# Extrapolative predictability of plasma turbulent transport via a multi-fidelity data fusion approach

ID: 2838

S. Maeyama[1], M. Honda[2], E. Narita[2], F. Miwakeichi[3], K. Yano[3], A. Okuno[3]
[1]NIFS, [2]Kyoto Univ., [3]Inst. Stat. Math.
maeyama.shinya@nifs.ac.jp

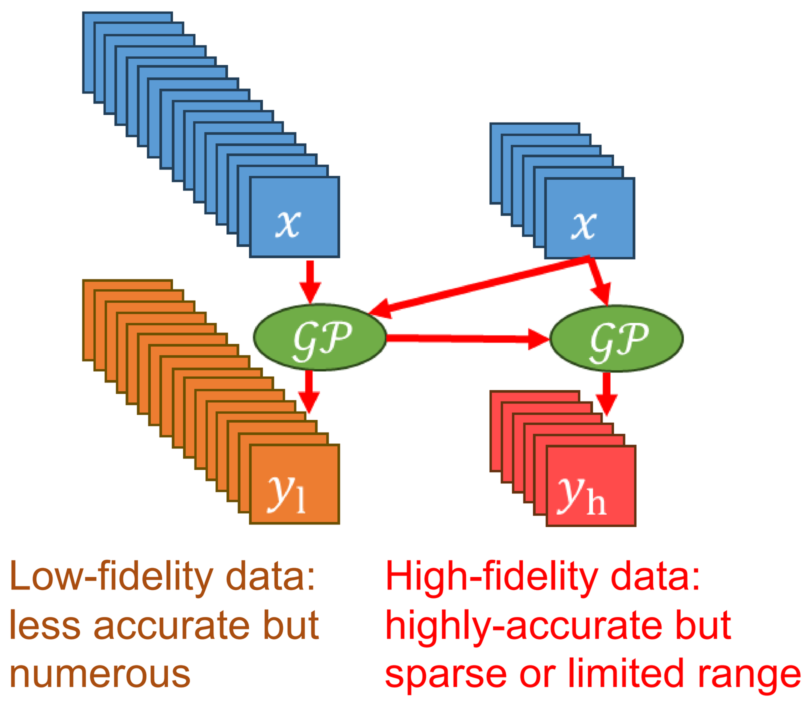
## MOTIVATION

- Modeling turbulent transport plays a crucial role in fusion development. Various data may have different levels of fidelity: experiments, simulations, and reduced models. We aim to improve the predictive accuracy of turbulent transport by multi-fidelity data fusion. [Maeyama (2024)]

## METHODS

**Multi-fidelity regression problem**

Estimate the high-fidelity function $y_h = f_h(x_h)$ by fully utilizing low- and high-fidelity dataset.




**Nonlinear Auto-Regressive Gaussian Process Regression (NARGP)**

- Lowest-fidelity: GP regression

$$f_l(x) \sim \mathcal{GP}\left(f_l | \mu_l, k_l(x, x'; \theta_l)\right)$$

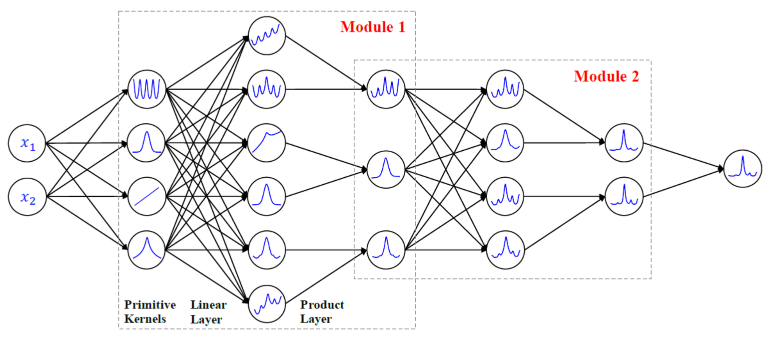
- High-fidelity: GP regression expressed by input $x$ and low-fidelity $f_{*l}(x)$

$$f_h(x) = g_h\left(x, f_{*l}(x)\right),$$
$$g_h \sim \mathcal{GP}\left(f_h | \mu_h, k_h\left((x, f_{*l}(x)), (x', f_{*l}(x'); \theta_h\right)\right)$$

NARGP is a successive composition of GPs, which captures nonlinear and $x$-dependent correlation between multi-fidelity datasets. [Perdikaris (2017)]

**Neural kernel network (NKN)** [Sun (2018)]

NKN is a framework that composes basic (e.g., periodic, radial basis, linear, rational quadratic) kernel functions using a neural network architecture, realizing a highly-expressible kernel for GP regression.



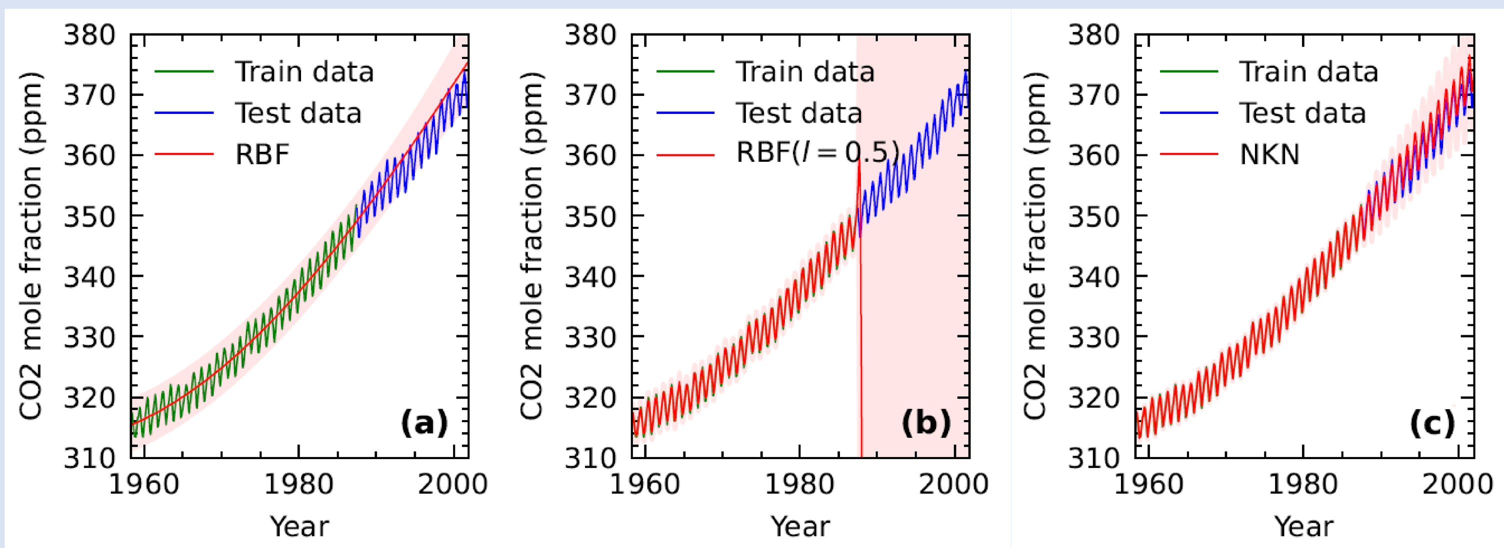




*Fig. 1 Comparison of expressiveness of kernel functions in single GP regression.*

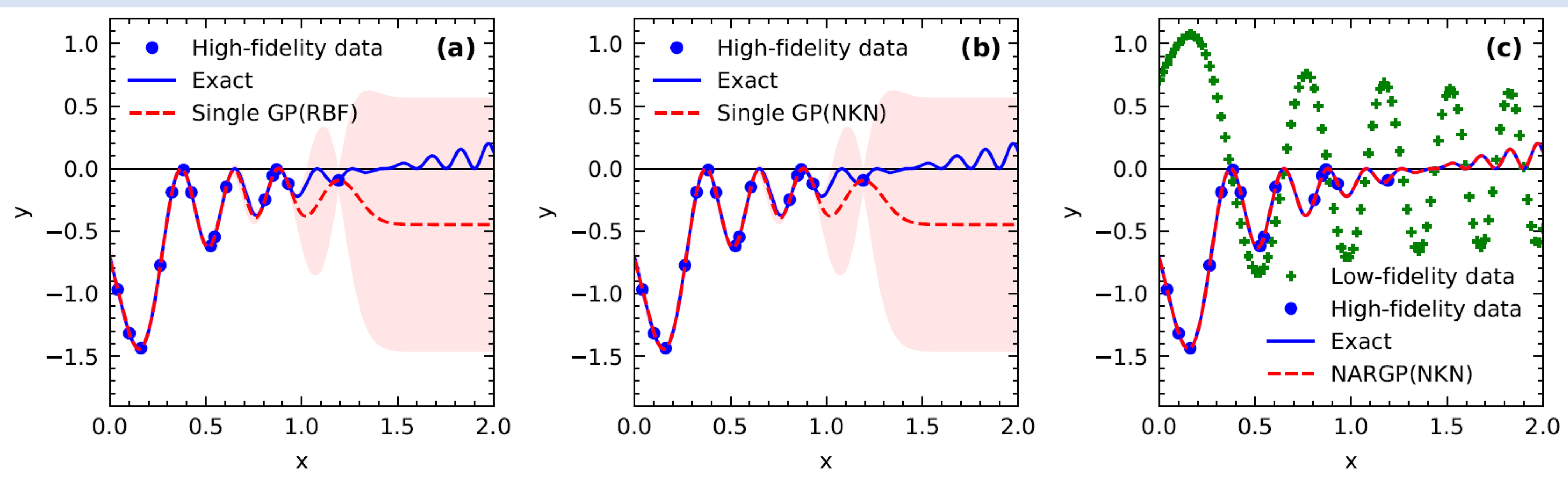



*Fig. 2 Extrapolative prediction by multi-fidelity regression for a 1D test problem.*

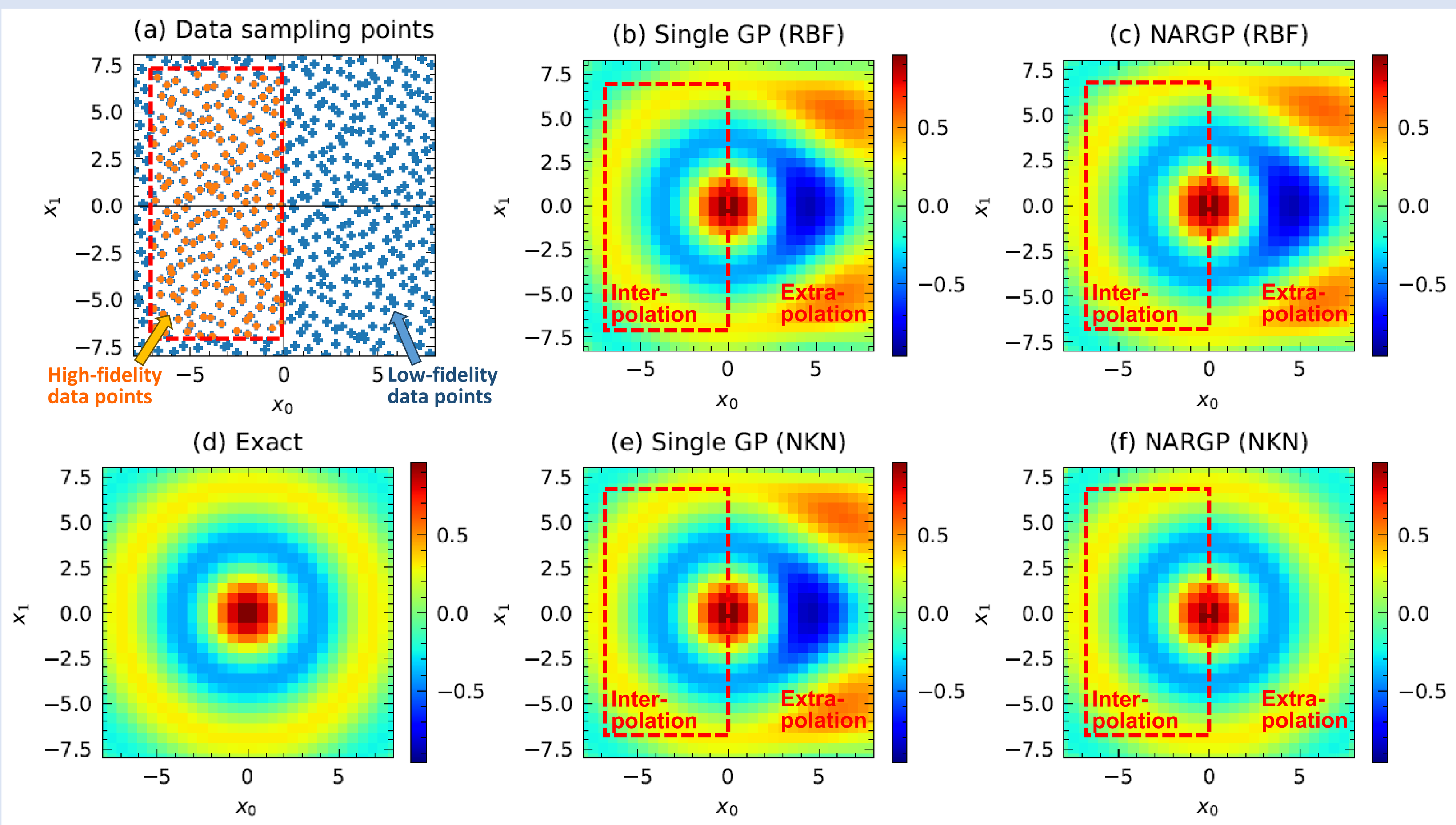



*Fig. 3 Extrapolative prediction by multi-fidelity regression for a 2D test problem.*

## RESULTS

- NKN captures short-scale variation and long-scale trend simultaneously, while a single-scale kernel (RBF) cannot (Fig. 1).
- NARGP learns the correlation between low and high-fidelity data and utilizes it for extrapolative prediction (Figs. 2 and 3). Employing both NKN and NARGP is essential for improvement (Fig. 3).
- We have applied the multi-fidelity regression to the real-world dataset of turbulent diffusion coefficient from JET experiments [Narita (2023)], mimicking extrapolative prediction for (future) better-confinement plasma by using the existing training data (Fig. 4-1). A combination of NKN and NARGP achieved improved extrapolative prediction (Fig. 4-2).

| Input (12 dims.) | Low-fidelity data | High-fidelity data |
|---|---|---|
| $R/L_n$, $R/L_{Te}$, $R/L_{Ti}$, $n_i/n_e$, $T_e/T_i$, β, $\nu_{ee}$, $q$, $s$, $\varepsilon$, $\kappa$, $\delta$ | Linear growthrate $\gamma$ and its wavenumber $k$ (135 points) | Experimental turbulent diffusion coefficient $D_{exp}$ (67 points) |

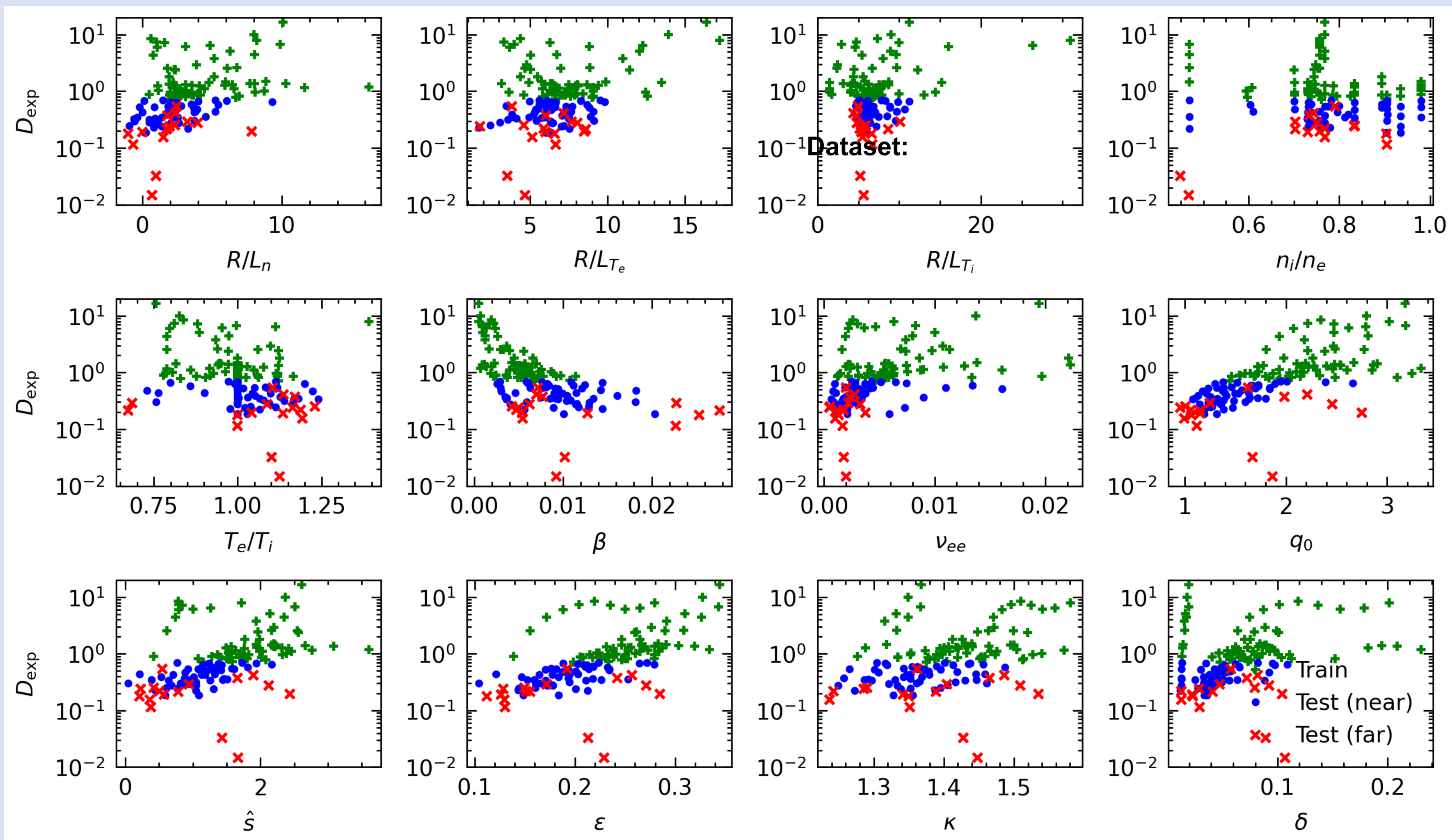



*Fig. 4-1 Extrapolative problem setting for the JET dataset. The training data showing worse confinement ($D_{exp} > 0.7$) is used to predict the test data (showing better confinement, $D_{exp} < 0.7$).*

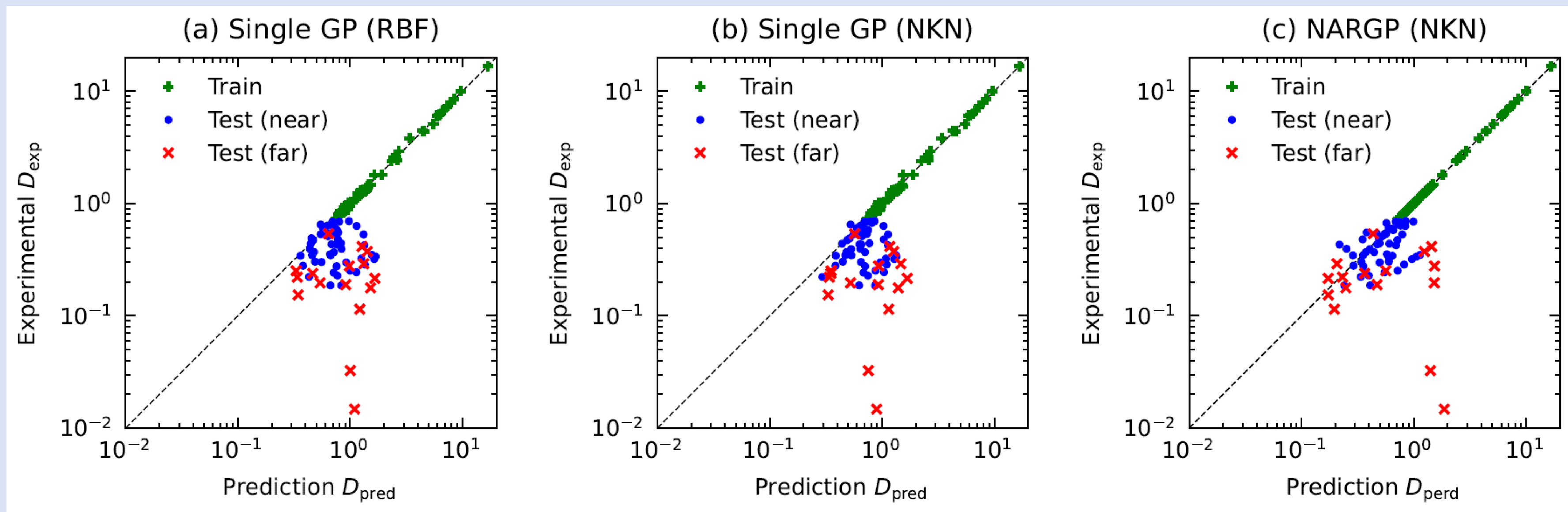



*Fig. 4-2 Extrapolative prediction by multi-fidelity regression for the JET dataset.*

## CONCLUSION

- Multi-fidelity data fusion provides a novel approach that combines the quantitativeness of experimental data with the physics-based extrapolability of theory/simulation.
- By integrating a highly-expressible NKN kernel and a multi-fidelity regression algorithm of NARGP, our model captures complex nonlinear correlations across multi-fidelity data and achieves improved extrapolative prediction.

**Key tricks**

- **A complex kernel** (such as NKN) is crucial for extrapolation by capturing short-time variation and long-scale trend simultaneously.
- **Dependence simplification by NARGP**: Express complexity of high-fidelity function $f_h(x)$ by low-fidelity data $f_l(x)$.
- **Interpolation recasting by NARGP**: Eliminate the dependence on extrapolative parameter, e.g., $f_h(x_0, x_1) = g_h(x_0, x_1, f_l) \simeq g_h(x_1, f_l)$.