Efficiency in contamination-free machining using microfluidic structures

Carlo Ferri[a,1], Timothy Minton[a,*], Saiful Bin Che Ghani[a,2], Kai Cheng[a]

[a] *Brunel University, AMEE - Advanced Manufacturing and Enterprise Engineering, Kingston Lane, Uxbridge, Middlesex, UB8 3PH, UK*

Abstract

The plastic deformation of the material in the chip formation and the friction when the chip slides on the rake face of the insert generate heat. The heat generation is responsible for a temperature rise of the chip, of the insert and of the newly created surface on the workpiece. Adhesion and diffusion between the chip and the insert are thus facilitated with detrimental effects on the tool wear. A cooling system based on microfluidic structures internal to the insert is considered in this study as a means of controlling the temperature at the chip-insert interface. The coolant and the part never enter in contact. Hence contamination of the part by coolant molecules is prevented. The aim of this study is to identify and to quantify the effect of the cutting parameters on the effectiveness of the internal cooling system. To measure this effectiveness an efficiency ratio r is defined as the percentage of the

*Corresponding author. Tel.: +44 1895 267945; fax: +44 1895 267583.
Email address: `timothy.minton@brunel.ac.uk` (Timothy Minton)
[1] Present address: Via XI Febbraio 40, 24060 Castelli Calepio, BG, Italy
[2] Permanent address: Universiti Malaysia Pahang - Mechanical Engineering , Lebuhraya Tun Razak, 26300 Gambang, Malaysia

mechanical power actually needed at the tool to remove material that is thermally dissipated by the internal flow of the coolant. Similarly, a specific efficiency ratio r' is also defined by considering the mechanical power per volume flow rate of the material removed and the dissipated thermal power per volume flow rate of the coolant. Both r and r' are then analysed in a 3^3 factorial experiment within the space of the technological variables depth of cut, feed rate and cutting speed. The cutting trials were conducted in turning operations of AA6082-T6 aluminium alloy. Linear Mixed-effects models were fitted to the experimental results using the maximum likelihood method. The main finding was that the efficiency ratio r depends only on the feed rate and the cutting speed but not on the depth of cut. An interaction effect of the feed rate and the cutting speed on the efficiency was also found significant. Higher efficiency is attainable by decreasing cutting speed and feed rate. The maximum efficiency predicted in the technological region investigated was 10.96 %. The specific efficiency once log-transformed was found linearly increasing with the depth of cut and the feed rate, whereas being insensitive to the cutting speed.

Keywords: Cutting temperature, internally-cooled tool, contamination-free machining, dry machining, Linear mixed-effects statistical models

---

1. Introduction

Dry cutting of key engineering materials is the epitome of sustainability in metal cutting. The removal of metal working fluids (MWF) from the machining processes is of benefit to the machine operator, swarf recycling and ultimately the environment. Reducing the temperature of the cutting tool and workpiece is one of the main purposes of the MWF, together with facilitating the removal of the chip from the machining area. Using an external supply of coolant makes it difficult for the fluid to penetrate into the tool-chip contact area. It is also difficult to quantify the amount of heat transferred between the cutting edge and the MWF. Dry machining removes the externally supplied coolant from the machining process at the expense of the cooling effect it provides. Although this method is acceptable for certain materials like aluminium, it may be problematic for high strength materials and certain grades of aluminium which contain harder elements like silicon. High temperatures which are uncontrolled due to lack of cooling can cause high wear rates and can dramatically reduce the useful life of the tooling insert. In some extreme cases the tool can become damaged not via traditional wear mechanisms but through deformation of the cutting edge [1]. Monitoring of the cutting temperature is a well-established research goal and has been presented using many differing technologies including an embedded thermocouple [2], the tool-work thermocouple [3], the calorimetric method [4], an embedded sensor film [5] and optical methods [6, 7]. Some of these methods are not applicable when using an external coolant supply. Dry machining allows the monitoring of the tool/chip temperature via the tool-work thermocouple [3] or optical methods [6, 7]. These methods however require time consuming setups or expensive auxiliary equipment and are hence better suited to a laboratory environment.

The method of indirect cooling is known in the area of metal cutting and has been steadily increasing in popularity since 1970 when Jefferies published the idea of an internally cooled single-point cutting tool [8]. The main benefit of the internally cooled tool is the indirect application of a cooling effect to the tool-chip interface. Previous research in the field of indirect cooling methods has shown that it is possible to reduce significantly the cutting temperature. In particular, Ferri *et al.* [9] compared the chip temperature in dry turning of the aluminium alloy AA6082-T6 when using conventional and internally cooled tools. Their main finding was that the internally-cooled tools appeared increasingly effective in containing the chip temperature while increasing the depth of cut. In a research effort jointly sponsored by the US Environmental Protection Agency and the Department of the Army, Rozzi *et al.* [10] patented a device to cool indirectly the tool-chip interface by creating micro-channels and a finned heat exchanger within the tool suitable for the use with cryogenic fluids (typically liquid Nitrogen). Sanchez *et al.* [11] proposed a similar apparatus where the cooling fluid flowing within the tool evaporates in proximity of the cutting edge, with the latent heat being provided by heat transfer with the tool-chip interface. In a condenser outside the tool holder, the fluid is then condensed again. The resulting liquid phase is re-conveyed within the tool, thus realising a close-loop circulation of the coolant. Liang *et al.* [12] studied the use of the heat pipe technology in turning operations. A heat pipe is a heat conductor in which the latent heat of evaporation is used for heat transfer purposes in experimental situations where differences in temperature are small. Moreover, a heat pipe operates without any external power supply. Shu *et al.* [13] presented a study based on the finite element method to simulate numerically turning operations in presence of both liquid coolant flowing in channels internal to the tool and a heat pipe. Uhlmann *et al.* [14] compared wet machining, dry machining and machining with an internally-cooled tool. They investigated the influence of different coolant temperatures on the tool flank wear (VB) and on the workpiece surface roughness. Their main finding is that the tool wear in dry machining appears larger than in the other cases. They tested internally-cooled tools with coolant temperatures of 20 °C and -10 °C. The tool flank wear in both these cases and in the wet machining were most similar. The internally-cooled tool with coolant at 20 °C appeared only slightly less worn (cf. figure 3 in Uhlmann *et al.* [14]).

Moreover, internally cooling the tool also provides the unique possibility to manipulate the cutting temperature without necessarily changing core machining parameters such as the cutting speed, the feed rate or the depth of cut. Whilst specifically focusing on a closed loop coolant supply within the tool shank, the introduction of two additional control variables such as the coolant supply flow rate and the coolant temperature can be deployed to affect the metal removal process. The concept of a coolant supply within the cutting tool itself also presents a great opportunity to quantitatively assess the thermal energy that the coolant conveys away from the cutting zone. The metal cutting process generates high heat and large thermal gradients [3]. According to Micheletti (cf page 203 in [15]), heat is almost instantaneously generated where work is done during cutting. Thus, the location of the heat sources is identified in the areas where the work due to the plastic deformation of the metal and to the friction of the chip on the rake face happen. If the tool is not in ideal conditions, i.e. if it is not perfectly sharpened, friction work also happens between the surface of the workpiece and the clearance face of the tool (also known as flank face) [15]. Boothroyd [16] measured the temperature distribution and constructed isotherm patterns in the workpiece, the chip and the tool by making joint usage of infra-red photography and thermocouples. From those measurements, Boothroyd was also able to derive the heat transferred into the chip, the tool and the workpiece. Boothroyd's results, displayed in the table on page 797 in [16], appear consistent with those reported by Micheletti (cf page 209 in [15]): most of the heat generated during the cutting process is transferred into the chip, say about 60 and 80 %, depending on the machining conditions; the remaining part is transferred into the tool and into the workpiece in similar proportions.

When the coolant flows internally to the insert and close to the cutting edge, a part of the generated heat is transferred into the coolant and away from the cutting zone. The heat transfer occurred is evidenced through the increment of the coolant temperature which is also instrumental to its measurement. This can all be achieved without the contamination of the tool and of the workpiece which instead occurs with external coolant supplies. For this reason the authors used in the title and elsewhere the terms 'contamination-free machining'. At first sight, this may appear as an oxymoron. In fact, for a metal cutting process to happen a tool must enter in contact with the workpiece. The cutting edge of the insert must be harder than the material to cut. Thus cutting edge and workpiece are of different materials. It is a reasonable expectation that during the cutting process a proportion of the material worn off the flank face (clearance face) of the tool will contaminate the workpiece at least on a sub-micrometre scale. Thus, strictly speaking, as long as flank wear exists on the tool, a cutting process is always most likely to pollute the workpiece with tool material. The term 'contamination-free' is therefore to be considered within these limitations.

In some cases reducing the temperature of the workpiece or cutting insert by too great a margin might be a problem. For example, if there is a strong work-hardening effect on the material the cutting forces may increase dramatically and induce additional issues with the surface finish and the surface integrity [17]. Another issue might be a thermal shock of the cutting insert. However, the manipulation of the coolant flow rate and/or the coolant temperature would make the management of these events possible. The benefits of a reduced cutting temperature appear to out-weigh the potential troubles by far. An increase in tool life is possible and a control of the critical temperature above which thermally induced wear mechanisms take place is achievable [18]. In this study, a tool system is designed and manufactured to cool the cutting insert by the adduction of the coolant in the proximity of the cutting insert via microfluidic structures within the tool. These structures prevent any possible contact between the coolant and the part. A cooling efficiency ratio is then defined and computed in a range of experimental conditions defined by the triplets of machining parameters cutting speed (v_c), feed rate (f) and depth of cut (a_p). This efficiency ratio denotes the portion of the total machining power which is transferred to the coolant in the form of thermal power. From a conceptual point of view, establishing experimentally how this efficiency ratio depends on (a_p, f, v_c) provides other researchers a further potential means of validating their theories regarding the thermal characteristics of the machining process. From a practitioner's point of view, this efficiency ratio can become a useful instrument in the selection of the coolant flow rate and coolant temperature at the inlet of the tool system. For example, cutting speed, feed rate and depth of cut may be set to comply with productivity requirements and/or the optimisation of some cost function. By setting the triplet (a_p, f, v_c), the power request for machining a given geometry from a given blank is uniquely determined. The knowledge of the efficiency ratio of the cooling system for the selected triplet (a_p, f, v_c) allows then the practitioner to know how much thermal power would be transferred away by the cooling system, had he or she set the flow rate and the inlet temperature of the coolant to the same values of this investigation. Prior to any actual machining, the efficiency ratio can therefore suggest to the practitioner whether the flow rate and the inlet temperature of the coolant may need increasing or decreasing in order to balance the mechanical power and have a thermally steady machining condition. More in general, this study of the efficiency ratio may constitute a stepping stone towards the formulation of a performance objective function (e.g. cost, profit) to be optimised in the newly established penta-dimensional technological space of depth of cut, feed rate, cutting speed, coolant flow rate and coolant inlet temperature.

2. Experimental set-up

The tool has been assembled and secured to a dynamometer as shown in Figure 1. The dynamometer was a three component Kistler type 9257B which had been attached to the tool turret of an Alpha Colchester Harrison 600 Group CNC lathe.

[Figure 1 about here.]

The workpiece material chosen for this study was Aluminium 6082-T6 (0.7-1.3 % Si and 0.6-1.2 % Mg). This aluminium alloy is readily available and widely used in numerous applications, an additional benefit is the low mechanical property demands on the tooling insert and therefore yields a low wear rate. A cylindrical workpiece of 65 mm diameter and 450 mm length was used. The internally cooled tool was enhanced in its measuring capability by mounting K-type thermocouples. These were installed within the inlet and the outlet pipes, close to where these pipes enter the tool body. These sensors measured the inlet/outlet coolant temperatures. They were linked to a PC via a National Instruments NI 9213 thermocouple input device. Data from the thermocouples and the dynamometer were collected and transferred to Labview prior to the analysis.

The internally cooled tool was comprised of the tool shank, a cooling adaptor and a hollow insert, as shown in Figure 1. The tool shank was an off the shelf model manufactured by Sandvik (CSBNR 2525M 12-4) which had been enhanced with designed fluid channels machined inside it. The adaptor block has been custom machined in mild steel. The cutting inserts were once again an off the shelf -item produced by Hertel (SNUN 120408, Tungsten Carbide WC with 6 % Cobalt). These were modified using electro discharge machining to create a hollow with a 1 mmwall thickness. The coolant was flowing from a central reservoir which contains approximately one litre of coolant. From here it flowed through silicone tubing to a micro-diaphragm pump from KNF-Neuberger (NFB 60 DCB). Upon exiting the pump, the coolant then flowed to and around the part of the circuit enclosed within the tool and finally back to the reservoir.

The volume flow rate of the coolant (Q) was approximately 0.3 L/min for all the tests, i.e. in SI units $Q = 0.3/60\,000$ m^3/s. The coolant was a 25 % in volume liquid solution of Ethylene Glycol in water. The specific heat (C_p) and the density (ρ) of the coolant were considered essentially constant and approximately equal to 3850 J/kg K and 1040 kg/m^3, respectively. The choice of using a 25 % Ethylene Glycol aqueous solution rather than water was conservatively made to benefit from the ebullioscopic elevation of the boiling point of the mixture. A bi-phase vapor-liquid flow within the internal microfluidics structures is in this way slightly less likely to take place. This choice however adversely affects the efficiency of the cooling system. For the same volume flow rate and for the same increment of temperature, a coolant comprised of the Ethylene Glycol solution would exchange heating power with the insert less than water would do. In the range of the tested experimental conditions, clean water has in fact comparable density but higher specific heat than the mixture used (approximately $\rho_{water} = 1000$ kg/m^3 and $C_{p,water} = 4184$ J/kg K , albeit they both are not constant).

3. Design of the Experiment

The temperature of the coolant at the inlet (T_{in}) and at the outlet (T_{out}) of the insert, together with the cutting and the thrust forces (F_c and F_t, respectively) were measured in a set of experimental conditions defined by three technological variables: the depth of cut (a_p), the feed rate (f) and the cutting speed (v_c). These variables assume numerical values. They have been therefore considered as continuous rather than categorical variables.

Each variable was assigned three values (Table 1). Thus a limited region was identified in the space (a_p, f, v_c).

[Table 1 about here.]

Cutting trials were performed in the resulting 3^3 experimental conditions (treatments). In each treatment, the cutting test was replicated three times. thus the total number of tests accrued to 81. A unique label was given to each treatment. Then, a permutation of the 27 labels was randomly generated out of 27! possible label permutations. The treatments were run in the order defined by such a permutation. All the three cutting trials for a given treatment were performed in the same machine set-up. A full randomisation of the cutting tests would have requested a new machine set-up (different or equal to the latest) for each single cutting test. The set-up time of the machine made a full randomisation of the 81 tests impracticable.

4. Modelling and Analysis

The thermal power exchanged between the coolant and the insert during machining ($\dot{Q}$) causes the temperature of the coolant at the insert outlet (T_{out}) to be higher than at the insert inlet (T_{in}, which is approximately equal to the ambient temperature). By the application of the first law of thermodynamics to the open system made of the coolant flowing in the microfluidic structures within the insert, the following equation is derived for the steady state:

$$\dot{Q} = Q\,\rho\,C_p\,(T_{out} - T_{in}) \tag{1}$$

From the measurements of the cutting force (F_c) and the thrust force (F_t), the cutting power $P_c = F_c\ (v_c/60)$ and the thrust power $P_t = F_t\ (f/1000)\ (n/60)$ were calculated. In these expressions, n denotes the angular speed of the blank in revolutions per minute, whereas the other coefficients have been introduced to express the power in watt. To explore the relationship between the efficiency of the internally-cooled tool and the machining conditions, a definition of efficiency ratio r is introduced as follows:

$$r = 100\,\frac{\dot{Q}}{P_c + P_t} \tag{2}$$

In equation (2), the efficiency ratio r represents the percentage of the power needed to remove material from the blank that is thermally transferred by the flow of the internal coolant. Alternatively, r can be described as the scaled ratio of the heat transfer rate associated with the flow of the coolant and the mechanical power used at the tool to remove material from the workpiece. In other words, The coefficient r does not represent some measurement of efficiency of the cutting process, but a measurement of efficiency of the internal cooling system. The idea behind this approach is that the internal cooling apparatus is more efficient the more thermal power it can remove from the system tool/chip/workpiece per unit of power in input to such a system, regardless of how this input power is then distributed between the workpiece, the chip and the tool. In this view, the efficiency of a machine tool in converting electrical power into mechanical power available at the tool is also not relevant.

A specific efficiency ratio r' is also introduced as follows:

$$P_s = \frac{P_c + P_t}{(a_p/1000)\,(f/1000)\,(v_c/60)} \tag{3}$$

$$r' = 100\,\frac{\dot{Q}/Q}{P_s} \tag{4}$$

The numerical coefficients in Equation (3) were introduced to convert the measured technological variables to the SI units (m, m/rev and m/s). In equation (4), the ratio *r'* represents the percentage of the total machining power per unit of volume (m^3) of material removed from the blank in the unit of time (s) that is thermally dissipated by a unit of volume flow rate (m^3/s) of the coolant. Both the dimensionless ratios r and *r'* have been considered as two response variables separately analysed. The measuring procedure for r and *r'* is the same for all the treatments.

Improving the efficiency merit by increasing the coolant mass flow rate ($\dot{m} = Q\rho$), by identifying more efficient coolant fluids (with higher C_p), by refrigerating the coolant (i.e. reducing T_{in} in Equation (1)) are all actions that can be thought of, but that were not within the scope of this study. Hence such actions were not taken. For example, the usage of cryogenic media such as nitrogen and carbon dioxide has been reported in other cooling systems such as high pressure jet cooling systems (cf page 311 – 338 in [19]). Opposite to the internally-cooled tool presented in this investigation, in those systems the cryogenic coolant is a consumable: it evaporates rather than being re-circulated in a closed-loop.

The parameters involved in the construction of a statistical model may have desirable statistical proprieties if the independent variables are centred around zero. Typically, intercepts and slopes are more likely to be uncorrelated if the independent variables are centred (cf. for example Pinheiro and Bates [20], page 34). Also, dimensionless independent variables facilitate the transformation of the data, which is often necessary in the construction of a model. For these reasons, dimensionless, centred, independent variables were defined as follows:

$$a'_p = 100\,\frac{a_p - 0.35}{0.35}; \quad f' = 100\,\frac{f - 0.15}{0.15}; \quad v'_c = 100\,\frac{v_c - 300}{300} \tag{5}$$

The equations (5) define the per cent deviations from the central point $(a_{p,c},\, f_c,\, v_{c,c}) = (0.35\,\text{mm},\, 0.15\,\text{mm/rev},\, 300\,\text{m/min})$, which is the centre of the investigated region in the space of the technological variables.

[Figure 2 about here.]

The diagram of the ratio r versus a'_p, f' and v'_c is displayed in Figure 2. The abscissae of the data have been increased by a random amount to avoid overlapping points and thus increasing the readability of the figure (a procedure called jittering). In the same figure the sample mean of the data for each value of the pertinent independent variable has been designated by a cross. A qualitative visual analysis of Figure 2 raises the suspicion that the dimensionless depth of cut a'_p does not significantly affect the efficiency ratio r, whereas the dimensionless feed rate f' and the dimensionless cutting speed v'_c may do. When either f' or v'_c increases the efficiency ratio r' appears to deteriorate. Also, the variability of r may be significantly inflated at high a'_p, low f' and low v'_c. Interaction plots (not shown here for brevity) were also constructed but they did not exhibit any pattern either strongly pointing to or strongly ruling out any significant second order interaction.

Running the experiment in 27 experimental units (alias blocks), each coincident with a treatment, suggests introducing a random effect in the model to account for physical events or circumstances that may lurk within an experimental unit while the tests are performed. For example, the portion of the blank being machined in an experimental unit may have micro-structural and mechanical proprieties slightly different from those of other experimental units. Without the introduction of a random effect, the likely effect of these properties on the measured response would then be unduly attributed in part to the independent variables.

A preliminary tentative model of the experimental data is as follows:

$$r_{ijkl} = \beta_0+\beta_1\, a'_{p,i}+\beta_2\, f'_j+\beta_3\, v'_{c,k}+\beta_4\, a'_{p,i}\, f'_j+\beta_5\, a'_{p,i}\, v'_{c,k}+\beta_6\, f'_j\, v'_{c,k}+\beta_7\, a'_{p,i}\, f'_j\, v'_{c,k}+b_{ijk}+\varepsilon_{ijkl} \tag{6}$$

where the subscripts $i = 1, \ldots, 3$, $j = 1, \ldots, 3$, $k = 1, \ldots, 3$ and $l = 1, \ldots, 3$ represent the different depths, feed rates, cutting speeds and replications of the tests, respectively. The β's are eight unknown parameters of the model, b_{ijk}'s are the 27 non-observable random variables associated with the corresponding experimental units, ε_{ijkl} are the 81 non-observable random variables that model the random error. It is then assumed that all the random variables in equation (6) are independent, identically distributed and normal with constant variance, namely: $b_{ijk} \sim N(0, \sigma_b^2)$, $\varepsilon_{ijkl} \sim N(0, \sigma^2)$, where the standard deviations σ_b and σ are two further unknown parameters of the model. Under these assumptions, the ten model parameters are estimated using the maximum likelihood method (ML) as implemented in the library `nlme` [21, 20] of R, a free language and run-time environment for statistical computing and graphics [22]. The significance of the terms associated to the technological variables that enter Equation (6) by the β's has been tested sequentially in the order they appear in the model and conditionally on the estimate of σ_b (cf. Pinheiro and Bates [20], 89-92). A term is added in the model only if such an inclusion reduces significantly the variability of the predicted errors. The test was performed using the `anova()` function of the `nlme` library.

[Table 2 about here.]

The results of the tests displayed in Table 2 support the conclusion that Equation 6 does not fit the data any better than the following simpler model equation, which is thus to be preferred:

$$r_{ijkl} = \beta_0 + \beta_2\, f'_j + \beta_3\, v'_{c,k} + b_{ijk} + \beta_6\, f'_j\, v'_{c,k} + \varepsilon_{ijkl} \tag{7}$$

The library `nlme` allows the experimenters to predict the observed response values by the fitted model, both at population level, i.e. $\hat{\mathrm{E}}\,[r_{ijkl}] = \hat{\mathrm{E}}\,[r_{ij}] = \hat{\beta}_0 + \hat{\beta}_2\, f'_j + \hat{\beta}_3\, v'_{c,k} + \hat{\beta}_6\, f'_j\, v'_{c,k}$ and at experimental unit level, i.e. $\tilde{\mathrm{E}}\,[r_{ijkl}|b_{ijk}] = \tilde{\mathrm{E}}\,[r_{ijk}|b_{ijk}] = \hat{\beta}_0 + \hat{\beta}_2\, f'_j + \hat{\beta}_3\, v'_{c,k} + \hat{\beta}_6\, f'_j\, v'_{c,k} + \tilde{b}_{ijk}$ (with $\mathrm{E}[X]$ designating the expected value of X, $\hat{\alpha}$ the estimate of the parameter α and $\tilde{X}$, the predictor of the random variable X). In this second case, the best linear unbiased predictors $\tilde{b}_{ijk}$ of the random effects are also calculated (*BLUEs*, cf. Pinheiro and Bates [20], 94). In turn, predictions of the non-observable errors can thus be computed and are usually referred to as residuals, namely: $\tilde{\varepsilon}_{ijkl} = r_{ijkl} - \tilde{\mathrm{E}}\,[r_{ijk}|b_{ijk}]$. Departures from the hypotheses underlying the model are diagnosed by the graphical analysis of the residuals.

[Figure 3 about here.]

In part (a) of Figure 3 the dispersion of the residuals around the zero appears to increase with the values fitted by the model of Equation (7). Such an observation is inconsistent with the assumed equal variance of the errors (σ^2). To overcome the violation of this hypothesis, the response is logarithmically transformed in the following new model:

$$\log(r_{ijkl}) = \beta_0 + \beta_2\, f'_j + \beta_3\, v'_{c,k} + \beta_6\, f'_j\, v'_{c,k} + b_{ijk} + \varepsilon_{ijkl} \tag{8}$$

An equivalent representation of equation (8) is given by its multiplicative form:

$$r_{ijkl} = e^{\beta_0}\, e^{\beta_2\, f'_j}\, e^{\beta_3\, v'_{c,k}}\, e^{\beta_6\, f'_j\, v_{c,k}}\, e^{b_{ijk}}\, e^{\varepsilon_{ijkl}} \tag{9}$$

More details regarding suitable transformations of the response to overcome observed departures of the assumed homoscedasticity of the errors in the case of linear models are presented by Faraway (cf pages 53–58 in [23]). The parameters in Equation (8) and (9) have been estimated as in the previous cases using the `nlme` library (Table 3). The adequacy of the fitted model has been assessed with the Akaike Information Criterion (AIC), formally defined by $AIC = -2\,\text{logLik} + 2\,n_{par}$, where logLik $= 29.70$ is the log-Likelihood of the fitted model (i.e. the maximum log-Likelihood) and $n_{par} = 6$ is the number of parameters estimated in the model, thus $AIC = -47.39$ (cf Pinheiro and Bates [20], pages 10, 83, 84).

[Table 3 about here.]

In part (b) of Figure 3 the residuals of the model involving the log-transformed efficiency ratio r appear to have a dispersion around zero that is markedly less dependant on the fitted values than in the original model with untransformed response (part (a) of Figure 3). Also, two residuals labelled '66 a' and '66 c' in part (b) of the same figure are noticeably lying quite far apart from the majority of the others. The two labels indicate that these two residuals have been obtained as the first and third replicate of the treatment 66, which corresponds to $a_p = 0.5$ mm, $f = 0.1$ mm/rev and $v_c = 300$ m/min ($a'_p = 42.86$, $f' = -33.34$, $v'_c = 0$). No specific reason has been identified for the two associated experimental results to cause this outlying situation. Thus there was no reason for excluding the two experimental results from the analysis. Moreover, even doing so, the resulting fitted model did not lead to significantly different estimates of the parameters. Namely, the confidence intervals for corresponding parameters in the two models were overlapping. The fact that these two residuals were obtained in the same experimental unit instils the suspicion that the uncontrollable unknown reason causing the outlying of the two residuals may be related to the specific experimental unit. In this sense, the two outlying residuals reinforce the motivations for introducing the random effects b_{ijk} in the model of the experimental results. Without random effects as in the following model equation:

$$\log(r_{jk}) = \beta_0 + \beta_2\, f'_j + \beta_3\, v'_{c,k} + \beta_6\, f'_j\, v'_{c,k} + \varepsilon_{jk} \tag{10}$$

the residuals appear inconsistent with the assumption of errors (ε_{ijkl}) characterised by zero mean and equal variance.

[Figure 4 about here.]

In Figure 4, when the random effects b_{ijk} are part of the model (cf. part (a) of the figure), the three residuals corresponding to each experimental condition (treatment) have a sample mean that is close to zero. Otherwise, they have not (cf. part (b) of the figure). The deviation of such a sample mean from zero is what the random effect of a treatment is specifically meant to account for. Moreover, in part (b) of the figure, 15 of these sample means are negative, whereas 12 are positive. This symmetry in the distribution of the realised random effects is consistent with the assumed normality of the random effects. Q-Q plots have also been constructed and did not contradict dramatically the assumed normality of both residuals and random effects for the model of Equation (8). The figures were not included for sake of brevity. In addition, in Figure 4 the dispersion of the realised residuals around their mean is visibly smaller when the random effects are included in the model (part (a) of the figure). All these qualitative observations have been substantiated by testing the hypothesis $\sigma_b = 0$. Under the not-disproved assumption of normality of both random effects and errors, a likelihood ratio test was conducted using a Monte Carlo approach. A short script was implemented in R to obtain an empirical distribution of the test statistics. 50 000 realisations of the test statistics were simulated in pseudo-random numerical tests. The p-value obtained was less than 0.00002 and led therefore to reject the hypothesis $\sigma_b = 0$.

The values of the specific efficiency ratio *r'* versus the dimensionless technological variables a'_p, f' and v'_c are displayed in Figure 5. From the observation of this figure, there is some strong suspicion that the specific efficiency ratio *r'* increases substantially with the dimensionless depth of cut. Possibly, also increments of the dimensionless feed rate may moderately improve *r'*, whereas the dimensionless speed of cut appears as hardly having any effect on *r'*. In Figure 5 it can also be noticed that increasing the dimensionless depth of cut a'_p appears to inflate the dispersion of the *r'* values around their a'_p mean.

[Figure 5 about here.]

A quantitative analysis confirmed these initial intuitions. By following the same methods and procedures as in the case of the efficiency ratio r, Such an analysis ultimately led to the following model equation:

$$\log(r'_{ijkl}) = \beta_0 + \beta_1\, a'_{p,i} + \beta_2\, f'_j + b_{ijk} + \varepsilon_{ijkl} \tag{11}$$

The ML estimates of the parameters for the model in Equation (11) are displayed in Table 4. The corresponding AIC is -45.28, the maximum log-Likelihood is 27.64 and $n_{par} = 5$.

[Table 4 about here.]

5. Discussion

The fixed effects part of the model of Equation (8) and (9) allows predictions to be made regarding the typical efficiency ratio $\hat{\mathrm{E}}\,[r_{ijkl}]$ when the technological variables are set within the experimental region investigated. Figure 6 provides an operational graphical representation of this model to assist its interpretation.

[Figure 6 about here.]

In such a figure, the yellow or light-grey transparent area respectively in colour and black-and-white print represents the region of the technological parameters experimentally explored. For any dimensionless feed rate in that area, increasing the cutting speed deteriorates the expected efficiency r. The maximum expected efficiency ratio in the area is 10.96 % and is obtained at the minimum feed rate and minimum cutting speed investigated (point A, at the corner of the yellow/light-grey region in Figure 6). The variable v_c' enters the model with a coefficient that is approximately the double in absolute value of that associated with f' ($\hat{\beta}_3/\hat{\beta}_2 \cong 2$). This supports the idea that the efficiency ratio r is more sensitive to per cent variations in cutting speed rather than in feed rate. The positive interaction coefficient ($\hat{\beta}_6$) is about one fifth of that of f' and one tenth of that of v_c' (both taken in absolute value). Hence for positive f' the degree of sensitivity of the expected efficiency ratio r to v_c' is slightly less than what implied by $\hat{\beta}_3$ alone. In the experimental region investigated, however, this sensitivity to v_c' is always larger than that to f'. When both f' and v_c' are positive or both are negative, the increment in efficiency ratio obtained by reducing both f' and v_c' is less than the sum of the increments that can be obtained by reducing f' and v_c' separately. The situation is reversed when f' and v_c' are of opposite sign. Any statement based on the extrapolation of the model outside of the experimental region investigated needs per se further experimental campaigns to be substantiated. However, an examination of the behaviour of the model outside the region investigated experimentally (the yellow/light-grey highlighted area in Figure 6) may assist the planning of future experiments. In Figure 6, it is observed that when considering $f' < -33.333$ the sensitivity of the expected efficiency to the cutting speed is increased greatly. When instead $33.333 < f' < 62.798$, increments in cutting speed still decrease the efficiency, but less and less. The value $\bar{f}' = -\hat{\beta}_3/\hat{\beta}_6 = 62.798$ is where any v_c' is expected to be equally efficient, namely $\bar{r} = e^{\hat{\beta}_0 - \frac{\hat{\beta}_2 \hat{\beta}_3}{\hat{\beta}_6}} = 5.1122$. For values $f' > 62.798$ the expected efficiency ratio

r is increasing and no longer descreasing with v_c'. The effect of v_c' on the efficiency ratio is reversed because of the interation term in the model. The point B in Figure 6 is the stationary saddle point of the model.

The above analysis indicates that in the investigated area and likely in large areas beyond it (up to $f' < 62.798$), the cooling system is more efficient, the smaller the cutting speed and the feed rate are. Hence, the cooling system is more efficient the smaller the mechanical power needed for the machining operation is. A decrease in machining power is accompanied with a less than proportional decrease in power dissipated by the cooling system.

Opposite to the case of the efficiency ratio r, the expected values of the specific efficiency ratio r' synthesised in Equation (11) do not exhibit any dependence on the cutting speed v_c'. They do however display a dependence on the depth of cut a_p' which does not exist for the ratio r. In contrast with the ratio r, the log-transformed specific efficiency r' does appear to be linear in the significant independent variables. Otherwise stated, there is no significant interaction between the two independent variables.

The model of Equation (11) shows that a unit volume flow rate of coolant dissipates more thermal power out of the mechanical power needed to generate a unit volume flow rate of chip when the depth of cut and the feed rate are larger. This conclusion seems consistent with the intuition that when the contact tool-workpiece is larger the thermal exchange between workpiece and tool is facilitated. Therefore more power can be dissipated into the tool and then into the cooling system. Large depths of cut and large feed rates increase the theoretical cross section of the chip (i.e. the cross section prior to actual removal of the chip from the part). So therefore they do increase the contact region tool-workpiece. The expected specific efficiency *r'* is sensitive to variations of depth of cut approximately twice as much it is to variations of feed rate ($\hat{\beta}_1/\hat{\beta}_2 \cong 2$). Whereas the depth of cut does not have any significant effect on the efficiency *r*, increasing it appears to improve the specific efficiency *r'*.

6. Conclusions

Microfluidic structures internal to the tool have been designed and manufactured to convey the flow of coolant in the near proximity of the cutter edge. The part and the coolant never enter in contact. Contamination of the part by molecules of the coolant is thus prevented.

The designed and manufactured internally-cooled tool system enabled heat transfer from the cutting zone of the insert to the flow of the liquid coolant. Measurements of cutting force, thrust force, coolant temperature at the inlet and at the outlet of the tool system were taken in a 3^3 experimental conditions defined by the depth of cut, the feed rate and the cutting speed. Each condition was replicated three times.

An efficiency ratio *r* and a specific efficiency ratio *r'* were respectively defined as the percentage of the whole machining power that is transferred to the coolant and as the percentage of machining power per volumetric flow rate of material removed that is transferred to a unit volume flow rate of the coolant.

Linear mixed-effects statistical models were fitted to the experimental results using the maximum likelihood method. The analysis revealed that the efficiency ratio r depends exponentially on the cutting speed and on the feed rate, whereas it does not depend on the depth of cut. Within the investigated experimental region, the less the cutting speed and the feed rate are, the higher the expected efficiency ratios r are. The maximum expected efficiency is therefore obtained at $f_{min} = 0.10$ mm/rev and $v_{c,min} = 250$ m/min and is equal to 10.96 %. A significant interaction effect of cutting speed and feed rate on the efficiency ratio r was also identified. The specific efficiency ratio r' was instead found exponentially depending on the depth of cut and the feed rate with no significant interaction effect. In other words, the $\log(r')$ was found to be linearly increasing with the depth of cut and the feed rate.

Acknowledgements

This study is dedicated to the memory of Gualberto Ricci Curbastro for no small amount of personal inspiration. The investigation was performed within the scope of the collaborative research project 'Self-learning control of tool temperature in cutting processes'(ConTemp) funded by the European Commission 7th Framework Programme (Contract number: NMP2-SL-2009-228585). The authors gratefully acknowledge the committed support of all the technical staff in the AMEE Department at Brunel University. Particular gratitude goes to Mr Paul Yates for his help in the cutting trials.

[1] V. P. Astakhov, The assessment of cutting tool wear, International Journal of Machine Tools and Manufacture 44 (6) (2004) 637 – 647.

[2] J. M. Longbottom, J. D. Lanham, Cutting temperature measurement while machining a review, Aircraft Engineering and Aerospace Technology 77 (2005) 122–130.

[3] E. M. Trent, P. K. Wright, Chapter 5 - heat in metal cutting, in: Metal Cutting (Fourth Edition), fourth edition Edition, Butterworth-Heinemann, Woburn, 2000, pp. 97 – 131.

[4] Y. Quan, Z. He, Y. Dou, Cutting heat dissipation in high-speed machining of steel based on the calorimetric method, Frontiers of Mechanical Engineering in China 3 (2) (2008) 175–179.

[5] A. Basti, T. Obikawa, J. Shinozuka, Tools with built-in thin film thermocouple sensors for monitoring cutting temperature, International Journal of Machine Tools and Manufacture 47 (5) (2007) 793 – 798.

[6] G. Cohen, P. Gilles, S. Segonds, M. Mousseigne, P. Lagarrigue, Thermal and mechanical modeling during dry turning operations, The International Journal of Advanced Manufacturing Technology 58 (1-4) (2012) 133–140.

[7] H. Young, T. Chou, Investigation of edge effect from the chip-back temperature using IR thermographic techniques, Journal of Materials Processing Technology 52 (24) (1995) 213 – 224.

[8] N. Jeffries, R. Zerkle, Thermal analysis of an internally-cooled metal-cutting tool, International Journal of Machine Tool Design and Research 10 (3) (1970) 381 – 399.

[9] C. Ferri, T. Minton, S. Bin Che Ghani, K. Cheng, Internally-cooled tools and cutting temperature in contamination-free machining, Proceedings of the Institution of Mechanical Engineers, part C: Journal of Mechanical Engineering Science(accepted for publication. OnlineFirst, 13 March 2013). doi:10.1177/0954406213480312.

[10] J. C. Rozzi, W. Chen, E. E. J. Archibald, Indirect cooling of a cutting tool, USA Patent US 8,061,241 B2 (2011).

[11] L. E. A. Sanchez, V. L. Scalon, G. G. C. Abreu, Cleaner machining through a toolholder with internal cooling, in: Proceedings of the 3rd International Workshop Advances in cleaner production, Vol. 3, San Paulo, 2011.

[12] L. Liang, Y. Quan, Z. Ke, Investigation of tool-chip interface temperature in dry turning assisted by heat pipe cooling, The International Journal of Advanced Manufacturing Technology 54 (1-4) (2011) 35–43.

[13] S. Shu, S. Chen, K. Cheng, Investigation of a novel green internal cooling in turning application, in: Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 International Conference on, Vol. 3, 2011, pp. 1156–1159.

[14] E. Uhlmann, P. Fürstmann, M. Roeder, S. Richarz, F. Sammler, Tool wear beaviour of internally cooled tools at different cooling liquid temperatures, in: Proceedings of the 10^{th} Global Conference on Sustainable Manufacturing, Istanbul, Turkey, 2012.

[15] G. F. Micheletti, Tecnologia Meccanica - Il taglio dei metalli, Vol. 1, UTET, 1977, (in Italian).

[16] G. Boothroyd, Temperatures in orthogonal metal cutting, Proceedings of the Institution of Mechanical Engineers 177 (29) (1963) 789–810.

[17] E. M. Trent, P. K. Wright, Chapter 9 - machinability, in: Metal Cutting (Fourth Edition), fourth edition Edition, Butterworth-Heinemann, Woburn, 2000, pp. 251 – 310.

[18] J. Kopac, M. Sokovic, S. Dolinsek, Tribology of coated tools in conventional and {HSC} machining, Journal of Materials Processing Technology 118 (13) (2001) 377 – 384.

[19] E. M. Trent, P. K. Wright, Metal Cutting, 4th Edition, Butterworth-Heinemann, Oxford, 2000.

[20] J. C. Pinheiro, D. M. Bates, Mixed-Effects Models in S and S-Plus, Springer, 2000.

[21] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Development Core Team, nlme: Linear and Nonlinear Mixed Effects Models, r package version 3.1-103 (2012).

[22] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2011).

[23] J. J. Faraway, Linear Models with R, 1st Edition, Texts in Statistical Science, Chapman & Hall/CRC, 2004.

List of Figures

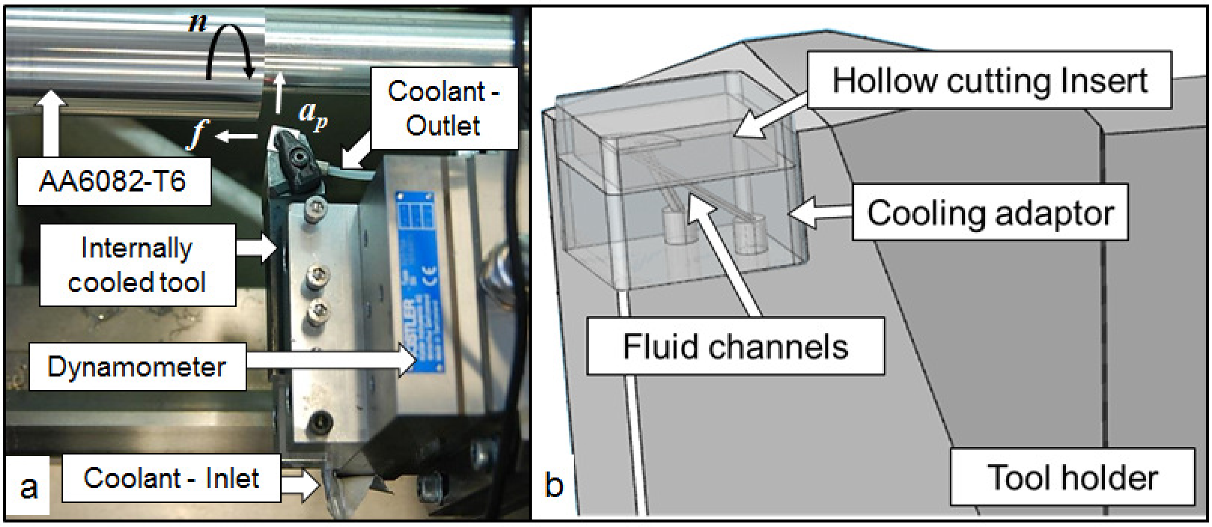

Figure 1: (a) The experimental set-up for the cutting trials. (b) A 3-D model of the assembled internally cooled tool system.

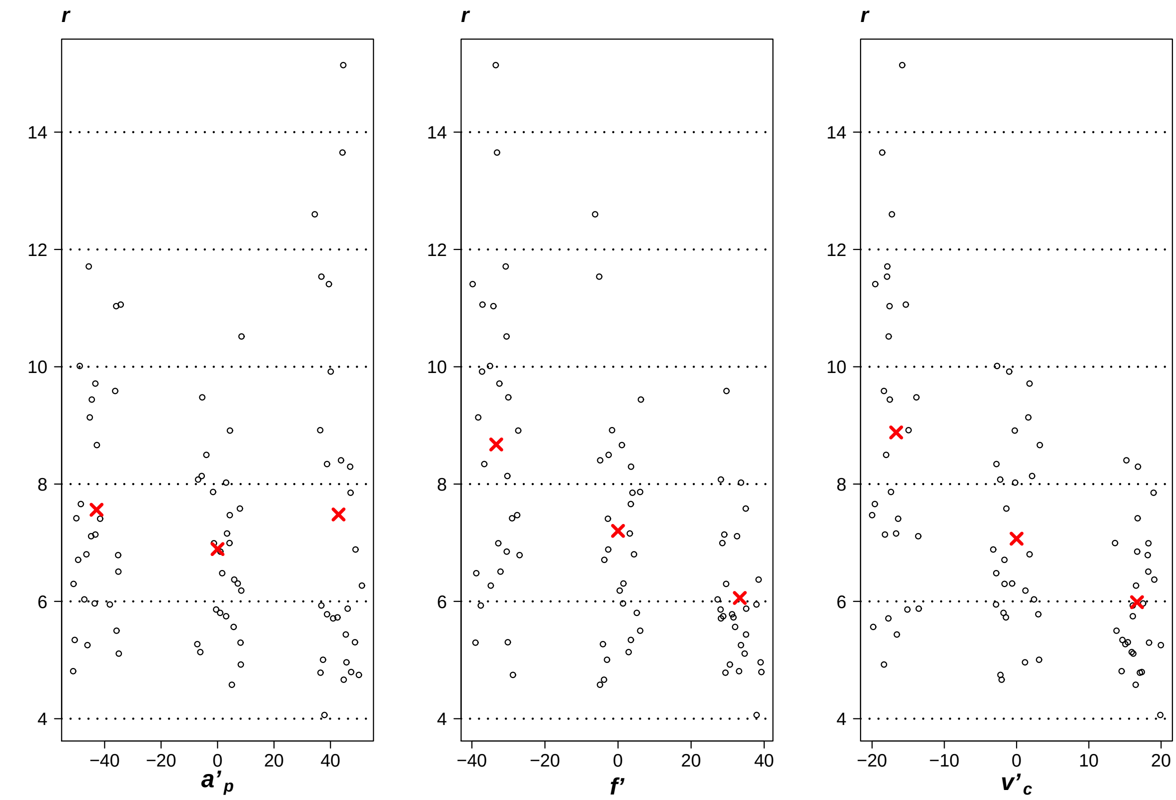

Figure 2: The ratio r versus the dimensionless technological variables a'_p, f' and v'_c. The abscissae have been jittered. The cross designates the sample mean for each of the three groups of data in each panel.

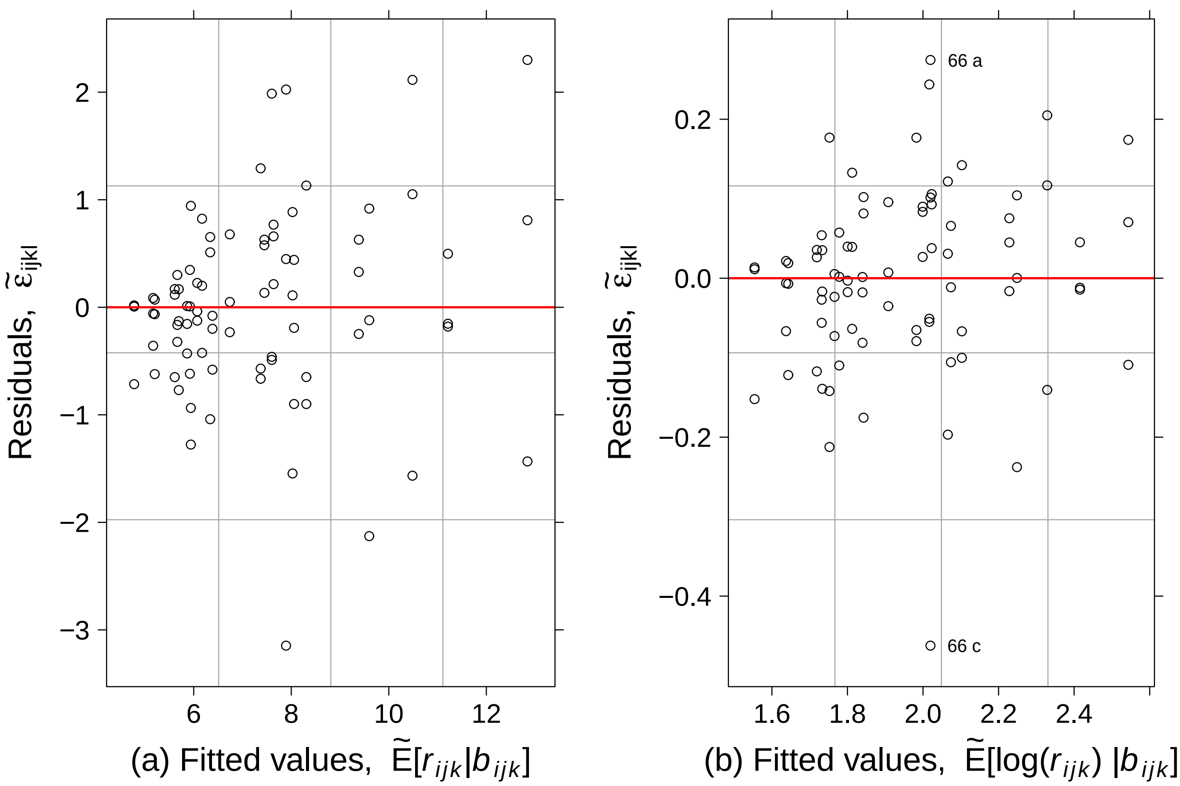

Figure 3: Residuals versus fitted values (alias predicted values) (a) of the simplified model (Equation (7)) (b) of the log-transformed simplified model (Equation 8 or 9).

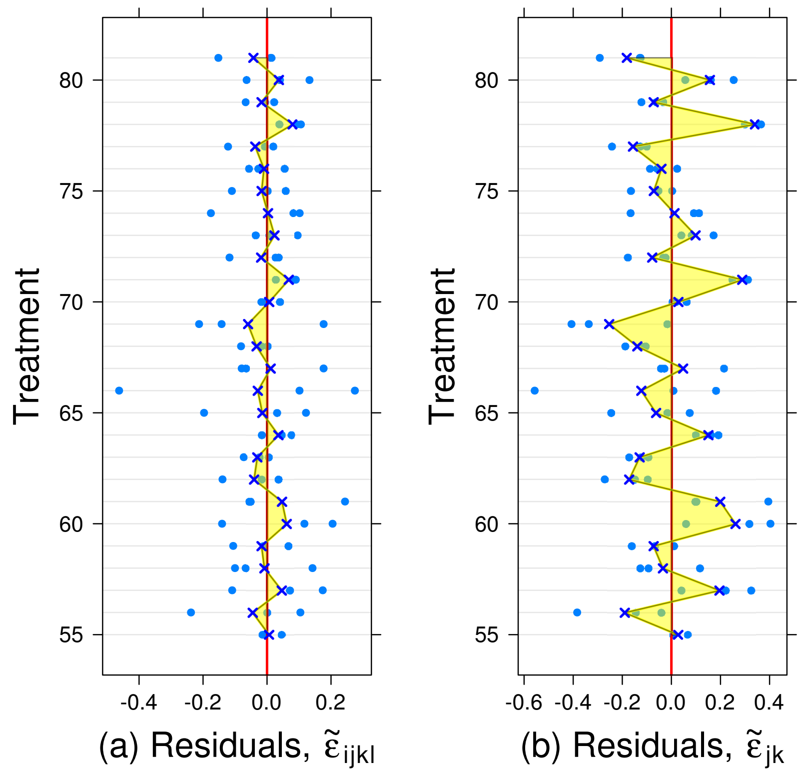

Figure 4: The realisations of the non-observable residuals versus the values fitted by (a) the model that includes the random effects (Equation 8) (b) the model that does not include the random effects (Equation 10). The average for each treatment is identified by the points 'X'. The shadow area underlying the segments joining these averages facilitate the visualisation of the different amount of violation of the assumed zero mean for errors of the two models.

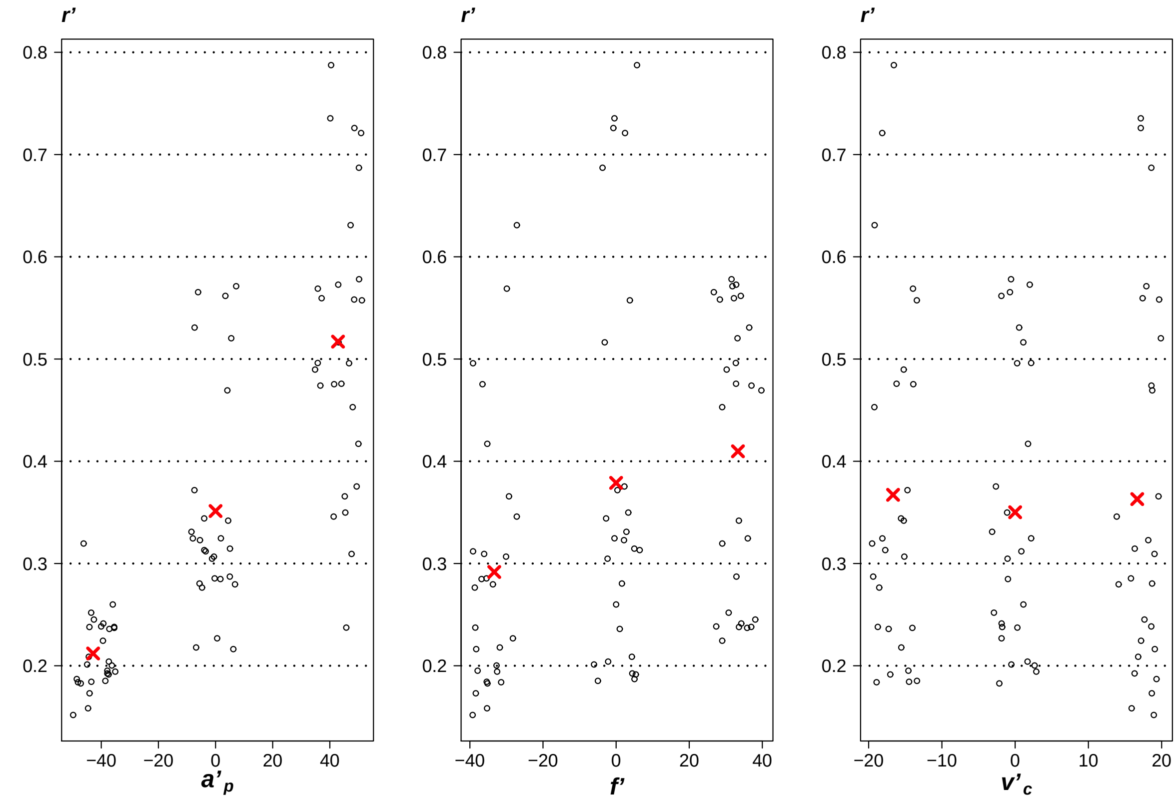

Figure 5: The ratio r' versus the dimensionless technological variables a'_p, f' and v'_c. The abscissae have been jittered. The cross designates the sample mean for each of the three groups of data in each panel.

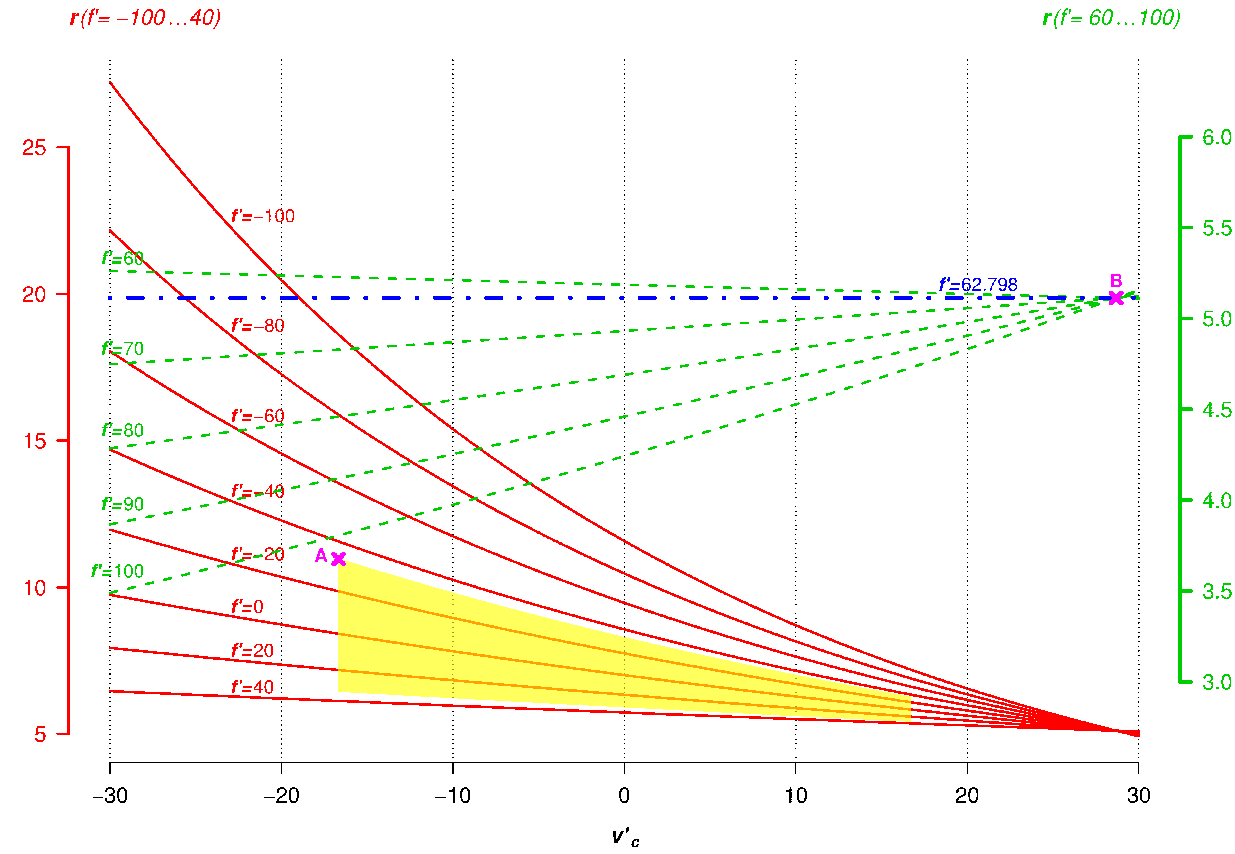

Figure 6: The expected value of the efficiency ratio r versus the dimensionless cutting speed v_c' for selected f' values ranging from -100 to 100. The highlighted area in yellow/light-grey in colour/black-and-white print represents the experimental region investigated. The point A identifies the maximum efficiency in that area. The point B identifies the stationary saddle point. The thicker dotted and dashed line is horizontal (iso-efficient line).

List of Tables

Table 1: Technological variables and their values.

Variable	Unit	values
depth of cut, a_p	mm	0.20, 0.35 and 0.50
feed rate, f	mm/rev	0.10, 0.15 and 0.20
cutting speed, v_c	m/min	250, 300 and 350

	numDF	denDF	F-value	p-value
(Intercept)	1	54	995.5	<0.0001
a'_p	1	19	0.01980	0.8896
f'	1	19	21.25	0.0002
v'_c	1	19	25.99	0.0001
$a'_p\, f'$	1	19	0.3581	0.5566
$a'_p\, v'_c$	1	19	0.1678	0.6866
$f'\, v'_c$	1	19	7.753	0.0118
$a'_p\, f'\, v'_c$	1	19	2.1257	0.1612

Table 2: Sequential tests of the hypotheses for the significance of the independent variables and their interactions both listed in the first column. 'numDF' and 'denDF' are the numerator and denominator degrees of freedom, respectively. The p-values are expressed in fractions of the unity rather than in per cent.

Parameter	ML estimate	standard error
β_0	1.947	0.03105
β_2	-0.005020	0.001141
β_3	-0.01099	0.002282
β_6	0.0001750	0.00008384
σ_b	0.1378	
σ	0.1316	

Table 3: ML estimates of the parameters for the model with Equation 8 or 9. For the estimators of the β's the standard errors are also shown.

Parameter	ML estimate	standard error
β_0	-1.116	0.03329
β_1	0.01012	0.0009513
β_2	0.005377	0.001223
σ_b	0.1518	
σ	0.1316	

Table 4: ML estimates of the parameters for the model with Equation 11. For the estimators of the β's the standard errors are also shown.