Journal of Manufacturing Systems 45 (2017) 195–200

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Metamodeling technique for designing reengineered processes by historical data

Francesco Gagliardi*, Giuseppina Ambrogio, Claudio Ciancio, Luigi Filice

University of Calabria, Department of Mechanical, Energy and Management Engineering, Italy

ARTICLE INFO

Article history:
Received 8 May 2017
Received in revised form 3 August 2017
Accepted 18 September 2017

Keywords:
Optimization
Manufacturing
Metamodeling technique
Response surface
Dynamic system
Extrusion

ABSTRACT

Physical-driven or simulation-driven experiments and data mining algorithms are combined for the identification of input–output relationships in complex manufacturing processes. To overcome time and cost consuming procedures, mathematical models have been applied finding and evaluating the factors that mostly affect the examined responses. In this context, a novel metamodeling technique was developed. This is able to use historical information on similar problems minimizing the amount of data necessary to the design of reengineered processes. The procedure was validated by applying it to the porthole extrusion optimizing the die geometries used for processing profiles characterized by various cross sections.

1. Introduction

Nowadays, the process and product improvements in advanced manufacturing companies are attained by applying analytical tools, such as Statistical Process Control (SPC) [1,2] and Design of Experiments (DoE) [3,4]. This action mode can be however costly and time consuming particularly if experimental data have to be collected. Finite Element Analysis (FEA) permits arbitrary combinations of input variables including design parameters and process conditions to be investigated limiting the number of experimental data [5]. Anyway, solving of 3D numerical simulations usually requires high computational time, too [6]. Optimization models, combined with the highlighted tools, allow one to define relationship between process parameters and outputs reducing investigation time and efforts [7]. In general, an optimization problem consists of three components, i.e. the objective function to be maximized or minimized, the decision variables to be selected and the constraints to be fulfilled [8]. Various optimization techniques have been developed to form processes characterized by strong non-linearity [9]. The construction of an approximation model, named metamodel, has been a solution widely utilized to study processes such as extrusion, rolling, forging, etc. Optimization problems in forming concern: (a) the design of the tool shape, (b) the process technology and (c) the final product quality. The main two step to analyze a process using statistical or artificial intelligence techniques are given by the definition of a design of experiment strategy (DoE) and by the choice of the algorithm used to analyze the results of the experiments. The next two tables analyze the main features of the most used DoE and regression approaches highlighting advantages and drawbacks.

Design of experiments is an approach used in numerous industries for conducting experiments to develop and/or improve products and processes. As reported in Table 1 several techniques have been tested by other researchers to perform analysis about the extrusion process. The full factorial and the CCD are usually two reliable designs to perform accurate analysis that allow to analyze the interactions among different process parameters. However they both require a large number of data. Moreover these approaches are not customizable in order to collect more data around the most interest area of the process domain or to analyze more deeply the influence of a particular input or interaction. The random sampling offers a good flexibility regards the number of total data but on the other hand, typically results in many experiments being made in the wrong direction and precious resources being wasted. In this work we proposed a 2-phases dynamic DoE which is able to interact with the metamodel in order to collect mode data in the most critical area and to minimize the number of experiments to obtain a certain desired accuracy. This approach can be integrated with several metamodel techinques. Also in this case, as presented in Table 2, the literature shows how different methodologies can be used for the same purpose. Neural

* Corresponding author.
E-mail addresses: francesco.gagliardi@unical.it (F. Gagliardi), giuseppina.ambrogio@unical.it (G. Ambrogio), claudio.ciancio@unical.it (C. Ciancio), luigi.filice@unical.it (L. Filice).

http://dx.doi.org/10.1016/j.jmsy.2017.09.004
0278-6125/

Table 1
Design of experiments (state of the art).

DoE	Sample size	Space filling	Flexibility	References
Full factorial	High	High	Low	[10]
Taguchi	Low	Average	Average	[11]
Latin hypercube	Customizable	High	Average	[12]
CCD	Very high	High	Low	[13]
Random sampling	Customizable	Not optimized	Average	
Proposed methodology	Customizable	High and optimized	High	

Table 2
Regression techniques (state of the art).

Technique	Non linear abilities	Interaction evaluation	Overfitting risk	Sample size required	References
Neural networks	Very high	High	Average	High	[14]
Linear regression	Null	Null	Very low	Low	[15]
Polynomial model	High	High	Average	Average	[13]
Kriging	Very high	High	Average	High	[16]
Fuzzy logic	Average	Low	Average	Average	[17]
Proposed methodology	Very high	High	Average	Low	

Networks and Kriging are prediction tools that can be used to map highly not linear relationships. However, both the techniques usually require a large number of experiments to reach a satisfactory accuracy. On the other hand, linear regression and fuzzy logic can be used to manage different parameters with a smaller number of data. Nevertheless both the techniques have several disadvantages that discourage their usage in this domain. Linear regression, in fact, cannot map accurately the complex relationship that regulates the process, while fuzzy logic make usually use of linguistic variables based on subjective opinions. In general, all the techniques previously presented can provide some insights into different process phenomena and interactions between process parameters. However there are all characterized by some drawbacks (flexibility, number of required data, accuracy) that limits their potential. In this work, a new methodology was proposed. In detail, RS models were integrated with new developed functions, named similarity functions (SFs), in a system able to achieve fast identifications of the factors, which mostly affect the investigated responses. This system is dynamic because, learning from the acquired data, changes adjusting itself during the investigation. The methodology, therefore, can be especially useful in designing of reengineered processes. This is important considering the current market demand with continuous product variations. The developed methodology can be applied to any forming processes. Its potentiality was tested and verified optimizing a porthole die geometry [18] used to extrude profiles with various cross sections. This bulk forming process was chosen considering the complexity of the die with a high number of input variables to be analysed and different outputs to be evaluated for a proper process execution [13].

2. The proposed metamodeling technique

The technique aims to define a relationship between a set of input parameters and a set of output variables using historical data, which are collected on variants of the investigated process. Explaining this concept mathematically, a model $f_j^{k'}$ of a new process variant k', function of n input parameters, for the jth output:

$$y'_j = f_j^{k'}(x_1, x_2, \ldots, x_n) \quad (1)$$

is assessed by means of the q data used to define the model f_j^k before its reengineering phase for the same output:

$$y_j = f_j^k(x_1, x_2, \ldots, x_n) \quad (2)$$

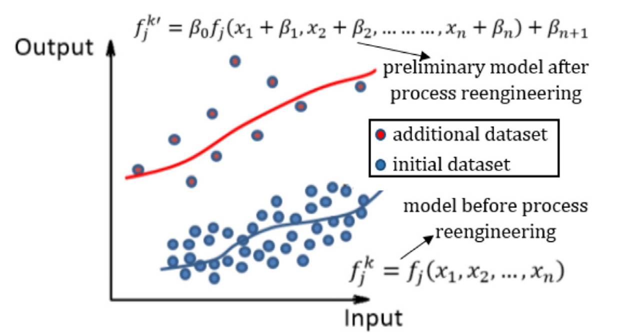

Fig. 1. Model before reengineering (f_j^k) and its preliminary surrogate form ($f_j^{k'}$) for the reengineered process variant.

and additional r data with $r < q$. The number of these additional data depends on the consistency of the available historical data. Furthermore, the detected model is composed of SFs, which have to be as simple as possible to limit overfitting risks. The highlighted methodology is performed through the below detailed steps:

- first sampling strategy
- preliminary analysis for surrogate model definition
- surrogate model definition
- second sampling strategy
- stopping criteria.

2.1. *First sampling strategy*

A small dataset, selected by a latin hypercube design [19], is collected like additional data for optimizing a reengineered process. This is done to capture first indications regarding the interactions between input parameters and output responses useful in the following Sections 2.2 and 2.3. Moreover, these first data are used to estimate the feasible process window for the new process variant if one or more constraints have to be considered.

2.2. *Preliminary analysis for surrogate model definition*

The small dataset, collected in Section 2.1, is used to determine the optimal value of a vector β

$$\beta = \{\beta_0, \beta_1, \ldots, \beta_{n+1}\} \quad (3)$$

which allows the output error, E_0, of the new model, $f_j^{k'}$, to be minimized. This new model is built as reported in Fig. 1.