## Tailoring DOE Constraints to the Problem

ENBIS Annual Conference June 28<sup>th</sup>, 2022

Volker Kraft, Chris Gotwalt - JMP



## Tailoring DOE Constraints to the Problem Abstract

There are often constraints among the factors in experiments that are important not to violate but are difficult to describe in mathematical form. In this presentation, we illustrate a simple workflow of creating a simulated dataset of candidate factor values. From there, we identify a physically realisable set of potential factor combinations that is supplied to the new Candidate Set Design capability in JMP 16. This then identifies the optimal subset of these filtered factor settings to run in the experiment. We also illustrate the Candidate Set Designer's use on historical process data, achieving designs that maximize information content while respecting the internal correlation structure of the process variables. Our approach is simple and easy to teach. It makes setting up

experiments with constraints much more accessible to practitioners with any amount of DOE experience.



## Starting point Custom Design > Factor Constraints

- Why an optimal design?
  - We want to tailor the design to the problem at hand!
- JMP Custom Designer workflow:
  - Define response(s) and factors
  - Define the model (incl. main effects, interactions and polynomial terms)
  - Define the budget = number of runs
- If factors can be changed independently, we are all set. If not, we have to
  - Define the constraints



## Starting point

### Custom Design > Factor Constraints

	• •		DOE - Custo	m Design			
•	Custom Design						
►	Responses						
Ŧ	Factors						
	Add Factor  Remove Add N Factors	s 1					
	Name Role		Changes	Values			
	X1 Con	tinuous	Easy	-1		1	
	▲X2 Con	tinuous	Easy	-1		1	
	▼X3 Cate	egorical	Easy	L1	L2		L3
	Covariate/Candidate Runs						
►	Define Factor Constraints						
	Model						
	Main Effects Interactions   RSM	Cross Powers	Remove Ter	m			
	Name		Estimability				
	Intercept		Necessary				
	X1		Necessary				
	X2		Necessary				
	X3		Necessary				
►	Alias Terms						
Ŧ	Design Generation						
	Group runs into random blocks of size:	2					
	Number of Contor Points:						
	Number of Center Folits. 0						
	Number of Runs:						
	Minimum 5						
	O Default 12						
	User Specified 12						
	Make Design						

## Starting point Custom Design > Factor Constraints



#### **Define Factor Constraints** None Specify Linear Constraints Use Disallowed Combinations Filter **Use Disallowed Combinations Script Disallowed Combinations** Clear Reset selection **X1** $\boxtimes$ -1 0 X3 (3) $\boxtimes$ L2 L1 L3 OR AND

STATISTICAL DISCOVERY

## More tailored constraints Custom Design > Covariate/Candidate Runs

#### **Covariate/Candidate Runs**

Select Covariate Factors

Load a set of candidate runs for covariates from the current data table.

- In a nutshell, think about covariates as:
  - Input variable(s) we want to account for, but don't have complete control.
  - Uncontrolled, but observable ahead of time.

These covariates were measured on runners who are candidates for an experiment on running shoe wear. (JMP Sample Data)

🖲 😑 🗧 Rur	ners Cova	ariate	<b>^</b>	»
▼ 3/0 Cols ▼				
■ 100/0	Miles	Weight	Strike Point	
1	10,4	143	Midfoot	
2	12,4	140	Midfoot	
3	10,4	147	Heel	
4	7,6	150	Forefoot	
5	8,4	145	Heel	
6	11,2	145	Midfoot	
7	14,0	139	Midfoot	
8	13,2	145	Midfoot	
9	9,2	146	Midfoot	
10	12,0	142	Heel	
11	7,6	145	Midfoot	
12	9,2	152	Midfoot	
13	12 2	1/12	Midfoot	



## What's next? Custom Design > Covariate/Candidate Runs

- Look at examples defining constraints by candidate runs

   first one as live demo
- Mention other use cases of covariate factors
- Share some technical insight into optimal design using covariates
  - JMP and JMP Pro, version 16+



## Use cases of candidate set designs Custom Design > Covariate/Candidate Runs

- The range of a continuous factor depends on the levels of a categorical factor
- Physical constraints on factors that cannot vary completely independently (i.e. temperature and pressure)
  - Historical data design
  - Filtered design
- Mixture problems
  - Especially sub-mixture problems
- Nonlinear constraints via the filtering method
  - Formula column based filtered designs





## #1: Suppliers with different operator regions

- 12-run DOE with two factors:
  - Machine Supplier (A and B)
  - Temperature
- Supplier A:
  - Temperature controls from 150C to 170C controllable to 5C
- Supplier B:
  - Temperature controls from 140C to 180C controllable to 10C



# #1: Suppliers with different operator regions

	DOE - Cust	om Design		
ustom Design				
sponses				
ctors				
dd Factor  Remove Add N Factors 1				
lame Role	Changes	Values	1	
Nachine Type Covariate	Easy	A	В	
emperature Covariate	Easy	140	180	
ovariate/Candidate Runs				
efine Factor Constraints				
odel				
ain Effects Interactions  RSM Cross Po	wers  Remove Te	erm		
Name	Estimability			
ercept	Necessary			
ichine Type	Necessary			
nperature	Necessary			
chine Type*Temperature	Necessary			
nperature*Temperature	Necessary			
ias Terms				
esign Generation				
clude all selected covariate rows in the design				
llow covariate rows to be repeated				
nder of Runs: 12				
ake Design				

Copyright © JMP Statistical Discovery LLC. All rights reserved.

## #1: Suppliers with different operator regions Chosen candidate runs selected

• • • Machine and Temperature Candidate Set.jmp		<b>^</b>		y 22 20 20 20 20 20 20 20 20 20 20 20 20	1
Machine and Temperature Candidate Set.jmp					
DOE leads to missing factors		Machine Type	Temperature		
		A B	180		
			140		
	1	А	150		
	2	A	155		
	3	A	160		
Columns (2/0)	4	А	165		
Machine Type	5	А	170		
Temperature	6	В	140		
	7	В	150		
	8	В	160		
	9	В	170		
	10	В	180		
	11	A	150		
	12	A	155		
	13	A	160		
	14	A	165		
	15	А	170		
	16	В	140		
■ Rows	17	В	150		
All rows 20	18	В	160		
Selected 10	19	В	170		
Hidden 0	20	В	180		
Labeled 0					



## #1: Suppliers with different operator regions Custom Design table

🖲 🗧 Custom Design				<b>^</b>		x 23	
■Custom Design					Covariate Row		
Design Custom Desigi	ı i	Machine Type	Temperature	Y	Index		
Criterion D Optima <ul> <li>Model</li> <li>Evaluate Design</li> </ul>		A B	180		1 5 6 8		
Covariate Data Table			140		5 others		
Generalized Regression	1	В	140		6		
DOE Dialog	2	В	180		10		
	3	А	150		1		
⊂Columns (4/0)	4	А	170		5		
📥 Machine Type \star	5	В	160		18		
Temperature *	6	А	150		1		
▲Y* ■ Covariate Row Index	7	В	140		16		
	8	В	160		8		
	9	В	180		20		
Rows	10	А	170		5		
All rows 12	2 11	А	150		11		
Selected (	) 12	А	170		15		
Excluded ( Hidden (	)						
	,						



## #1: Suppliers with different operator regions Forcing a missing value into the design (Machine A at 160°C)



#### Covariate/Candidate Runs







#### 🖲 😑 Powder Metallurgy Historical Data.jmp

#### ↑ ♣ 〒 ₩ ► / ₩ ≥

												,
Powder Metallurgy Historica	al Data.jmp 🔹 🕨	● 9/0 Cols										
Source: Philip J. Rams Used	with Permission			Formation	Compaction	Compaction	Sinterina	Sinterina	Sintering	Powder	Powder	
Notes: A company manufactu	ares steel drive sh		Shrinkage	Method	Method	Pressure	Time	Temp	Method	Source	Туре	
			0,76	Die Compaction	Cold	115	30,8	1221	Atmosphere	D	3	
⊂Columns (18/0)				Extrusion	VVarm				Vacuum	N	2	
⊿ Shrinkage											5	
📥 Formation Method \star			-2,2			60	17,5	1092				
Lompaction Method		1	0,19	Die Compaction	Cold	90	27,48	1133	Vacuum	N	1	
Compaction Pressure		2	-0,37	Extrusion	Cold	90	27,48	1143	Atmosphere	D	3	
▲ Sintering Time		3	-0,16	Extrusion	Cold	100	27,48	1121	Atmosphere	D	3	
Sintering Temp		4	0,41	Extrusion	Cold	95	27,48	1122	Atmosphere	D	3	
Bowder Source		5	-0,65	Extrusion	Cold	85	27,48	1118	Atmosphere	D	4	
Powder Type		6	0,41	Extrusion	Cold	95	27,48	1122	Atmosphere	D	4	
		7	-0,23	Extrusion	Cold	100	27,48	1137	Atmosphere	D	3	
⊿ Unit 😎		8	0,12	Extrusion	Cold	100	27,48	1126	Atmosphere	D	3	
🔺 Dwell 🕱		9	0.41	Extrusion	Cold	105	20.85	1149	Atmosphere	D	4	
🔺 Particle Size 😎		10	-0.72	Extrusion	Cold	85	27.48	1126	Atmosphere	D	3	
⊿ P1 🕫		11	-1.08	Extrusion	Cold	85	27.48	1117	Atmosphere	D	5	
		12	-0.65	Extrusion	Cold	80	27.48	1124		D	3	
		12	_0.79	Extrusion	Cold	85	27,40	11/0	Atmosphere	D	1	
▲ P5 ↔		11	-0,77	Extrusion	Cold	100	27,40	1120	Atmosphere		4 5	
		14	-0,10	Extrusion	Cold	100	27,40	1127	Atmosphere		5	
Rows		15	-0,65	Extrusion	Cold	85	27,48	1118	Atmosphere	D	2	
All rows	6 253	16	0,12	Extrusion	Cold	100	27,48	1126	Atmosphere	D	2	
Selected	0.233	17	-0,72	Extrusion	Cold	85	27,48	1126	Atmosphere	D	3	
Excluded	0	18	-0,44	Extrusion	Cold	85	27,48	1135	Atmosphere	D	5	
Hidden	0	19	-0,65	Extrusion	Cold	80	27,48	1124	Atmosphere	D	3	
Labeled	0	20	-1,15	Extrusion	Cold	80	27,48	1120	Atmosphere	D	5	

## #2: Using historical process data Implicit constraints suggest DOE





	ection Limit
Responses       Add Response * Remove       Number of Responses         Response Name       Goal       Lower Limit       Upper Limit       Importance       Lower Detection Limit       Upper Detection Limit         Shrinkage       Minimize       Importance       Lower Detection Limit       Upper Detection Limit	ection Limit
Add Response *       Remove       Number of Responses         Response Name       Goal       Lower Limit       Upper Limit       Importance       Lower Detection Limit       Upper Detection         Shrinkage       Minimize       Importance       Lower Detection       Importance	ection Limit
Response Name         Goal         Lower Limit         Upper Limit         Importance         Lower Detection Limit         Upper Detection           Shrinkage         Minimize         Importance         Impo	ection Limit
Shrinkage Minimize	
interities per	
Factors	
Add Factor  Remove Add N Factors 1	
Name Role Changes Values	
*Formation Method Covariate Easy Die Compaction Extrusion Isostatic Press	
Compaction Method Covariate Easy Cold Warm	
Compaction Pressure Covariate Easy 60 115	
<sup>4</sup> Sintering Time Covariate Easy 17,5347487421324 30,801247903554	
Sintering Temp Covariate Easy 1091,769 1220,625	
<sup>®</sup> Sintering Method Covariate Easy Atmosphere Vacuum	
<sup>®</sup> Powder Source Covariate Easy D N	
*Powder Type         Covariate         Easy         1         2         3         4         5	
▶ Covariate/Candidate Runs	
▶ Define Factor Constraints	
• Model	
Main Effects Interactions v DSM Cross Downers v Remove Term	
Name Estimability	
Intercept Necessary	
Formation Method If Possible	
Compaction Method If Possible	
Compaction Pressure If Possible	
Sintering Time It Possible	
Sintering lemp It Possible	
Sintering Method If Possible	
Powder Source If Possible	
Alias Ierms	
▼ Design Generation	
C Include all selected covariate rows in the design	
Value covariate rows to be repeated	
Make Design	

- Set response
- Load (controllable) covariate factors
- Choose RSM model, estimate all terms except intercept "If possible" (Baysian Ioptimal candidate set design)
- Allow optimally chosen replication
- Choose # of runs to 100
- Make design



Distributions (all data)



viaitiv			aj					
Corre	lations	;						
		Со	mpact	tion Pr	essure	Sinte	ering Time Sin	tering Temp
Compa	action Pre	ssure		1	,0000		-0,6489	0,4945
Sinterir	Sintering Time			-0	),6489		1,0000	-0,7343
Sinterir	ng Temp			C	),4945		-0,7343	1,0000
The cor	relations	are estim	lated k	oy Row	/-wise			
method	l.							
Scatt	erplot	Matrix						
110			•	•	•			•
100			•	•	:			•
90	Comr	action		:	:		•	:
80-	Pres	Pressure			•	•		•
70-					•	•		:
60				•	•	•		
30-		•						•
28								•
26-								
24 -	• • •		S	Sinterir	ng Tim	ne	•••	•
22								
20-								
18-		• •	•					•
1220	••••	••••	•	•	• •	•		
1190								
1140							Sinterina T	emp
1140	. i	illi					Sintening i	cinp
1120				I		i		
1100-				I	:			
4	0 80	100		• • • • •	- 	30 2	1100 1140 11	80
0.						00		

#### Multivariate (Candidate set design)

#### Correlations

	Compaction Pressure S	Sintering Time S	Sintering Temp
Compaction Pressure	1,0000	-0,4367	0,4464
Sintering Time	-0,4367	1,0000	-0,4100
Sintering Temp	0,4464	-0,4100	1,0000

The correlations are estimated by Row-wise method.

#### Scatterplot Matrix



## #3: Machine suppliers with temperature and pressure

Space Filling Design						
Responses						
Add Response  Remove	Number of Responses					
Response Name	Goal	Lower Limit	Upper Limit	Importance	Lower Detection Limit	Upper Detection Limit
Υ	Maximize					
Factors						
Name Rol	e \	′alues				
Temperature Co	ntinuous 1	080	1160			
Pressure Con	ntinuous <u>6</u>	0	120			
Machine Cat	tegorical [4	\	B			
Define Factor Constra	ints					
Space Filling Design Methods Number of Runs: 1000 Sphere Packing Latin Hypercube Uniform Minimum Potential Maximum Entropy	Optimal Packing of Sph Inside of a Cube. Latin HyperCube with C Uniform Design Minimum energy desigr Maximum entropy desig process	eres Pptimal Spacing Is in a spherical regior Ins for a Gaussian	n.			
Gaussian Process IMSE Optimal	Integrated mean square for a Gaussian process	error optimal designs	5			
Fast Flexible Filling	Space filling design thro	ough clustering.				
Back						

Copyright © JMP Statistical Discovery LLC. All rights reserved.

STATISTICAL DISCOVERY

## #3: Machine suppliers with temperature and pressure

💿 🔘 🔹 Machine Tem	perature Pressure Candidate Set For	Filtering.jmp				● ● Machine Temperat
<ul> <li>Machine Temperature</li> </ul>	Pressure Candidate Set For Fi ►	•			,	<ul> <li>Multivariate</li> </ul>
Design	Fast Flexible Filling Design		Machine	Temperature	F	Correlations
Model			B	1160		Temperature Pressure
DOL Dialog						Temperature 1,0000 -0,0137
	Data Filter for Machine Tem		$\langle  $	1080		Pressure -0,0137 1,0000
	▼ <b>⊡</b> Data Filter	•	1 B	1137 5105639	96	
	Clear Favorites -	•	2 A	1139 8400955	98	The correlations are estimated by Row-wise
	✓ Select Show Include	•	3 A	1139,4502326	5	method.
	216 matching rows		4 B	1140,2937075	95	<ul> <li>Scatterplot Matrix</li> </ul>
	Inverse	•	5 B	1144.3101743	98	1160-
	✓ Machine (2)	•	6 A	1142,1004341	99	
Columns (4/0)		•	7 A	1141.6982023	10	1140-
▲ Machine <b>*</b>	🔄 🔄 lemperature 📓 – 🗕	•	8 B	1141.3187741	10	
▲ Temperature *		•	9 B	1143.7561592	10	1120- Temperature
Pressure *	1000 01 10077	•	10 A	1145,8828239	10	
▲Y*	1080,0140977 1140	•	11 B	1150,253102	96	1100-
		•	12 A	1150,6716359	98	
	Pressure	•	13 A	1147,5931964	9	1080 -
		•	14 B	1146,6808648	97	120-
		•	15 B	1143,0305454	96	110-
	85 119,99085322	•	16 A	1145,2015951	95	100
	AND OR	•	17 B	1139,1444061	10	Proseuro
		•	18 A	1140,6396273	1	90-
		•	19 B	1142,904167	10	80-
Rows		•	20 A	1142,0195644	10	70-
All rows	1.000	•	21 B	1145,2767042	10	60-
Selected	216	•	22 A	1144,151537	10	
Excluded 0 Hidden 0		•	23 B	1148,842558	10	1080 1100 1120 1140 1160 60 70 80 90 100 110 120
		•	24 A	1147,1874032	10	
Labeled	0				• •	

Copyright © JMP Statistical Discovery LLC. All rights reserved.

ĩm

STATISTICAL DISCOVERY

## #3: Machine suppliers with temperature and pressure Sample design





## Covariate Factors More use cases: Using all rows

- We have observable, uncontrollable factors for our experimental units that we want to account for
  - Assign two treatments to a group of patients.
- We want to force a particular structure for a subset of inputs (that may even be controllable)
  - 12 runs, 5 two-level factors, but factor A needs 2/3 of the levels at "L1" and 1/3 at "L2".
- DSD-type designs
  - 3-level categorical factor
- Complex block structures
- Account for other covariate information (time trend, for example)



## Covariate Factors

More use cases: Using a subset of rows

- Pick an optimal subset (NOT a random subset)
  - Pick a subset of patients from a larger group
  - Hard-to-change covariates (such as multiple measurements per student)
- Augmentation force original runs into the design
- Allow runs to be used more than once
- Constraints: Disallowed combinations alternative
  - Create candidate set (Full Factorial, Space-Filling, Simulation) > Filter & select runs

$$X1^{2} + 4 \cdot X3 + X4 \cdot X5 \ge 100$$

## Covariate Factors "Behind the scenes" in JMP 16+

- JMP versions 10 15: Two-step optimization
  - 1. Row-exchange on the covariates (D-optimality)
    - Switching rows, starts with candidate set
  - 2. Coordinate-exchange on the controllable factors
    - Swapping values element-by-element, no candidate set needed
- JMP versions 16+: Hybrid approach
  - Allows for different optimality criteria to be considered for the covariates.
  - More flexibility for your design creation (best of both worlds).
  - Option much more visible to the user.



## Covariate Factors High flexibility

- Covariates can be combined with other Custom Design settings
  - Other factor types, e.g. mixtures
  - Split-plot designs, e.g. hard-to-change covariates
  - Blocking
  - Model effects, e.g. interactions
- No JMP Pro needed

# Design Generation Include all selected covariate rows in the design Allow covariate rows to be repeated Number of Runs: 20

Factors	Factors								
Add Factor	Remove	Add N Factors	; 1						
Name	Ro	le Ch	anges	Values					
• Machine	Туре Со	ovariate Ea	sy	A	В				
Temperat	ture Co	ovariate Ea	sy	140	180				
<ul> <li>Covaria</li> </ul>	ate/Candid								
Select Co	variate Factors	Load a set of o	candida	ate runs for covariates					
Machine		nom the curre							
Туре	Temperature								
А	150								
А	155								
Δ	160								

Factors

Continuous Discrete Numeric

Categorical

Add Factor **v** 

Remov

>

>

>

STATISTICAL

## Suggested resources

- JMP Developer Tutorial: <u>"Handling Covariates Effectively when</u> <u>Designing Experiments"</u>, Ryan Lekivetz, Principal Research Statistician Developer, Oct 2021
- JMP Blog: <u>"What is a covariate in design of experiments?"</u>, Ryan Lekivetz, Principal Research Statistician Developer, Feb 2021
- JMP Help: Design of Experiments Guide >
  - Custom Designs > Build a Custom Design > Factor Types
  - Examples of Custom Designs > Experiments with Covariates
- Goos, P., and Jones, B. (2011). "Optimal Design of Experiments: A Case Study Approach", chapter 9. New York: John Wiley & Sons.





jmp.com