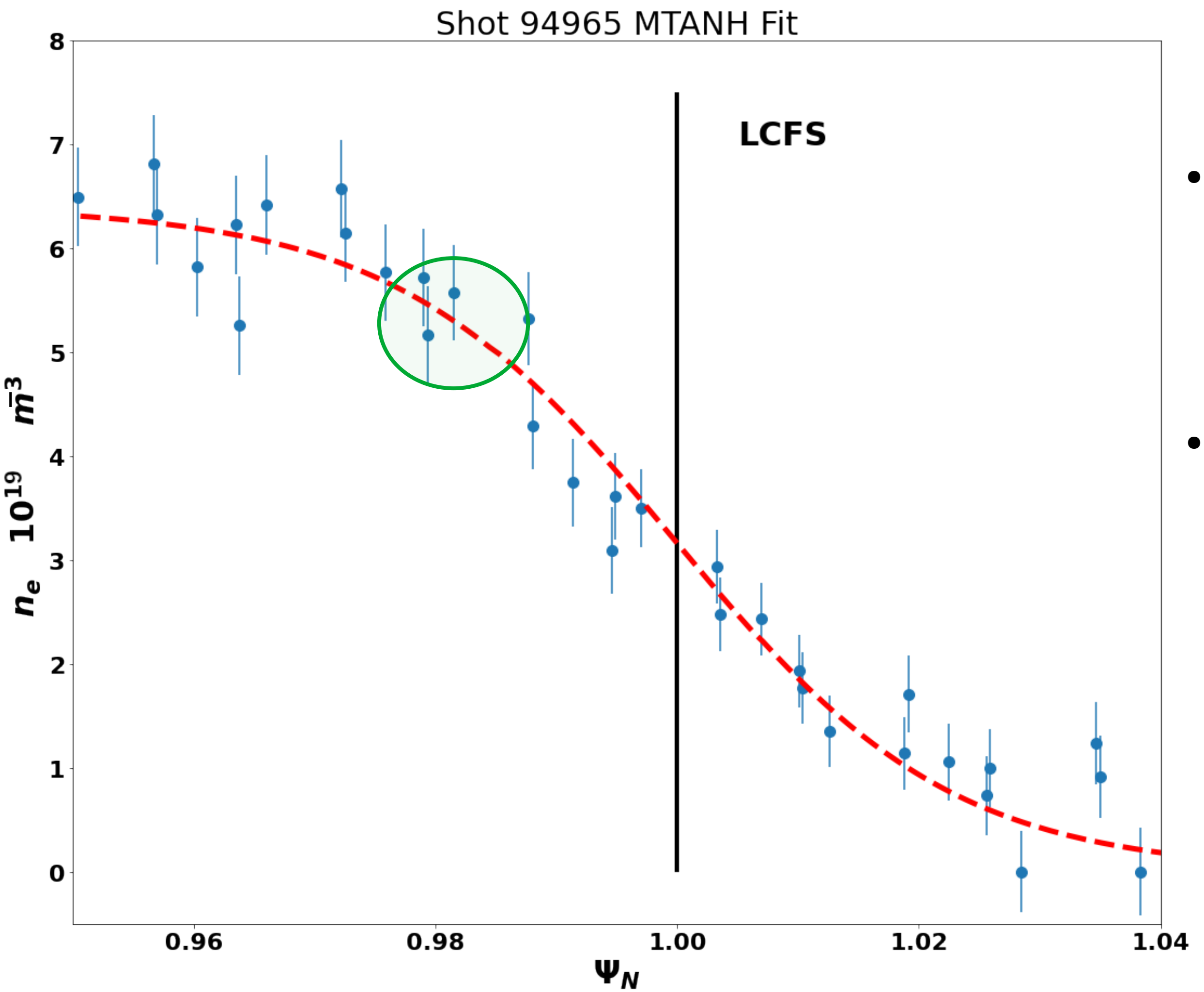
Supervised Machine Learning Analysis on the JET Pedestal Database

A Kit and A Järvinen

and the WP2 team of ENR 08

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

JET pedestal database contains 3000+ entries of H-mode, pre-ELM plasmas



- Pedestal parameters determined from mtanh fit (red) of HRTS measurements (blue)
- Focus on predicting n_e^{ped} height (green)

Current predictive models are limited by their input parameters and simplifying assumptions

- EPED
  - Pressure gradient in pedestal is limited by kinetic balooning modes (KBM) and total pressure by ideal-MHD peeling-balooning modes
  - These are not justified for a large fraction of JET pedestal database L. Frassinetti et al. *Pedestal Structure, stability and scalings in JET-ILW: the EUROfusion JET-ILW pedestal database* (2020)
  - Takes n_e^{ped}, global params β and Z_{eff} as inputs
- EUROPED
  - EPED like pedestal model which uses other models to get around using experimental inputs
  - Neutral penetration model (NPM) or log-linear regression to determine n_e^{ped} S. Saarelma et al. *Self consistent pedestal prediction for JET-ILW in preparation of the DT campaign* (2019)

Use only main engineering parameters to predict neped height

Inputs

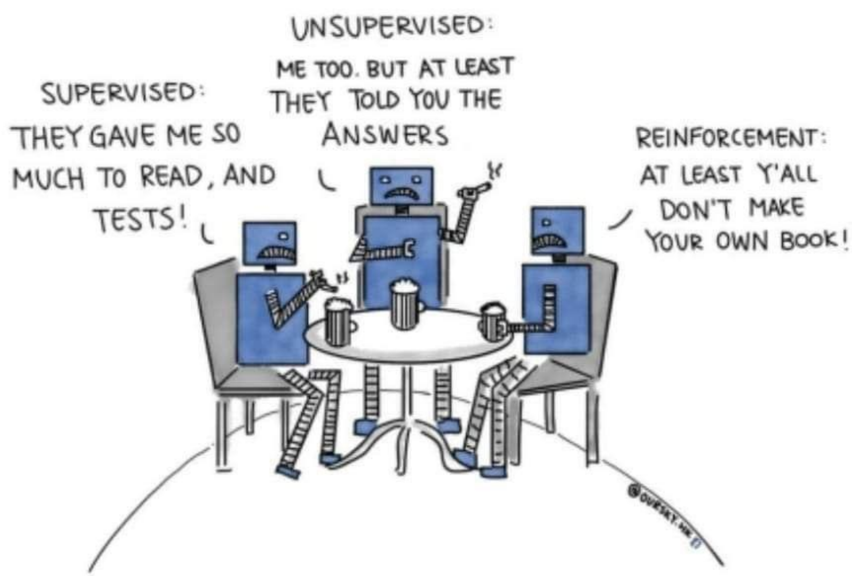
- δ
- a
- I_P
- V_P
- Γ
- B_T
- P_{ICRH}
- P_{TOT}
- P_{NBI}
- $H_{isotope}$
- M_{eff}
- q_{95}

Supervised Learning

Outputs

n_e^{ped} +/- Model uncertainty

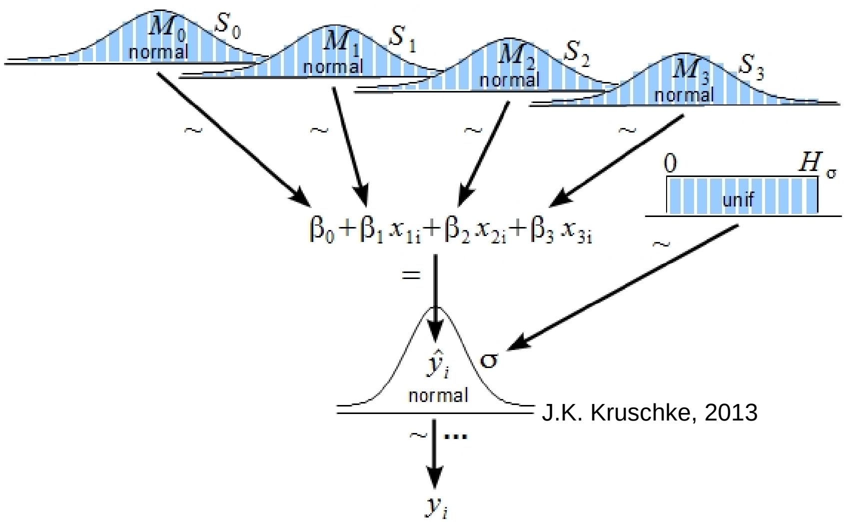
- Excluded from inputs:
  - major radius (not useful for single machine database)
  - Categorical (Divertor config, Wall type)
  - Sub-divertor pressure (not enough values)



https://forum.onefourthlabs.com/t/supervised-vs-unsupervised-vs-reinforment-learning/4269

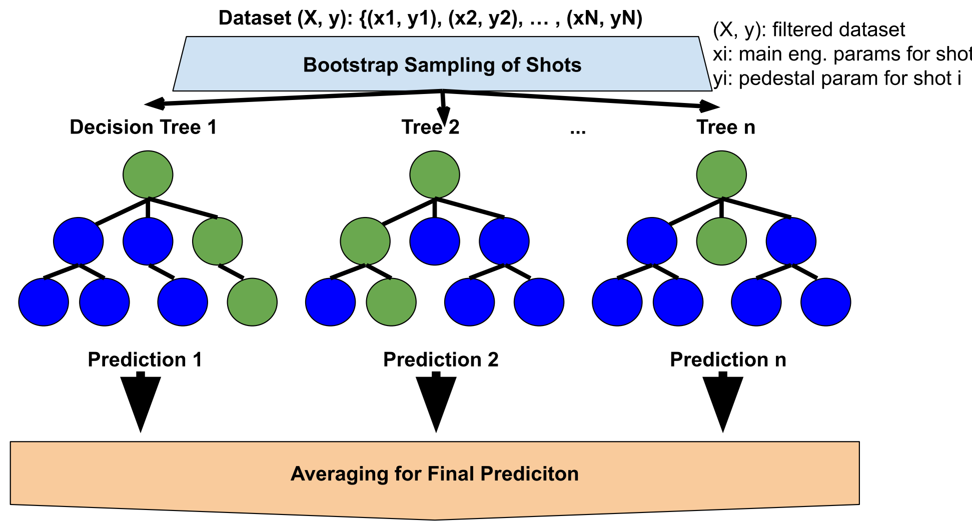
Crash course of supervised machine learning tools utilized to improve emperical predictions of neped

Bayesian Linear Regression

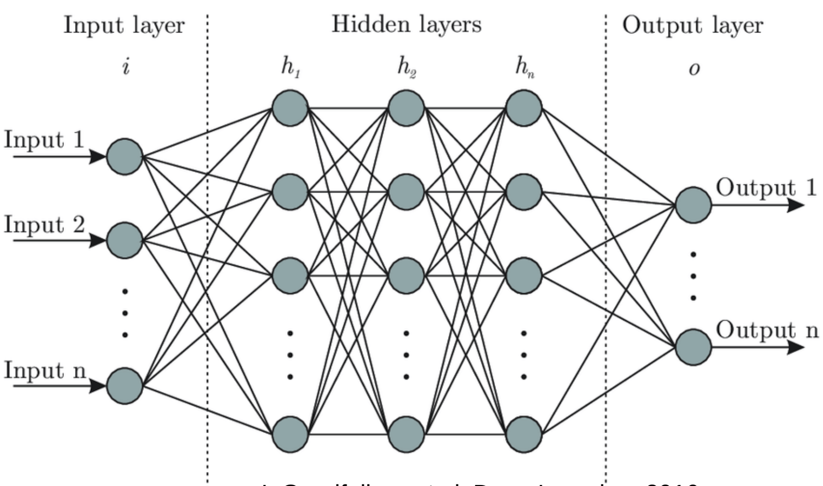


J.K. Kruschke, 2013

Random Forests (RFs) & Extreme Random Trees (ERTs)

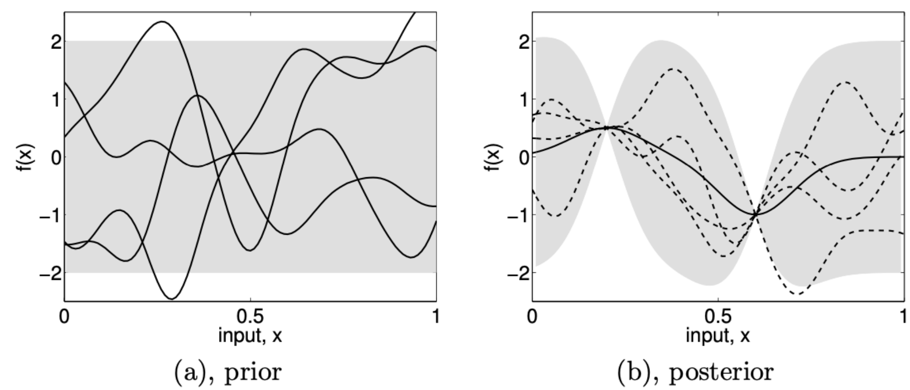


Artificial Neural Networks (ANNs)



I. Goodfellow, et al. Deep Learning. 2016

Gaussian Processes (GPs)

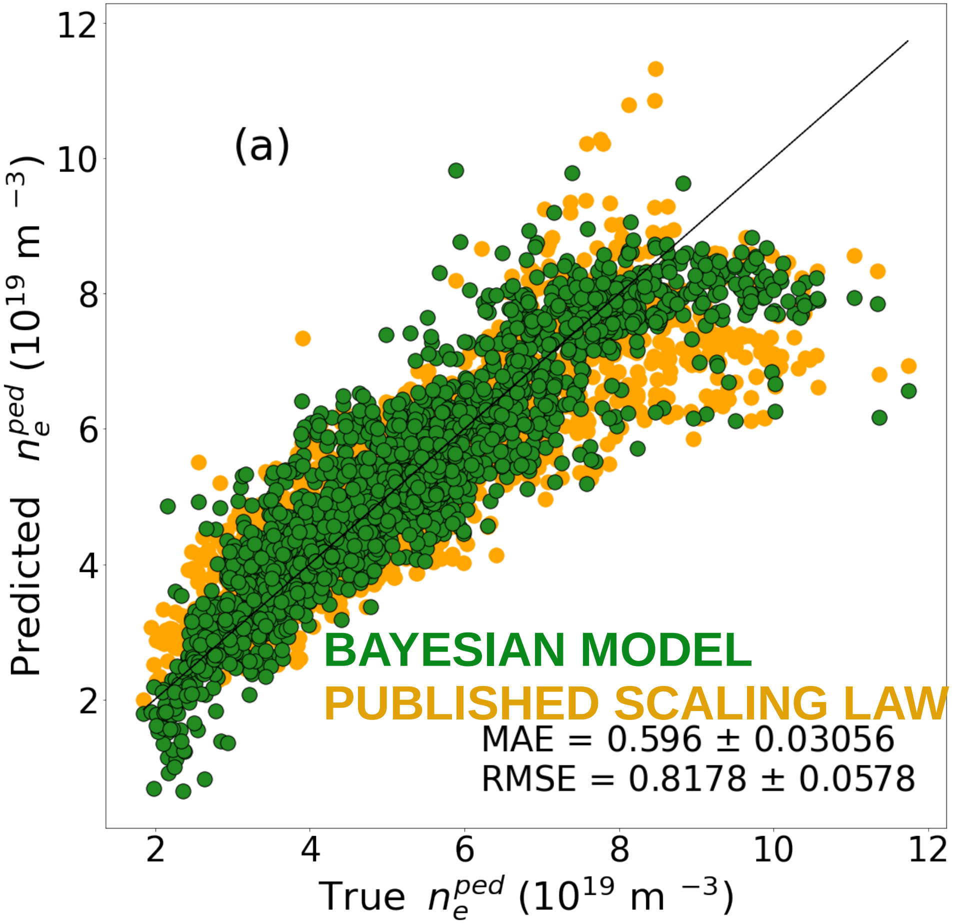


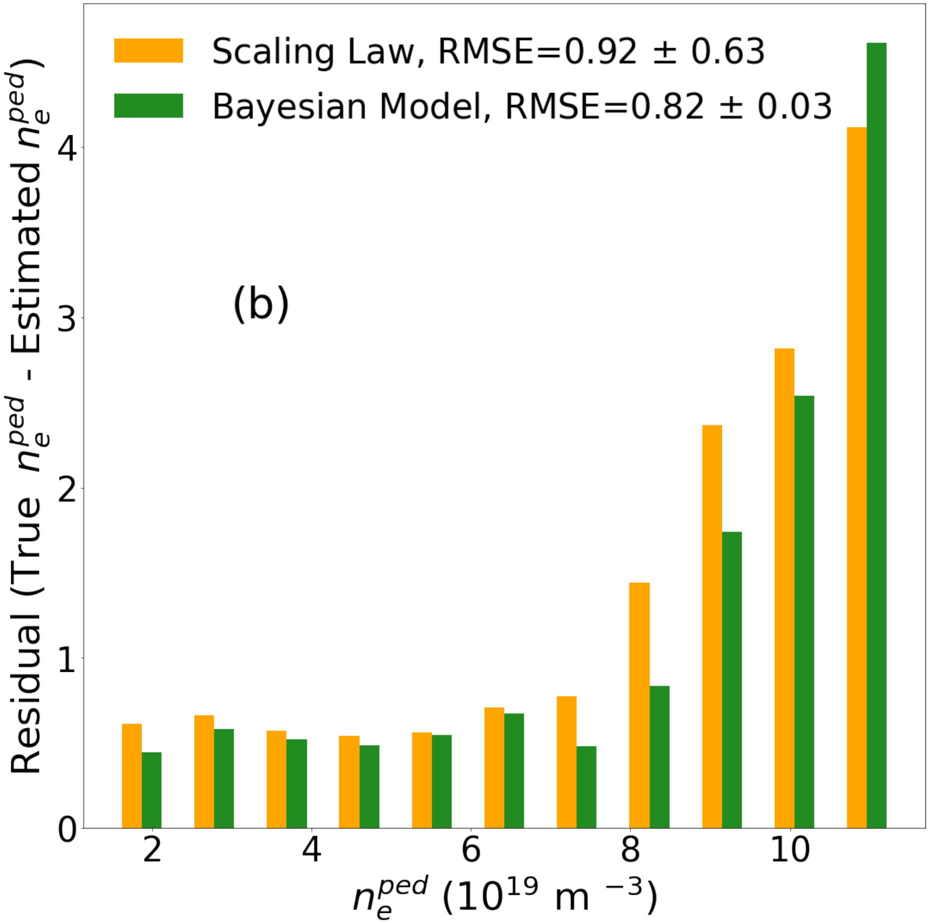
T. Hastie, et al. The Elements of Statistical Learning. 2001

Utilizing more input parameters yields higher prediction quality than existing scaling law

$$n_e^{ped} = (9.9 \pm 0.3) I_p^{1.24 \pm 0.19} P_{TOT}^{-0.34 \pm 0.11} \delta^{0.62 \pm 0.14} \Gamma^{0.08 \pm 0.04} M_{eff}^{0.2 \pm 0.2}$$

L. Frassinetti et al. *Pedestal Structure, stability and scalings in JET-ILW: the EUROfusion JET-ILW pedestal database* (2020)





JET

Including more input parameters reduces interpretability through cross-correlation

$$n_e^{ped} = (9.9 \pm 0.3) I_p^{1.24\pm0.19} P_{TOT}^{-0.34\pm0.11} \delta^{0.62\pm0.14} \Gamma^{0.08\pm0.04} M_{eff}^{0.2\pm0.2}$$

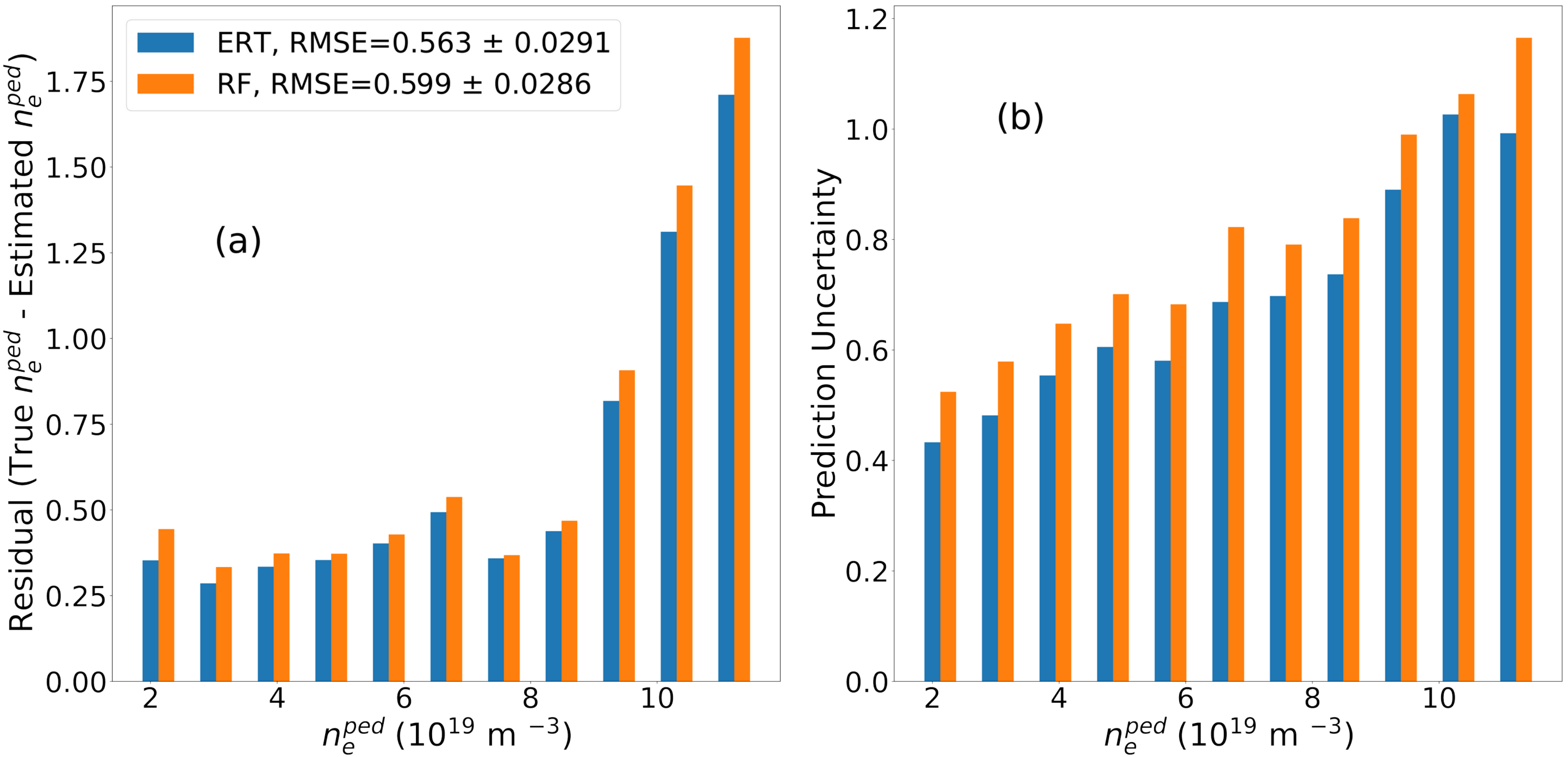
L. Frassinetti et al. *Pedestal Structure, stability and scalings in JET-ILW: the EUROfusion JET-ILW pedestal database* (2020)

- Bayesian Model coefficients μ are normally distributed with std. σ^2
- Published scaling law tries to avoid cross-correlation in variables, while Bayesian Model does not

Feature	μ	σ^2
I_p	0.15	0.06
B_T	0.956	0.072
a	2.966	0.479
δ	12.95	0.154
V_P	-0.05	0.007
q_{95}	-1.064	0.0542
P_{NBI}	-1.911	0.0546
P_{ICRH}	-1.976	0.0561
P_{TOT}	1.926	0.0557
Γ	0.125	0.007
H	-4.016	0.374
M_{eff}	1.369	0.053

JET

Non-parametric models (Random Forests and Extreme Randomized Trees) outperform linear models (Bayes)

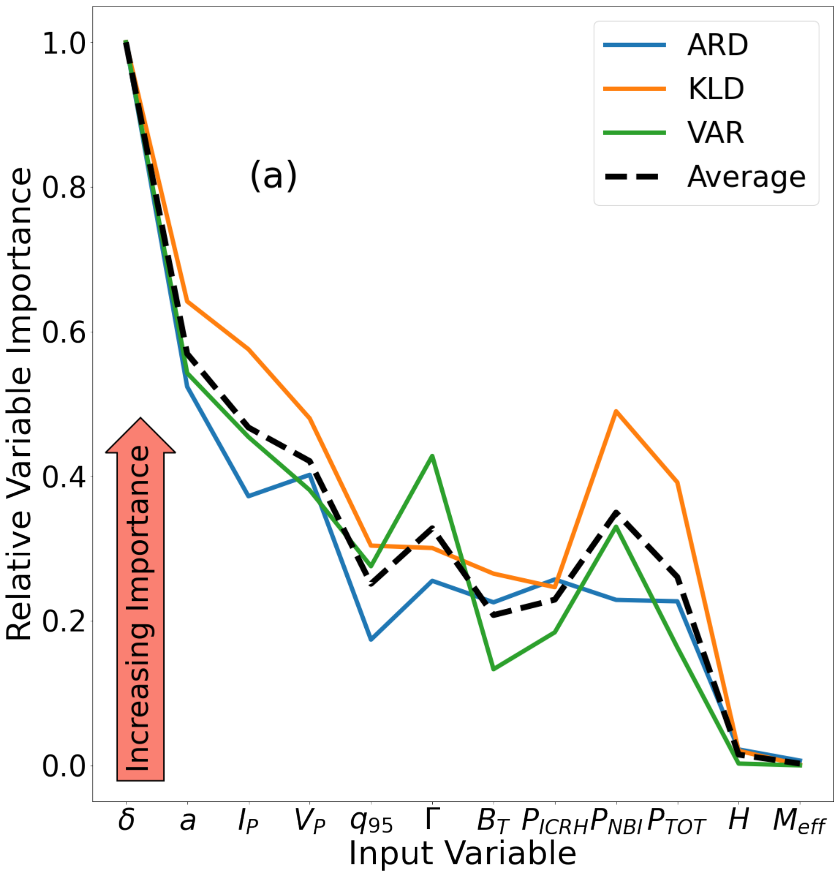


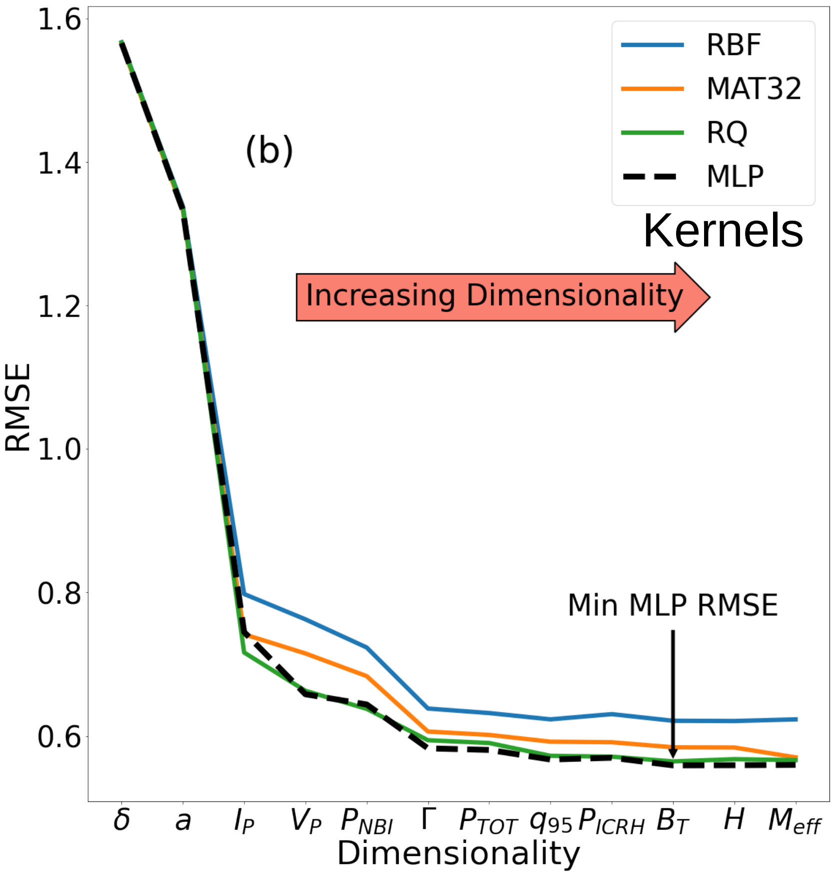
(b) Prediction uncertainty determined from standard deviation of the predictions from each decision tree in forest

Model	Scaling	Bayes
RMSE	0.92	0.82

JET

Dimensionality reduction for Gaussian processes using sensitivity analysis yields removal of isotope ratio (H), Meff from input space



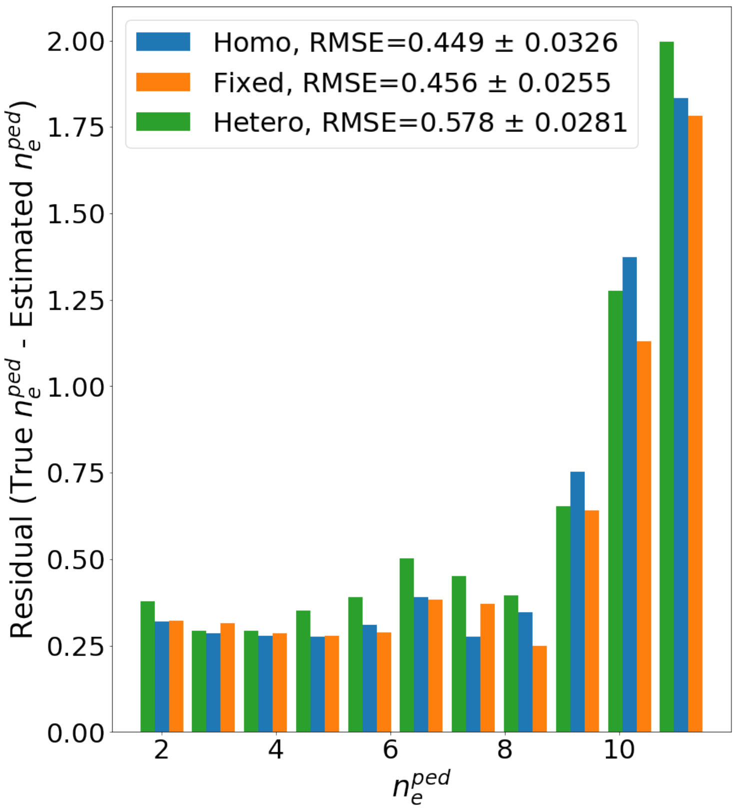


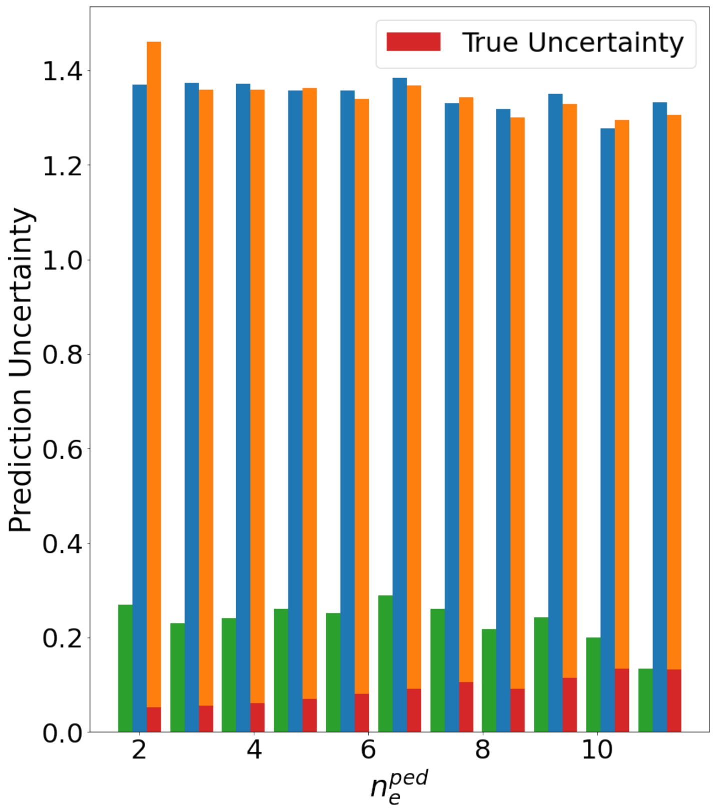
- Sensitivity Analysis: how much n_e^{ped} prediction changes with small variations to input variables
- ARD, KLD, and VAR are different methods of measuring the change in n_e^{ped} prediction in GPs

T. Paananen et al. *Variable selection for Gaussian processes via sensitivity analysis of the posterior predictive distribution* (2019)

JET

GPs outperform RFs and ERTs, and heteroscedastic models can capture local uncertainty



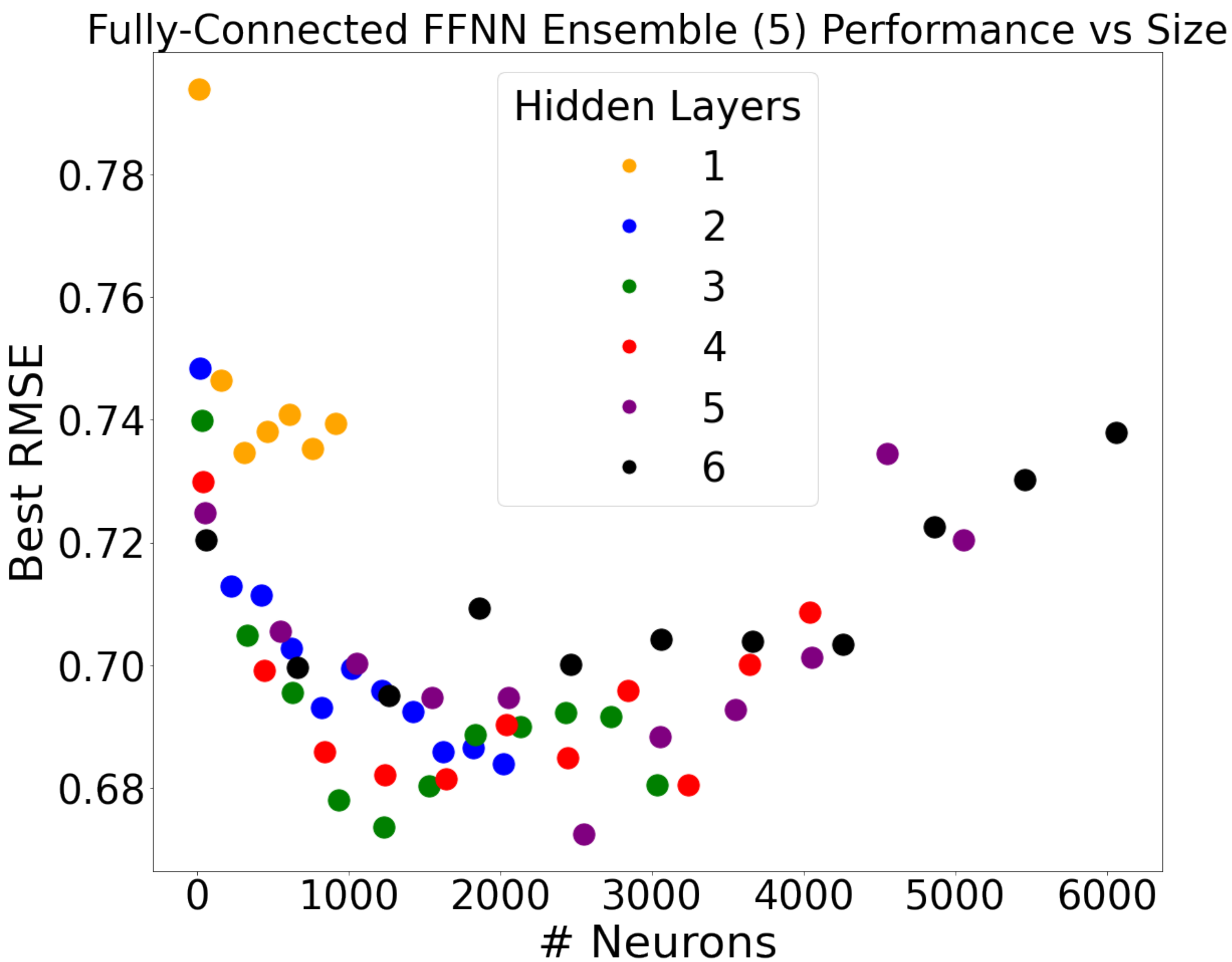


- Heteroscedastic models have the uncertainty in n_e^{ped} propagated alongside input parameters, unlike homoscedastic models which are "vanilla" GPs

Model	Scaling	Bayes	RF	ERT
RMSE	0.92	0.82	0.6	0.56

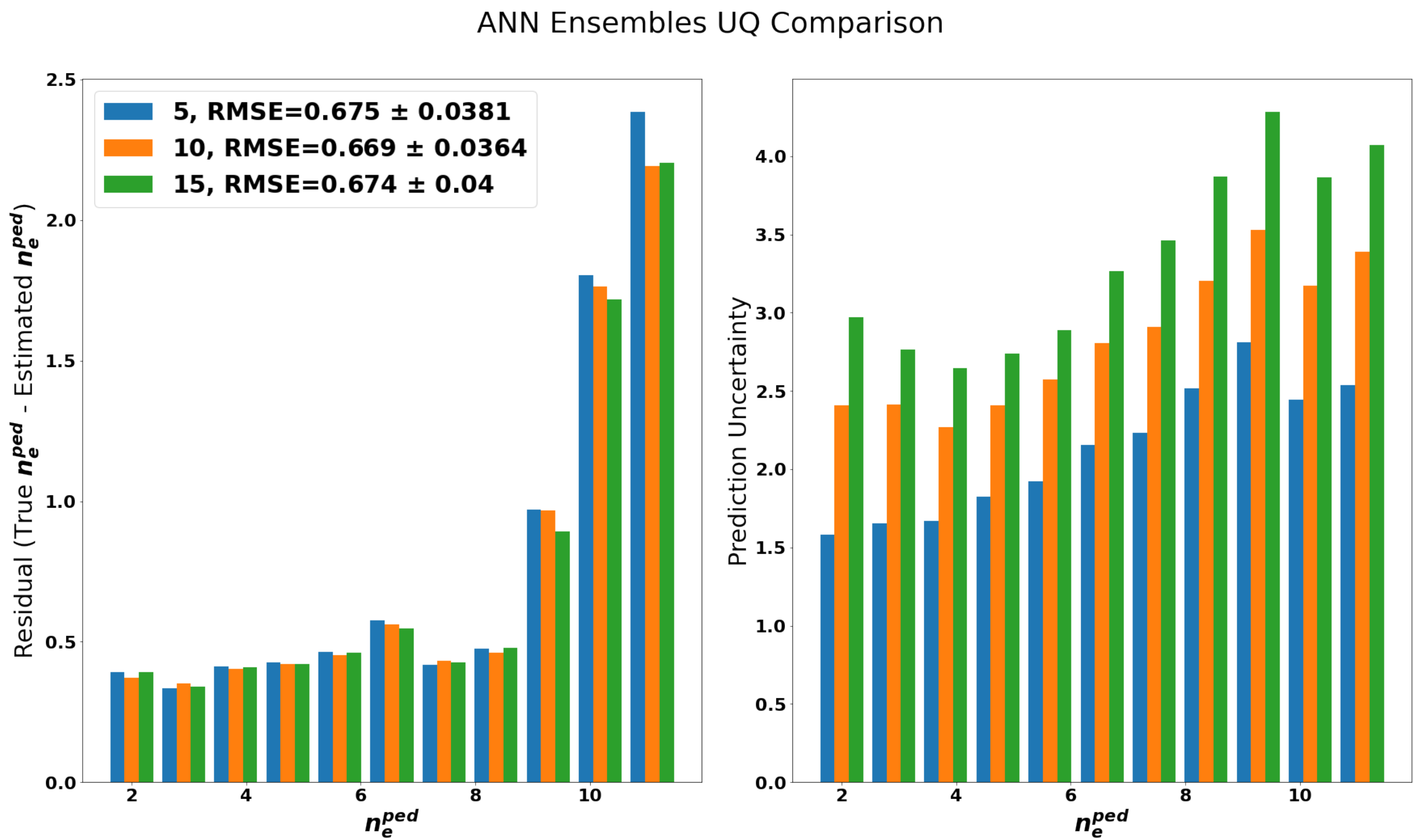
JET

Shallow Artificial Neural Networks outperform larger networks



JET

Prediction uncertainty increases with size of ensemble with similar performance



JET

Conclusions on predicting neped through emperical machine learning models

Model	Scaling	Bayes	RF	ERT	**GP**
RMSE	0.92	0.82	0.6	0.56	**0.449**

- New ML tools could improve the predictive capabilities of pedestal modeling
- Non-parametric models outperform linear models in terms of minimizing chosen error measures, with GPs showing most promise
- Heteroscedastic GPs can map local uncertainty while homoscedastic GPs do not
- Neped >= 9 is elusive to non-linear models
- Next Step
  - Separatrix quantities, ratio of separatrix to pedestal top, and connection between pedestal and SOL
  - Inverse UQ for a informed numerical model of EUROPED database