SymReg
JOSEF RESSEL CENTER FOR
SYMBOLIC REGRESSION

Integration of Physical Knowledge in Empirical Models – A New Approach to Regression Analysis

G. Kronberger, C. Haider, M. Kommenda, (University of Applied Sciences Upper Austria)
S. Scheidel (AVL List GmbH), M. Kordon (AVL Commercial Driveline & Tractor Engineering GmbH)

12. Nov. 2019

Summary

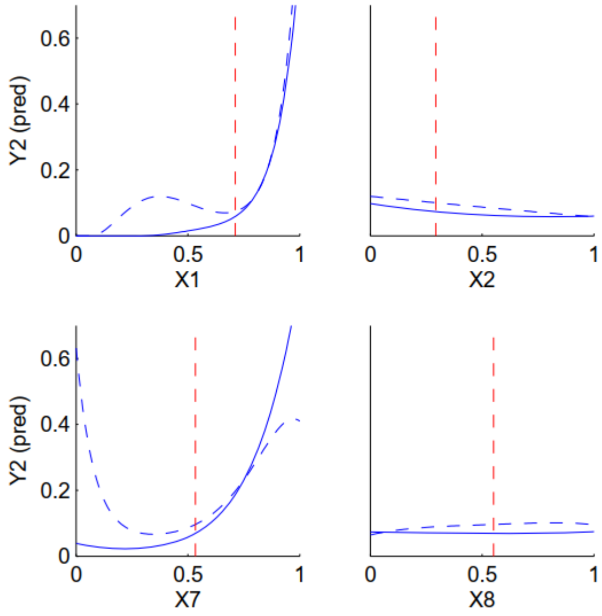
We include prior knowledge into empirical models by enforcing monotonicity over selected inputs

Method: Shape-constrained polynomial regression

Application: modelling diesel engine emission

Results:

- Model conforms to prior knowledge
- Extrapolation is plausible
- Slightly larger fitting errors



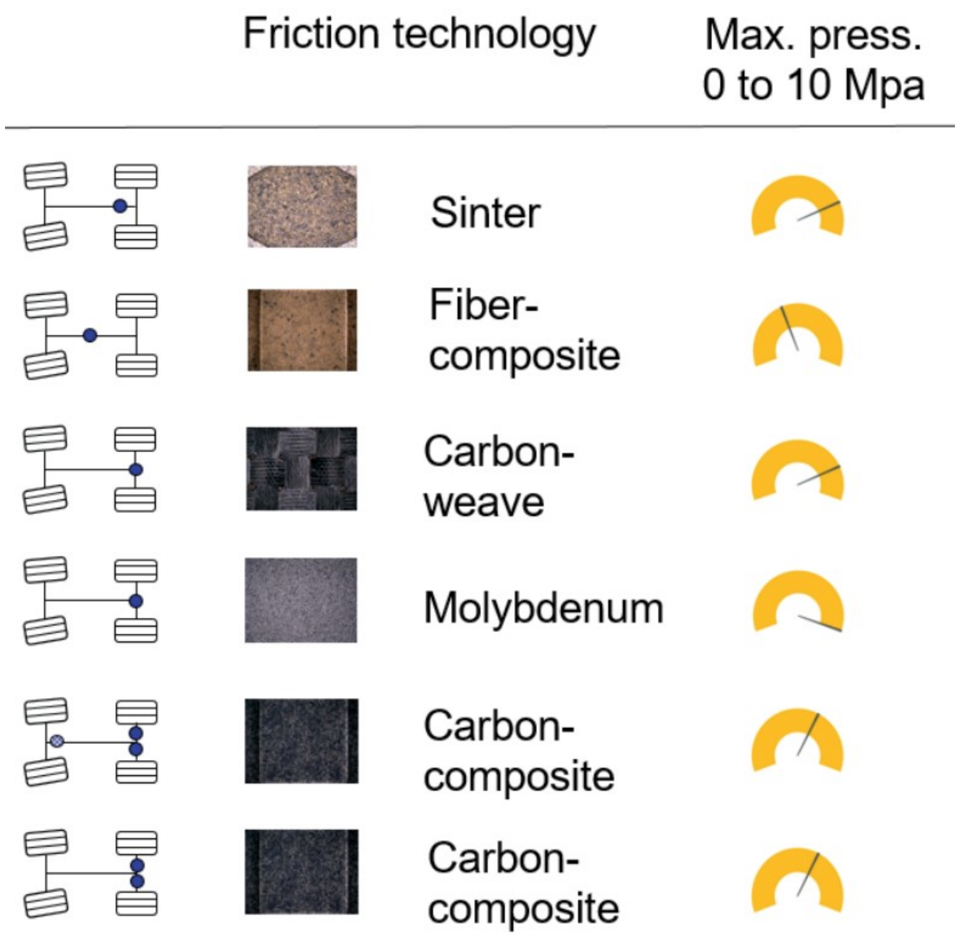
Background

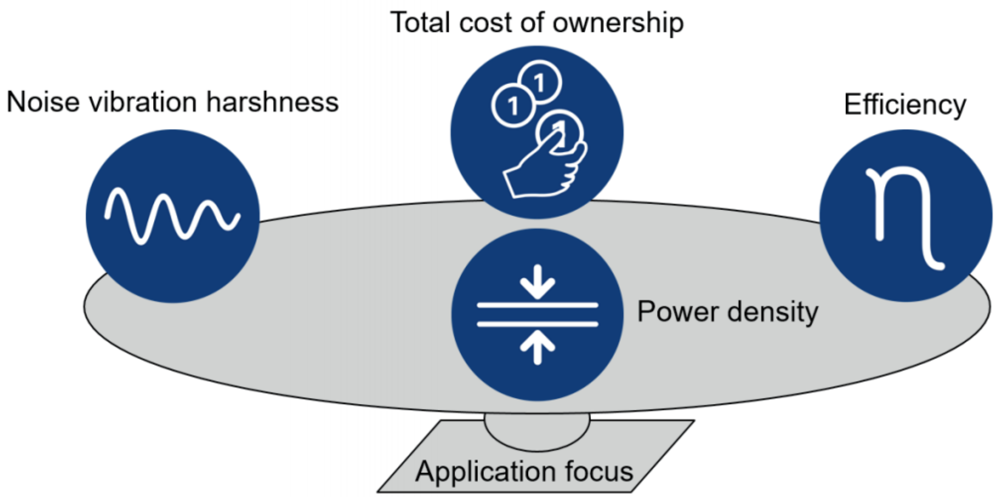
Empirical models and model-based optimization are widely known and applied in engineering.

Facts:

- No analytical model required
- Relatively low computational effort (compared to CFD, FEM)
- Large set of methods from statistical learning
- Broadly applicable
- Data availability and quality is essential
- Extrapolation is (almost) impossible or risky

Example: friction systems

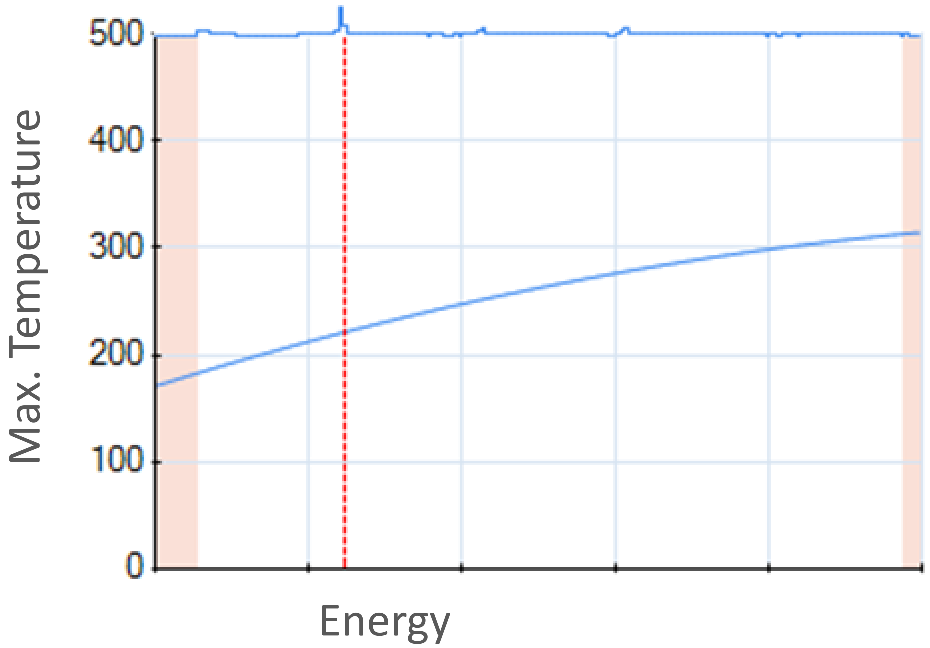


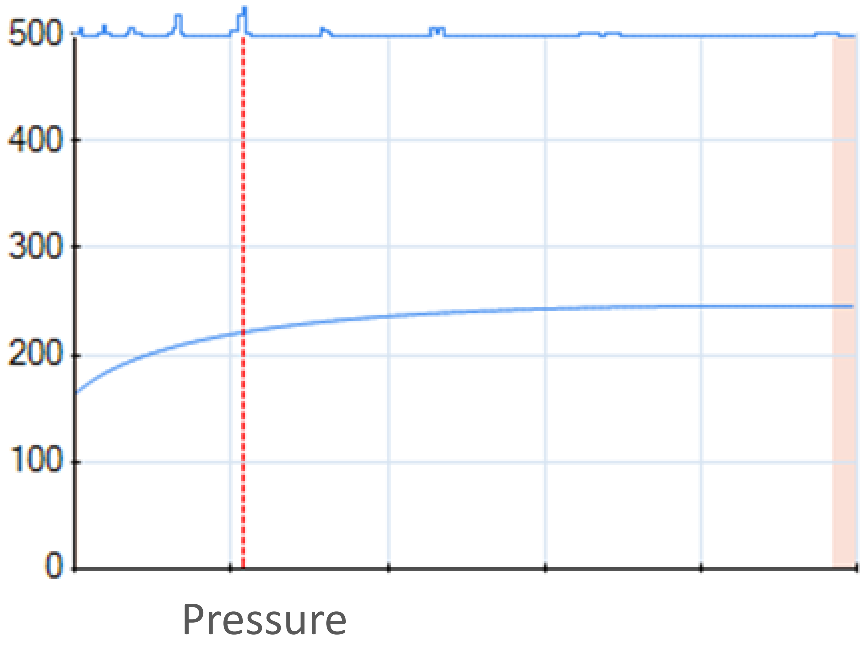


Miba AG, Falk Nickel, 16th International CTI Symposium Automotive Transmissions, HEV and EV Drives, Berlin, December 2017

A simple example

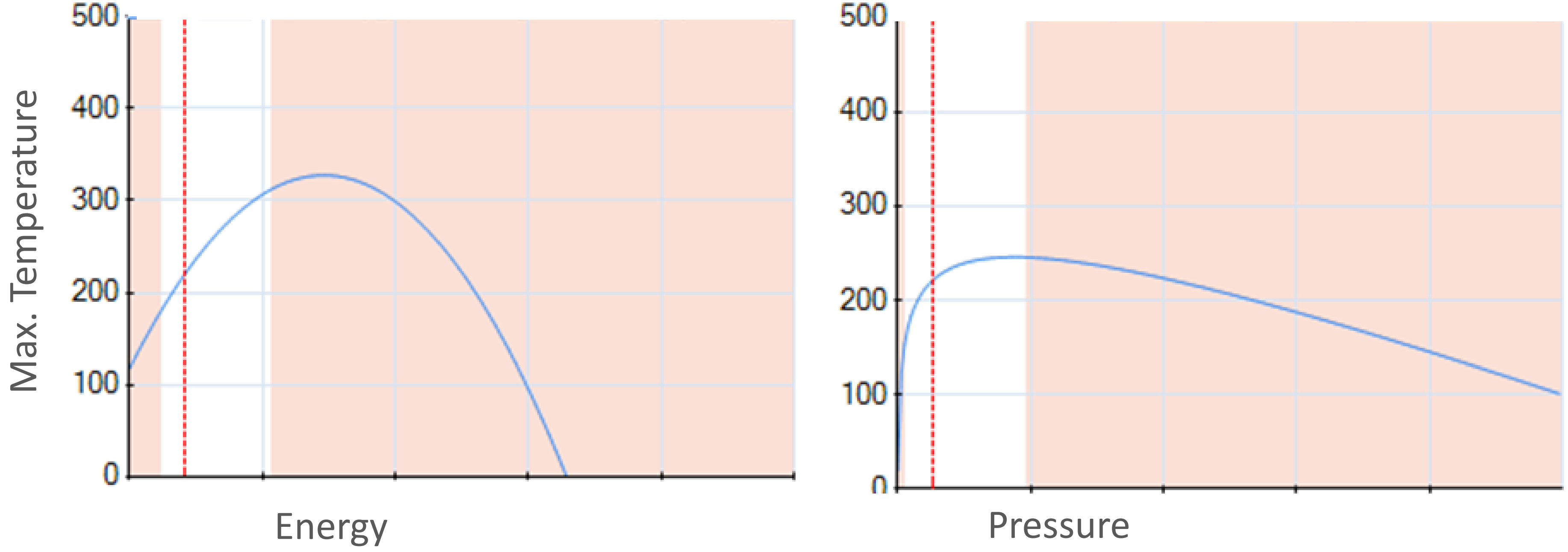
Empirical model for the maximum temperature of a disc brake



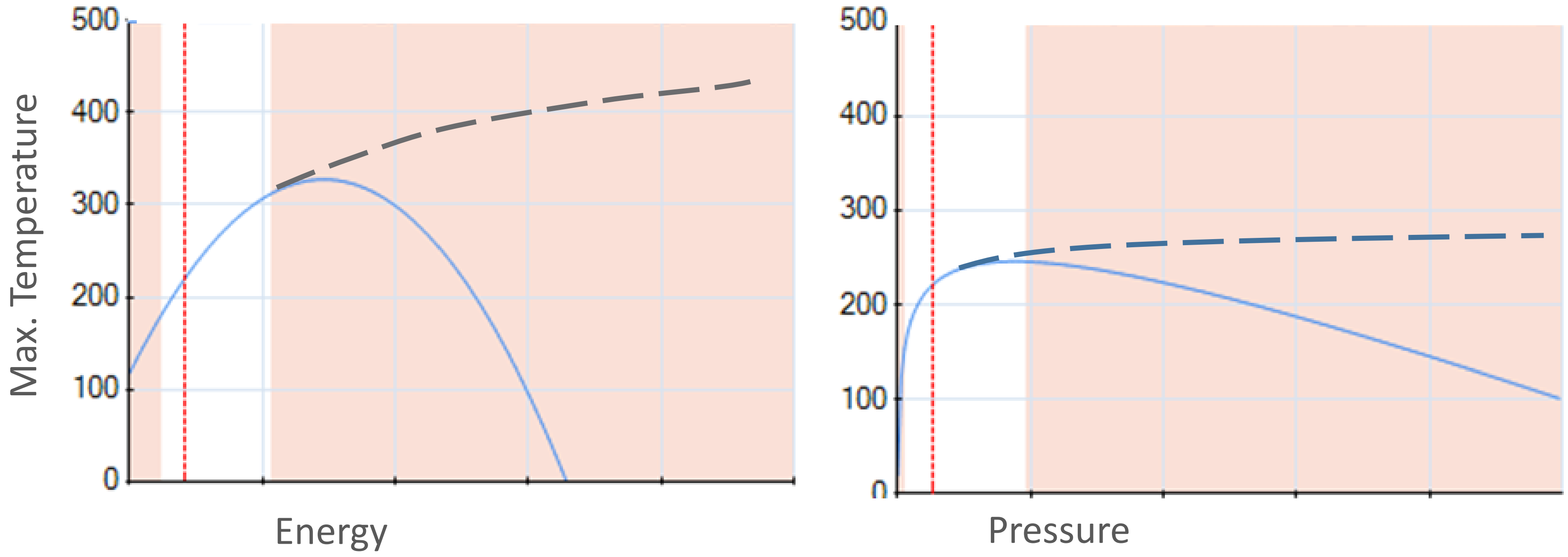


Extrapolating...

The prediction does not conform to physical intuition.

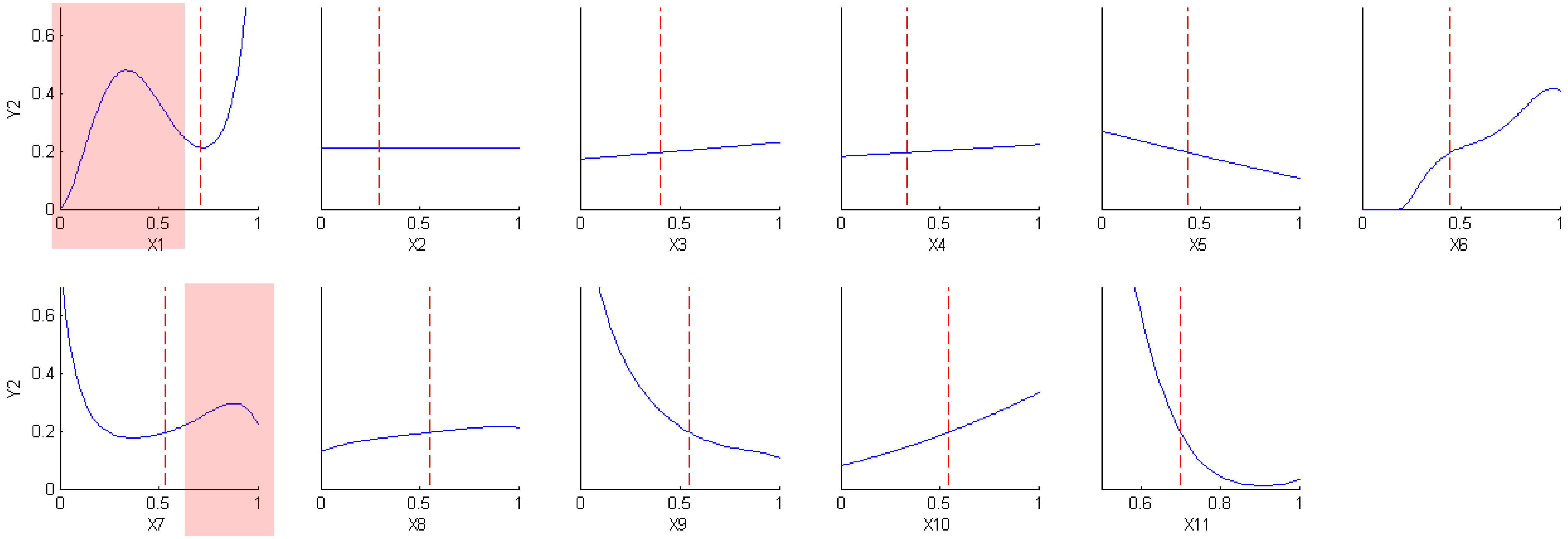


How can we enforce monotonicity?

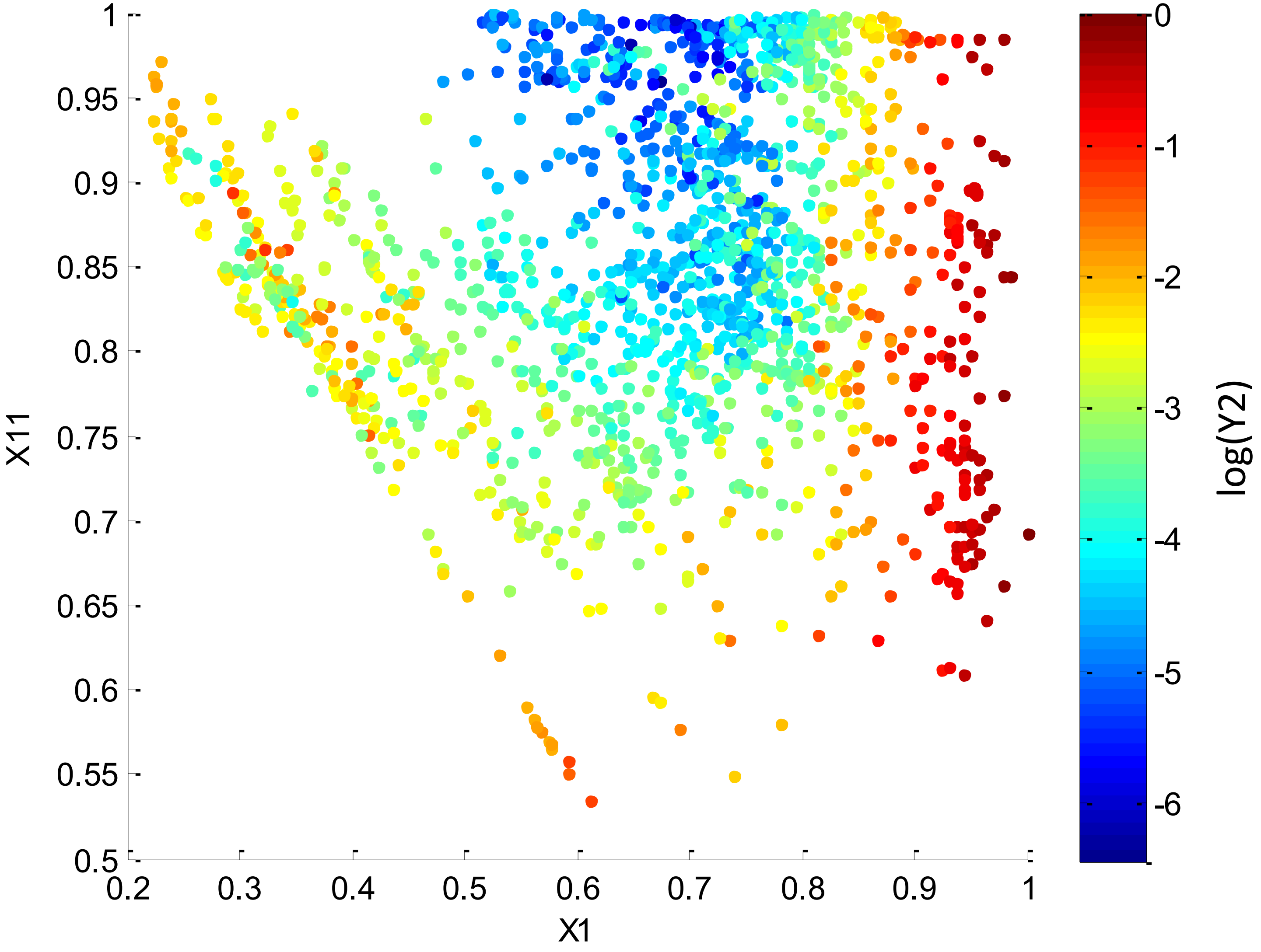


Another example

Empirical model for diesel engine emissions

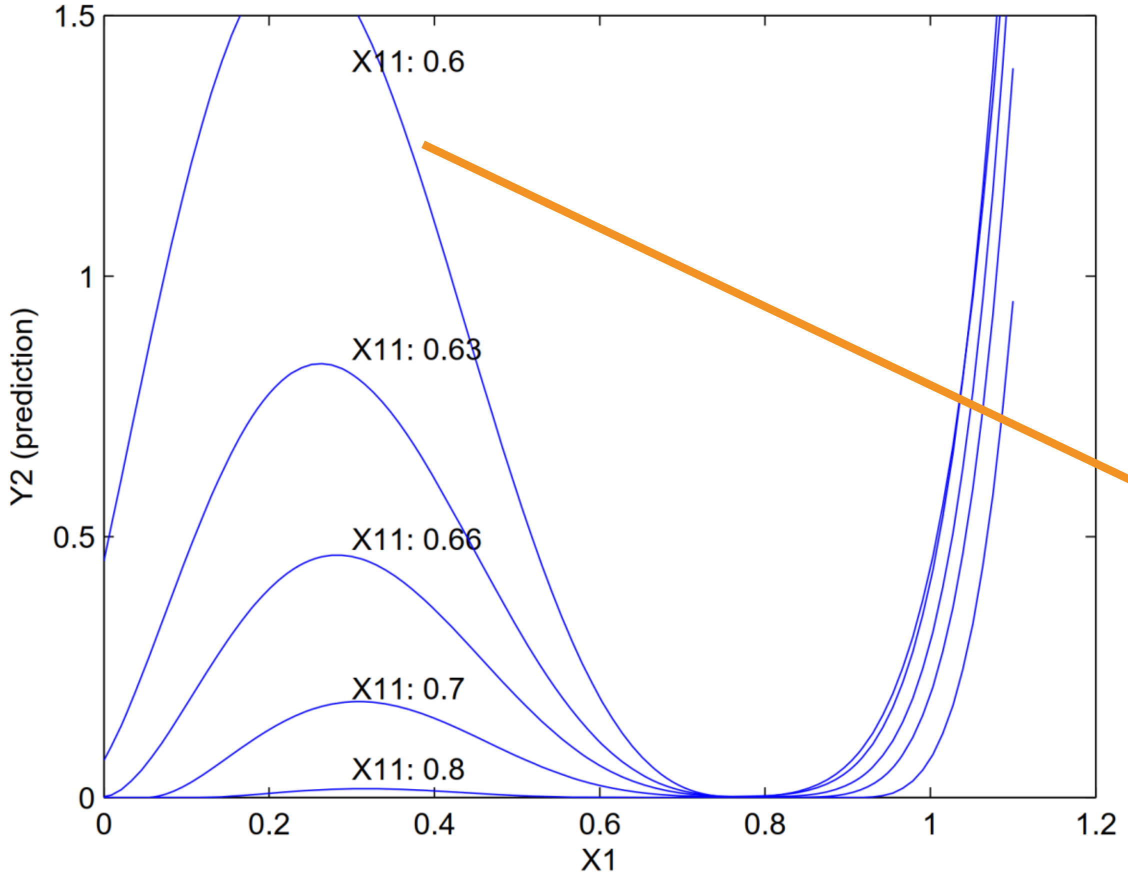


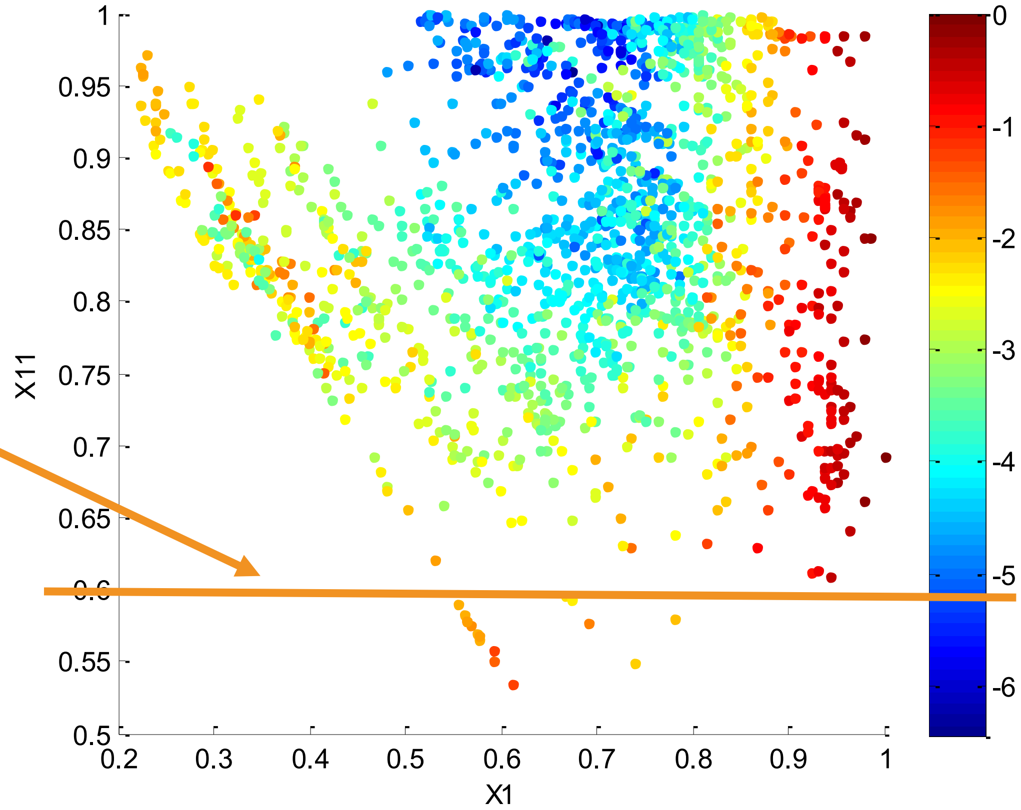
A potential cause for the problem



SymReg

Further investigation





SymReg

Objectives

- Integrate vague knowledge into empirical models.
- The additional effort should be minimal.
- Better or equal model accuracy
- Low computational overhead
- Applicable to many methods for empirical modelling:
  - Polynomial regression, artificial neural networks, kernel regression, symbolic regression, …

The concept of shape-constrained regression

$$f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}}\, \mathcal{L}(f, X, \boldsymbol{y})$$

$\mathcal{L}(f, X,\ y)$ is the loss function (e.g. sum of squared errors)

$\mathcal{F}$ is a model class e.g.:

- polynomials of given degree
- neural network architecture

The concept of shape-constrained regression

$$f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \mathcal{L}(f, X, \boldsymbol{y})$$

$$s.t.:$$

$$\begin{array}{lcccc} l_f & & f(x_f) & & u_f \\ l_{Jac} & \leq & \nabla f(x_{Jac}) & \leq & u_{Jac} \\ l_{Hess} & & \nabla^2 f(x_{Hess}) & & u_{Hess} \end{array}$$

$\mathcal{L}(f, X,\ y)$ is the loss function
(e.g. sum of squared errors)

$\mathcal{F}$ is a model class e.g.:

- polynomials of given degree
- neural network architecture

∇f is the vector of partial derivatives of f over all inputs

f must be differentiable

The concept of shape-constrained regression

$$f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}}\, \mathcal{L}(f, X, \boldsymbol{y})$$

$s.t.$:

$$\begin{array}{ccccc} l_f & & f(x_f) & & u_f \\ l_{Jac} & \leq & \nabla f(x_{Jac}) & \leq & u_{Jac} \\ l_{Hess} & & \nabla^2 f(x_{Hess}) & & u_{Hess} \end{array}$$

$$\forall x_f, x_{Jac}, x_{Hess} \in \mathbb{R}^d,$$

$$\begin{array}{ccccc} l_{x_f} & & x_f & & u_{x_f} \\ l_{x_{Jac}} & \leq & x_{Jac} & \leq & u_{x_{Jac}} \\ l_{x_{Hess}} & & x_{Hess} & & u_{x_{Hess}} \end{array}$$

$\mathcal{L}(f, X,\ y)$ is the loss function (e.g. sum of squared errors)

$\mathcal{F}$ is a model class e.g.:

- polynomials of given degree
- neural network architecture

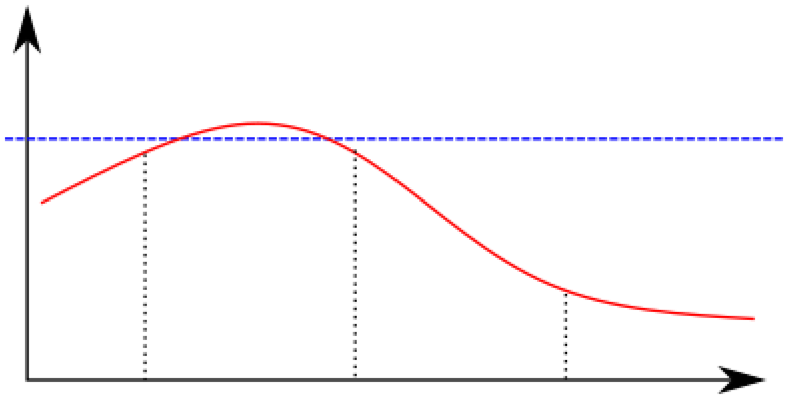
∇f is the vector of partial derivatives of f over all inputs

f must be differentiable

The optimistic and pessimistic approach

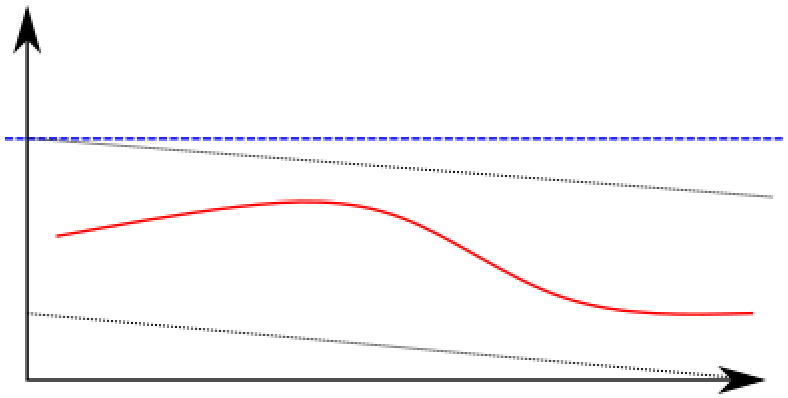
Optimistic

- Evaluate constraint on finite sample from input space
- Easy to implement
- Suffers from curse of dimensionality
- Result can violate the constraints



Pessimistic

- Approximate constraint for the infinite set
- Feasible for certain model classes
- Result conforms to constraints
- Potentially rejects optimal solution



Shape-constrained polynomial regression

A polynomial is positive if it can be expressed as a sum of squares.

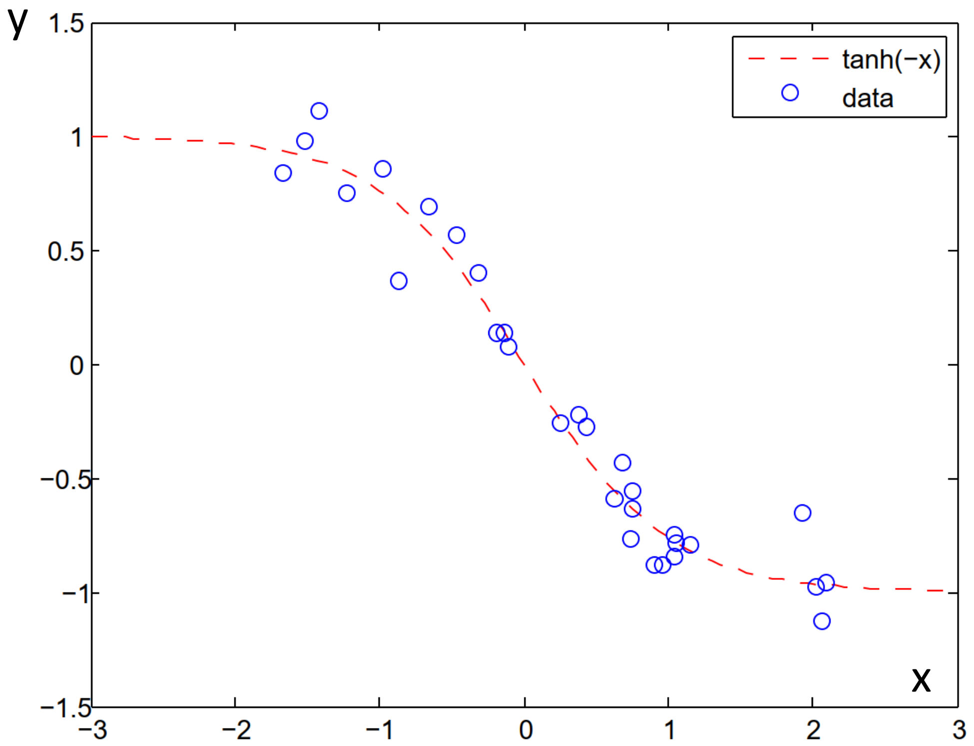
$$P(\boldsymbol{x}, \boldsymbol{\theta}) = T_1(\boldsymbol{x}, \boldsymbol{\theta_1})^2 + \cdots + T_n(\boldsymbol{x}, \boldsymbol{\theta}_n)^2 \Rightarrow P(x) \geq 0$$

Derivatives of polynomials are polynomials.

→ Shape-constrained polynomial regression is a quadratic optimization problem with sum-of-squares constraints.

Shape-constrained polynomial regression

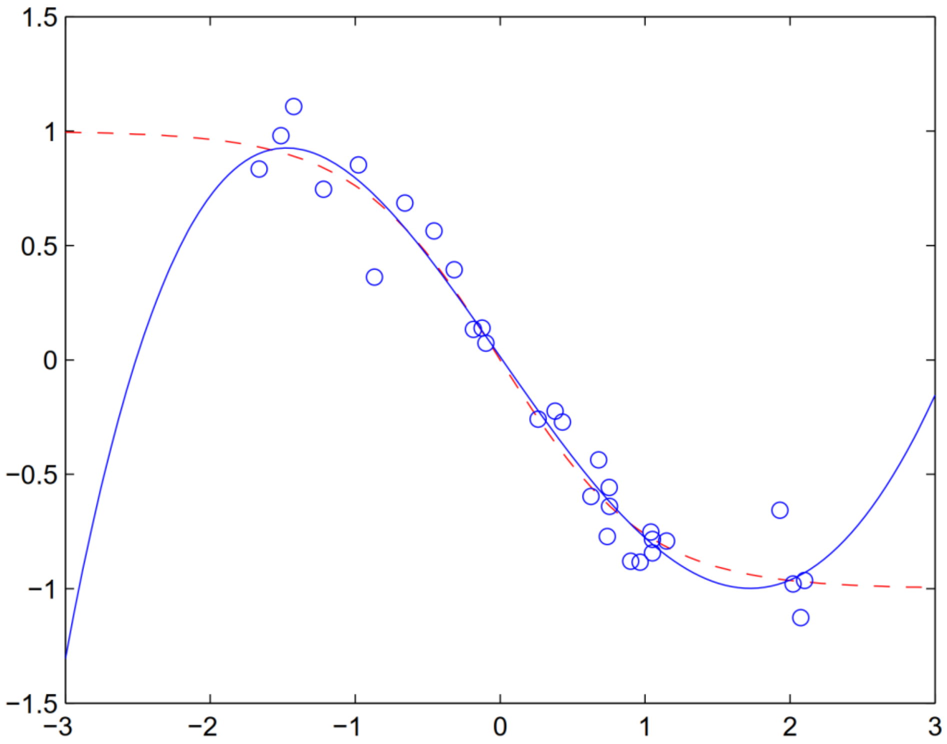
Toy-Problem:



$$y = \tanh(-x) + \epsilon$$
$$x \sim N(0,1)$$
$$\epsilon \sim N(0, 0.15)$$

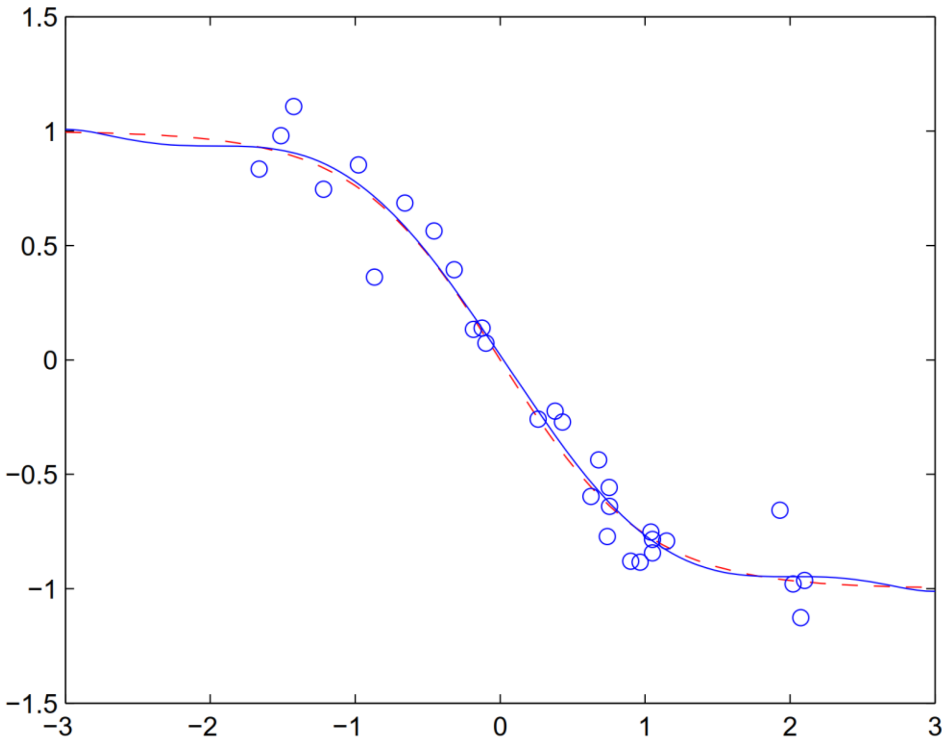
Shape-constrained polynomial regression

Fitting a polynomial with degree 4:



Shape-constrained polynomial regression

Fitting a polynomial with degree 9 and constraints:

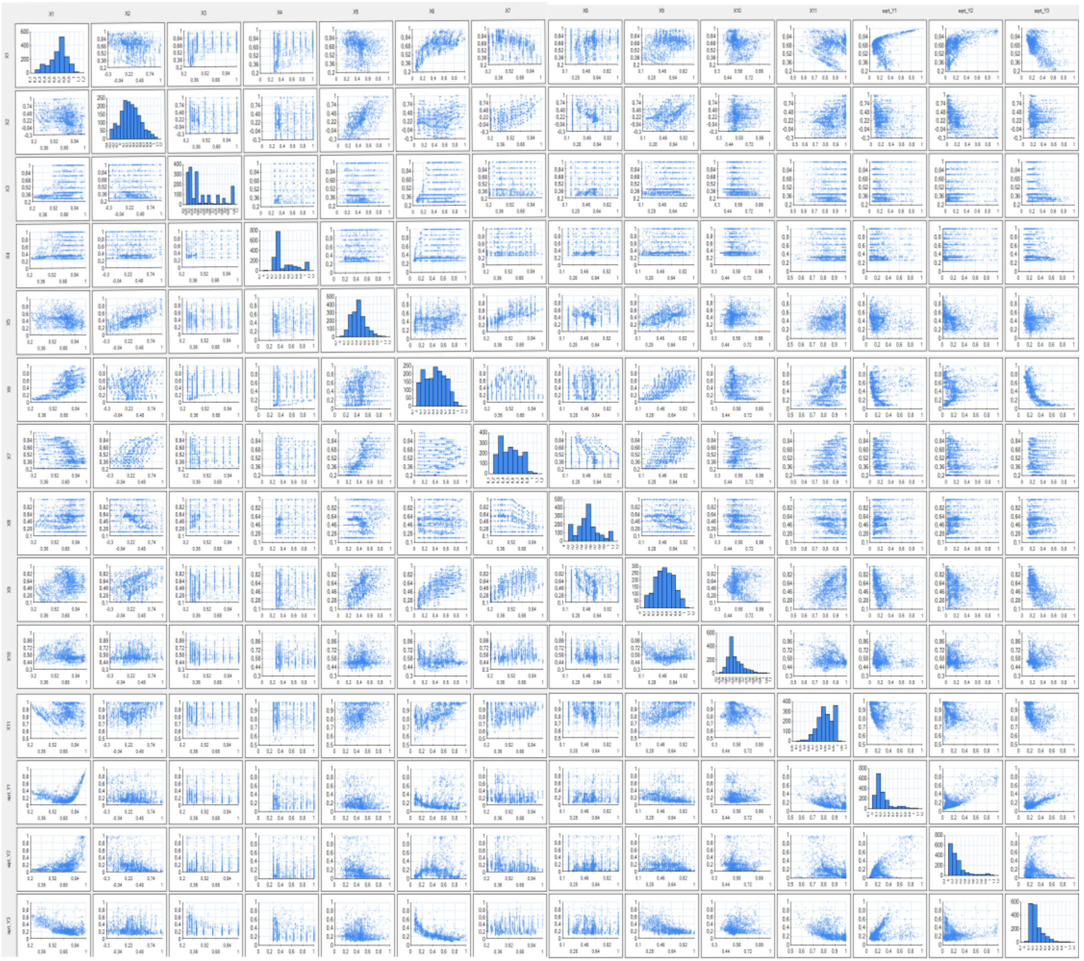


$$\forall_{x \in R} - 3 \leq x \leq 3:$$
$$-1 \leq P(x) \leq 1$$
$$\frac{\partial P(x)}{\partial x} \leq 0$$

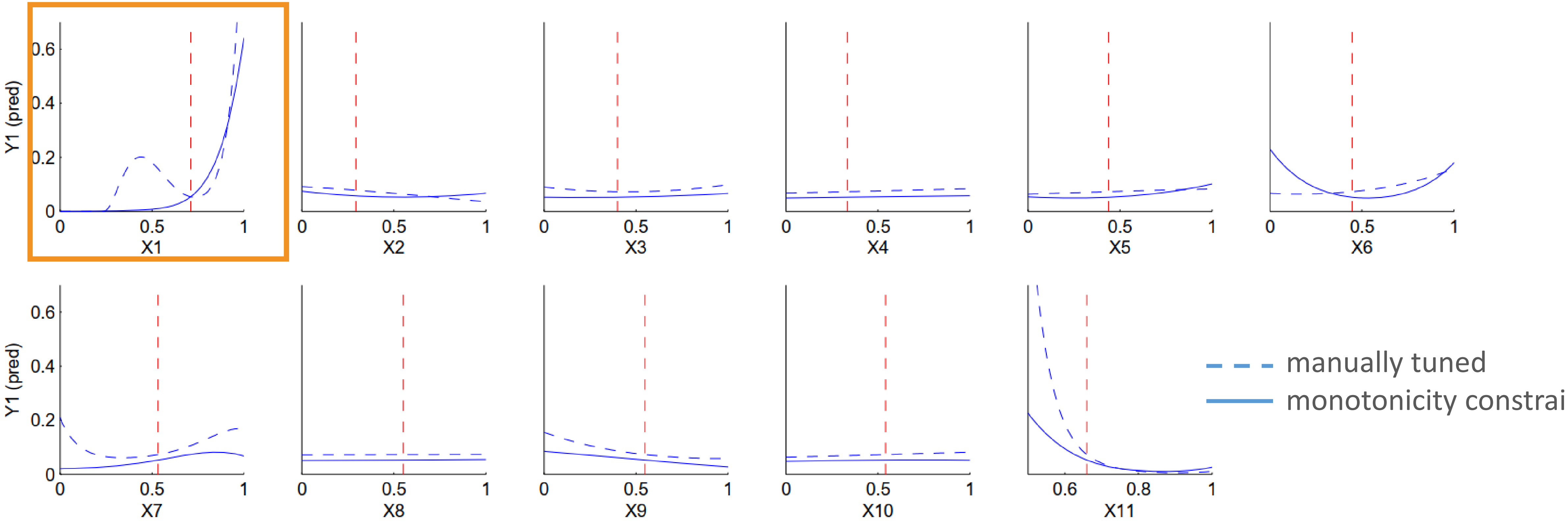
Application to Emission Models

Methodology

1. Gather data from test bench using DOE
   - 11 parameters
   - 3 targets
   - 1500 observations
2. Find and fit polynomial model
   - 3rd degree
   - two-way interactions
   - approx. 50 terms
3. Add monotonicity constraint $\frac{\partial}{\partial X_{11}} P(x) \geq 0$
4. Fit parameters with constraints

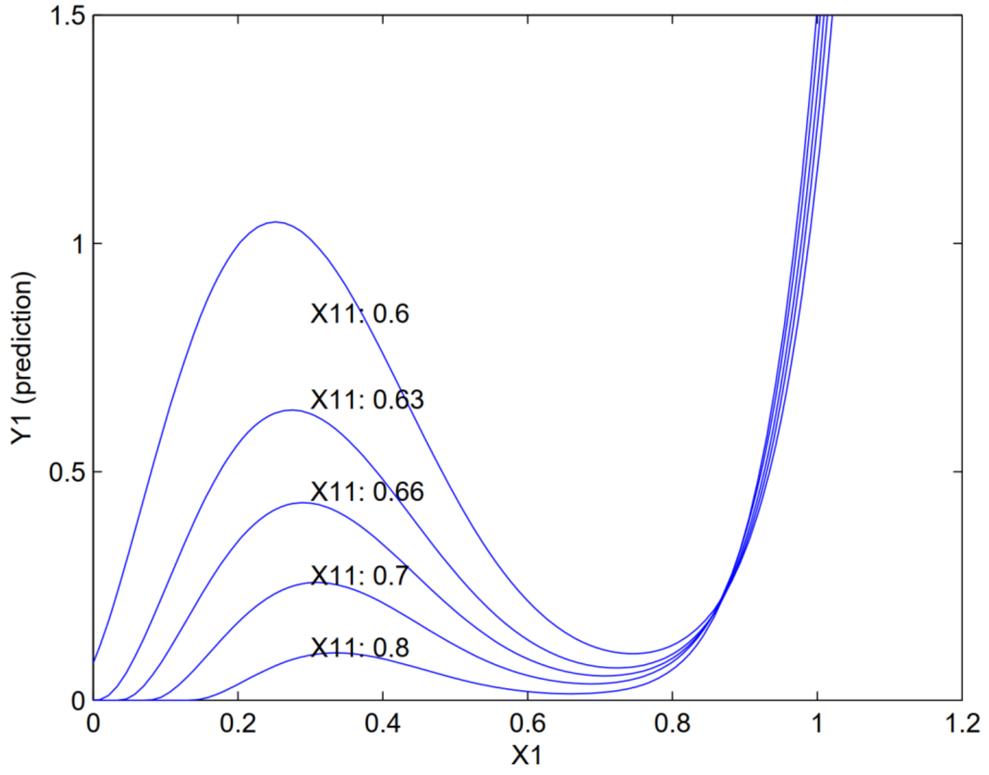


Model dependence plots for target Y1

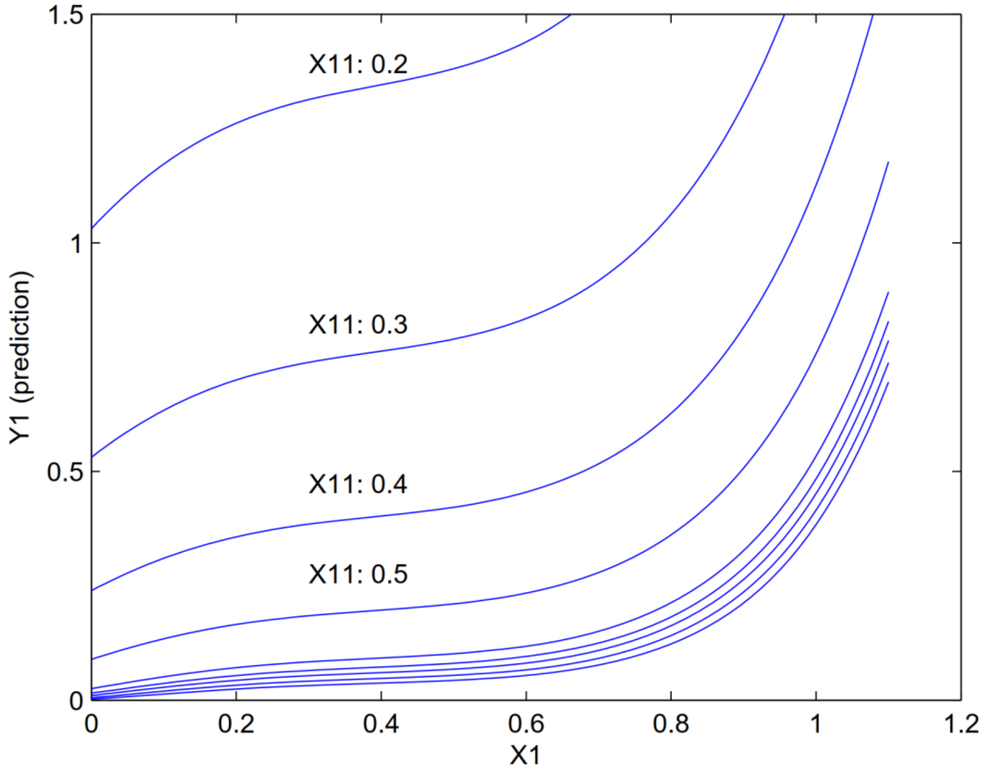


SymReg

Detail of X1 and X11 interaction for Y1

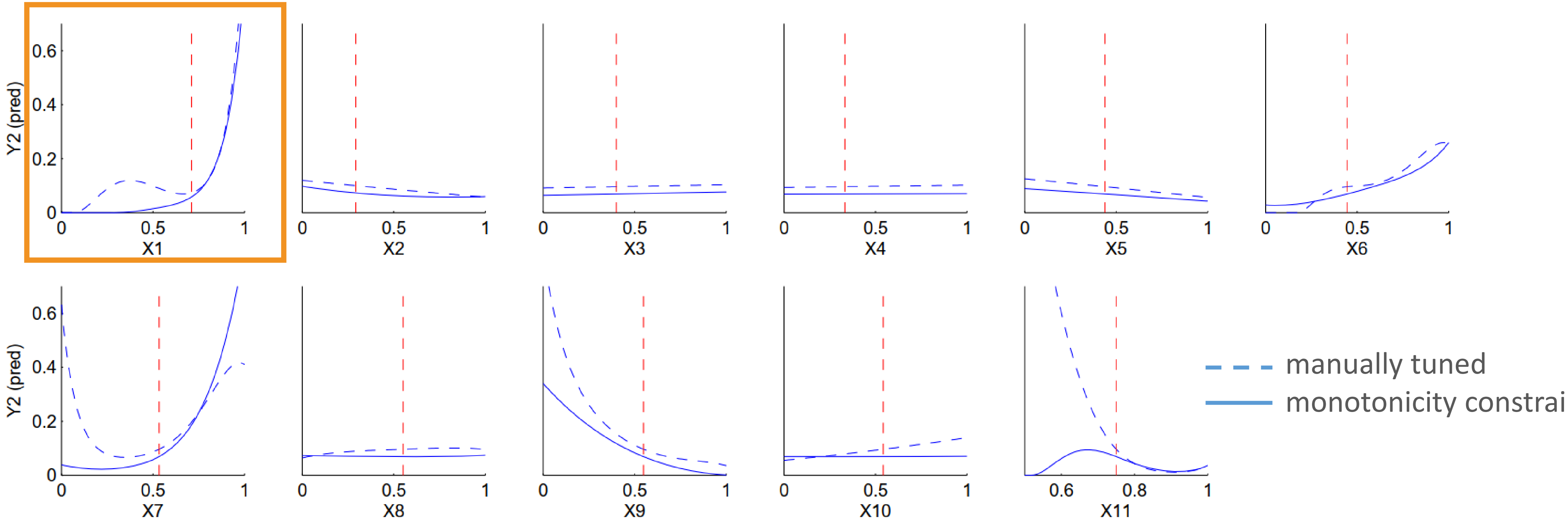


Manually tuned model

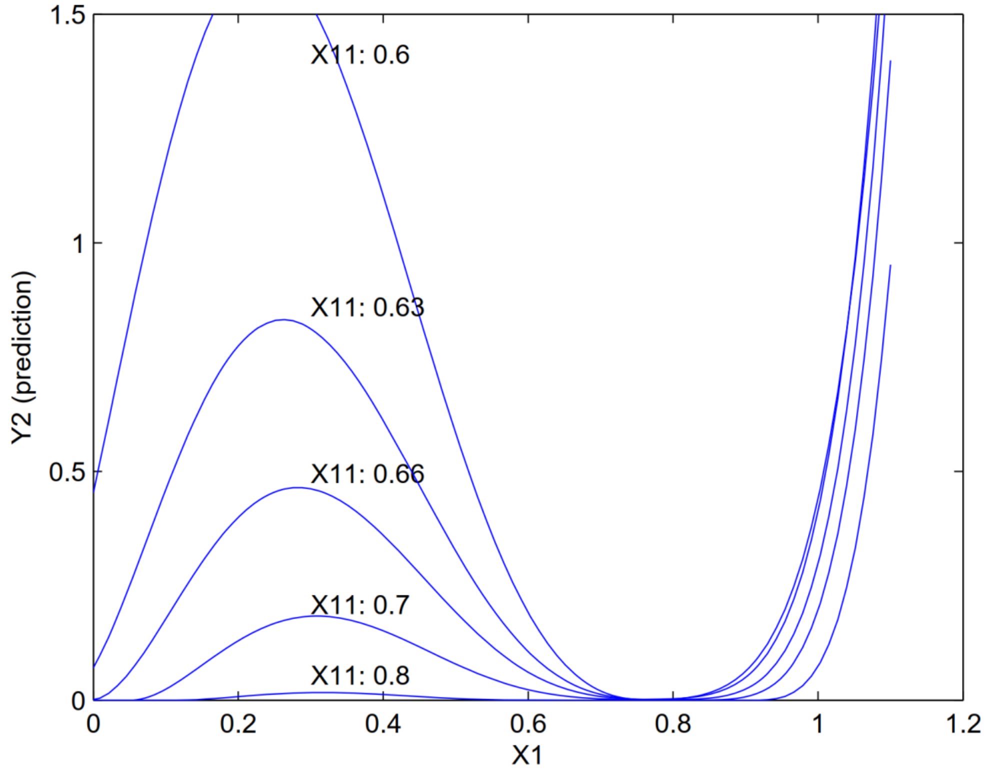


With monotonicity constraint

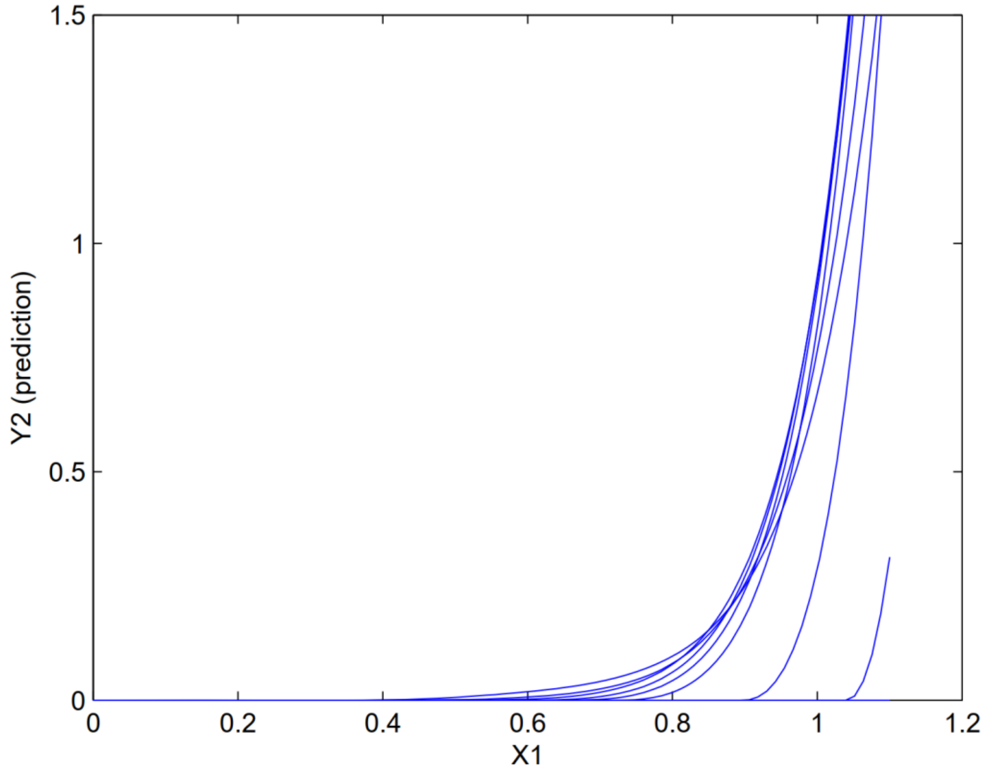
Model dependence plots for target Y2



Detail of X1 and X11 interaction for Y2

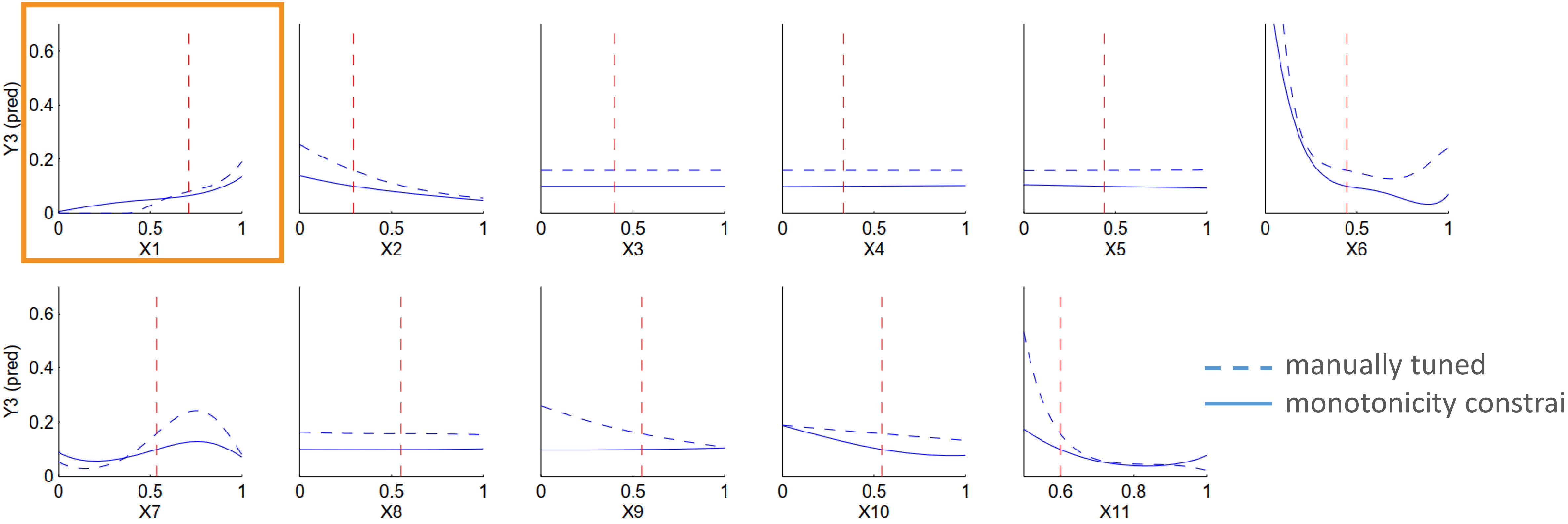


Manually tuned model



With monotonicity constraint

Model dependence plots for target Y3



SymReg

Model accuracy and runtime

Variable	RMSE (no constraints)	RMSE (with constraints)	Run time (with constraints)
Y1	0.0228	0.0405	117 s
Y2	0.0444	0.0475	44.5 s
Y3	0.0509	0.0521	108 s

- Higher training error for shape-constrained regression
- Runtime: 1 – 2 minutes
  - Office PC
  - a single monotonicity constraint
  - ~ 1500 data points

Conclusions

- Shape constraints can be used to improve polynomial regression models.
- Successful for monotonicity constraints.
- Concavity / convexity can be enforced similarly.
- Runtime depends mainly on the number of constraints.

Limitations:

- Works only for polynomial models.
- Improved extrapolation has not been quantified

SymReg

JOSEF RESSEL CENTER FOR SYMBOLIC REGRESSION

https://symreg.at

Integration of Physical Knowledge in Empirical Models – A New Approach to Regression Analysis

G. Kronberger, C. Haider, M. Kommenda, (University of Applied Sciences Upper Austria)
S. Scheidel (AVL List GmbH), M. Kordon (AVL Commercial Driveline & Tractor Engineering GmbH)