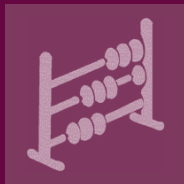


Area

metodi quantitativi



MARIA PIA D'AMBROSIO, FRANCO ANZANI, SERGIO VENTURINI

DFSS E SIMULAZIONE MONTE CARLO: PROBLEMATICHE DI IMPLEMENTAZIONE NELLA TOLERANCE ANALYSIS

La prassi abitualmente seguita di simulare poche migliaia di valori e da questi estrapolare conclusioni circa l'ottimalità del processo (in termini di difettosità dello stesso), consente di ottenere solo risposte approssimate e parziali. Per arrivare a risultati corretti è necessario disporre di strumenti che possano gestire con facilità simulazioni di milioni di dati

DFSS E SIMULAZIONE MONTE CARLO: PROBLEMATICHE DI IMPLEMENTAZIONE NELLA TOLERANCE ANALYSIS



La prassi abitualmente seguita di simulare poche migliaia di valori e da questi estrapolare conclusioni circa l'ottimalità del processo (in termini di difettosità dello stesso), consenta di ottenere solo risposte approssimate e parziali. Per arrivare a risultati corretti è necessario disporre di strumenti che possano gestire con facilità simulazioni di milioni di dati

MARIA PIA D'AMBROSIO
SixSigmaIn.it

FRANCO ANZANI
SixSigmaIn.it

SERGIO VENTURINI
SDA Bocconi

Come implementare con successo la tecnica di simulazione Monte Carlo alla metodologia di analisi nota con il nome di Design for Six Sigma (DFSS)?

Il DFSS è una metodologia per la progettazione di nuovi processi/prodotti, o per il ridisegno di processi/prodotti esistenti, il cui obiettivo è il raggiungimento di requisiti di performance, affidabilità e costi propri delle metodologie Six Sigma. Nel presente contributo, al fine di rendere la trattazione meno tecnica, si farà riferimento a un esempio specifico riguardante un processo di stampaggio di un componente in plastica e ad una sola delle fasi del DFSS, nota come Tolerance Analysis (TA).

La TA studia la variabilità dell'output di un processo in funzione degli input del processo stesso. Il fine ultimo della TA è identificare quali input abbiano l'impatto maggiore sull'output. Lo strumento principale per portare a termine una TA è la *funzione di trasferimento*, ovvero la rela-

zione che lega tra loro input e output del processo.

Per condurre una TA sono disponibili diverse tecniche quantitative, tra cui le più comuni sono:

- *worst case*, in cui gli input del processo vengono fissati ai loro limiti estremi conosciuti; questo approccio ha il vantaggio di essere molto semplice e veloce, ma spesso risulta molto conservativo;
- *Root Sum of Squares (RSS)*, in cui si assume che la funzione di trasferimento sia nota e che gli input del processo siano tra loro indipendenti; nonostante sia un metodo semplice, il suo limite maggiore è che prevede che gli input siano distribuiti secondo una distribuzione normale, ipotesi non sempre soddisfatta nei casi concreti;
- *propagazione degli errori (POE)*, che si applica anche a funzioni di trasferimento non lineari, ma che presenta in sostanza gli stessi limiti della RSS;
- *simulazione Monte Carlo*, in cui gli input si assumono tra loro indipendenti e ognuno distribuito secondo una certa distribuzione di probabilità,¹ il cui limite maggiore in alcune situazioni è rappresentato dai tempi di elaborazione.

Tra queste tecniche, la simulazione Monte Carlo è certamente la più efficiente. Nata durante la seconda guerra mondiale, il suo obiettivo è la soluzione di problemi complessi, sia di natura deterministica sia di natura intrinsecamente aleatoria. Gli ingredienti richiesti sono l'assegnazione di una distribuzione di probabilità a ognuna delle variabili in input del problema e la specificazione di una funzione di trasferi-

1. Rubinstein R.Y., *Simulation and the Monte Carlo Method*, Wiley, New York.
2. Ragsdale C.T., *Spreadsheet Modeling & Decision Analysis*, Thompson.
3. Montgomery D.C., *Design & Analysis of Experiments, 5th Edition*, Wiley, New York.

mento. Campionando in maniera casuale (usando opportuni algoritmi) valori dalle distribuzioni in input, attraverso la funzione trasferimento si ottiene una corrispondente distribuzione per l'output. Il vantaggio principale della simulazione Monte Carlo consiste nel poter valutare la precisione della stima come funzione del numero di valori simulati.

Anche se più efficiente di altre tecniche, la simulazione Monte Carlo non sempre è impiegata in pratica. Ciò è dovuto alla relativa semplicità di implementazione delle altre tecniche e al fatto che per condurre una simulazione Monte Carlo sono necessari strumenti software specifici. I software statistici tradizionali per il Six Sigma, infatti, non comprendono al loro interno nessuna capacità esplicita, ovvero immediatamente utilizzabile dall'utente, di simulazione di modelli. Per effettuare una simulazione Monte Carlo sono oggi disponibili sul mercato *add-in* di Microsoft Excel, le quali hanno il vantaggio di aggiungere alla completa modellazione propria di Excel² quelle funzionalità di simulazione Monte Carlo che Excel non possiede, o possiede in modo molto limitato. Tali software sono nati in origine per sviluppare applicazioni nell'ambito della Risk Analysis e solo recentemente sono stati utilizzati anche per applicazioni tecniche nell'ambito del DFSS. Questi strumenti soffrono tuttavia di alcuni limiti, riconducibili alla loro genesi:

- ▶ mancanza di parser matematico, ovvero tutti i calcoli vengono direttamente eseguiti da Excel, il che rende

- la simulazione spesso molto lenta;
- ▶ limitazioni dei runs possibili, dovute alla modalità di gestione della memoria e alla limitata velocità di calcolo proprie di Excel;
- ▶ necessità di reinterpretare i risultati della variabile risposta con riferimento a una data distribuzione;
- ▶ perdita del contenuto informativo dei valori degli input generati durante la simulazione.

Tali limitazioni divengono critiche nel momento in cui si passa all'interpretazione dei risultati: il rischio che si corre nella mera applicazione delle tecniche di simulazione Monte Carlo al DFSS con gli strumenti software oggi maggiormente impiegati è quello di arrivare a conclusioni parziali e a volte fuorvianti del problema di ottimizzazione del processo in esame. Per illustrare questo punto utilizzeremo una simulazione alternativa dell'esempio Simulation with Design of Experiments disponibile sul sito <http://www.decisioneering.com>. Il caso riguarda un processo di stampaggio a iniezione in cui sono coinvolti tre input, o fattori (Mold Temperature, Cycle Time e Hold Pressure), che congiuntamente contribuiscono a determinare la variabilità dell'output (la variabile Length, ovvero la lunghezza del manufatto stampato). La funzione di trasferimento utilizzata nell'esempio originale è stata generata usando un modello di regressione multipla i cui coefficienti sono la soluzione di un piano fattoriale³ con due livelli per ognuno dei tre fattori e cinque replicazioni (figura 1). Nell'esempio si assume che i tre fattori in input siano distribuiti secondo una distribuzione normale (figura 2). Lo scopo dell'applicazione della simulazione Monte Carlo è mostrare come sia possibile ottimizzare la variabilità della risposta in funzione dei fattori controllabili e di conseguenza ridurre il numero di difetti ottenibili durante il processo di stampaggio.

Figura 1

Term	Equation Terms	Coef
Const	Constant	10,1625
A	HoldPres	0,27625
B	CycleTime	0,176875
C	MoldTemp	0,08275
D	MS Residual Error	1
A*B	HoldPres*CycleTime	-0,0010625
B*C	CycleTime*MoldTemp	0,0004125
A*B*C	HoldPres*CycleTime*MoldTemp	-0,0000025

Variable Setup		Response	A	B	C	D
			1 Normal	1 Normal	1 Normal	1 Normal
			2 Uniform	2 Uniform	2 Uniform	2 Uniform
			3 Weibull	3 Weibull	3 Weibull	3 Weibull
			4 LnNormal	4 LnNormal	4 LnNormal	4 LnNormal
			5 Exponential	5 Exponential	5 Exponential	5 Exponential
Distribution						
Name		Length	HoldPres	CycleTime	MoldTemp	Residual Error
1* Par (Nominal Mean)			130,0000	100,0000	150,0000	0,0000
2* Par (Nominal Std Dev)			6,0000	5,0000	7,0000	0,2121
3* Par (Nominal Value)						
Lower Spec		60,6750	112,0000	85,0000	129,0000	-0,6364
Upper Spec		66,6750	148,0000	115,0000	171,0000	0,6364

Figura 2

Figura 3

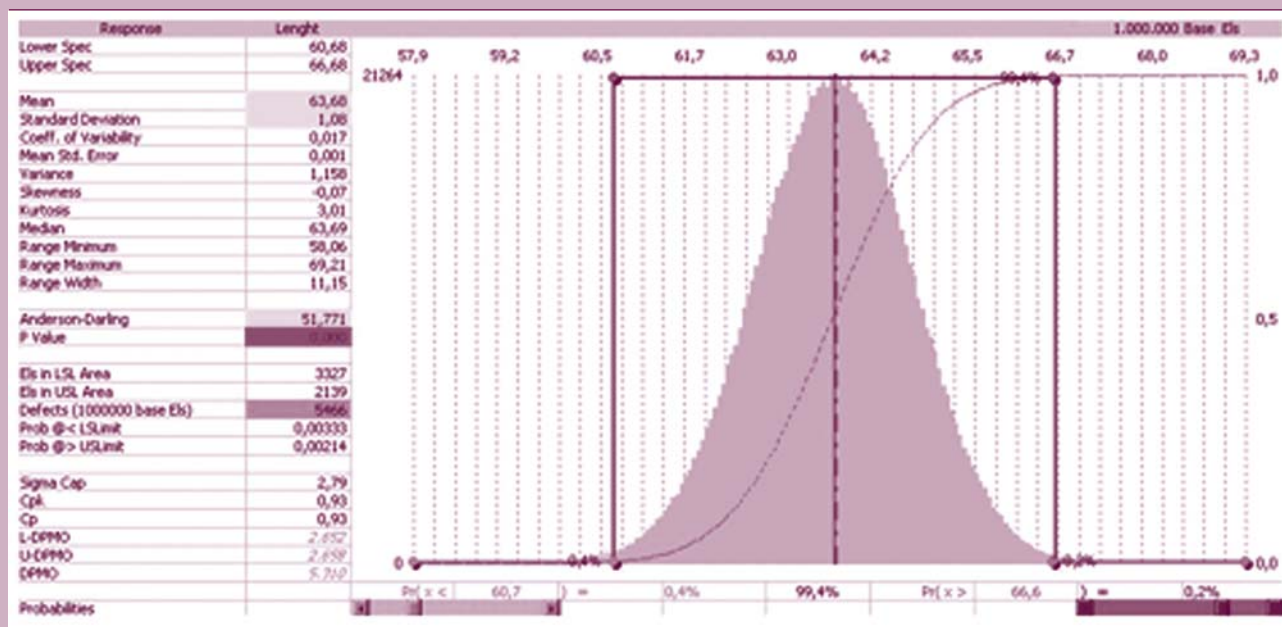
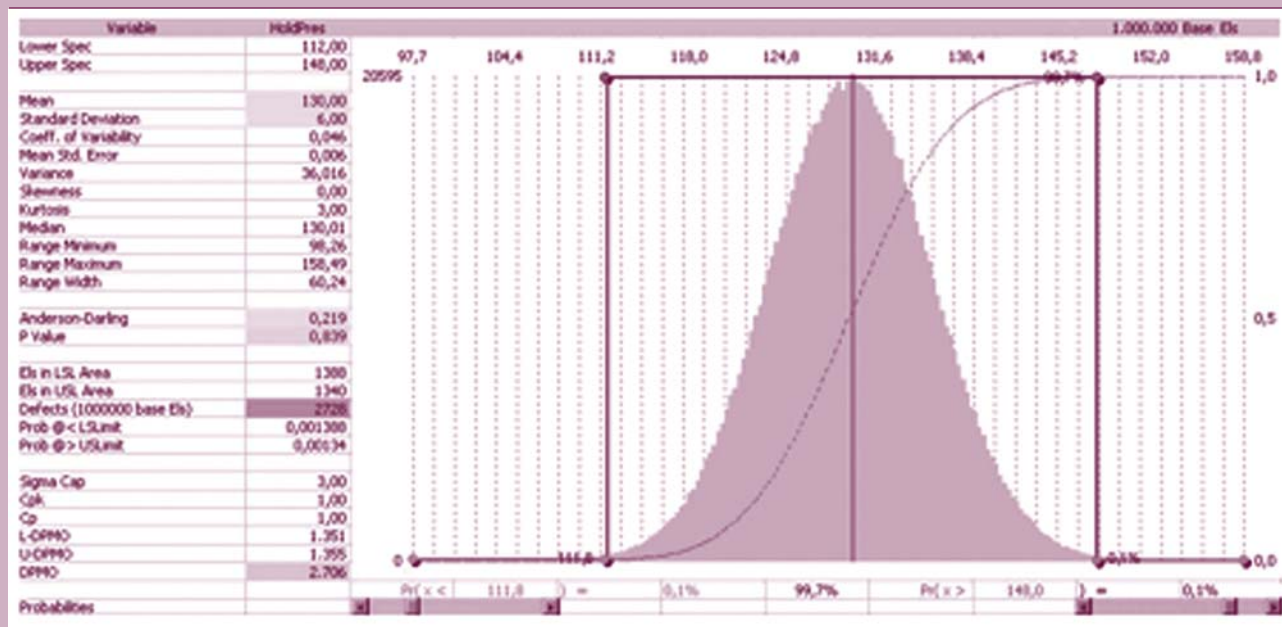


Figura 4

A tale fine è stata effettuata la simulazione di un milione di valori per gli input, i quali, attraverso la funzione di trasferimento, diventano un milione di valori

simulati per l'output.⁴ In figura 3 si riporta la distribuzione dei valori simulati per uno dei fattori del modello (A = Hold Pressure), mentre in figura 4 è riprodotta

ta la medesima distribuzione ma per la variabile di risposta (Length).

Dalle precedenti figure si può notare che, a differenza dei fattori in input, per i quali il test di normalità di Anderson-Darling⁵ deve essere superato come presupposto, in quanto si ipotizza che ognuna di tali variabili segua una distribuzione normale, lo stesso test di normalità non è invece superato dalla variabile risposta. Questo è ovviamente dovuto al fatto che, a causa delle interazioni tra i fattori presenti nel modello, la funzione di trasferimento non è una pura funzione lineare dei fattori stessi. In particolare è possibile dimostrare che, in conseguenza della marginalità dei coefficienti relativi alle interazioni tra i fattori, simulazioni a basso numero di runs (<2000) possono generare una risposta la cui distribuzione è interpretabile come una normale, ma un numero superiore di runs genererà sempre una distribuzione in cui il test di normalità non viene superato. Tale conclusione diviene tanto più evidente quanto più elevato è il numero di valori simulati. Ne consegue che la pratica abituale di estrapolare i risultati ottenuti su una simulazione a basso numero di runs all'intera popolazione, assumendo che la risposta sia distribuita in modo normale, risulta essere molto rischiosa. Ciò è particolarmente importante qualora vi sia la necessità di avere una stima precisa per una certa quantità, quale il numero di pezzi difettosi ottenuti nella simulazione (ovvero di unità con valori della risposta esterni ai limiti di specifica). Nel nostro esempio, infatti, il numero di unità con risposta esterna ai limiti di specifica risulta essere pari a 5466 (rispettivamente 3327 minori del limite di specifica inferiore e 2139 maggiori del limite superiore), mentre assumendo erroneamente che la risposta sia normale otterremmo un numero di pezzi difettosi pari a 5311 (rispettivamente 2652 e 2659). Questa differenza di difettosità nelle due code della distribuzione della

risposta non è marginale. Nell'esempio illustrato, infatti, una difettosità nella lunghezza del manufatto oltre il limite superiore di specifica è ancora tecnicamente superabile sottoponendo il manufatto stesso a una riduzione meccanica della lunghezza eccedente. Dall'altro lato, però, una difettosità nella lunghezza inferiore al limite inferiore di specifica comporta probabilmente la sola possibilità di riciclo del materiale estruso (qualora il polimero utilizzato lo permetta). Poiché serve stimare il numero di difettosità del processo, sempre espresso in parti per milione (PPM), risulta dunque rischioso simulare solo qualche migliaio di valori ed estrapolare i risultati ottenuti al milione richiesto dalla metodologia Six Sigma. Chiaramente, simulare milioni di dati comporta avere a disposizione strumenti potenti e altamente efficienti, ma il vantaggio in termini di precisione delle conclusioni che se ne possono trarre è decisamente elevato.

Un secondo aspetto di notevole importanza su cui si è posta l'attenzione riguarda la possibilità di tenere correttamente traccia dei valori simulati per gli input. I software per la simulazione Monte Carlo ad oggi maggiormente impiegati e citati precedentemente non consentono tale memorizzazione. Ciò non rappresenta affatto un'informazione trascurabile, poiché da quei valori è possibile risalire a quali configurazioni (simulazioni) hanno originato le diverse unità difettose risultanti e trarne interessanti considerazioni in termini sia di analisi sia di ottimizzazione. Nel nostro esempio è possibile infatti calcolare (figura 5) che la maggior parte delle unità difettose ottenute nella simulazione sono generate da valori dei fattori input che soddisfano i loro limiti di specifica (nella figura un valore simulato che si trova entro i limiti di specifica è indicato da un segno di uguale). Quindi l'obiettivo di bassa difettosità risulta difficilmente raggiungibile se si continuano ad accettare come vali-

4. I calcoli di questo esempio sono stati effettuati con un software per Microsoft Excel sviluppato dai primi due autori del presente contributo.

5. D'Agostino R.B., Stephens M.A., Goodness of Fit Techniques, Marcel Dekker.



di gli attuali limiti di specifica per gli input del processo di stampaggio. Il problema si traduce allora in come scegliere i nuovi limiti di specifica, ovvero la variabilità degli input, al fine di ottenere un numero prestabilito di unità difettose. La nostra proposta consente di dare una risposta definitiva anche a questo problema. Infatti, ripetendo l'intera simulazione un certo numero di volte (nel nostro caso sono esemplificate sei ripetizioni), è possibile stimare la relazione che lega la variabilità (le deviazioni standard) degli input a quella della risposta. In questo modo, una volta fissata la variabilità ritenuta accettabile per la risposta (che si traduce in un certo numero di pezzi difettosi), è possibile risolvere tale relazione in senso inverso, ottenendo così i valori delle deviazioni standard dei fattori in input che consentono di raggiungere la precisione desiderata per la risposta (figura 6).

Per arrivare a un'efficace implementazione della simulazione Monte Carlo nella fase di Tolerance Analysis del DFFS occorre dunque rilevare come la prassi abitualmente seguita – dovuta in gran parte alle capacità dei software oggi maggiormente diffusi per l'implementazione della simulazione Monte Carlo – di simulare poche migliaia di valori, e da questi estrapolare conclusioni circa l'ottimalità del processo (in termini di difettosità dello stesso), consenta di ottenere solo risposte approssimate e parziali. Per arrivare a risultati e a conclusioni corrette è necessario disporre di strumenti che possano gestire con facilità simulazioni di milioni di dati. Infine, risulta fondamentale tenere traccia di tutti i valori simulati per tutte le variabili coinvolte nel modello, al fine di non perdere informazioni rilevanti per ottimizzare il processo in esame.

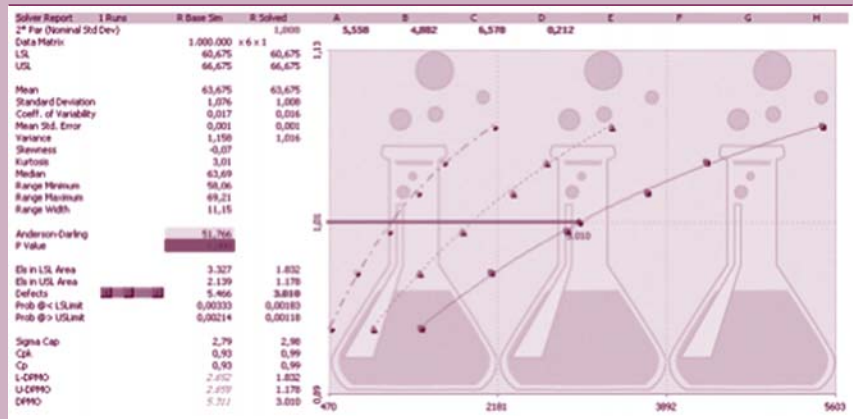
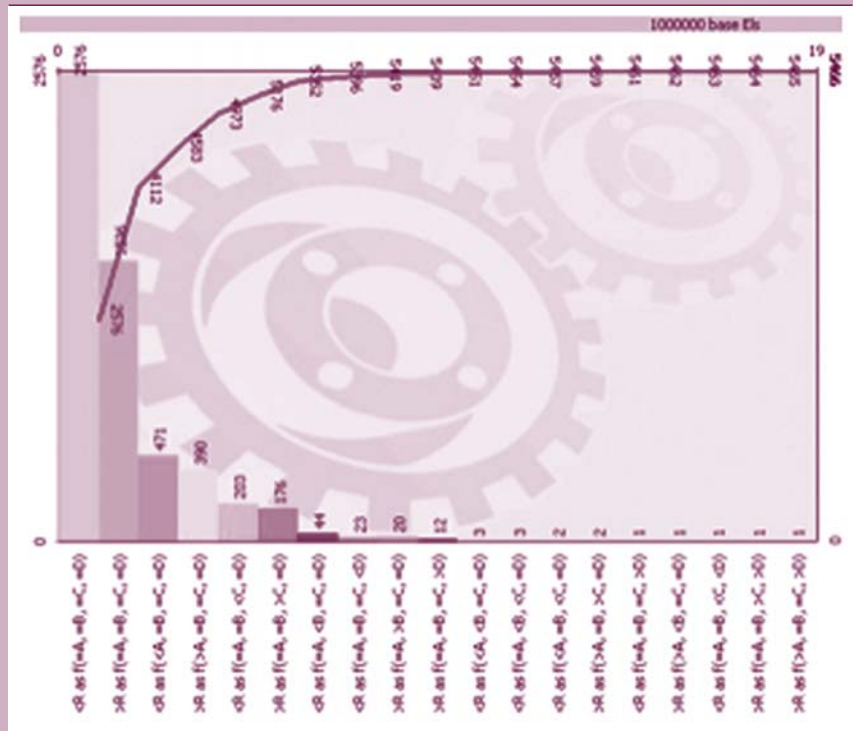


Figura 5 e 6