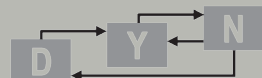


SCHRIFTENREIHE DES
LEHRSTUHLs FÜR
SYSTEMDYNAMIK UND PROZESSFÜHRUNG

Band 1/2022

Corina Nentwich

**Surrogate modeling of phase
equilibrium calculations using
adaptive sampling**



Surrogate modeling of phase equilibrium calculations using adaptive sampling

Zur Erlangung des akademischen Grades eines

Dr.-Ing.

von der Fakultät Bio- und Chemieingenieurwesen
der Technischen Universität Dortmund
genehmigte Dissertation

vorgelegt von

Corina Nentwich, M. Sc.

aus

Recklinghausen

Tag der mündlichen Prüfung: 19.05.2021

1. Gutachter: Prof. Dr. Sebastian Engell
2. Gutachter: PD Dr. Michael Bortz

Dortmund **2022**

Schriftenreihe des Lehrstuhls für
Systemdynamik und Prozessführung
herausgegeben von Prof. Dr.-Ing. Sebastian Engell

Band 1/2022

Corina Nentwich

**Surrogate modeling of phase equilibrium
calculations using adaptive sampling**

D 290 (Diss. Technische Universität Dortmund)

Shaker Verlag
Düren 2022

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Zugl.: Dortmund, Technische Univ., Diss., 2021

Copyright Shaker Verlag 2022

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8440-8486-3

ISSN 1867-9498

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Phone: 0049/2421/99011-0 • Telefax: 0049/2421/99011-9

Internet: www.shaker.de • e-mail: info@shaker.de

Vorwort

Die Ergebnisse dieser Arbeit entstanden während meiner Zeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Systemdynamik und Prozessführung der TU Dortmund unter der Leitung von Professor Engell.

Daher möchte ich zunächst Herrn Professor Engell danken. Ohne Ihr Vertrauen in mich, vor allem in der Anfangszeit, hätte ich diese Arbeit nie geschrieben. Die vielen fachlichen Gespräche haben mir sehr geholfen den schließlich gefundenen Kurs zu halten und einen Abschluss zu finden. Ich danke auch Dr. Michael Bortz für die Begutachtung.

Für die finanzielle Unterstützung und die Möglichkeit, im Rahmen des Sonderforschungsbereichs Transregio 63 viele hervorragende Forscher*innen kennen zu lernen, möchte ich der Deutschen Forschungsgemeinschaft danken. Die Runde der Doktorand*innen hat eine Gemeinschaft geschaffen, in der ich mich sehr wohl gefühlt habe, Danke auch an euch.

Meinen Kolleg*innen am Lehrstuhl für Systemdynamik und Prozessführung möchte ich für die sehr angenehme Arbeitsatmosphäre und die vielen fachlichen und nicht-fachlichen Gespräche danken, besonders Simon Wenzel und Clemens Lindscheid. In meiner Zeit am Lehrstuhl habe ich auch sehr ambitionierte Student*innen betreuen dürfen, die mich tatkräftig unterstützt haben. Insbesondere möchte ich mich an dieser Stelle bei Joschka Winz, Christopher Varela, Anoj Winston Gladius, Shreya Bhatia und Roderich Wallrath bedanken.

Ich möchte mich auch bei meinen Eltern und meinem Bruder bedanken. Der größte Dank gilt hierbei meiner Mutter, Maria Mrukwa, die mich mein gesamtes Studium mit Herz und Seele durch Höhen und Tiefen begleitet hat und mich in schwierigen Zeiten immer mit viel Liebe und Makkaroni aufgefangen hat.

Für seine unendliche Geduld und den besten Rückhalt, den ich mir vorstellen kann, möchte ich mich bei meinem Partner Ingo Schmutzler bedanken.

Prepublished content

This thesis is the result of research activities at the group of Process Dynamics and Operations of the TU Dortmund University. Parts of this thesis were published in journal articles and conference proceedings, applied code implementations and results were partly generated in supervised theses, as declared in the following.

Chapter 2	introduction	partly modified from	[A]
	2.1.1	partly modified from	[A]
	2.1.2	partly modified from	[A]
	2.1.3	partly modified from	[A,B]
Chapter 3	2.2	partly modified from	[A]
	3.1.2	partly modified from	[A]
	3.1.3	partly modified from	[B]
	3.2	partly modified from	[A]
	3.2.1	partly modified from	[A]
	3.2.2	partly modified from	[A]
	3.2.3	partly modified from	[A]
	3.3.1	partly modified from	[B]
Chapter 5	3.3.2	partly modified from	[A,B]
	5.1	partly modified from	[B]
Chapter 6	5.2	partly modified from	[A]
	6.2.1	partly modified from	[A,B]
Chapter 7	6.2.2	partly modified from	[B]
	6.3	partly modified from	[A]
Chapter 8	7.1	partly modified from	[A,B,C]
Chapter 9	8.1.1	partly modified from	[B]
	8.1.1	implementation and results partly from	[a]
	8.1.2	implementation partly from	[a]
	8.2	applied PC-SAFT parameters fitted in	[d]
Chapter 10	8.2	implementation partly from	[a]
	9.1.1	implementation partly from	[b]
	9.1.2	implementation partly from	[b]
	9.2	applied PC-SAFT parameters fitted in	[d]
Chapter 10	9.2	implementation partly from	[b]
	10.2	implementation partly from	[D,c]
	10.3	implementation partly from	[D,c]

Publications

The previously referenced list of journal articles and conference proceedings articles is given in the following.

Journal articles

- [A] Nentwich, C.; Engell, S. Surrogate modeling of phase equilibrium calculations using adaptive sampling. *Comput. Chem. Eng.* **2019**, *126*, 204–217.*
- [B] Nentwich, C.; Winz, J.; Engell, S. Surrogate Modeling of Fugacity Coefficients Using Adaptive Sampling. *Ind. Eng. Chem. Res.* **2019**, *58*, 18703–18716.†

Conference proceedings

- [C] Nentwich, C.; Engell, S. Application of surrogate models for the optimization and design of chemical processes. Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN). Vancouver, BC, Canada, 2016; pp 1291–1296, ©2016 IEEE.
- [D] Nentwich, C.; Varela, C.; Engell, S. Optimization of chemical processes applying surrogate models for phase equilibrium calculations. Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary, 2019; pp 1–8, ©2019 IEEE.

Supervised theses

- [a] Winz, J. *Adaptive sequential sampling for surrogate modeling of fugacity coefficients*, Bachelor's thesis TU Dortmund, 2018.
- [b] Wallrath, R. *Adaptive sequential sampling for surrogate modeling of liquid-liquid phase equilibrium calculations*, Bachelor's thesis TU Dortmund, 2018.
- [c] Varela Barreno, C. O. *Modeling and optimization of a reaction-separation system with surrogate models for the phase equilibrium calculations*, Master's thesis TU Dortmund, 2018.
- [d] Wu, M. *Application of surrogate models for the activity coefficient calculation in a reactive distillation column*, Master's thesis TU Dortmund, 2018.

*This article was published in *Computers & Chemical Engineering*, *126*, Corina Nentwich, Sebastian Engell, Surrogate modeling of phase equilibrium calculations using adaptive sampling, 204–217, Copyright Elsevier (2019).

†Reprinted (adapted) with permission from [B]. Copyright 2019 American Chemical Society.

Kurzfassung

Die Wahl des thermodynamischen Modells für Phasengleichgewichtsberechnungen spielt im Rahmen der computergestützten Prozesssimulation eine zentrale Rolle. Für stark nichtideale Systeme werden insbesondere Zustandsgleichungen wie das PC-SAFT-Modell (Perturbed-Chain Statistical Associating Fluid Theory) zur genauen Beschreibung von Phasengleichgewichten im Rahmen einer Prozesssimulation verwendet. Um solche komplexen Modelle in eine Prozesssimulation zu implementieren, wird das thermodynamische Modell entweder direkt in die Prozesssimulationssoftware implementiert oder in eine externe Bibliothek ausgelagert, die vom Prozesssimulator aufgerufen wird. Der Einsatz von PC-SAFT-Modellen kann aufgrund ihres iterativen Charakters für beide Optionen zu einem hohen Rechenaufwand für die Simulation und Optimierung chemischer Prozesse führen.

Diese Arbeit schlägt vor, Surrogatmodellierung – Ersatz eines komplexen Modells durch ein schneller auswertbares Black-Box-Modell – zu verwenden, um den Rechenaufwand für Phasengleichgewichtsberechnungen zu verringern. Es werden zwei Methoden vorgeschlagen: Die indirekte Methode wendet Surrogatmodelle innerhalb einer expliziten Formulierung des Phasengleichgewichtsproblems an, sodass die Phasenberechnungen während der Prozesssimulation effizient gelöst werden, während die direkte Methode Surrogatmodelle anwendet, um direkt Phasenzusammensetzungen zu berechnen. Vor dem Trainieren der Surrogatmodelle werden Trainingsdaten durch das Originalmodell berechnet.

Um den Rechenaufwand für dieses Samplingverfahren so gering wie möglich zu halten, wird ein adaptives Samplingverfahren vorgeschlagen, das Sampling und Training der Surrogatmodelle iterativ kombiniert. Diese Methode ermöglicht bei gleicher Anzahl an Trainingspunkten wesentlich bessere Surrogatmodelle im Vergleich zu einer herkömmlichen raumfüllenden Samplingmethode, was für verschiedene Surrogatmodeltypen gezeigt wird.

Das Training und die Anwendung der Surrogatmodelle werden für verschiedene Systeme demonstriert: Ein ternäres und ein quaternäres Flüssig-Flüssig-Gleichgewichtssystem, ein senäres Gas-Flüssigkeits-Gleichgewichtssystem und ein quinäres Dampf-Flüssigkeits-Gleichgewichtssystem. Die trainierten Surrogatmodelle werden zur Simulation und Optimierung des Prozesses zur Hydroformylierung von 1-Dodecen angewendet und die Ergebnisse mit Simulationen unter Verwendung des Originalmodells PC-SAFT verglichen.

Abstract

The choice of thermodynamic models for phase equilibrium calculations plays a central role in the context of process simulation. For highly non-ideal systems, equations of state as the Perturbed-Chain Statistical Associating Fluid Theory (PC-SAFT) model are preferred for an accurate description of phase equilibria in a chemical process simulation. In order to implement such complex models, the thermodynamic model either is directly implemented into the process simulation software or it is outsourced into an external library, which is called by the process simulator. For PC-SAFT models, due to the iterative solution process, both options can lead to a high computational effort for the process simulation and optimization.

This thesis suggests to use surrogate modeling – replacing a complex model by a more simple black box model – in order to reduce the computational effort for complex phase equilibrium calculations. Two methods are proposed: the indirect method applies surrogate models within explicit formulation of the phase equilibrium to efficiently solve the phase calculations during process simulation, while the direct method applies surrogate models to directly determine phase compositions. Prior to training the surrogate models, samples using original model calculations are drawn.

In order to reduce the computational effort for sampling, an adaptive sampling method is proposed, which combines sampling and training of the models in an iterative manner. This method provides superior surrogate models for the same number of samples compared to a conventional space-filling sampling design, which is shown for different surrogate model types.

The training and the use of the surrogate models are demonstrated for different systems: a ternary and a quaternary liquid-liquid equilibrium, a senary gas-liquid equilibrium, and a quinary vapor-liquid equilibrium. The surrogate models are applied to the simulation and optimization of the hydroformylation process of 1-dodecene and the results are compared to simulations using the original PC-SAFT model.

Contents

1. Introduction	1
I. Theoretical foundations	5
2. Phase equilibrium modeling	7
2.1. Modeling of the fugacity	8
2.1.1. Modeling the fugacity via the fugacity coefficient	8
2.1.2. Modeling the fugacity via the activity coefficient	10
2.1.3. PC-SAFT	12
2.2. Two-phase flash calculations	14
3. Surrogate modeling	17
3.1. Surrogate models	17
3.1.1. Linear regression models	18
3.1.2. Kriging	18
3.1.3. Artificial Neural Networks (ANN)	20
3.1.4. Support Vector Machines (SVM)	23
3.2. Sampling	25
3.2.1. One-shot designs	26
3.2.2. Sequential designs	28
3.2.3. Mixed adaptive sequential design	29
3.3. Validation	30
3.3.1. Assessment of the surrogate model quality	30
3.3.2. Sample distribution metric	32
4. Literature review	33
4.1. Surrogate modeling in chemical engineering	33
4.2. Surrogate modeling for thermodynamic computations	35

II. Surrogate modeling of phase equilibria	39
5. Surrogate modeling methods for phase equilibria	41
5.1. Indirect surrogate modeling method	41
5.2. Direct surrogate modeling method	41
6. Adaptive sampling	45
6.1. Exploration	45
6.2. Exploitation	45
6.2.1. Prediction variance	46
6.2.2. Candidate selection	47
6.3. Combination of exploration and exploitation	49
III. Case studies and results	51
7. Case studies	53
7.1. Hydroformylation of 1-dodecene	53
7.2. Transesterification of dimethyl carbonate	55
8. Surrogate modeling using the indirect method	57
8.1. Hydroformylation of 1-dodecene	57
8.1.1. LLE	57
8.1.2. GLE	75
8.2. Transesterification of dimethyl carbonate (VLE)	78
8.3. Summary and discussion	85
9. Surrogate modeling using the direct method	89
9.1. Hydroformylation of 1-dodecene	89
9.1.1. LLE	89
9.1.2. GLE	101
9.2. Transesterification of dimethyl carbonate (VLE)	104
9.3. Summary and discussion	108
IV. Application of surrogate models in simulation and optimization	111
10. Simulation and optimization of the hydroformylation process	113
10.1. Model quality comparison for the two surrogate modeling methods	113
10.2. Application of the surrogate models in process simulation	118
10.3. Application of the surrogate models in process optimization	122

10.4. Summary and discussion	127
11. Conclusion and outlook	129
11.1. Conclusion	129
11.2. Outlook	131
V. Appendix	133
A. Model parameters and sampling ranges	135
B. Other results for the indirect method	143
C. Other results for the direct method	147
Bibliography	158