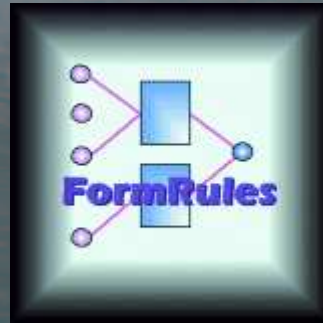


# FormRules



Better Products **Faster** – Finding  
Models & Rules for  
Product Formulation

# What is FormRules?

- A Windows-based software package developed by Intelligensys
  - Designed to be used by product formulators themselves
- 'Learns' cause-and-effect relationships within experimental data, automatically
  - and expresses these as **rules** in the form IF (ingredients are)...THEN (properties are...)

# Why? Worldwide Competition Requires Better Products... *Faster*

- improved product properties
- consistent quality
- more efficient processing
- lower cost
- shorter time to market
- conforming to new environmental constraints and legislation

# This means...

- You might need to formulate in new product domains, where you have not had time to develop a deep expertise
  - And you can't afford to spend that time!
- You want to accelerate the process of finding out about your new formulations and requirements

# FormRules can help you...

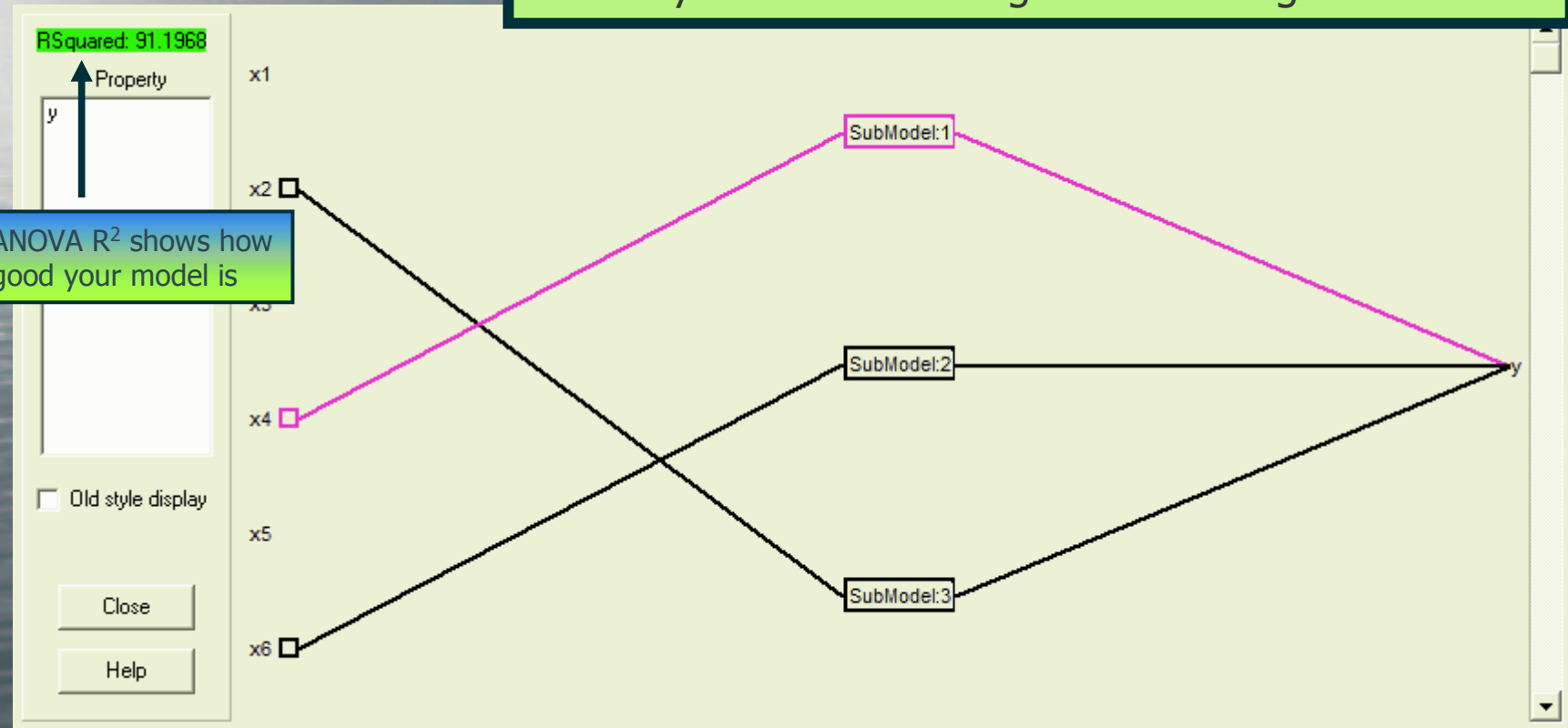
- Extract maximum knowledge from your data quickly and easily
  - Find which inputs (ingredients and process conditions) REALLY affect the properties of your product/formulation
  - Find how changes in ingredients change properties
- Make sure you are measuring and controlling **all** important variables – reduces 'hit and miss'
- Eliminate unnecessary ingredients
- Decide where to take your next experiments to get the most fruitful results

# How does FormRules do this?

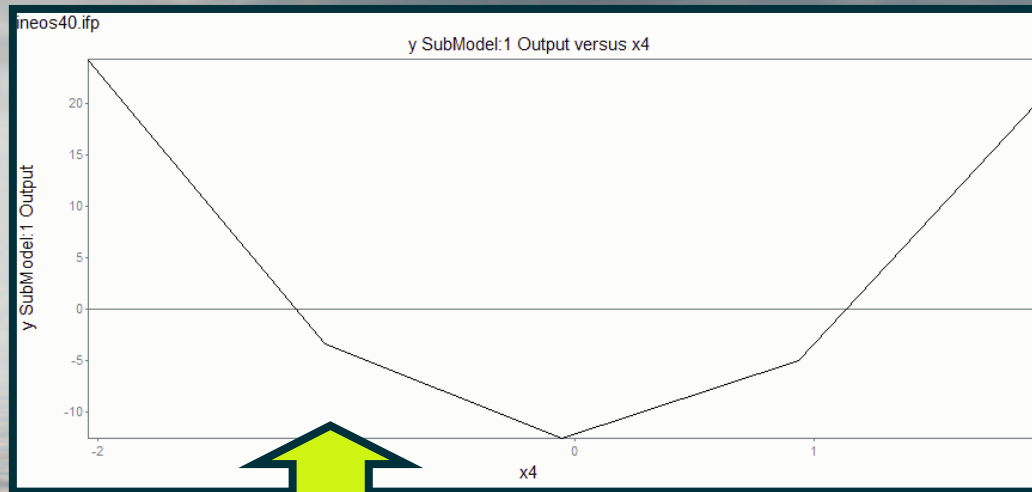
- Using a proven computational technique (neurofuzzy logic) to discover models and rules within your own experimental data
- Presenting the results clearly to you, the user – with graphical display and statistical validations to help you understand them

# What does that mean??

Here's an example. **FormRules** has identified that only x2, x4 and x6 are important for property y. x1, x3 and x5 play no role, so we don't need to worry about measuring or controlling them.

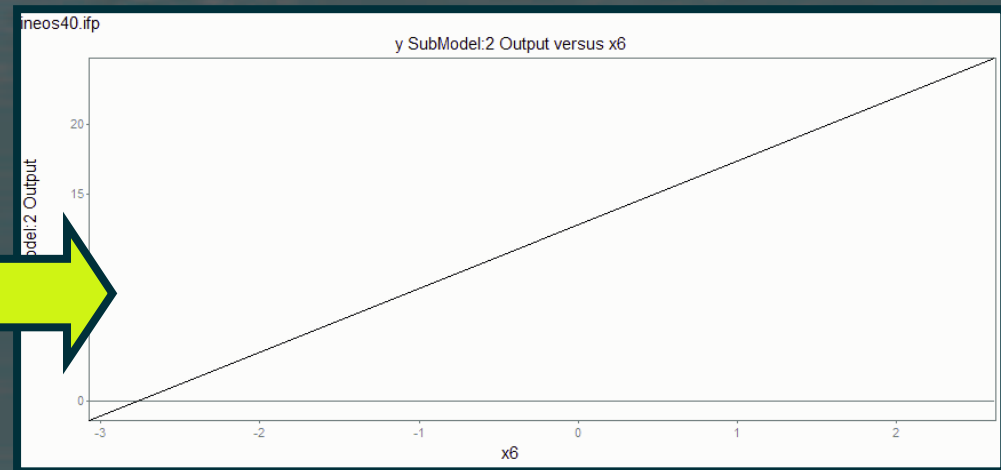


# But there's more...



The dependence on  $x_4$  is approximately quadratic (i.e.  $y$  depends on the square of  $x_4$ )

However, the dependence on  $x_6$  is linear.



# Advantages?

- You don't need to know in advance which variables are important
- You don't need to know the functional form of the dependence
- FormRules will discover all this for you **automatically** and, at most, **in minutes**
- **Later in this presentation, you'll see exactly how**

# A Hidden Advantage!

- If FormRules can't build a good cause-and-effect model from your data....
- this means that there is some 'cause' that is not being measured, so you have an uncontrolled variable that is affecting your results – a crucial piece of information for you to know!

# Who uses FormRules?

- Companies like AstraZeneca, Cabot Microelectronics, Syngenta and others...
  - Formulating diverse products in the chemicals, pharmaceuticals and healthcare fields
- Academic groups in Europe, America and Asia

# How does FormRules work?

- You collect experimental data by varying the 'recipe' (ingredients and process conditions) and measuring end-use properties
- Then import the data into FormRules
- FormRules will try many cause-and-effect models relating inputs and outputs to find the one that best represents *your* system
- Results are shown graphically, and in words as IF...AND...THEN rules

# What's inside FormRules?

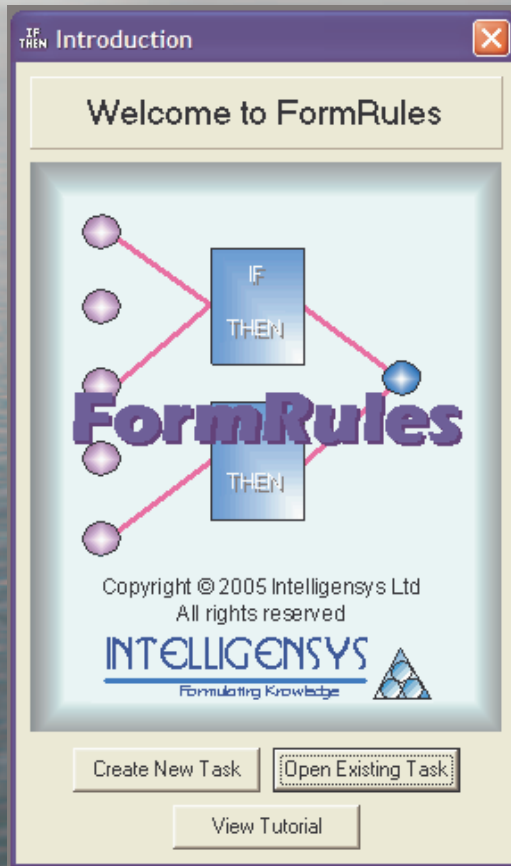
- FormRules uses neurofuzzy logic to develop cause-and-effect models
  - Works well even for non-linear and complex situations
  - Automatic 'machine learning' tool that discovers rules for itself, directly from your data
- Integrated graphics and statistics help you to interpret results easily

Let's look at an example....

# What you'll see here

- This presentation will show you some of the key aspects of FormRules, illustrated using a data set for controlled release pharmaceutical tablet formulation
- Because FormRules works on your own data, though, it can be applied much more widely to many formulations and processing situations. See the on-line application notes for some examples.

# When you launch FormRules...



- When you first enter FormRules, you can start a new project or load up a previous one
- Or, as a new user, you might also want to look at the on-line Tutorial
- In the next slides, we'll look at a New task... one where we have first to collect our data
  - Some issues are discussed next....

# Collecting Data – the first step

- Save your data in a spreadsheet or tab-delimited file
- Each row corresponds to a separate experiment or 'recipe'
- First row contains the 'field names' describing your inputs and outputs

In our example, the data is taken from the literature, so you can see what's involved.

# Data

- The following slide shows part of a data set for a tablet formulation, used in our example
  - 4 ingredients: Polymer A, Polymer B, Dextrose, and Lubricant
  - 3 'process' variables: hardness, % moisture, particle size
  - 10 outputs: amount of drug released at different time intervals (1 hr, 2 hrs and so on)
- 22 different formulations were made and studied

# Part of a typical data set

PolyA	PolyB	Dextrose	lubricant	Hardness	size	%moisture	1hr	2hr	4hr	6hr	8hr	10hr
28.8	9.6	89.6	2.8	11.1	156	3.88	7.7	15.6	34.5	60.7	87	99.2
57.6	19.2	179.2	5.4	9.77	156	3.88	3	11	26.9	46.4	64	85.2
86.4	28.8	268.8	2	8.97	156	3.88	6.4	11.9	25.1	40.4	57	76
19.2	19.2	89.6	2.8	9.53	141	3.85	10.8	20.6	36.9	59.4	82	96.4
38.4	38.4	179.2	5.4	8.79	141	3.85	9.6	15.7	28.7	42.8	58	74.5
57.6	57.6	268.8	2	8.67	141	3.85	5.7	11	22.1	35	50	67.3
9.6	28.8	89.6	2.8									70.5
19.2	57.6	179.2	5.4									50.4
28.8	86.4	268.8	2									47.3
28.8	9.6	38.4	1.8									72.1
57.6	19.2	76.8	3.4									77.3
86.4	28.8	115.2	4.9									77.3
19.2	19.2	38.4	1.8									101
38.4	38.4	76.8	3.4									66.4
57.6	57.6	115.2	4.9	2.96	115.6	5.16	5.5	10.1	19.9	32.2	44	59.3
9.6	28.8	38.4	1.8	2.78	129	4.93	33.2	55.4	83.5	97.2	106	105
19.2	57.6	76.8	3.4	3.15	129	4.93	16.7	24.9	35.3	41.5	48	53.9
28.8	86.4	115.2	4.9	2.28	129	4.93	8.8	13.6	20.8	25.8	31	35.4

Each row describes a separate formulation. Ingredient names, process conditions and properties are given in the first row.

You can have just a few experiments – even fewer than the number of ingredients. In that case, simpler models will be developed.

# Loading Data

- When you select New Task from the welcome window, you'll see a spreadsheet. (You'll see this on the next slide.)
- Just copy-and-paste your own data...
- You can have up to 150 inputs/outputs, and as many experiments as you want

# First, enter your data

The screenshot shows the FormRules software interface. The title bar reads 'FormRules - [Enter/Edit Data]'. The menu bar includes 'Task', 'Model', 'Use', 'Reports', 'Window', and 'Help'. The toolbar contains icons for file operations and a 'Data Sheet' button. The 'Data Sheet' menu is open, showing options like 'Import...', 'Save As...', 'Auto Adjust Column Widths', 'Freeze Data Labels', 'Freeze Record Labels', 'Clear', 'Copy', 'Paste', 'Insert Row', 'Insert Column', 'Clear Sheet', 'Delete Row', 'Delete Column', and 'Print'. The 'Graph' button is also visible. The 'OK' button is highlighted in a blue box. The spreadsheet area is empty. The Windows taskbar at the bottom shows the start button and several open applications: 'Inbox - Outlook ...', 'powdermatrix.or ...', 'Microsoft PowerP...', 'FreeCell', 'FormRules', and 'Microsoft Excel - I...'. The system clock shows '22:02'.

1. Data Sheet button gives you 2 options

- i. Copy and paste from a spreadsheet or
- ii. Import a tab-delimited file

2. If you wish, you can graph your data, or do simple statistics on it.

3. Click on the OK button to move on to the next stage.

# Then, set inputs and outputs

The screenshot shows the 'FormRules - [Set Field Types]' application window. The main area contains a table with the following data:

	Field Name	Field Type	Units	Min	Max
1	PolymerA	Ingredient		9.6000	86.4200
2	PolymerB	Not Used		9.6000	86.4200
3	Dextrose	Ingredient		32.9100	268.8000
4	lubricant	Processing Condition		1.8000	5.4000
5	Hardness	Property		1.3300	11.1000
6	Particle-size	Not Used		115.8000	156.0000
7	%moisture	Not Used		3.8500	9.0300
8	1hr	Not Used		3.0000	33.2000
9	2hr	Not Used		6.7000	55.4000
10	4hr	Not Used		16.8000	83.5000
11	6hr	Not Used		25.8000	97.2000
12	8hr	Not Used		30.8000	105.6000
13	10hr	Not Used		35.3500	105.4000
14	12hr				
15	16hr				
16	20hr				
17	24hr				

At the bottom of the window, there is a 'Quick Select' button circled in blue. A blue arrow points from this button to a text box. The task name is 'C:\Program Files\Intelligensys\FormRules\invitro2'.

Set each variable as an Ingredient, Processing Condition or Property.

The Quick Select feature lets you pick blocks of ingredients, processing Conditions and properties, making it quick to set up all of the fields.

You can also get these options by right-clicking with the mouse anywhere over the screen.

# Setting inputs and outputs

FormRules - [Set Field Types]

Task Model Use Reports Window Help

	Field Name	Field Type	Units	Min	Max
1	PolymerA	Ingredient		9.6000	86.4200
2	PolymerB	Ingredient		9.6000	86.4200
3	Dextrose	Ingredient		32.9100	268.8000
4	lubricant	Ingredient		1.8000	5.4000
5	Hardness	Ingredient		1.3300	11.1000
6	Particle-size	Ingredient		115.8000	156.0000
7	%moisture	Ingredient		3.8500	9.0300
8	1hr	Property		3.0000	33.2000
9	2hr	Property		6.7000	55.4000
10	4hr	Property		16.8000	83.5000
11	6hr	Property		25.8000	97.2000
12	8hr	Property		30.8000	105.6000
13	10hr	Property		35.3500	105.4000
14	12hr	Property		39.9000	105.4000
15	16hr	Property		50.8000	110.0000
16	20hr	Property		61.0000	119.6000
17	24hr	Property		72.3000	119.5000

Quick Select OK Cancel Help Task Name: C:\Program Files\Intelligensys\Forr Browse

When inputs (ingredients and processing conditions) and outputs (properties) are set, the screen looks like this...

Click on **OK** to move on to Data Analysis.

# Data Analysis

FormRules - C:\Program Files\Intelligensys\FormRules\invitro2 - [Data Analysis]

Task Model Use Reports Window Help

Data Sheet Graph Statistics Analyze OK Cancel Help

Complete Data Missing Data

	A	B	C	D	E	F	G	H	I	J	K	
1		PolymerA	PolymerB	Dextrose	lubricant	Hardness	Particle-size	%moisture	1hr	2hr	4hr	6hr
2	Record#1	28.8	9.6	89.6	2.8	11.1	156	3.88	7.7	15.6	34.5	60.7
3	Record#2	57.6	19.2	179.2	5.4	9.77	156	3.88	3	11	26.9	46.4
4	Record#3	86.4	28.8	268.8	2	8.97	156	3.88	6.4	11.9	25.1	40.4
5	Record#4	19.2	19.2	89.6	2.8	9.53	141	3.85	10.8	20.6	36.9	59.4
6	Record#5	38.4	38.4	179.2	5.4	8.79	141	3.85	9.6	15.7	28.7	42.8
7	Record#6	57.6	57.6	268.8	2	8.67	141	3.85	5.7	11	22.1	35
8	Record#7	9.6	28.8	89.6	2.8	6.54	116	6.65	21.1	32.6	46.6	56
9	Record#8	19.2	57.6	179.2	5.4	5.45	116	6.65	13.1	19.7	28.8	35
10	Record#9	28.8	86.4	268.8	2	5.37	116	6.65	12.6	18.6	27.7	34.7
11	Record#10	28.8	9.6	38.4	1.8	4.85	137	5.84	10.2	18.1	34.2	51.7
12	Record#11	57.6	19.2	76.8	3.4	3.72	137	5.84	4.7	11.1	25.4	43.0
13	Record#12	86.4	28.8	115.2	4.9	3.88	137	5.84	4.7	11.1	25.4	43.0
14	Record#13	19.2	19.2	38.4	1.8	5.03	115.8	5.16	19.6	31.5	48.8	65.2
15	Record#14	38.4	38.4	76.8	3.4	3.07	115.8	5.16	6.7	13.6	23.5	39.8
16	Record#15	57.6	57.6	115.2	4.9	2.96	115.8	5.16	5.5	10.1	19.9	32.2
17	Record#16	9.6	28.8	38.4	1.8	2.78	129	4.93	33.2	55.4	83.5	97.2
18	Record#17	19.2	57.6	76.8	3.4	3.15	129	4.93	16.7	24.9	35.3	41.4
19	Record#18	28.8	86.4	115.2	4.9	2.28	129	4.93	8.8	13.6	20.8	25.8
20	Record#19	57.59	19.2	32.91	2.4	1.84	120.9	7.65	6.4	11.1	22	36.8
21	Record#20	86.42	28.81	49.38	3.6	3.01	120.9	7.65	5.1	6.7	16.8	29.6
22	Record#21	38.4	38.4	32.91	2.4	1.55	140.18	8.8	9.4	14.4	24.1	34.4
23	Record#22	57.61	57.61	49.39	3.6	1.96	140.18	8.8	5.7	9.6	17.6	27.3
24	Record#23	19.2	57.59	32.91	2.4	1.33	121.4	9.03	14.7	22.4	31.6	37.5
25	Record#24	28.81	86.42	49.38	3.6	1.76	121.4	9.03	14.1	16.2	23.9	29.5

Data points are split into 'complete' and 'missing data' sets

Data analysis allows the options for:

Graphical display of data

Statistical analysis

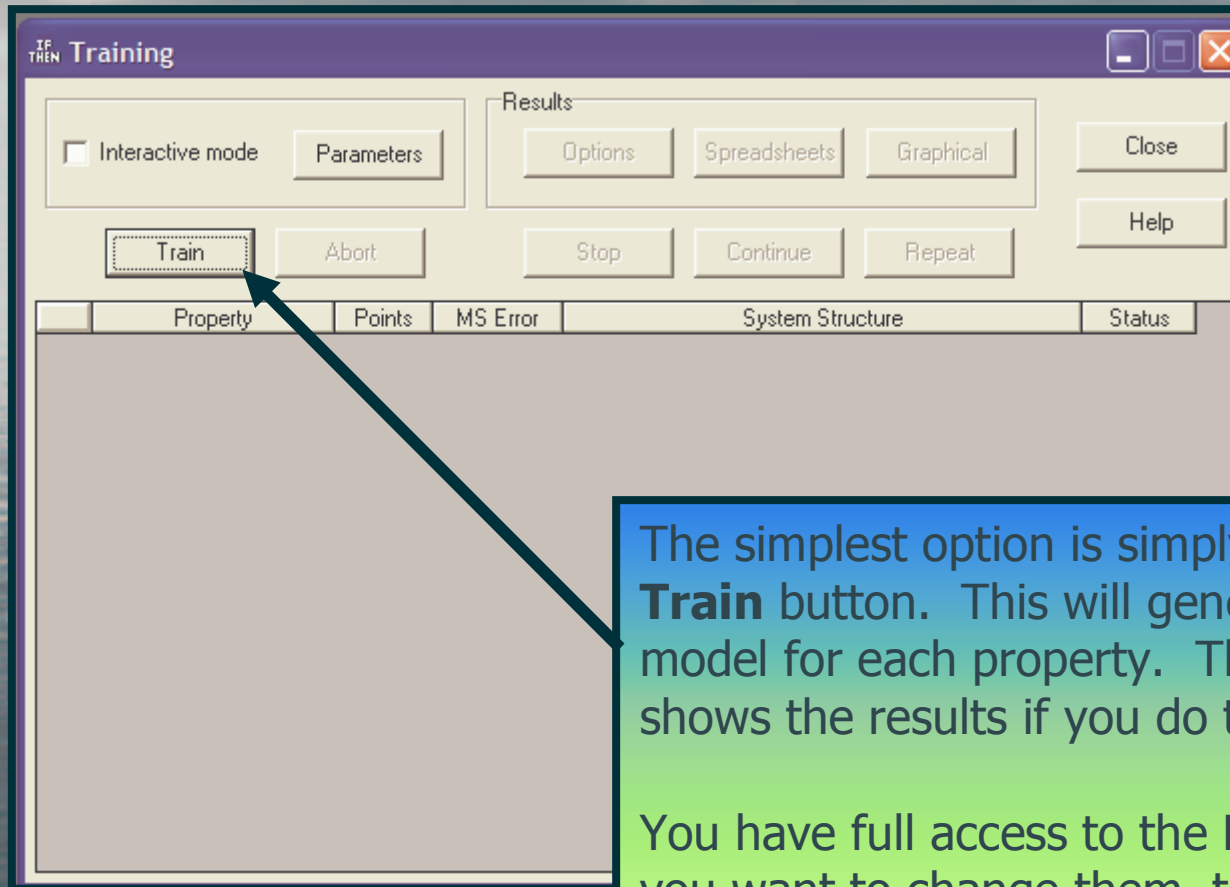
Examination for outliers with **Analyze** button

Click on **OK** to move on to Training

# What is 'Training'?

- Defined as the process of building a model that connects your inputs to outputs
  - Given specific inputs, the outputs are predicted in line with experimental findings
- FormRules tries many possible models
  - Uses statistical 'assessment criteria' to balance complexity of model with goodness of fit to your known experimental values

# The Training screen



The simplest option is simply to press the **Train** button. This will generate a separate model for each property. The next screen shows the results if you do this.

You have full access to the Parameters if you want to change them, though.

# After training...

The screenshot shows the 'IF THEN Training' software interface. At the top, there are buttons for 'Interactive mode', 'Parameters', 'Options', 'Spreadsheets', 'Graphical', and 'Help'. Below these are buttons for 'Train', 'Abort', 'Stop', 'Continue', and 'Repeat'. The main area contains a table with the following data:

	Property	Points	MS Error	System Structure	Status
1	1hr	24	0.007506	1(4)	TRAINED
2	2hr	24	0.007145	1(4)	TRAINED
3	4hr	24	0.010033	1(3)+2(2)	TRAINED
4	6hr	24	0.014630	2(2)+1(3)	TRAINED
5	8hr	24	0.007725	2(2)+1(2)4(2)+7(0)	TRAINED
6	10hr	24	0.015347	2(2)+7(2)	TRAINED
7	12hr	24	0.019239	2(2)+7(2)	TRAINED
8	16hr	24	0.031387	2(3)	TRAINED
9	20hr	24	0.028507	2(3)	TRAINED
10	24hr	24	0.046240	2(2)	TRAINED

A 'FormRules: Message' dialog box is overlaid on the table, displaying the text 'Training is complete!' and an 'OK' button.

Now, results are available.

These have been generated using the default training parameters.

**FormRules** tries lots of possible models, to see which one is best for modeling your data.

MS (mean-squared) error gives idea of how well model fits the training data.

# Assessing Model Quality

	A	B	C	D	E
1	Train Set				
2	1hr				
3	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
4	Model	910.61	4	227.652	26.3254
5	Error	164.305	19	8.64762	
6	Total	1074.92	23		
7					
8	Train Set R-squared	84.7147			
9					
10	2hr				
11	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
12	Model	2067.88	4	516.969	24.1525
13	Error	406.683	19	21.4044	
14	Total	2474.57	23		
15					
16	Train Set R-squared	83.5655			
17					
18	4hr				
19	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
20	Model	3371.6	4	842.901	14.9504
21	Error	1071.22	19	56.3799	
22	Total	4442.84	23		
23					
24	Train Set R-squared	75.8889			
25					
26	6hr				
27	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
28	Model	3851.72	4	962.931	10.2208
29	Error	1790.04	19	94.2126	
30	Total	5641.77	23		
31					
32	Train Set R-squared	68.2717			
33					

The Spreadsheets button on the Training screen gives you ANOVA (ANalysis Of VAriance) statistics. As long as the  $R^2$  value is good, you can feel confident that the model reflects the cause-and-effect.

If you can't get a good model, maybe it's because you aren't measuring an important input variable – valuable information in saving you wasted effort doing pointless experiments.

Here, the models are fine, so let's go on to look at more results.

# Getting Results

IF THEN Training

Interactive mode   Parameters

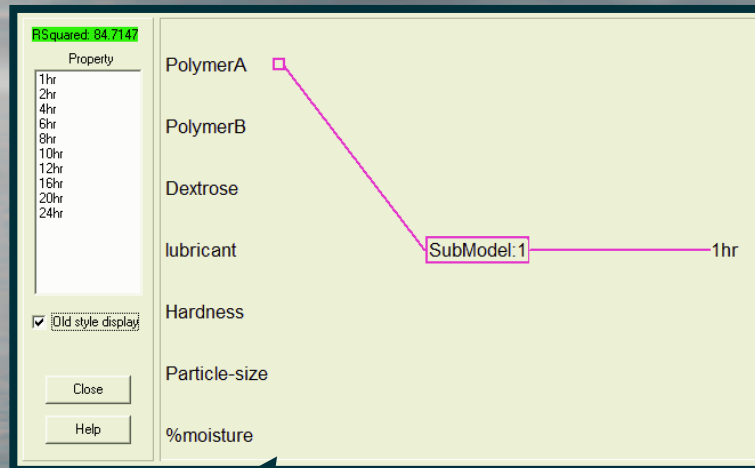
Results:   Options   Spreadsheets   Graphical   Close

Train   Abort   Stop   Continue   Repeat   Help

	Property	Points	MS Error	System Structure	Status
1	1hr	24	0.007506	1(4)	TRAINED
2	2hr	24	0.007145	1(4)	TRAINED
3	4hr	24	0.010033	1(3)+2(2)	
4	6hr	24	0.014630	2(2)+1(3)	
5	8hr	24	0.007725	2(2)+1(2)4(2)+7(2)	
6	10hr	24	0.015347	2(2)+7(2)	
7	12hr	24	0.019239	2(2)+7(2)	
8	16hr	24	0.031387	2(3)	
9	20hr	24	0.028507	2(3)	
10	24hr	24	0.046240	2(2)	

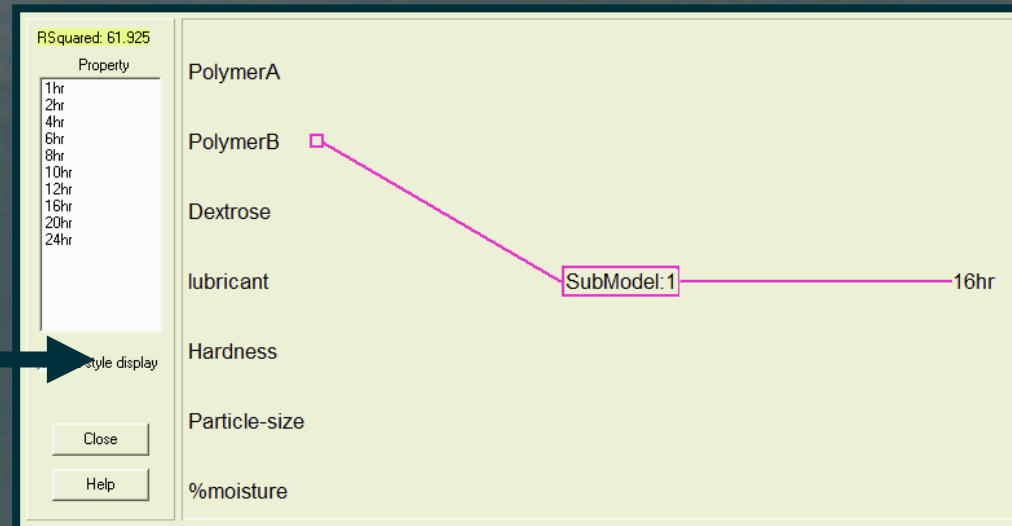
Pressing the **Graphical** button on the Training screen gives you a graphical display of the results, as the next slide shows.

# Graphical Display of Results

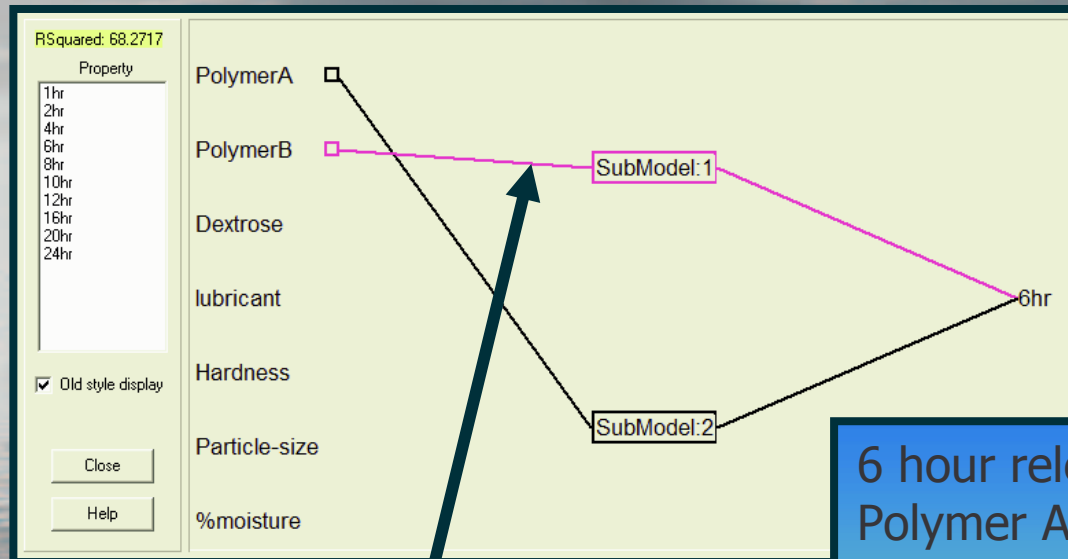


1 h release depends only  
On Polymer A.

But 16 hr release depends  
only on Polymer B.



# More Results



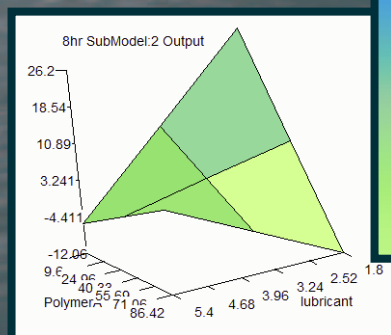
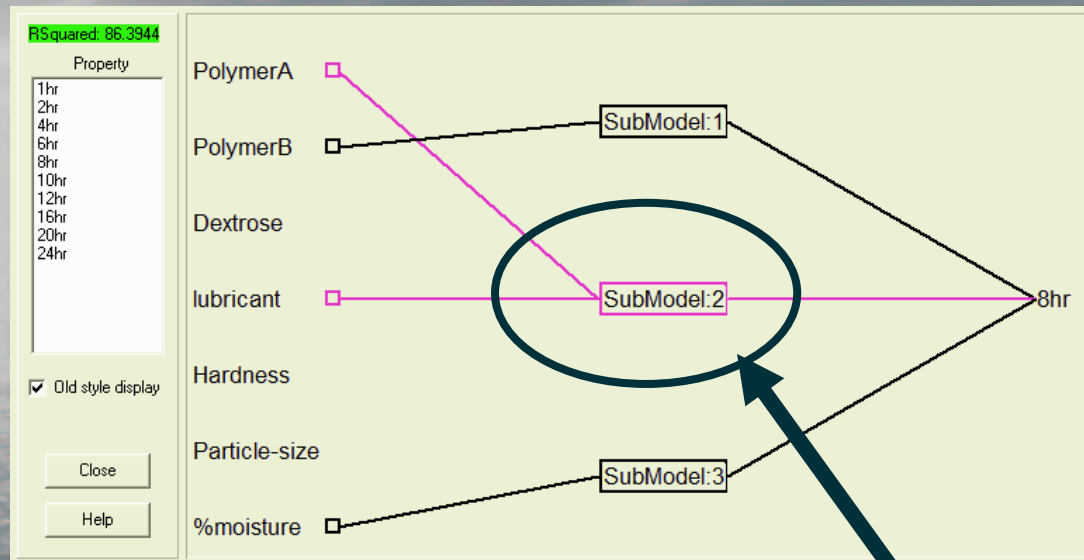
(Polymer B has a bigger effect than Polymer A – so it is shown in purple)

6 hour release depends on both Polymer A and Polymer B.

So we can say that short-term release is controlled by Polymer A, and long term release by Polymer B. At intermediate times, the amounts of both polymers are important.

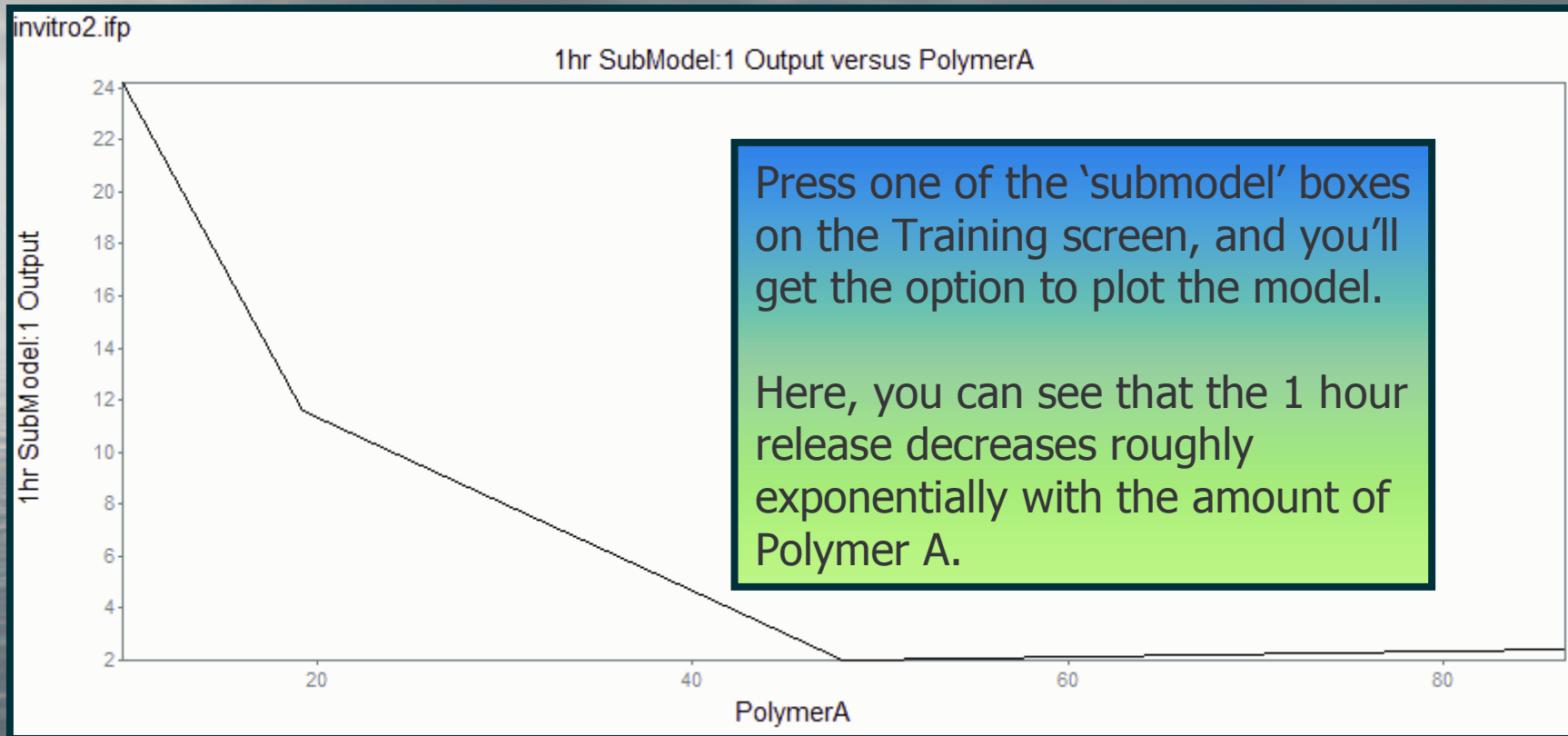
This has been discovered automatically from the data!

# And even more powerfully...



**FormRules can discover interactions between inputs.** Here, submodel 2 has 2 inputs, showing that there is an interaction between the amount of Polymer A and the amount of lubricant. The plots and rules show exactly what this interaction is.

# You can get even more information



So, with **FormRules**, you can discover how each property is affected by the ingredients – whether the relationship is linear or complex.

# And you even get rules

IF Polymer A is LOW AND lubricant is LOW  
THEN 8hr is HIGH (1.00)  
IF Polymer A is LOW AND lubricant is HIGH  
THEN 8hr is LOW (1.00)  
IF Polymer A is HIGH AND lubricant is LOW  
THEN 8hr is LOW (1.00)  
IF Polymer A is HIGH AND lubricant is  
HIGH THEN 8hr is LOW (0.77)

Here are the rules  
for the amount of  
drug released at  
8 hours.

Within **FormRules**, the models are expressed as rules, with 'confidence levels' that show exactly how high or low the property is. Now that you know the trends in behaviour, you can focus your experimentation on the most important inputs, and on the ranges most likely to give the properties you want.

# Minimizing your learning curve

- Help and support is available throughout your session
  - On-line Help via the Help buttons
  - Options of having 'Tips' displayed to tell you what's important about each of the screens – the next slide shows this for the Enter/Edit Data screen

# "Tips" for use...

The screenshot shows the FormRules software interface. The main window is titled "FormRules - C:\Program Files\Intelligensys\FormRules\invitro - [Set Field Types]". It features a menu bar (Task, Model, Use, Reports, Window, Help) and a toolbar. The central area contains a table with the following data:

	Field Name	Field Type	Units	Min	Max
1	PolymerA	Ingredient		9.6000	86.4200
2	PolymerB	Ingredient		9.6000	86.4200
3	Dextrose	Ingredient		32.9100	268.8000
4	Lubricant	Ingredient		1.8000	5.4000
5	tablet_weight	Not Used		88.2000	395.6000
6	drug/polymer	Not Used		0.0770	0.2000
7	PolymerA/PolymerB	Not Used		0.3300	3.0000
8	Hardness	Ingredient		1.3300	11.1000
9	Particle-size	Ingredient		115.8000	156.0000
10	%moisture	Ingredient		3.8500	9.0300
11	1hr	Property		3.0000	33.2000
12	2hr	Property		6.7000	55.4000
13	4hr	Property		16.8000	83.5000
14	6hr	Property		25.8000	97.2000
15	8hr	Property		30.8000	105.6000
16	10hr	Property		35.3500	105.4000
17	12hr	Property		39.9000	105.4000
18	16hr	Property		50.8000	110.0000
19	20hr	Property		61.0000	119.6000
20	24hr	Property		72.3000	119.5000

At the bottom of the main window are buttons for "Quick Select", "Previous", "Next", and "Help".

Overlaid on the right side is a "Tips" window titled "Set Field Types Screen" with the following content:

1. Set each data type as an Ingredient, Processing Condition, Property or Not Used, by clicking in the Field Type and selecting the required option from the pull-down menu.
2. Processing Conditions and Ingredients are both inputs. Properties are the outputs. If Not Used is selected, that variable plays no part in developing the model.
3. After setting a Field Type, right-clicking anywhere above the main screen area, or pressing the Quick-Select button allows you to set all fields below the current one to the same type as you have just set. This speeds up the field selection process.
4. Units are optional and can be left blank.
5. By default, the Task Name will be the same as the imported file name, if data from a tab-delimited file was imported into the Enter/Edit data screen.
6. You can edit the Task Name to avoid overwriting the data in an existing task directory.
7. You can press the Browse button to select a new task directory using the Windows file browser. Please note you are selecting the directory to which the task will be saved, not a file.
8. If you have accidentally closed the Set Field Types screen (the main screen area is empty), select Set Field Types from the Model menu to get it back.
9. Press the Help button for further help.
10. Click the Next button to move onto the next screen.

At the bottom right of the Tips window is a "Browse" button.

Once you're an expert, you can switch the Tips off, of course.

# Remember...

- FormRules is versatile
  - It works on your own data, so as long as you have cause-and-effect information hidden in that data, FormRules can find this information and present it to you clearly
- FormRules does not require specialist artificial intelligence expertise
  - Defaults have been tested so that they are generally reliable for formulation problems
  - You are given full information that enables you to assess your models
  - Tips and on-line Help are available throughout the program

# FormRules

- Can be supplied as a networkable licence, allowing anyone on your LAN to access the software and share results
- Is backed by a full range of training, maintenance and support options
- Is continually being developed in collaboration with customers and academic groups worldwide

# Expand Your Horizons

- For more information on FormRules, contact Intelligensys

- <http://www.intelligensys.co.uk>

- [sales@intelligensys.co.uk](mailto:sales@intelligensys.co.uk)

- Download the demo at

[www.intelligensys.co.uk/Products/Downloads/downdemo.htm](http://www.intelligensys.co.uk/Products/Downloads/downdemo.htm)

