Application of Gaussian Process Regression to the Current-Voltage Characteristics of a Langmuir Probe

Yuichi KAWACHI*, Atsushi OKAMOTO, Koki SATO,
Yuto YAMADA, Kota TAKEDA, Takaaki FUJITA

Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

(Received 6 March 2025 / Accepted 6 April 2025)

We evaluate the performance of Gaussian Process Regression (GPR) in estimating the current-voltage characteristics of a Langmuir probe, as well as its first and second derivatives. The results show good agreement between the estimated and measured data. When comparing GPR with the conventional Savitzky-Golay filter, we find that GPR is comparable to the Savitzky-Golay filter in terms of the accuracy of the estimated data. The uncertainty of the estimated data is also evaluated, and the results indicate that GPR underestimates the uncertainty of the electron current. This is likely due to the assumption of a homoscedastic noise model in the standard GPR.

Keywords: Langmuir probe, Gaussian process regression, uncertainty quantification, electron energy distribution function

DOI: 10.1585/pfr.20.1201030

Phase space dynamics is crucial in various fields of plasma physics, such as space and magnetospheric plasma and fusion research [1–3]. To elucidate the understanding of the phase space dynamics, energy distribution measurements of electrons and ions are essential. The Langmuir probe and energy analyzer are widely used diagnostic tools for measuring the electron energy distribution function (EEDF) in basic plasma experiments and satellite measurements. The EEDF is evaluated by taking the first derivative of the current-voltage characteristics obtained from the energy analyzer or the second derivative from the Langmuir probe [4].

However, evaluating the EEDF is often challenging because derivative measurements are highly sensitive to noise. Several methods have been proposed to estimate the EEDF from current-voltage characteristics, such as analog differentiation, Fourier decomposition with an additional superimposed sinusoidal wave, and numerical differentiation after applying a digital filter, including the Savitzky-Golay filter. The effectiveness of each method has been discussed [4, 5]. However, little attention has been given to evaluating the uncertainty of the differentiated values. In this study, we propose a method to estimate the first and second derivatives of the current-voltage characteristics of a Langmuir probe using Gaussian Process Regression (GPR) [6]. GPR is a Bayesian, nonparametric regression method that considers the correlation between data points and has the potential to evaluate the uncertainty of the estimated values and its derivatives [7]. We evaluate the first and second derivatives of the current-voltage characteristics using GPR with experimental data and compare the results with those obtained using the Savitzky-Golay filter. Additionally, we assess the uncertainty of the estimated values using GPR and discuss its effectiveness in estimating the EEDF.

Test data for this study were obtained from Langmuir probe measurements in a magnetized plasma device, NUMBER. NUMBER is a pulsed one-sided mirror device that produces ECH plasma using a 2.45 GHz microwave [8]. The Langmuir probe was inserted into the plasma, and the current-voltage characteristics were measured in argon plasma with a mass flow rate of 10 sccm. The voltage and current dataset was obtained by downsampling to 200 points from a single voltage sweep, as shown in Fig. 1.

The GPR model was trained using the voltage and current dataset and implemented using the Python library GPflow [9]. The kernel function was set to the Radial Basis Function (RBF) kernel defined as,

$$k(x, x') = \theta_1 \exp\left(-\frac{(x-x')^2}{2\theta_2^2}\right) + \sigma^2\delta(x, x'). \tag{1}$$

Here, x and x' represent observed data points. θ_1 is a parameter of the RBF kernel that controls the amplitude of the estimated function, while θ_2 determines the covariance scale of the data, some what similar to the window size in the Savitzky-Golay filter. σ^2 is the noise variance, which accounts for uncertainty. These parameters were optimized using the maximum likelihood method.

Figure 1(a) shows the estimated current-voltage characteristics obtained using GPR. For comparison, the Savitzky-Golay filter, a conventional and reliable method capable of analytically evaluating the n-th derivative, was also applied.

Corresponding author's e-mail: y-kawachi@energy.nagoya-u.ac.jp

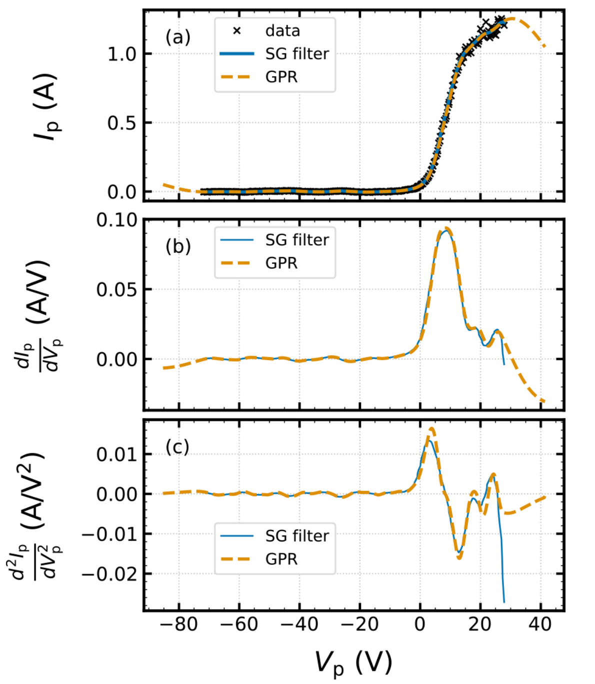

Fig. 1. (a) I-V characteristics of the data obtained from the Langmuir probe. X marks indicate the measured data, while the solid blue line and the dashed orange line represent the estimated results obtained using the Savitzky-Golay method and Gaussian Process Regression, respectively. The (b) first and (c) second derivatives were also estimated by both methods and are indicated in the same manner.

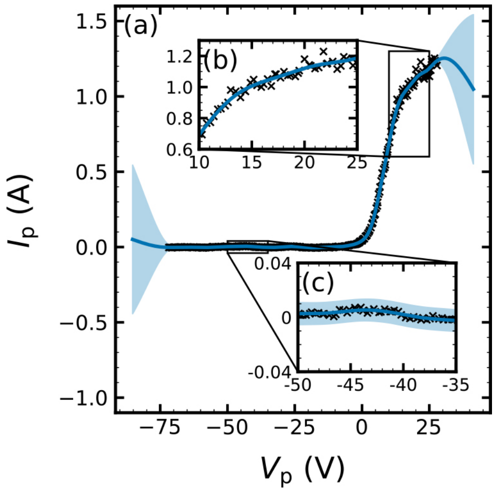

Fig. 2. (a) Evaluated uncertainty of the estimated data using GPR. The solid blue line represents the estimated data, while the shaded area indicates the uncertainty. The measured data is also indicated by X marks. Zoomed-in views of the regions of (b) electron current and (c) ion current are shown.

The parameter settings for the Savitzky-Golay filter were set to a polynomial order of 3 and a window size of 40. It can be seen that GPR accurately estimates the I-V characteristics. Furthermore, when compared to the Savitzky-Golay filter, GPR demonstrates comparable performance. This trend is also observed for the first- and second-order derivatives, where the estimation results exhibit nearly the same shape. Focusing on the second-order derivative evaluation, the two methods exhibit a slight difference, with the peaks being relatively shifted by approximately 1 V. When discussing the EEDF, the tail and integral values are important, so the extent to which this shift affects the analysis needs to be considered carefully. The parameters of Savitzky-Golay filter is arbitary chosen by users while the kernel function of GPR is optimized by the maximum likelihood method. So the GPR might be closer to the true solution than the Savitzky-Golay filter. It is difficult to make a definitive judgment on the factors determining this, so it would be desirable to compare it with an analog filter or mock-up data, or to increase the ensemble, which is left for future work.

Here, we evaluate the uncertainty of the estimated data using GPR, as shown in Fig. 2. The uncertainty is defined as the standard deviation of the posterior distribution of the estimated data. We note that regions where data is lacking exhibit large uncertainty, which is consistent with the characteristics of GPR. Figure 2(b) shows an enlarged view of the region where the ion current is collected when the bias voltage is sufficiently negative to repel electrons. We confirm that the uncertainty is in good agreement with the measured data. However, the uncertainty of the electron current, where the bias voltage is sufficiently positive to repel ions, is underestimated, as shown in Fig. 2(c). This is likely due to the assumption of a homoscedastic noise model in standard GPR. In the case of the Langmuir probe, the electron current is much larger than the ion current, and the noise level is also higher. Therefore, the uncertainty of the electron current should be larger than that of the ion current. However, standard GPR assumes a uniform noise level across all data points, which may lead to an underestimation of the electron current uncertainty. To address this issue, we need to extend the method to estimate the uncertainty of the electron current by considering a heteroscedastic noise model [10]. This approach will be discussed in future work.

In summary, we evaluated the performance of GPR in estimating the current-voltage characteristics of a Langmuir probe, as well as its first and second derivatives. The results show good agreement between the estimated and measured data. When comparing GPR with the conventional Savitzky-Golay filter, we find that GPR performs comparably to the Savitzky-Golay filter in terms of estimation accuracy. We also evaluated the uncertainty of the estimated data, and the results indicate that GPR underestimates the uncertainty of the electron current. This is likely due to the assumption of a homoscedastic noise model in standard GPR. Future work will focus on extending the method to improve uncertainty estimation for the electron current by incorporating a heteroscedastic noise model.

This work was supported in part by JSPS KAKENHI under Grant Nos. JP23K13082, and JP23K25845, NIFS collaboration research program under Grant No. NIFS24KIPS008, the Nitto Foundation, and the TOYOAKI SCHOLARSHIP FOUNDATION. We also thank fruitful discussions with F. Kin, T. Kobayashi, and T. Nishizawa.

[1] Y. Kasahara *et al.*, Earth Planets Space **70**, 1 (2018).
[2] T. Kobayashi *et al.*, Proc. Natl. Acad. Sci. **121**, e2408112121 (2024).
[3] Y. Kawachi *et al.*, Rev. Sci. Instrum. **94**, 103505 (2023).
[4] A. Caldallel *et al.*, Phys. Plasmas **30**, 040501 (2023).
[5] F. Magnus and J.T. Gudmundsson, Rev. Sci. Instrum. **79**, 073503 (2008).
[6] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, 2006).
[7] E. Solak *et al.*, Adv. Neural Inf. Process. Syst. **15**, 1209 (2002).
[8] D. Hamada *et al.*, Plasma Fusion Res. **13**, 3401004 (2018).
[9] A.G. de G. Matthews *et al.*, J. Mach. Learn. Res. **18**, 1299 (2017).
[10] S. Krishnan *et al.*, *Proceedings of the 24th international conference on Machine learning* (Association for Computing Machinery, New York, 2007) 393–400.