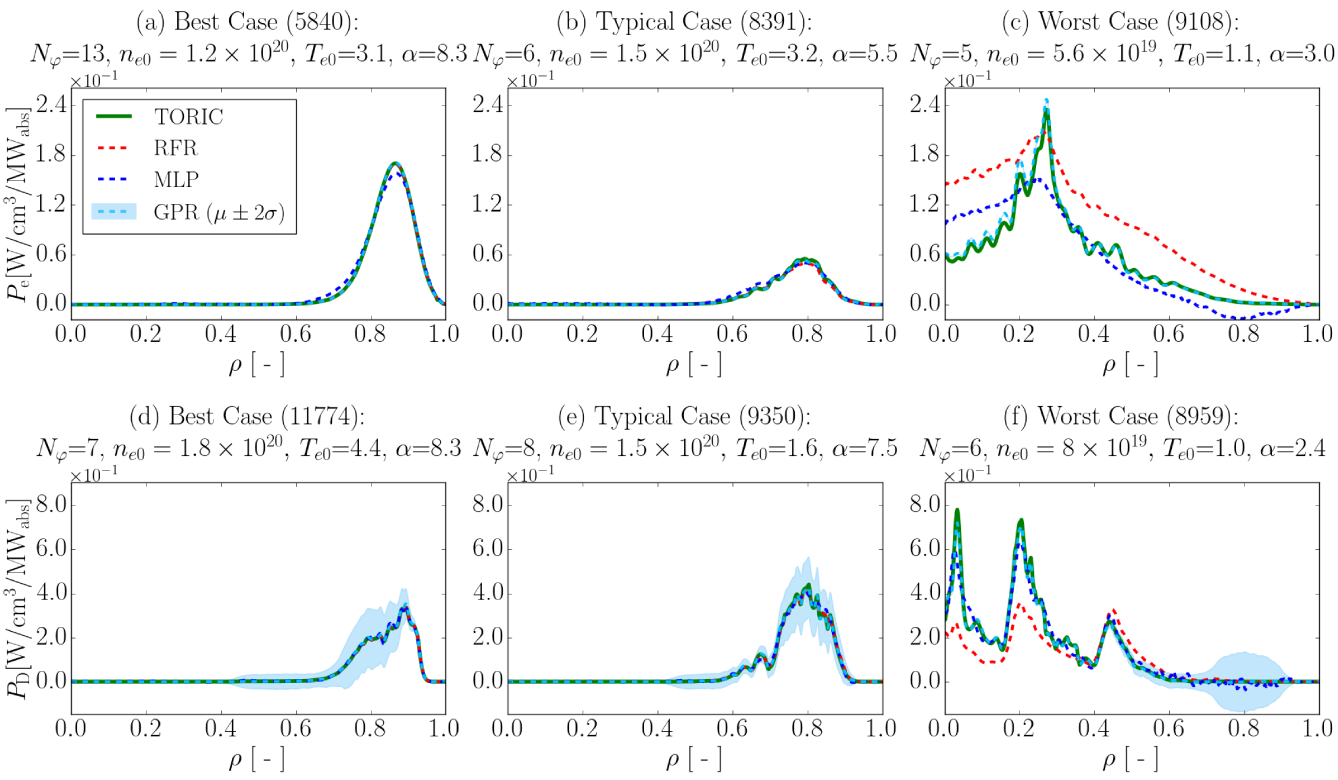

FIG. 5: TORIC-ML RFR (red dashed), MLP (blue dashed) and GPR (light-blue dashed) predictions of electron (top) and deuterium (bottom) power deposition density in (a,d) best, (b,e) typical, and (c,f) worst cases in terms of RFR MSE value. TORIC solutions are shown with with WLC inactive (green solid). Key input parameters (n_{e0}[m^{-3}] and T_{e0}[keV]) are shown in the title for each case.

specific scenarios, carry out inter-shot predictive modeling, or enable their incorporation to real-time control frameworks.

Moreover these architectures provide a catalog with necessary trade-offs to be carried out by the developer. RFR models offer simplified implementation, specifically in terms of hyperparameter tuning and low training times, and exhibit high accuracies, making them an ideal starting point for novice developers. To achieve the highest accuracies, GPR models are the most desirable. However, their development is firstly hindered by a complex hyperparameter tuning process and, more importantly, by the required training times which can scale up to more than three orders of magnitude with respect to the other architectures. MLP-based regressors provide a middle ground, offering good balance between training cost and prediction accuracy. In terms of inference time, although GPR models are roughly two orders of magnitude slower than MLPs, their total prediction time remains under 1 ms, which makes them compatible with real-time application requirements. This trade-off may be acceptable given the added benefit of uncertainty quantification. Each model architecture thus serves a distinct role, depending on the application priorities and available resources.

The GPR models—shown predicting accurately both outlier scenarios when trained on filtered WLC-active TORIC data in Fig. 4, and cases from the WLC-inactive TORIC dataset (free of outliers) in Fig. 5—also provide a standard deviation along with each prediction, which serves as a metric for uncertainty quantification (UQ). UQ is essential to provide reliable, robust and interpretable predictions, as it reflects the noise and inheret variability in the dataset. In practice, we report uncertainty as a 95% confidence interval (i.e. $\mu \pm 2\sigma$) around the predicted mean μ, as shown shadowed for the GPR predictions in Figs. 4 and 5. As a practical guideline, predictions with a signal-to-noise/uncertainty ratio (SNR), defined as μ/σ, below approximately 3.3 (i.e. $\sigma/\mu > 0.3$) are considered increasingly unreliable. Although this threshold is not absolute, it provides a useful measure for identifying low-confidence predictions, particularly in extrapolated scenarios. In contrast, predictions with SNR>5 are generally regarded as trustworthy. By comparing the degree of uncertainty observed in interpolated inference (see Fig. 5) and extrapolated inference (see Fig. 4), we observe that the uncertainty increases for extrapolated inference. This indicates that although effective, the extrapolation capabilities might be limited by the "distance" from the data or information used during training.

In this work we show that UQ is intrinsically provided

by the Gaussian Process Regressors due to their statistical or Bayesian-based nature. In contrast, deterministic models such as the Random Forest Regressors and the standard Multi-Layer Perceptrons are unable to offer such probabilistic insights directly. However, development of ensembles for MLP and the use of the different tree predictions of the RFR model could be employed to extract the statistics of the inferred profiles for those architectures. Here, for simplicity, we focus on the UQ capabilities of GPR models. By combining UQ with cross-validation and the verification of the surrogates accuracy, we present a comprehensive Verification, Validation, and Uncertainty Quantification (VVUQ) methodology. Such approach is critical for the development of robust and reliable surrogate models.

The results shown in Fig. 4, further quantified by the accuracy metrics shown in Table II, highlight the practical value of training surrogate models on curated, physics-consistent subsets of data, even when some regimes are excluded due to numerical artifacts. Such models can still provide reliable predictions in regimes affected by simulation instabilities, offering a robust strategy for handling imperfect data–a scenario common in high-fidelity computational physics workflows. By leveraging surrogates, we can achieve accurate predictions in scenarios where the original model exhibits numerical instabilities, such as that of outliers described in §II. Specifically, we demonstrate that surrogates can not only remove the outlier features, smoothing over them, but also accurately predict the physical portion of the heating profiles.

These results confirm that the surrogate models trained only on filtered data can, in some scenarios, successfully generalize to previously excluded regimes affected by numerical artifacts. This behavior supports the viability of using carefully curated datasets to build robust surrogate models even in the absence of full domain coverage. This correction has been validated using an improved version of the original model that is free of such instabilities. While surrogate models are widely recognized for accelerated interpolative inference, our findings suggest they may also generalize effectively to nearby scenarios not present in the training set.

A promising direction is to explore whether surrogate models trained on data from one device or configuration, such as HHFW heating in NSTX, can be adapted to predict behavior in the upgraded device, NSTX-U. Such an approach, conceptually related to transfer learning, could significantly extend the applicability of surrogate models across different machines or experimental conditions. As new fusion devices are developed, predictive modeling becomes essential for guiding experiments and anticipating device performance. In this context ML models that are progressively refined with incoming experimental data may help bridge the gap between idealized first-principle simulations and real plasma behavior. Over time, this adaptive framework could enhance the accuracy and relevance of surrogate models for emerging devices.

V. SUMMARY AND CONCLUSION

In this work, we developed surrogate models for ICRF heating that are not only real-time capable, but also demonstrably robust in regimes previously affected by numerical artifacts. Building on our earlier work[27], we identified the root cause of spurious outliers in HHFW scenarios as IBW-like modes, which arise due to the wavelength control (WLC) routine in TORIC. The algorithm controls the coefficient of the FLR equations for HHFW scenarios based on the value of the quasi-local xx-component of the dielectric tensor, which we here showed that can exhibit sign reversal after harmonic resonances under certain conditions. We modified the TORIC code to disable WLC in HHFW scenarios and generated a new HHFW-NSTX database, which we confirmed to be free of such artifacts and provides a recommended change in the code and setup for future simulations.

This new database not only enabled the training of final surrogate models, but also served to verify the predictions of RFR models trained on a curated subset of WLC-active data[27], where outlier scenarios were removed using a outlier identification metric based on Gaussian filtering. Despite never having seen these outlier cases during the training, the surrogate models successfully predicted heating profiles in those challenging scenarios. The predictions not only eliminated the spurious features but also correctly estimated HHFW heating in the low-field-side region characteristic of single-pass absorption. While the original TORIC simulations underestimated heating in that region, the surrogate predictions closely matched those from the corrected, artifact-free simulations. These results demonstrate a practical and generalizable approach for using machine learning surrogates to overcome simulation-induced artifacts via data curation, surrogate architecture selection, and targeted verification. We showed that unlike MLP models, GPR-based models also exhibit this capability, in addition to the RFR models.

We demonstrate that GPR models can be successfully trained to provide accurate ICRF heating predictions, achieving the highest accuracy among all architectures investigated in this work—albeit with significantly longer training times compared to RFR and MLP models. These models also provide a standard deviation in the prediction, and thus means for UQ. This allows for an improved understanding of the model reliability, particularly when extrapolating to parametric regimes or scenarios unseen during training, where in fact an increased prediction uncertainty is observed. The addition of UQ to these models enhances interpretability of the predictions, and in combination with the verification and cross-validation methodology presented, constitutes a comprehensive VVUQ methodology in surrogate modeling. The methodology presented here is not limited to ICRF heating but is extendable to other RF heating problems and fusion physics challenges requiring surrogate modeling. Future work will focus on expanding the application of

these methodologies, exploring the potential capabilities of ML-based surrogates to complement physics-based computational models. Beyond their well-known accelerated inference capabilities, these models also demonstrate promising generalization behavior in scenarios not represented in the training set. The ability to predict outcomes in numerically challenging regimes—where the original simulations exhibit limitations—offers a fast and practical alternative to full-wave modeling. Such models may also provide a foundation for transferability across different operating scenarios or devices. Additionally, while the present work relies on a Maxwellian assumption for the hot plasma dielectric, future extensions of this surrogate modeling approach to incorporate non-Maxwellian distribution functions—such as those obtained from kinetic or Fokker–Planck solvers—would enable more accurate modeling of minority species and broaden the applicability of the framework. However, these applications require careful consideration of the underlying physical constraints, training data coverage, and the uncertainty associated with the predictions.

Overall, ML-based surrogate models present an opportunity to complement physics-based tools, enabling the development of advanced simulation frameworks that support the accurate design and operation of high-performance fusion reactors. Further work will focus on the automated surrogate training and optimization and approach to investigate IBW mode conversion in these challenging scenarios.

ACKNOWLEDGMENTS

The authors A. Sanchez-Villar, N. Bertelli, and S. Shiraiwa thank R. Bilato and M. Brambilla for the fruitful discussion on the TORIC code and the outlier scenarios found. This work was supported by the U.S. Department of Energy under contract number DE-AC02-09CH11466. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award FES m3716 for 2023. The United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Princeton Data Commons at https://doi.org/10.34770/h8gb-be77.

REFERENCES

[1]F. W. Perkins, Nucl. Fusion **17**, 1197 (1977).
[2]F. W. Perkins, IEEE Trans. Plasma Sci. **12**, 53–63 (1984).
[3]J. R. Wilson and P. T. Bonoli, Phys. Plasmas **22**, 021801 (2015).
[4]P. Lamalle, Plasma Phys. Control. Fusion **39**, 1409 (1997).
[5]M. Brambilla, Plasma Phys. Control. Fusion **41**, 1 (1999).
[6]M. Brambilla, Plasma Phys. Control. Fusion **44**, 2423 (2002).
[7]M. Brambilla and R. Bilato, Nucl. Fusion **46**, S387 (2006).
[8]E. F. Jaeger, L. A. Berry, E. D'Azevedo, D. B. Batchelor, and M. D. Carter, Phys. Plasmas **8**, 1573–1583 (2001).
[9]E. F. Jaeger, L. A. Berry, E. F. D'Azevedo, R. F. Barrett, S. Ahern, D. W. Swain, D. B. Batchelor, R. Harvey, J. Myra, D. D'ippolito, *et al.*, Phys. Plasmas **15** (2008).
[10]R. J. Dumont, Nucl. Fusion **49**, 075033 (2009).
[11]S. Shiraiwa, J. C. Wright, P. T. Bonoli, T. Kolev, and M. Stowell, EPJ Web Conf. **157**, 03048 (2017).
[12]N. Bertelli, S. Shiraiwa, and M. Ono, Nucl. Fusion **62**, 126046 (2022).
[13]D. Humphreys, A. Kupresanin, M. D. Boyer, J. Canik, C. Chang, E. C. Cyr, R. Granetz, J. Hittinger, E. Kolemen, E. Lawrence, *et al.*, J. Fusion Energ. **39**, 123–155 (2020).
[14]J. Seo, S. K. Kim, A. Jalalvand, R. Conlin, A. Rothstein, J. Abbate, K. Erickson, J. Wai, R. Shousha, and E. Kolemen, Nature **626**, 746–751 (2024).
[15]S. Chouchene, F. Brochard, M. Desecures, N. Lemoine, and J. Cavalier, Sci. Rep. **14**, 27965 (2024).
[16]V. Mehta, J. Barr, J. Abbate, M. D. Boyer, I. Char, W. Neiswanger, E. Kolemen, and J. Schneider, Nucl. Fusion **64**, 046014 (2024).
[17]A. Jalalvand, M. Curie, S. K. Kim, J. Seo, P. Steiner, Q. Hu, A. Nelson, and E. Kolemen, Bull Amer. Phys. Soc (2024).
[18]C. Rea, K. J. Montes, K. G. Erickson, R. S. Granetz, and R. A. Tinguely, Nucl. Fusion **59**, 096016 (2019).
[19]K. Rath, D. Rügamer, B. Bischl, U. von Toussaint, C. Rea, A. Maris, R. Granetz, and C. G. Albert, J. Plasma Phys. **88**, 895880502 (2022).
[20]O. Meneghini, S. P. Smith, P. B. Snyder, G. M. Staebler, J. Candy, E. Belli, L. Lao, M. Kostuk, T. Luce, T. Luda, *et al.*, Nucl. Fusion **57**, 086034 (2017).
[21]O. Meneghini, G. Snoep, B. C. Lyons, J. McClenaghan, C. S. Imai, B. Grierson, S. P. Smith, G. M. Staebler, P. B. Snyder, J. Candy, *et al.*, Nucl. Fusion **61**, 026006 (2020).
[22]K. L. van de Plassche, J. Citrin, C. Bourdelle, Y. Camenen, F. J. Casson, V. I. Dagnelie, F. Felici, A. Ho, S. Van Mulders, and JET Contributors , Phys. Plasmas **27** (2020).
[23]P. Seleson, M. Mustafa, D. Curreli, C. D. Hauck, M. Stoyanov, and D. E. Bernholdt, Comput. Phys. Commun. **279**, 108436 (2022).
[24]R. Shousha, J. Seo, K. Erickson, Z. Xing, S. Kim, J. Abbate, and E. Kolemen, Nucl. Fusion **64**, 026006 (2023).
[25]M. D. Boyer, S. Kaye, and K. Erickson, Nucl. Fusion **59**, 056008 (2019).
[26]G. M. Wallace, Z. Bai, R. Sadre, T. Perciano, N. Bertelli, S. Shiraiwa, E. W. Bethel, and J. C. Wright, J. Plasma Phys. **88**, 895880401 (2022).
[27]Á. Sánchez-Villar, Z. Bai, N. Bertelli, E. W. Bethel, J. Hillairet, T. Perciano, S. Shiraiwa, G. M. Wallace, and J. C. Wright, Nucl. Fusion **64**, 096039 (2024).
[28]A. P. Smirnov, R. W. Harvey, and K. Kupfer, Bull Amer. Phys. Soc. **39**, 1626 (1994).
[29]R. Harvey and M. McCoy, in *Proceedings of the IAEA Technical Committee Meeting on Simulation and Modeling of Thermonuclear Plasmas* (1992) pp. 489–526.
[30]M. Ono, Phys. Plasmas **2**, 4075–4082 (1995).
[31]J. E. Menard, J. P. Allain, D. J. Battaglia, F. Bedoya, R. E. Bell, E. Belova, J. W. Berkery, M. D. Boyer, N. Crocker, A. Diallo, *et al.*, Nucl. Fusion **57**, 102006 (2017).
[32]T. H. Stix, Nucl. Fusion **15**, 737 (1975).

[33]J. Bucalossi, J. Achard, O. Agullo, T. Alarcon, L. Allegretti, H. Ancher, G. Antar, S. Antusch, V. Anzallo, C. Arnas, *et al.*, Nucl. Fusion **62**, 042007 (2022).
[34]L. Breiman, Machine learning **45**, 5–32 (2001).
[35]P. Geurts, D. Ernst, and L. Wehenkel, Machine learning **63**, 3–42 (2006).
[36]D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature **323**, 533–536 (1986).
[37]C. M. Bishop, *Neural networks for pattern recognition* (Oxford university press, 1995).
[38]L. A. Berry, E. F. Jaeger, C. K. Phillips, C. H. Lau, N. Bertelli, and D. L. Green, Phys. Plasmas **23** (2016).
[39]T. H. Stix, *Waves in plasmas* (Springer Science & Business Media, 1992).
[40]M. Brambilla, *Kinetic theory of plasma waves: homogeneous plasmas*, 96 (Oxford University Press, 1998).
[41]M. Brambilla, Plasma Phys. Control. Fusion **31**, 723 (1989).
[42]R. J. Hawryluk, "An empirical approach to tokamak transport," *Phys. Plasmas Close Thermonucl. Cond.*, , 19–46 (1981).
[43]P. T. Bonoli, M. Brambilla, E. Nelson-Melby, C. K. Phillips, M. Porkolab, G. Schilling, G. Taylor, S. J. Wukitch, R. L. Boivin, C. J. Boswell, *et al.*, Phys. Plasmas **7**, 1886–1893 (2000).
[44]M. Brambilla and R. Bilato, Plasma Phys. Control. Fusion **66**, 075009 (2024).
[45]G. Taylor, J. C. Hosea, C. E. Kessel, B. P. LeBlanc, D. Mueller, C. K. Phillips, E. J. Valeo, J. R. Wilson, P. M. Ryan, P. T. Bonoli, *et al.*, Phys. Plasmas **19** (2012).
[46]C. Williams and C. Rasmussen, Adv. Neural Inf. Process. Syst. **8** (1995).
[47]A. G. d. G. Matthews, M. Van Der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. Le, Z. Ghahramani, J. Hensman, *et al.*, Journal of Machine Learning Research **18**, 1–6 (2017).
[48]K. P. Murphy, *Machine learning: a probabilistic perspective* (MIT press, 2012).
[49]C. K. I. Williams and C. E. Rasmussen, *Gaussian Processes For Machine Learning*, Vol. 2 (MIT press Cambridge, MA, 2006).
[50]P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "Scipy 1.0: fundamental algorithms for scientific computing in python," Nature Methods **17**, 261–272 (2020).
[51]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research **12**, 2825–2830 (2011).