Anatune Ltd
Unit 4, Wellbrook Court, Girton Road, Cambridge, CB3 0NA, UK
Tel: +44 (0) 1223279210 Fax: +44 (0) 1223279253
Email: info@anatune.co.uk Internet: www.anatune.co.uk

Design of Experiment (DoE) and GERSTEL MultiPurpose Sampler (MPS): a match made in heaven

Camilla Liscio, Anatune Ltd., Girton, Cambridgeshire (UK).

Introduction

In the scientific community the general approach to method development and optimisation is what is called the one-variable-at-a-time approach (OVAT), which varies only one experimental variable at a time keeping the other variables fixed. However, statistically designed experiments which can vary several variables simultaneously are often a more efficient way to explore a complex experimental space. Design of experiment (DoE) is a controlled set of tests designed to model and explore the relationship between experimental variables and one or more responses. The foundations of DoE were developed by Sir Ronald Fisher in the UK in the first half of the 20th century. Immediately following World War II, the first industrial era marked another resurgence in the use of DoE highlighting the importance of statistical driven quality control.

DoE advantages are several:

- It requires less resources (experiments, time, material) for the amount of information obtained
- It can estimate the effects of each variable on the response more precisely since more observations of an effect result in higher precision and reduced variability
- It can investigate interactions between variables systemically while this information is completely overlooked in the OVAT approach
- It provides information of a larger region of the experimental space

In this scenario, automated sample preparation finds its perfect application. In fact, DoE requires the change of several variables at the time for each experiment with randomization of experiments to eliminate potential biases. Automation offers very accurate control of all the variables at the same time and changes to experimental variables are programmed and scheduled without requiring the constant presence of an operator. Furthermore, the robotic control of variables further reduces the occurrence of biases which could jeopardise the outcome of the study.

This application note describes the standard workflow using DoE for the optimisation of an analytical challenge. The derivatisation of fatty acids to picolinyl derivatives was chosen as an example. In fact, derivatisations always involve optimisation of several parameters to achieve the best performances and therefore offer a very good situation for the DoE to shine. Traditionally, fatty acids are analysed by GC-MS using their methyl ester derivatives (FAMEs). However, FAMES mass spectra cannot provide ions indicative of structural features such as double bonds and branch–points. On the other hand, fatty acids 3-pyridylcarbinol esters also known as picolinyl derivatives are a better option for structural elucidation purposes. In fact, in these derivatives the carboxyl group is derivatised with a reagent containing a nitrogen atom. During molecular ionisation, the nitrogen will carry the charge rather than the alkyl chain and this will minimise double bonds ionisation and migration helping providing very informative structural information.

The following paragraphs will describe the use of automated sample preparation for the optimisation of the picolinyl derivatisation using the DoE approach.

Instrumentation

Autosampler: GERSTEL MPS xt Dual Rail, Left MPS 10uL syringe /Right MPS 250uL syringe

Modules: Agitator, GERSTEL MultiPosition Vortexer (mVorx), GERSTEL MultiPosition Evaporation station (mVAP), Anatune CF-200 Robotic Centrifuge

GC-MS: Agilent GC 7890- MSD 5975C, Inert Ion Source

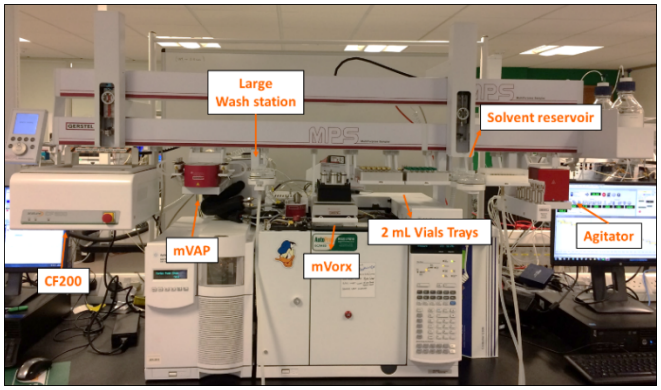

Figure 1: Online fully automated solution for optimisation of fatty acids derivatisation using Design of Experiment

Methods

Optimised Automated Sample Prep

FAME Mix certified reference material in dichloromethane (Sigma-Aldrich) was added into a 2mL high recovery vial. Derivatising reagent was added to the sample and the sample was vortexed. Samples were then incubated to promote derivatisation. Once derivatisation was completed, water and hexane were added to quench the derivatisation and to back extract the picolinyl derivatives. Samples were vortexed to allow efficient partitioning between the two liquid phases. After centrifugation, 2 µL of the top organic layer were directly injected on the GC-MS.

GC-MS Analysis

GC-MS (MSD):
Column: HP-35MS Ultra inert 30 m x 0.25 mm x 0.25 µm
Injection mode: Pulsed Splitless at 250 °C
Flow: 1 mL/min
GC ramp: 50 °C held for 2 min, 40 °C/min to 210 °C held for 1 min, 2°C/min to 280 °C held for 5 min, 40 °C/min to 300 °C held for 1.5 min
Runtime: 49 min
Auxiliary temperature: 300 °C
Inert Source, EI mode at 300°C, Mass range 50-650 *m/z*

anatune
Collaborate • Innovate • Automate

Results and Discussion

The ideal experimentation approach when using Design of Experiment should be sequential. In fact, the knowledge of the investigated process develops as the experimental plan progresses and investing all resources in a one-off large experiment would result inefficient in many cases.
DoE standard workflow usually encompasses an initial step for planning and scoping to identify goals, responses and variables, followed by a screening step to generate a design and evaluate its suitability. Once the model that best fits the experimental data has been attained, the optimisation and robustness assessment are performed to optimise the variables settings and predict process performances.

Design of Experiment Step 1: Planning and Scoping

Planning and Scoping are the first two steps required in the DoE workflow and they are crucial to lead to a successful experimentation. Planning allows to determine the goal of the experiment and to identify relevant responses and variables. Scoping conducts pilot trials which have several advantages:

- Practice of the experimental process to investigate
- Estimation of the experimental random variation
- Evaluation of the experimental design space and goals to assess achievability

During planning, the experimental variables (also called factors) which will be varied systematically from trial to trial are selected and levels are chosen for each variable to explore broadly the experimental space. The pilot trials during scoping usually comprise a minimum of four trials: low (all variables are set at the lowest level), center points (all variables are set in the middle of the experimental range) and high (all variables are set at the highest level). Centre points allow to test for non-linearity and also assess random variation therefore a minimum of 2 replicates can be performed.

Planning identified 4 relevant factors for the investigation of the fatty acids picolinyl derivatization: derivatising reagent volume, incubation mixing speed, incubation temperature and incubation time. Area of each target picolinyl derivatives was chosen as Response. Table 1 summarises the pilot trial plan for the scoping of the fatty acids picolinyl derivatisation with the selected factors and choses levels for each of them.

	Factors			
Scoping trials	Reagent [µL]	Speed [rpm]	Temperature [°C]	Time [min]
LOW	25	250	30	30
Centre point	35	500	45	45
HIGH	45	750	60	60

Table 1: Pilot trial plan for the scoping of the fatty acids picolinyl derivatisation

As an example, Figure 2 shows the results obtained for the scoping trails for three of the investigated fatty acids: picolinyl hexanoate (alkyl chain C6), picolinyl pentadecanoate (alkyl chain C15) and picolinyl methyl arachidoate (alkyl chain 19). Response is shown on the y axis (chromatographic peak areas).

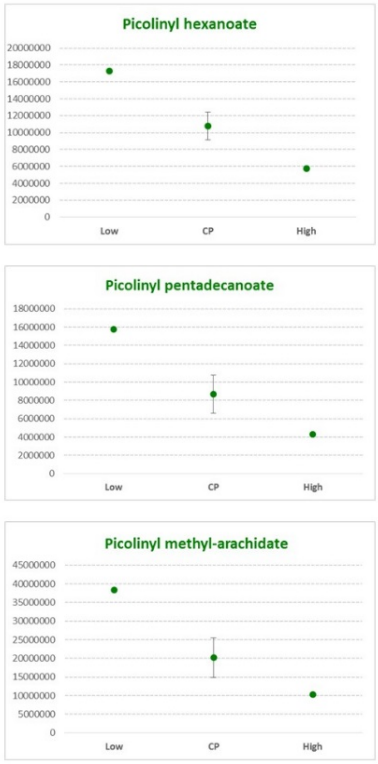

Figure 2: Pilot trails (low, center points CP, high) for the picolinyl derivatization of three fatty acids: picolinyl hexanoate (top), picolinyl pentadecanoate (middle) and picolinyl methyl arachidate (bottom).

As shown in Figure 2, the response changed significantly between the low and the high trials suggesting that relevant factors were selected and levels were including appropriate experimental space. Centre points result showed that experimental variability was acceptable (within 20% RSD, *n*=3).

These preliminary results encouraged to move to the following step: screening.

Design of Experiment Step 2: Screening

Screening designs are mainly used to differentiate between key factors (statistically significant) and less important factors. Furthermore, screening can estimate the magnitude of the significant factors and it can identify existing interactions between the investigated factors. Several types of screening design are available. In this study, a two-level full factorial design was selected. A full factorial design defines an experiment where trials are run at all possible combinations of factor settings. In this design an experimental run is performed at every combination of the factor levels and the sample size is the product of the number of levels of the factors.

For this application, a two level full factorial design was used investigating four factors (2^4=16 runs) plus 3 centre points for a total of 19 runs. Figure 3 shows the experimental matrix for the selected design, factors, levels and number of centre points. Each row corresponds to a trial to be performed experimentally to acquire the response of interest (in our case the picolinyl derivative chromatographic peak area).

	Pattern	Reagent	Speed	Temperature	Time
1	----	25	250	30	30
2	---+	25	250	30	60
3	--+-	25	250	60	30
4	--++	25	250	60	60
5	-+--	25	750	30	30
6	-+-+	25	750	30	60
7	-++-	25	750	60	30
8	-+++	25	750	60	60
9	+---	45	250	30	30
10	+--+	45	250	30	60
11	+-+-	45	250	60	30
12	+-++	45	250	60	60
13	++--	45	750	30	30
14	++-+	45	750	30	60
15	+++-	45	750	60	30
16	++++	45	750	60	60
17	0000	35	500	45	45
18	0000	35	500	45	45
19	0000	35	500	45	45

Figure 3: Experimental matrix for a full factorial design at two level with 3 centre points for the investigation of the picolinyl derivatisation of fatty acids

This is where the capability of automated on-line solutions (e.g. an MPS autosampler mounted or linked to a dedicated instrumental technique such as GC-MS or LC-MS) can be extremely beneficial to perform the experiment time/cost effectively and robustly. In fact, automated solutions allow the exploitation of the PrepAhead function within the GERSTEL Maestro software. This software feature includes the ability to complete sample preparation for each sample immediately preceding the GC-MS injection, whilst the previous sample is running. Figure 4 shows the PrepAhead function in the timeline for the preparation of the full factorial design. The green and yellow multi-coloured bands represent the sample preparation and the light orange bands the GC run-time.

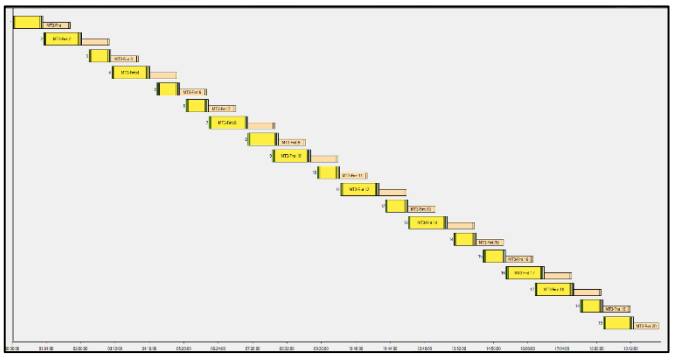

Figure 4: Maestro timeline for the online automated sample preparation and GC-MS analysis for the full factorial design (n=19 samples, estimated duration 20 hours 04 min)

Pareto Plots and Half Normal Probability plots are very useful statistical tools to help differentiating significant factors. In the Pareto Plot effects estimates labelled as contrasts are calculated and effects whose individual p-value is less than 0.1 are highlighted as significant. In the Half Normal Probability Plots the contrasts are plotted against standardised normal scores. The points which deviate from an approximate straight line represent the most significant effects. Figure 5 shows the Pareto Plots and Half Normal Probability plot as for the picolinyl derivatisation of picolinyl hexanoate as an example.

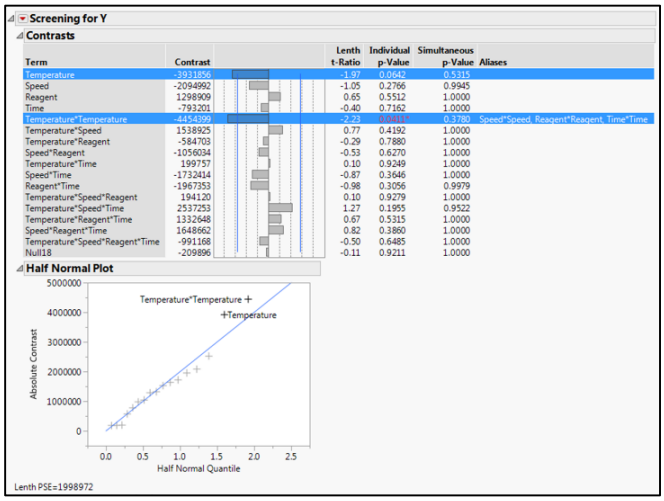

Figure 5: Pareto Plot and Half Normal Plot for the full factorial design for the picolinyl derivatisation of picolinyl hexanoate

As shown in both the Pareto Plot and in the Half Normal plot in Figure 5, temperature was a significant factor for the derivatisation of the picolinyl hexanoate. In fact, the design suggested the presence a quadratic effect of temperature (shown as temperature*temperature). Similar results were obtained for the other target analytes. However, full factorial cannot fully investigate quadratic surfaces. In order to do that a response surface design must be used. Response surface designs can identify points where a minimum or a maximum of the response occurs using three levels factors. One of the most commonly used response surface design is the Central Composite Design (CCD). CCD combines a two-level fractional factorial and two other types of points: centre points with all factors set in the middle of the range and axial points, where one factor is set to a high or low value (known as axial value) and all other factors are set to the midrange value. Figure 6 shows the experimental matrix for the central composite design for the selected factors, levels and number of centre points.

	Pattern	Reagent	Speed	Temperature	Time
1	-+-+	25	600	40	60
2	+-++	45	300	60	60
3	+---	45	300	40	30
4	+++-	45	600	60	30
5	-+--	25	600	40	30
6	00a0	35	450	30	45
7	0A00	35	750	50	45
8	0a00	35	150	50	45
9	++--	45	600	40	30
10	+--+	45	300	40	60
11	-+++	25	600	60	60
12	000A	35	450	50	75
13	+-+-	45	600	40	60
14	00A0	35	450	70	45
15	--+-	25	300	60	30
16	-++-	25	600	60	30
17	---+	25	300	40	60
18	----	25	300	40	30
19	--++	25	300	60	60
20	a000	15	450	50	45
21	+-+-	45	300	60	30
22	000a	35	450	50	15
23	A000	55	450	50	45
24	++++	45	600	60	60
25	0000	35	450	50	45
26	0000	35	450	50	45
27	0000	35	450	50	45

Figure 6: Experimental matrix for a central composite design at three levels

with 3 centre points for the investigation of the picolinyl derivatisation of fatty acids

Figure 7 and Figure 8 show the Half Normal Plot and the response surface profile obtained using the CCD design for the picolinyl derivatisation of picolinyl hexanoate as an example

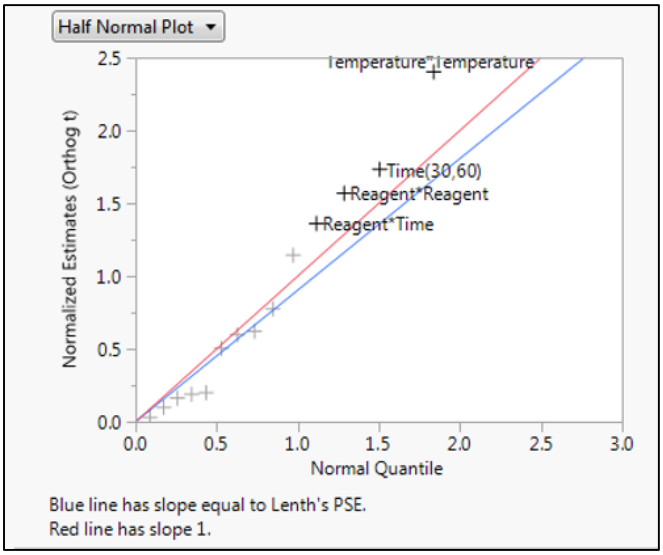

Figure 7: Half Normal Plot for the central composite design for the picolinyl derivatisation of picolinyl hexanoate

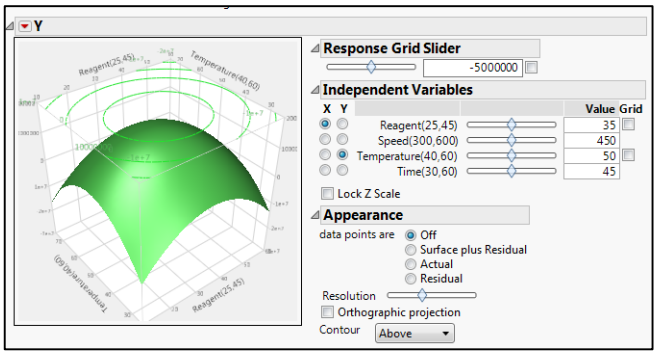

Figure 8: Response surface profiler for the central composite design for the picolinyl derivatisation of picolinyl hexanoate

The response surface design confirmed the presence of a quadratic effect for temperature for picolinyl hexanoate. The other analytes showed similar trends but in order to find the best optimal derivatisation conditions for the majority of the fatty acids present in the mixture, contour profiles tools were used to find the best compromise (see ***Optimised Automated sample prep*** in **Methods**).

Design of Experiment Step 3: Optimisation and Robustness

To evaluate the performances of the optimised derivatisation, two batches of three analytical replicates were run over two separate days to evaluate *intra*-day (repeatability) and *inter*-days (reproducibility) variability of the method. Repeatability for the tested mixture of fatty acids picolinyl derivatives ranged between 2% and 20% and reproducibility between 6% and 24%. Figure 9 shows the TIC chromatographic peaks for the 6 analytical replicates for picolinyl hexanoate (alkyl chain C6), picolinyl pentadecanoate (alkyl chain C15) and picolinyl methyl arachidoate (alkyl chain 19).

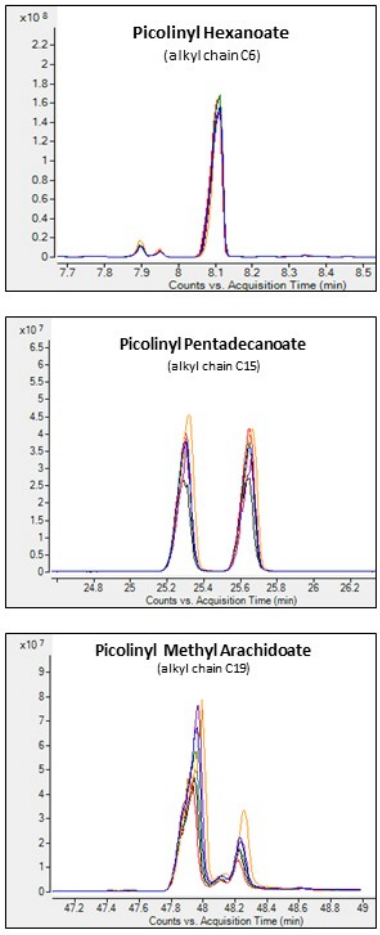

Figure 9: Overlaid TICs of the six analytical replicates for picolinyl hexanoate (top), picolinyl pentadecanoate (middle) and picolinyl methyl arachidoate (bottom)

Conclusions

A robust and time effective solution for the automated derivatisation of fatty acids to picolinyl derivatives was developed in our labs and optimised using design of experiment. DoE helped in developing a deeper understanding of the process and defining an optimum operating region of the experimental space.

Acknowledgements

Nathan Hawkins, Anatune Account Manager, South East of England and Metabolomics Market Specialist.