Bayesian optimization of disruption scenarios with fluid-kinetic models

Citation for the original published paper (version of record):
Ekmark, I., Pusztai, I., Hoppe, M. et al (2023). Bayesian optimization of disruption scenarios with fluid-kinetic models. 49th EPS Conference on Plasma Physics, EPS 2023

Bayesian optimization of disruption scenarios with fluid-kinetic models

I. Ekmark[1], I Pusztai[1], M Hoppe[2], P Jansson[3], T Fülöp[1]

[1] *Department of Physics, Chalmers Univ. Technology, Göteborg, Sweden*

[2] *Department of Electrical Engineering, KTH Royal Inst. Technology, Stockholm, Sweden*

[3] *Department of Computer Science and Engineering, Chalmers Univ. Technology, Göteborg, Sweden*

Tokamak disruptions can damage the device due to localized heat loads, mechanical stresses and impact of energetic runaway electron (RE) beams. Using Bayesian optimization and the simulation tool DREAM [1], we study massive material injection (MMI) to find what densities of deuterium and neon result in successful disruption mitigation. Besides a more systematic construction of the cost function, we go beyond the proof-of-concept study of [2] by utilizing both fluid and kinetic plasma models. The fluid model allows the exploration of a large parameter space, and promising regions are studied in higher physics fidelity using kinetic simulations.

Simulation setup The disruption simulations are performed in an ITER-like tokamak setup with a deuterium plasma of density $n = 10^{20}\,\mathrm{m}^{-3}$ and a magnetic field of $B_0 = 5.3\,\mathrm{T}$ on axis. The major radius $R_0 = 6\,\mathrm{m}$, minor radius $a = 2\,\mathrm{m}$ and the radius of the first toroidally closed conducting structure $b = 2.833\,\mathrm{m}$. At $t = 0$, the plasma current $I_\mathrm{p} = 15\,\mathrm{MA}$, and plasma temperature is parabolic with $T(r = 0) = 20\,\mathrm{keV}$. The deuterium and neon of the MMI are assumed to consist of cold atom populations, deposited instantaneously and homogeneously. During the thermal quench (TQ), spatially and temporally constant magnetic perturbations of $\delta B/B = 0.5\,\%$ lead to a transport of REs and heat according to the Rechester–Rosenbluth model [3], with a diffusion coefficient $D = \pi R_0(\delta B/B)^2$.

DREAM divides the electrons in up to three populations. In our simulations, the bulk and RE populations are evolved as fluids. The bulk electrons are characterized by their density n_e, temperature T_e and current density j_Ω, while the REs are described by their density n_re, which also yields the RE current density $j_\mathrm{re} = ecn_\mathrm{re}$. Additionally, in our kinetic simulations the superthermal electrons are represented by the distribution function f_hot, which is evolved using the pitch-angle averaged Fokker-Planck equation [1]. Remaining plasma parameters are evolved as fluids. In the kinetic model, the Dreicer and hot-tail generation mechanisms are naturally present as velocity space particle fluxes, while in the fluid model, they are both implemented as generation rates. The Dreicer generation rate is computed by a neural network [4], and the hot-tail generation rate by the model in [5]. In both models, the avalanche generation is determined by the model in [6]. Since the plasma fuel consists only of deuterium, the activated generation sources from tritium decay and Compton scattering are not present.

In order to perform the optimizations, the impact of the disruption is quantified by a representative maximum RE current $I_{\rm re}^{\rm max}$, the final ohmic current $I_{\Omega}^{\rm final}$, conducted heat load $\eta_{\rm cond}$ and current quench (CQ) time $\tau_{\rm CQ}$. The representative maximum RE current is the maximum RE current $\max_t I_{\rm re}$, if it happens before $I_{\rm re}(t) = 0.95 I_{\rm p}(t)$, otherwise the latter is used. Ideally for ITER, $I_{\rm re} < 150\,{\rm kA}$ [7], which we have used as the upper limit of safe operation for both $I_{\rm re}^{\rm max}$ and $I_{\Omega}^{\rm final}$. The conducted heat load and CQ time are defined as in [2]. For safe operation in ITER, $\eta_{\rm cond} < 10\,\%$ and $\tau_{\rm CQ} \in [50,\ 150]\,{\rm ms}$ [8].

The cost function $\mathscr{L}(I_{\rm re}^{\rm max}, I_{\Omega}^{\rm final}, \eta_{\rm cond}, \tau_{\rm CQ})$ used in the optimization was designed to be differentiable and yield a value below one if all components are within their intervals of safe operation. During safe operation, minimizing $I_{\rm re}^{\rm max}$ and $I_{\Omega}^{\rm final}$ was deemed more important than minimizing $\eta_{\rm cond}$, which, in turn, is more important than minimizing $|\tau_{\rm CQ} - 100\,{\rm ms}|$. To achieve this order of importance, the function $f_i = x_i^{k_i}$ if $x_i \leq 1$, and $f_i = k_i x_i + 1 - k_i$ otherwise, was used; for the currents $k_I = 1$ and $x_I = I/150\,{\rm kA}$, for the conducted heat load $k_{\eta_{\rm cond}} = 3$ and $x_{\eta_{\rm cond}} = \eta_{\rm cond}/10\,\%$, and for the CQ time $k_{\tau_{\rm CQ}} = 6$ and $x_{\tau_{\rm CQ}} = |\tau_{\rm CQ} - 100\,{\rm ms}|/50\,{\rm ms}$. Finally, these function values where combined using the Euclidean norm, to obtain the cost function $\mathscr{L}(I_{\rm re}^{\rm max}, I_{\Omega}^{\rm final}, \eta_{\rm cond}, \tau_{\rm CQ}) = \sqrt{(cf_{I_{\rm re}^{\rm max}})^2 + (cf_{I_{\Omega}^{\rm final}})^2 + (cf_{\eta_{\rm cond}})^2 + (cf_{\tau_{\rm CQ}})^2}$, where the weights $c = 0.5$ ensure values below one for scenarios of safe operation.

The Bayesian optimization was performed using the Python package [9] which utilizes Gaussian processes (GPs). The Matérn kernel was used for the GPs and the estimated improvement acquisition function was used as the optimization strategy. The optimization was performed on the MMI densities within $\log\big(n_{\rm D}/[1\,{\rm m}^{-3}]\big) \in [20, 22.5]$ and $\log\big(n_{\rm Ne}/[1\,{\rm m}^{-3}]\big) \in [15, 20]$ for the fluid model, and within $\log\big(n_{\rm D}/[1\,{\rm m}^{-3}]\big) \in [21.5, 22.25]$ and $\log\big(n_{\rm Ne}/[1\,{\rm m}^{-3}]\big) \in [18.5, 19.75]$ for the kinetic model. When using the fluid (kinetic) model, the optimization was performed using 500 (200) disruption simulations.

Results As shown in 1a, regions of safe operation are found for both the fluid and kinetic models, but the kinetic model predicts a much larger safe region. The optimum found using the fluid model is located at $n_{\rm D} = 1.2 \times 10^{22}\,{\rm m}^{-3}$ and $n_{\rm Ne} = 5.9 \times 10^{18}\,{\rm m}^{-3}$ with a value of $\mathscr{L} = 0.4$, while for the kinetic model, the optimum located at $n_{\rm D} = 8.8 \times 10^{21}\,{\rm m}^{-3}$ and $n_{\rm Ne} = 2.4 \times 10^{19}\,{\rm m}^{-3}$ with a value of $\mathscr{L} = 0.1$. In the vicinity of the optimum found using the fluid model both models exhibit similar behaviour with the kinetic model yielding $\mathscr{L} = 0.5$. However, close to the optimum found using the kinetic model, the models yield significantly different results, with the fluid model yielding $\mathscr{L} = 9$.

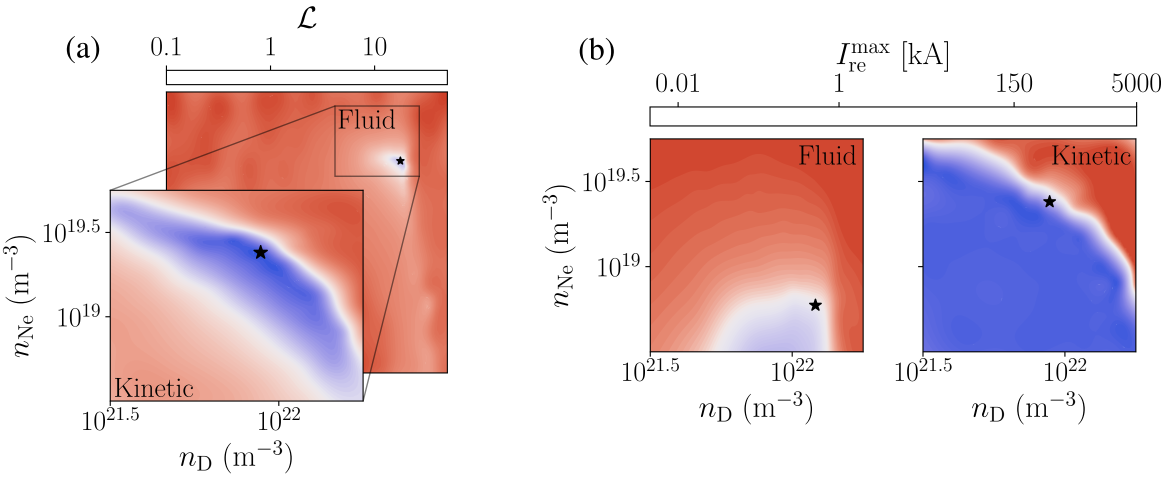

Figure 1: Logarithmic contour plots of the (a) cost function, and (b) representative maximum RE current. Note that the colour mapping is adapted such that blue shades represent regions of safe operation. The black stars indicate the optimal samples found in each optimization.

In order to analyze the differences between the models further, the cost function components were compared between the two models within the bounds used for the optimization using the kinetic model. The maximum RE current exhibits the most significant differences out of the four cost function components; the maximum RE currents are shown for the two models in figure 1b. For the fluid model, $I_{\mathrm{re}}^{\mathrm{max}} < 150\,\mathrm{kA}$ only in a small region, and here it is of the order $100\,\mathrm{kA}$. On the other hand, the kinetic model predicts an insignificant current for a majority of the region. This makes it possible for a better trade-off between $I_{\mathrm{re}}^{\mathrm{max}}$ and the other plasma parameters, resulting in the different location and the lower value of the optimum found using the kinetic model. At the optimum found using the fluid model, the fluid model yields $I_{\mathrm{re}}^{\mathrm{max}} = 83\,\mathrm{kA}$ and $\eta_{\mathrm{cond}} = 8.6\,\%$, while at the optimum found using the kinetic model, the kinetic model yields $I_{\mathrm{re}}^{\mathrm{max}} = 3.5\,\mathrm{kA}$ and $\eta_{\mathrm{cond}} = 5.8\,\%$.

To examine why the RE current is so large when using the fluid model, fluid simulations of the optimum found using the kinetic model were performed with the Dreicer and hot-tail generation rates disabled (separately). Only disabling the hot-tail generation rate resulted in a significant difference, with $I_{\mathrm{re}}^{\mathrm{max}} = 160\,\mathrm{kA} \ll 2.7\,\mathrm{MA}$, which is what the fluid model with hot-tail generation yields. This suggests that the hot-tail generation rate is overestimated for this scenario, which accounts for most of the discrepancy in RE currents between the two models. The hot-tail generation rate in [5] is derived based on the assumption of $Z_{\mathrm{eff}} \gg 1$, which is not satisfied in our simulations, which results in bursts of hot-tail generation in the beginning of both the TQ and CQ, not present during the kinetic simulation.

However, as noted above, the RE current is still significant, $I_{\rm re}^{\rm max} = 160\,{\rm kA} \gg 3.5\,{\rm kA}$. For this scenario, the Dreicer generation rate is overestimated as well, even though not to the same degree as the hot-tail generation. The maximum Dreicer generation rate $\gamma_{\rm Dreicer} \sim 10^3\,{\rm m}^{-3}{\rm s}^{-1}$ on axis while the hot-tail generation peaks at $\gamma_{\rm hot\text{-}tail} \sim 10^{11}\,{\rm m}^{-3}{\rm s}^{-1}$. The normalized Dreicer generation rate $\bar{\gamma}_{\rm Dreicer} = \tau_{\rm ee}\gamma/n_{\rm e}$ (where $\tau_{\rm ee}$ is the thermal electron collision time) is determined by a neural network, which was trained on kinetic simulations performed with deuterium densities $n_{\rm D} \in [10^{18},\ 10^{21}]\,{\rm m}^{-3}$ and normalized electric fields $E/E_{\rm D} \in [0.01, 0.13]$. For values of $E/E_{\rm D}$ below the training range, the neural network incorrectly predicts a very low, constant level of Dreicer RE generation $\bar{\gamma}_{\rm Dreicer} \sim 10^{-30}$, as seen in figure 3 of [4]. In our case, $n_{\rm e} \sim n_{\rm D} \sim 10^{22}$, $E/E_{\rm D} \sim 0.0001$ and $\tau_{\rm ee} \sim 10^{-11}\,{\rm s}$, which gives a $\bar{\gamma}_{\rm Dreicer}$ of the same order as predicted by the NN for low $E/E_{\rm D}$. Our results suggest that for $E/E_{\rm D}$ below the training range, the error predicted by the neural network will scale with the particle density in the plasma. Due to the strong avalanching, even such small errors in the Dreicer generation are amplified to a substantial final RE current. In fact, we find that only a small proportion of the REs in the original simulations are generated by the hot-tail and Dreicer generation rates – the large majority of the REs are generated in the avalanche process. The avalanche mechanism starts to dominate the RE generation already when $I_{\rm re} \sim 10^{-3}\,{\rm A}$, which indicates its sensitivity to small errors in the other generation mechanisms.

Conclusions A region of safe operation was discovered in the MMI injected density space of deuterium and neon. The optima found using both the fluid and kinetic plasma models lie close to a deuterium density of $n_{\rm D} \sim 10^{22}\,{\rm m}^{-3}$ and a neon density of $n_{\rm Ne} \sim 10^{19}\,{\rm m}^{-3}$, but the kinetic model favoured lower deuterium and higher neon densities. In general, the kinetic model was more optimistic than the fluid model regarding the success of the disruption mitigation. This is primarily due to the significantly larger RE current predicted with the fluid model, caused by overestimated hot-tail and Dreicer RE rates. Already small absolute errors in the hot-tail and Dreicer generation rates result in significant magnification of the RE seed due the avalanche mechanism.

[1] M. Hoppe et al., CPC, **268**, 108098 (2020).
[2] I. Pusztai et al., JPP, **89**, 905890204 (2023).
[3] A. B. Rechester and M. N. Rosenbluth, PRL, **40**, 38 (1978).
[4] L. Hesslow et al., JPP, **85**, 475850601 (2019).
[5] I. Svenningsson, Master thesis (2020).
[6] L. Hesslow et al., NF, **59**, 084004 (2019).
[7] M. Lehnen et al., Presentation, PPPL (2021).
[8] E. M. Hollmann et al., PP, **22**, 021802 (2015).
[9] F. Nogueira, Github (2014).

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them