May 18, 19, 20, 2022
Hannover, Germany

Review of Machine Learning and Experimental Design in Rubber Compounding

Dr. Hans-Joachim Graf, H-JG Consulting

- **Content of this Presentation**

  - **Introduction**
    - **What is Design of Experiments (DoE)**
    - **What is Artificial Intelligence Software**
  - **Analysis of Data for Predictions with Software “Feed Forward AI”**
    - **DoE (Design of Experiments) Prediction Tool**
    - **Analysis of Data with AI Software**
  - **The Problem with Recurrent AI Software**
  - **Conclusion**

[redacted]

➔ Definition:
Statistic Experimental Design

- The (statistical) design of experiments (DOE) is **an efficient procedure for planning experiments** so that the data obtained can be analyzed to yield valid and objective conclusions. ... An Experimental Design is the laying out of a detailed experimental plan in advance of doing the experiment.
- Statistic is a formal science, whose methods are applied to a wide variety of natural and engineering sciences. The interpretation of the results only make sense if they have been verified for their plausibility in the context of the intrinsic sciences.
- In engineering, it is often necessary to work with small samples, so the treatment of small samples or series of experiments and the presentation of distribution-independent test procedures is of particular importance.

Steps of Design of Experiments

Source: T. Duever; UofWaterloo

- An investigation consists of the following steps applied in a <u>sequential</u>, <u>iterative</u> manner:
  - **HYPOTHESIZE**
  - **DESIGN**
  - **ANALYZE**
- The statistic analysis answers the question whether test results are to be considered as random phenomena or may be treated as characteristic.

Regression Analysis

Experimentation: Variation of Factors

Measurements: Responses

F_1 →

F_2 →

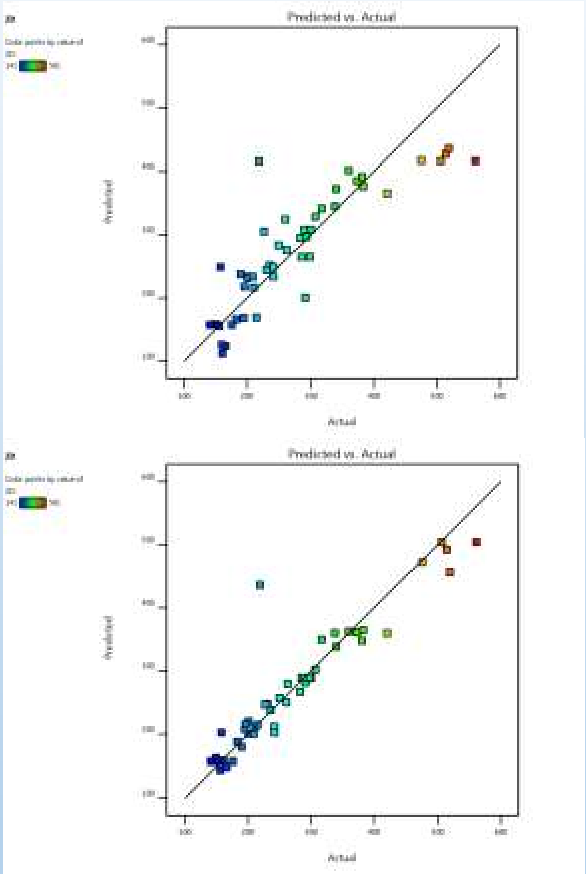
F_3 →

Compound Process → $R_1, R_2,.. R_n$

Objective of the Experiment is the identification of the factors ($F_{1,}..F_n$) type of influence on the responses ($R_{1,}...R_n$) and description with mathematical equations for further processing.
ANOVA is used for statistical evaluation.

$$R_{i(1...n)} = f(A_0 + A_1F_1+....A_nF_n +....)$$

EPDM Curative study



TB (Tensile at Break): Predicted over actual, top – linear, bottom – 2FI

DoE: Point Prediction

Source: Cabot TB RG-133

Properties of Compound is determined by Polymer, CB and Oil content and the ratio of CB and Oil.

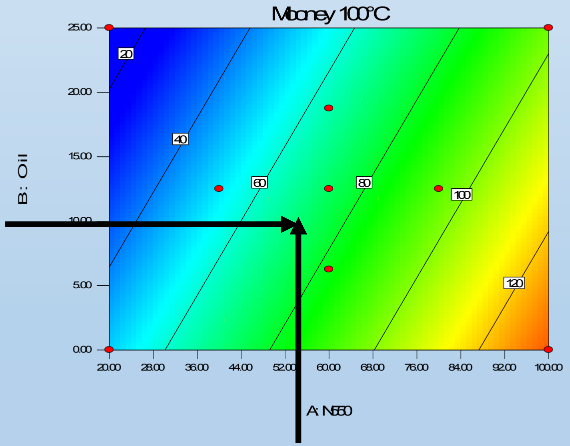
Data of DoE converted into a contour plot

- **CB 550: 55 phr**
- **Oil: 10 phr**

- **Mooney Viscosity: 71 M-Units**
- **Hardness: 60 ° ShA**
- **Tensile: 21 MPa**
- **Elongation: 460 %**
- **C-Set: 28%**

NR Compound

- *SMR 5CV – 100 phr*
- ***CB – Var***
- ***Oil – Var***
- *ZnO – 5 phr*
- *StAc – 1 phr*
- *AO – 1 phr*
- *NR 100 phr*
- *MBTS – 0.6*
- *S – 2.5 phr*

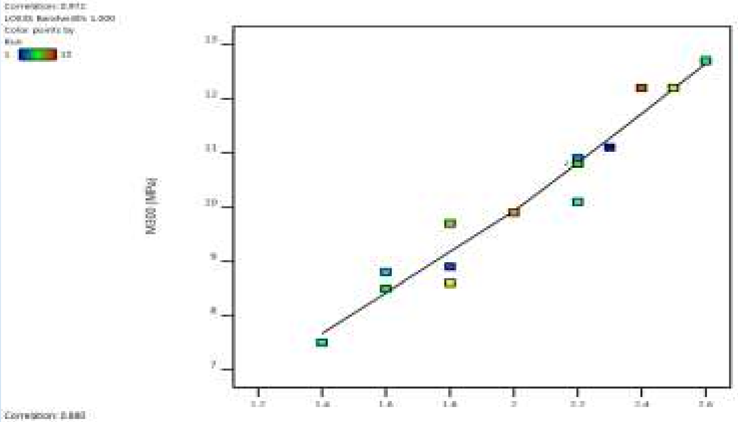


DoE: Tools in Modern DoE Software

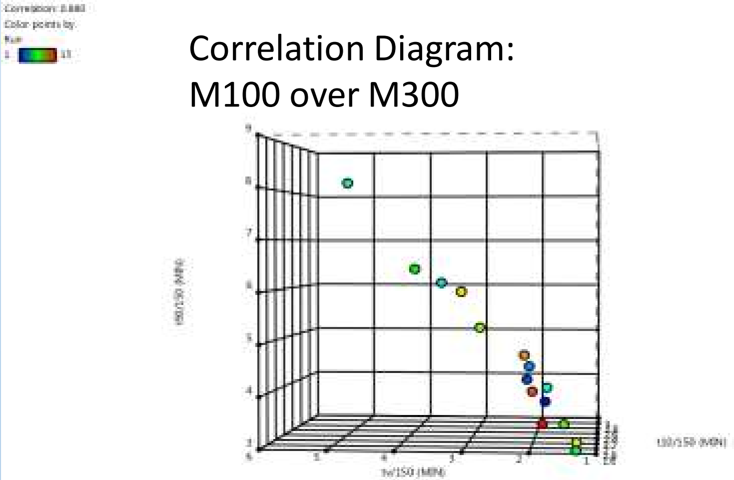
Statistic experimental design software tools in Design Expert®

- **Correlation diagrams allowing a first evaluation of the data:**
  - **Following inherent logic of rubber principles**
- **Histograms**
- **Scatter plots**

- **Further tools are**
  - **Contour plots**
  - **Desirability plots, whether target is met / or failed**
  - **Prediction**



Correlation Diagram:
M100 over M300



Correlation Diagram:
t_{10}/150C, T_v 150C, and t_{90}/150C

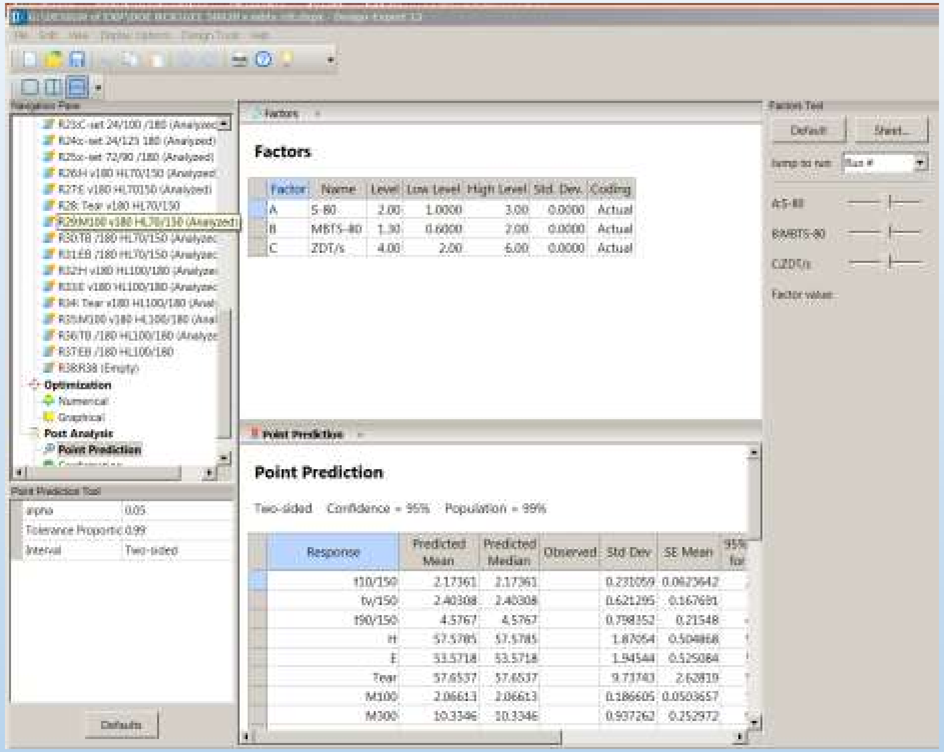
DoE: Point Prediction based on Regression

The Prediction is calculated with the

- **Intercepts and**
- **Regression Factors:**

(*Table shows case for linear regression*)

Response	Intercept	F1	F2	Fn
R1	A_1	$A_{F1.1}$	$A_{F2.1}$	$A_{Fn.1}$
R2	A_2	$A_{F1.2}$	$X_{F2.2}$	$A_{Fn.2}$
.....	.....	.....	.....	.....
Rn	A_n	$A_{F.n}$	$A_{Fn.n}$	$A_{Fn.n}$

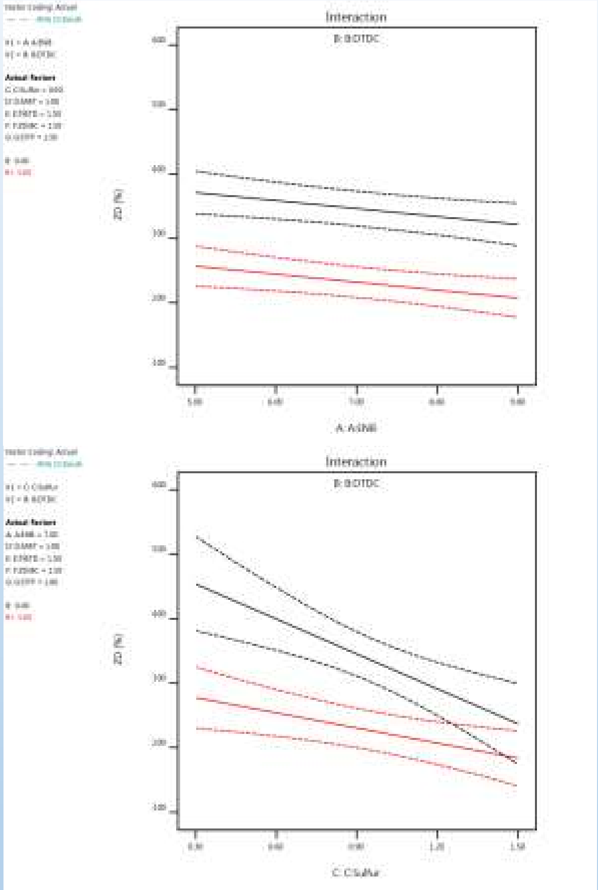


Point Prediction Tool in Design Expert®12 Software (Screen shot)

DoE: Advantages and Limits

- **Advantage**
  - Investigation of New, Unknown Material
  - Results are statistically sound
  - Estimation of Measurement Error
  - Identification of Outliers
  - Possibility to Schedule and Budget Development

- **Disadvantage**
  - Limited Number of Factors in a Compound
    3 – 4 Factors preferred
  - Effect of Variation of Factors among themselves
  - Extrapolation not recommended
  - Results are limited to the compound used for evaluation



TB: Interaction Graph: TB over Sulfur: DTDC
top – linear, bottom – 2FI

AI: Random Datasets

[redacted]

➔ Definition:
Artificial Intelligence

- The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. (google 15.10.21)
- Deep learning, meanwhile, is a subset of machine learning that enables computers to solve more complex problems.
- Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.

AI: Random Datasets

Source: G Roden; Heise Medien 2021

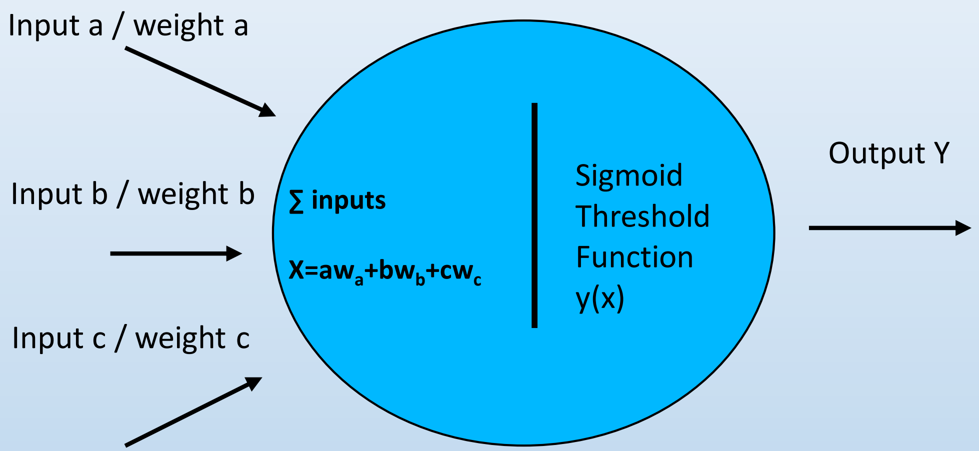
Algorithm used in AI Programming

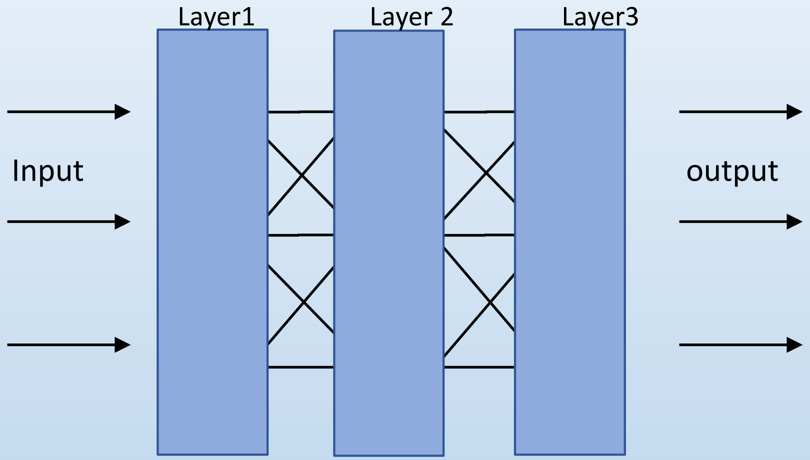
- **k-Means-Algorithm**
  - **Creates cluster, calculates mean (Centroid) reorganization, Iteration**
- **Genetic Algorithm**
  - **Proposed Solution, iteration and mutation towards optimum, selection of results with “fitness” function**
- **Neuronal Net**
  - **Complex algorithm (Neuron): Calculation of a weighted sum, normalization to emphasize or dampen the summands. „Input – Hidden – Output“ layer. Feed Forward Networks / Recurrent Neuronal Networks**

AI: Schematic Structure

Source: T. Rashid, Neuronal Nets, 2017

Artificial Intelligence: Neuronal Network and Algorithm





Extended model of an artificial neuron with moderated input weights

A model with multiple layers of neurons, each connected to the preceding and succeeding layer.

- Weights could be assigned to the connections between nodes

AI: Basic Algorithm

Source: T. Rashid, Neuronal Nets, 2017

Matrix multiplication with the terminology of neuronal nets

*(**w** – weight, **I** = Input, **O** – X-Matrix multiplied sigmoid function)*

$$\begin{bmatrix} I\ 1 & \\ I\ 2 & \end{bmatrix} \quad \begin{bmatrix} w\ 1{,}2 & w\ 2{,}1 \\ w\ 1{,}2 & w\ 2{,}2 \end{bmatrix} = \begin{bmatrix} I\ 1\ * w\ 1{,}1 & I\ 2\ * w\ 2{,}1 \\ I\ 2\ * w\ 1{,}2 & I\ 2\ * w\ 2{,}2 \end{bmatrix}$$

The input values for the subsequent layer is in each case:

X = w * I

Each x must be treated with the sigmoid, respective activation function:

Y = $1/(1+e^{-x})$

Values for subsequent layers change to:

O = sigmoid (**X**)

AI: Old Algorithm on Fast Computers

Source: T. Rashid, Neuronal Nets, 2017

Basic Idea of Neuronal Network Calculation:

- The calculation of signals passing through a neural network can be modeled as a **matrix calculation**.
- The links can be specified more precisely independent of the size of the network.
- Programming languages designed for matrix calculations can perform such calculations efficiently and quickly.

- Calculation speed on todays Machines is fast and able to handle large amounts to calculations in mill seconds

AI: Rubber Compound Formula

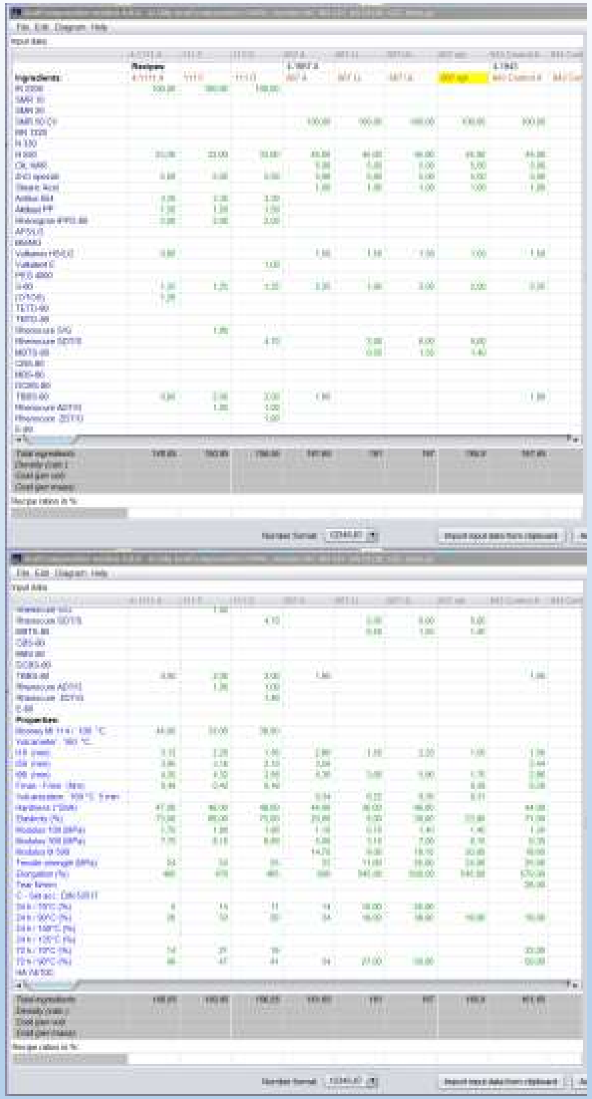
Ingredients

Properties

Rubber Recipe Data Stag

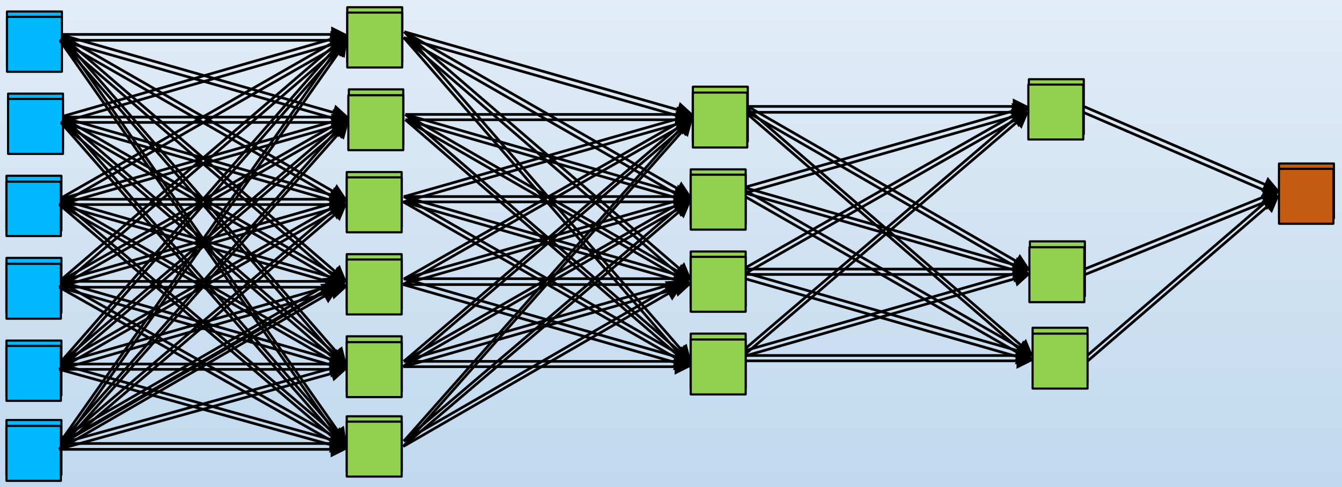
- **Ingredients**
  - **Normalized to 100 parts polymer = phr**

- Properties
  - Rheological properties
  - Physicals
  - Other
    - Appearance
    - Dynamics



Feed Forward AI Program

A square represents one dataset = data stag consist of n data: upper stage – Ingredients; lower stage - properties



Input **1. Layer** **n.th Layer** **n.th + 1 Layer** **Output**

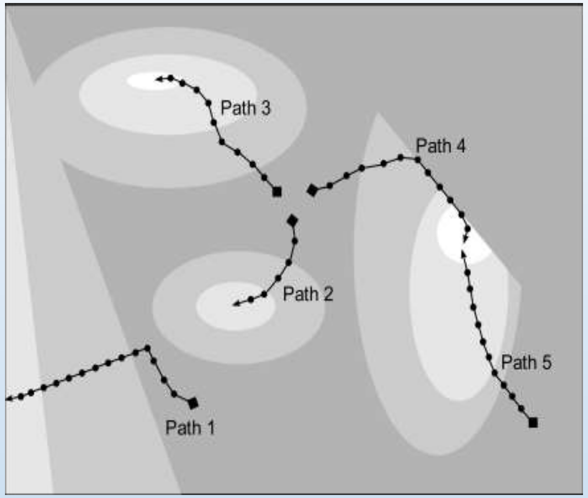
Each square represents recipe and property information

- ❑ **Solutions closer to targets survive**
    - ❑ **Walking gradient method**

Feed Forward AI Program

- Database
  (Historic, most likely incomplete, unorganized Data)
- Input: Data + Multi target query
  - Search for the best compromise with K-Mean / Multi Objective Evolutionary Algorithm (MOEA)
  - Numerical Solution with a "Gradient Walking Method." Data treatment simultaneously.
  - Approximation / Desirability Function calculates distance from target = "Survival Function."
  - Finally selects Solution
- Output:

  Recipe with Ingredients and its Properties



Starting Points:
Wherever cluster of Data can be Identified which shortens the distance to one of the targets

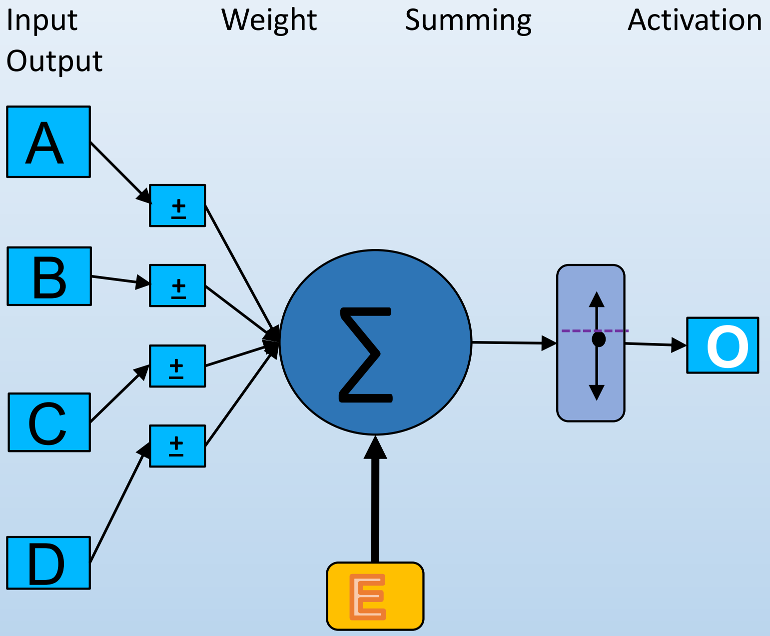
ML: Recurrent AI Program

Source: S. Velasco in Spektrum 11/21

Challenges and Hints to Machine Learning [ML] with “Back Propagation” in rubber compounding:

- **Time consuming testing**
  specifically long term aging / dynamic
- **Compound preparation to generate data with high effort.**
  (Tests with data generated in the laboratory failed in the past)

- Database standardization
  - Basic data collective
  - Specific data collective
- Identification of data errors
  - Errors, if Data transferred
  - Compound preparation / measurement errors



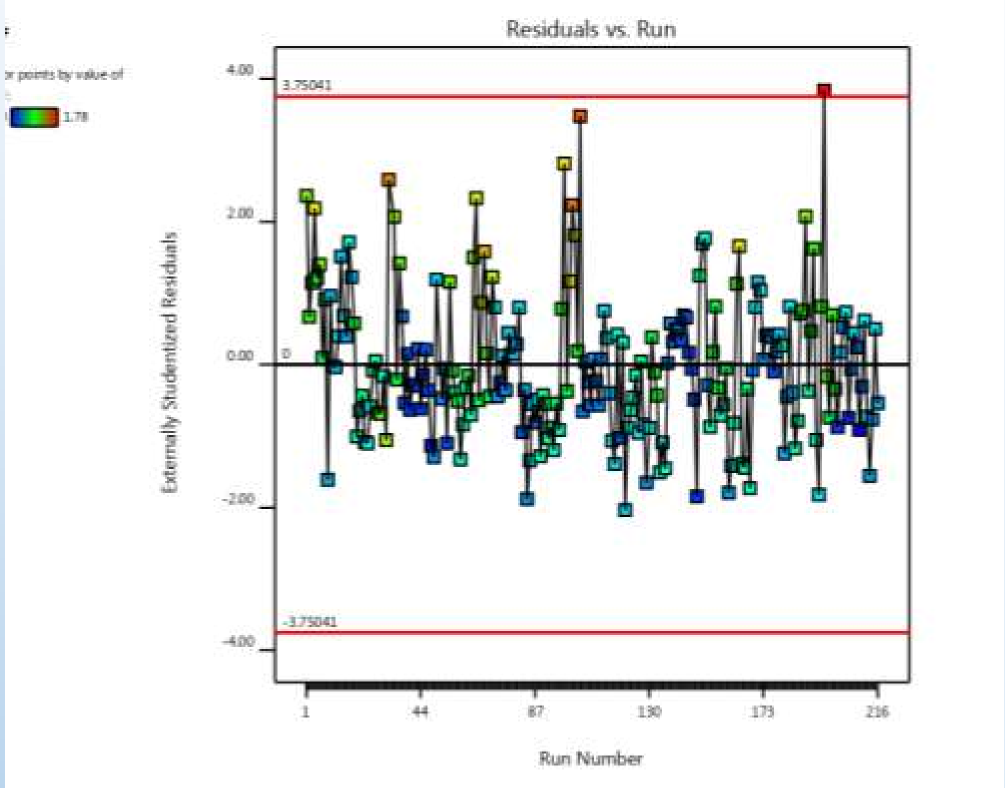
Errors:
The nerve cell learns via errors to optimize the weighting and the values of the input function.

ML: Recurrent AI Program

How Neuronal Nets learn?

The weight functions between knots is adjusted through:

- **Training with the difference between calculated and true value**
- **This difference (error) is not obvious**
    - **Fitting to the dimension of the value**
    - **The error must be a number**
      Each Property needs another error number assigned
- **Machine Learning:**
  Error must be constant, but not subject to fluctuation or variation

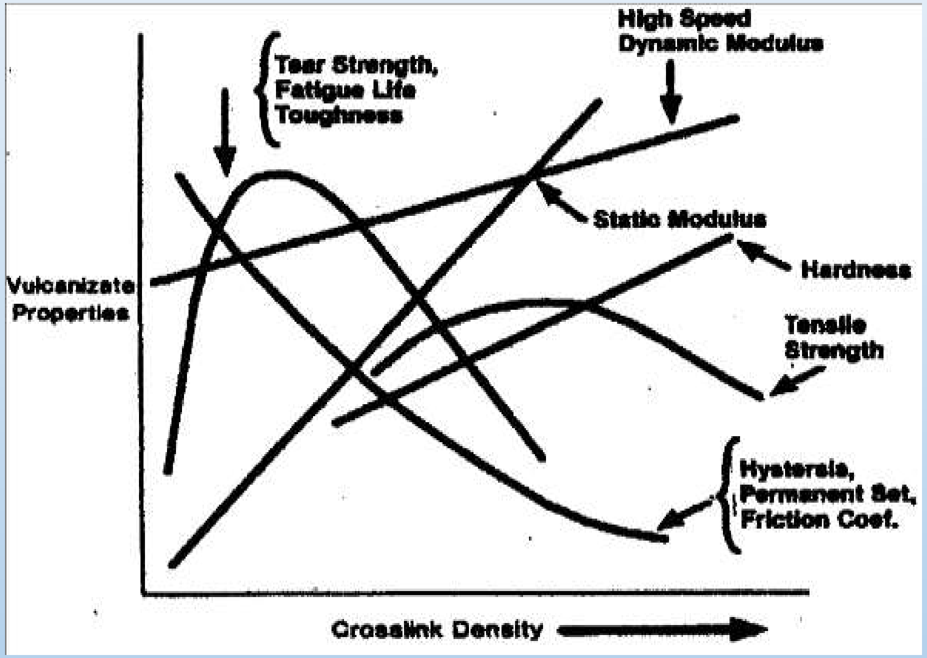


Diagnostics of Hardening Factor

Feed Forward AI
How to improve Dataset

- **Prediction of Compound inside 95% Confidence Interval**
  - Elimination of Outliers
    - Correlation analysis of Properties
    - Properties correlated to Crosslink Density
    - Properties correlated among themselves by nature

Correlation between crosslink density and physical properties

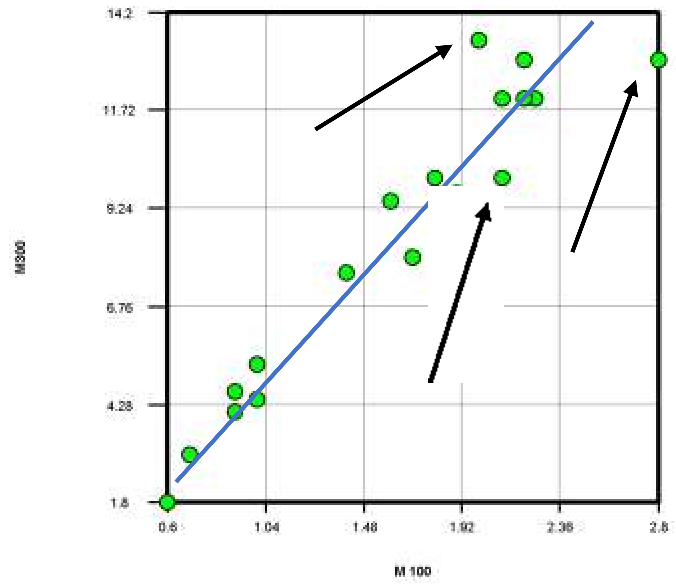


Source: D. Hertz, Elastomerics 1984,
A.Y. Coran, in Science Technology of Rubber, § Vulcanization, Academic Press 1994

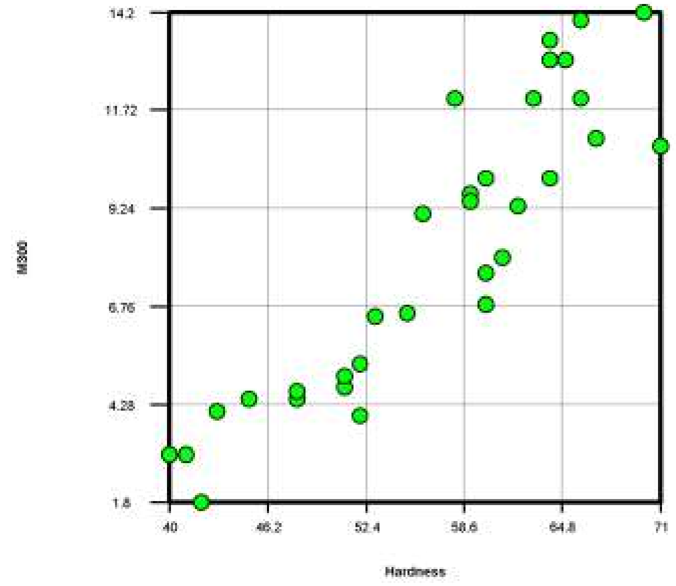
Correlation in Incidental Database

Source: NR-Data MRPRA

- ➢ Module 100 over Module 300



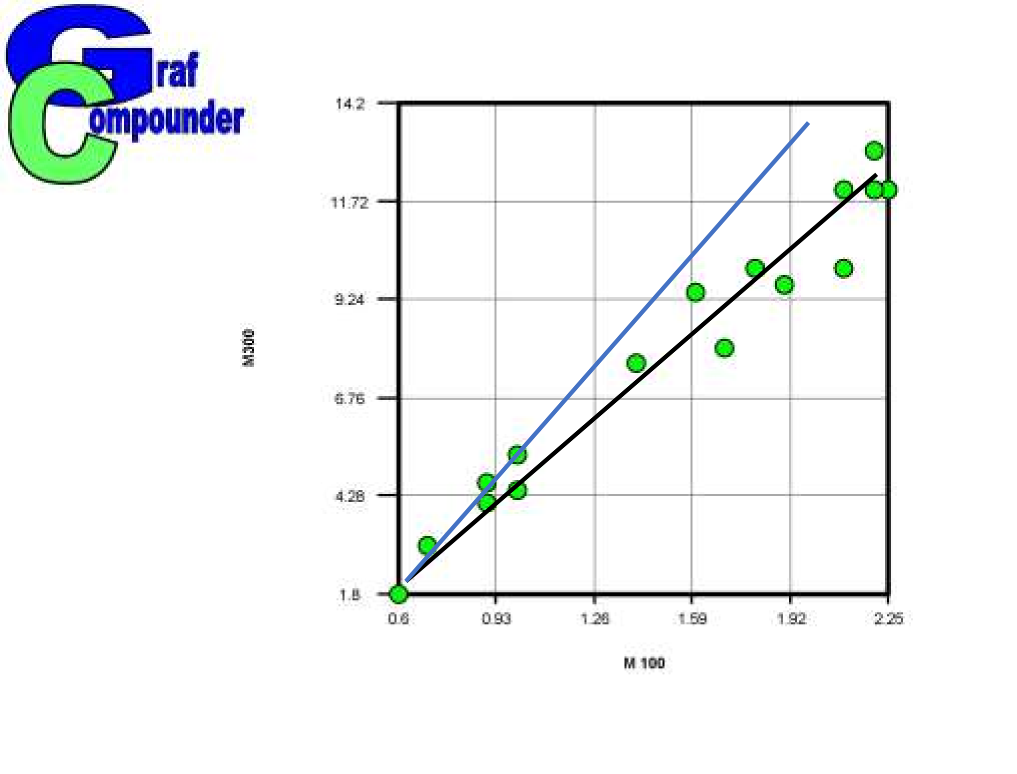
- ➢ Hardness over Module 300



Correlation in Incidental Database

Source: NR-Data MRPRA

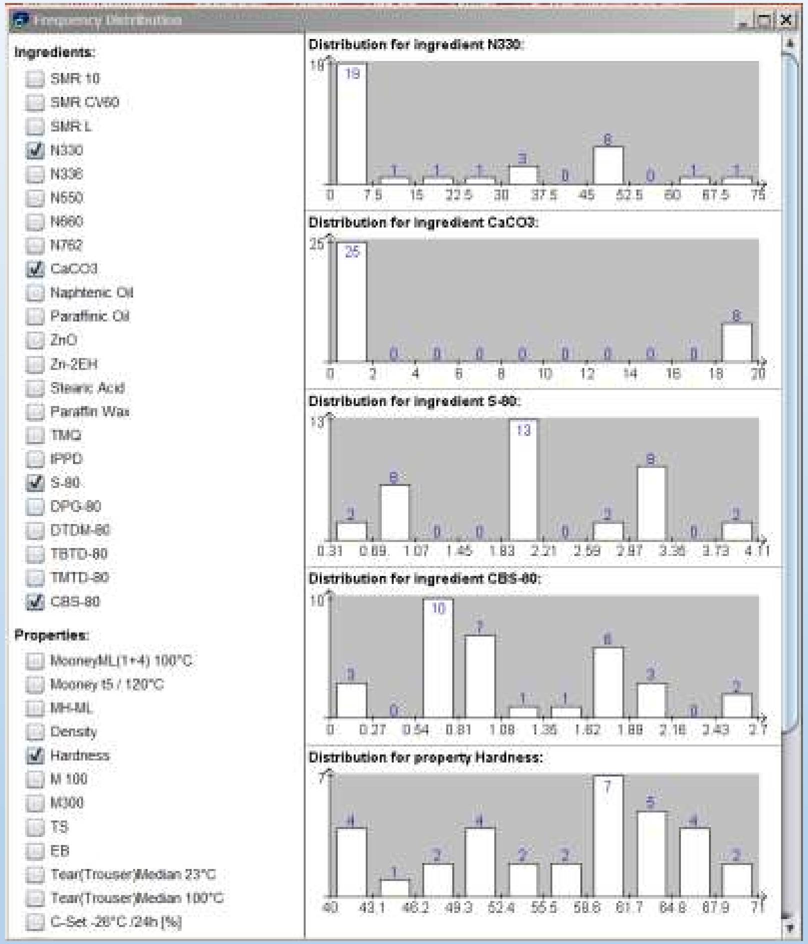
- Module 100 over Module 300



- Cleaning the Data
- Inactivation of Datasets, which large distance from regression line
  - Mouse click – right
  - Click on Refresh recipes
  - Observe shift in regression line

Spread of Data in Incidental Database

- Frequency Distribution of Ingredients
- Distribution of Properties
- Standardization of Properties:
  - Testing according Specification (ISO, ASTM)
  - Additional Testing Procedures (OEM)
- Documentation Standards

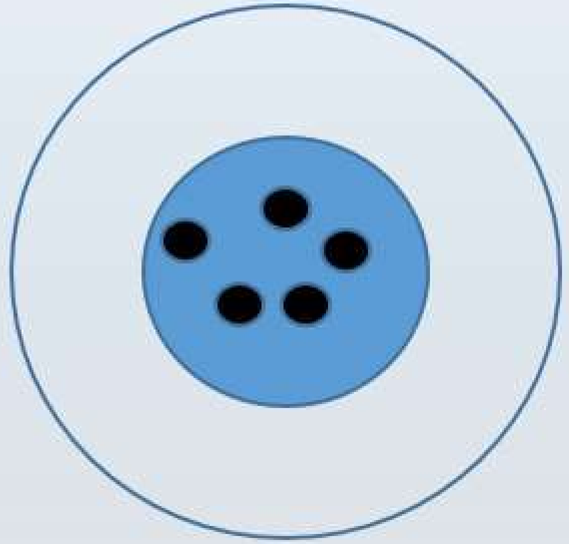


Experimentation:
Decrease of Variation, Increase Precision

- **Laboratory Mixing Machine: Variation**
  - Raw Material / Process Protocol Influences
- **Testing**
  - Test Protocol: 5 instead of 3 samples per Test
  - Evaluation of Measurement Error

    (Man – Machine – Material)
  - Design of Experiments: Number of Repeats
- **Round Robin Testing with external Laboratories**

Source: C.E. McCormick, Evaluating and Utilizing the Precision of a Test Method; Elastomerics 1983



Good Accuracy
Good Precision

Prediction and Repeat of Experiments in the Laboratory

	EPM Compound No1			EPM Compound No2		
	Predicted	Test results	Deviation	Predicted	Test results	Deviation
Hardness	69.99	70.00	0.01%	60.28	61.00	0.59%
Tensile at Break	19.56	18.75	2.16%	15.74	18.08	6.47%
Elongation at Break	338.61	339.00	0.06%	513.48	476.00	3.94%
M 100	3.87	4.75	9.26%	2.04	3.05	16.56%
C-Set 23°C / 22h	10.28	10.91	2.89%	10.00	9.97	0.15%

Data in good agreement except Modulus 100

Feed Forward AI:
What is the advantage?

- **Compound Prediction** is possible with **any** data set
    - Data set can be small or large
- **All significant parameters** should be included in the query simultaneously
    - 2-3 Properties in the query makes no sense.
    - **Most specification values, but at tighter limits**
        - Improve query with “Weights” and “Trdoff”

Criteria:

Name	Min	Max	From	To	Weight	Trdoff
ZnO	1.2	4				
TMQ	0	1.2				
BDMA/S	0	1.1				
TAC-50	0	1.4				
S- 80	0	0.3				
Oil	0	50				
CB 02	0	78				
CB 01	0	95				
Filler 01	0	30				
Perkadox 14/40	6	6				
Density	1.082	1.17				
Hardness	53	79.5		70		
Tensile Strength	13.65	22.8				
Elongation at	247	605				
M100	1.28	6.24				
M300	5.05	16.58				
CSET-22h/23C	7.41	14.55		9	20	
CSET-22h/150C	9.8	32.38		25		
Delta H	1	9	-5	3	10	
Delta TS	-13.21	31.14	-5	10		
Delta EB	-14.74	18.62	0	8	50	
Delta M100	-9.17	65.68	-5	8	20	
Delta M300	-8.13	18.94				
Low Temp	-59.09	49.9		-50	10	
Tear	3.92	12.44				

Criteria: / Output:

Name	Min	Max	Output
Polymer 1	0	[illegible]	
Polymer 2	0	0	
Polymer 3	0	[illegible]	
Polymer 4	0	0	
Polymer 5	0	0	
Polymer 6M	0	50	
Stearic Acid	0.4	0.4	0.4
Process Aid 1	0	4	3.165
PEG	0	2	0.255
Process Aid 2	0	4	0.8775
MgO	0	4	2.069
ZnO	1.2	4	3.2
TMQ	0	1.2	1.2
BDMA/S	0	1.1	0.85975
TAC-50	0	1.4	0.3135
S- 80	0	0.3	0.06475
Oil	0	50	11.8875
CB 02	0	78	58.8125
CB 01	0	95	19.85
Filler 01	0	30	1.9125
Perkadox 14/40	6	6	6
Density	1.082	1.17	1.117973
Hardness	53	79.5	70.9475
Tensile Strength	13.65	22.8	19.31475

AI: Conclusion

Feed Forward AI: GRAFCOMPOUNDER

- Compound History
- Analyze
- Simulate
- Select
- Confirm

EXPERIMENTAL DESIGN

- Explore Materials
- Evaluate
- Decide
- Conclude
- Confirm

Conclusion

- **<u>Feed Forward AI: Compounder Program</u>**
  - Compound Cost target included
  - Starting Formula for Compound Development
  - Simulation of Recipe for Specification Adjustments
  - Historic Data usage and improved Data storage in GC Format

- **<u>Statistic Experimental Design</u>**
  - Evaluation of New Materials
  - Cost optimization in Compounding
  - Design / Improvement of New Processes
  - Process Window to Control Quality according SPC

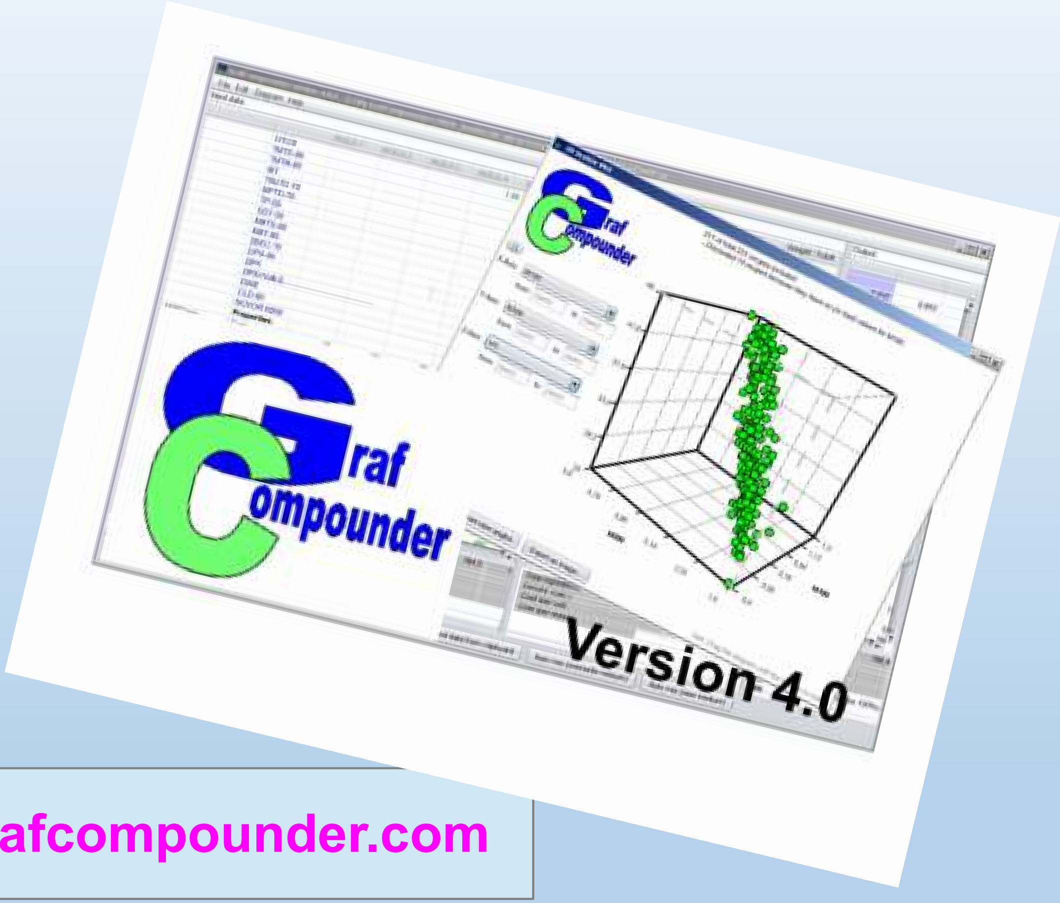
- **Preferred Rubber Compound Development Strategy:**
  Combination of Compound Simulation from Database with Statistic Experimental Design Experimentation Procedure

  Both methods have their justification.

28

Thank you for joining this presentation.

➔Questions, Remarks, Discussion ?



More information under: **www.grafcompounder.com**