UNIVERSITEIT GENT

FACULTY OF ENGINEERING AND ARCHITECTURE

Using Artificial Neural Networks to Reconstruct the Composition of Complex Feedstocks

S.P. Pyl, K.M. Van Geem, M.-F. Reyniers, G.B. Marin

Ghent University – Department of Chemical Engineering
Laboratory for Chemical Technology, Technologiepark 918, Ghent, Belgium

Feedstock Reconstruction

Goal

Derive the detailed composition of a complex mixture from a limited set of macroscopic characteristics

Feedstock Properties → **Detailed Feedstock Composition**

- *Fast*
- *Robust*
- *Standard*

Average Molecular Weight
Hydrogen/Carbon ratio
Specific Density
Global PIONA analysis
Global SARA analysis
Group-type GC-MS analysis
Sulfur content
Nitrogen content
^{1}H and/or ^{13}C NMR data
D86 Distillation data
D2887 SimDist data
Etc.

The Need for Detail...

Fundamental Process Modeling = Molecule-based Modeling
→ Accurate feedstock composition is crucial!

Detailed Feedstock Composition → **Fundamental Model** → **Detailed Product Composition**

Microkinetic Model — **Reactor Model**

These models are indispensible for state-of-the-art process design, optimization and control

Artificial Neural Network

Modelling a complex non-linear relalationship between multiple input and output variables

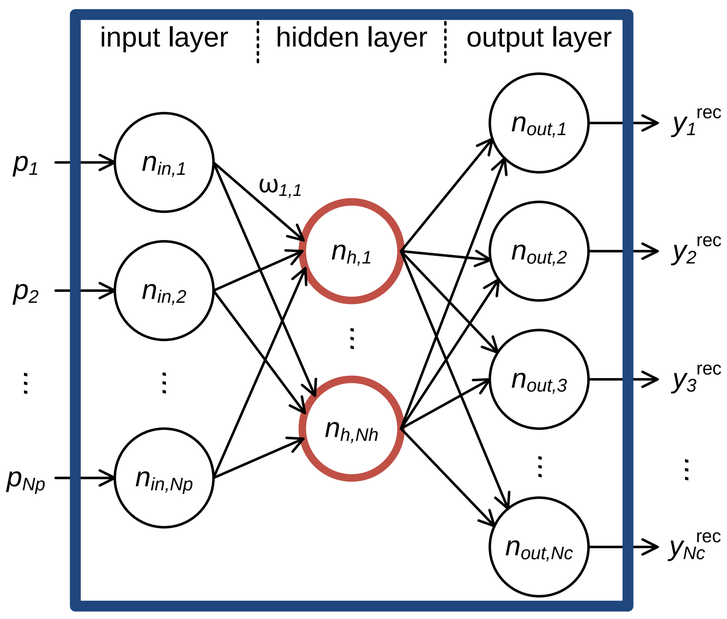
ANN for Naphtha Reconstruction

Input

9 Feedstock Properties

⬇

Global PIONA, Density, ASTM D86 IBP, 50%BP, and FBP



Output

Detailed PIONA $C_3 - C_{13}$

⬇

Weight fractions of 28 components

ANN = nodes + weighted connections

The Hidden Nodes

Input
weighted sum of input signals

$$S_{in,i} = \sum_j \omega_{i,j} \cdot S_j$$

Ouput
transformed signal → Sigmoid funtion

$$S_{out,i} = \frac{1}{1 + e^{-(S_{in,i} - \alpha_i)}}$$

(α_i: activation value)

✓ Restricts the node output to a value between 0 and 1
✓ Introduction of the non-linear modeling properties

Training

How to determine the weights and activation values?
→ A lot of *training data* = examples of the relationship to be modeled

Back-propagation algorithm
→ Gradient based optimization of weights and activation values
→ Minimizing the ***training error*** $TE \sim \sum_i \sum_j (y_{i,j}^{rec} - y_{i,j}^{exp})^2$

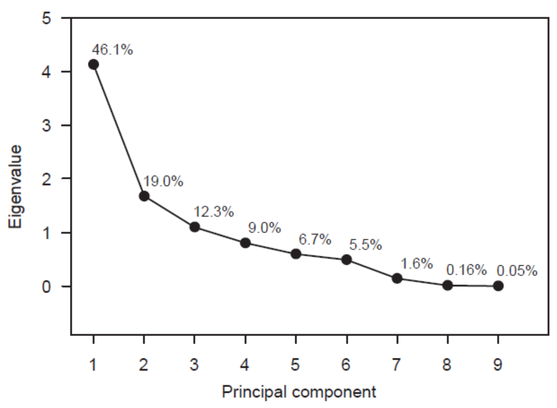
Acknowledgement
The Long Term Structural Methusalem Funding by the Flemish Government – grant number BOF09/01M00409

Application Range

The application range of the Artificial Neural Network is determined by the training data

Principal Component Analysis of the Training Data

Naphtha training set
→ Contains the macroscopic properties and detailed composition of approximately 300 naphthas with a wide range of properties



9 feedstock properties

$$\underset{n\times 3}{\mathbf{Z}} = \underset{n\times 9}{\tilde{\mathbf{P}}} \; \underset{9\times 3}{\mathbf{V}}$$

3 principal components

→ Accounting for approximately 80% of the total variance
→ Graphical representation of the training data in the 3D principal component space

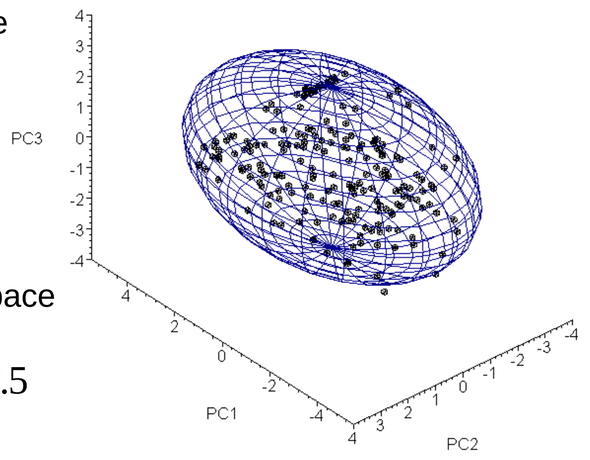
Mahalanobis Distance d_M

→ d_M is a measure for the distance from the data set center, taking into account the variance of the data in each direction

$$d_M = \mathbf{z}^T (\mathbf{Cov})^{-1} \mathbf{z}$$

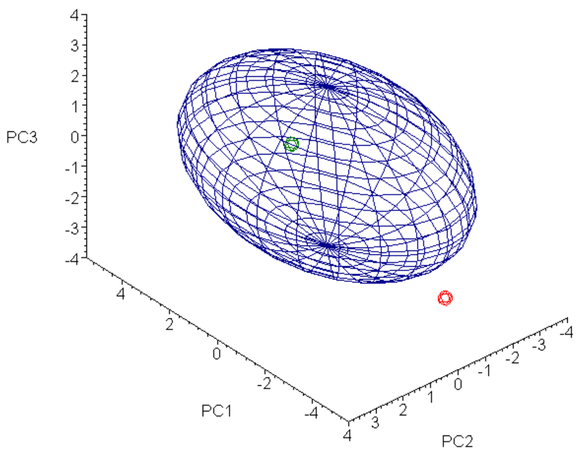
→ d_M follows a Hotelling T^2-distribution
→ critical value defines an ellipsoidal subspace

$$d_M \sim T^2_{p,n,\alpha} = \frac{p(n-1)}{n-p} F_{p,n-p,\alpha} \Rightarrow d_{M,crit} = 2.5$$



Unknown naphthas

A priori classification of the naphtha based on its principal component representation



Naphtha A – in range
d_M = 0.9

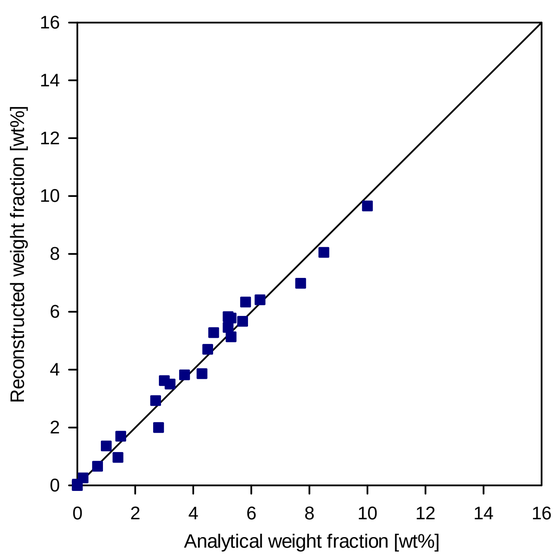
Naphtha B – out of range
d_M = 3.2

Pyl, et al. *AIChE* J. 2010

Validation

Comparison between reconstructed composition and detailed analytical composition, determined using gas chromatography

Naphtha A – in range



Naphtha B – out of range