UNIVERSITY OF HOHENHEIM

Improving Food Processes using Intelligent Systems

Design, development, and validation of hybrid digital twins for spray drying processes using explainable artificial intelligence

CSH Young Researcher Seminar

Stuttgart 05.06.2025

Florian Kaltenecker, M.Sc.

University of Hohenheim – Department of Food Informatics (150L)

Spray Drying

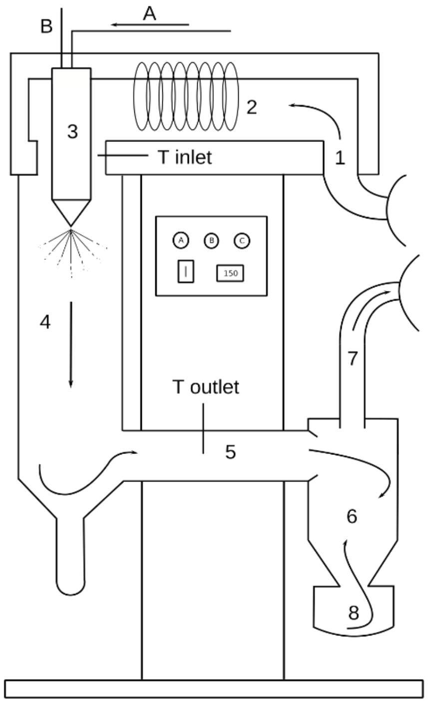
- = rapid method of drying by injecting thinly dispersed droplets into hot gas.
- Used for sensitive materials
  - Chemicals
  - Pharmaceuticals
  - Food Products → Milk Powder, Instant coffee, etc…

Operation Principle

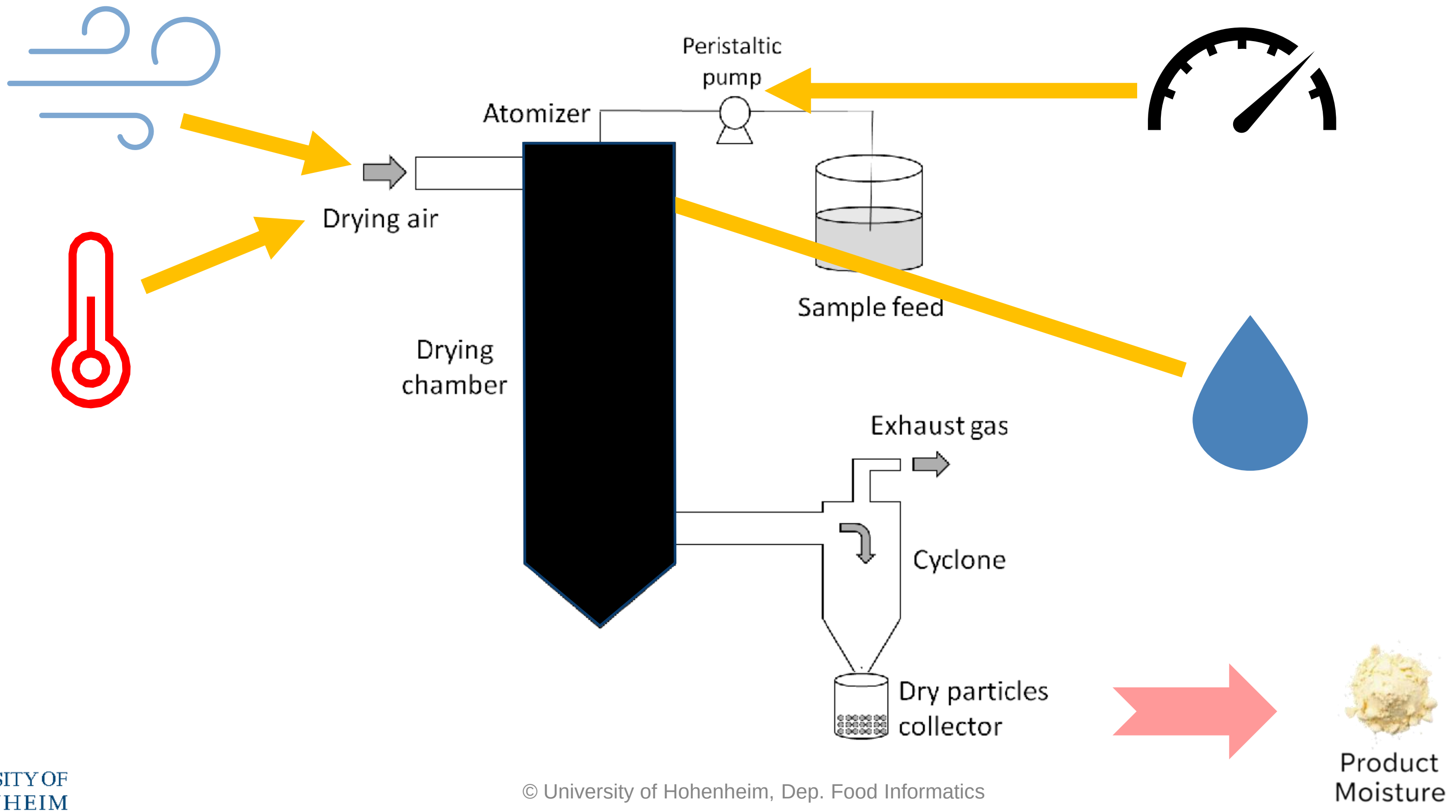
A) Product Feed

B) Atomization gas

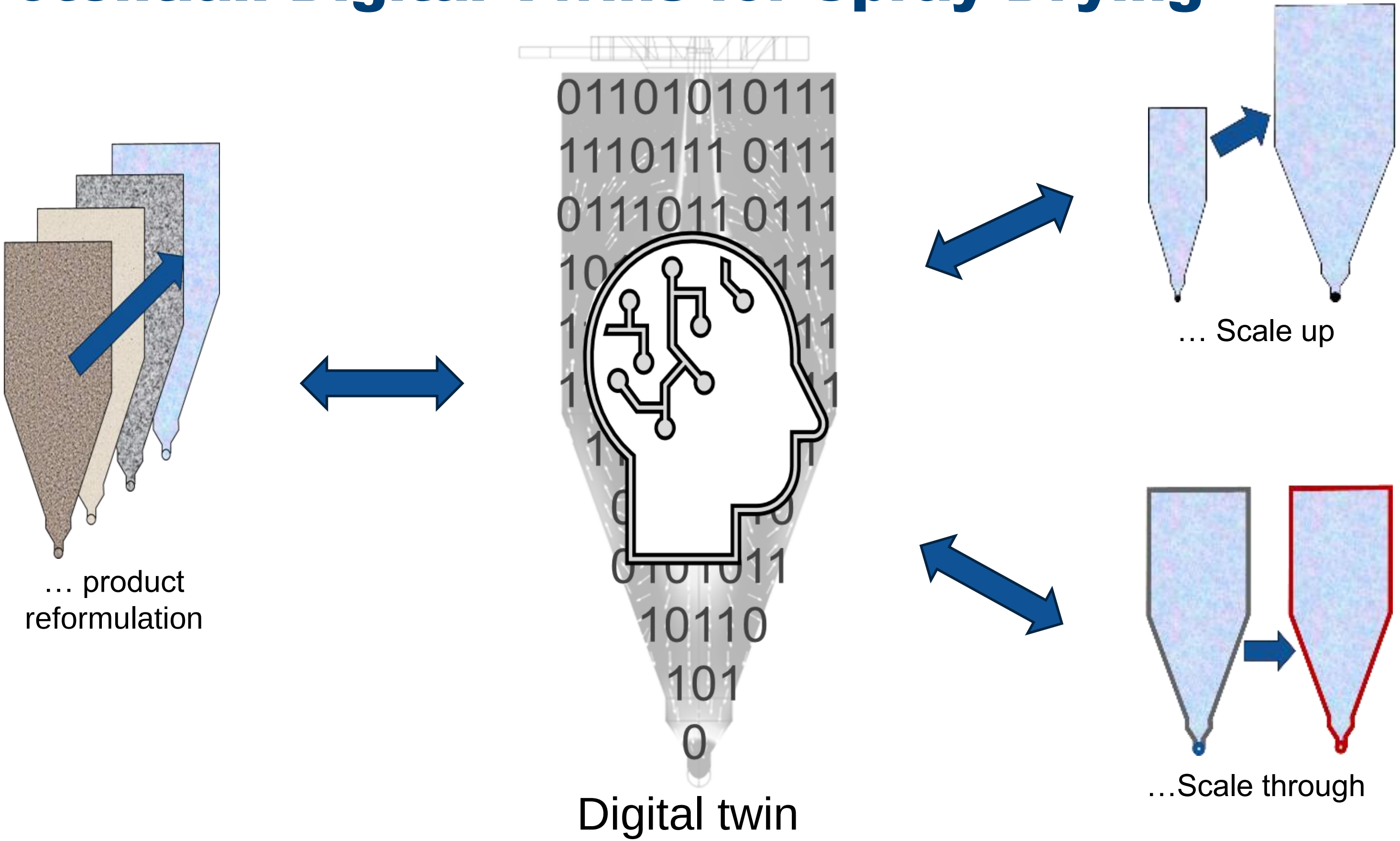
1. Drying gas inlet
2. Heater
3. Nozzle
4. Drying chamber
5. Outlet
6. Cyclone separator
7. Air outlet
8. Product



Controllable Parameters in Spray Drying



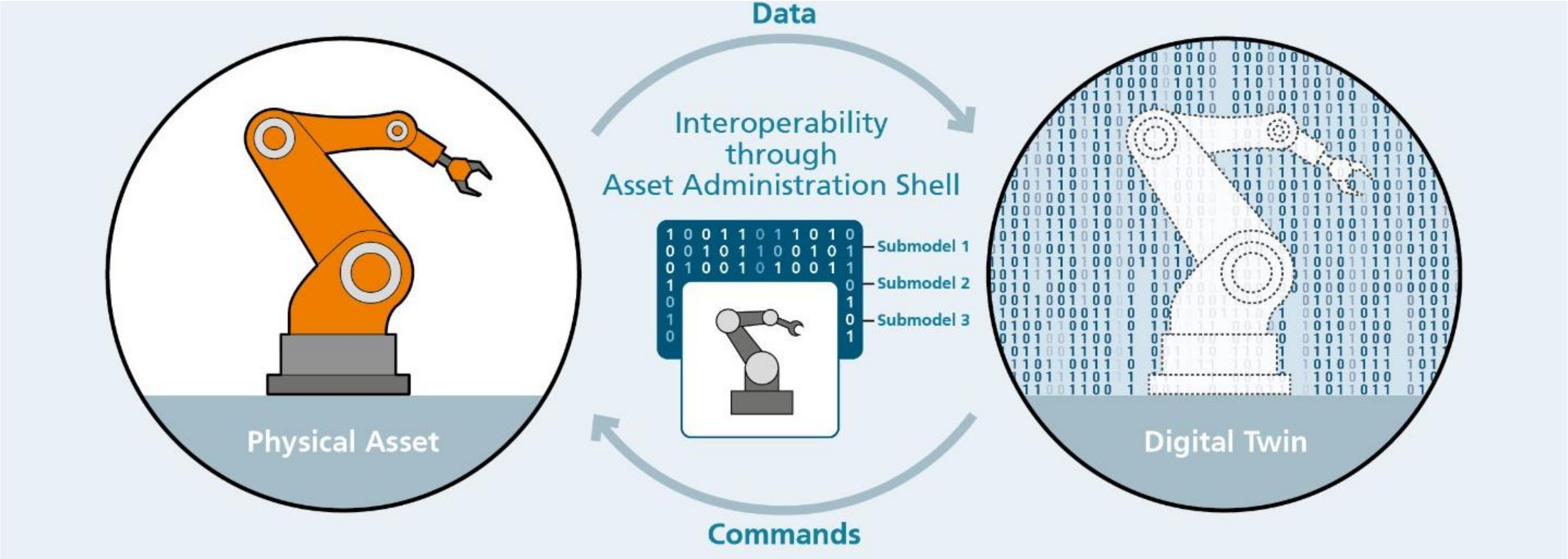
Potential: Digital Twins for Spray Drying



Definition of a Digital Twin

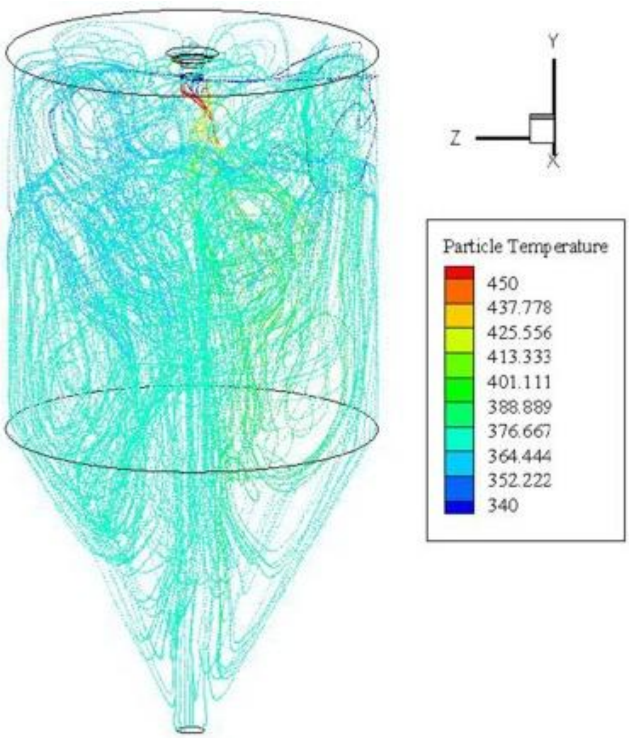
A **digital twin** is a digital representation of an unique product that comprises its selected *characteristics*, *properties*, *conditions*, and *behaviors* by means of models, information, and data within a single or even across multiple life cycle phases.

[Henrichs et al., 2021], adapted from [Stark & Damerau, 2019]



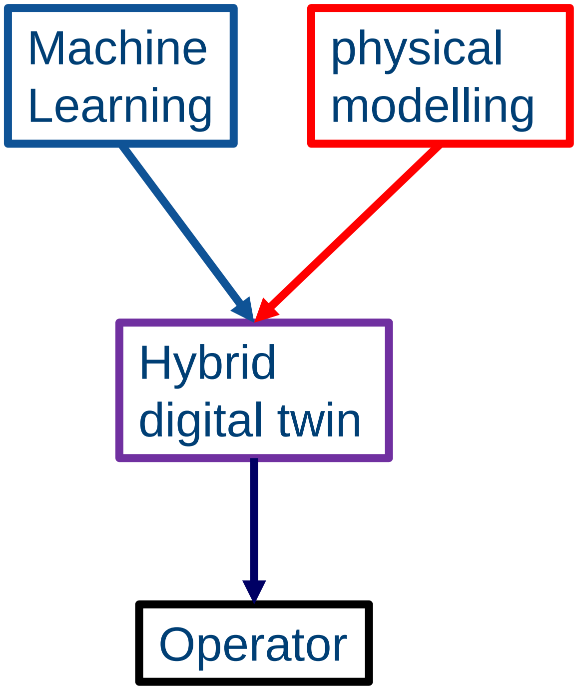
Physical Modeling for Spray Drying

- Physical models based on partial differential equations (PDE)
- Validity of the model depends on the correctness of assumptions
    - E.g. Idealized droplets, boundary conditions
- Current models are slow and computationally expensive



Project Goal

- A hybrid digital twin consisting of physical modeling and machine learning (ML) models.
- This model should …
  - Reflect the system behavior of the spray dryer
  - Gives the operator recommended actions
  - Be transferable between different systems
  - Not operate as a black box

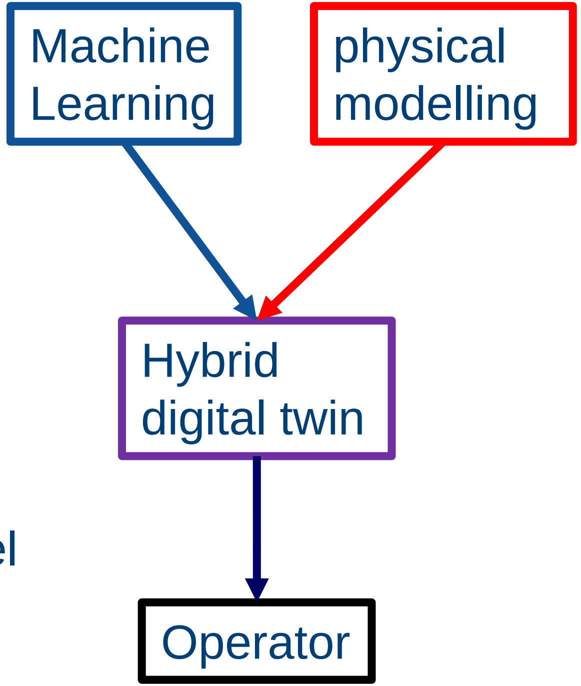


Research Questions

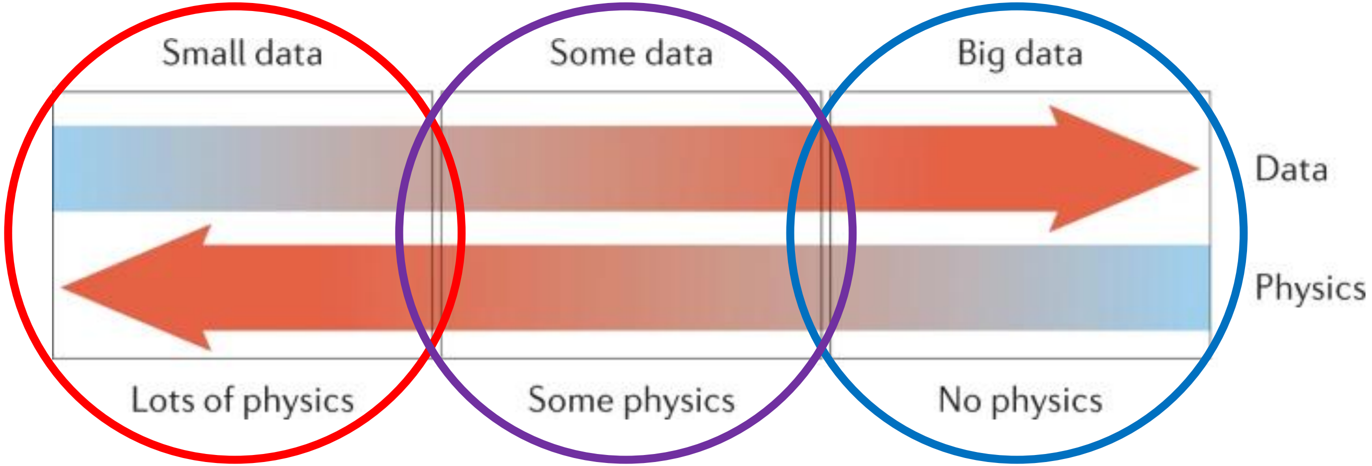
1. **How can machine learning and physics-based modelling be combined to simulate spray drying processes?**

Physics or ML First?

- A hybrid digital twin consisting of physical modeling and machine learning models
- 2 Options:
  - Integrating physics into the ML algorithms predicting the product properties
  - Using ML to parameterize the physical model



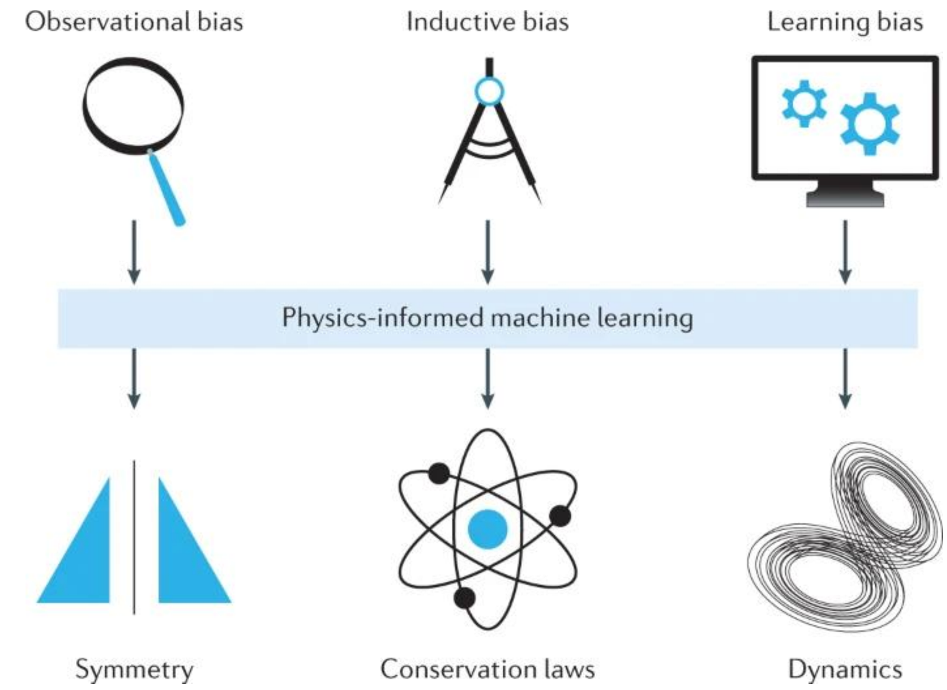
Data Requirements & Physics Informed ML



Physical modeling Physics-informed Modelling Pure ML approach

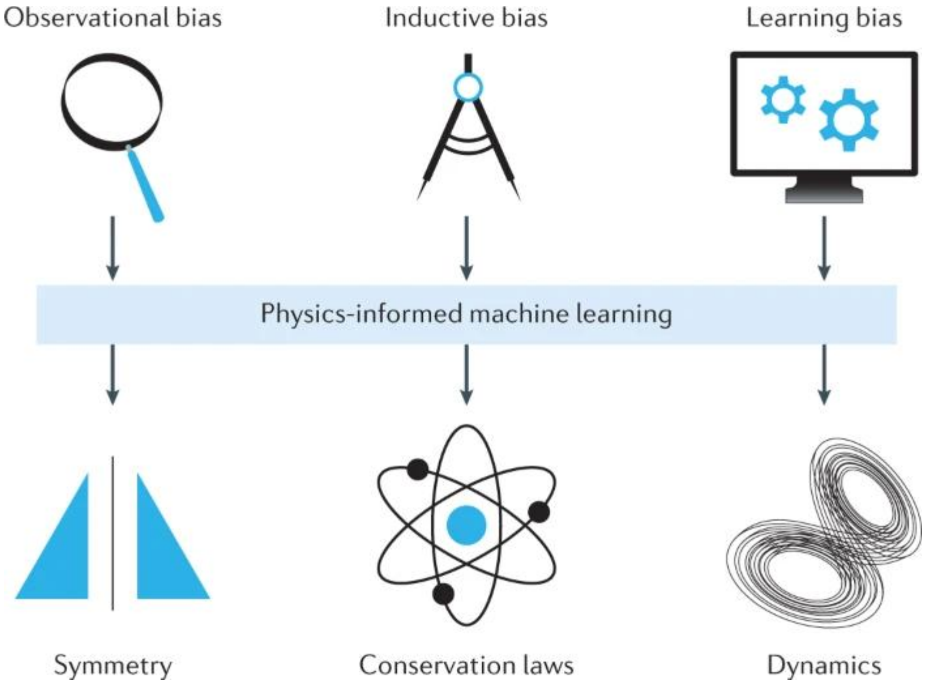
Principles of Physics-Informed Machine Learning

- Observational biases
- Inductive Biases
- Learning biases



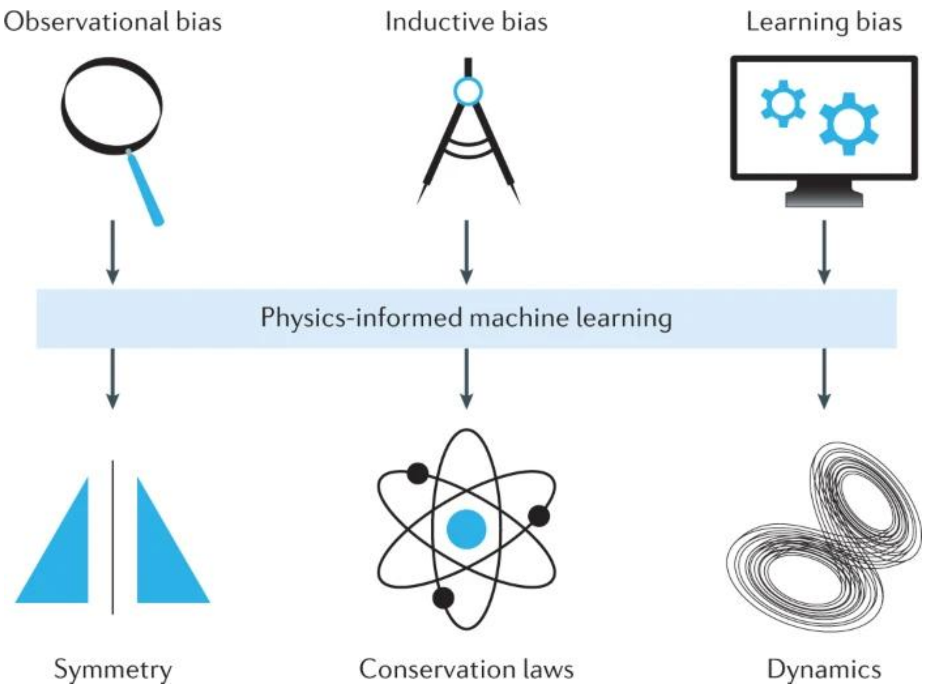
Principles of Physics-Informed Machine Learning

- Observational biases
  - Data embodies the underlying physics
  - Data must cover the input domain of a learning task
  - Data requirements increase with the number of parameters
- Inductive Biases
- Learning biases



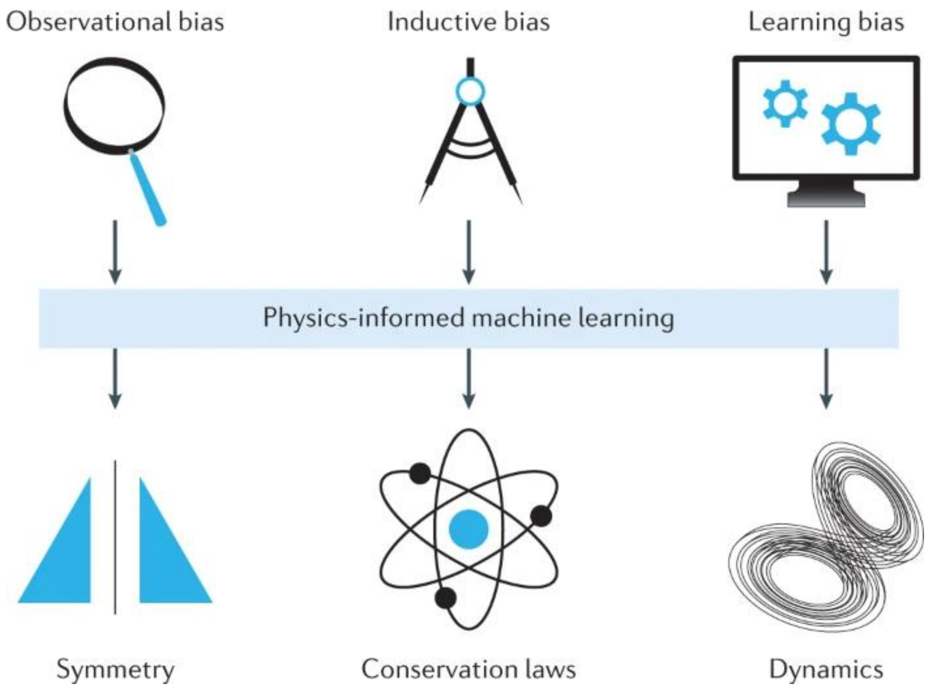
Principles of Physics-Informed Machine Learning

- Observational biases
- Inductive Biases
  - Prior assumptions are integrated in ML-model design
  - Physical laws expressed as mathematical constraints
  - Predictions strictly satisfy the laws of physics
- Learning biases

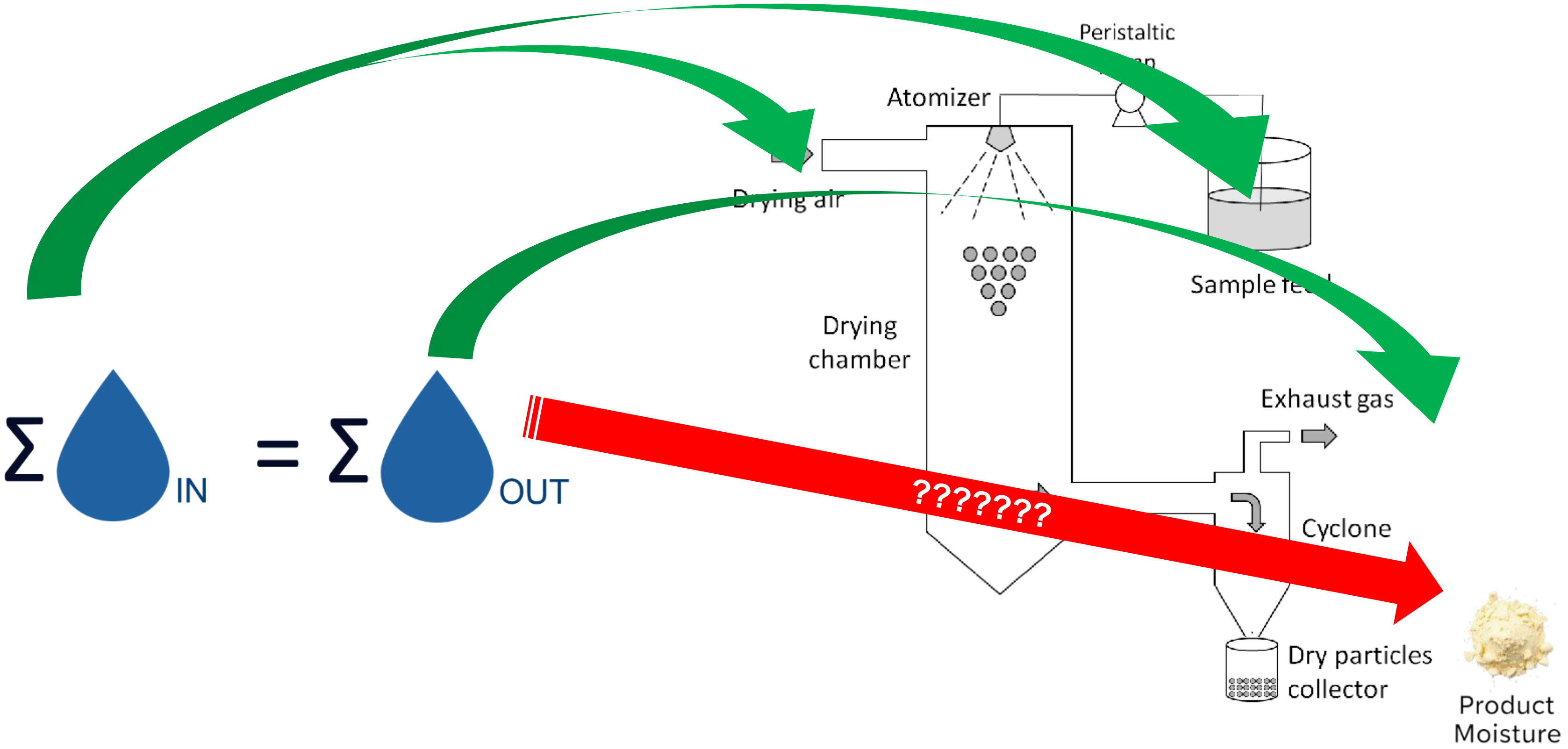


Principles of Physics-Informed Machine Learning

- Observational biases
- Inductive Biases
- Learning biases
  - Imposing physics as soft constraints via the loss function
  - $\mathcal{L} = w_{data}\mathcal{L}_{data} + w_{PDE}\mathcal{L}_{PDE}$
  - Model has to satisfy both the observed data and physical constraints



Example: Moisture Prediction – Mass Balance



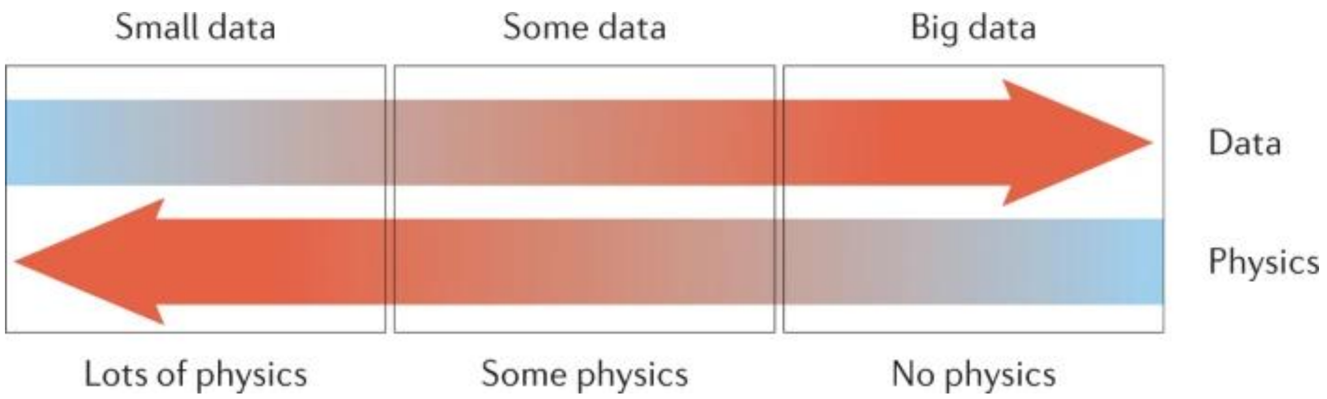
Advantages of Physics-Informed Machine Learning

- Work well on imperfect data & models
- Strong generalization on small data
- Physics-informed models can extrapolate
- Mesh free → less computationally expensive
- Easy implementation due to available libraries

Research Questions

1. How can machine learning and physics-based modelling be combined to simulate spray drying processes?
2. **How can we obtain the necessary data?**

Data Requirements & Physics



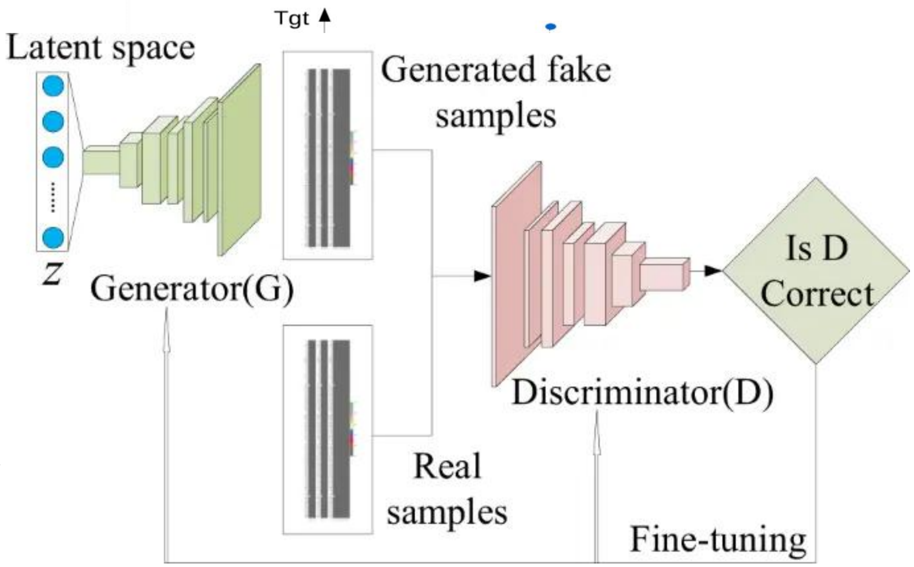
Karniadakis, G.E., Kevrekidis, I.G., Lu, L. *et al.* Physics-informed machine learning. *Nat Rev Phys* **3**, 422–440 (2021). https://doi.org/10.1038/s42254-021-00314-5

Planned 1000 to 2000 measurement points within 18 months

→ Data Basis in the beginning will be limited.

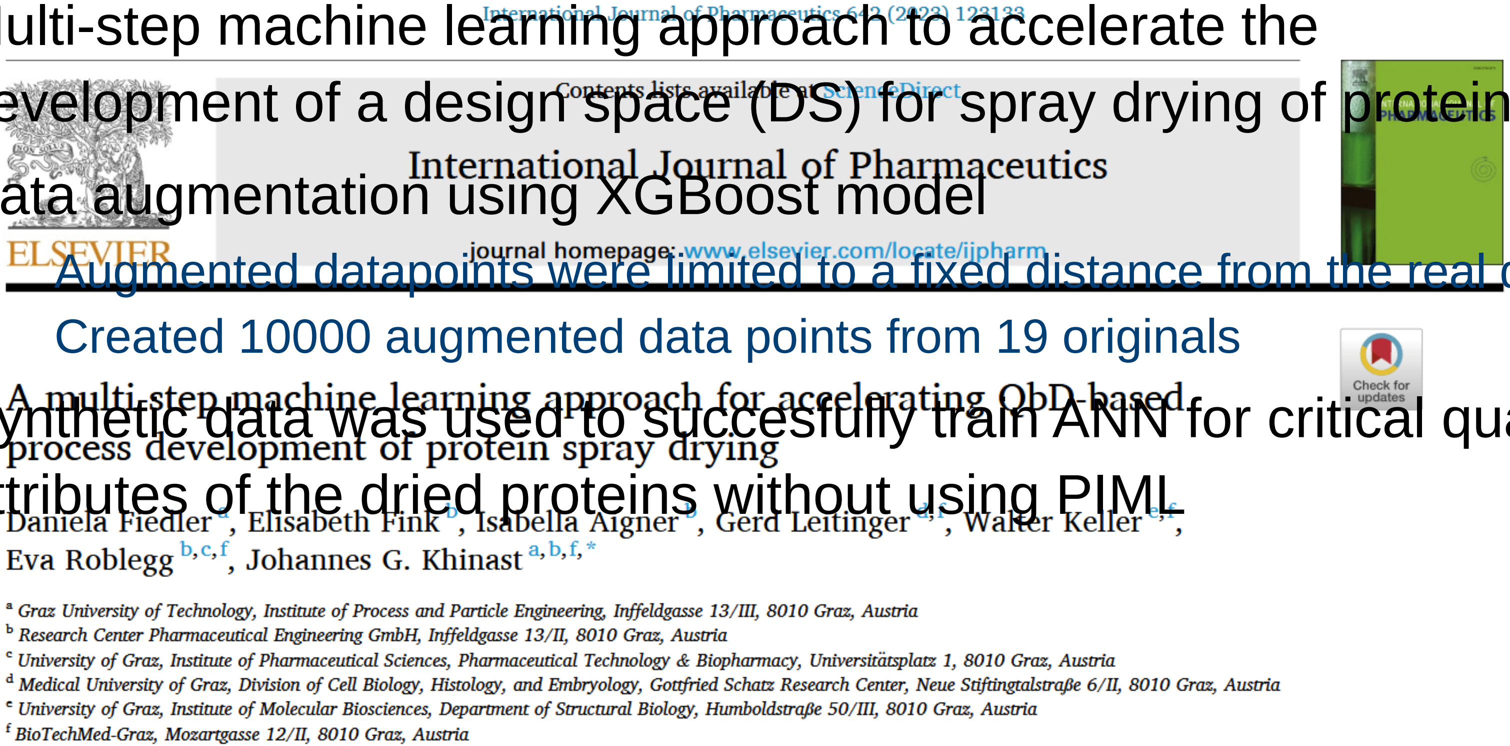
Data Augmentation

- Limited database requires augmentation to create additional data points
- Wide variety of options:
  - Numeric Transformations
    → adding Noise, changing scaling
  - Permutation of feature values
  - Interpolation → linear interpolation of multiple data points to form new ones e.g. SMOTER/SMOGN
  - Generative Models for creating additional data e.g. GAN



Multistep Machine Learning

- Multi-step machine learning approach to accelerate the development of a design space (DS) for spray drying of proteins.
- Data augmentation using XGBoost model
  - Augmented datapoints were limited to a fixed distance from the real data
  - Created 10000 augmented data points from 19 originals
- Synthetic data was used to succesfully train ANN for critical quality attributes of the dried proteins without using PIML


International Journal of Pharmaceutics 642 (2023) 123133

Contents lists available at ScienceDirect

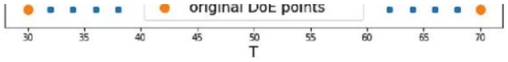
International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

A multi-step machine learning approach for accelerating QbD-based process development of protein spray drying

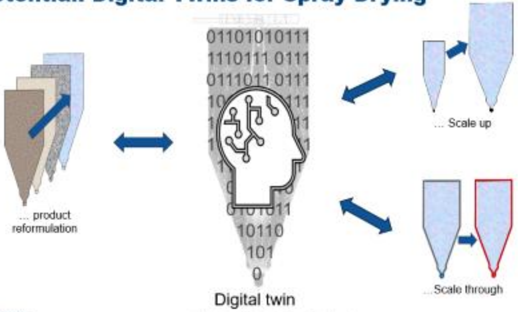
Daniela Fiedler[a], Elisabeth Fink[b], Isabella Aigner[b], Gerd Leitinger[d,f], Walter Keller[e,f], Eva Roblegg[b,c,f], Johannes G. Khinast[a,b,f,*]

[a] Graz University of Technology, Institute of Process and Particle Engineering, Inffeldgasse 13/III, 8010 Graz, Austria
[b] Research Center Pharmaceutical Engineering GmbH, Inffeldgasse 13/II, 8010 Graz, Austria
[c] University of Graz, Institute of Pharmaceutical Sciences, Pharmaceutical Technology & Biopharmacy, Universitätsplatz 1, 8010 Graz, Austria
[d] Medical University of Graz, Division of Cell Biology, Histology, and Embryology, Gottfried Schatz Research Center, Neue Stiftingtalstraße 6/II, 8010 Graz, Austria
[e] University of Graz, Institute of Molecular Biosciences, Department of Structural Biology, Humboldtstraße 50/III, 8010 Graz, Austria
[f] BioTechMed-Graz, Mozartgasse 12/II, 8010 Graz, Austria



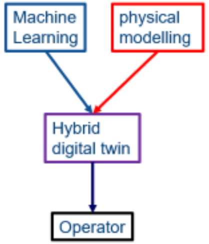
Summary

Potential: Digital Twins for Spray Drying



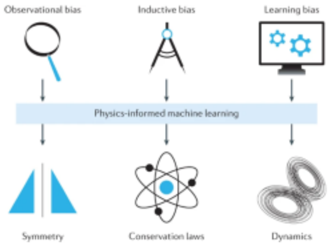
Project Goal

- A hybrid digital twin consisting of physical modeling and machine learning (ML) models.
- This model should …
  - Reflect the system behavior of the spray dryer
  - Gives the operator recommended actions
  - Be transferable between different systems
  - Not operate as a black box



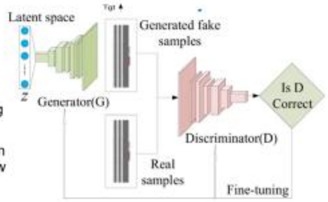
Principles of Physics-Informed Machine Learning

- Observational biases
- Inductive Biases
- Learning biases



Data Augmentation

- Limited database requires augmentation to create additional data points
- Wide variety of options:
  - Numeric Transformations → adding Noise, changing scaling
  - Permutation of feature values
  - Interpolation → linear interpolation of multiple data points to form new ones e.g. SMOTER/SMOGN
  - Generative Models for creating additional data e.g. GAN



References

Mohamad Baassiri, Vivek Ranade, Luis Padrela, CFD modelling and simulations of atomization-based processes for production of drug particles: A review, International Journal of Pharmaceutics, Volume 670, 2025, 125204, ISSN 0378-5173, https://doi.org/10.1016/j.ijpharm.2025.125204.

Daniela Fiedler, Elisabeth Fink, Isabella Aigner, Gerd Leitinger, Walter Keller, Eva Roblegg, Johannes G. Khinast, A multi-step machine learning approach for accelerating QbD-based process development of protein spray drying, International Journal of Pharmaceutics, Volume 642, 2023, 123133, ISSN 0378-5173, https://doi.org/10.1016/j.ijpharm.2023.123133.

Goodfellow, Ian & Pouget-Abadie, Jean & Mirza, Mehdi & Xu, Bing & Warde-Farley, David & Ozair, Sherjil & Courville, Aaron & Bengio, Y.. (2014). Generative Adversarial Networks. Advances in Neural Information Processing Systems. 3. 10.1145/3422622.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L. et al. Physics-informed machine learning. Nat Rev Phys 3, 422–440 (2021). https://doi.org/10.1038/s42254-021-00314-5

Kasilingam, S., Yang, R., Singh, S. K., Farahani, M. A., Rai, R., & Wuest, T. (2024). Physics-based and data-driven hybrid modeling in manufacturing: a review. Production & Manufacturing Research, 12(1). https://doi.org/10.1080/21693277.2024.2305358

Paula Branco, Luís Torgo, Rita P. Ribeiro, SMOGN: a Pre-processing Approach for Imbalanced Regression, Proceedings of the First International Workshop on Learning with Imbalanced Domains: Theory and Applications, PMLR 74:36-50, 2017.

References cont‘d

M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics, Volume 378, 2019, Pages 686-707, ISSN 0021-9991, https://doi.org/10.1016/j.jcp.2018.10.045.

Scott M. Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777.

Seyedalborz Manavi, Ehsan Fattahi, Thomas Becker, A parameterized physics-informed machine learning approach for solving heat and mass transfer equations in the drying process, International Communications in Heat and Mass Transfer, Volume 158, 2024, 107897, ISSN 0735-1933, https://doi.org/10.1016/j.icheatmasstransfer.2024.107897.

Seyedalborz Manavi, Ehsan Fattahi, Thomas Becker, A trial solution for imposing boundary conditions of partial differential equations in physics-informed neural networks, Engineering Applications of Artificial Intelligence, Volume 127, Part B, 2024, 107236, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2023.107236.

Shuo Zhang, Nan Liu, Yanqiu Pan, Wei Wang, Yizhe Li, Yanni Zhu, Three-dimensional modelling of two-phase flow and transport in a pilot centrifugal spray dryer, Chemical Physics Letters, Volume 765, 2021, 138309, ISSN 0009-2614, https://doi.org/10.1016/j.cplett.2020.138309.

Suraj Pawar, Omer San, Burak Aksoylu, Adil Rasheed, Trond Kvamsdal; Physics guided machine learning using simplified theories. Physics of Fluids 1 January 2021; 33 (1): 011701. https://doi.org/10.1063/5.0038929

List of Figures

- Folie 2: File:Spray-Dried-Coffee.jpg. (2024, July 17). Wikimedia Commons. Retrieved 10:28, May 23, 2025 from https://commons.wikimedia.org/w/index.php?title=File:Spray-Dried-Coffee.jpg&oldid=899852416.
- Folie 3:File:Labspraydryer.svg. (2020, October 6). Wikimedia Commons. Retrieved 11:05, May 23, 2025 from https://commons.wikimedia.org/w/index.php?title=File:Labspraydryer.svg&oldid=482975188.
- Folie 3: Florian Kaltenecker (2025)
- Folie 4: Adapted from Silva, Neiton & Andrade, Gabriela & Barrozo, Marcos. (2024). Spray-Dried Jaboticaba Powder as Food Resource. Resources. 13. 102. 10.3390/resources13080102.
- Folie 5: Christian Krupitzer (2025)
- Folie 6: https://www.ipa.fraunhofer.de/de/referenzprojekte/interopera.html
- Folie 7: Shuo Zhang, Nan Liu, Yanqiu Pan, Wei Wang, Yizhe Li, Yanni Zhu, Three-dimensional modelling of two-phase flow and transport in a pilot centrifugal spray dryer, Chemical Physics Letters, Volume 765, 2021, 138309, ISSN 0009-2614, https://doi.org/10.1016/j.cplett.2020.138309.
- Folie 11 & 19: Karniadakis, G.E., Kevrekidis, I.G., Lu, L. et al. Physics-informed machine learning. Nat Rev Phys 3, 422–440 (2021). https://doi.org/10.1038/s42254-021-00314-5
- Folie 12-15: Karniadakis, G.E., Kevrekidis, I.G., Lu, L. et al. Physics-informed machine learning. Nat Rev Phys 3, 422–440 (2021). https://doi.org/10.1038/s42254-021-00314-5
- Folie 16: Adapted from Silva, Neiton & Andrade, Gabriela & Barrozo, Marcos. (2024). Spray-Dried Jaboticaba Powder as Food Resource. Resources. 13. 102. 10.3390/resources13080102.
- Folie 20: Paula Branco, Luís Torgo, Rita P. Ribeiro, SMOGN: a Pre-processing Approach for Imbalanced Regression, Proceedings of the First International Workshop on Learning with Imbalanced Domains: Theory and Applications, PMLR 74:36-50, 2017.
- Folie 20: Dan, Y., Zhao, Y., Li, X. et al. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. npj Comput Mater 6, 84 (2020). https://doi.org/10.1038/s41524-020-00352-0
- Folie 21: Daniela Fiedler, Elisabeth Fink, Isabella Aigner, Gerd Leitinger, Walter Keller, Eva Roblegg, Johannes G. Khinast, A multi-step machine learning approach for accelerating QbD-based process development of protein spray drying, International Journal of Pharmaceutics, Volume 642, 2023, 123133, ISSN 0378-5173, https://doi.org/10.1016/j.ijpharm.2023.123133.
- Folie 29: Wikimedia Commons contributors, 'File:Transfer learning.svg', Wikimedia Commons, 13 January 2024, 05:44 UTC, <https://commons.wikimedia.org/w/index.php?title=File:Transfer_learning.svg&oldid=841017966> [accessed 3 June 2025]
- Folie 31: Shap in https://github.com/shap/shap [accessed 4 June 2025]

Relevant PDE

- Airflow → Navier-Stokes equation, mass balance
- Heat and moisture transfer → energy and mass balances, Ficks Law
- Droplet dynamics → Newton‘s law
- Particle size distributions → Population balance equation

Libraries for Physics Informed Machine Learning

Software Name	Usage	Language	Backend
DeepXDE	Solver	Python	TensorFlow
SimNet	Solver	Python	TensorFlow
PyDEns	Solver	Python	TensorFlow
NeuroDiffEq	Solver	Python	PyTorch
NeuralPDE	Solver	Julia	Julia
SciANN	Wrapper	Python	TensorFlow
ADCME	Wrapper	Julia	TensorFlow
GPyTorch	Wrapper	Python	PyTorch
Neural Tangents	Wrapper	Python	JAX

Research Questions

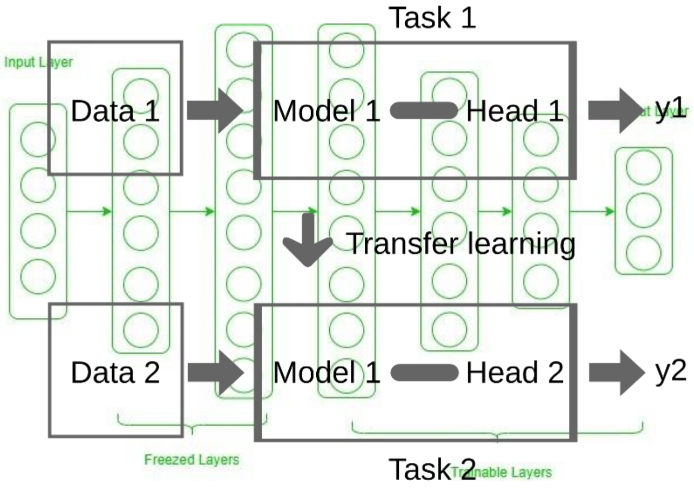
1. How can machine learning and physics-based modelling be combined to simulate spray drying processes?
2. How can we obtain the necessary data?
3. **How can we transfer the results between different (sized) spray drying setups?**

Why Transferring Results?

- The model should be applicable to commercial spray dryers
- A lab-scale spray dryer might help to speed up the data collection
  - Faster change of conditions = more samples in the same time
  - Faster cleanup
  - Reduces operating cost
- Challenges:
  - Results not directly transferable to larger setups
  - Physics differ between small and large scale setups
  - cannot recreate all experimental conditions of a large spray dryer

Transfer Learning

- = Repurposing a trained model for a second task
- Avoids costly retraining
- Often requires small changes to the model
- Example - Transferlearning for ANN:
  - Freezing weights of pre-trained layers
  - Changes to predicting layers allowed
  - Reduces training time

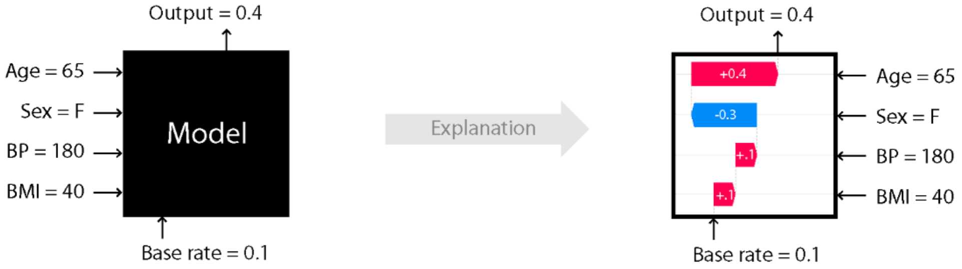


Research Questions

1. How can machine learning and physics-based modelling be combined to simulate spray drying processes?
2. How can we obtain the necessary data?
3. How can we transfer the results between different (sized) spray drying setups?
4. **How can we make the AI model's recommendation transparent for the Operator?**

XAI – Explainable AI

- Recommendations should be verifiable to ensure trust in the system
- Explainability can result from:
  - Intrinsically explainable models
  - Post hoc methods e.g. feature importances, SHAP
- Issue: Explainable Models are often less performant as black box models



Timeline

10/2024
Start as research assistant

02/2025
Start PW David Vollmar

05/2025
Official start FEI Project

05/2025
Begin Data Collection

09/2025
Poster presentation PARTEC 2025

10-12/2025
Paper based on Masters thesis results