Experimental Techniques (2018) 42:319–331
https://doi.org/10.1007/s40799-017-0230-1

Use of Factorial Designs and the Response Surface Methodology to Optimize a Heat Staking Process

Antonio Faria Neto [1,2] · Antonio Fernando Branco Costa [1] · Michel Floriano de Lima [2]

Received: 14 February 2017 / Accepted: 20 December 2017 / Published online: 3 January 2018

Abstract

The demand from the automotive industry for lighter and more resistant structures produced at lower costs has shifted the development focus of production processes toward hybrid components. A problem that arises from hybrid components is the necessity to join dissimilar materials, e.g., polymers and metals. A method to achieve this joining involves a process known as heat staking, in which a metal insert is heated and pushed against a thermoplastic surface. At the end of this process, the metal component may not be level with the thermoplastic surface; rather, it may be over flushed, and this discrepancy is known as the Insertion Height. This paper aims to apply the design of experiments and the response surface methodology to develop a model for the Insertion Height, considering the Heating Temperature and the Insertion Time as independent variables. The experiments revealed that the Insertion Height is most affected by the Heating Temperature. There are several combinations of the factors that can keep the Insertion Height within the specifications; therefore, it is possible to increase productivity by decreasing the Insertion Time and to save energy by reducing the Heating Temperature while considering the process constraints and specifications.

Keywords Design of experiments · Factorial designs · Response surface · Heat staking · Plastic welding

Introduction

A problem that arises from the production of hybrid components is the necessity to join dissimilar materials, e.g., polymers and metals [1, 2]. In certain cases, the solution involves the use of threads. However, most thermo-plastics are too soft to be properly held by threads [3, 4]. To address this issue, brass or steel threaded inserts can be added by means of a heat induction process. The metallic threaded insert is heated by induction to a temperature higher than the melting temperature of the thermoplastic, and it is subsequently pushed for a short period by means of a mechanical device against a properly prepared cavity in the thermoplastic material. Such a process, known as heat staking, has been widely used in the automotive industry.

At the end of the insertion process, the metal insert may not be level with the thermoplastic surface; in other words, the insert may be over flushed. The difference between the top of the insert and the thermoplastic surface is known as the Insertion (or Installation) Height. Figure 1 illustrates a heat staking process and the Insertion (Installation) Height (*h*).

The Insertion Height is an important design parameter from the perspective of the overall product quality. The over flushed condition decreases the tensile strength, creating unacceptable effects on the final product [5, 6].

According to [7, 8], the Insertion Height is related to the Heating Temperature of the insert, the Insertion Time, and the Insertion Load. These factors are directly linked to the material of the insert, the amount of polymer pushed during the insertion process, and the melting point of the polymer [9–11]. However, the equipment used in this research does not have an adjustment valve; as a result, the operator cannot change the value of the Insertion Load, which is kept constant at 5 bar by a lever system driven by a pneumatic cylinder. The lever pulls the insert as far as the end of the cavity. When the insert reaches the bottom of the cavity, a mechanical device stops the insertion process.

✉ Antonio Faria Neto
antfarianeto@gmail.com

[1] São Paulo State University (UNESP), School of Engineering, Guaratinguetá, Av. Dr. Ariberto Pereira da Cunha, 333, Pedregulho, Guaratinguetá, SP CEP:12516-410, Brazil

[2] Professional Master's Degree Program, University of Taubaté (Unitau), R. Daniel Danelli, s/n, Vila da Juta, Taubaté, SP CEP:12060-440, Brazil

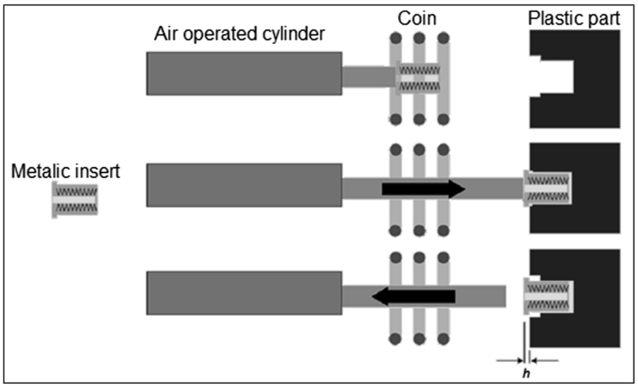

Fig. 1 Insertion (Installation) Height [5]

Although a great deal is known regarding the insertion processes, little is known concerning how much the Heating Temperature and Insertion Time affect the Insertion Height and if there is any interaction between these factors.

This paper presents a model for predicting the Insertion Height for a brass insert in a thermoplastic joined by a heat staking process; the model is obtained using the design of experiments (DOE) and the response surface methodology (RSM).

The Insertion Height must remain within the specified tolerances to ensure the perfect assembly of the parts and to reduce the issues of reworking of the parts, setup changes in the production line, and client complaints. Determining the lowest Heating Temperature and Insertion Time for which the Insertion Height is within the specified tolerance enables energy saving and increases productivity because more parts can be assembled for a given period of time.

The experiments were conducted on intake manifolds used in combustion engines, as illustrated in Fig. 2.

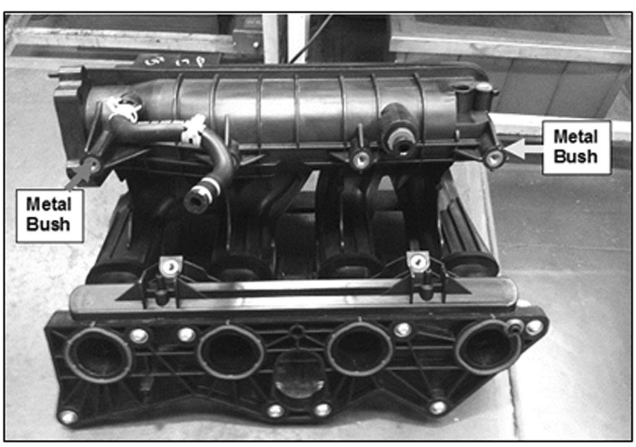

Fig. 2 Engine intake manifold collector with several brass threaded inserts

Materials and Methods

Materials

The inserts used in this research were metallic bushes made of CuZn36Pb3, which can be heated from 250 °C to 370 °C. Figure 3 shows an insert employed in this research. The metal insert is housed into a polyamide base PA6-GF30.

Methods

The goal of this study is to determine the effects of the Heating Temperature and the Insertion Time on the Insertion Height (*h*). Factorial design methods are the most efficient approach to study the joint effect of two or more factors on a response. The most important class of design is the 2^k factorial design, i.e., k factors, with each being at only two levels. This class of design is particularly useful in the early stages of experimental work [12].

In running a two-level factorial experiment, one assumes that the response is approximately linear over the range of the chosen factors levels. This assumption is a reasonable assumption at the beginning of the study; however, it is necessary to be alert to the possibility that a linear model for the response is not adequate because it is possible that the response variable does not behave linearly [12].

Before proceeding, it is necessary to check for curvature in the response variable by adding n runs at the center point of the 2^k factorial design. Next, if the linear model is adequate, then it will be necessary to use more replicates at the factorial points to improve the estimation of the regression coefficients; otherwise a new design, with more points over the factor range, is required to fit a second-order model [12].

Factors range and response variable

The range of the Heating Temperature was chosen according to the polyamide characteristics, such as the melting temperature (260 °C) and the plastic decomposition temperature (340 °C). The range of the Insertion Time was chosen according to the equipment specifications, ranging from 1.00 s to 3.70 s. The center point (300 °C, 2.35 s) is equidistant from

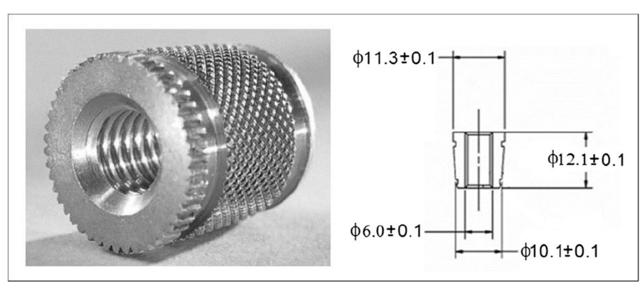

Fig. 3 Metal insert (units in mm)

these extreme points. The current setup before this research was 320 °C and 3.70 s. Table 1 summarizes the level of the factors adopted for this experiment.

The response variable, the Insertion Height, is the distance between the plane of the bushing and the top surface of the cavity, as illustrated in Fig. 1; this variable was measured by means of a digital dial gauge with a resolution of 0.001 mm. The specification for the Insertion Height ranges from 0.000 to 0.300 mm.

Procedures and assumptions

Before every new run, twenty bushes were disposed to ensure that the Heating Temperature was stabilized at the new level.

Hypothesis testing is a classical statistical procedure used to determine the probability that a given hypothesis is true. This approach was used several times in this study. Hypothesis testing consists of two hypothesis: the null hypothesis (H_0), the alternative hypothesis (H_1), and a test statistics, which is used to assess the truth of the null hypothesis.

The hypothesis testing is associated with two types of statistical errors: type I error and type II error. The type I error is the probability of rejecting the null hypothesis, given it is true, it is formally referred to as the test significance level (α). Whenever the calculated probability of type I error occurrence (P-value) is greater than α, there is no statistical evidence to reject the null hypothesis. Otherwise, if the P-value is lower than α, then there is statistical evidence to reject the null hypothesis because the risk of rejecting a valid hypothesis is lower than the one considered acceptable [12–14].

The type II error, also known as β risk, is the probability of accepting the null hypothesis, given it is not true. The β risk is inversely related to the sample size [12–14].

In all hypotheses tests performed in this research, the adopted significance level (α) was 1%; a β risk less than 10% is acceptable for the main effects of the factors, and a β risk approximately 20% is acceptable for the interaction between factors, as suggested by [12–14].

Results and Discussion

The experimental procedure was conducted in two stages. The first stage was a 2^2 factorial design with five replicates at the center point. The number of replicates at the center points should range from three to five [12]. Five replicates were chosen to be conservative. At the end of this phase, evidence of a curvature in the response function was detected over the region of exploration; thus, a new design with more points within the domain of the factors was required to fit a second-order regression model.

Table 1 Level of the factors

Level	Heating Temperature (°C) (Factor A)	Time Insertion (s) (Factor B)
-1	260	1.00
0	300	2.35
+1	340	3.70

2^2 Factorial Design with Five Replicates at the Center Point

Because the main goal of this stage of the experimental procedure was to evaluate a quadratic curvature in the response function, a 2^2 factorial design with a single replicate of each factorial point augmented with five replicates at the center point was employed. According to Montgomery [12], the number of replicates at the center point ranges from 3 to 5; thus, to be conservative, 5 replicates were adopted. The curvature in the response function is evaluated by comparing the average of the runs at the factorial points, $\overline{y}_F$, to the average of the runs at the center point, $\overline{y}_C$. If the averages are different, then there is statistical evidence of a quadratic curvature in the response function over the domain of the factors. Such a check is performed by means of the hypothesis test (1) [12].

$$\begin{cases} H_0 : \overline{y}_F - \overline{y}_C = 0 \\ H_1 : \overline{y}_F - \overline{y}_C \neq 0 \end{cases} \quad (1)$$

If the averages are different, then the null hypothesis must be rejected. According to [12, 13], the appropriate test statistics for hypothesis test (1) is the Student's t-score (t_0), which is given by equation (2).

$$t_0 = \frac{\overline{y}_F - \overline{y}_C}{\sqrt{S^2\left(\frac{1}{n_F} + \frac{1}{n_C}\right)}} \quad (2)$$

Where n_F and n_C are the number of factorial points and the replicates at the center point, respectively; and S^2 is the estimate of variance obtained using the center points.

The hypothesis H_0 is to be rejected if $|t_0| > t_{\alpha/2; n_C - 1}$, where $t_{\alpha/2}; n_C - 1$ is the critical value of the Student's t-distribution with $n_C - 1$ degrees of freedom for a significance level α. According to [13], for a significance level of 0.01, as stated in "Procedures and assumptions" section, and $n_C = 5$ the critical value of the Student's t-distribution is $t_{0.01/2;4} = 4.604$.

The runs of this experiment are summarized in Table 2.

Substituting the numerical values into equation (2), results $t_0 \cong 140.77$, which is greater than $t_{0.01/2;4} = 4.604$; thus, the null hypothesis H_0 can be rejected, suggesting the existence of a quadratic curvature. Therefore, a new design with more points within the domain of the factors will be required.

Table 2 2^2 factorial design with five replicates at the center point

	Heating Temp. (°C)	Insertion Time (seconds)	Insertion Height (mm)
Factorial Points $n_F = 4$	260	1.00	0.862
	260	3.70	0.638
	340	1.00	0.574
	340	3.70	0.493
		$\overline{y}_F =$	0.64175
Replicates at the Center Point $n_C = 5$	300	2.35	0.192
	300	2.35	0.206
	300	2.35	0.216
	300	2.35	0.219
	300	2.35	0.188
		$\overline{y}_C =$	0.2042
		$S^2 =$	0.000193

Several designs exist that are quite good for fitting second-order response models. In particular, the Central Composite Design (CCD) and the Faced-centered Cube Design (FCD) are worthy alternatives [14]. The 3^2 Factorial Design is also a possible choice; although it is not the most efficient method to model the quadratic relationship [12], it was employed because it is a natural extension of the 2^k factorial design, and this issue of the choice of design is not relevant when there are only two factors [14].

3^2 Factorial Design with Five Replicates

This is a factorial arrangement with 2 factors, each at three levels. The low, intermediate, and high levels are coded as −1, 0, and +1, respectively. A very important decision to make at this point was the choice of the number of replicates at each treatment combination. The number of replicates depends on the size of effects that are intended to be detected, i.e., the smaller the effect to be detected, the greater the number of replicates [12]. In general, the number of replicates ranges from 2 to 5; thus, to be conservative, 5 replicates were chosen.

The results from running all combinations of the chosen factors, with each being at three levels with five replicates, are shown in Table 3.

Sample size adequacy checking

Before proceeding, it is worth verifying whether the number of replicates chosen is adequate to keep the β risk within the acceptable range, as previously specified in "Procedures and assumptions" section.

The β risk and the sample size are related by the operating characteristic curves [12–14]. In other words, the operating characteristic curves are employed to verify whether the selected number of replicates is adequate to make the design sensitive to a given difference between treatments.

The link between the β risk and the operating characteristic curves is determined by means of a dimensionless parameter φ given by equation (3) in the case of the main effect of the factors, and by equation (4) for the interaction between factors [12].

$$\phi^2 = \frac{nbD^2}{2aS^2} \tag{3}$$

$$\phi^2 = \frac{nD^2}{2S^2[(a-1)(b-1)+1]} \tag{4}$$

where n is the number of replicates, b and a are the numbers of levels of each factor, D is the minimum difference between two treatment means, and $S^2$is the estimate variance of the response variable.

From Table 3, it is possible to observe that the difference between two treatment means range from 0.054 to 0.131 mm; therefore, it is reasonable to check whether the design is sensitive to a difference of $D = 0.050$ mm between any two treatment means. The estimate variance of the Insertion Height is the average variance of the treatments, which is approximately 0.000288 mm^2. Both factors are studied in three levels.

The results for equations (3) and (4) are summarized in Table 4, considering that the number of replicates (n) is equal to five.

Referring to the operating characteristic curves [12, 13] for the values of φ presented in Table 4, it can be seen that five replicates are sufficient to obtain a β risk less than 1%, i.e., 99% chance of rejecting the null hypothesis if the difference in mean Insertion Height at two Heating Temperatures or two Insertion Times is as high as 0.050 mm. In the same manner, five replicates provide a β risk of approximately 20% for the same difference between any two interactions effects, as seen in Table 4.

Graphical interpretation of the results

To assist in interpreting the results in Table 3, it is helpful to construct a graph of the average Insertion Height at each treatment combination, as shown in Fig. 4.

The Insertion Height decreases from low to intermediate temperature (260 to 300 °C), because, at low temperatures, the polyamide is not adequately melted to enable a perfect insertion [15]. The minimum value for the Insertion Height occurs at approximately 300 °C. From the intermediate to high temperature (300 to 340 °C), the Insertion Height increases, because as the temperature increases, the amount of gas produced in the polyamide also increases, pushing out the insert [5].

Thus, the insertion height follows the order of 1.00 s > 2.35 s > 3.70 s for all temperature. A shorter

Table 3 Insertion Height (mm) for the 3^2 design

Heating	Insertion Time (s)					
Temperature (°C)	1.00		2.35		3.70	
	(−1)		0		(+1)	
	0.851	Average	0.734	Average	0.646	Average
	0.862	0.869	0.739	0.738	0.668	0.650
260 (−1)	0.866		0.754		0.673	
	0.876	Variance	0.745	Variance	0.621	Variance
	0.891	0.000229	0.716	0.000201	0.641	0.000448
	0.281	Average	0.217	Average	0.134	Average
	0.307	0.301	0.209	0.204	0.139	0.139
300 (0)	0.328		0.193		0.161	
	0.304	Variance	0.181	Variance	0.147	Variance
	0.285	0.000358	0.218	0.000260	0.116	0.000275
	0.591	Average	0.554	Average	0.504	Average
	0.586	0.586	0.551	0.531	0.479	0.477
340 (+1)	0.588		0.516		0.465	
	0.584	Variance	0.525	Variance	0.484	Variance
	0.582	0.000012.	0.509	0.000419	0.455	0.000352

Insertion Height is attained at a longer Insertion Time, regardless of the Heating Temperatures.

To confirm the conclusions from the graphical analysis above, it is necessary to compare all the treatments means; the analysis of variance is the appropriate procedure for testing the equality of several means.

Analysis of variance

The analysis of variance (ANOVA) for data in Table 4 is summarized in Table 5.

It can be seen in Table 5 that the *P*-value for the main effects of Heating Temperature and for Insertion Time, as well as for the interaction between these two factors are less than 0.01; thus, they are statistically significant.

The analysis of variance is the decomposition of the variability in the observations through a purely algebraic relationship. However, such a procedure requires that experimental data can be represented by an empirical model, known as the Fixed Effect Model, which expresses the relationship between the observations (response variable) and the factors [12]. Therefore, before adopting the conclusions from the analysis of variance, it is necessary to check the adequacy of such a model by means of the analysis of the residuals.

Model adequacy checking

The primary diagnostic tool for the model adequacy checking is the residual analysis. The residuals (or errors) are defined as the difference between the actual and predicted values for the response variable.

Table 6 presents the standardized residuals for the Insertion Height data in Table 3.

The residual analysis is performed by means of a graphical analysis of the residuals to check whether they are normally and independently distributed with a mean zero and constant variance [12].

A check of the normality assumption can be made by constructing a normal probability plot of the standardized residuals, as shown in Fig. 5. If the underlying error distribution is normal, then this plot will resemble a straight line.

The normal probability plot shows a reasonably linear pattern in the center of the data. However, the tails show deviation from the fitted line. In other words, the middle of the

Table 4 Choice of sample size

($S^2 = 0.000288$ mm^2)	φ^2	n	φ	Numerator degrees of freedom	Denominator degrees of freedom	β risk
Main effects	$5n$	5	4.658	2	36	< 0.01
Interaction effects	$1n$	5	2.083	4	36	≈ 0.20

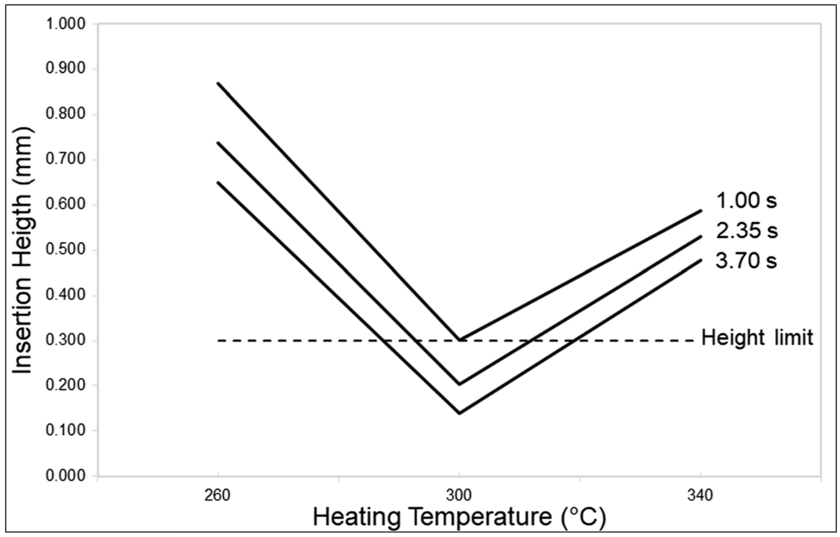

Fig. 4 Average Insertion Height at each treatment combination

residuals shows a mild S-like pattern. The first few points show increasing deviation from the fitted line below the line, and the last few points show increasing deviation from the fitted line above the line; thus, this pattern suggests the possibility of a distribution with long tails relative to the normal distribution. However, according to [12], moderate deviation from normality is of little concern in the fixed effects analysis of variance because the analysis of variance is robust to the normality assumption. Thus, there were no considerable violations of normality.

To be more objective, three classical goodness-of-fit tests were used to check whether the residuals come from a normally distributed population, the Kolmogorov-Smirnov test [16], the Anderson-Darling test [16], and the Ryan-Joiner test [17]. All of these tests are hypothesis testing, as stated in (5).

$$\begin{cases} H_0 : \textit{the data}\text{ are }\textit{normally distributed} \\ H_1 : \textit{the data}\text{ are }\textit{not normally distributed} \end{cases} \quad (5)$$

In this paper, the statistics for the Kolmogorov-Smirnov test, the Anderson-Darling test, and the Ryan-Joiner test are denominated KS, AD, and RJ, respectively. Although these statistics can be compared to the respective critical values, most of the data analysis software give the P-values associated with them. Therefore, if the P-value is greater than the significance level (α) adopted for this test (which is 0.01 in this paper), there is no statistical evidence to reject the null hypothesis, and the residuals should be considered normally distributed [12, 14].

Using the software Minitab, the Komolgorov-Smirnov test produced the statistics KS = 0.065 with a P-value >0.150, the Anderson-Darling test produced the statistics AD = 0.361 with a P-value = 0.430, and the Ryan-Joiner produced RJ = 0.998 with a P-value >0.100. Thus, according to the previous paragraph, the normality assumption of the residuals has not been violated.

The independence assumption can be evaluated by means of a plot of residuals in the time order of the data collection, as shown in Fig. 6.

The plot of the residuals versus the observation order is helpful in detecting a correlation between the residuals. Positive serial correlation exists when residuals tend to be followed by residuals of the same sign and approximately the same magnitude. Negative serial correlation exists when residuals of one signal tends to be followed by residuals of the opposite sign. If such a correlation exists, then the independence assumption on

Table 5 Analysis of variance for the insertion height data

Source of variation	Degrees of freedom	Sum of Squares	Mean square	$F_0 = \frac{\textit{Mean Square}}{\textit{Mean Square Error}}$	P-value
Heating Temperature	2	2.19020	1.09510	3861.27	0.000
Insertion Time	2	0.20164	0.10082	355.49	0.000
Heating Temp. x Time Insert.	4	0.01610	0.00403	14.19	0.000
Error	36	0.01021	0.00028		
Total	44	2.41815			

Table 6 Standardized residuals for the Insertion Height (the numbers in parentheses indicate the order of data collection)

Heating	Insertion Time (s)					
Temp. (°C)	1.00		2.35		3.70	
	−1.208(1)	0.451(29)	−0.239(44)	0.491(15)	−0.252(8)	−1.912(13)
260	−0.478(22)	1.447(23)	0.093(19)	−1.434(40)	1.208(32)	−0.584(28)
	−0.212(24)		1.089(37)		1.540(43)	
	−1.328(6)	0.199(17)	0.889(4)	−1.500(5)	−0.358(12)	0.504(45)
300	0.398(25)	−1.062(34)	0.358(9)	0.956(21)	−0.027(27)	−1.553(38)
	1.792(39)		−0.704(26)		1.434(42)	
	0.319(11)	0.119(14)	1.527(2)	−0.996(7)	1.766(3)	−0.823(10)
340	−0.013(30)	−0.146(36)	1.328(16)	−0.398(20)	0.106(18)	0.438(31)
	−0.279(41)		−1.460(35)		−1.487(33)	

the errors has been violated [12]. Figure 6 shows that the residuals bounce randomly around the zero line, suggesting that there is no correlation.

The homoscedasticity assumption can be verified by means of the plot of the residuals versus the fitted value [12], as shown in Fig. 7.

A constant variance implies the variation of observation is approximately constant as the magnitude of the observation increases. Figure 7 reveals that the residuals are structureless, i.e., they are unrelated to the predicted response. As a result, the assumption of homogeneity of variances has not been violated.

Because none of the assumptions was violated, the conclusions based on ANOVA (Table 5) remain valid, including the fact that the interaction between the factors is significant. In such cases, comparisons between the means of one factor may be obscured by the interaction. Therefore, it is recommended to apply a multiple comparisons procedure, such as Tukey's test, to evaluate the effect of each factor [12].

Tukey's test

Tukey's test is used in conjunction with Analysis of Variance to find the means that are significantly different from each other. The test compares all possible pairs of means by fixing one factor at a specific level and testing the means of the other factors at that level. This test is based on a studentized range distribution (q).

Therefore, first, it is necessary to determine the differences among the means of all factors combinations, which are shown in Table 7, and afterwards, it is necessary to test these means.

For the data in Table 4, Tukey's test declares that two means are significantly different if the absolute value of their sample mean differences exceeds $T_{0.01} \approx 0.033$ mm, which can be calculated according to Montgomery [12] for $q_{0.01}(3;36) \approx 4.39$ (by interpolation) and considering that the error variance is estimated as the Error Mean Square in Table 5 ($MS_E = 0.000280$).

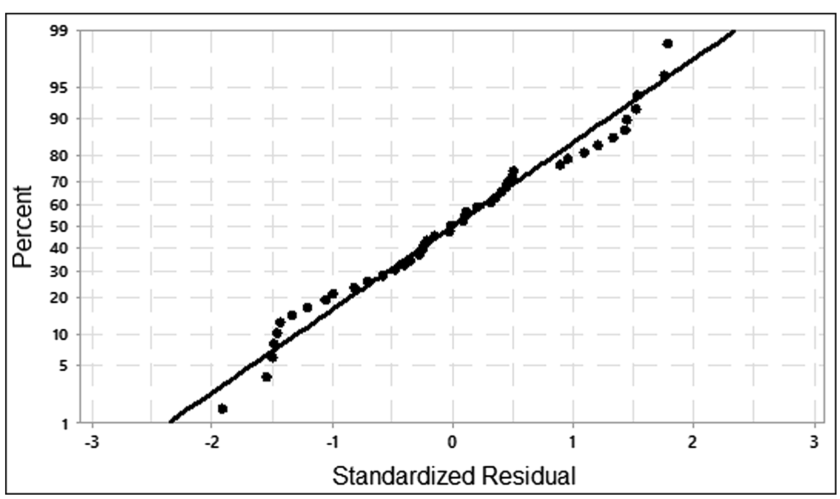

Fig. 5 Normal probability plot of the residuals from the fixed effects model

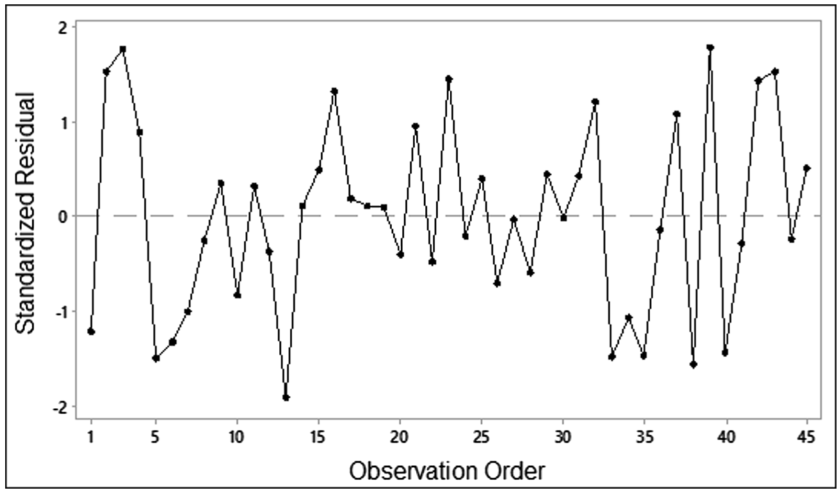

Fig. 6 Residuals versus observation order

Therefore, Tukey's test indicates that the Insertion Height is different for all treatment combinations in Table 4; thus, the previous conclusions are confirmed.

Recall that the tolerance for the Insertion Height for the process under analysis ranges from 0.000 to 0.300 mm; as a result, a trade-off exists between the Heating Temperature and the Insertion Time to increase productivity and save energy while ensuring that the Insertion Height remains within the specifications (see Fig. 4).

To perform a better analysis of this trade-off between Insertion Time and Heating Temperature, a regression model for the Insertion Height as a function of these factors is constructed.

Response Surface Model

A response surface or regression model is useful for prediction, process optimization, or control processes. Equation (6) is the second-order regression model for two variables [14].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon \quad (6)$$

where y represents the response function, and x_1 and x_2 are the factors of the experiment, with the independent variables often called the predictor variables or the regressors. The parameters β_i, $(i = 0, 1, \cdots)$ are called the regression coefficients, and ε is the model error.

Equation (6) is called the multiple linear regression model; the most popular method to estimate the regression coefficients β's is the least squares method.

The fitted regression model can be written as equation (7) [14]:

$$\hat{y} = B_0 + B_1 x_1 + B_2 x_2 + B_{12} x_1 x_2 + B_{11} x_1^2 + B_{22} x_2^2 \quad (7)$$

where $\hat{y}$ estimates y, $B_i\,(i = 0, 1, \cdots)$ estimates$\beta_i\,(i = 0, 1, \cdots)$, x_1 is the Heating Temperature, and x_2 is the Insertion Time.

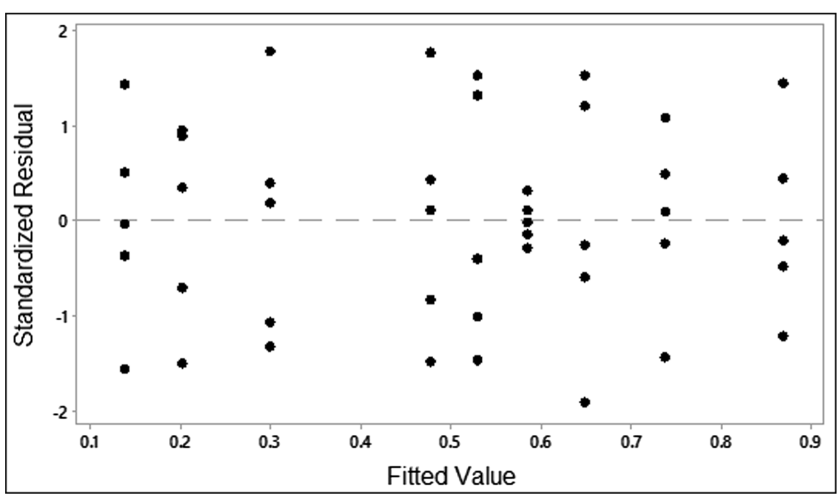

Fig. 7 Residuals versus fitted value

Table 7 Differences among the means of all factor combinations

Insertion Time (s)	Heating Temperature (°C)		
	260	300	340
1.00 → 2.35	0.131	0.097	0.055
1.00 → 3.70	0.219	0.162	0.109
2.35 → 3.70	0.088	0.065	0.054
Heating Temperature (°C)	Insertion Time (s)		
	1.00	2.35	3.70
260 → 300	0.568	0.534	0.511
260 → 340	0.283	0.207	0.173
300 → 340	0.285	0.327	0.338

Table 8 presents a summary of the information concerning the regression model for the data in Table 3, listed in coded variables. The regression coefficients are presented followed by the respective *P*-value. In this case, the *P*-value is associated with a hypothesis test for which the null hypothesis state that the regression coefficient is zero [14]. Therefore, if *P*-value > α, then the null hypothesis cannot be rejected, i.e., the corresponding population regression coefficient is zero.

The adjusted R^2 statistics, R^2(adj), the maximum value of which is 1, is used to assess the model performance. R^2(adj) is a measure of the amount of reduction in the variability of the estimated variable obtained by using the regressor variables in the model. In opposition to the classical statistics R^2, its value does not always increases as variables are added to the model. In fact, if unnecessary terms are added, its value decreases [12, 14]. Thus, the model exhibits a good performance because R^2(adj) is very close to 1 and the residual error is notably low, approximately 0.0168. The performance analysis of the model can be complemented by the ANOVA presented in Table 9.

The ANOVA indicates that the model fits the data well; in other words, the model is adequate to describe the Insertion Height. Before relying on ANOVA conclusions, it is necessary to investigate the model adequacy.

Table 8 Regression model in coded variables for data in Table 4

Regressor	B_i	*P-value*
Intercept	0.20593	< 0.001
x_1	−0.11033	< 0.001
x_2	−0.08163	< 0.001
$x_1 \cdot x_2$	0.02765	< 0.001
${x_1}^2$	0.42720	< 0.001
${x_2}^2$	0.01310	= 0.018
R^2(adj) =	0.9949	
Res. Error =	0.0168	
Reg. *P*-value=	< 0.001	

Model adequacy checking

As performed previously, this verification can be accomplished by performing a residual analysis based upon the standardized residuals presented in Table 10.

The full graphical analysis is presented in Fig. 8. In Fig. 8(a), the normal probability plot shows a moderate deviation from the fitted line. However, according to the normality tests of Kolmogorov-Smirnov (KS = 0.106; *P*-value >0.1500), Anderson-Darling (AD =0.457; *P*-value =0.254), and Ryan-Joiner (RJ = 0.986; *P*-value >0.100), the residuals above can be considered normally distributed, i.e., the normality assumption of the residuals has not been violated. Figure 8(b), (c) and (d) indicate that the variance of the observed Insertion Height is stable with respect to the predicted Insertion Height, Heating Temperature and Insertion Time, respectively. Thus, the homoscedasticity of the residuals has not been violated.

Once the model presented in Table 8 is confirmed, the next step is to study the behavior of the response function.

Study of the response surface

The regression model presented in Table 8 can be rewritten as follows:

$$\hat{y} = 0.2059 - 0.1103x_1 - 0.08163x_2 + 0.02765x_1x_2 + 0.42720x_1^2 + 0.01310x_2^2 \quad (8)$$

and, for the uncoded variable, as

$$\hat{y} = 25.6062 - 0.1642x_1 - 0.2479x_2 + 0.0005120x_1x_2 + 0.0002670x_1^2 + 0.007188\,x_2^2 \quad (9)$$

as illustrated in Fig. 9.

The response surface is a parabolic cylinder, which has one critical point solely, and it is at 301.17 °C and 6.51 s. Although this point is a minimum point, it is outside the domain of the Insertion Time (1.00–3.70 s); therefore, it is necessary to identify the extreme value of $\hat{y}$ at the response function boundary. The minimum point is the vertex of the parabola resulting from the intersection of the surface with a plane parallel to the Insertion-Height - Heating-Temperature$(x_1O\hat{y})$ plane at the Insertion Time equal to 3.70 s. Thus, substituting

Table 9 ANOVA for the regression model presented in Table 8

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	$F0 = \frac{Mean\ Square}{Mean\ Square\ Error}$	*P*-value
Regression	5	2.40713	0.48143	761.15	0.000
Error	39	0.01103	0.00028		
Lack-of-Fit	3	0.00081	0.00027	0.95	0.423
Pure Error	36	0.01021	0.00028		
Total	44	2.41816			

$x_2$by 3.70 s in equation (9), it will become a single variate function, with a local minimum at $x_1 = 303.87^{\circ}C$ which is $\hat{y}_{\min} = 0.133\,mm$.

The Insertion Height contour, Fig. 10, is useful to determine the region over which the Insertion Height is within the specified tolerance (0 to 0.300 mm).

Several combinations of the Heating Temperature and the Insertion Time correspond to an Insertion Height less than or equal to 0.300 mm. The combination that maximizes productivity and minimizes energy consumption is the Heating Temperature of approximately 300 °C and the Insertion Time equal to 1.00 s.

Although the optimum point has been achieved, it is necessary to check the capability of the process.

Process Capability

Among the several statistics that can be used to measure the capability of a process, the Process Capability Index, C_{pk}, will be used in this work. According to Montgomery [18], the index C_{pk} can be estimated as follows:

$$C_{pk} = \min\left(\frac{USL-\hat{y}}{3S}, \frac{\hat{y}-LSL}{3S}\right) \tag{10}$$

Table 10 Residuals for the second-order model in Table 8

Order	y_i	$\hat{y}_i$	Residual	Stand. Res.	Order	y_i	$\hat{y}_i$	Residual	Stand. Res.
1	0.851	0.866	−0.015	−0.964	24	0.866	0.866	0.000	0.010
2	0.554	0.523	0.031	1.968	25	0.307	0.301	0.006	0.399
3	0.504	0.482	0.022	1.434	26	0.193	0.206	−0.013	−0.816
4	0.217	0.206	0.011	0.698	27	0.139	0.137	0.002	0.101
5	0.181	0.206	−0.025	−1.573	28	0.641	0.647	−0.006	−0.408
6	0.281	0.301	−0.020	−1.241	29	0.876	0.866	0.010	0.659
7	0.516	0.523	−0.007	−0.429	30	0.586	0.590	−0.004	−0.252
8	0.646	0.647	−0.001	−0.083	31	0.484	0.482	0.002	0.135
9	0.209	0.206	0.003	0.193	32	0.668	0.647	0.021	1.345
10	0.465	0.482	−0.017	−1.098	33	0.455	0.482	−0.027	−1.748
11	0.591	0.590	0.001	0.073	34	0.285	0.301	−0.016	−0.988
12	0.134	0.137	−0.003	−0.214	35	0.509	0.523	−0.014	−0.870
13	0.621	0.647	−0.026	−1.707	36	0.584	0.590	−0.006	−0.382
14	0.588	0.590	−0.002	−0.122	37	0.754	0.743	0.011	0.664
15	0.745	0.743	0.002	0.097	38	0.116	0.137	−0.021	−1.350
16	0.551	0.523	0.028	1.779	39	0.328	0.301	0.027	1.724
17	0.304	0.301	0.003	0.210	40	0.716	0.743	−0.027	−1.733
18	0.479	0.482	−0.003	−0.189	41	0.582	0.590	−0.008	−0.512
19	0.739	0.743	−0.004	−0.282	42	0.161	0.137	0.024	1.489
20	0.525	0.523	0.002	0.139	43	0.673	0.647	0.026	1.670
21	0.218	0.206	0.012	0.761	44	0.734	0.743	−0.009	−0.597
22	0.862	0.866	−0.004	−0.250	45	0.147	0.137	0.010	0.606
23	0.891	0.866	0.025	1.633					

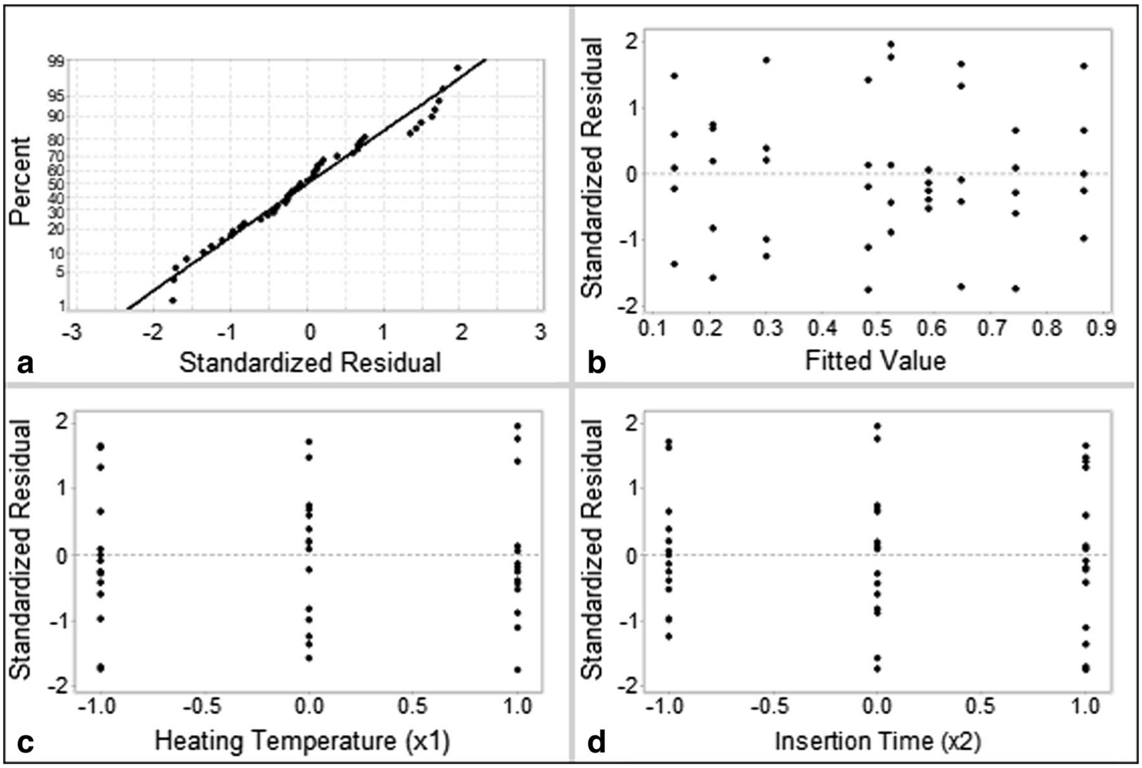

Fig. 8 (**a**) Normal probability plot. (**b**) Residual versus fitted value. (**c**) Residual versus Heating Temperature. (**d**) Residual versus Insertion Time

Considering the upper limit specification (USL) is equal to 0.300 mm, the lower specification limit is equal to 0.000, and the standard deviation estimate (S), as the residual error of the model from Table 8, is 0.0168 mm, (10) can be rewritten as follows:

$$C_{pk} = \min\left(\frac{0.300-\hat{y}}{0.0504}, \frac{\hat{y}}{0.0504}\right) \qquad (11)$$

Unfortunately, the process is not capable of achieving the optimal setup (300 °C, 1.00 s), once the C_{pk} is very low, i.e., the probability that the Insertion Height for an insert bush is out of specification is very high. A standard acceptance criterion for a process in the automotive industry [19] is C_{pk} is equal to 1.33, which corresponds to an Insertion Height of 0.233 mm. By replacing $\hat{y}$ with 0.233 in equation (9), one obtains equation (12), which represents the ellipse illustrated in Fig. 11.

$$0.0002670\, x_1^2 + 0.007188\, x_2^2 - 0.1642\, x_1 - 0.2479\, x_2 + 0.0005120\, x_1 x_2 + 25.3732 = 0 \qquad (12)$$

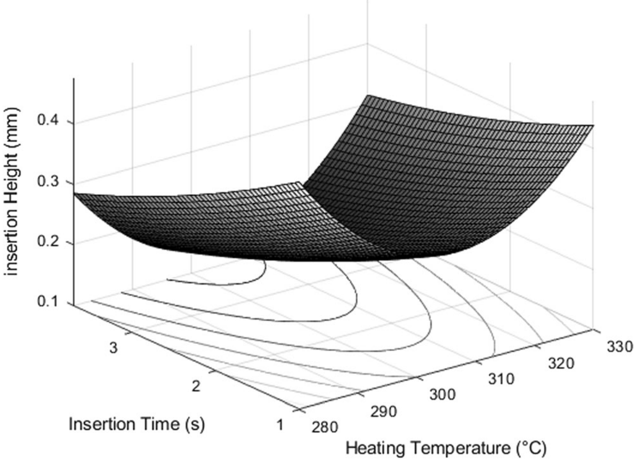

Fig. 9 Insertion height surface

The contour line in Fig. 11 shows several combinations of the Heating Temperature and the Insertion Time, for which the Insertion Height is 0.233 mm, and therefore, it is within the specifications. Thus, it is possible to choose the combination that maximizes the productivity or minimize the energy consumption by moving the operating point along such a curve.

Productivity Improvement and Energy Saving

The current operating point is (320 °C, 3.70 s), and it can move along the curve in Fig. 11, depending on the production requirements. According to the plant manager, a decrease of 1 s in the Insertion Time saves approximately 366 h/yr. Thus, if the main concern is to improve the process productivity, the Insertion Time must be minimized in equation (12), leading to

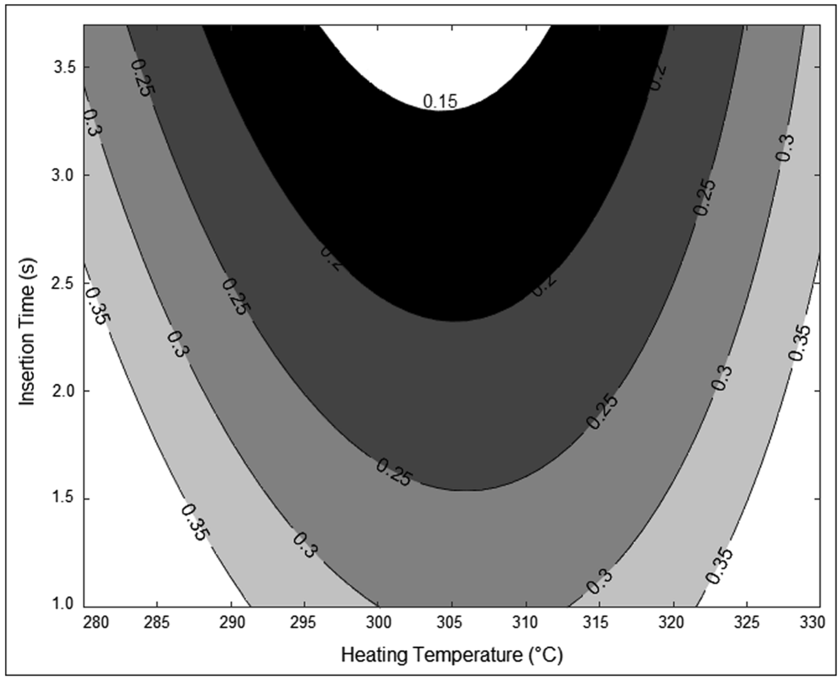

Fig. 10 Insertion height contour plot (mm)

the minimum point (305.9 °C, 1.62 s), as illustrated in Fig. 12, resulting in a time saving of approximately 760 h/yr.

However, if the main concern is energy saving, a decrease of approximately 0.020 kWh/yr. in the electric energy consumption accompanies a reduction of 1 °C in the Heating Temperature. Thus, the operating point can be moved along the curve towards to the minimum value of the Heating Temperature, 283.5 °C, which corresponds to 3.70 s, resulting in a saving of 0.730 kWh/yr as illustrated in Fig. 13.

Conclusions

This paper presented a model for predicting the Insertion Height for the joining of a brass insert to a thermoplastic form using a heat staking process. First, factorial designs were applied to investigate the behavior of the Insertion Height with respect to the Heating Temperature and Insertion Time, which were the two controllable factors of the process. Afterwards, the RSM was used to obtain the desired model.

A 2^2 factorial design with a central point with five replicates demonstrated the existence of curvature in the response surface. As a result, it was necessary to run a new design to provide extra points to model the response function. The 3^2 factorial design was chosen. Although the Insertion Height was affected by both factors and their interaction, the experiments revealed that the Insertion Height was affected primarily by the heating.

From the Insertion Height contour plot, it was possible to define a region where the Insertion Height is within the specifications, i.e., less than or equal to 0.300 mm. The combination of 300 °C and 1 s was found to correspond to an Insertion Height of 0.300 mm. Although this set point allows for the highest productivity, the process cannot be implemented because it showed a very low C_{pk}.

C_{pk} equal to 1.33 is a standard acceptance criterion for processes in the automotive industry. A C_{pk} equal to 1.33

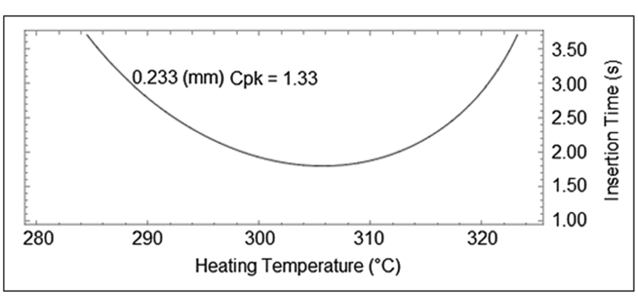

Fig. 11 Contour plot for insertion height 0.233 mm

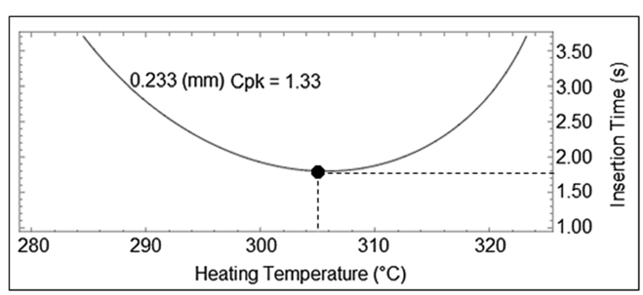

Fig. 12 Point of maximum productivity

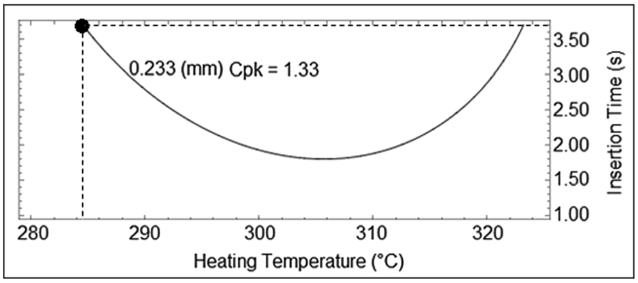

Fig. 13 Minimum power consumption point

corresponds to an Insertion Height equal to 0.233 mm; as a result, a contour line for the Insertion Height equal to 0.233 mm (C_{pk} equal to 1.33) was determined. Thus, the operating point can be set on this line, and it can be chosen based on the optimization criterion, maximizing energy saving or productivity.

Changing the current setpoint from 320 °C and 3.70 s to 283.55 °C and 3.70 s results in an energy saving of 0.730 kWh/yr. However, the new operating point of 305.9 °C and 1.62 s results in a time saving of 760 h/yr. Therefore, the benefits of the optimized settings can be even greater, depending on the produced items, production level, and number of machines operating simultaneously.

Acknowledgements This work was supported by the São Paulo Research Foundation (FAPESP) [grant number 2014/18628-9].

References

1. Amancio Filho ST, Santos JF (2009) Joining of Polymers and Polymer–Metal Hybrid Structures: Recent Developments and Trends. Polym Eng Sci 49(8):1461–1476. https://doi.org/10.1002/pen.21424
2. Tamrin KF (2013) Laser lap joining of dissimilar materials: a review of factors affecting joint strength. Mater Manuf Process 28(8): 857–871. https://doi.org/10.1080/10426914.2013.792413
3. Grujicic M, Sellappan V, Omara MA, Seyr N, Erdmannc M, Holzleitner J (2007) An Overview of the Polymer-to-metal Direct-adhesion Hybrid Technologies For Load-bearing Automotive Components. J Mater Process Technol 197(1–3): 363–373. https://doi.org/10.1016/j.jmatprotec.2007.06.058
4. Rojek M, Stabik J, Muzia G (2010) Thermography in plastics welding processes assessment. Journal of Achievements in Materials and Manufacturing Engineering 41(1–2):40–47
5. Dahake, G. (2011) Heat Staking: Inserting Metal into Plastic. Precision Induction Heating 9
6. Jegadeesan AJ, Karunamoorhty L, Arunkumar N (2015) Investigation of Hot-Staking Process Parameters and Evolution of Improvement in Strength. Mater Manuf Process 30(2):155–161. https://doi.org/10.1080/10426914.2014.961479
7. Aktas S (2012) Effect of spot welding parameters on tensile properties of DP 600 steel sheet joint. Mater Manuf Process 27(7):756–764. https://doi.org/10.1080/10426914.2011.647940
8. Nizamzul N (2011) Investigating the effects of injection molding parameters on the mechanical properties of recycled plastic parts using Taguchi method. Materials and Manufacturing Process 26(2): 202–209. https://doi.org/10.1080/10426914.2010.529587
9. Barghash MA, Alkaabneh FA (2014) Shrinkage and Warpage Detailed Analysis and Optimization for Injection Molding Process Using Multistage Experimental Design. Qual Eng 26(3):319–334. https://doi.org/10.1080/08982112.2013.852679
10. Marré M (2012) An experimental study on the groove design for joints produced by hydraulic expansion considering axial or torque load. Mater Manuf Process 27(5):545–555. https://doi.org/10.1080/10426914.2011.593232
11. Lambiase F, Durante M (2017) Mechanical behavior of punched holes produced on thin glass fiber reinforced plastic laminates. Compos Struct 173:25–34. https://doi.org/10.1016/j.compstruct.2017.04.003
12. Montgomery DC (2012) Design and Analysis of Experiments, 8th edn. John Wiley & Sons, New York
13. Pearson, E. S.; Hartley, H. O. Bibliometrika Tables for Statisticians, third ed. Cambridge Unuversity Press, Cambridge
14. Myers RH, Montgomery DC, Anderson-Cook CM (2009) Response Surface Methodology, 3rd edn. John Wiley & Sons, New York
15. Sullivan G, Crawford L (2003) The Heat Staking Advantage. Plastic Decorating Magazine 70:11–12
16. Chakravarti IM, Laha RG, Roy J (1967) Handbook of Methods of Applied Statistics, 1st edn. John Wiley & Sons, New York
17. Ryan Jr., T. A.; Joiner, B. L. (1976) Normal Probability Plots and Tests for Normality, Technical Report, Statistics Department, The Pennsylvania State University
18. Montgomery DC (2013) Introduction to Statistical Quality Control, 7th edn. John Wiley & Sons, New York
19. Albing, M. (2006) Process Capability Analysis with Focus on Indices for One-sided Specification Limits. Luleå University of Technology - Department of Mathematics, ISSN 1402–1757