Application of Robust Design and Multi-Regression in Pharmaceutical Development

Dr. Erh-Rong Wu (Ph.D., P.E.)
erwu@excite.com

Abstract

This article presents a new computer software package for medicinal chemical development at experiment or design stage. Multi-regression method creates a pseudo analytical model equation from test data or simulation constituents. Robust Design is then used to perform probabilistic optimization of desired quantity and identify the best pharmaceutical components and processing parameters of a drug substance meeting their specifications. The software may be interfaced with other structure-based drug design software. An imitated example is given to illustrate the capability of the software package.

Introduction

The process of drug design and development is very complex. It involves with many variables. Those variables include chemical compounds, environmental conditions and manufacturing parameters. They are complicated with inherent random characteristics. Computer-aided design modeling and molecular structure based analytical techniques have been developed. Nevertheless, on the macroscopic level, the identification of optimal design components such as chemical compounds and processing factors, for example temperature, pressure and duration time, is still needed. The drug development process can be considered as a multivariate problem involving many random variables. Robust design is aimed at minimizing the deviation from a target, the design point which satisfies a required condition. A limit-state function is defined as the difference between the target and the experimental data or data from an available analytical model. When experimental data are available or simulation constituents are known, the statistical approach, the multi-regression method can be utilized to generate a pseudo analytical equation to render data for establishing the limit-state function. With the limit-state function defined, we proceed to carry out robust design procedures using the UNIPASS software package to obtain the optimal design variables and their sensitivities associated with the design outcome.

As the powerful combined software package can be extended to broad applications for product developments, such as medical devices, the general concepts of multi-regression method and robust design are explained. Then, an imitated example for producing an effective drug substance is demonstrated.

Construction of Limit-State Function

The limit-State function can be expressed as

$$g\left(\underline{X}_i\right) = R_{req} - R_{act}\left(\underline{X}_i\right) \qquad \textbf{(1)}$$

where R_{req} is the required value and R_{act} is the calculated or measured value and the index presents the i-th measurement.

When an analytical solution, e.g., a closed form formula or a finite-difference/finite element solver, is available, $R_{act}(\underline{X}_i)$ can be calculated from the analytical model. Otherwise, $R_{act}(\underline{X}_i)$ should be constructed from measured data through experiments. If the later is the case, the experimental data will be used and the multi-regression method will be utilized to establish the relationship between the function $R_{act}(\underline{X}_i)$ and the selected random variables $\underline{X}_i$. Experimental data can be exemplified by the following table:

Table 1 Measured Data from Experiments

Measured Data index j	$R_{act}(\underline{X}_j)$	X_{1j}	X_{2j}	X_{3j}	X_{4j}	X_{5j}	X_{6j}	X_{7j}
1	R_1	X_{11}	X_{21}	X_{31}	X_{41}	X_{51}	X_{61}	X_{71}
2	R_2	X_{12}	X_{22}	X_{32}	X_{42}	X_{52}	X_{62}	X_{72}
3	R_3	X_{13}	X_{23}	X_{33}	X_{43}	X_{53}	X_{63}	X_{73}
4	R_4	X_{14}	X_{24}	X_{34}	X_{44}	X_{54}	X_{64}	X_{74}
5	R_5	X_{15}	X_{25}	X_{35}	X_{45}	X_{55}	X_{65}	X_{75}
6	R_6	X_{16}	X_{26}	X_{36}	X_{46}	X_{56}	X_{66}	X_{76}
7	R_7	X_{17}	X_{27}	X_{37}	X_{47}	X_{57}	X_{67}	X_{77}
8	R_8	X_{18}	X_{28}	X_{38}	X_{48}	X_{58}	X_{68}	X_{78}
9	R_9	X_{19}	X_{29}	X_{39}	X_{49}	X_{59}	X_{69}	X_{79}

Table 1 shows, as a typical example, the relations between $R_{act}(\underline{X}_j)$ and $\underline{X}_j = \{X_{1j}, X_{2j}, \ldots, X_{7j}\}$ where j represents experimental data index.

Based on the principle of Least Square Sum of Deviation, from those data presented in Table 1, a linear multi-regression equation can be obtained as

$$R_{act}\left(\underline{X}_j\right) \approx A_0 + \sum_{i=1}^{N} A_i X_{ij} \qquad (2)$$

Where N is the number of variables. In general, i and j can be large numbers, Table 1 shows i=7 and j=9. In Eq. (2) A_0 and A_i are determined through linear algebraic computation using the data given in Table 1. Once Eq. (2) is established, the limit-state function, Eq. (1), can be defined.

Robust Design Process

Because the inherent randomness in the variables $\underline{X}_j$, $R_{act}(\underline{X}_j)$ is also a random quantity. A Design is robust when the product performance is minimum sensitive (or impact) to the uncertainties in the design variables. In other words, a Robust design has minimum standard deviation on the product performance, e.g., $R_{act}(\underline{X}_j)$. In the model presented by Eq. (1), The robust point should lie on the surface represented by the limit-state function and has the minimum norm of the gradient vector.

UNIPASS software package has the capability of robust design. The production process of any product can be depicted in the following figure:

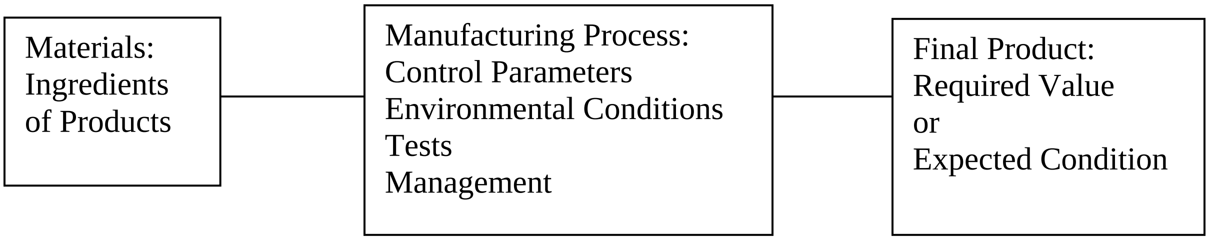

Figure 1 Production of A Product

Figure 1 shows the simplified process for producing a product. For the robust design process, the input consists of the random variables, $\underline{X}_{i,}$ selected from the first two boxes, i.e., uncertainties of material and variables involved in the manufacturing process. The outputs are the robust design point that gives the best design with the minimum variation in the product performance. The following figure depicts the sequence of the input, UNIPASS robust design module and the output.

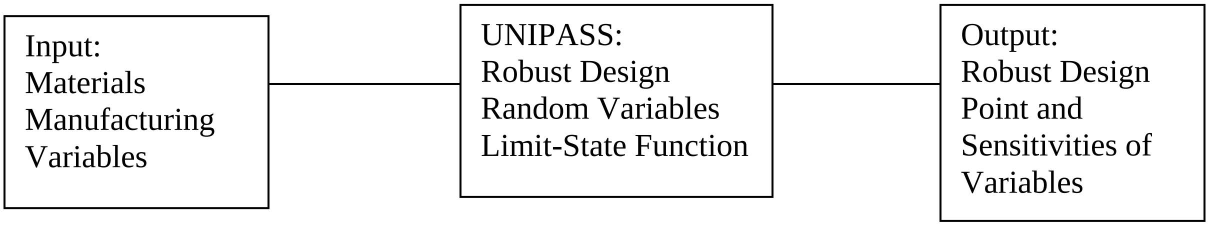

Figure 2 Typical Robust Design Process

Example

The process of producing a drug substance through medicinal chemical development can be typically represented by the following figure:

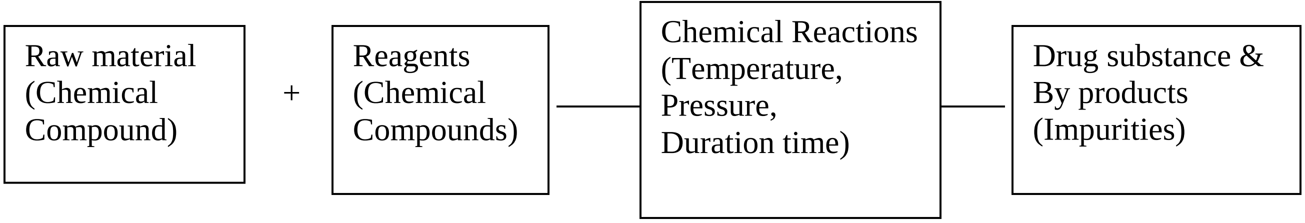

Figure 3 Typical Medicinal Chemical Developments

Imitated experimental data are given in Table 1. The variables are identified as:
G_{act} : drug substance (g) ,
X_{1j} : raw material (g),
X_{2j} : reagent A (g),
X_{3j} : reagent B (g),
X_{4j} : temperature (F),
X_{5j} : pressure (psi),

X_{6j} : duration time (hour),
X_{7j} : number of impurities.

Table 2 Dimensional Data of Medicinal Chemical Development

j	G_{act}	X_{1j}	X_{2j}	X_{3j}	X_{4j}	X_{5j}	X_{6j}	X_{7j}
1	128.7	256.7	35.8	45.8	189.5	24.8	3.5	4
2	168.5	358.9	26.8	35.6	195.8	35.6	4.6	5
3	156.4	268.5	31.5	40.3	210.4	48.5	5.4	3
4	178.4	482.7	37.9	48.2	225.6	41.9	3.2	3
5	155.8	319.5	29.6	41.8	199.4	22.6	2.8	4
6	163.8	278.4	33.7	29.6	178.5	32.9	3.9	3
7	142.8	234.9	25.8	42.6	188.4	38.4	4.3	5
8	133.8	305.3	26.8	35.9	180.5	39.5	3.7	4

Normalizing data of Table 2 with the mean values of each column, $G_{act,m}$, $X_{ij,m}$, i=1,6, where $G_{act,m} = 153.53$, $X_{1j,m} = 313.11$, $X_{2j,m} = 30.99$, $X_{3j,m} = 39.88$, $X_{4j,m} = 196.01$, $X_{5j,m} = 35.53$, $X_{6j,m} = 3.93$, $X_{7j,m} = 3.88$, Table 2 becomes:

Table 3 Dimensionless Data of Medicinal Chemical Development

j	G'_{act}	X'_{1j}	X'_{2j}	X'_{3j}	X'_{4j}	X'_{5j}	X'_{6j}	X'_{7j}
1	0.838	0.820	1.155	1.146	0.967	0.698	0.892	1.032
2	1.098	1.146	0.865	0.891	0.999	1.002	1.172	1.290
3	1.019	0.858	1.017	1.008	1.703	1.365	1.376	0.774
4	1.162	1.542	1.223	1.256	1.151	1.180	0.815	0.774
5	1.015	1.020	0.955	1.046	1.017	0.636	0.713	1.032
6	1.067	0.889	1.088	0.741	0.911	0.926	0.994	0.774
7	0.930	0.750	0.833	1.066	0.961	1.081	1.096	1.290
8	0.872	0.975	0.865	0.898	0.921	1.112	0.943	1.032

Employing the multi-regression software, MULTGRSN, with data given in Table 3, we obtain equation (2) as shown in the following

$$G'_{act} = -2.589 - 0.751 X'_{1j} + 1.317 X'_{2j} - 2.172 X'_{3j} + 4.485 X'_{4j} + 0.611 X'_{5j} - 1.035 X'_{6j} + 1.134 X'_{7j} \quad (6)$$

Supposedly, we want to produce 200 grams of the drug substance, then the limit-state function should be

$$\begin{aligned} G'(X'_{ij}) &= (200/153.53) + (2.589 + 0.751 X'_{1j} - 1.317 X'_{2j} + 2.172 X'_{3j} - 4.485 X'_{4j} - 0.611 X'_{5j} + 1.035 X'_{6j} - 1.134 X'_{7j}) \\ &= 3.892 + 0.751 X'_{1j} - 1.317 X'_{2j} + 2.172 X'_{3j} - 4.485 X'_{4j} - 0.611 X'_{5j} + 1.035 X'_{6j} - 1.134 X'_{7j} \quad (7) \end{aligned}$$

With this limit-state function and 7 random variables, we can use UNIPASS software to carry out robust design.

The final results are:

Required Drug Substance= 200 g
Raw Material = 279.89 g
Reagent A =35.76 g
Reagent B = 37.21 g
Temperature = 180.23 F
Pressure = 42.72 psi
Duration Time = 3.04 hr.
Impurity = 4

Conclusion

We illustrated the power of the software package combining robust design methodology and the multi-regression technique for producing a drug substance. Since the pseudo-analytical equation created by the multi-regression method can be extended to any product in any industry, the whole design process can be applied to execute robust design for any product existed (for improvement), or being developed, or futuristic in nature (to be developed).

The number of random variables theoretically can be a large number to cover many realistic affecting factors associated with a product development and manufacturing process.

The final design parameters not only lead to a product meeting requirements but also optimize the qualities of the product. Robust design minimizes the deviation of the target function from its expected value and results in high stability.

The technique presented here will realistically cut the product development time and cost.