ELSYS Note

Go with the flow

This ELSYS note investigates different methodologies for modeling a pump system. The considered approaches include a physics-based model, two data-driven models, and a hybrid model. All models are trained and evaluated on measured operating data with respect to predictive accuracy and generalization capability under different operating conditions. The results highlight the strengths and limitations of each modeling strategy and illustrate the trade-offs between physical interpretability and data-driven accuracy.

Context

Pumping systems account for a considerable share of industrial electricity demand and therefore offer significant potential for improving energy efficiency. Optimizing their operation is, however, challenging, as experiments on real facilities are often not feasible due to safety, availability, or economic constraints.
To enable the analysis and optimization of operating strategies, digital representations of real systems in the form of models are required. Such models allow investigations to be carried out in a virtual environment without interfering with the physical system.
Models can be developed either from operating data or based on physical principles. Both approaches have specific advantages and limitations with respect to accuracy, interpretability, and generalization. In this work, physics-based, data-driven, and hybrid modeling approaches are investigated and compared using a real pump system.

System and task

The considered pump system consists of two vertically arranged tanks connected by piping and a centrifugal pump. A bypass line with ball valves allows different operating conditions to be investigated. The pump is driven by a variable-speed induction motor. Water is used as the working fluid.
The models are developed for steady-state operation in a speed range from 350 to 2900 rpm. Input parameters are the pump speed, the fill level in the lower tank, the valve positions, and the fluid temperature. The outputs are the flow rate, the shaft torque, and the pressures at the pump inlet and outlet. A hydraulic diagram of the system is shown in Fig. 1.

Compared models

Different modeling approaches are investigated to describe the pump system behavior. These include a physics-based model, a hybrid model, and two data-driven models. The approaches differ in their underlying assumptions, required data, and generalization capabilities.

Physics-based model

A bottom-up model is formulated from the pressure–flow relations of the hydraulic components. The pump characteristic is extended over the full speed range using affinity laws [1] and interpolation. The resulting nonlinear system is solved numerically to obtain flow, pressures, and torque.

Hybrid model

The physics-based model is extended by correction terms that compensate model errors. These terms are initialized with zero and identified from measurement data using gradient-based optimization.

Data-driven model

Two purely data-driven approaches are considered: a decision tree and a multilayer perceptron. Both models are trained on measured operating data. To ensure robust training, cross-validation and hyperparameter tuning are applied.

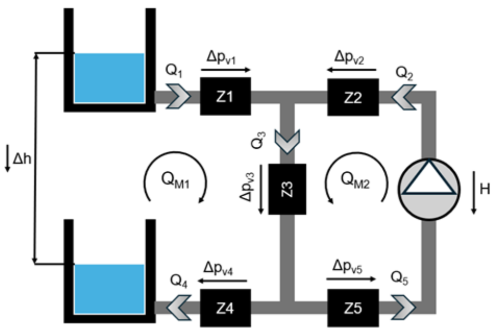

Fig. 1: Hydraulic diagram

 ✉ Author: F. Augenstein

Output / operating condition		Physics-based	Hybrid	Decision tree	Neural network
Flow	Non-cavitating	4.2	3.9	2.6	2.4
	Non-cavitating, unseen speed range	3.9	3.6	11.4	2.5
	Cavitating	40.6	27.3	1.1	0.7
Torque	Non-cavitating	3.0	1.8	1.0	1.1
	Non-cavitating, unseen speed range	2.8	1.6	9.2	3.9
	Cavitating	6.9	4.5	0.8	2.9
Suction pressure	Non-cavitating	4.1	2.8	1.8	1.7
	Non-cavitating, unseen speed range	2.7	2.5	7.7	2.5
	Cavitating	35.7	28.4	1.5	1.5
Discharge pressure	Non-cavitating	1.5	1.4	0.6	0.5
	Non-cavitating, unseen speed range	3.1	3.1	4.5	5.6
	Cavitating	4.2	3.8	0.7	1.5
			good compromise	Data-driven models	

RMSE in %: low medium high

Fig. 2: RMSE comparison of the different models and operating conditions. Lower values indicate better predictive accuracy.

Data and evaluation

To ensure a fair comparison of the different modeling approaches, all models are trained and evaluated on the same dataset. The evaluation focuses on predictive accuracy under different operating conditions as well as on the generalization capability of the models.

Data basis. A dataset was recorded at different rotational speeds and valve positions. For training the models, a preprocessing step is applied. Stationary operating points are selected, duplicates are removed, and all features are normalized to the interval $[0, 1]$. This results in a heterogeneous dataset with reduced bias.

Test cases. The models are evaluated on:

- non-cavitating operating points to assess general accuracy,
- cavitating operating points to evaluate prediction quality under special effects, and
- non-cavitating operating points in an unseen speed range to analyze generalization capability.

Performance metric. Predictive accuracy is assessed using the root mean square error (**RMSE**) for each output variable. The RMSE provides a quantitative measure of the deviation between model predictions and measured values.

Key findings

- **Physics-based model:** Moderate accuracy for non-cavitating conditions, but poor performance under cavitation since this effect is not included in the model.
- **Hybrid model:** Improves the physics-based predictions, particularly when cavitating operating points are included in the training data.
- **Decision tree:** High accuracy within the training domain, but poor generalization to unseen speed ranges.
- **Neural network:** Good overall accuracy and better generalization than the decision tree, although performance also deteriorates outside the training domain (cf. Fig. 2).

Take-away

Physics-based models are interpretable and require little data, but they rely on system knowledge and may miss relevant effects.
Data-driven models can achieve high accuracy, but their performance strongly depends on the coverage of the training data.
Hybrid models combine physical structure with data-driven correction and therefore offer a promising compromise.

References

[1] Johann Friedrich Gülich. *Kreiselpumpen. Handbuch für Entwicklung, Anlagenplanung und Betrieb.* Springer Vieweg Berlin, Heidelberg, 2020. DOI: 10.1007/978-3-662-59785-9.