Gradient boosting models for photovoltaic power estimation under partial shading conditions

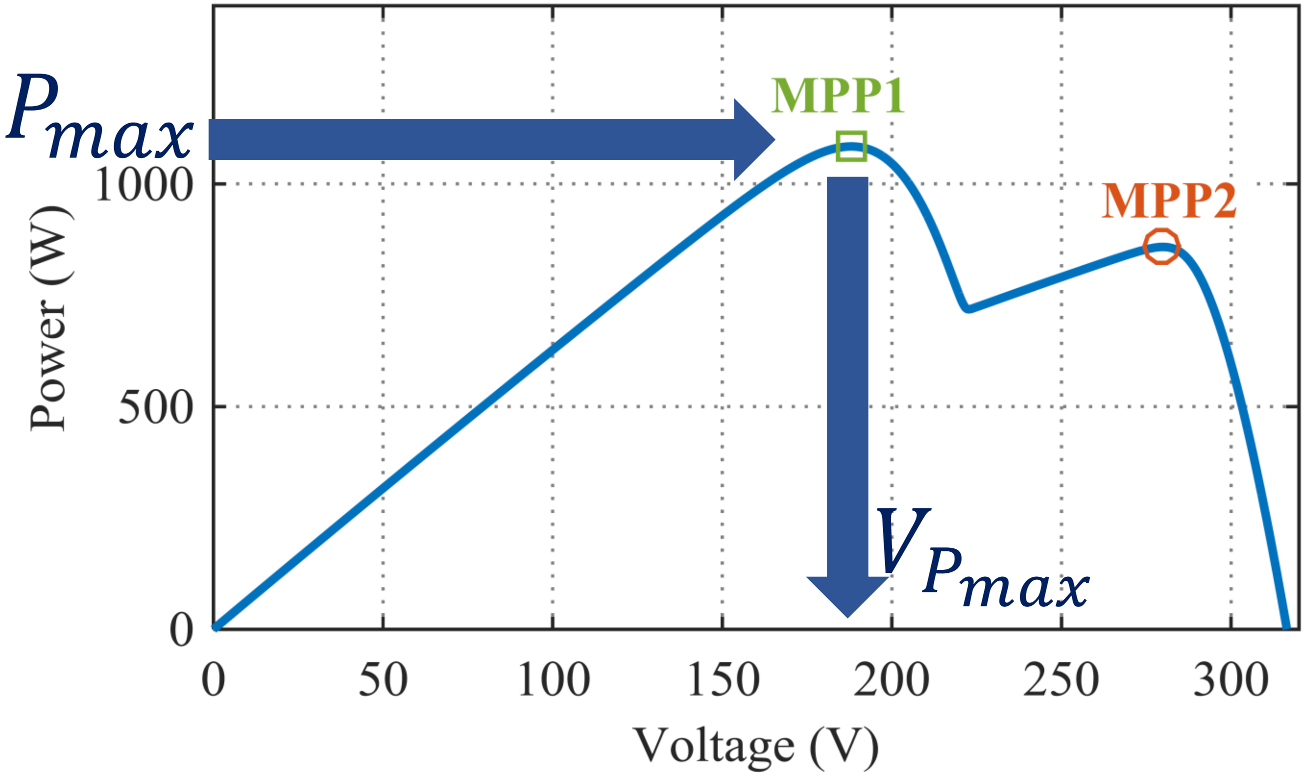
Nikolaos Nikolaou, Efstratios Batzelis, Gavin Brown

MANCHESTER 1824
The University of Manchester

Imperial College London

Partial shading in PV panel strings

Characteristic P-V curve of a partially shaded string



- Can have multiple **local** maximum-power-points (MPPs)
- As many MPPs as irradiance levels on a PV string
- Need to track **global** MPP $(P_{max}, V_{P_{max}})$ that provides maximal power output

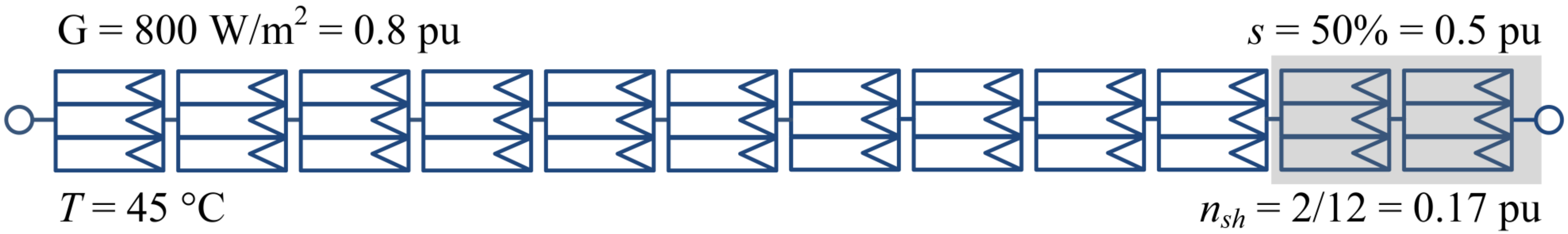
Main approaches Used

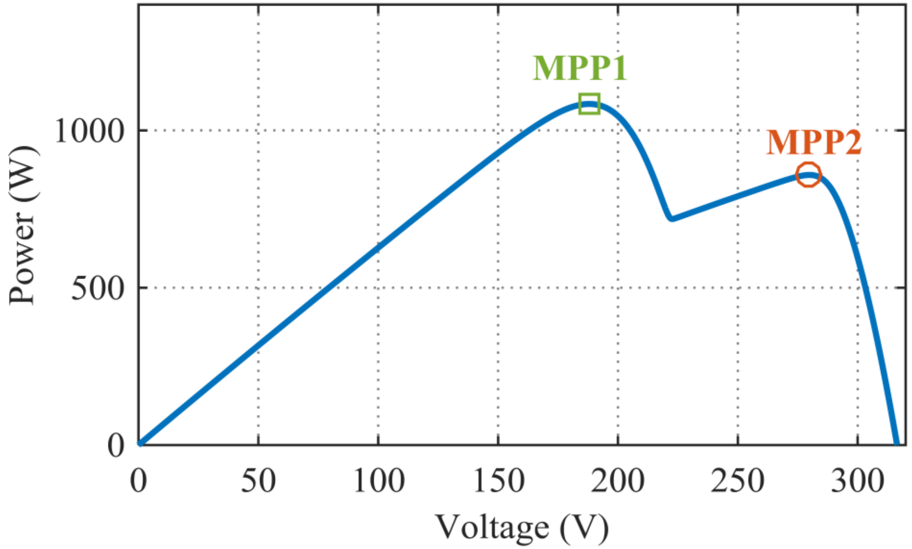
1. **Circuit-based** methods
   - Strong theoretical foundation
   - High accuracy
   - Require tedious simulations
2. **Heuristic** methods
   - Fast
   - Lower Accuracy

I. **Empirical formulas**
   - Simple formulation
   - Low accuracy

II. **Circuit equations**
   - Provide all MPPs
   - Good average accuracy
   - Occasionally high errors

Two irradiance levels (common case)



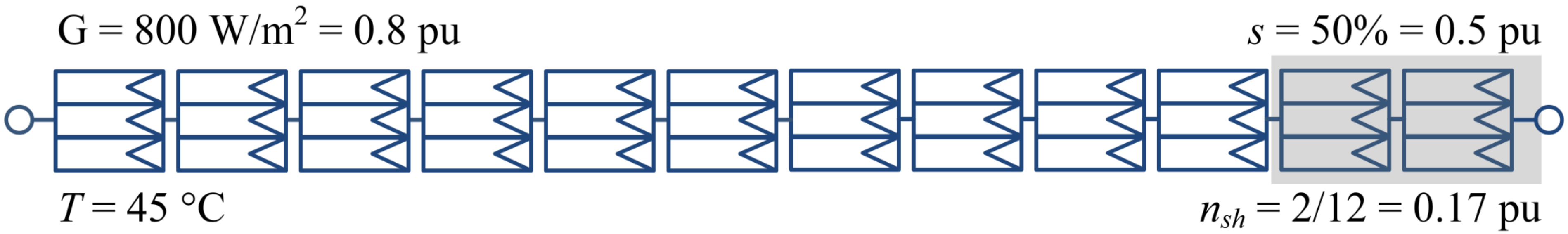


Operating conditions

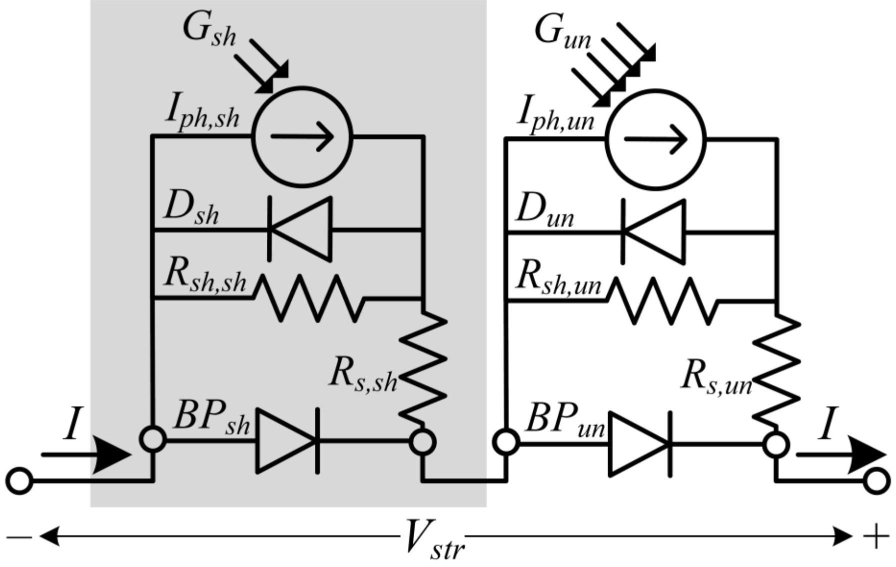
- G: irradiance (full)
- T: temperature
- s: shadow ratio
- n_{sh}: shadow extent

2 irradiance levels: P-V curve has up to 2 local MPPs

Equivalent circuit



- Simulate circuit under operating conditions to obtain P-V curve
- Find global MPP of P-V curve



Closed-form solution of the equivalent circuit

- Compute the local MPPs

$$MPP1 : \begin{cases} V_1 = N_{tot}[(1 - n_{sh})V_{mp}^T + n_{sh}\Delta V_D] \\ I_1 = GI_{mp}^T \\ P_1 = V_1 I_1 \end{cases}$$

$$MPP2 : \begin{cases} V_2 = N_{tot}[(1 - n_{sh})(sV_{mp}^T + (1 - s)V_{oc}^T) + n_{sh}V_{mp}^T] \\ I_2 = sI_{mp}^T[1 + \lambda(1 - nsh)] \\ P_2 = V_2 I_2 \end{cases}$$

- Then find global MPP

$$global\ MPP : \begin{cases} P_{max} = max\{P_1, P_2\} \\ V_{P_{max}} = \{V_{i^*} : i^* = \arg\max_{i \in \{1,2\}} P_i\} \end{cases}$$

Modelling using machine learning

- Train on examples of
  input (feature) $[G, T, s, n_{sh}]$ &
  output (target) $[P_1, V_1, P_2, V_2, P_{max}, V_{P_{max}}]$ vector pairs

- **Goal**: Better approximate $(P_{max}, V_{P_{max}})$ than closed-form equations

- Also included **'intermediate targets'** of MPP1 (P_1, V_1) & MPP2 (P_2, V_2)
  2 of the 3 models we train use these

Modelling the circuit using machine learning

- Models examined: **Gradient Boosted Trees** (Regression & Classification)

  FW will include Random Forests (initial results favourable) & Neural Networks

- In every case, **consider P & V independent**

  Correlation very low - verified by initial experiments

  FW will include relaxing this assumption

Model 1: Direct modelling of global MPP

- Do not use intermediate targets (local MPPs) – **directly model global MPP**

- Train **2 regressors** (can do in parallel, since independent):

$$[G, T, s, n_{sh}] \rightarrow P_{max}$$
$$[G, T, s, n_{sh}] \rightarrow V_{P_{max}}$$

- Given a reasonable amount of data **already beats closed-form equations**… But **can do better**!

Model 2: Stagewise modelling of global MPP

- First model local MPPs (intermediate targets), then predict global MPP

- Train **4 regressors** (can do in parallel, since independent):

$$\left.\begin{array}{l}[G, T, s, n_{sh}] \rightarrow P_1 \\ [G, T, s, n_{sh}] \rightarrow V_1\end{array}\right\} \text{MPP1}$$

$$\left.\begin{array}{l}[G, T, s, n_{sh}] \rightarrow P_2 \\ [G, T, s, n_{sh}] \rightarrow V_2\end{array}\right\} \text{MPP2}$$

- To predict global MPP: $\begin{cases} P_{max} = max\{P_1, P_2\} \\ V_{P_{max}} = \{V_{i^*} : i^* = \arg\max_{i \in \{1,2\}} P_i\} \end{cases}$

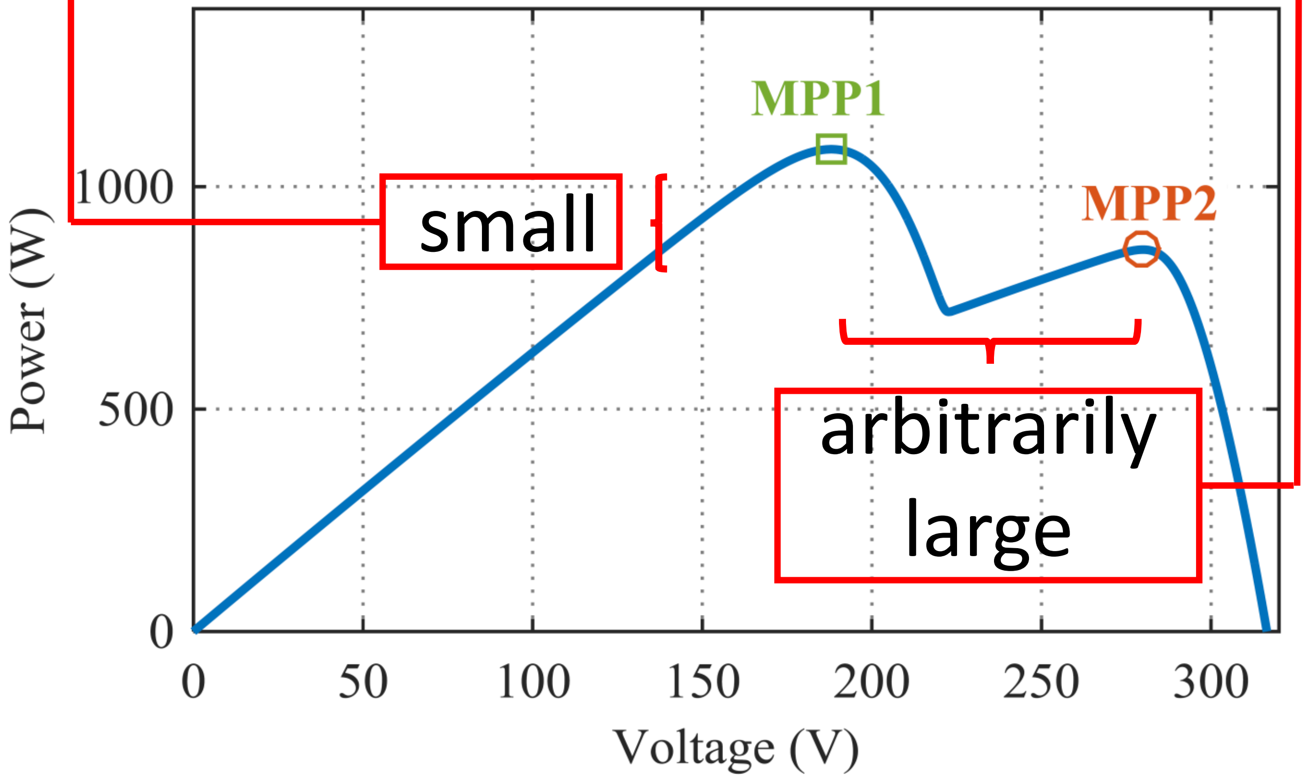
Model 2: Stagewise modelling of global MPP

- Given a reasonable amount of data **already beats closed-form equations**…
  Also **beats direct model** (Model 1)…
  But **can do even better**!

- **Why?** Because **intermediate outputs are FAR better estimated than those of closed-form equations**. In P_{max} this is reflected, but **why not** in $V_{P_{max}}$?

Model 2: Stagewise modelling of global MPP

- **Hypothesis**: In situations when $P_1 \cong P_2$, small estimation errors will have a small effect on P_{max}, but can cause us to predict the wrong V_i as $V_{P_{max}}$



- But $|V_1 - V_2|$ can be arbitrarily large, this can mean a large estimation error in $V_{P_{max}}$

$$\begin{cases} P_{max} = max\{P_1, P_2\} \\ V_{P_{max}} = \{V_{i^*} : i^* = \arg\max_{i \in \{1,2\}} P_i\} \end{cases}$$

Model 3: Stagewise, classifier-aided modelling of global MPP

- First model local MPPs (intermediate targets) & a mapping from input to the local MPP that is the global one, then predict global MPP

- Train **4 regressors & 1 classifier** (can do in parallel, since independent):

$$\left.\begin{array}{l}[G, T, s, n_{sh}] \rightarrow P_1 \\ [G, T, s, n_{sh}] \rightarrow V_1\end{array}\right\} \text{MPP1}$$

$$\left.\begin{array}{l}[G, T, s, n_{sh}] \rightarrow P_2 \\ [G, T, s, n_{sh}] \rightarrow V_2\end{array}\right\} \text{MPP2}$$

$$[G, T, s, n_{sh}] \rightarrow \{MPP1\ is\ global, MPP2\ is\ global\}$$

- To predict global MPP: If $MPP1\ is\ global$, then $P_{max} = P_1$ and $V_{P_{max}} = V_1$

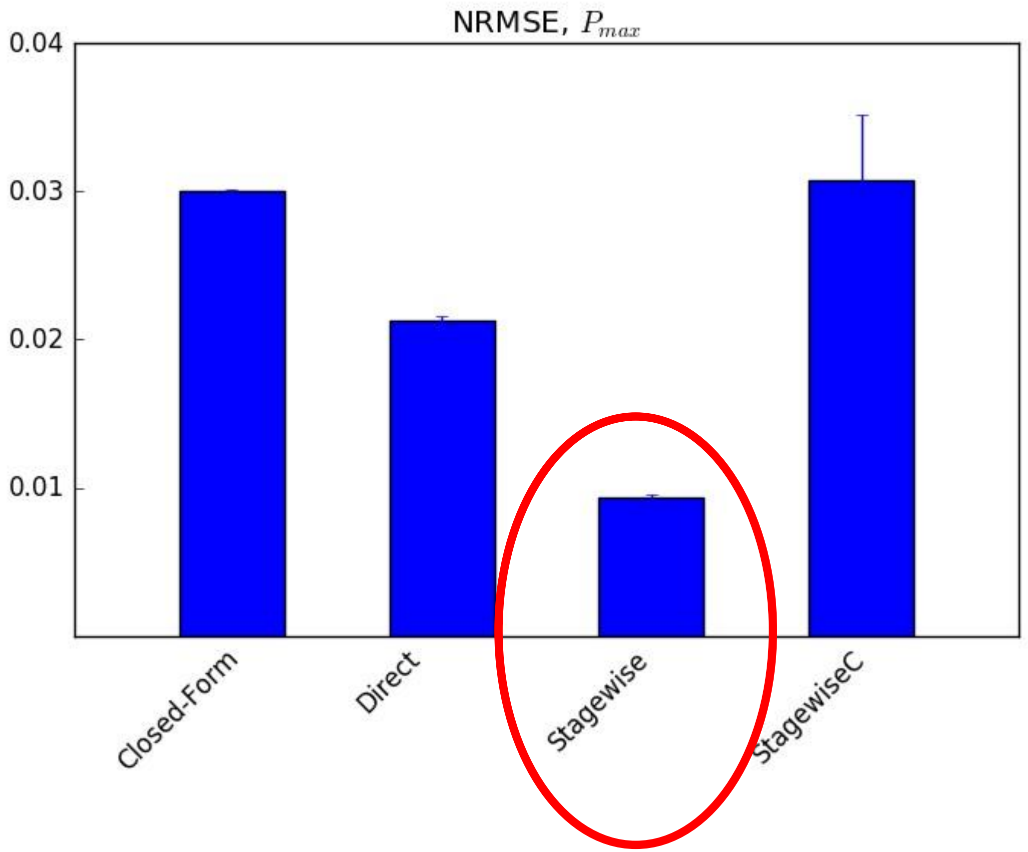
  Else $P_{max} = P_2$ and $V_{P_{max}} = V_2$

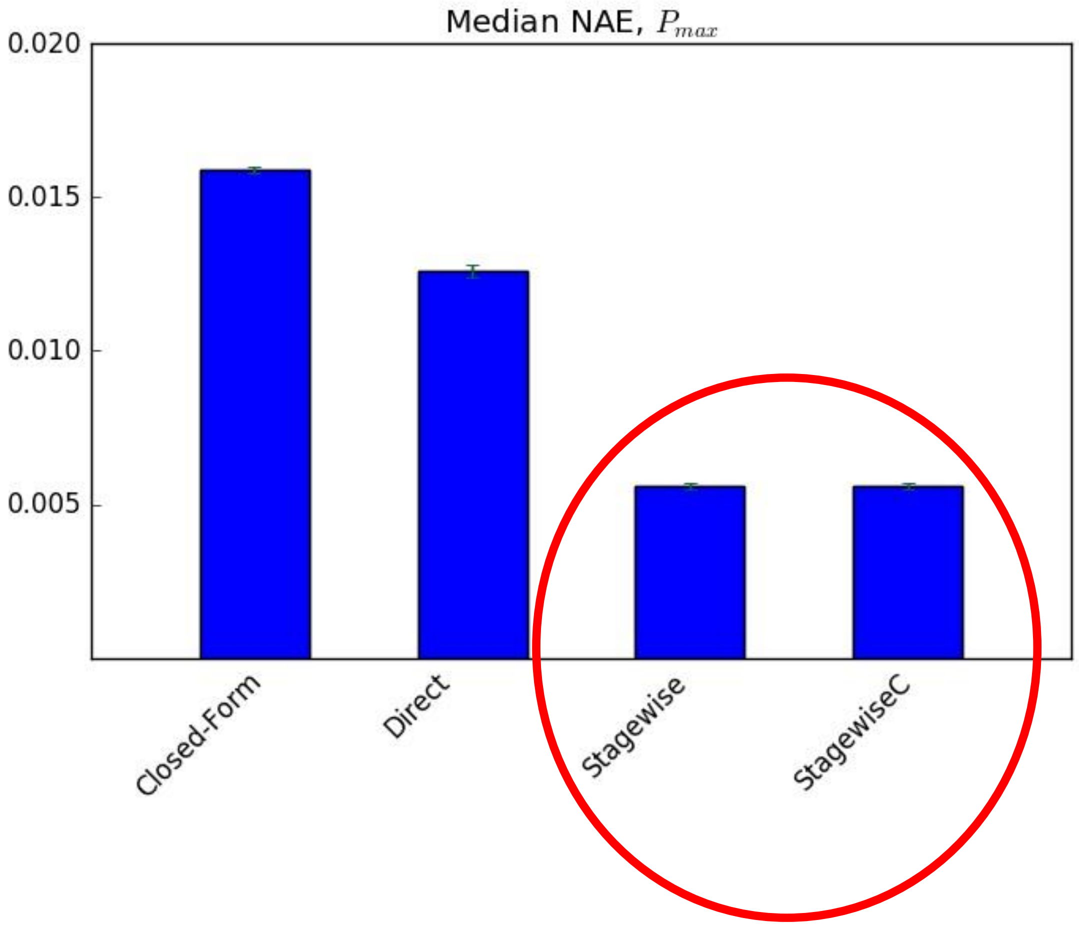
Experimental Setup

- 94905 datapoints generated by simulating circuit under various conditions
- Compare approximation of three models against closed-form estimates
- Ensemble size M=1000, tree depth d=3
- Trained on 75% of the data - have learning curves with fewer as well

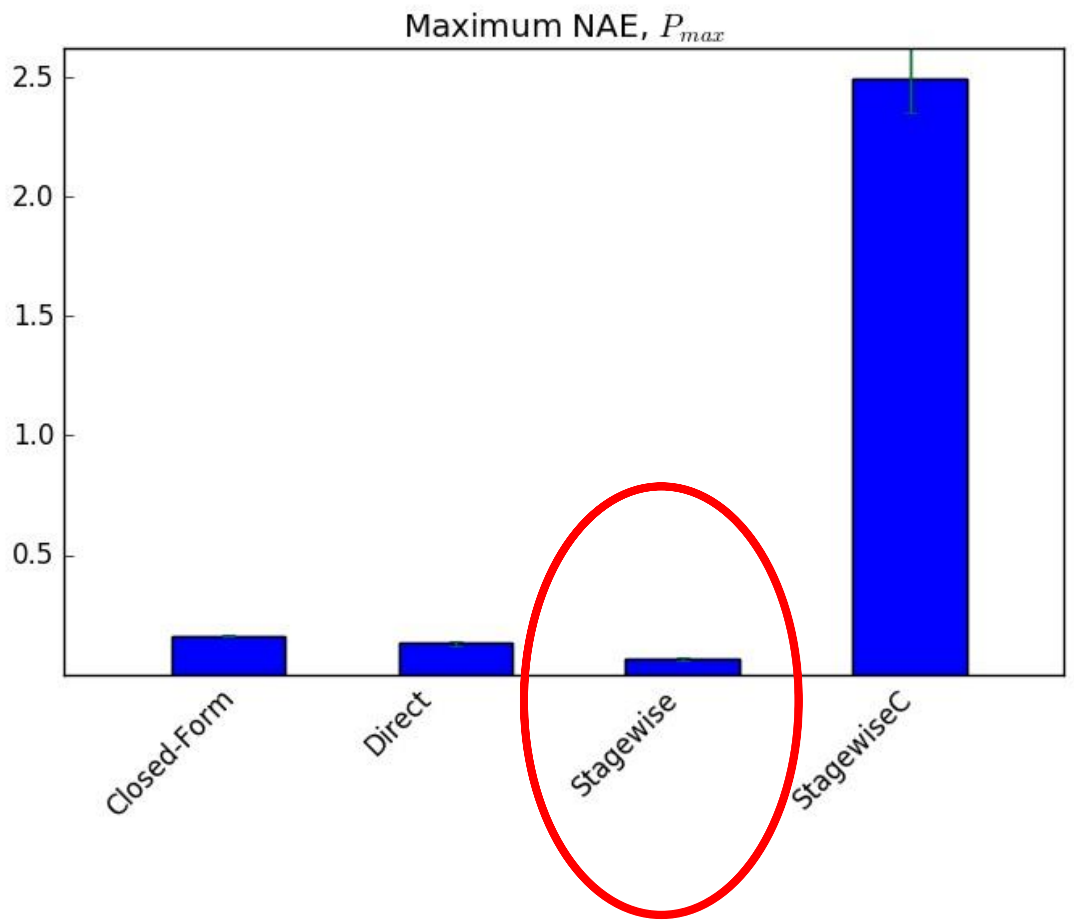
$$NRMSE = \sqrt{\sum_n (X_n - \hat{X}_n)^2} \Big/ \mu_X \qquad NAE_n = |X_n - \hat{X}_n| / \mu_X$$

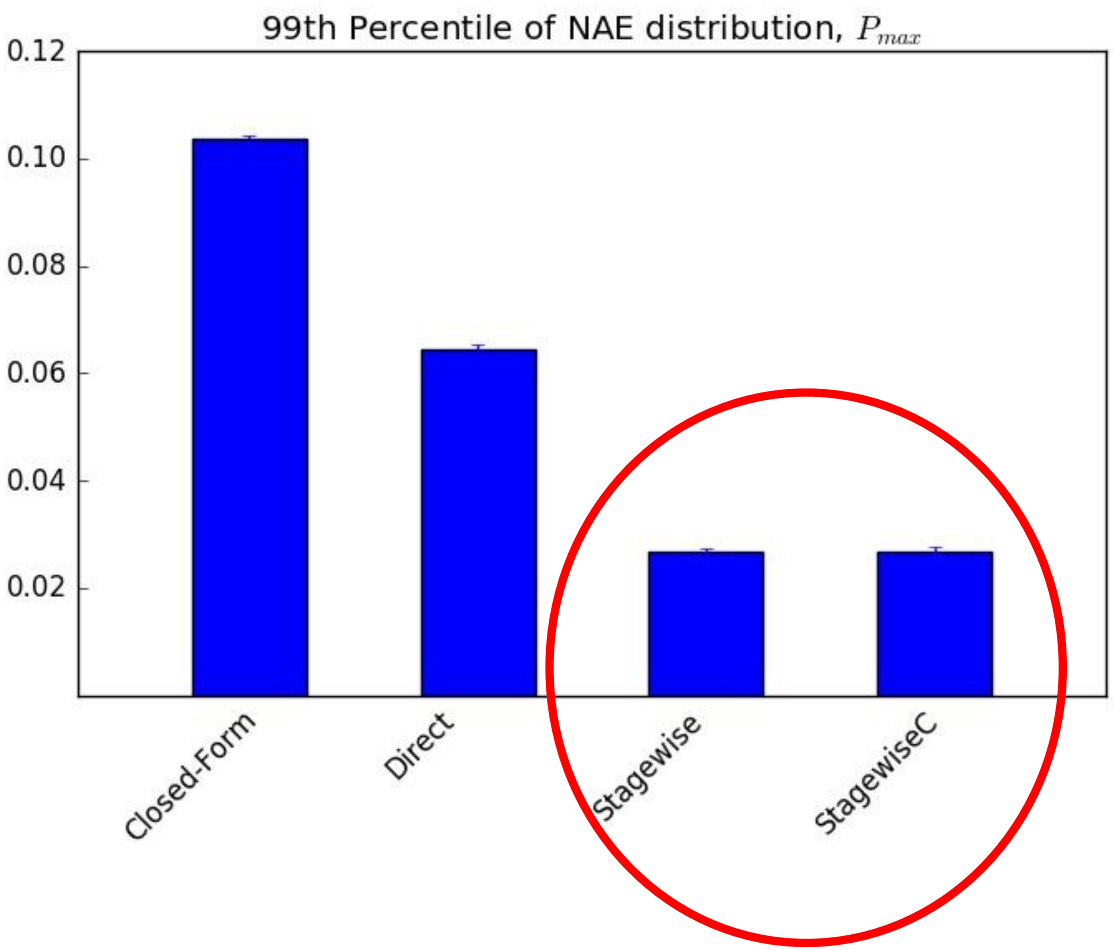
Results - P_{max} (Average NRMSE, Median NAE)



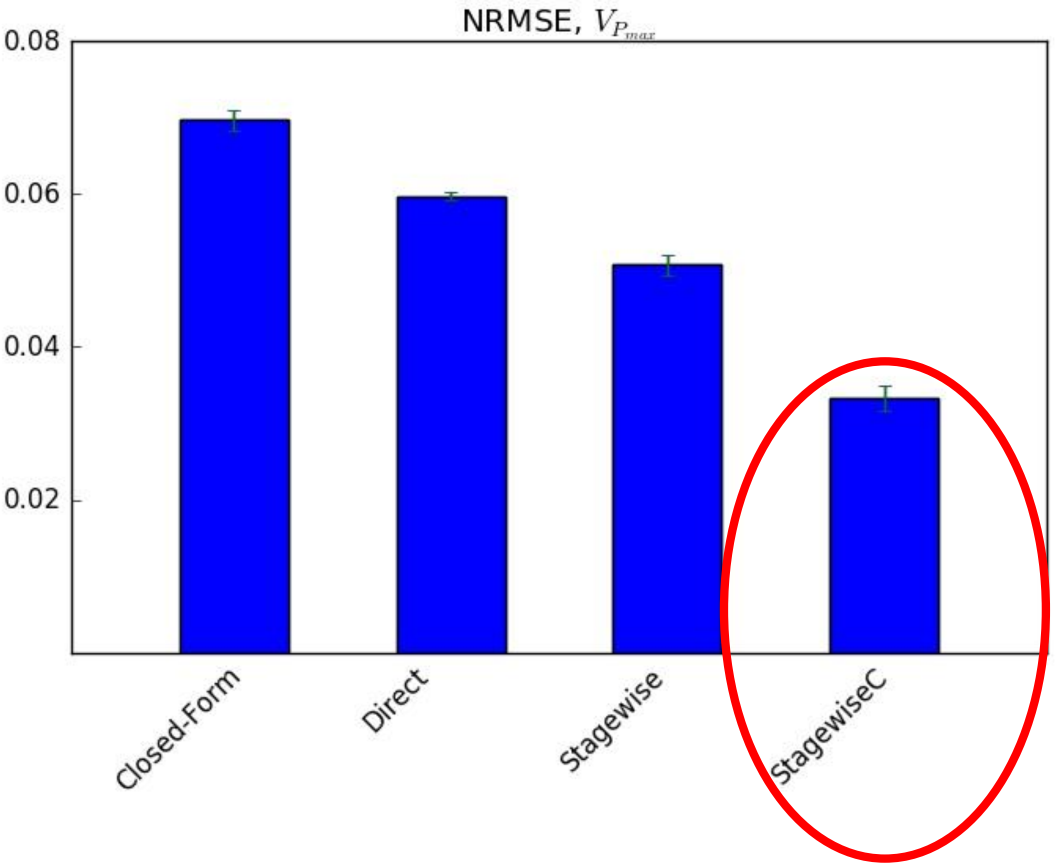


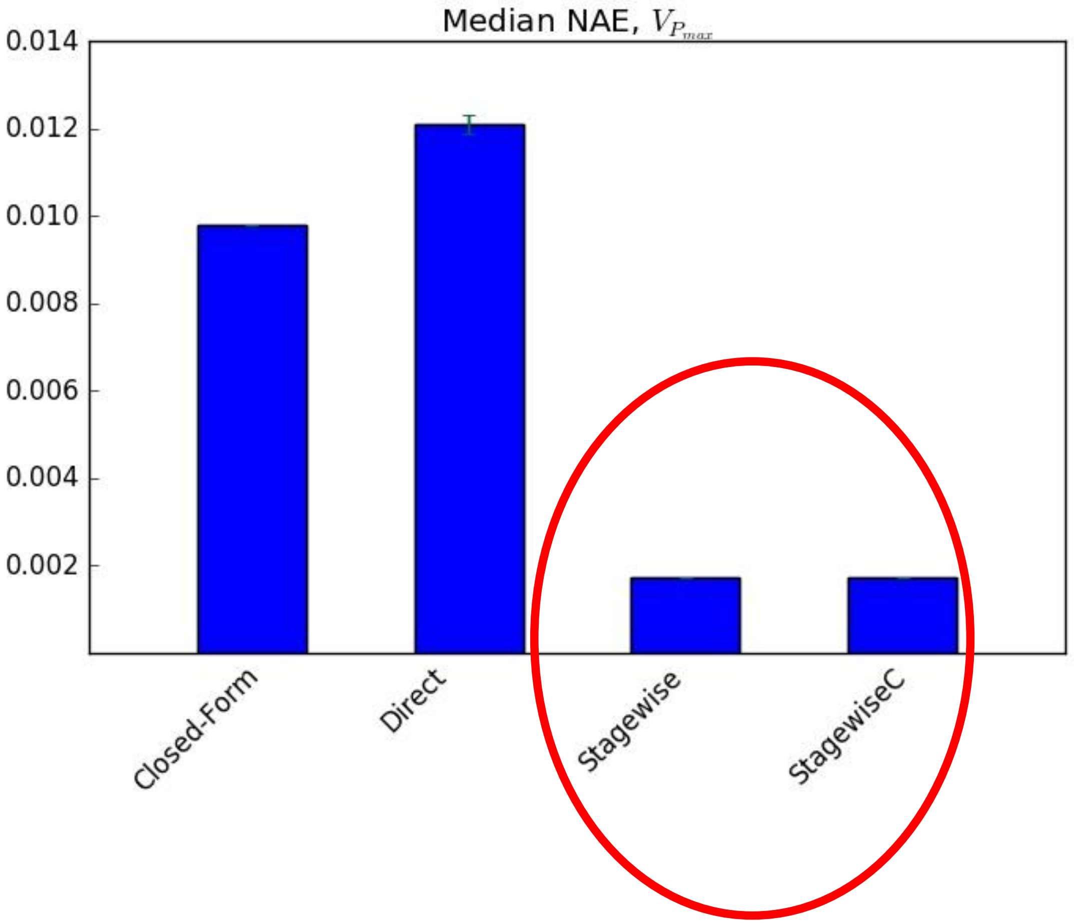
Results - P_{max} (Maximum NAE, 99th Percentile of NAE)



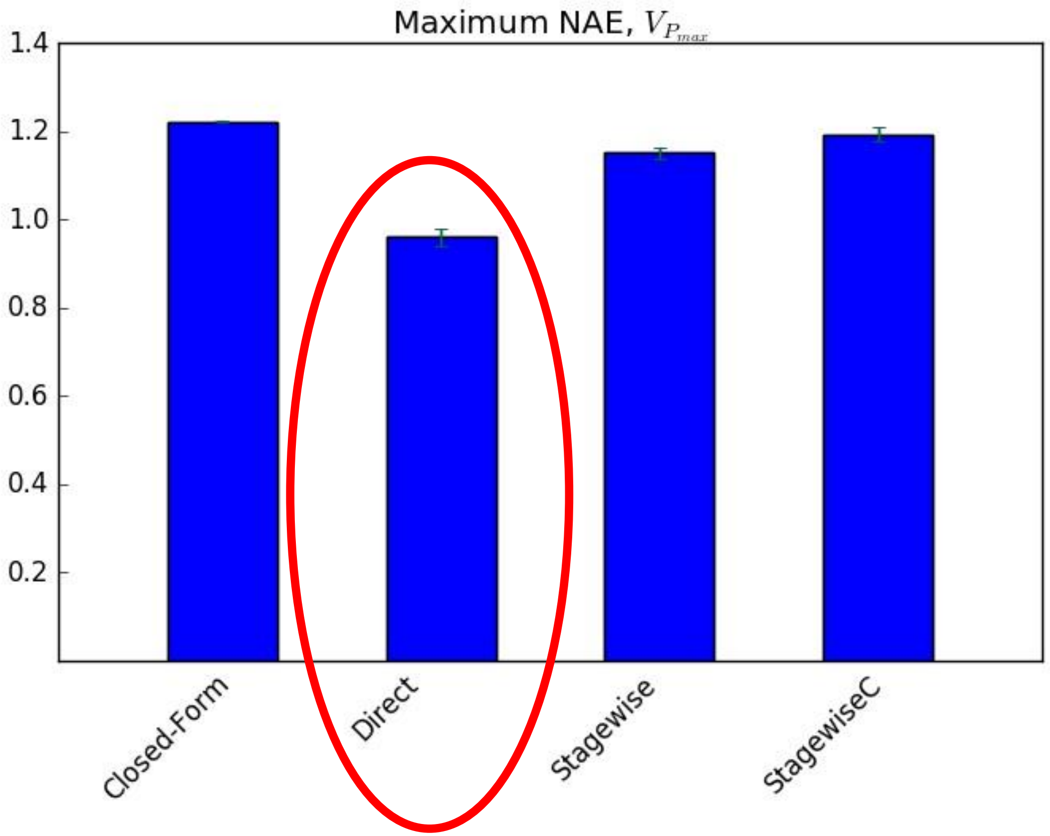


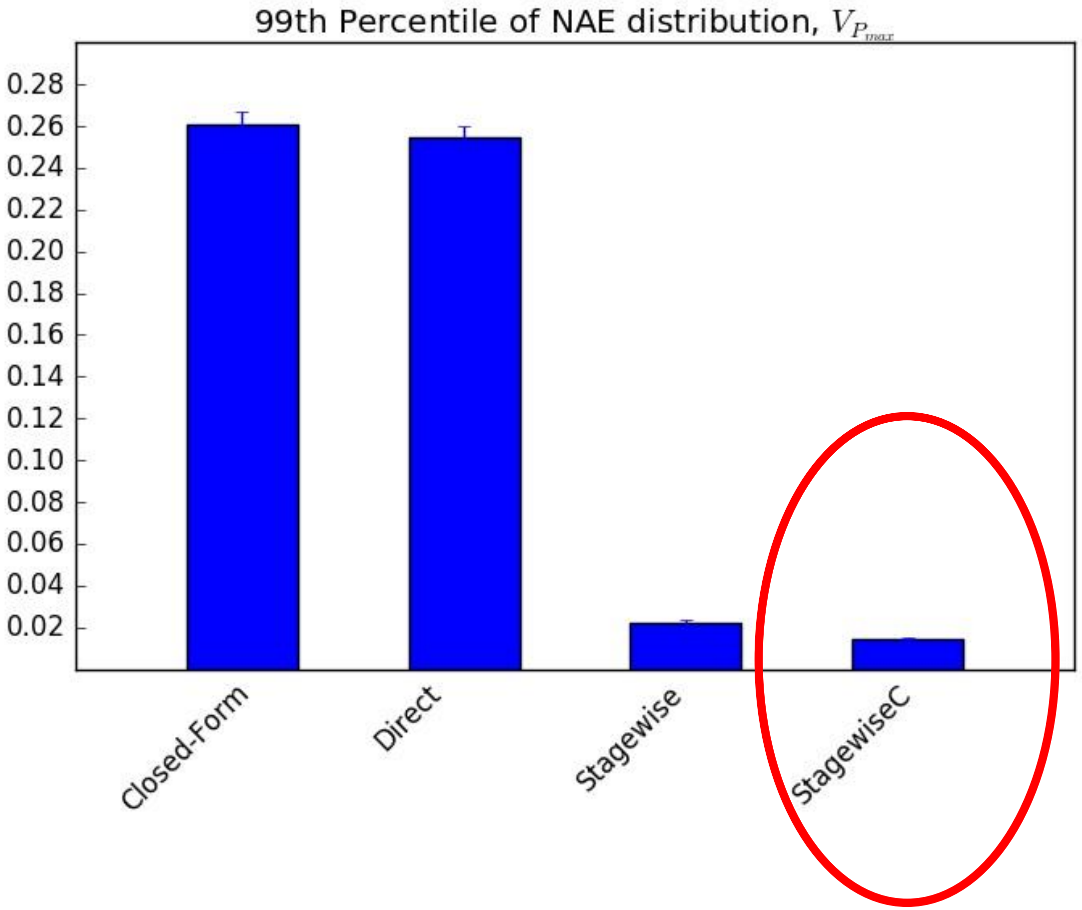
Results - $V_{P_{max}}$ (Average NRMSE, Median NAE)





Results - $V_{P_{max}}$ (Maximum NAE, 99th Percentile of NAE)

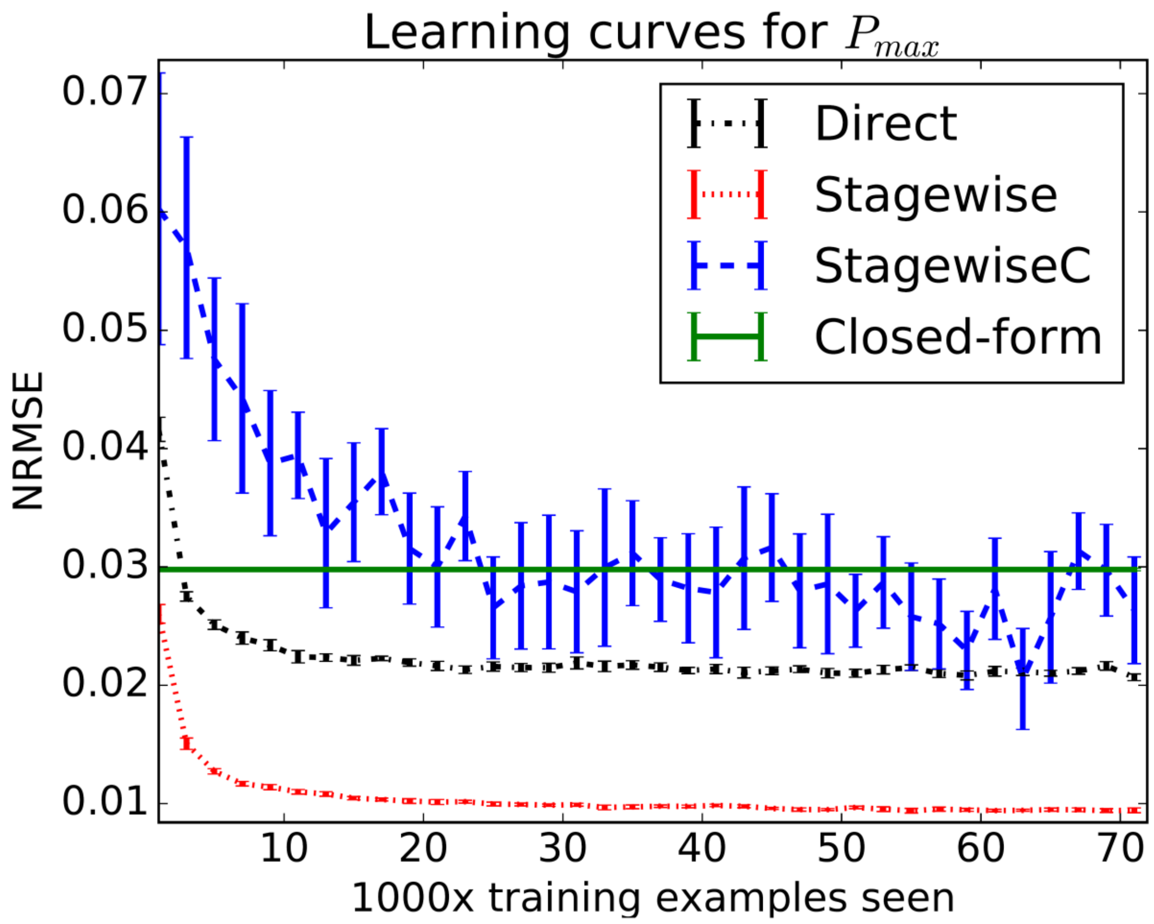


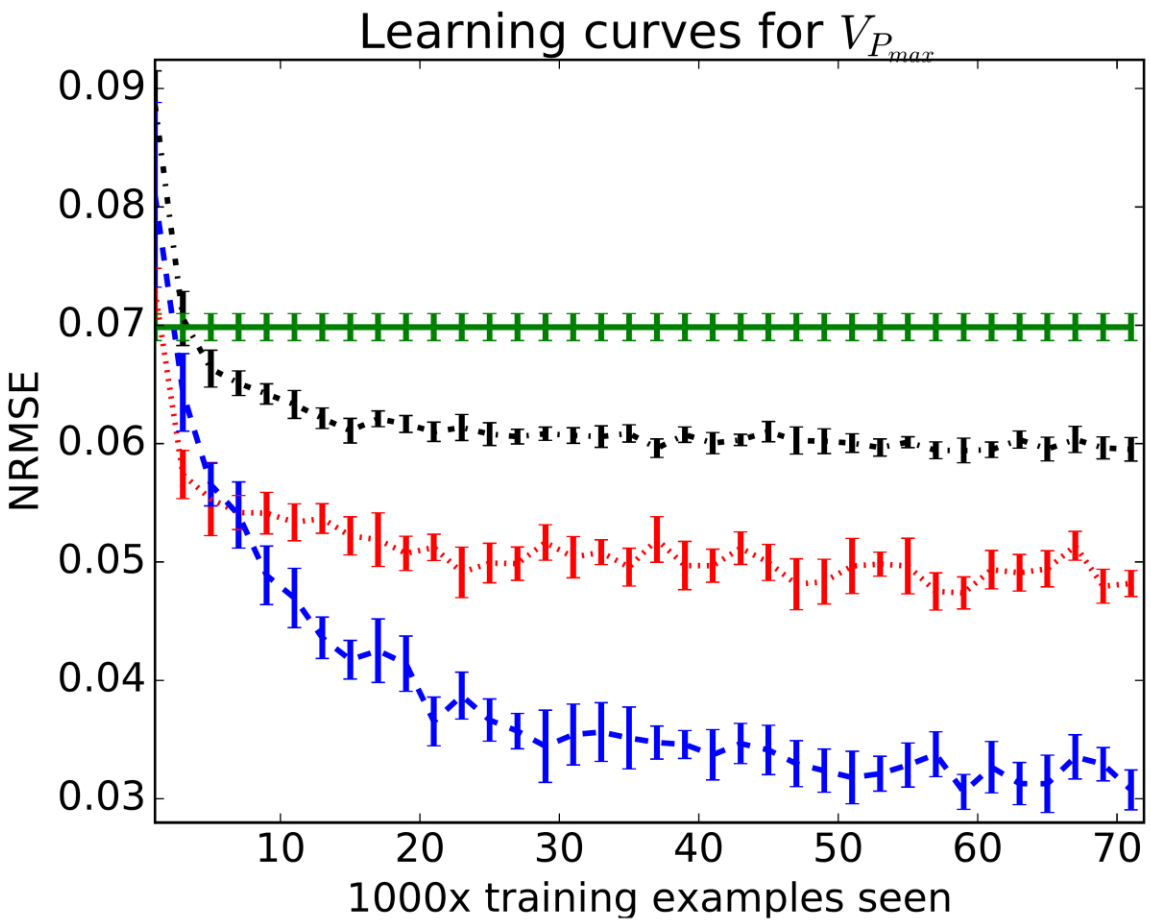


Results – advice for practitioners

- To minimize **maximum error in** $V_{P_{max}}$, use Model 1
- To minimize **average or maximum error in** P_{max} , use Model 2
- To minimize **average error in** $V_{P_{max}}$, use Model 3
- For **overall good** performance, use Model 2, or **combine** Models 2 & 3

Effect of training set size





Conclusions

- A **small training set is sufficient** to outperform closed-form…
- … **on average** (& median) **AND worst case**
- **Thin-tailed error distribution** - some statistical guarantees
- Increasing **ensemble size and/or tree depth** improves performance at increased computational cost
- To some degree **parallelizable**, **fast** to train - very fast to predict

Extensions

- More **irradiance levels** - up to 4-5 in practice
- Non-uniform **temperature**
- Different PV **configurations** **-** e.g. arrays (strings in parallel)
- More **machine learning methods** (Random Forests, Bagging, NNs)
- More experimentation with **hyperparameters,** multi-objective **optimization of cost & performance**
- Taking **correlations** into account
- **Interpreting models** - could relate them back to circuit theory

Thank you!