Lesson#4: **Empirical Models in Chemical-Technological Processes**

Goal

This lecture focuses on the principles and applications of empirical modeling in chemical-technological processes. Students will learn how experimental data are used to construct mathematical models that describe system behavior when first-principle (mechanistic) approaches are too complex or unavailable. The lecture emphasizes linear, non-linear, and multivariable regression techniques, as well as the emerging role of machine-learning-based empirical models in process optimization and control.

Empirical models are based on observed data rather than theoretical derivations.

In chemical-technological processes, empirical models are often used to predict system behavior where first-principle models are complex or unavailable.

They rely on fitting mathematical functions to experimental data.

Importance in Industry

- Often used for process optimization, control, and scaling up in various industries like petrochemicals, pharmaceuticals, and materials processing.
- Empirical models allow for faster decision-making when precise mechanistic models are not feasible or require excessive computation.

Key Differences with Mechanistic Models

- Mechanistic models depend on understanding fundamental chemistry and physics, while empirical models focus on the relationship between input and output variables without delving into the underlying mechanisms.

Types of Empirical Models

Linear Models

- Simple linear regression models often used when the relationship between variables is straightforward.

- Example: $Y = a + bX$, where Y is the output, X is the input, and a, b are constants derived from data.

Non-Linear Models

- When processes exhibit complex behaviors, non-linear empirical models are more appropriate.

- Example: Exponential or power law models used to describe chemical reaction rates or diffusion processes.

Multivariable Empirical Models

- Often, chemical processes depend on multiple input factors (temperature, pressure, concentration, etc.), requiring multivariable models like multiple linear regression (MLR) or response surface methods (RSM).

- Example: MLR model $Y = a + b_1X1 + b_2X_2 + \dots + b_nX_n$, where multiple inputs affect the output.

Machine Learning-Based Empirical Models

- Modern empirical modeling can leverage machine learning (ML) algorithms, such as neural networks, which can capture very complex input-output relationships without explicit mathematical forms.

- Widely used in process control and optimization today due to their adaptability.

Applications and Limitations of Empirical Models

Applications

- Process Optimization: Tuning operating conditions for desired product quality and yields.

- Quality Control: Predicting product quality from online measurements (e.g., temperature, pressure).

- Process Simulation: For scenario testing and sensitivity analysis where mechanistic understanding is incomplete.

- Fault Detection: Identifying when a process deviates from normal operating conditions based on empirical trends.

Limitations

- Data Dependency: Empirical models are only as good as the data used to develop them. Poor-quality or insufficient data can lead to unreliable models.

- Limited Extrapolation: They work well within the range of data they are built on but can fail outside this range.

- Lack of Physical Insight: Unlike mechanistic models, empirical models provide little insight into the underlying chemistry or physics of the process.

Balancing Empirical and Mechanistic Models

- Often, a hybrid approach combining empirical and mechanistic models provides the best of both worlds, where empirical models capture aspects of the process that are hard to describe theoretically, while mechanistic models bring physical insight.

Learning Outcomes

By the end of this lecture, students will be able to:

1. Explain the concept and importance of empirical modeling in chemical-technological processes and its difference from mechanistic modeling (related to LO 3, ID 3.1).

2. Apply basic linear and non-linear regression techniques to fit experimental data and predict process outcomes (related to LO 3, ID 3.3).

3. Construct and interpret multivariable empirical models, such as multiple linear regression (MLR) and response-surface methods (RSM), for process optimization (related to LO 3, ID 3.3).

4. Discuss the advantages, limitations, and reliability of empirical models in comparison with mechanistic and hybrid approaches (related to LO 3, ID 3.3).

Questions and Self-study Assignments

1. Define an empirical model and explain how it differs from a mechanistic model.

2. Using example data (temperature vs yield), construct a simple linear regression model and interpret the meaning of the coefficients aaa and bbb.

3. Explain when it is appropriate to use non-linear empirical models in chemical-process analysis. Provide one example.

4. Develop a multiple linear regression (MLR) model for a process where yield depends on temperature, pressure, and residence time.

5. Describe how response-surface methodology (RSM) can be used to find optimal operating conditions in a multivariable process.

6. Identify three advantages and three limitations of empirical models in process engineering.

7. Discuss how machine-learning methods can enhance empirical modeling, particularly for process monitoring and fault detection.

8. Read a recent journal article (published within the last 3 years) that applies empirical or hybrid modeling to a chemical or electrochemical system. Summarize:

- *the modeling technique used;*
- *main input and output variables;*
- *key findings regarding process optimization or prediction accuracy.*

References

1. Finlayson B.A. Introduction to Chemical Engineering Computing. Second Edition. - John Wiley & Sons, 2012. DOI: 10.1002/9781118309599

2. Pryor R.W. Multiphysics Modeling Using COMSOL5 and MATLAB. - Mercury Learning and Information, 2015. – 700 p.