The Power of the Curve

Stat-Ease DOE Summit - Oct 5, 2022

Presented by Arved Harding
Eastman

EASTMAN

Who is Arved J. Harding, Jr.?

Statistical Consultant at Eastman
>34 Years

Work at Kingsport, TN site

Senior Statistical Associate

Native of Wise, VA
Currently in Blountville, TN

Family Man

Wife + 2 boys

Active volunteer and leader in the Northeast TN Section of the American Society for Quality (ASQ)

Graduate of UVA College at Wise - B.S. in Math, 1985 + High School Math Teacher Certification

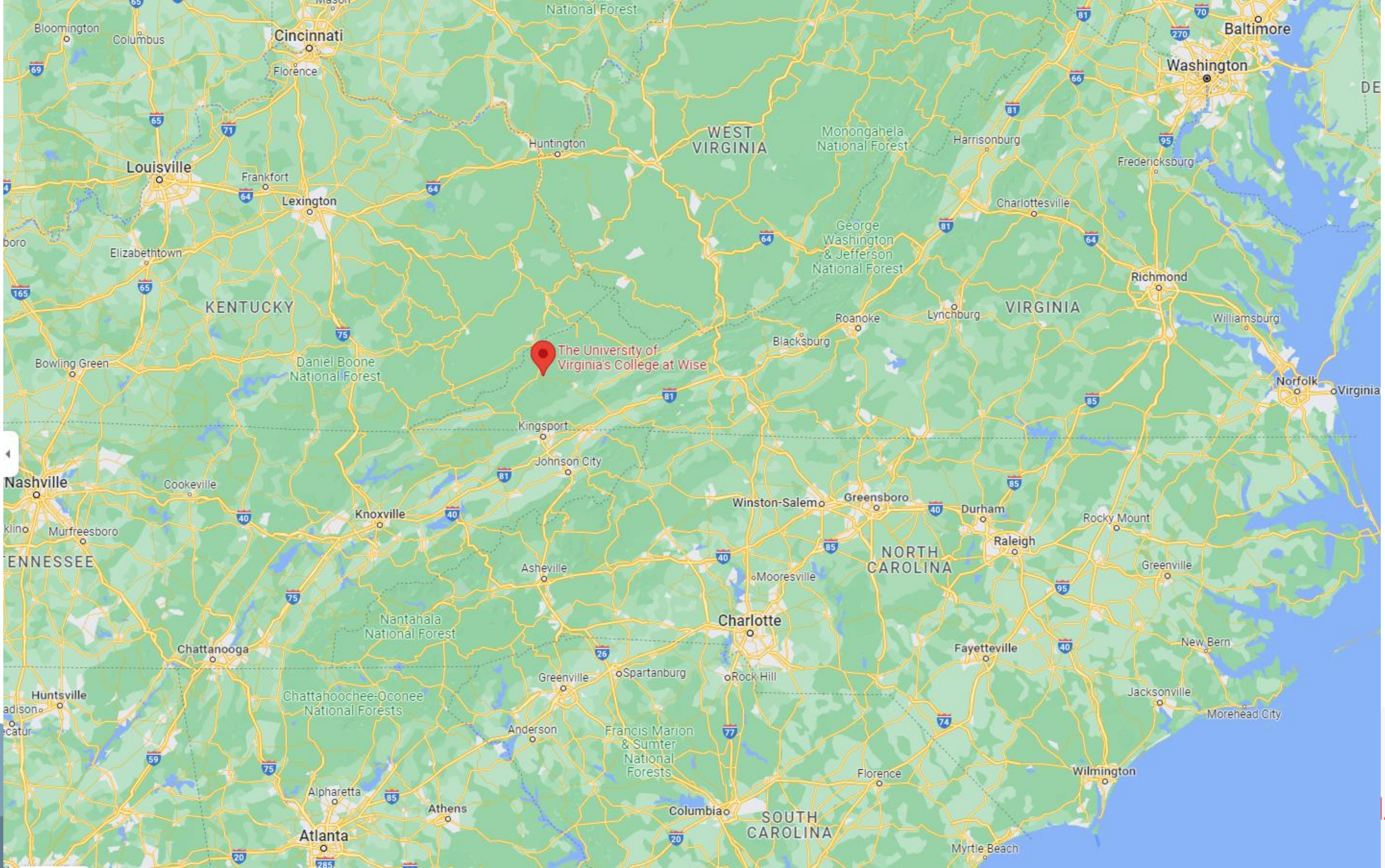
M.S. in Statistics, VA Tech, 1988

Have supported many organizations within Eastman, including research and development, technical service, manufacturing and analytical testing in the fields of Polymers, PVB Interlayers, Performance Films, Cellulose Esters, Coatings, Adhesives and more.

Current assignment:

Advanced Interlayers
Performance Films

Instructor for internal Eastman Statistics and Six Sigma courses


National Forest
Baltimore
Washington
Bloomington
Columbus
Cincinnati
Florence
DE
WEST VIRGINIA
Monongahela National Forest
Harrisonburg
Huntington
Fredericksburg
Louisville
Frankfort
Lexington
Charlottesville
George Washington & Jefferson National Forest
Elizabethtown
Richmond
KENTUCKY
VIRGINIA
Roanoke
Lynchburg
Williamsburg
Blacksburg
The University of Virginia's College at Wise
Bowling Green
Daniel Boone National Forest
Norfolk
Virginia
Kingsport
Johnson City
Nashville
Cookeville
Winston-Salem
Greensboro
Durham
Rocky Mount
Knoxville
Murfreesboro
Raleigh
TENNESSEE
NORTH CAROLINA
Greenville
Asheville
Mooresville
Charlotte
Nantahala National Forest
Fayetteville
New Bern
Chattanooga
Spartanburg
Rock Hill
Greenville
Huntsville
Chattahoochee-Oconee National Forests
Jacksonville
Morehead City
Anderson
Francis Marion & Sumter National Forests
Wilmington
Florence
Alpharetta
Athens
Columbia
SOUTH CAROLINA
Atlanta
Myrtle Beach

A global industry leader

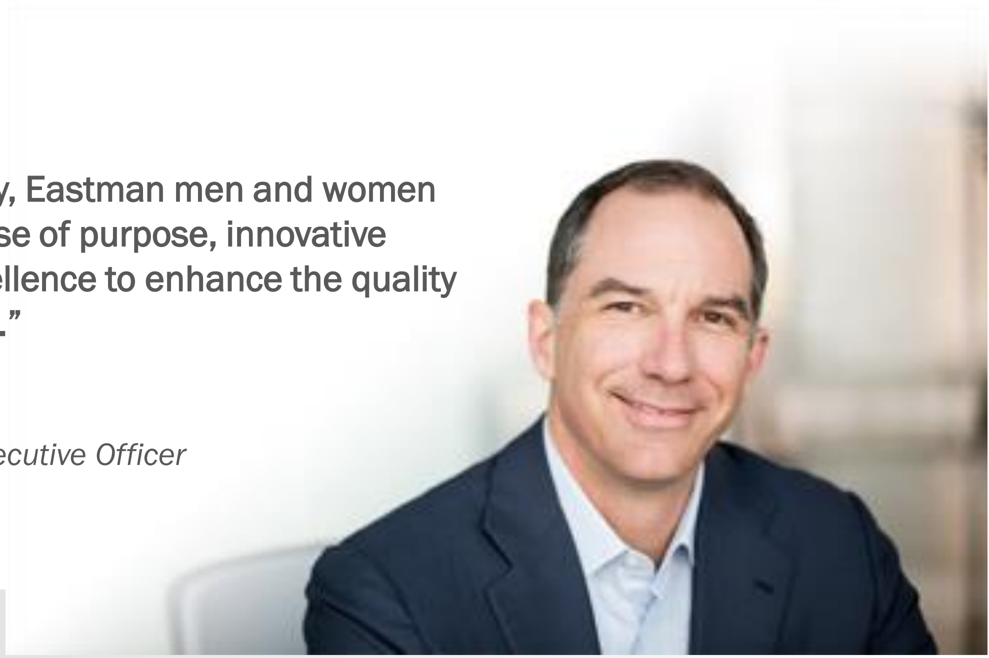
- Fortune 500 specialty materials company with 2021 revenue of ~$10.5B
- Global manufacturer and marketer of advanced materials and specialty additives
- Operates four business segments
- Global team of ~14,000
- Serving customers in >100 countries

A legacy of **innovation** and growth

A LEGACY THAT BEGAN MORE THAN A CENTURY AGO

"Throughout our history, Eastman men and women have focused their sense of purpose, innovative spirit and drive for excellence to enhance the quality of life in a material way."

– *Mark Costa*
Board Chair and Chief Executive Officer



EASTMAN

*Making life **safer***

*Making life **easier***

*Making life **better***

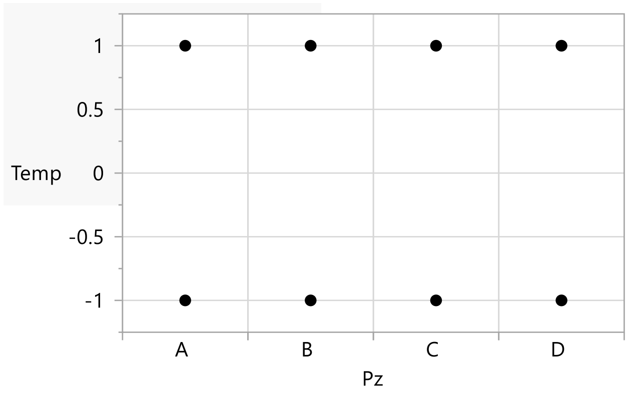
Enhancing the quality of life in a material way

EASTMAN

Designed Experiments Come in Various Shapes and Sizes

Factorial and Fractional Factorial

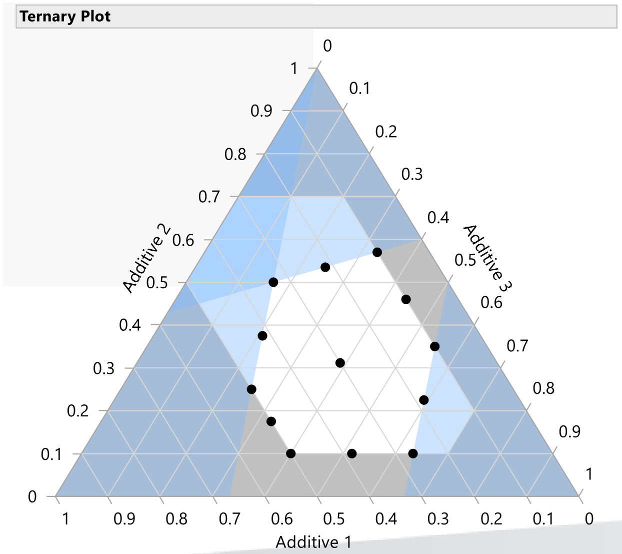
A 2x4 Factorial Experiment



8 runs + replicates

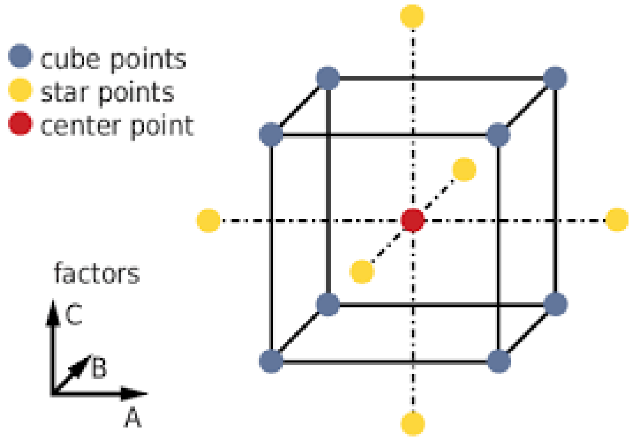
Mixture Experiments
Components Add to A Constant

A 3-component mixture experiment with constraints
13 runs + replicates



Response Surface Design

3 Factor Central Composite



15 runs + replicates

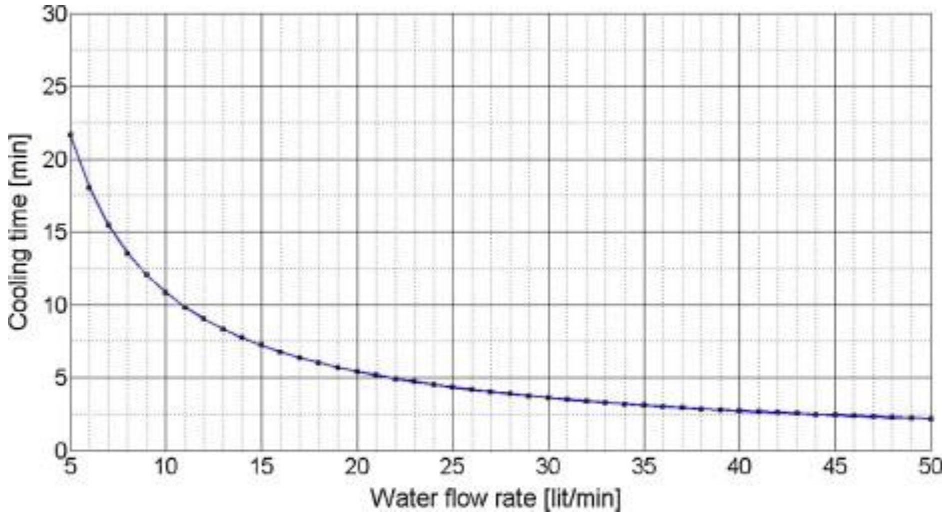
EASTMAN

Responses for Experiments Also Come in Various Shapes and Sizes

- A single measurement on a sample
- An average of several measurements on a sample
- An average of several samples taken at the same conditions
- A standard deviation of several measurements or several samples at the same condition
- Censored data (Examples: drop height, when the sample does not break at that height and test values below the detection limit)

Responses for Experiments Also Come in Various Shapes and Sizes

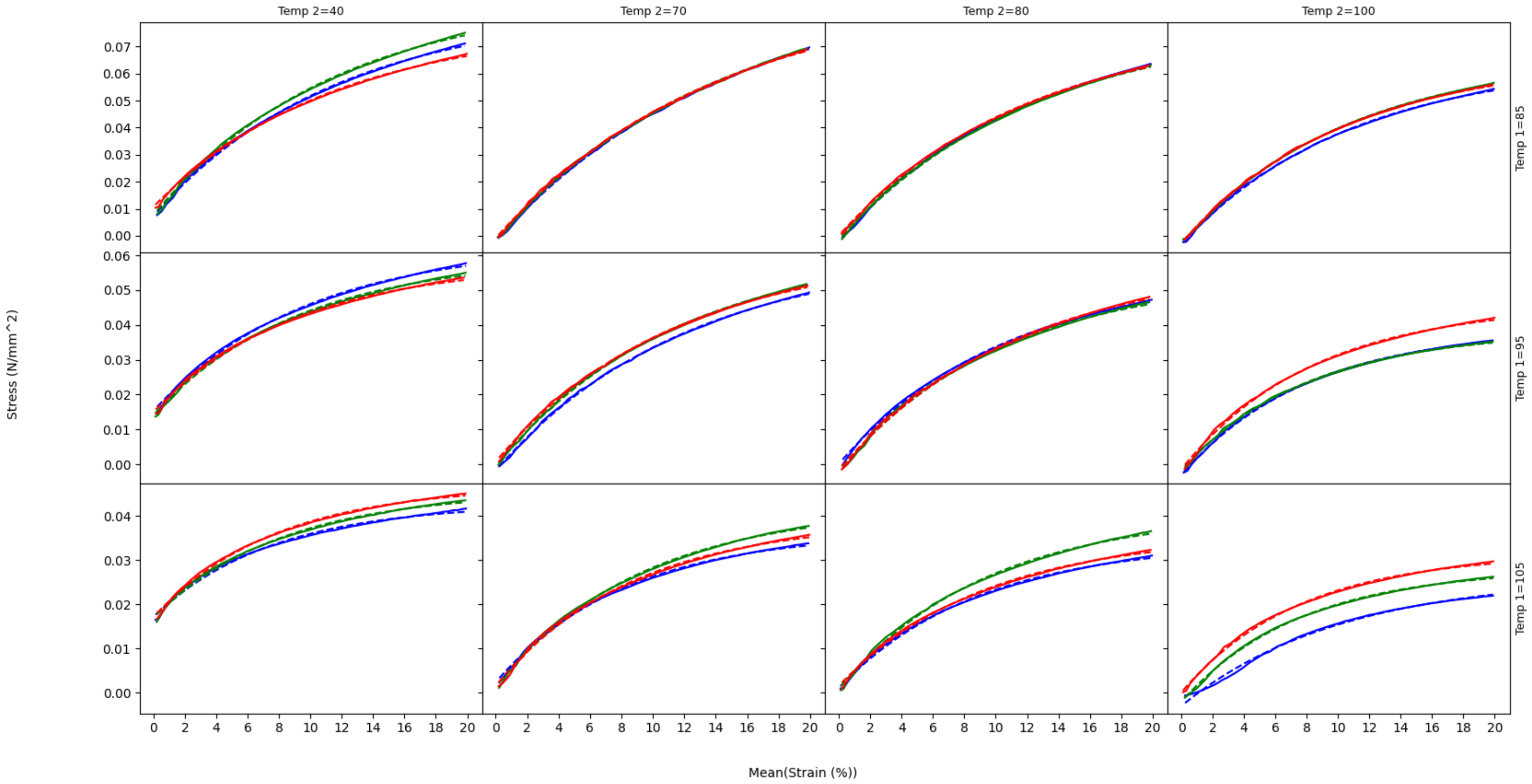
- A curve
  - Stress-Strain Curve – take the stress at a given strain level
  - Response over time – take the response at a given time or at the end of x hours
  - Viscosity as temperature rises or falls – take the response at a given temperature, or the temperature at a given viscosity as the response
  - Frequency response curve in sound– use the value at a certain frequency as the response
  - Dose-Response curves in drug development – use dose to achieve a certain %response



3 x4 x3 = 36 curves

The Power of the Curve

Experimentation to determine the effect of two processing temperatures on the **Stress@ 20 Strain**.



3x4 factorial with 3 complete replicates

EASTMAN

Acknowledgements

- Stijn Van de Vyver – Eastman Chemist
- Veerle De Praeter – Eastman Application Service Lab Technician
- Hank Anderson – Stat-Ease
- Martin Bezener - Stat-Ease
- Derek Harding – Eastman Data Scientist

EASTMAN

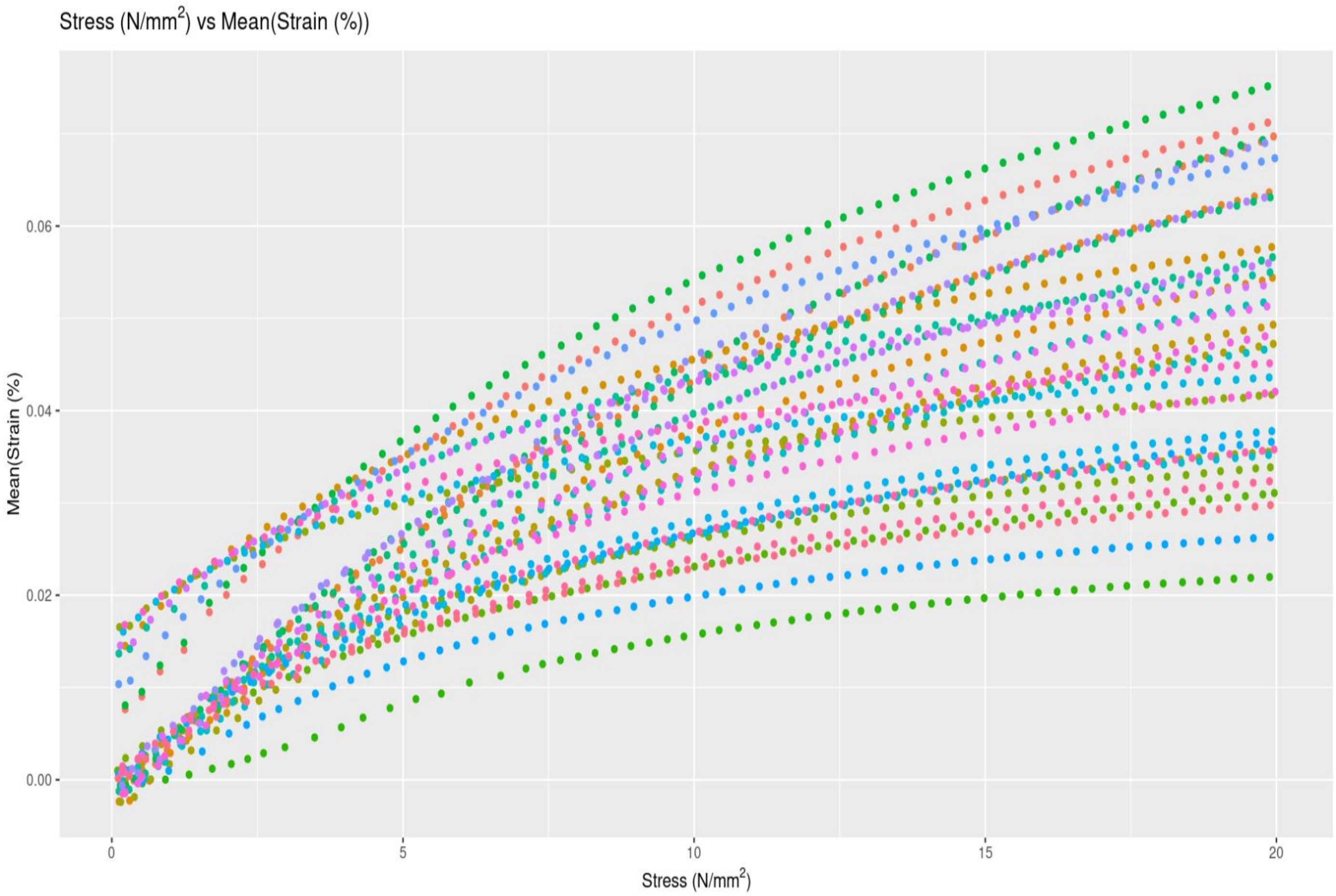
The Power of the Curve

- **Original response of interest: Stress at a Strain of 20%**
- What if we could develop a model to understand how the two processing temperatures affect the Stress/Strain curve?
- What if we could understand which aspects of the curve are most affected by which variable?
- What if we could predict the Stress/Strain curve anywhere inside our design space?
- And the predictions will be within +- 0.005 of the actual stress value for the conditions we have.

Process of Analysis

- Decide what curve to fit for each run
- Fit each run and save the parameters
- Fit the parameters as a function of the variables in the study
- Decide what the optimum combination of the parameters are
- Draw a curve with the optimum set of parameters

What curve should we fit?



To make this decision use process knowledge, literature searches, trying different fits.

Sometimes the non-linear parameters of the model can be scientific information, such as rates and asymptotes that mean something.

If nothing seems right, spline fits are a great option in software today, but may not provide as much scientific insight as non-linear models.

Functional data analysis (FDA) is a branch of statistics that analyzes data providing information about curves, surfaces or anything else varying over a continuum.

What curve should we fit?

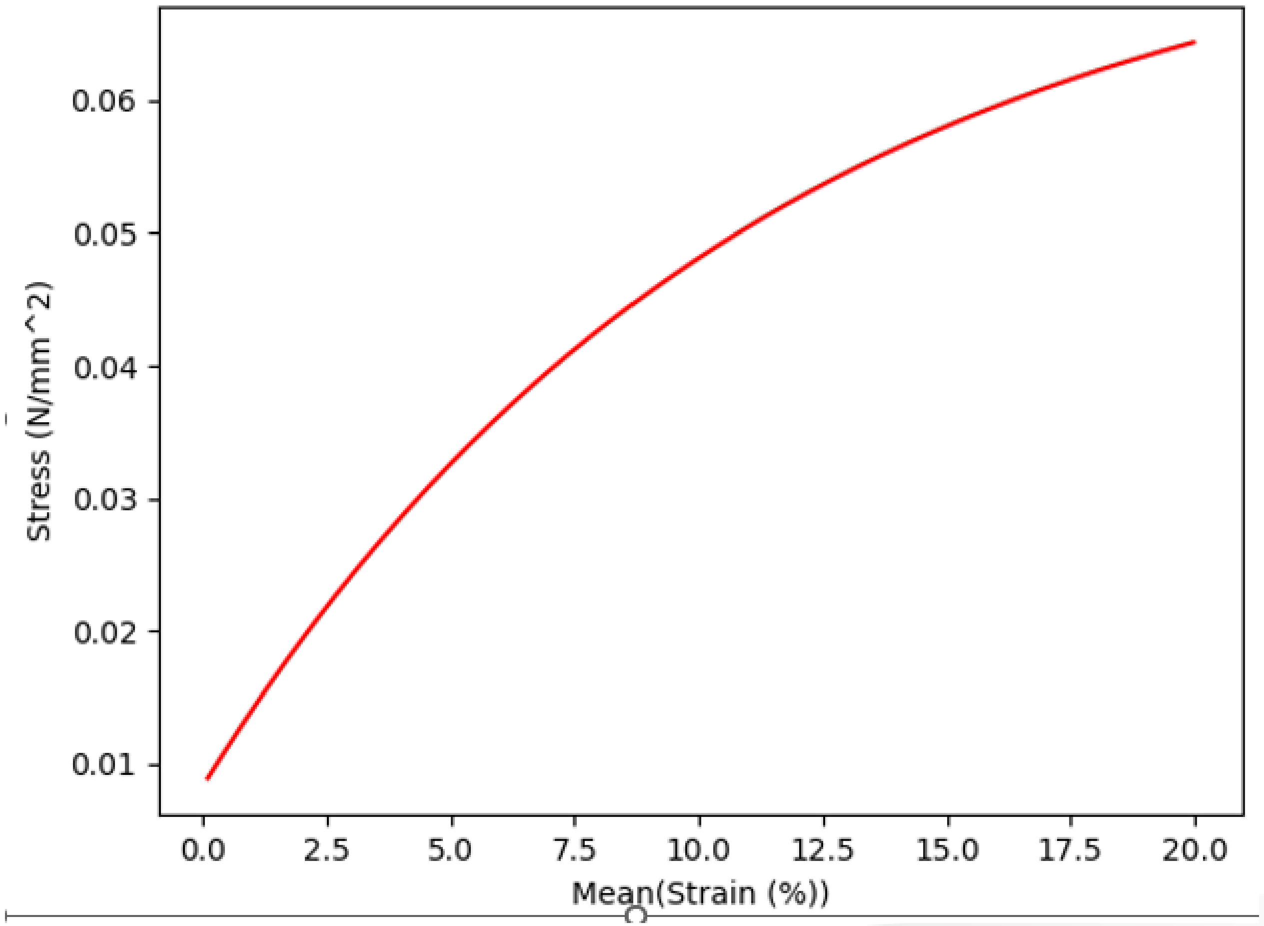
Prediction Model
Stress = a - b * Exp(-c*Strain)

a = Asymptote
-b = Scale
-c = Growth Rate

Starting Point = Asymptote + Growth Rate
Starting point = Stress at Strain = 0.

36 curves

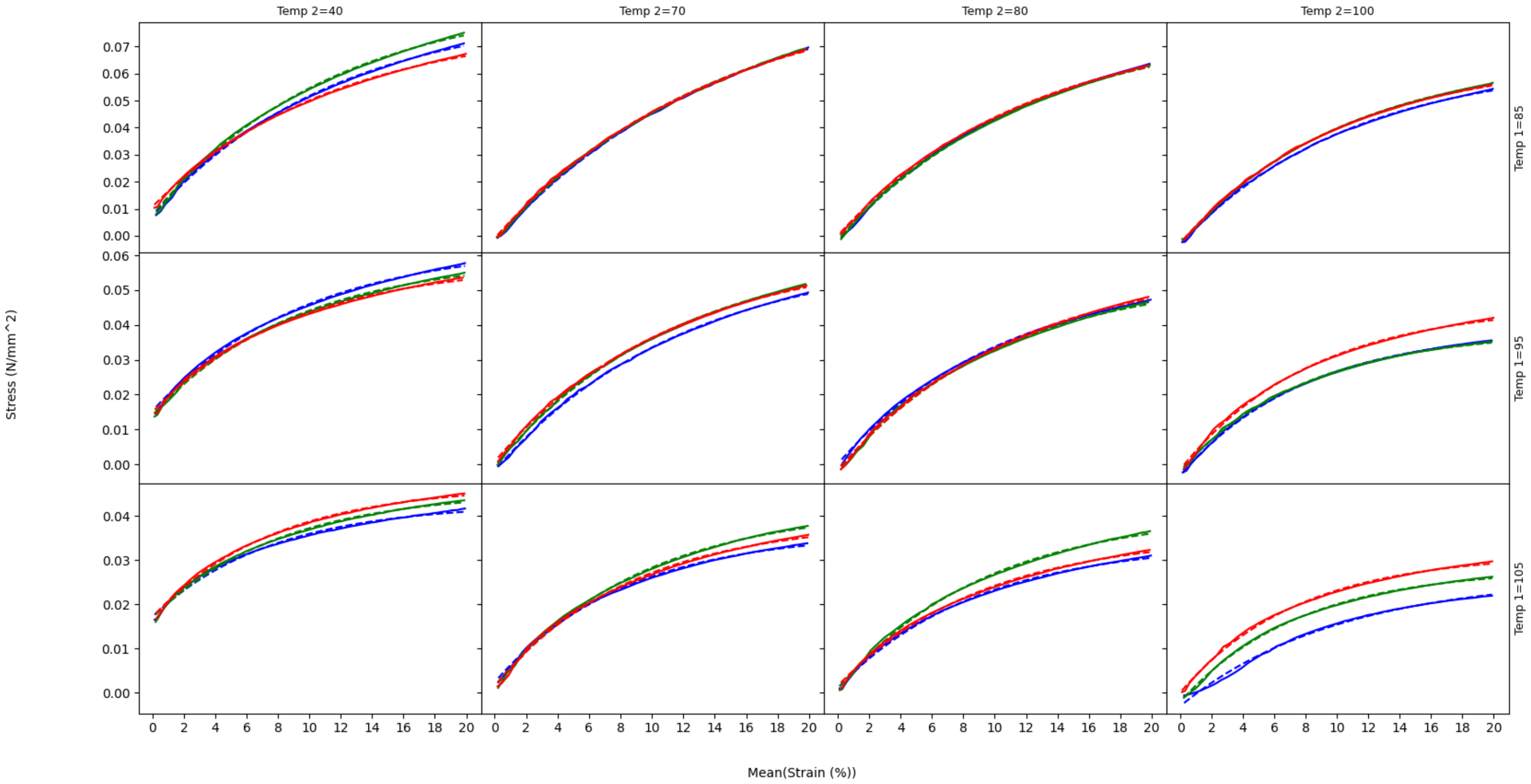
[Useful Nonlinear Models in Python • Juliano Garcia (robotenique.github.io)]



The Power of the Curve

3 x4 x3 = 36 curves

Experimentation to determine the effect of two processing temperatures on the **Stress@ 20 Strain**.



3x4 factorial with 3 complete replicates

Parameter Estimates Obtained for all 36 curves

		Factor 1	Factor 2	Response 1	Response 2	Response 3	Response 4	Response 5
Group	Run	a:Temp 1	B:Temp 2	Asymptote	Scale	Growth rate	Starting Point	Stress at 20% Strain
		°C	°C					
1	1	85	40	0.0841	-0.0767	-0.0865	-0.0024	0.0705
1	2	85	70	0.0936	-0.0956	-0.0683	0.0253	0.0692
1	3	85	80	0.0787	-0.0804	-0.0821	-0.0034	0.0631
1	4	85	100	0.0644	-0.0675	-0.0929	-0.0286	0.0539
2	5	95	40	0.0631	-0.0478	-0.1025	-0.0393	0.0570
2	6	95	70	0.0610	-0.0628	-0.0826	-0.0215	0.0490
2	7	95	80	0.0548	-0.0547	-0.0953	-0.0405	0.0467
2	8	95	100	0.0388	-0.0419	-0.1238	-0.0850	0.0353
3	9	105	40	0.0428	-0.0256	-0.1324	-0.0896	0.0410
3	10	105	70	0.0364	-0.0340	-0.1218	-0.0854	0.0334
3	11	105	80	0.0340	-0.0329	-0.1125	-0.0785	0.0306
3	*12*	*105*	*100*	*0.0263*	*-0.0293*	*-0.0992*	*-0.0729*	*0.0223*

Estimates for only 12 curves shown.

Notice split plot structure of the data table with groups of Temp 1.
And Temp 2 not randomized.

EASTMAN

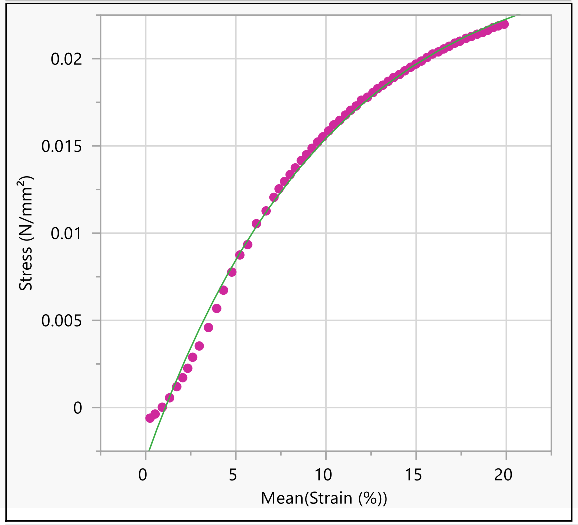
An Outlier?

Fit Curve Set=12

Model Comparison

Model		AICc	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-752.0763	-744.4262	1.0934e-5	1.9183e-7	0.000438	0.9962184

Plot



Exponential 3P

Parameter Estimates

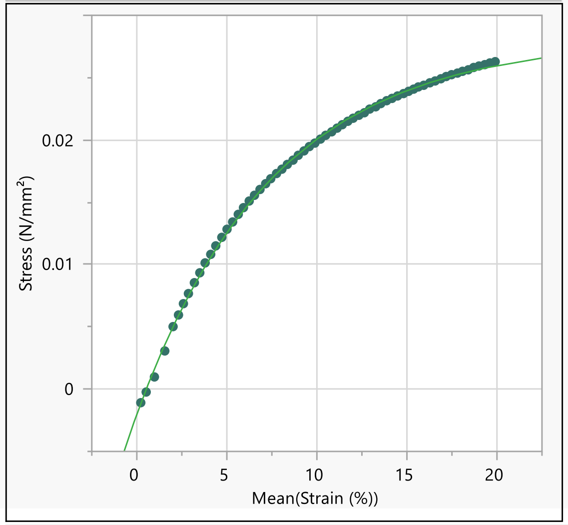
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.02631	0.0004231	3867.6207	<.0001 *	0.0254808	0.0271392
Scale	-0.029302	0.0003474	7115.3856	<.0001 *	-0.029983	-0.028622
Growth Rate	-0.099247	0.0034148	844.69812	<.0001 *	-0.105939	-0.092554

Fit Curve Set=24

Model Comparison

Model		AICc	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-894.9476	-886.9901	2.7659e-6	4.5343e-8	0.0002129	0.9991809

Plot



Exponential 3P

Parameter Estimates

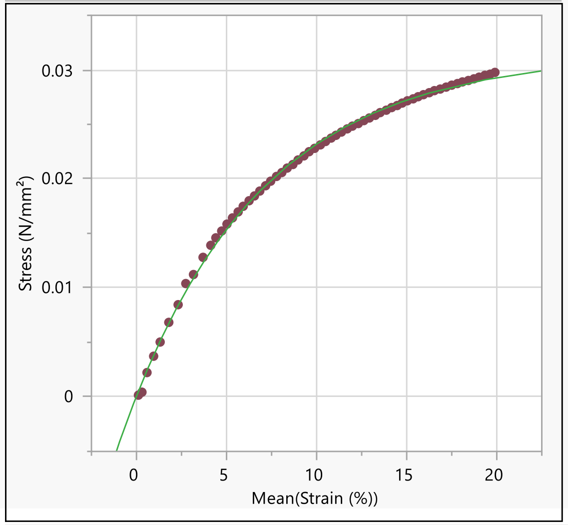
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.0281214	0.0001221	53033.253	<.0001 *	0.0278821	0.0283608
Scale	-0.030139	0.0001137	70207.921	<.0001 *	-0.030362	-0.029916
Growth Rate	-0.131964	0.0016016	6788.621	<.0001 *	-0.135103	-0.128825

Fit Curve Set=36

Model Comparison

Model		AICc	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-832.5864	-824.7035	5.8543e-6	9.7572e-8	0.0003124	0.9984995

Plot



Exponential 3P

Parameter Estimates

Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.031464	0.0001715	33663.97	<.0001 *	0.0311279	0.0318001
Scale	-0.031383	0.0001663	35605.987	<.0001 *	-0.031709	-0.031057
Growth Rate	-0.13344	0.0021347	3907.3199	<.0001 *	-0.137624	-0.129256

EASTMAN

How good do the models fit? An Example.

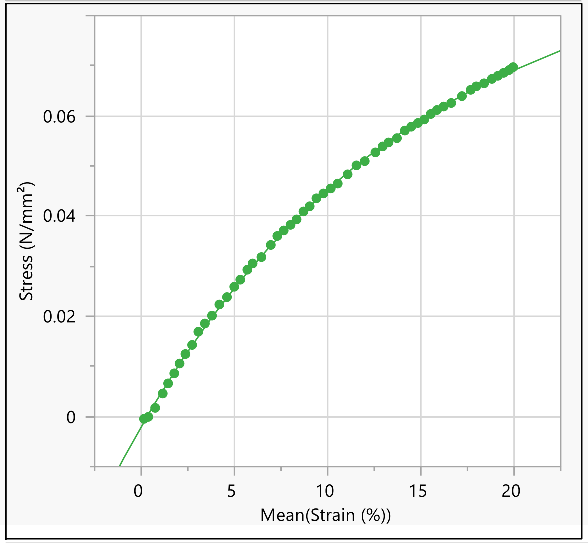
Temp 1=85, Temp 2 = 70
Runs 2, 14, 26

Fit Curve Set=2

Model Comparison

Model		AICc ^	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-687.6352	-680.4956	7.9198e-6	1.5529e-7	0.0003941	0.9996702

Plot



Exponential 3P

Parameter Estimates

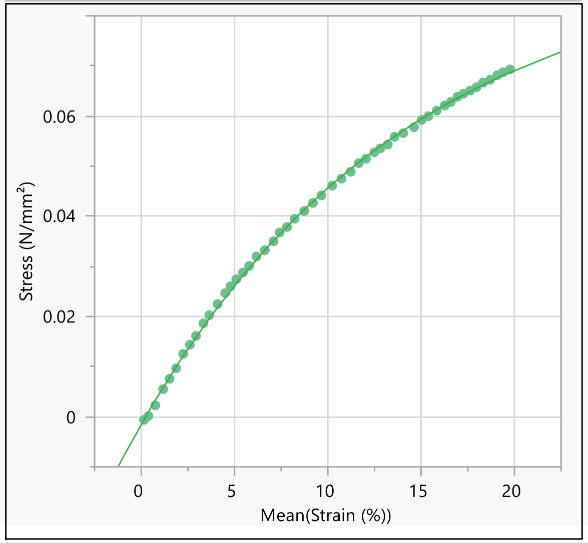
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.0935751	0.0007694	14792.662	<.0001 *	0.0920672	0.0950831
Scale	-0.095567	0.0006758	19999.459	<.0001 *	-0.096892	-0.094243
Growth Rate	-0.068286	0.0010246	4442.0288	<.0001 *	-0.070294	-0.066278

Fit Curve Set=14

Model Comparison

Model		AICc ^	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-625.7976	-618.9399	1.1764e-5	2.4507e-7	0.000495	0.9994732

Plot



Exponential 3P

Parameter Estimates

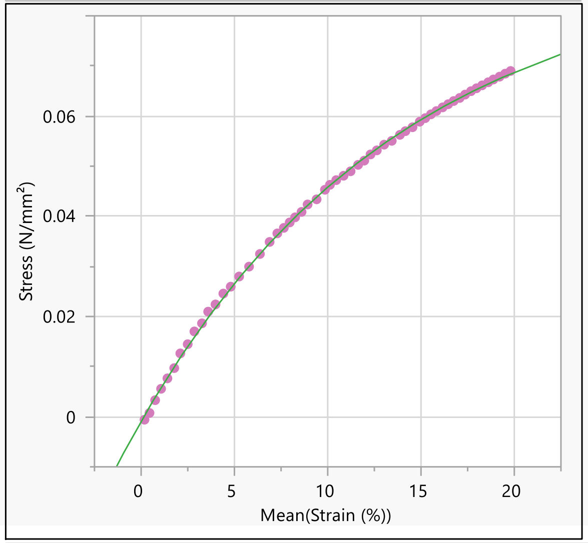
Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.0922522	0.000976	8933.5839	<.0001 *	0.0903392	0.0941652
Scale	-0.093731	0.000853	12074.69	<.0001 *	-0.095403	-0.09206
Growth Rate	-0.069947	0.0013729	2595.8254	<.0001 *	-0.072638	-0.067256

Fit Curve Set=26

Model Comparison

Model		AICc ^	BIC	SSE	MSE	RMSE	R-Square
Exponential 3P		-694.914	-687.5969	1.1438e-5	2.1581e-7	0.0004646	0.9995087

Plot

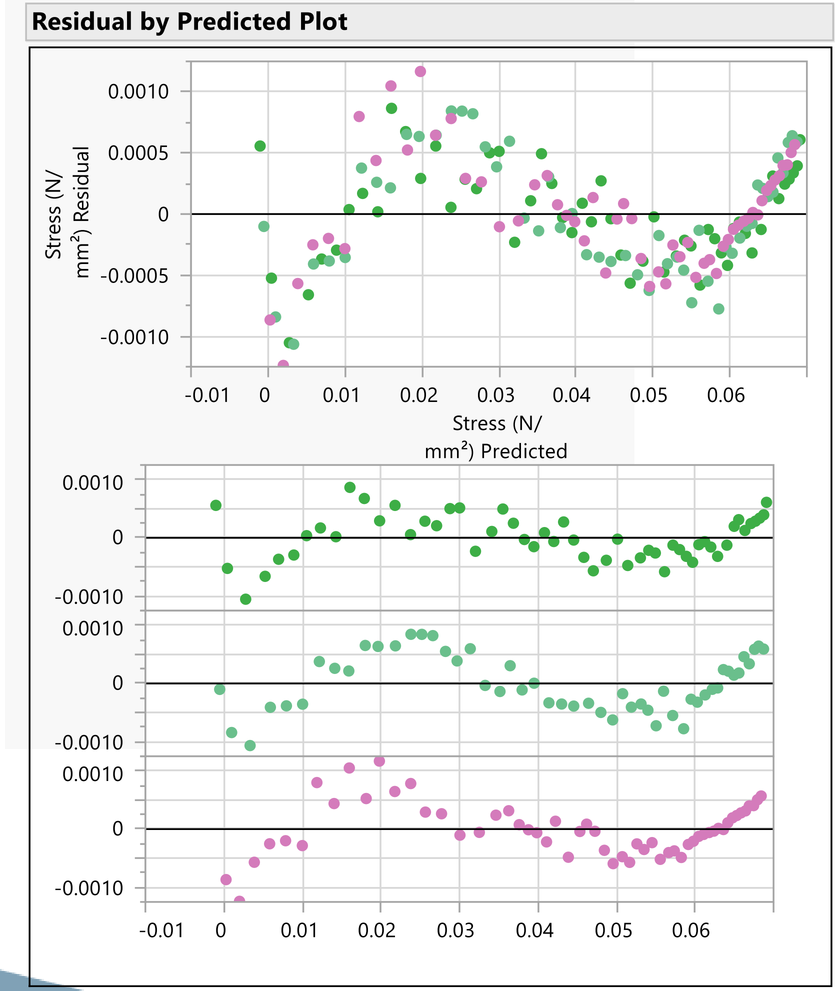


Exponential 3P

Parameter Estimates

Parameter	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Asymptote	0.0908299	0.0008427	11617.334	<.0001 *	0.0891782	0.0924816
Scale	-0.091724	0.000735	15572.954	<.0001 *	-0.093164	-0.090283
Growth Rate	-0.071176	0.0012495	3244.7052	<.0001 *	-0.073625	-0.068727

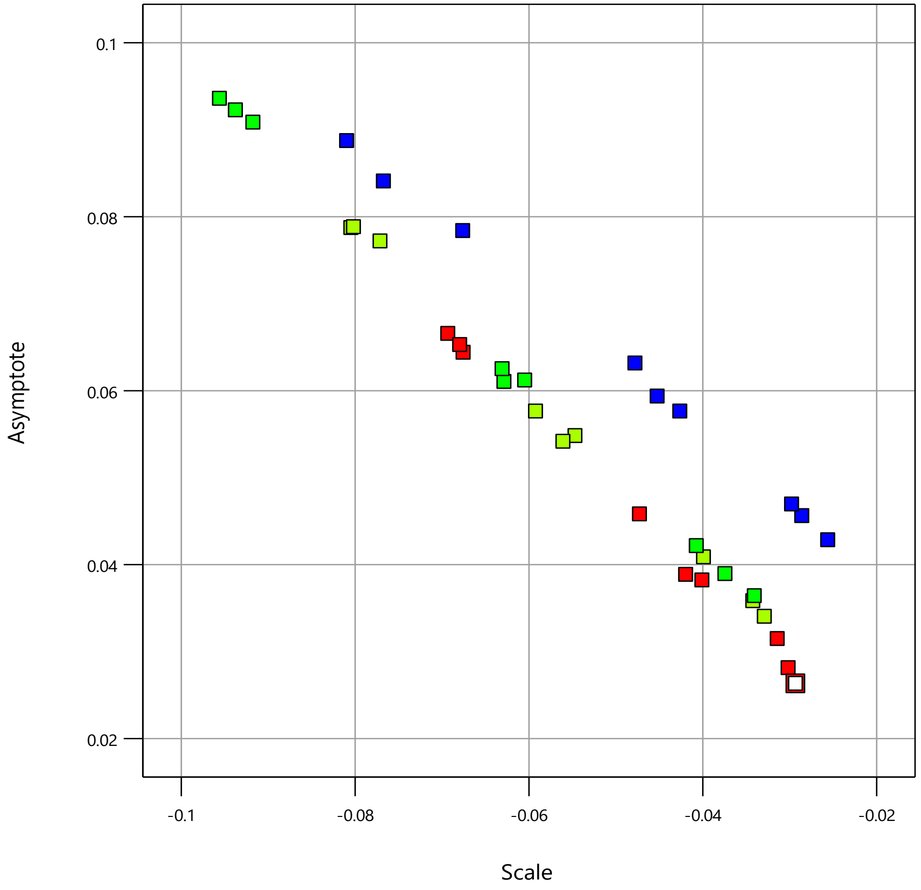
How good do the models fit? An Example.



Temp 1=85, Temp 2 = 70
Runs 2, 14, 26

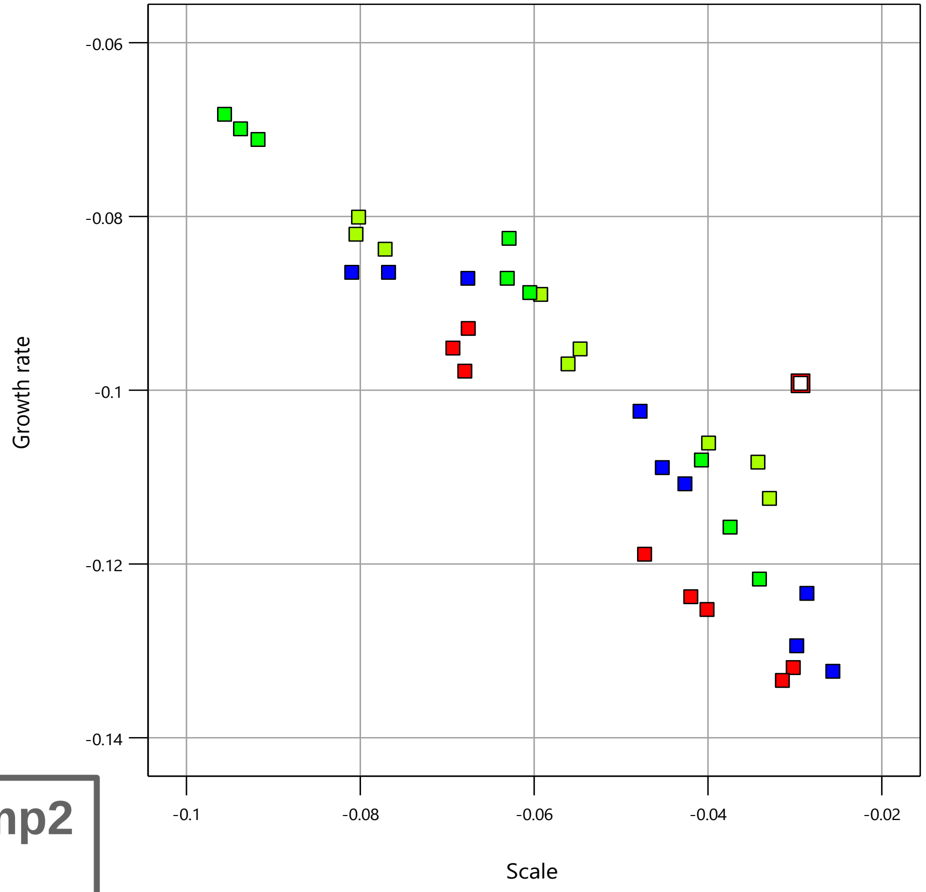
A little systematic error left, but the predictions are +- 0.0015.
R-squared > 0.99.

Relating the Exponential Parameters to One Another



Correlation = -0.946

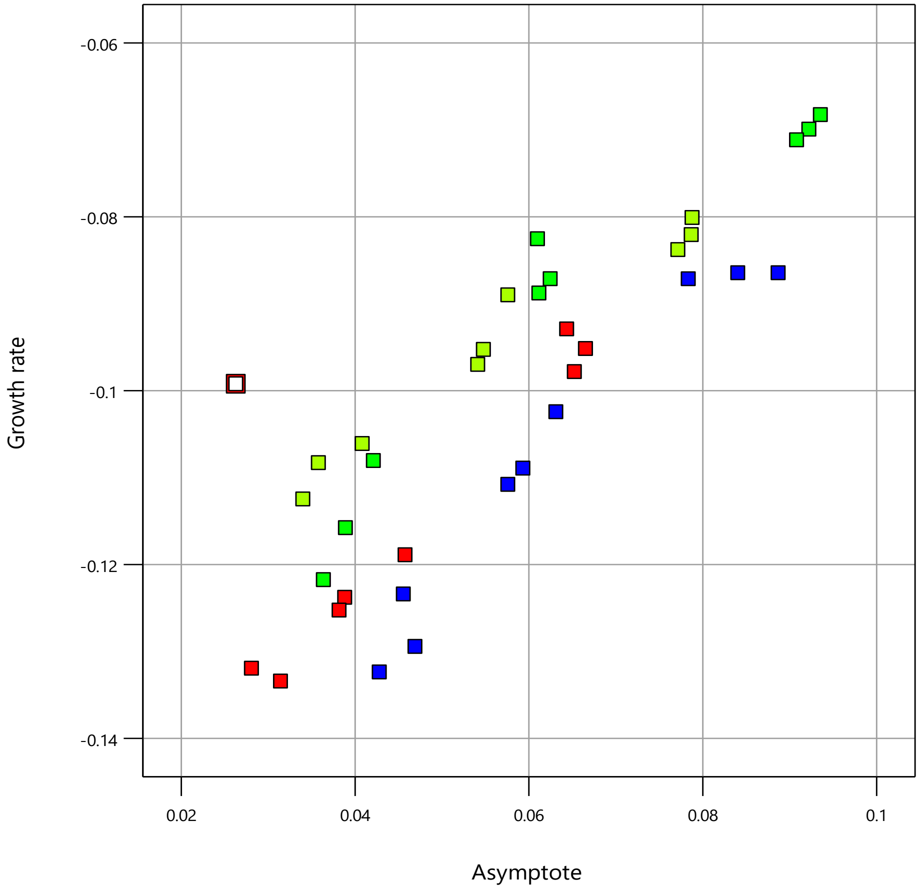
Color	Temp2
Blue	40
Green	70,80
Red	100



Correlation = -0.910

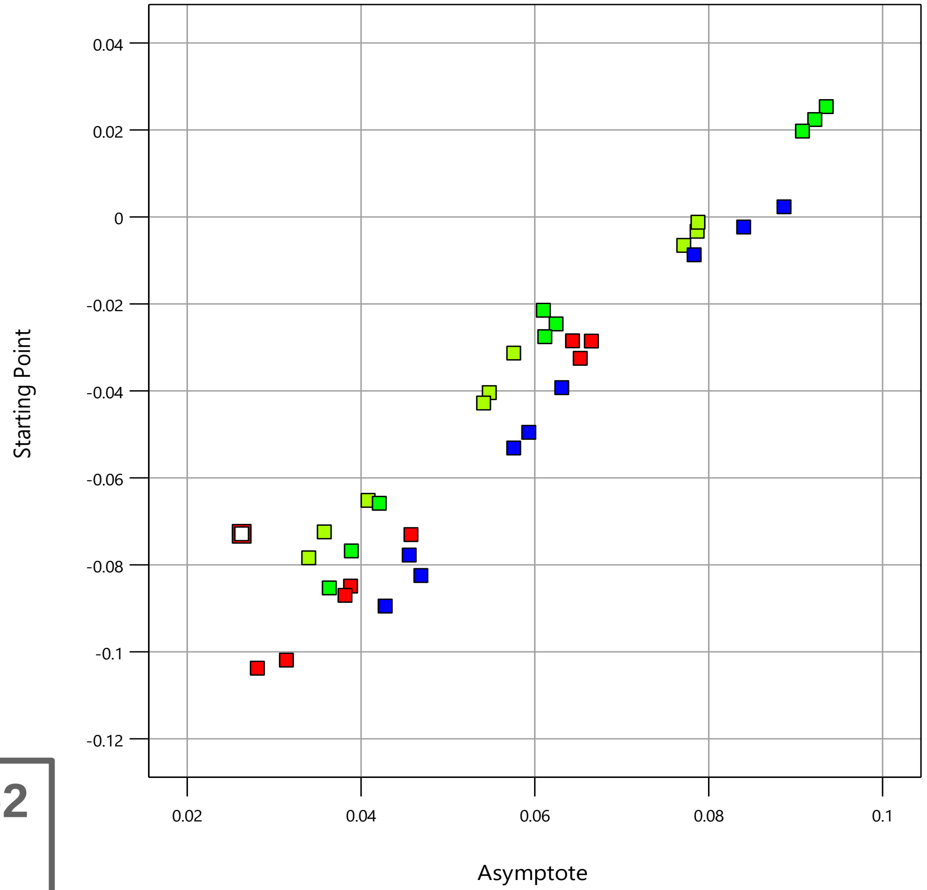
EASTMAN

Relating the Exponential Parameters to One Another



Correlation = 0.849

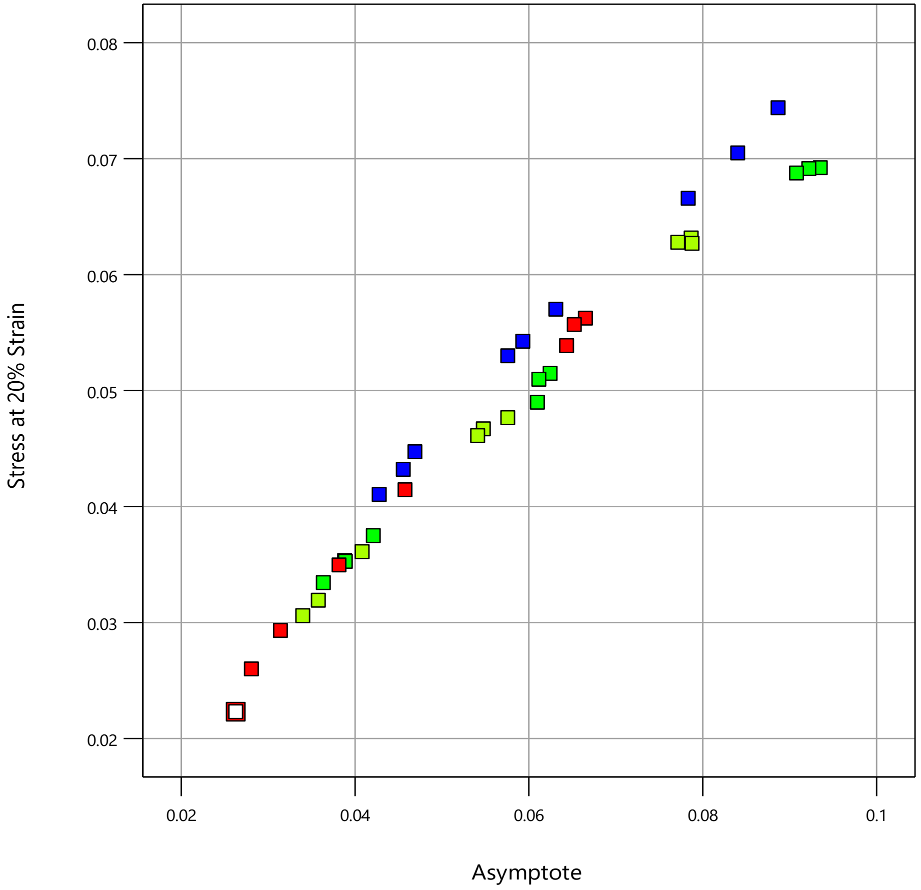
Color	Temp2
Blue	40
Green	70,80
Red	100



Correlation = 0.963

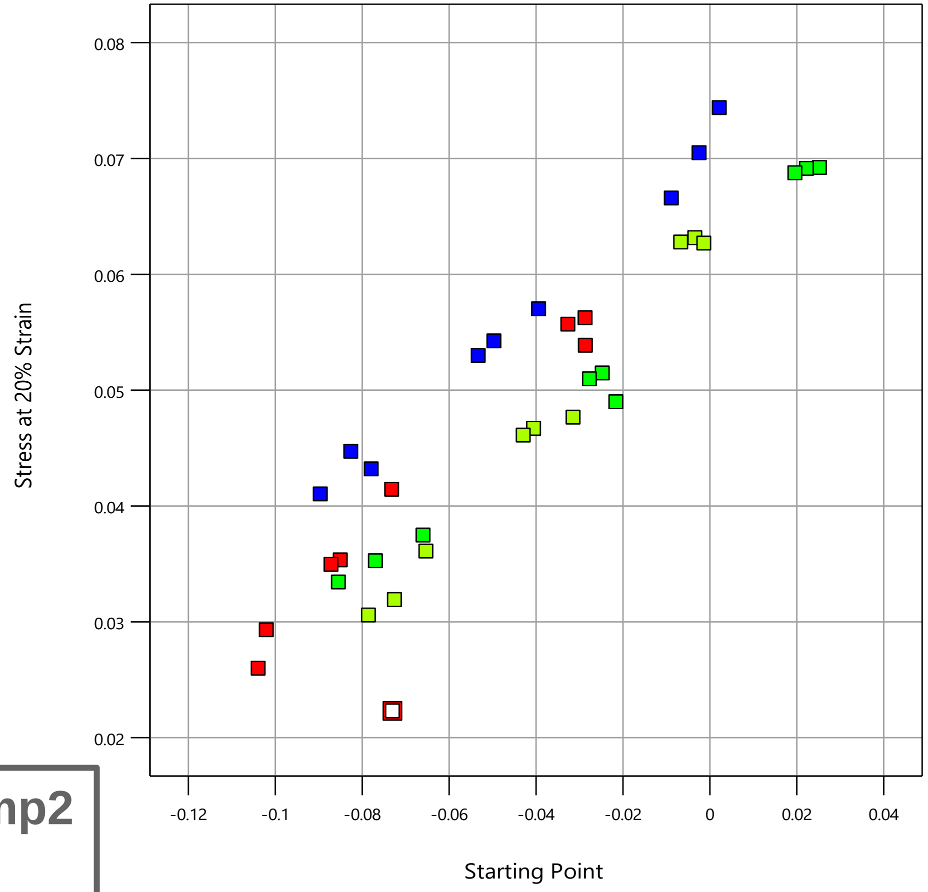
EASTMAN

Relating the Exponential Parameters to One Another



Correlation = 0.985

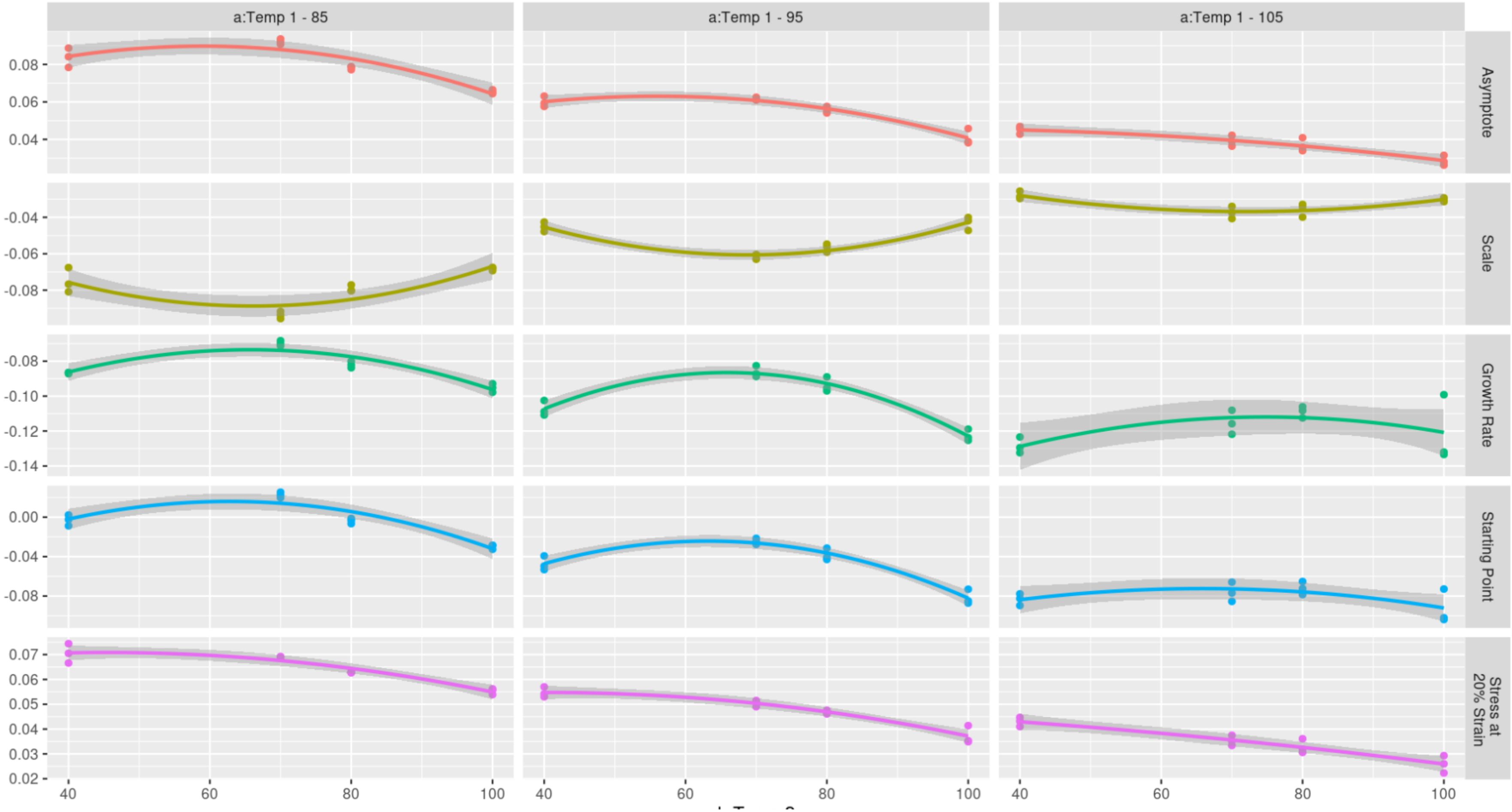
Color	Temp2
Blue	40
Green	70,80
Red	100



Correlation = 0.917

EASTMAN

A Great View of the Raw Data Table of Parameter Estimates



An individual quadratic fit of the parameter vs Processing Temperature1 is made to each of the three Processing Temperature 2's.

Can you see the effects?

$Temp1^2$ Quadratic effect of temp1

$Temp2^2$ Quadratic effect of temp1

Temp1* $Temp2^2$ – The degree of Quadratic effect of temp1 depends on the level of temp2

EASTMAN

Modeling in Stat-Ease 360

Show models live.

- Used split-plot analysis approach.
- Base model was

Source
Whole-plot
a-Temp 1
a^2
Subplot
B-Temp 2
aB
B^2
a^2B
aB^2

- Model was reduced as appropriate using backward regression.
- All models had R-squared>0.93 and R-squared adjusted > 0.92
- Transformation were considered but not needed when adding the aB^2 term.

EASTMAN

Model Coefficients

Coefficients Table

p-value shading: **$p < 0.05$** $0.05 \le p < 0.1$ $p \ge 0.1$

	Intercept	a	B	aB	a^2	B^2	a^2B	aB^2
Asymptote	0.0598908	**-0.0242506**	**-0.00957755**	0.0011293	**0.00421981**	**-0.00866812**	0.000681229	**0.00579371**
p-values		**< 0.0001**	**< 0.0001**	0.2867	**0.0021**	**< 0.0001**	0.6957	**0.0009**
Scale	-0.0591525	**0.0258346**	0.00153807	**-0.0027489**	**-0.00420093**	**0.0137787**		**-0.00474039**
p-values		**< 0.0001**	0.1290	**0.0329**	**0.0076**	**< 0.0001**		**0.0153**
Growth rate	-0.090602	**-0.0191023**	**-0.00495619**			**-0.0218372**		
p-values		**< 0.0001**	**0.0005**			**< 0.0001**		
Starting Point	-0.0277879	**-0.0405879**	**-0.0145747**			**-0.0311242**		
p-values		**< 0.0001**	**< 0.0001**			**< 0.0001**		
Stress at 20% Strain	0.0496747	**-0.0160782**	**-0.00818191**	-1.43812E-05	0.00218338	**-0.00320084**		**0.00222346**
p-values		**< 0.0001**	**< 0.0001**	0.9823	0.0560	**0.0004**		**0.0301**

In these models, hierarchy was followed so if a higher order term was statistically significant, then a lower order term is included regardless of its significance.

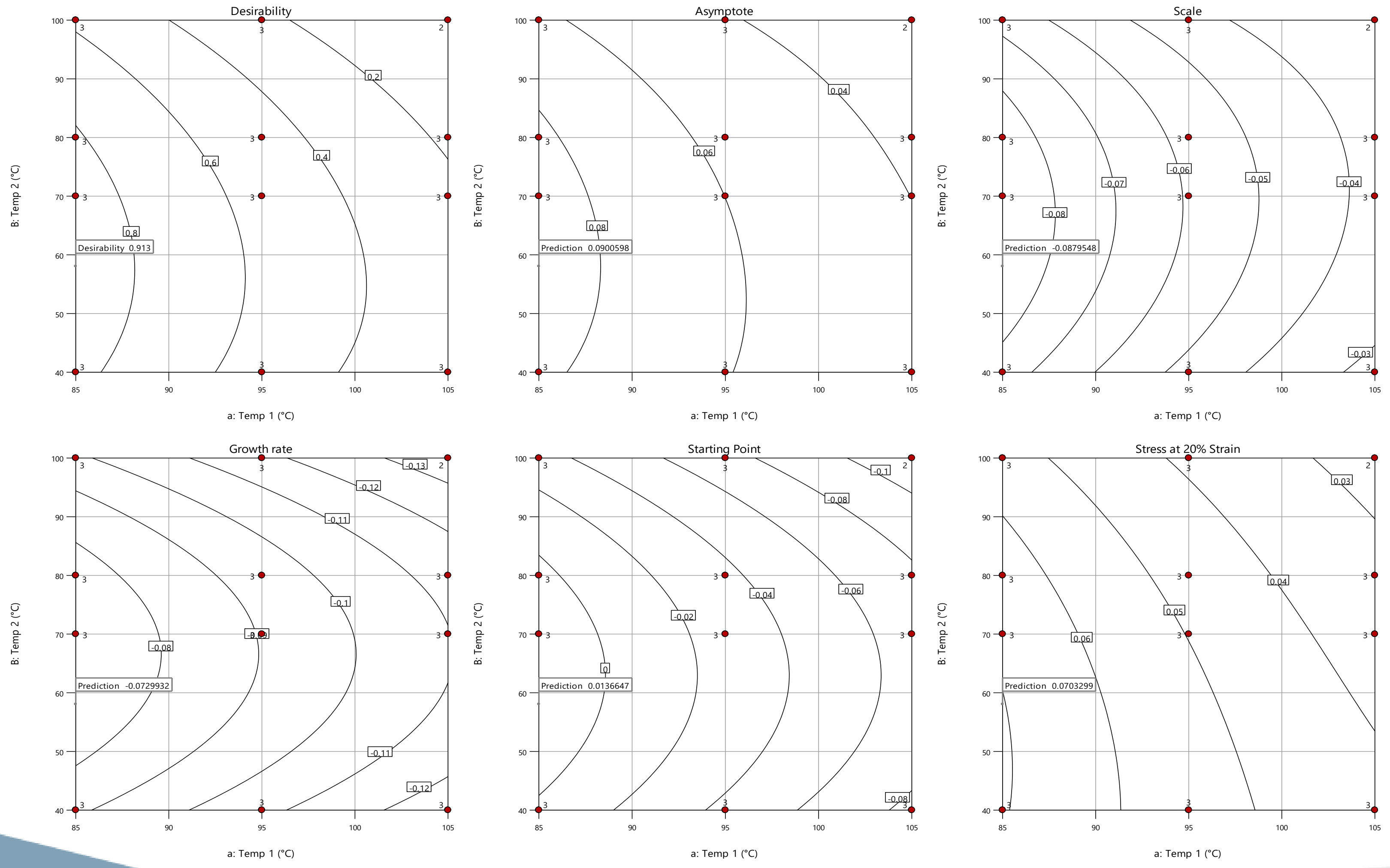
EASTMAN

Objective

- Maximize the starting point parameters, as well as the stress at 20% strain.

Number	Temp 1	Temp 2	Asymptote	Scale	Growth rate	Starting Point	Stress at 20% Strain	Desirability
1	**85.000**	**58.043**	**0.090**	**-0.088**	**-0.073**	**0.014**	**0.070**	**0.913**
2	85.000	57.788	0.090	-0.088	-0.073	0.014	0.070	0.913
3	85.000	58.451	0.090	-0.088	-0.073	0.014	0.070	0.913
4	85.000	57.387	0.090	-0.088	-0.073	0.013	0.070	0.913
5	85.000	58.991	0.090	-0.088	-0.073	0.014	0.070	0.913
6	85.000	56.828	0.090	-0.088	-0.074	0.013	0.070	0.913
7	85.000	56.363	0.090	-0.087	-0.074	0.013	0.071	0.913
8	85.000	55.021	0.090	-0.087	-0.074	0.012	0.071	0.911
9	85.000	54.500	0.090	-0.086	-0.075	0.012	0.071	0.911
10	85.000	65.500	0.090	-0.089	-0.071	0.014	0.069	0.902
11	85.000	46.692	0.087	-0.081	-0.081	0.005	0.071	0.887
12	85.000	45.983	0.087	-0.081	-0.082	0.005	0.071	0.884
13	85.000	45.483	0.087	-0.080	-0.082	0.004	0.071	0.881
14	85.000	74.500	0.087	-0.088	-0.073	0.010	0.067	0.860
15	85.000	80.500	0.083	-0.085	-0.076	0.004	0.064	0.814

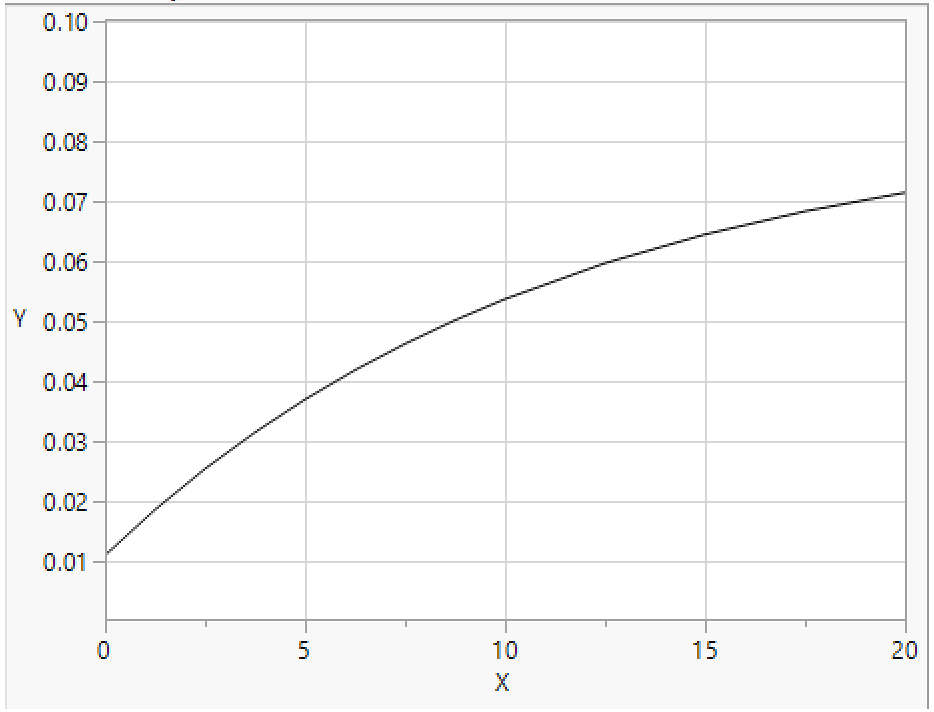
Optimum Area



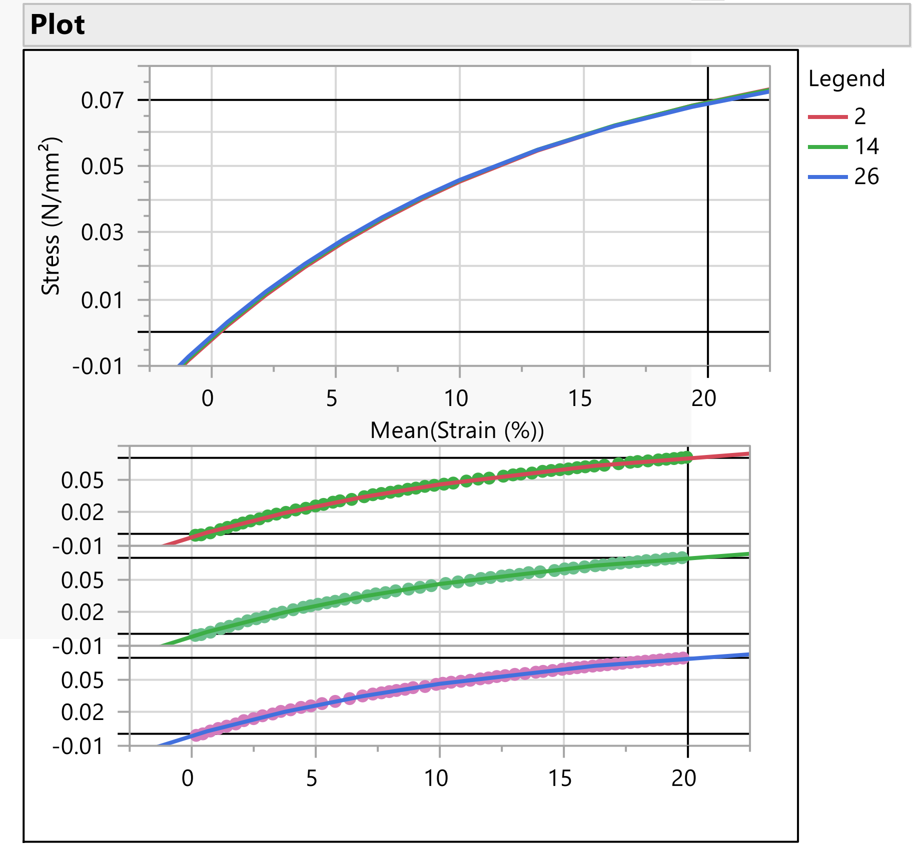
If the optimum is in a corner or edge, the optimum is not in the corner or edge.

Optimum Parameters

Temp 1=85, Temp 2 = 58



Temp 1=85, Temp 2 = 70
Runs 2, 14, 26



Starting point can be a little better at the new optimum.

Learnings

- **Original response of interest: Stress at a Strain of 20%**
- By recognizing the power of the curve, we can learn a lot more than our original objective.
- With this new knowledge we now have power over the curve, and can change the shape, starting values, asymptote and stress @ 20% strain, to what we want.