

從零到一的工程師DOE零基礎指南： 實驗設計賦能研發創新與製程優化



目錄

前言	3
實驗設計簡介	3
實驗設計類型	4
實驗設計應用場景	4
實驗設計實施階段	5
實驗設計工作流	5
實驗設計常用術語	6
實驗設計三大原則	6
案例實戰	7
案例 1：響應曲面設計優化 PCB 設計模擬實驗參數	7
案例背景	7
創建實驗方案	8
採集實驗資料	8
建立迴歸模型	8
應用迴歸模型	10
小結	11
案例 2：基於定製設計實現化工材料配方開發與製程優化	12
案例背景	12
創建實驗方案	12
採集實驗資料	12
建立迴歸模型	13
應用迴歸模型	14
小結	15
案例 3：結合空間填充設計與機器學習優化食品配方與製程	16
案例背景	16
創建實驗方案	16
採集實驗資料	16
標準最小二乘建模	17
SVEM 機器學習建模	18
模型比較	19
小結	20
總結	21
實驗設計方法選擇	21
經典 vs 最優 vs 現代	21
基於模型 vs 模型未知	22
常用實驗設計方法	24
設計診斷	24
JMP 實驗設計優勢	25



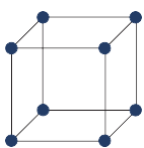
前言

實驗設計簡介

實驗設計 (Design of Experiment) 是研究和處理多因子與反應變數關係的一種科學方法，它透過合理地挑選實驗條件，安排實驗，並透過對實驗資料的分析，從而找出最優的改進方案。從上個世紀20年代費雪 (Ronald Fisher) 在農業實驗中首次提出DOE的概念，到六標準差管理在世界範圍內的蓬勃發展，DOE已經歷了近一個世紀的發展歷程，在學術界和企業界均獲得了崇高的聲譽。

雖然在台灣，實驗設計 (DOE) 這個詞可能聽起來較為陌生，但其實相關概念早在幾十年前就已廣泛應用於工業現場。當時，業界所熟悉的，是另一個名稱：最佳化方法。後來，像「均勻設計」這樣的方法也相繼被提出，並在我國航太領域發揮了關鍵效益。

然而，這些方法大多停留在學術層面，或者只侷限於少數高端產業。真正讓現場工程師能夠落地操作、用來提升產品品質與製程穩定性的 DOE 工具，仍有待普及。田口設計正是在這樣的需求下誕生：它的重點不只是找到「最佳組合」，而是設計出「不怕變異、穩健可靠」的條件，讓品質提升變得有系統、可重現，進而轉化為真正的生產力。



如何最大限度獲取洞察，同時減少實驗次數？

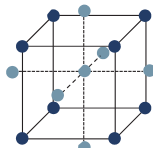


圖 1 實驗設計簡介

從本質上講，DOE是這樣一門科學：研究如何以最有效的方式安排實驗，透過對實驗結果的分析以獲取最大資訊量。所以，DOE有兩大技術支柱：**實驗規劃**和**分析方法**。

- **實驗規劃**：主動、預先地制定方案和計畫，在設計空間中合理布點進行實驗，進行科學、系統、高效的資料採樣
- **分析方法**：分析採集的實驗資料，對反應變數和因子之間的關係建模，輔助決策

一般的實際問題都是紛繁複雜、千變萬化的，但是透過現象看本質，所有實際問題的共同點也可以透過統一的模型來抽象概括。圖2就是一個高度簡化的過程模型，其中Y1、Y2等變數是我們關心的輸出變數，例如品質指標、生產能力和成本等，通常被稱為「反應變數」(Response)；X1、X2等變數是我們在工作中可以加以控制的輸入變數，例如人員、設備、原材料、操作方法和環境等，通常被稱為「可控因子」(Factor)，它們可以是連續型變數，也可以是離散型變數；中間的「暗箱」是「過程」(Process)，在前兩者之間起著銜接轉換的作用，它與不同產業、不同產品、不同技術密切相關，但整體都可以用 $y=f(x)$ 的數學模型來表示。這個數學模型的具體運算式越精準，說明我們對這個過程的理解越深刻，而DOE就是協助我們解密或驗證該數學模型運算式的利器。

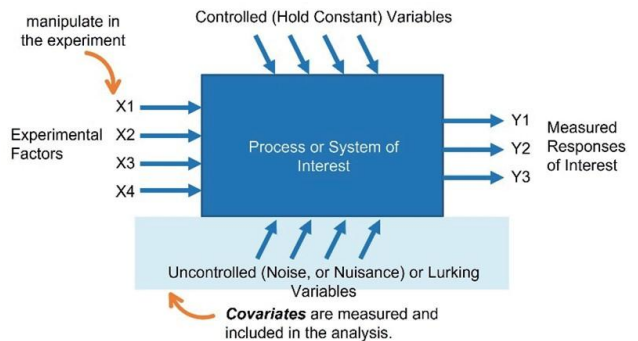


圖 2 過程模型

在某些要求不高的工作環境中，往往不需要用一個複雜的數學運算式來描述過程的全貌，但至少瞭解哪個或哪幾個因子對反應的影響顯著，哪些因子之間存在著相互影響的關係等。這

時，「主效應」(Main Effect)和「交互作用」(Interaction)可以幫助我們回答這些問題。主效應是指一個因子在不同水準下的變化導致反應變數的平均變化量。正如圖3所示，X在-1和+1兩個水準下Y值的落差反映的就是主效應。交互作用是指當其他因子的水準改變時，一個因子的主效應的平均變化量。正如圖4所示，圖4a中因子A對Y的影響沒有受因子B的變化而變化，兩組A與Y的回歸直線完全平行，表明因子A與B之間沒有任何交互作用；反之，圖4b中因子A對Y的影響受因子B的變化而變化，兩組A與Y的回歸直線明顯相交，表明因子A與B之間存在顯著的交互作用。

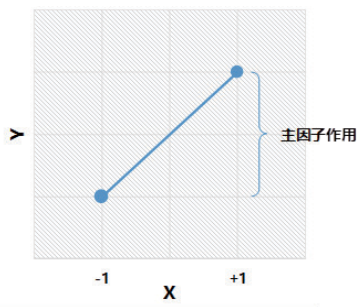


圖 3 主效應示意圖

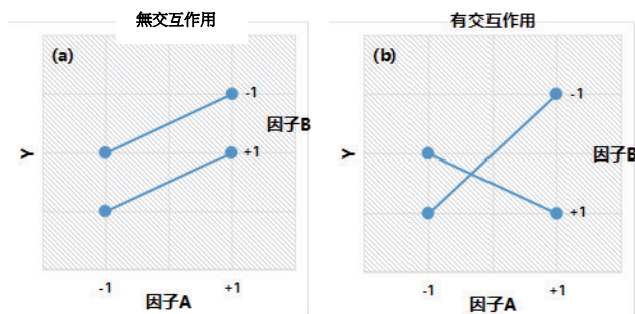


圖 4 因子交互作用示意圖

在DOE方法引入以前，傳統實驗方法往往是採用試錯法 (trial and error)或單因子實驗法 (OFAT, one-factor-at-a-time)⁹。OFAT易於理解和實施，但缺點也很明顯。使用OFAT方法時，因子之間的交互作用是未知的，非線性效應無法衡量，沒有數學方程或模型能夠預測反應變數值，並且由於實驗在設計空間中的分佈不均勻，效率很差。更重要的是，由於其反覆運算和幾乎隨機的樣式，OFAT方法可能需要大量的實驗才能運行⁹。

相比之下，DOE是一種強大的資料獲取和分析工具，更加有效和高效，它允許同時設置多個輸入因子，確定它們對所研究反應變數的影響，並識別在採用OFAT方法時可能遺漏的重要交互作用。DOE經過統計學方法設計，可以採用較少的實驗更好地找到用於最佳化反應的最佳設置，讓您理解因子之間的交互作用，估計

統計模型，以因子及其組合效應的函數形式來預測反應變數，並且允許實驗者根據研究目標確定適當的實驗次數。

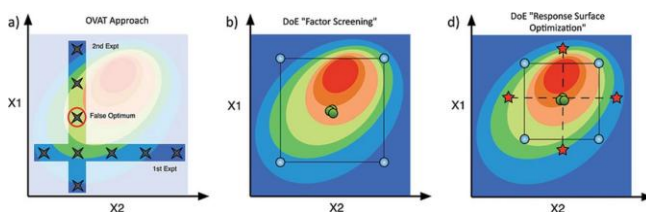


圖 5 OFAT/OVAT vs DOE⁹

實驗設計類型

傳統DOE包括全因子設計、篩選設計、反應曲面設計、混料設計等；與此對應，進階DOE則主要包括定製設計、擴充設計、確定性篩選設計、空間填充設計、非線性設計等。

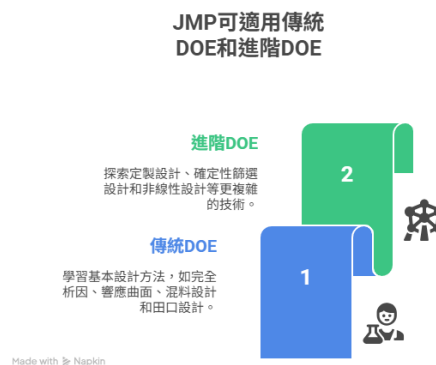


圖 6 實驗設計類型

實驗設計應用場景

無論是在工程技術、品質管制、產品研發等方面，還是在六標準差領域，DOE都是我們解決問題的好幫手，其應用可以說涵蓋了包括機械、半導體、電子、化工、汽車、煙草、醫藥、食品等眾多產業。眾所周知，各類高科技公司的產品本身及其製造工序千差萬別，小到英特爾公司生產的CPU晶片，大到喬治亞宇航中心研製的火箭系統，都離不開DOE的應用。

DOE方法能夠幫助您快速地、可預測地最佳化產品和製程，作出更好、更明智的決策，從而降低業務風險。您還將能夠在充分了解

⁹ Design of Experiments Introduction. Victor Guillier. Medium

⁹ 「A Design of Experiments (DoE) Approach Accelerates the Optimization of Copper-Mediated 18F-Fluorination Reactions of Arylstannanes」, Gregory D. Bowden, Bernd J. Pichler and Andreas Maurer, Sci Rep 9, 11370 (2019). <https://doi.org/10.1038/s41598-019-47846-6>

產品和製程的運作方式之後，再對它們進行轉移，減少投入生產後再改進或「補救」產品或過程的麻煩，讓研發團隊將更多的時間投入到新產品開發、研發創新和製程開發中。





-  **新品配方開發**
確定混料配方成分比例，減少試錯成本，加速新品研發與上市
-  **最佳化反應指標**
建立預測模型，統籌最佳化多個反應指標
-  **高效篩選因子**
最小化實驗規模，識別關鍵因子，指導研發設計與品質問題解決
-  **確定製程窗口**
探索設計空間和最佳製程視窗，滿足產品規格限要求

圖 7 實驗設計應用場景

實驗設計實施階段

在產品研發與製程最佳化過程中，DOE的實施通常包括4個階段。可以根據各個階段的分析目標選取適合的實驗設計方法。

- 第一步，篩選階段（浪裡淘沙）：從大量因子中高效篩選，識別可以顯著影響反應變數的重要因子。
- 第二步，描述階段：建立因子和反應變數之間關係的初步模型，對研究的系統或過程進行描述。
- 第三步，最佳化階段：重點最佳化少數關鍵因子，縮小設計空間範圍，考慮更複雜的模型效應，對過程模型進行更為精細的刻畫，實現反應變數最佳化。
- 最後，穩健階段：研究如何穩定、高效地生產出高品質的產品，降低雜訊因子的影響，減小產品品質的變異性，提高實驗的穩健性。

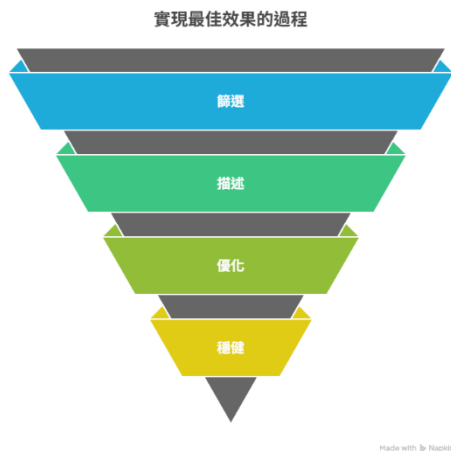


圖 8 實驗設計實施階段

實驗設計 workflow

儘管DOE的應用場景千變萬化，但實驗設計的執行步驟其實是類似的，基本上可以概括為五大步驟，即實驗設計的五部曲。在每個階段的DOE實施過程中，都可以依據DOE的工作流來生成實驗方案並構建模型，可透過反覆運算改進，持續最佳化產品或流程，並加深對過程的理解。在整個DOE流程中，工程知識和經驗的重要性不言而喻，數學方法和電腦分析工具是賦能這一流程的關鍵。透過精確地描述、設計、採集、擬合和預測，DOE能夠幫助工程師們更高效地達成實驗目標，最佳化產品設計或生產流程。

描述 (Describe)：確定實驗目標，全盤規劃實驗設計輸入，具體包括以下方面：

- 定義問題和目標；
- 識別一個（或多個）反應變數；
- 確定測量系統是否有能力；
- 識別實驗因子和因子水準；
- 識別對隨機化的限制以及其他約束；
- 描述因子和反應變數之間關係，確定可能顯著的效應項；
- 確定可以允許的實驗資源和次數等。

設計 (Design)：根據實驗設計輸入，生成實驗設計方案。

收集 (Collect)：根據實驗方案設定因子狀態，運行實驗，收集反應變數的實驗結果資料。

擬合 (Fit)：擬合和診斷統計模型，定量刻畫因子和反應變數的關係。

預測 (Predict)：應用模型，根據因子預測反應變數，準確地估計實驗結果。得出結論，並針對後續步驟提供建議。

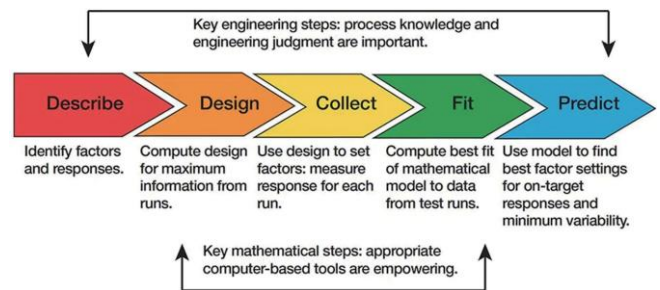


圖 9 實驗設計 workflow

實驗設計常用術語



圖 10 實驗設計常用術語

實驗設計三大原則

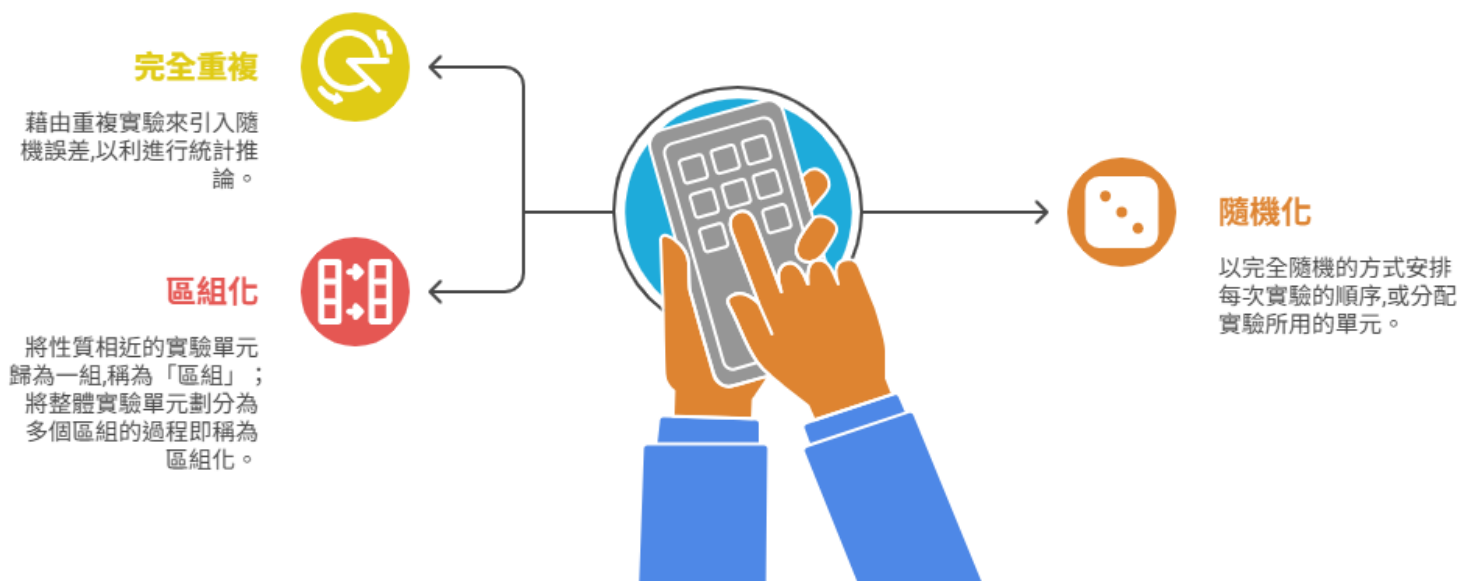


圖 11 實驗設計三大原則

JMP DOE隨選精選課程

DOE入門：從方法到實踐·60分鐘輕鬆掌握DOE

https://www.jmp.com/zh_cn/events/ondemand/local/jmp-doe-from-zero-on-demand.html

DOE經典案例實踐：產品研發與製程優化的3種實驗設計方法

<https://www.jmp.com/zh-hant/resources/on-demand/non-series/doe-2024>



DOE入門課



DOE經典案例實戰

案例實戰^④

案例1：應用反應曲面設計最佳化PCB設計模擬實驗參數

案例背景

高速數位PCB設計的元器件參數、元器件佈局等因素，都會影響信號完整性(Signal Integrity)。通常透過模擬實驗獲取數位信號的眼圖評估數位信號的品質。而模擬實驗參數眾多，模擬計算時間長，透過窮舉模擬參數，來預測量產系統的信號完整性是十分低效的，也是不現實的。應用實驗設計確定眼圖模擬實驗參數，可顯著降低實驗次數，並可透過建立代理模型加速高速數位PCB設計。

本案例基於10Gb/s差分PCB介面設計場景，目標為最佳化發送端端端阻抗、接收端端端端阻抗、差分阻抗、線路長度和均衡係數以最大化眼圖的眼高和眼寬。

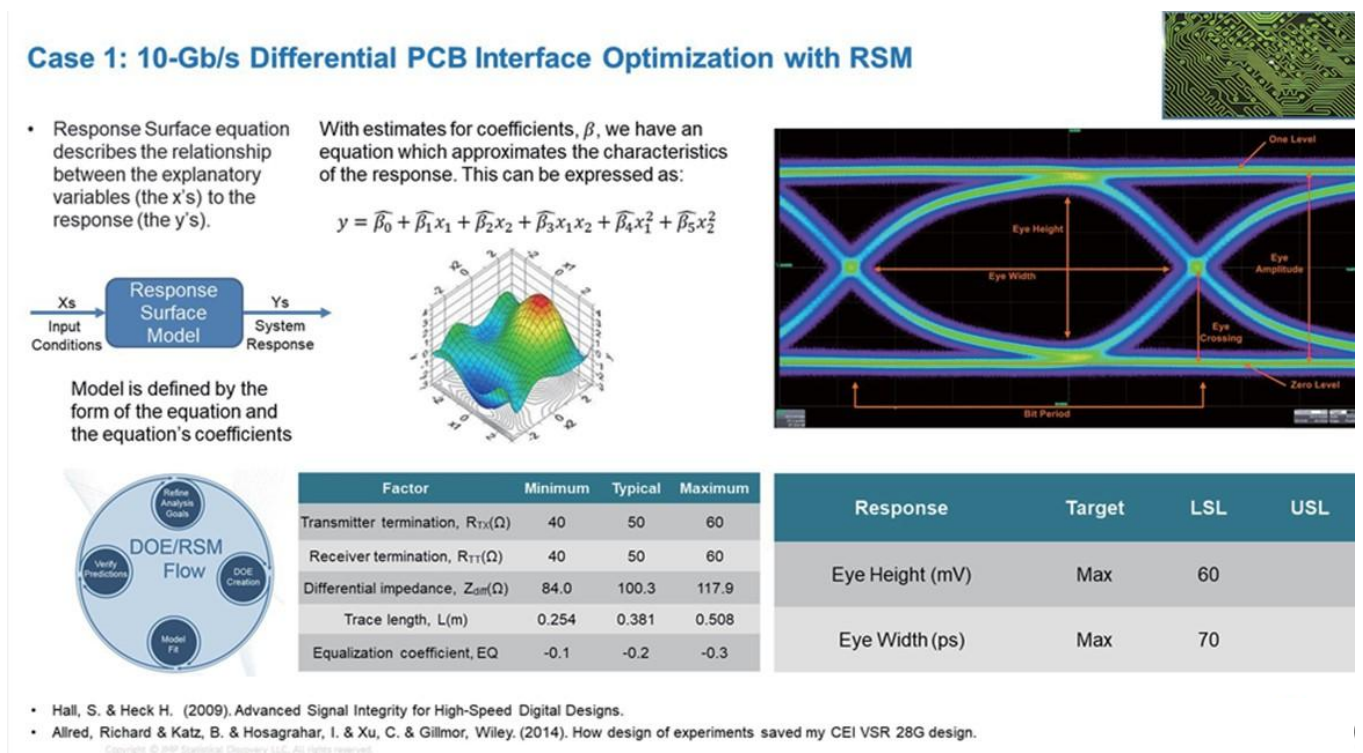


圖 12 案例背景

④ 關於本章節的3個實戰案例，可透過《DOE經典案例實戰：產品研發與製程最佳化的3種實驗設計方法》公開視頻課程瞭解詳細實現過程。

創建實驗方案

JMP功能表列-實驗設計-經典-反應曲面設計，根據實驗背景輸入下圖中的內容，創建反應曲面設計，採用中心複合表面設計 (CCF)，添加2個中心點，生成28次的實驗設計方案。



圖 13 反應變數曲面設計

收集實驗資料

根據實驗方案設置實驗參數，收集28次眼圖模擬實驗的眼高和眼寬資料。

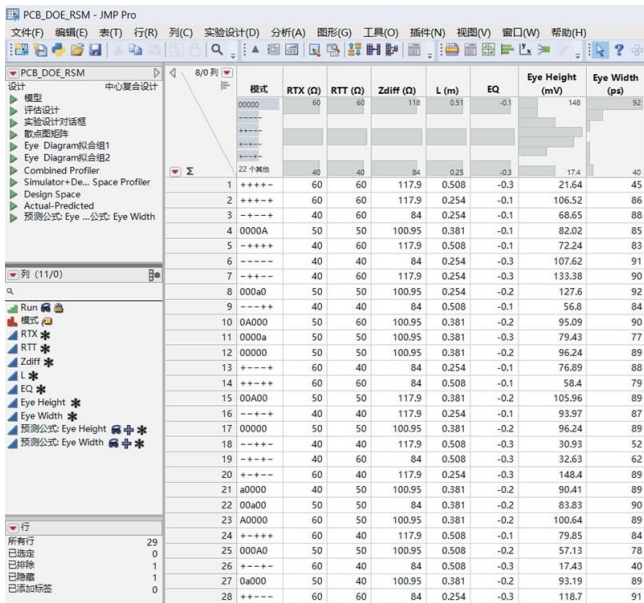


圖 14 採集實驗資料

透過散點圖矩陣查看設計空間布點情況。

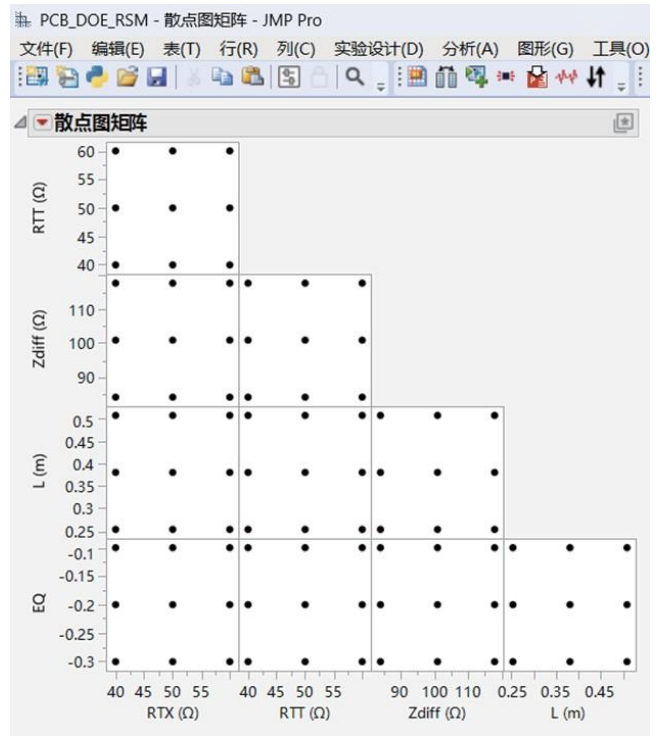


圖 15 實驗點分佈

建立迴歸模型

JMP功能表列-分析-擬合模型，將眼高和眼寬作為Y，採用默認的反應曲面設計的模型效應。採用預設的標準最小平方法建立回歸模型，因為具有眼高和眼寬兩個反應變數，建議勾選分別擬合。

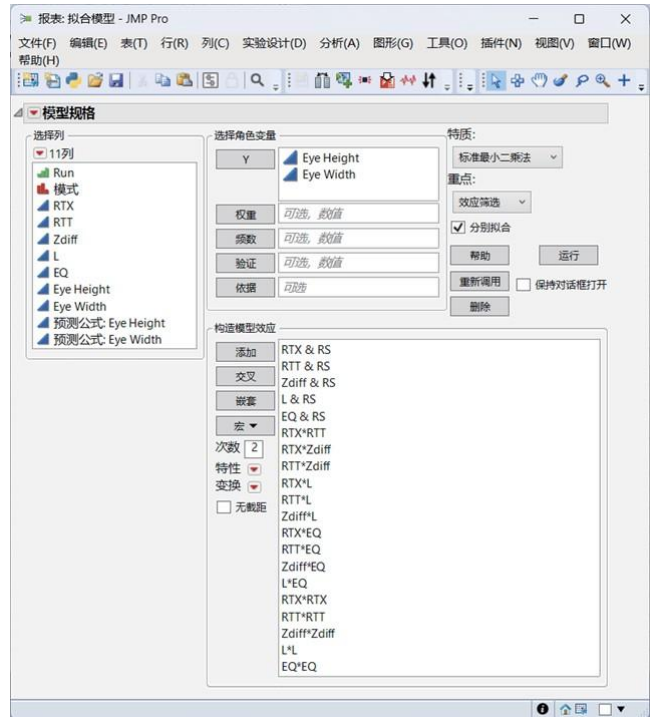


圖 16 擬合模型

對初始模型分別進行變數刪減。大多數情況下，以 p 值= 0.05為界，一般認為 p 值大於0.05的效應項是統計不顯著的，因此依次刪去 p 值大於0.05的效應項，注意層級原則 (Hierarchy Principle，一般不建議在刪除高階交互項之前，刪去所含的主效應)。

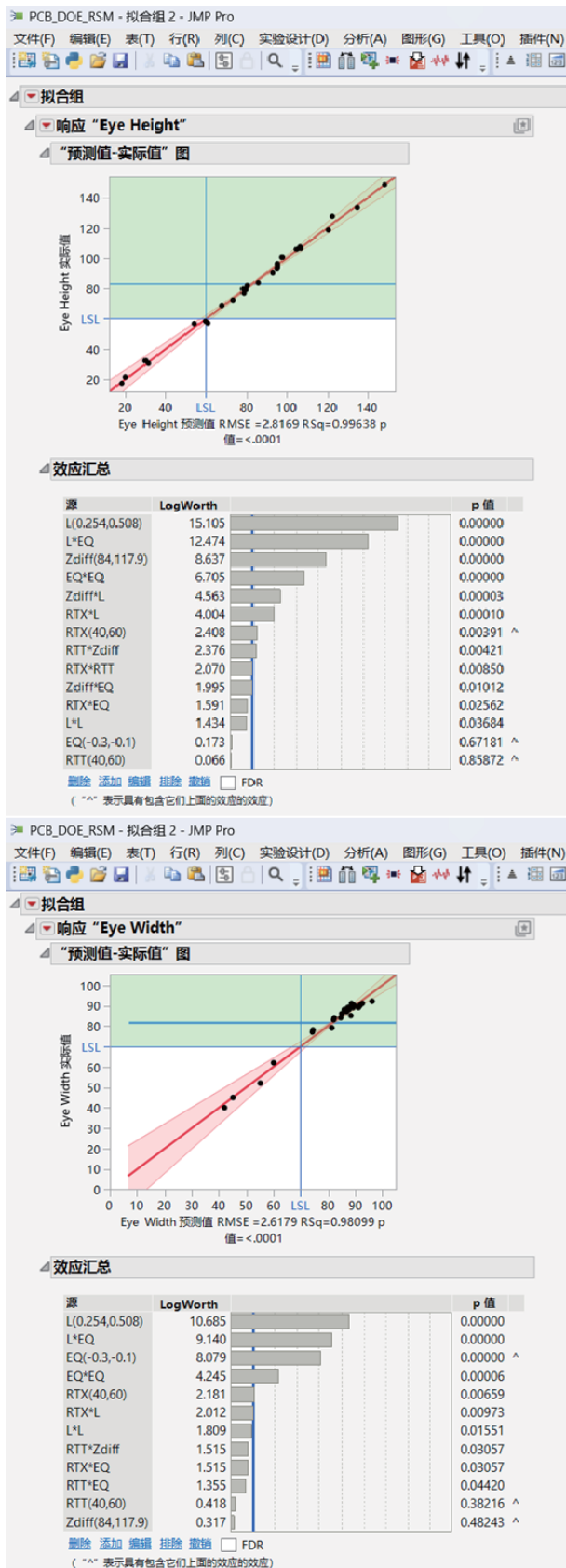


圖 17 效應匯總

對簡化後的模型，可以從R方、調整R方、均方根誤差、變異數分析報表F檢驗結果、失擬報表F檢驗結果、殘差圖等多個維度判斷模型的好壞。如下圖所示，對於眼高和眼寬，R方和調整R方都比較高 (>0.9)，變異數分析顯著 (p 值<0.05)，殘差隨機分佈滿足正態性，因此可進行下一步分析。

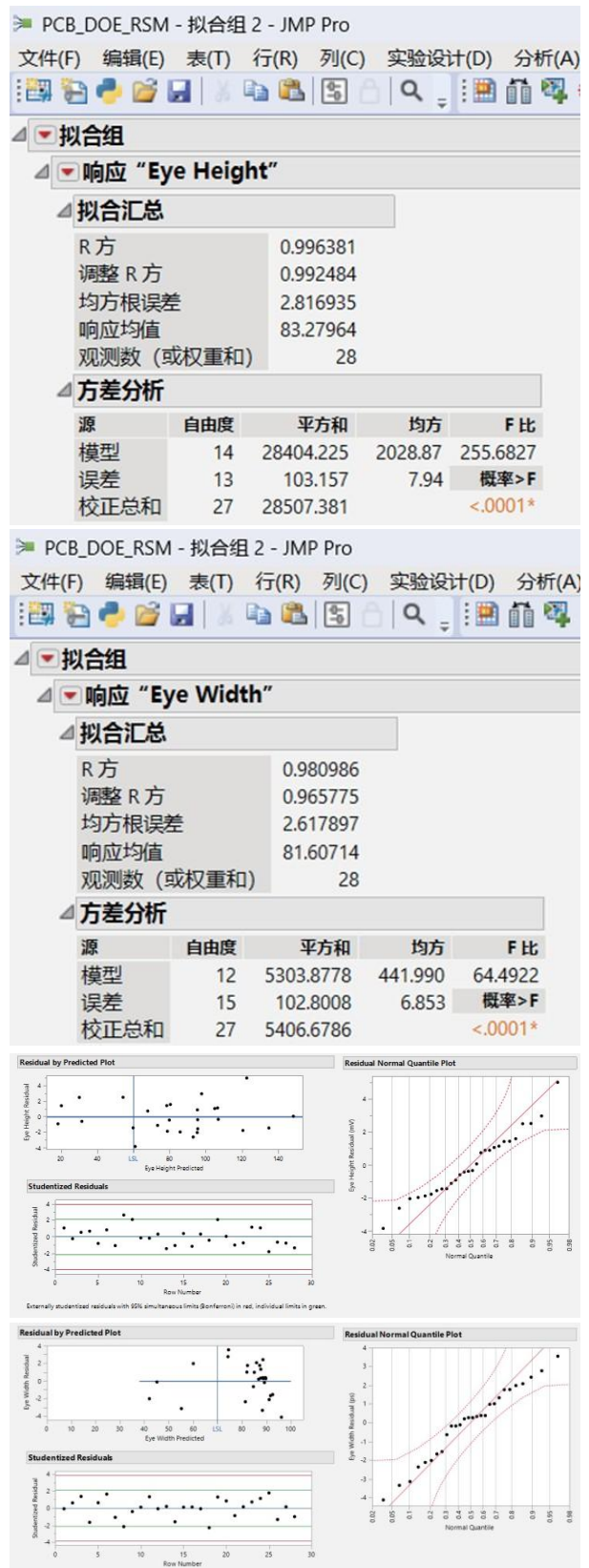


圖 18 模型評估

應用迴歸模型

反應變數預測與多反應變數同步最佳化

建立有效的回歸模型後，即可透過模型公式對反應變數進行預測，同時透過預測刻畫器確定最佳製程參數組合。

當研究中有多個反應存在時，一般很難找到一個參數組合的取值，讓多個反應的結果均達到最優。如本例，眼高和眼寬分別取最大值時，參數組合的取值會有一定的差異。這時我們可以利用 JMP 軟體的最大化意願功能，自動找到一個參數組合，使得多個反應變數盡可能同時接近要求的目標值，此外還可以透過為多個反應變數設定不同權重，實現定向的最佳化策略。

調出JMP刻畫器(擬合組紅三角-刻畫器)，尋找平衡最優解(預測刻畫器紅三角-最佳化和意願-意願函數、設置意願、最大化意願)，自動得到如下圖所示的最優參數組合(RTX=60, RTT=40, Zdiff=117.9, L=0.254, EQ=-0.215)下，各個反應變數都可以盡可能地滿足或超出預先設定目標。

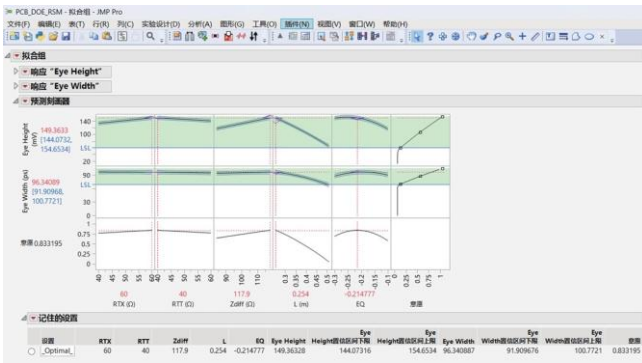


圖 19 多個反應的最優參數組合確認

透過分別保存眼高和眼寬回歸模型的預測公式，並透過預測刻畫器將最優參數組合設置追加至表(預測刻畫器紅三角-因子設置-將設置追加至表)，可以得到最優參數組合情況下的眼高和眼寬預測值，並可透過散點圖視覺化實現與實驗資料點的比較，對當前參數組合設置的可行性進行判斷。

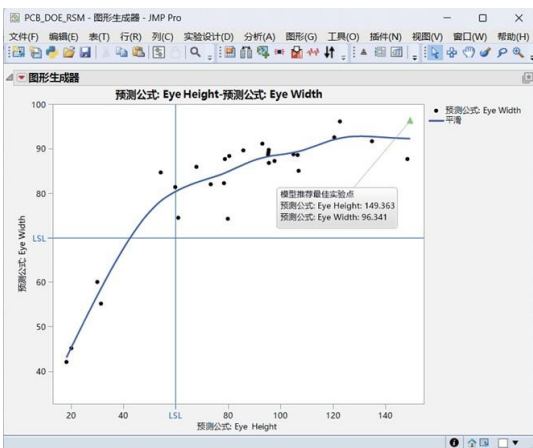
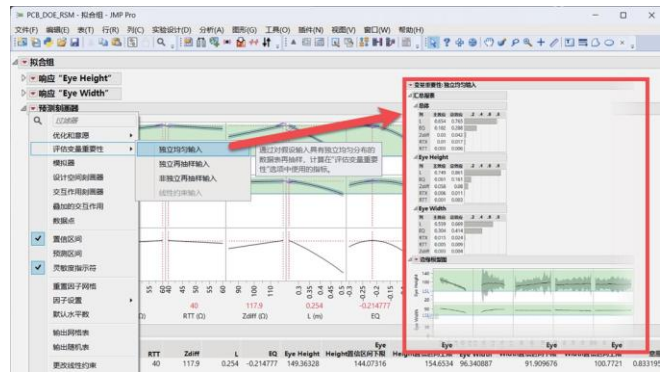


圖 20 最優參數組合反應變數預測

識別重要因子

如下圖所示，可以透過預測刻畫器評估變數重要性，識別出重要因子，以對其進行重點調控。



定義設計空間

在產品實際生產過程中，某些製程參數往往存在小範圍的波動，為了滿足規格限要求，需要定義可行的設計空間。

JMP 一般有三種方式定義設計空間，第一種是透過等高線圖，第二種是利用蒙特卡洛模擬，第三種是透過設計空間刻畫器。等高線圖由於是二維圖表，因此對於研究2個參數的場景比較有效，目前用的越來越少。而蒙特卡洛模擬的應用範圍更為廣泛，可以針對2個和多個參數的場景。設計空間刻畫器是 JMP 17版本推出的新功能，可以結合蒙特卡洛模擬自動縮小設計空間範圍，透過不斷點擊內移按鈕，軟體自動尋找能使缺陷比例最大程度減少的參數及收縮範圍，從而快速的確定符合要求的參數空間，同時評估當前參數組合的預測反應變數滿足規格限要求的情況。需要注意的是參數比較多的模型，往往最佳化的參數空間不止一個，所以可嘗試多執行幾次選擇最優。

方法一：選用蒙特卡洛模擬的方法，先將參數設定在最大化意願組合點(紅三角-最佳化和意願-最大化意願)，然後調出模擬器

(預測刻畫器紅三角-模擬器)，將每個參數的輸入都設為隨機，分佈類型選擇正態分佈或均勻分佈。以分佈類型設置為均勻分佈為例，在輸入規格限之後(模擬器紅三角-規格限)，按一下類比，透過缺陷率報表判斷當前參數組合的可行性，並可手動調節模擬器參數上下限，確定可行的製程視窗。

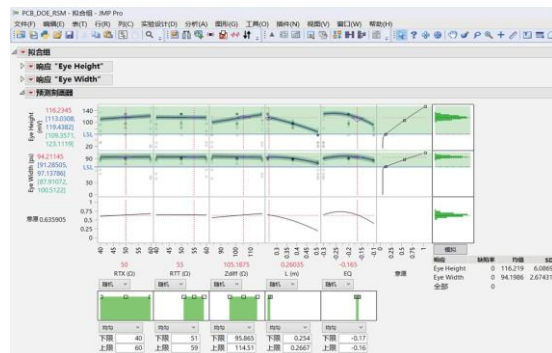


圖 21 定義設計空間

方法二：調出設計空間刻畫器（預測刻畫器紅三角-設計空間刻畫器），生成並連接隨機表，嵌入因子和反應變數空間散點圖。透過在設計空間刻畫器不斷內移，可以實現設計空間範圍的縮小，並可同步觀察到規格中對應部分比例的提升。縮小到合適的因子範圍後，可以結合預測刻畫器和模擬器再次確認當前參數的可行性。

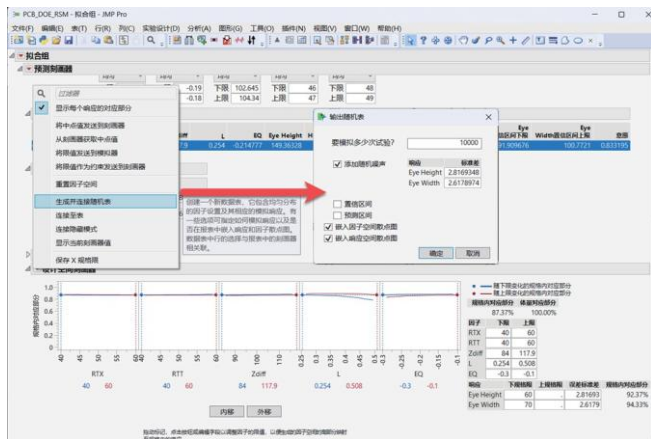


圖 22 生成並連接隨機表

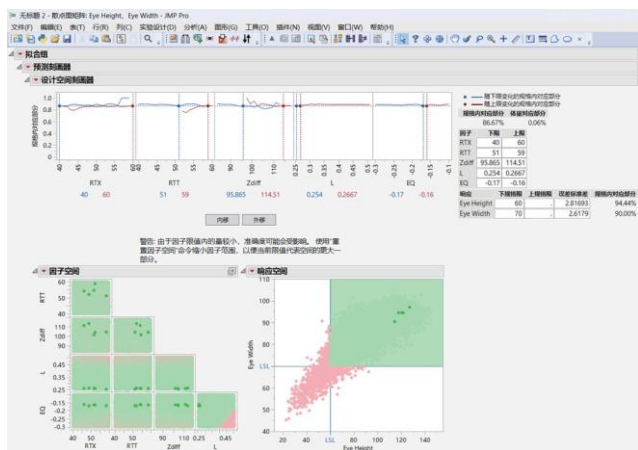


圖 23 設計空間刻畫

在確定設計空間時，上述提到的手動和自動方式都可使用，最終的參數範圍也可能不同。但手動的方式的一個優勢是可以透過人為評估參數調節的難易程度，來引導設計空間的確定，所以建議比較後擇優選用。

小結

本例透過應用反應曲面設計生成了28次的實驗方案，基於模擬結果建立了代理模型，識別出了關鍵影響因子，確定了最佳的實驗參數範圍，同時最佳化了眼高和眼寬。借助JMP軟體實驗設計平台可以輕鬆生成實驗設計方案，高效建立回歸模型，並透過預測刻畫器、模擬器、設計空間刻畫器進行模型應用，實現反應變數預測和多反應變數同步最佳化、識別重要因子和定義設計空間。



案例2：基於定製設計實現化工材料配方開發與製程最佳化

案例背景

在化工領域，配方開發非常具有挑戰性，如果不進行科學系統的研究，往往需要進行大量實驗，且難以摸清其中的規律，費時費力。採用實驗設計方法可以科學系統地研究化工過程，透過建立模型描述和理解其規律，從而加速產品研發和助力製程改進。

本案例基於化工行業導電薄膜配方開發和製程最佳化場景，目標為最佳化硫酸銅、亞硫酸鈉和乙二醛3種混料成分比例，並最佳化波長、溫度和時間3個過程變數參數，實現導電薄膜阻尼最大化和成本最小化。

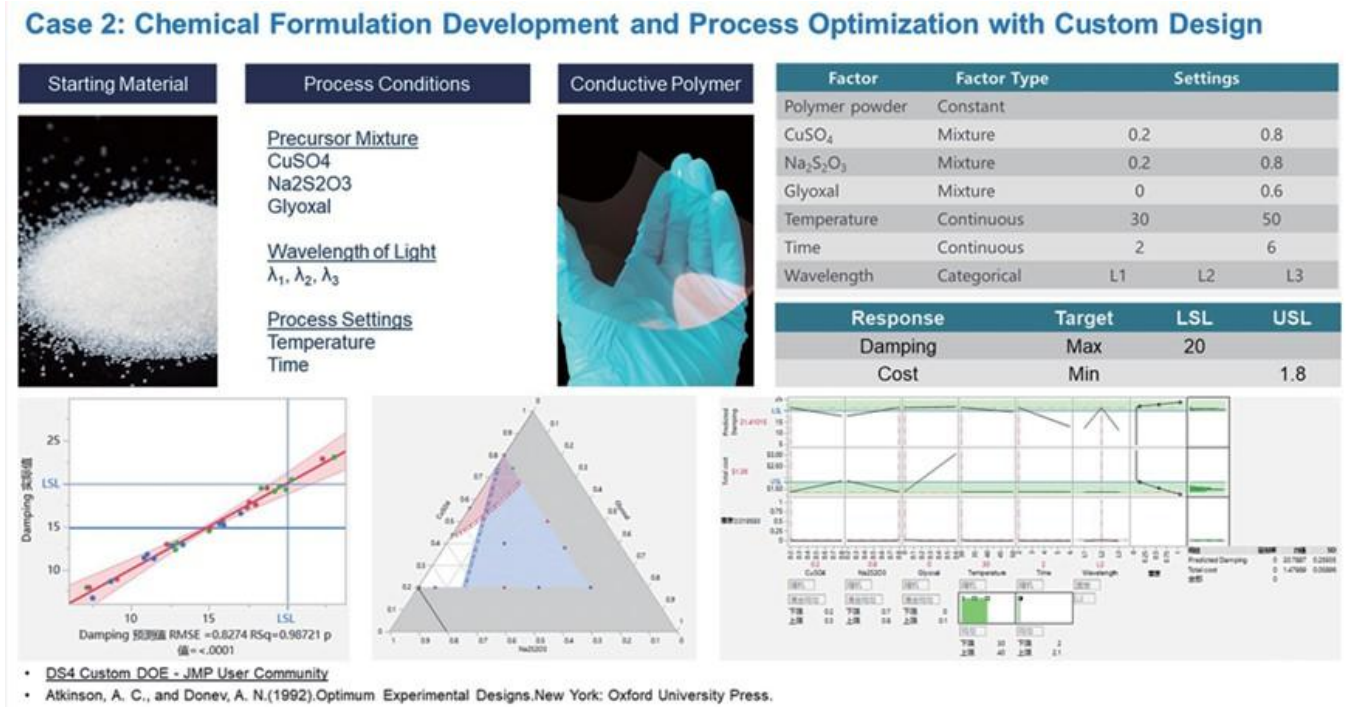


圖 24 案例背景

建實驗方案

本案例同時具有混料因子和過程變數因子，因此需要採用定制設計創建實驗方案。如下圖所示，設置因子和反應變數後，考慮二階交互作用，採用I最優性準則的定制設計，添加3個中心點，生成30次實驗的方案。



圖 25 定制設計

採集實驗資料

根據實驗設計方案進行實驗，採集導電薄膜阻尼資料，並根據混料成分比例和原料價格公式計算得到總成本資料。

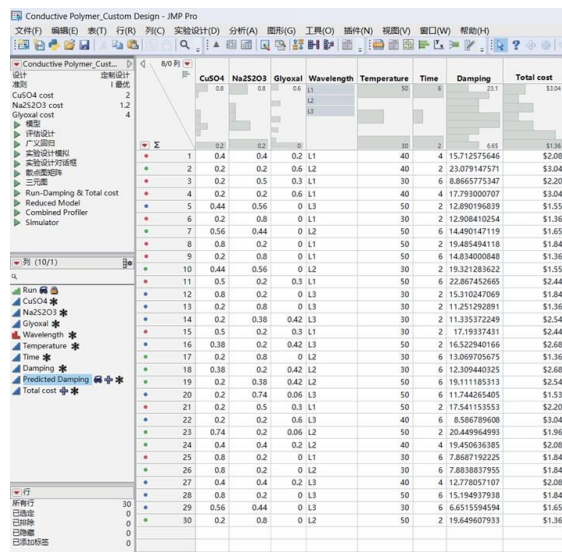


圖 26 採集實驗資料

透過散點圖矩陣可以直觀觀察到設計空間的布點情況。對於混料設計，可以透過三元圖觀察混料因子的成分比例。

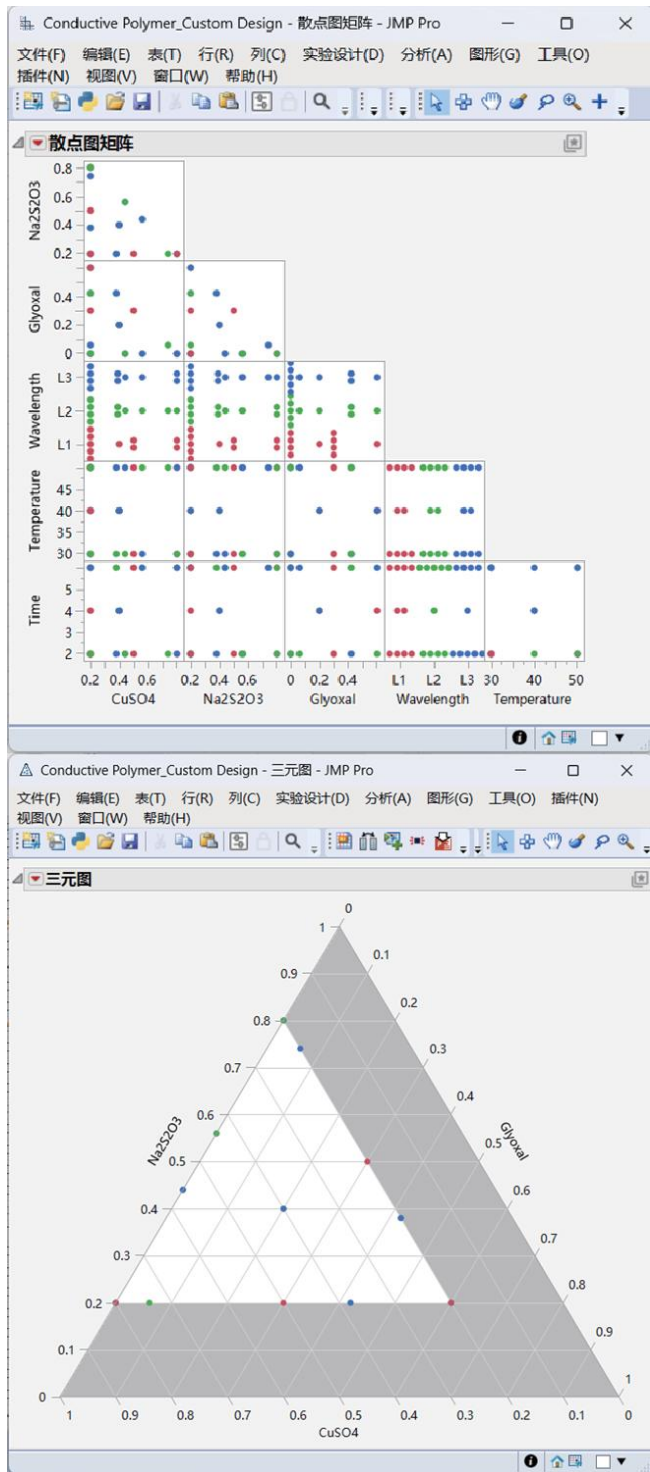


圖 27 實驗點分佈

建立迴歸模型

透過標準最小平方法建立阻尼的迴歸模型，並根據p值是否大於0.05對模型效應項進行刪減，得到簡化後的模型。

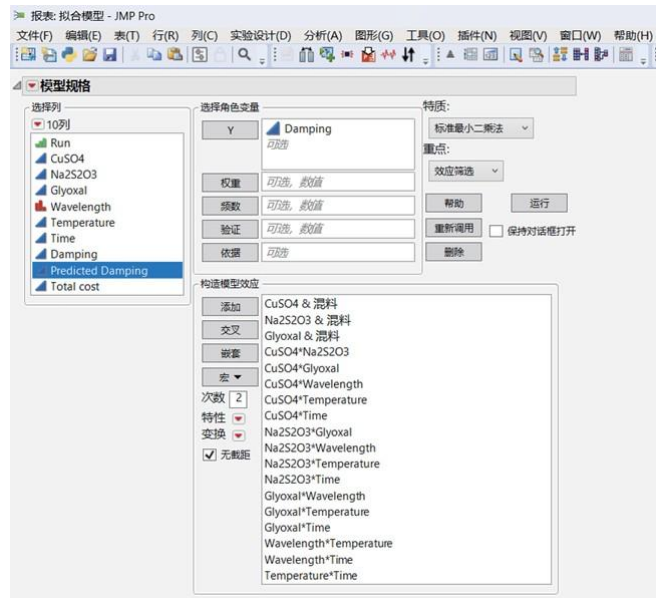


圖 28 擬合模型

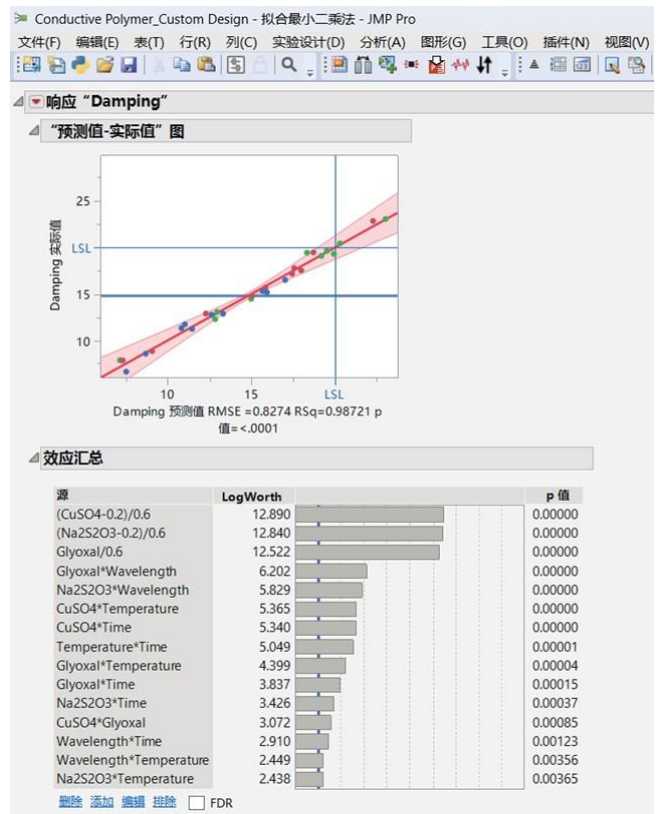


圖 29 效應匯總

透過R方、調整R方、均方根誤差、變異數分析報表F檢驗結果、失擬報表F檢驗結果、殘差圖等多個維度判斷模型的好壞。如下圖所示，對於阻尼，R方和調整R方都比較高(>0.9)，方差分析顯著 (p值<0.05)，殘差隨機分佈滿足正態性，因此可進行下一步分析。

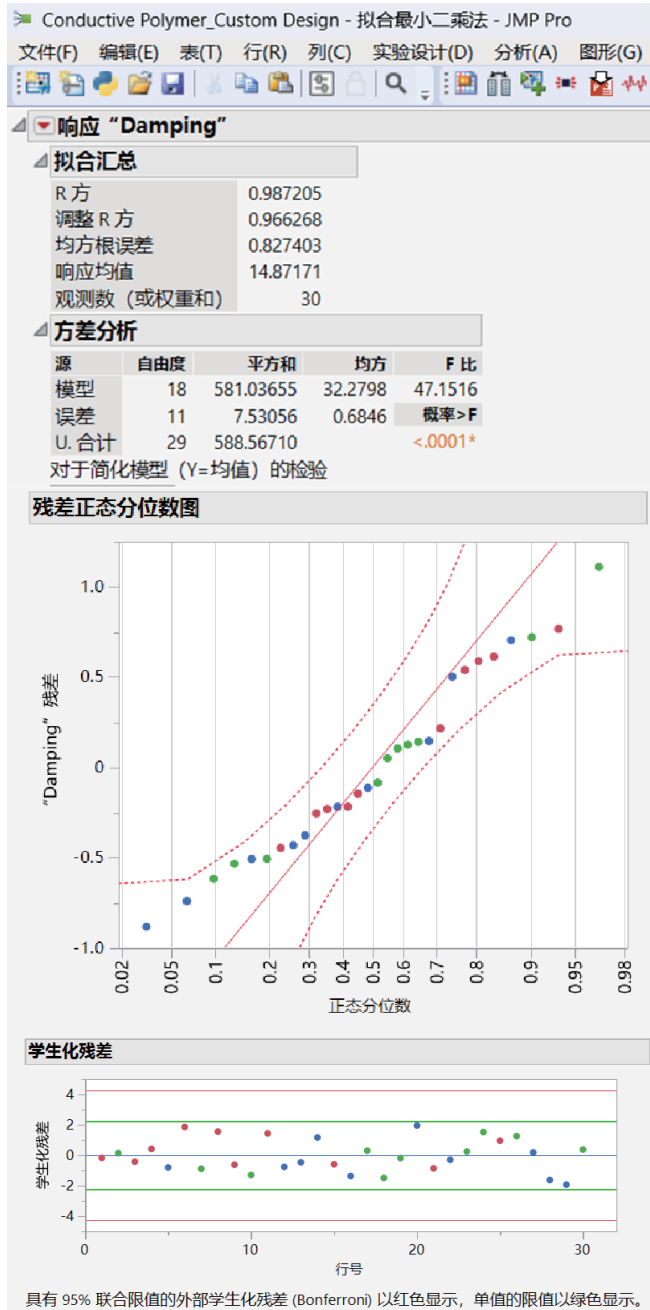


圖 30 模型評估

應用迴歸模型

反應預測與多反應同步最佳化

透過保存阻尼預測公式，並結合總成本的計算公式，可以形成組合刻畫器 (圖形-刻畫器-添加公式列)，以此尋找阻尼和總成本的平衡最優解 (預測刻畫器紅三角-最佳化和意願-意願函數、設置意願、最大化意願)。自動得到如下圖所示的最優參數組合，各個反應變數都可以盡可能地滿足或超出預先設定目標。

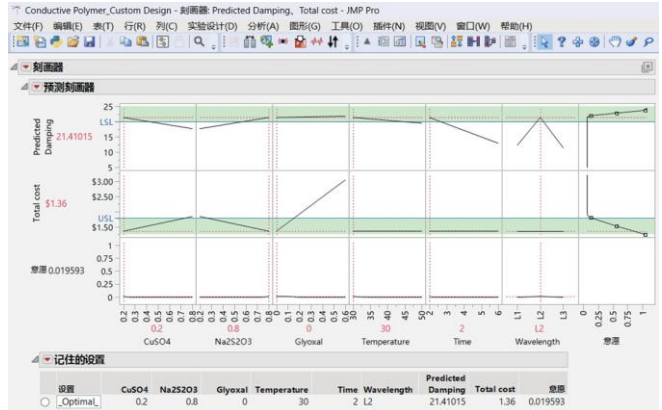


圖 31 多個反應的最優參數組合確認

透過預測刻畫器將最優參數組合設置追加至表 (預測刻畫器紅三角-因子設置-將設置追加至表)，可以得到最優參數組合情況下的阻尼預測值和總成本計算值，並可透過散點圖視覺化實現與實驗資料點的比較，對當前參數組合設置的可行性進行判斷。當前推薦參數組合的反應預測值可以同時滿足兩者的規格限要求，優於之前的實驗資料點。

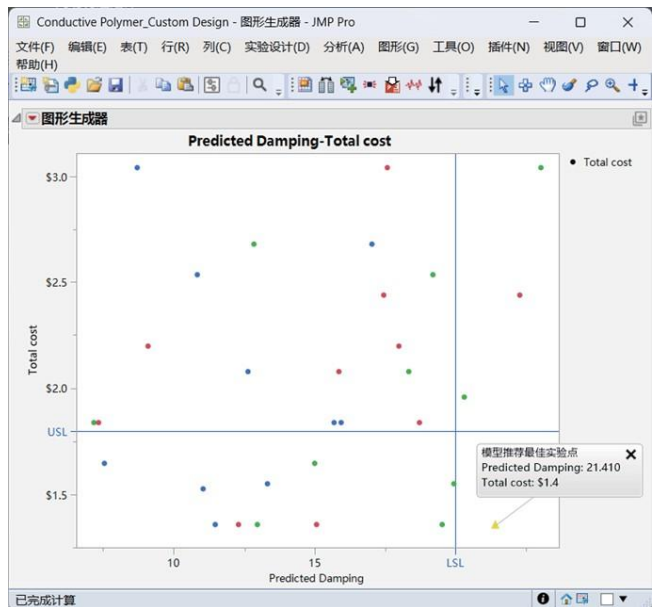


圖 32 最優參數組合反應變數預測

定義設計空間

結合預測刻畫器、模擬器和混料刻畫器，可以直觀地探索最佳參數範圍，透過關聯刻畫器和傳播因子設置，可以保持預測刻畫器和混料刻畫器的因子設置聯動，從而更高效地確定混料因子最佳成分比例。混料刻畫器的白色區域可以清晰直觀地輔助確定設計空間。

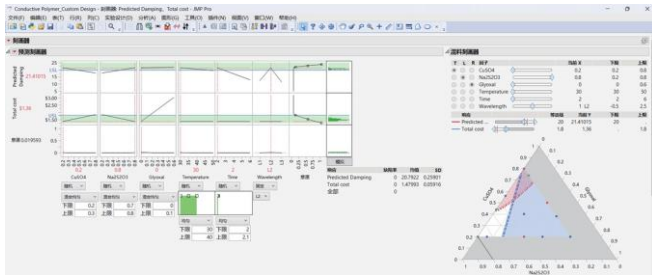


圖 33 定義設計空間

小結

目前化工行業研究的場景越來越複雜，基於傳統的實驗設計方式來進行製程開發和最佳化，往往是非常低效的，並且傳統的混料設計也無法應對存在過程因子的情況。借助JMP獨有的定制設計方法，可以應對更為複雜和有挑戰性的實驗設計場景，並有效地利用少量實驗達成研究目標，從而大大縮短研發時間。本案例應用定制設計生成了30次的實驗方案，透過擬合模型確定了最佳配方成分比例和過程變數參數設置，同時最佳化了阻尼和成本，顯著加速了導電薄膜材料研發與量產。



案例3：結合空間填充設計與機器學習最佳化食品配方與製程

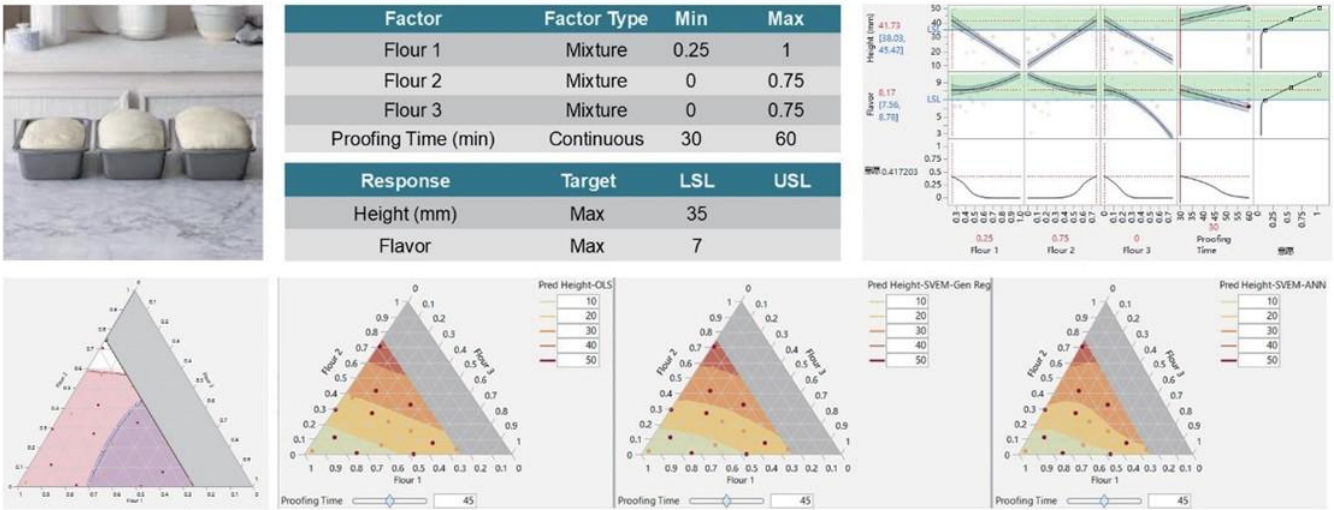
案例背景

在食品研究領域，食品的風味、口感、外觀是非常重要的評價指標，同時也要考慮加工難易程度、成本等一系列的因素，以滿足消費者口味需求，並實現高效品控。

本案例基於食品行業麵包配方研發和製程最佳化場景，目標為最佳化3種麵粉混料成分的比例和麵團醒發時間，實現麵包高度最大化和口感評分最大化。

Case 3: Combining Space Filling Design with Machine Learning to Optimize Food Formulation

Bread Experiment: Identify mixture proportions and process settings that will make bread that is “tall” and tasty



- [Designing Mixture Experiments - Part 2 - JMP User Community](#)
- [Accelerating Innovation with Space Filling Mixture Designs, Neural Networks and SVM](#)
- Snee, Ronald, and Roger Hoerl. 2016. *Strategies for Formulations Development: A Step-by-Step Guide Using JMP*. Cary, NC: SAS Institute Inc.

圖 34 案例背景

創建實驗方案

在模型未知的情況下，本案例創新性地應用快速靈活填充設計這種現代實驗設計方法，生成了18次的空間填充實驗設計方案。在實驗設計時，無需提前考慮模型效應項。



圖 35 空間填充設計

採集實驗資料

根據實驗設計方案進行實驗，採集麵包高度和口感評分資料。

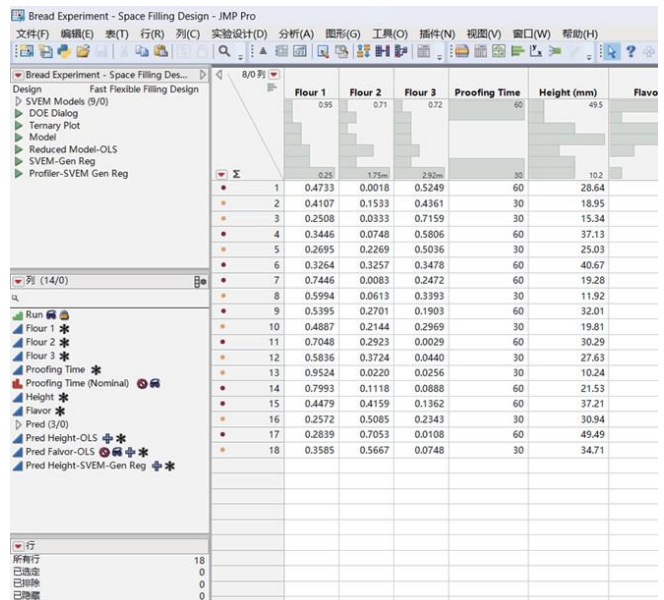


圖 36 採集實驗資料

透過三元圖觀察3種麵粉的混料成分比例，可以看到空間填充設計的實驗點更均勻地分佈在設計空間內部。

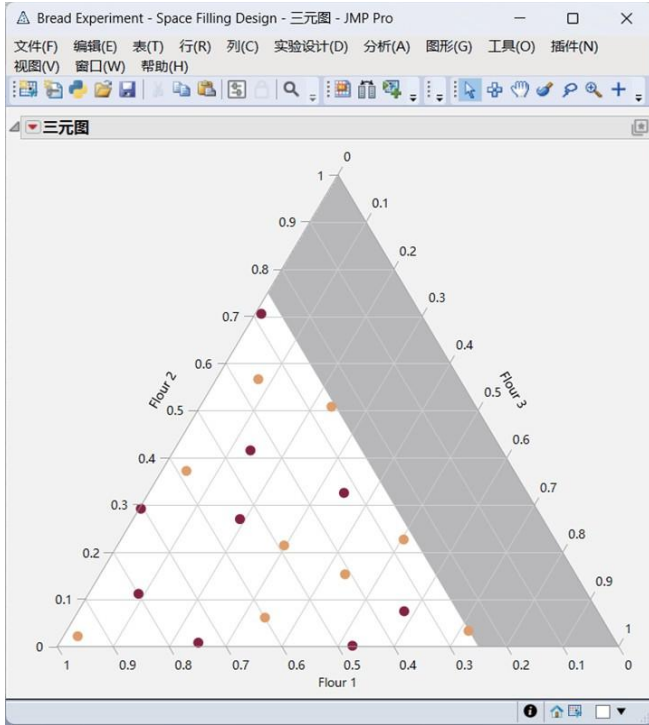


圖 37 實驗點分佈

標準最小平方法建模

採用標準最小平方法建立回歸模型，建立了配方成分比例、醒發時間與麵包高度、評分之間的預測公式，並透過了模型診斷，採用預測刻畫器實現了反應變數預測與多反應變數同步最佳化，並透過混料刻畫器確定了設計空間。

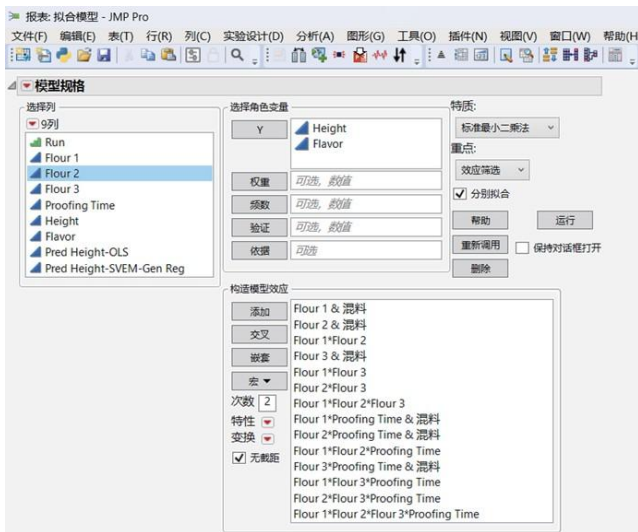


圖 38 擬合模型

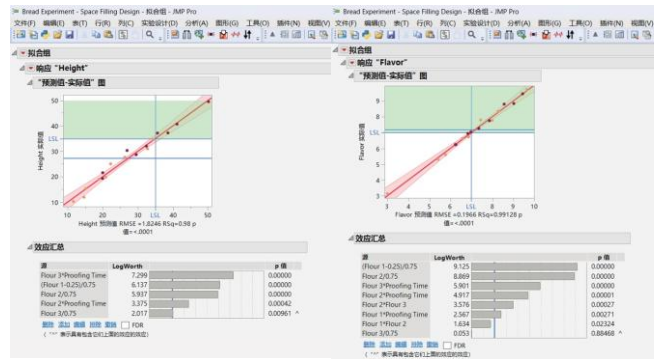


圖 39 效應匯總

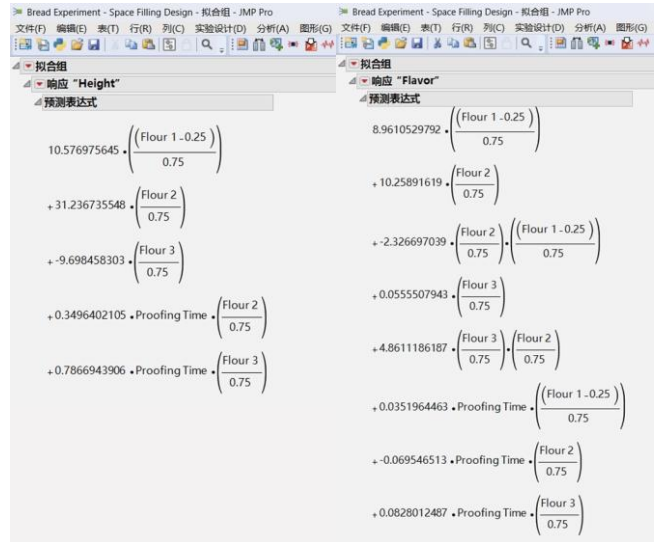


圖 40 模型預測運算式



圖 41 模型評估

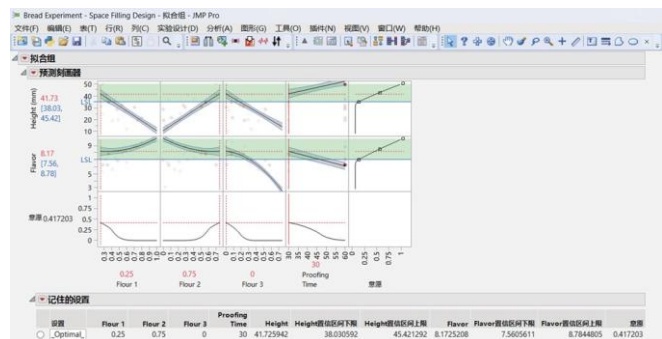


圖 42 多個反應的最優參數組合確認

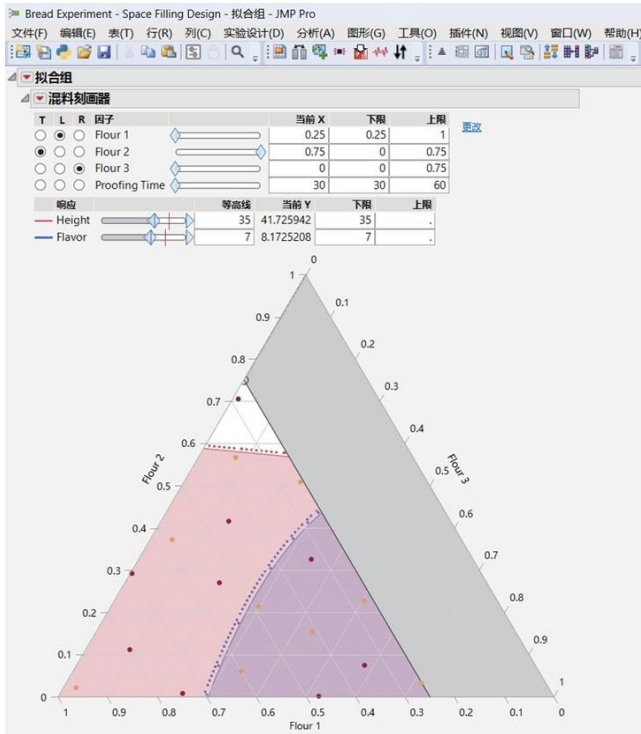


圖 43 定義設計空間

SVEM機器學習建模

由於模型未知，本案例進一步採用機器學習方法建立模型，以對比標準最小平方、SVEM廣義回歸和SVEM神經網路的建模結果。對於實驗設計產生的小樣本資料集，採用SVEM(自驗證集合模型)無需劃分資料集，而是對不同實驗資料點賦予不同的訓練和驗證權重，從而充分利用實驗設計資料。透過建立集成模型，可以提高模型預測精度和穩健性。

SVEM廣義回歸模型

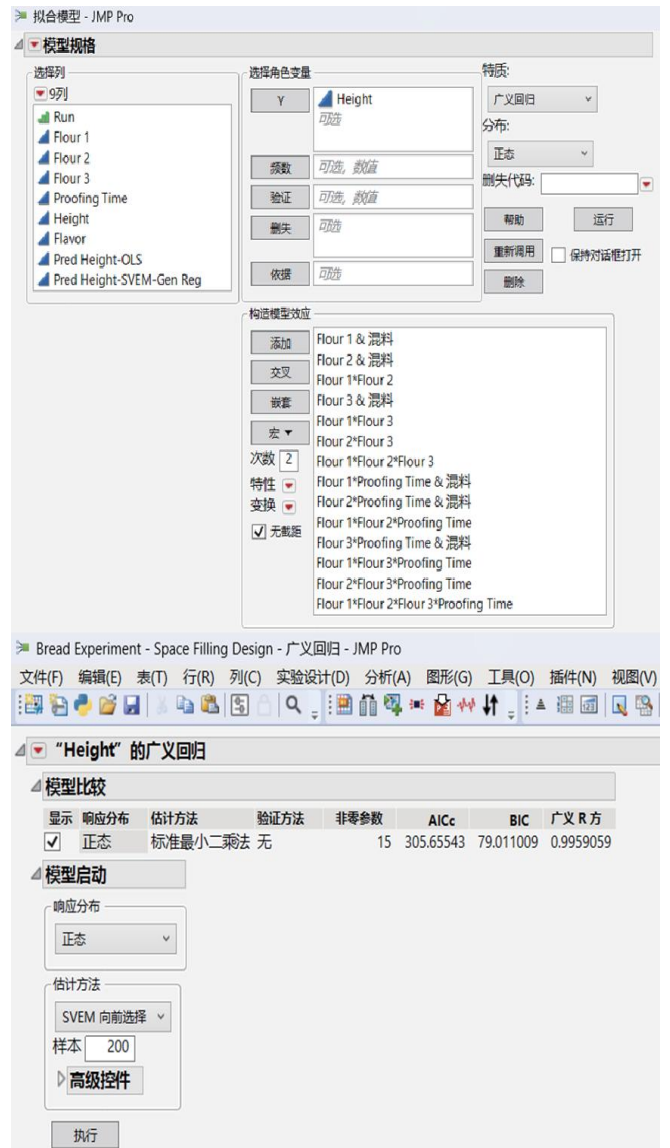


圖 44 擬合SVEM廣義回歸模型

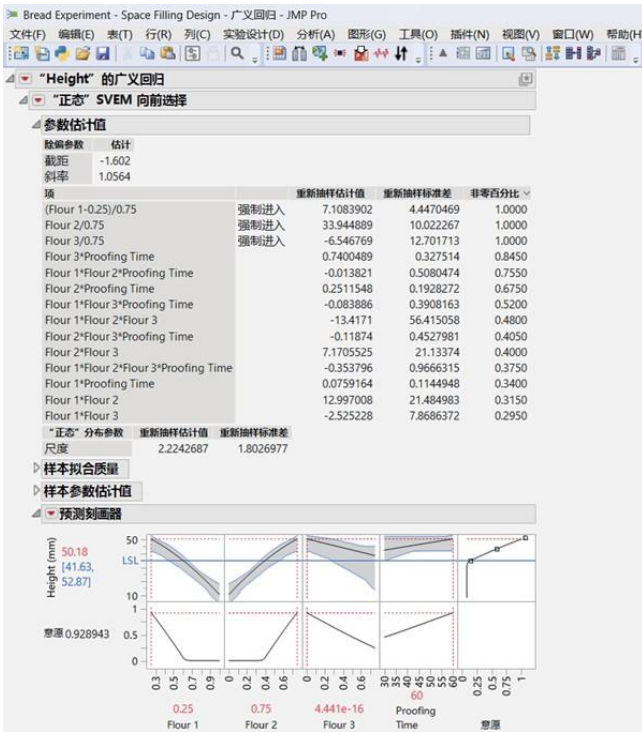


圖 45 SVEM廣義回歸模型

SVEM神經網路模型

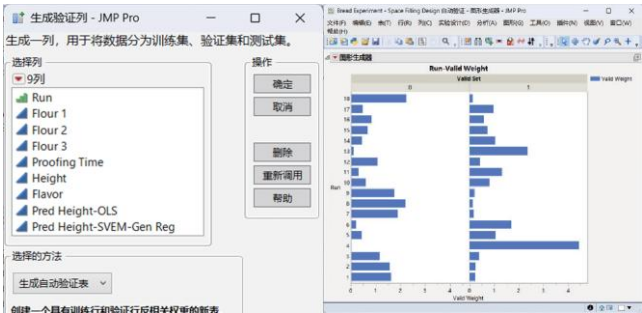


圖 46 生成自動驗證列表

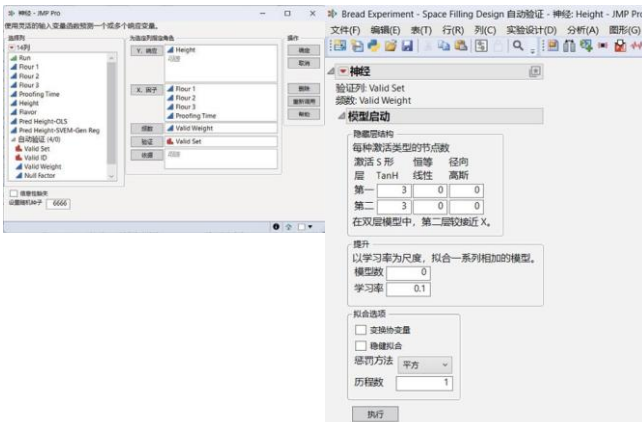


圖 47 拟合 SVEM 神经网络模型

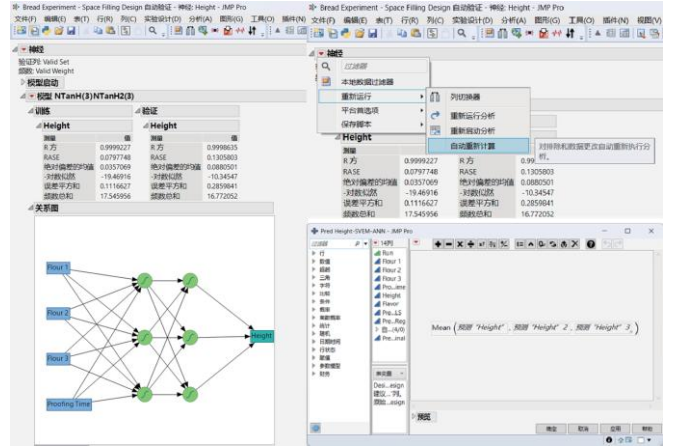


圖 48 多個 SVEM 神经网络模型集成

模型比較

保存標準最小平方、SVEM廣義回歸、SVEM神經網路模型的預測公式，並應用模型比較，可以發現本案例中，SVEM神經網路模型的擬合效果更優。



圖 49 模型比較

組合三種不同建模方法的預測公式形成刻畫器，可以發現相比於標準最小平方模型，SVEM廣義回歸模型和SVEM神經網路模型可以捕捉到更多非線性關係，進一步反映了機器學習建模方法結合空間填充設計的實驗資料，可以更精細地刻畫因子和反應變數之間的複雜尤其是非線性的關係，從而實現更高精度的預測。

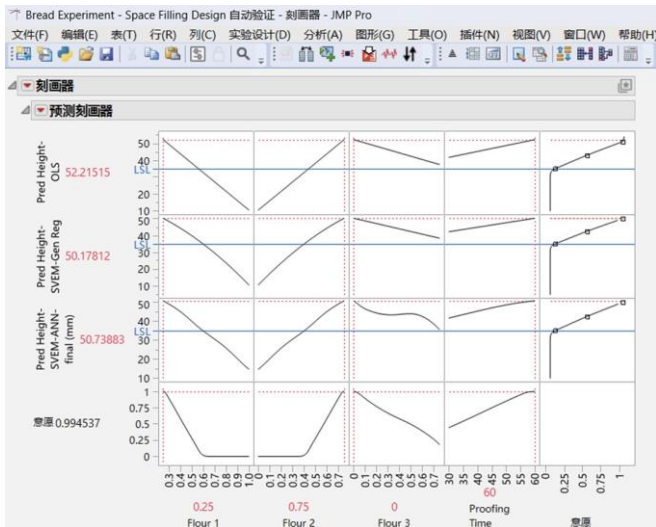


圖 50 預測刻畫器比較

透過三元圖結合三種模型的預測公式，可以更直觀地觀察到，SVEM神經網路模型能夠描述因子和反應變數之間的複雜非線性關係。

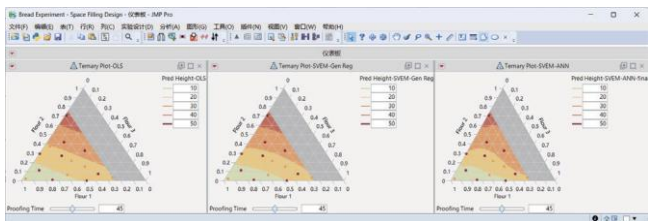


圖 51 模型預測公式比較

小結

本案例應用空間填充設計生成了18次的實驗方案，分別應用標準最小平方模型、SVEM廣義回歸模型和SVEM神經網路模型，建立了配方成分比例、醒發時間與麵包高度、評分之間的預測公式。在模型未知的情况下，創新性應用快速靈活填充設計這種現代實驗設計方法，並基於小樣本資料集進行訓練並驗證了模型的準確性，為實驗設計應用打開了新思路。綜合對比，對於本案例中採用空間填充設計獲得的實驗資料，結合SVEM神經網路建模方法更優，同時其模型的複雜性提高，可解釋性降低。但在模型未知的情况下，空間填充設計結合機器學習方法不失為一種創新性的實用解決方案。

JMP的空間填充設計平台可以支援更為複雜的實驗設計，無需模型效應項的先驗知識即可生成實驗方案，尤其適合電腦模擬實驗、早期實驗探索和複雜非線性關係建模研究，結合JMP預測建模平台的眾多機器學習方法，可以建立高精度的預測模型，實現反應變數預測和最佳化。





實驗設計方法選擇

實驗設計方法是科學研究、產品開發、製程最佳化等領域中不可或缺的一部分，其核心在於透過系統地安排實驗條件，以最小的成本、最短的時間獲取最準確、最有價值的資料和資訊。在選擇實驗設計方法時，需要考慮多個因素，如實驗目的、實驗規模、約束限制等。選擇適當的實驗設計方法對於確保實驗的成功、提高資料準確性、最佳化資源利用、提升決策品質、促進知識創新以及降低風險等方面都具有重要意義。因此，在進行實驗設計時，應綜合考慮各種因素，選擇最適合的實驗設計方法。

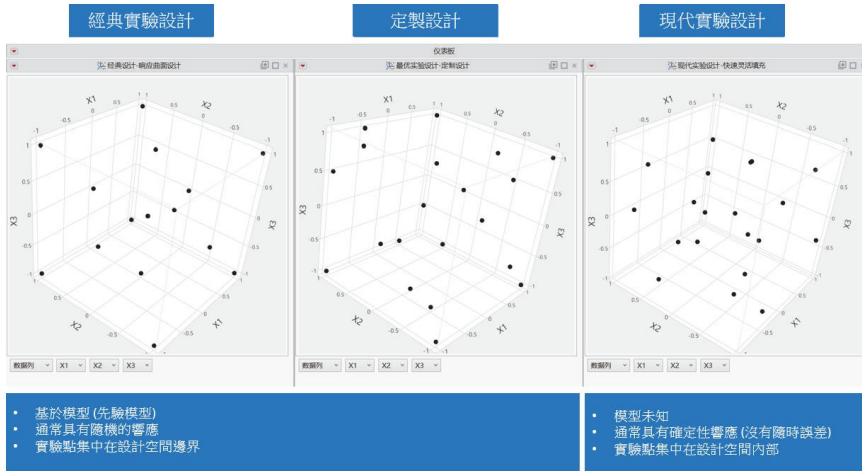


Made with Napkin

圖 52 實驗設計方法選擇

經典 vs 最優 vs 現代

根據實驗設計的發展歷程和設計思維，實驗設計可以分為經典實驗設計、最優實驗設計和現代實驗設計三類方法。對於案例實戰章節的三個案例，反應變數曲面設計屬於經典實驗設計，定制設計屬於最優實驗設計，空間填充設計屬於現代實驗設計。三類實驗設計方法的對比總結如下。



經典實驗設計	定製設計	現代實驗設計
廣泛應用於傳統真實實驗領域	相比於經典實驗設計，定製設計更適合於現實中較為複雜和具有挑戰性實驗設計方案的制定	廣泛應用於電腦仿真模擬實驗，響應通常沒有隨機性
完全析因設計/部分析因設計/響應曲面設計/混料設計/田口設計	支持多種因子類型（如連續、離散、分類、混料、協變量、區組等）	涉及更多的系統誤差，而不是傳統真實實驗中遇到的隨機誤差
多應用於低階多項式建模	考慮工程經驗和產業知識，更靈活地評估指定的效應項，探索有約束的設計空間	在設計空間內部進行採樣，減少近似模型中的偏差
旨在最大限度地減少傳統實驗中未知或不受控辯量引起的隨機誤差的影響，實驗點多傾向於分布在設計空間邊界及其附近	支持多種客製準則	空間填充設計具有以下特點： <ul style="list-style-type: none"> 普遍存在於電腦仿真模擬實驗領域（沒有隨機誤差） 實驗點多分布在設計空間內部 適合模型未知的實驗 適合構建高度複雜和非線性的機器學習模型 更適用於低維
可以進行序貫實驗，透過擴充設計支持更複雜模型的構建	根據實驗資源指定實驗次數	可以支持構建線性和複雜的高度非線性模型

圖 53 實驗設計方法對比總結

基於模型 vs 模型未知

根據實驗設計是否具有先驗知識，實驗設計方法可以分為兩種思維方式：基於模型的與模型未知^e。

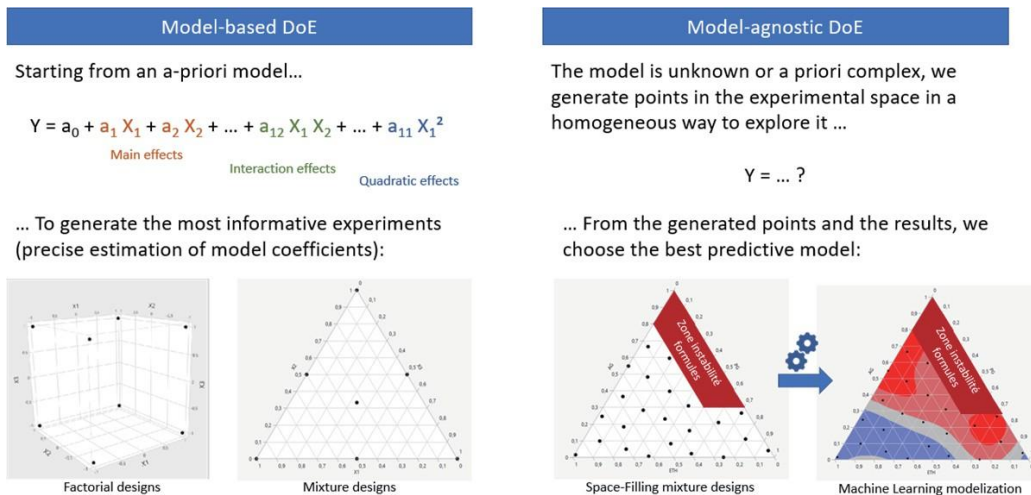


圖 54 基於模型 vs 模型未知^e

^e 聽說你們在玩一種很新的 DOE ?

^⑥ Solved: Space filling versus optimal DOE - JMP User Community

基於模型：這種類型的實驗設計先驗地假設一個特定的模型（具有主效應、交互作用項等），然後生成實驗點以最佳化模型係數的估計或最小化模型預測方差。這種方法假設了一種特定的模型形式（通常是多項式模型），當底層模型是眾所周知的，或者對所研究的系統有先驗知識時，這種方法很有用。它可能需要更少的實驗點，如果想輕鬆獲得可解釋的結果，基於模型的實驗設計是理想的選擇。（範例：反應Y受到因子A和B的統計顯著主效應以及A和B之間統計顯著的雙因子交互作用的嚴重影響）

模型未知：這種類型的實驗設計在實驗空間中均勻和隨機分佈實驗點，然後將不同的模型擬合到這些實驗點上，以獲得最佳的預測模型。這種方法可能需要大量的實驗點，具體取決於所構建模型的複雜性（線性迴歸、支援向量機、隨機森林、神經網路等），不假設任何特定的模型，在底層模型未知或複雜時非常有用。但是這種方法可能對雜訊很敏感，使用機器學習模型可能產生過擬合的情況（示例：實驗初期探索階段，對模型知之甚少；底層模型存在複雜非線性關係，簡單多項式模型難以精準描述）

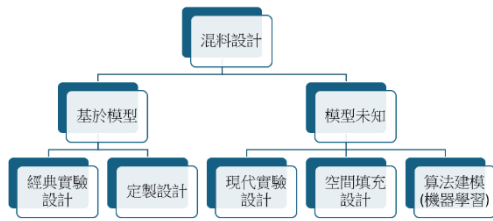


圖 55 多種實驗設計方法在混料設計中的體現

經典實驗設計和定製設計會假定某種先驗模型，是基於模型的實驗設計。而空間填充設計沒有先驗模型，在設計空間內部合理均勻布點採樣，是模型未知的實驗設計。

基於模型與模型未知，兩種實驗設計方法各有所長。基於模型的實驗設計可能更高效，需要更少的資料，但它們更容易受到模型設定錯誤 (Model Misspecification) 的影響，在面對複雜或未知的底層模型時可能不那麼穩健。模型未知的實驗設計可以更靈活、更穩健地應對模型設定錯誤，但它們可能需要更多的資料來實現與基於模型的實驗設計相當的精度。這兩種實驗設計思維方式之間的評估也可能不同，基於模型的實驗設計更強調統計顯著性（反應曲面設計和混料設計除外），模型未知的實驗設計試圖在設計空間中獲得最佳預測，而不是在模型中找到一個重要因子或效應項，更強調預測準確性 (RMSE和其他預測性能指標)。

基於模型的實驗設計強調在設計空間的邊緣採樣（降低隨機誤差的影響），通常具有較少的因子水準，更適合構建比較簡單的多項式模型；模型未知的實驗設計（空間填充設計）則更強調在設計空間內部合理均勻布點（通常沒有隨機誤差，降低偏差），具有更多的因子水準，適合構建複雜的精細模型^⑥。

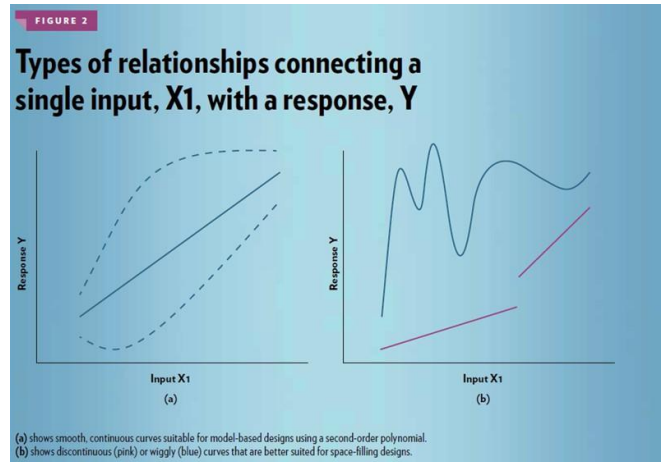
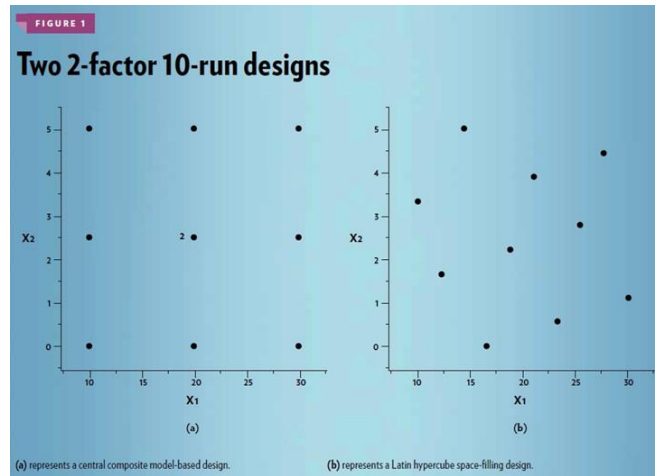


圖 56 基於模型 vs 模型未知

兩種實驗設計方法的選擇，沒有固定的範式，需要根據分析目標具體情況具體分析。瞭解每種方法的優點和局限性，靈活運用，方能應對紛繁複雜的分析難題。選擇基於模型還是模型未知，只是實驗設計旅程中的第一個關鍵岔路口，實驗設計方法的選擇需要綜合權衡眾多因素，才能最大限度確保實驗設計的成功。

常用實驗設計方法

常見的實驗設計方法包括全因子設計、反應變數曲面設計、混料設計、確定性篩選設計、定制設計等。每種方法都有其特點和適用範圍，應根據具體情況進行選擇。

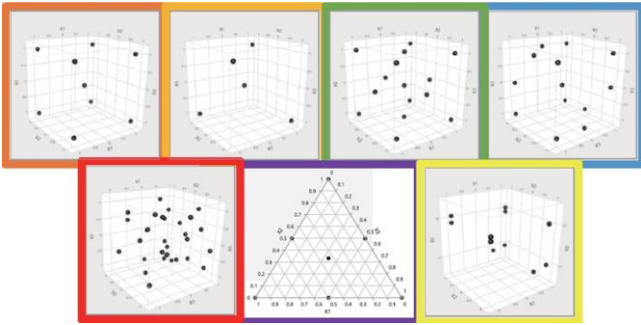


圖 57 常用實驗設計方法^h

Easy DOE：以實驗設計工作流形式，透過引導式介面完成定義、模型、設計、資料登錄、分析、預測、報表各步驟的操作，使用者友好

全因子設計：經典的實驗設計方法，描述階段使用，通常在因子數量較少（ 2^6 個）時用於描述和理解系統特性，可以估計主效應和交互效應，實驗次數相對適中（ $n=2^k$ ），但隨著因子個數的增加，實驗次數成幾何級增長

部分因子設計：經典的實驗設計方法，篩選階段使用，通常在因子數量較多（ 3^15 個）時用於識別關鍵因子，基於效應稀疏假設，存在效應混雜問題，需要權衡實驗次數（ $n=2^{(k-p)}$ ）和實驗設計的分辨率

反應曲面設計：經典的實驗設計方法，最佳化階段使用，通常在需要重點最佳化少量關鍵因子（ 2^3 個）時使用，實現對反應變數的高精度預測，可以估計主效應、交互效應和非線性效應，要求所有因子均為連續型變數，且實驗區域已接近最優區域

確定性篩選設計：JMP原創的實驗設計方法，篩選階段使用，適用於早期實驗工作，用於從大量潛在因子（ >4 個）中高效識別關鍵因子，僅需要少量試驗（ $n \approx 2k$ ），使用連續因子和兩水準分類因子，允許估計主效應、某些交互作用和二次項，可避免效應混雜，並且可以標識對反應變數造成非線性效應的因子

定製設計：JMP原創的實驗設計方法，構造為您特定的實驗場景量身定制的最優設計，可以定義多種因子類型、添加參數空間約束、結合經驗知識研究特定效應、根據實驗資源指定實驗次數等，靈活適應不同特定場景，適合在實驗任務稍複雜時使用，通常定制設計比使用其他方法獲得的設計更為經濟高效

混料設計：常用於配方研發，研究混合物中各組分相對比例對反應變數的影響，滿足混料約束，各組分的相對比例之和為 1

空間填充設計：常用於電腦模擬實驗或非線性複雜系統建模，實驗次數通常較多，適合與機器學習方法結合使用構建複雜模型

擴充設計：透過序貫實驗擴充已設計的實驗，向現有實驗設計添加實驗點，解決現有設計的不明確性

設計診斷ⁱ

生成多個實驗設計方案後，可以採用評估設計和比較設計平台，對實驗設計方案進行診斷，結合多個評價指標和報表綜合比較，選取最適合當前研究目標的實驗設計方案。

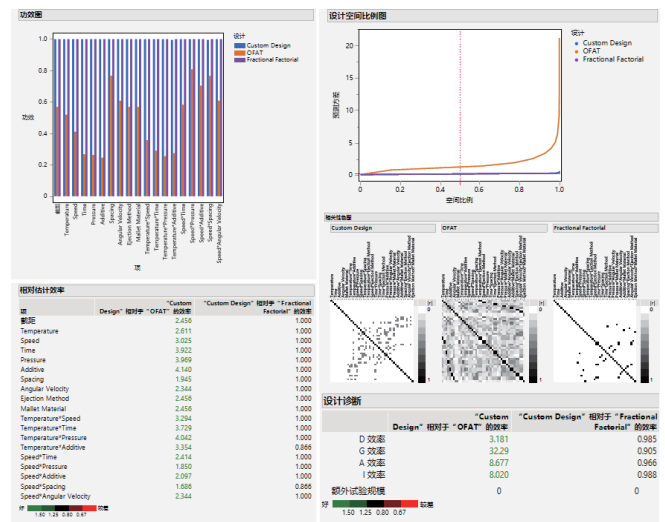


圖 58 評估與比較設計

^h Experimental Design Choices. Phil Kay. #DoEbyPhilKay.

ⁱ 可參考《DOE入門課：從方法到實踐，60分鐘輕鬆掌握DOE》公開視頻課程瞭解詳細內容。

JMP實驗設計的優勢

同時支援經典、最優和現代實驗設計方法，適用常見或複雜場景的產品研發與製程最佳化。

具備全面的統計建模和機器學習平台，支援實驗設計資料高效建模分析。

具備一系列交互、動態的刻畫器、模擬器和視覺化平台，輕鬆實現模型應用，加速產品研發與製程參數最佳化。

JMP的實驗設計平台秉承JMP「以解決問題為中心」的理念，力求根據實際問題構建實驗計畫，一舉顛覆傳統軟體需要修正實際問題以適應理論模型的「削足適履」的做法，並將實驗設計平台與專業的模擬功能(Simulation)完美結合，幫助用戶直擊問題的實質並控制新方案實施的失敗風險。JMP以完整的實驗設計方法集完全滿足客戶需要，同時以靈活多樣的定制設計功能幫助用戶高效地利用實驗資源，大幅節省實驗成本。

全面的實驗設計方法集

使用JMP，用戶不僅能實現因子設計、反應曲面設計、混料設計等傳統實驗設計方案，還能輕鬆實現目前業界其他同類軟體難以實現的定制設計、擴充設計、確定性篩選設計、非線性設計、空間填充設計等高級實驗設計方案，用戶對實驗設計方法的各種需求基本都能得到滿足。

獨有的高級實驗設計(Advanced DOE)

JMP強大的實驗設計平台獨具空間填充設計、非線性設計以及定制設計等高級實驗設計方案，讓用戶獲得解決實際問題更加高效的方法。

如果所研究的過程相對穩定，在一定的實驗預算內，空間填充設計能實現最佳的實驗覆蓋面，從而獲得信息量最豐富的實驗結果資料。

最佳化的傳統實驗設計

即使對於一般軟體都能實現的傳統實驗設計，JMP的實驗設計平台也不乏過人之處。JMP不僅讓使用者在既有的理論模型和分析結果的基礎上獲得更為強大的擴展功能，而且用更完整的實驗計畫制定方案幫助用戶提高效率，降低風險。

高度集成的專業模擬(Simulation)功能

JMP的實驗設計平台與專業模擬功能高度集成，幫助用戶預測最優方案的實施效果，減小失敗風險；同時使用者也不必另行購買任何類比軟體，既節省了成本，又降低了因為切換於不同軟體之間而產生錯誤的幾率。

靈活易用的實現方式

JMP的實驗設計平台充分考慮用戶的要求，完全按照分析和解決問題的思路引導用戶完成實驗方案，便於用戶學習以大幅降低使用門檻；運用JMP實驗設計平台豐富的定制功能，用戶可以完全按需定制個性化的實驗設計方案，甚至可以突破傳統軟體對實驗次數的限定，根據使用者實際的可用資源指定最佳實驗次數。

關於 JMP

自1989年第一版JMP軟體問世以來，JMP一直致力於透過互動式視覺化的統計分析，發現資料背後的價值，被廣泛應用於探索性資料分析（EDA）、資料探勘、建模預測、實驗設計、產品研發、生物統計、醫學統計、可靠度分析、資料視覺化、市場調研、品質管制及改善等領域，裨益半導體和電子、高科技製造、醫藥、臨床、公共衛生、基因研究、新能源、鋼鐵、化工、食品、快速消費品、金融和服務、政府和教育等各行各業。作為全球領先的資料分析方法及諮詢供應商，JMP致力於說明客戶從資料中獲取價值，最佳化決策，驅動創新，成就未來。



聯繫我們

JMP Taiwan

台北市民生東路三段10號12樓
信箱：jmpmarketing@jmp.com

關注「JMP資料分析」

