Click on the tab to open the drop-down menu

Methodology: Thermophysical Properties (experiment &modeling) Oral

---

Predicting thermophysical properties with uncertainty quantification – probabilistic inference

Martin Hangaard Hansen, Bjørn Maribo-Mogensen

Hafnium Labs, Copenhagen, Denmark

www.hafniumlabs.com

Corresponding author e-mail: mh@hafniumlabs.com

Introduction

Large and costly uncertainties in process design and plant construction often result from small uncertainties in thermophysical properties. As an example an *assumed* deviation of +/-5% in the value of the infinite dilution coefficient of could propagate to require a 15 % increase in reboiler duty when designing a two column distillation system for the acetone-chloroform-benzene mixture, showed by P. M. Mathias [1]. Accurate confidence intervals are therefore highly valuable, but not trivial to obtain or to propagate from thermophysical data, which is usually experimental and therefore inherently noisy.

Probabilistic inference methods learn from noisy data and infer with uncertainty. They are built on Bayesian statistical theory and their roots go back many decades. [2] In the past decade, however, probabilistic inference methods have developed and proliferated to many scientific and engineering disciplines at a rapid pace.[3]

Gaussian Processes

A popular probabilistic learning algorithm is the Gaussian process (GP). The modern GP is a kernel machine, which emerged in the 00's. [2] Fitting a GP regressor corresponds to fitting a Gaussian distribution of models to the data, and every prediction is therefore a Gaussian distribution with a mean and a variance, from which confidence intervals on predictions can be obtained. [2] GP's have yet to become widely applied in the thermophysical properties community. They do, however, present excellent behaviour for regression and interpolation of smooth and noisy datasets. Furthermore, they have shown superior performance in probabilistic optimization and data acquisition routines i.e. automated design of experiments. [3, 4] This talk will introduce Gaussian processes and showcase applications to thermophysical properties. Assumptions, advantages and limitations will be discussed, and we will talk about pitfalls and discuss metrics for assessing performance of predictions and uncertainties. Finally, we will discuss an algorithm for outlier detection, based on gaussian process regression.

[1] P. M. Mathias, 2016, *Effect of VLE Uncertainties on the Design of Separation Sequences by Distillation – Study of the Benze-Chloroform-Acetone System*, Fluid Equilibria 408, 265-272.

[2] A. G. Wilson, 2014, *Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes,* Doctoral dissertation, University of Cambridge.

[3] Z. Ghahramani. 2015, *Probabilistic machine learning and artificial intelligence*, Nature 521(7553), 452.

[4] Torres, J. A. G., Jennings, P. C., Hansen, M. H., Boes, J. R., & Bligaard, T., 2019, *Low-Scaling Algorithm for Nudged Elastic Band Calculations Using a Surrogate Machine Learning Model*, Physical review letters *122*(15), 156001.

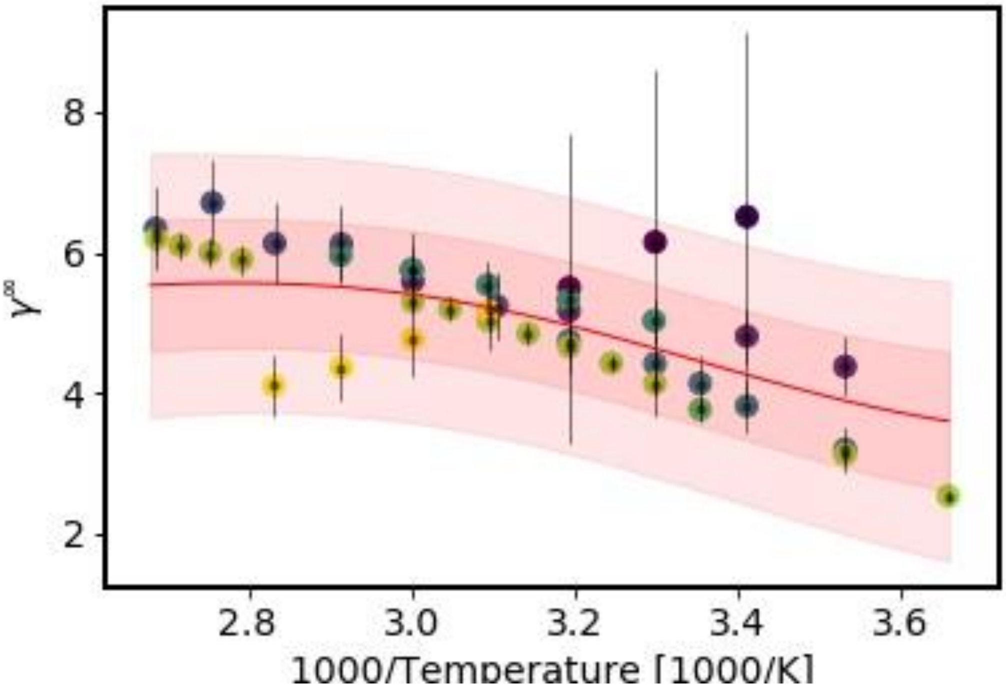

Figure 1 A Gaussian process interpolating (red line) a noisy dataset with a smooth underlying trend. +/- 1 and 2 standard deviations are shown as shaded regions. The color codes on the data points denote the sources of the data.