Towards a generalized energy prediction model for machine tools

Raunak Bhinge[a], Jinkyoo Park[b], Kincho H. Law[b], David A. Dornfeld[a], Moneer Helu[c] and Sudarsan Rachuri[d]

[a]Laboratory for Manufacturing and Sustainability, University of California, Berkeley, CA, USA
[b]Engineering Informatics Group, Stanford University, Stanford, CA, USA
[c]Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA
[d]Advanced Manufacturing Office, Office of Energy Efficiency and Renewable Energy (EERE), Department of Energy, Washington, DC, USA

Abstract

Energy prediction of machine tools can deliver many advantages to a manufacturing enterprise, ranging from energy-efficient process planning to machine tool monitoring. Physics-based, energy prediction models have been proposed in the past to understand the energy usage pattern of a machine tool. However, uncertainties in both the machine and the operating environment make it difficult to predict the energy consumption of the target machine reliably. Taking advantage of the opportunity to collect extensive, contextual, energy-consumption data, we discuss a data-driven approach to develop an energy prediction model of a machine tool in this paper. First, we present a methodology that can efficiently and effectively collect and process data extracted from a machine tool and its sensors. We then present a data-driven model that can be used to predict the energy consumption of the machine tool for machining a generic part. Specifically, we use Gaussian Process (GP) Regression, a non-parametric machine-learning technique, to develop the prediction model. The energy prediction model is then generalized over multiple process parameters and operations. Finally, we apply this generalized model with a method to assess uncertainty intervals to predict the energy consumed to machine any part using a Mori Seiki NVD1500 machine tool. Furthermore, the same model can be used during process planning to optimize the energy-efficiency of a machining process.

Keywords : Computer-integrated manufacturing, Machining processes, Sustainable manufacturing

Email:
Raunak Bhinge : raunakbh@berkeley.edu
Jinkyoo Park : jkpark11@stanford.edu
Kincho H. Law : law@stanford.edu
David A. Dornfeld : dornfeld@berkeley.edu
Moneer Helu : moneer.helu@nist.gov
Sudarsan Rachuri : sudarsan.rachuri@hq.doe.gov

1. INTRODUCTION

Over 22% of greenhouse gas emissions in the US come from the industrial sector, which is also the highest consumer of electrical power in the US (EIA, 2014). To reduce energy use in the manufacturing sector, manufacturers need energy prediction models that can estimate electricity costs and peak power demand for their equipment based on a production plan. These models can help manufacturers reduce their energy costs and environmental footprint and respond to new regulations and business drivers. For example, the Smart Grid and carbon cap-and-trade may incentivize manufacturers to adjust their operations to respond to load adjustments in the grid or take advantage of lower energy or carbon prices during specific time windows. These models can also improve process monitoring since deviations in the power demand and energy consumption can be related to component wear, tool breakage, or collisions (Gutowski, 2006, Vijayaraghavan, 2010). The first step towards building such models is to understand the energy consumption patterns of machine tools and manufacturing operations. In this paper, we use the data collected from a machine tool to determine how different operational strategies influence the energy consumption pattern of a machine tool and to derive the most energy-efficient strategy to machine a part.

Models for predicting energy consumption and optimizing manufacturing processes have been a subject of research interest for over 50 years. Most of these efforts are physics-based, which means that the models are built upon the physical laws that govern manufacturing operations. Based on the energy transfer from an electrical system to a mechanical system, Neugebauer et al. (2007) formulated a mechatronic representation for computing the total energy consumption of a metal-cutting machine tool. Using energy conservation, Dietmair and Verl (2009) categorized and derived the energy consumption of a metal-cutting machine tool using its two basic operations: moving axes and removing material. Although these methods are based on the physics of a machine tool, they are difficult to implement because they often require a large number of physical parameters that are often hard to compute or estimate. It is also difficult to properly incorporate the stochastic nature of a manufacturing process into a physics-based model. These difficulties challenge the construction of physics-based models that account for different mechanical characteristics of different machine tools.

To address the challenges presented by physics-based models, a number of studies have explored characterizing the energy consumption of a machine tool using experimental data. Draganescu et al. (2003) used experimental data to construct statistical regression models based on machining parameters, such as the feed rate, spindle speed, and depth of cut. Diaz et al. (2009) used experimental data on face-milling operations to show that the material removal rate is one key indicator of energy consumption in a machine tool. Gutowski et al. (2006) developed machine-tool characterization techniques by studying the effects of different process parameters on the total energy consumption.

Both the physical and experimental applications in literature have been developed for specific machining operations, parameter spaces, and tool-workpiece material combinations, which limits the broad applicability of these approaches. Furthermore, they may be insufficient for machine tools with relatively high tare power demand (tare power refers to the power required for non-cutting operations and auxiliary equipment). Most modern machine tools fit this description since the energy needed for material removal is a fraction of the overall energy consumed. Most of the limitations in the literature are due to a limited access to data, lack of standardized data-collection systems, and inadequate post-processing techniques.

Advances in machine automation and sensing have begun to address such limitations by allowing continuous measurements of the operating conditions and energy consumption of a machine tool. Such advances provide new opportunities to build data-driven models to characterize a machine-tool and its performance. Teti et al. (2010) gave an extensive survey of sensor technologies, signal processing, and decision-making methodologies for machine-tool monitoring. One recent advancement is MTConnect, which is an XML-based standard that has been developed to facilitate archiving, accessing, and retrieving operational data from various manufacturing equipment (MTConnect Institute, 2014, Vijayaraghavan et al. 2008). MTConnect enables aggregation of raw power data and machining operational information, which provides a means to track variations in energy consumption by different machining operations (Vijayaraghavan and Dornfeld, 2010). MTConnect has also been used to study the effects of different process parameters on the energy consumption of a machine tool and to construct statistical regression

models for energy consumption (Diaz et al., 2011). Although these studies have clearly illustrated the possibility of collecting real-time operational and energy consumption data for future data analysis, they have so far dealt primarily with data collected from slotting operations. In practice, machining a part requires a variety of machining operations with many different combinations of operational parameters. However, the principles of data-driven process planning and machine tool monitoring based on energy consumption that come across from these studies are the key motivators of this work.

In this paper, we construct a generalized, energy prediction model for different machining operations using various combinations of process parameters. The constructed energy prediction model can be used for several tasks. First, by establishing the correlation among machining parameters and the resultant energy consumption, the energy prediction model can help gain a better understanding of the energy consumption pattern of a target machine tool. Furthermore, the energy prediction model can be used to facilitate operating a target machine more efficiently, such as reducing the total energy consumption of the target machine by selecting an energy-efficient toolpath. Finally, an energy prediction model allows monitoring of a target machine by observing sudden, unexpected events that may show deviations between the predicted and actual energy consumption, which could be an indication of failure or deterioration of a certain machine component.

This paper is organized as follows: We first describe a data-processing methodology that uses MTConnect to extract data from a machine tool controller and add-on sensors efficiently and effectively. We used this methodology to collect data from an automated milling machine tool (Mori Seiki NVD1500DCG), which allowed us to contextualize energy-consumption data with the corresponding machining operation and its process (control) parameters. To explore all possible combinations of process parameters for different machining operations, we collected data from 18-machined parts with different machining strategies (NC codes). Using this data, we developed a generalized, data-driven, energy prediction model that can determine the energy consumption of the machine tool for machining a generic part. We applied the Gaussian Process (GP) regression model, a non-parametric regression model, to model the complex input and output relationship. Finally, we illustrate the use of the energy prediction model to evaluate the optimal strategy for machining a generic part.

2. DATA COLLECTION AND POST-PROCESSING

The first step to construct an accurate energy prediction model of a machine tool is to collect and process data with minimum noise from the target machine. This data includes the process parameters collected from a wide range of machining operations, which are inputs to the model, and the corresponding energy-consumption measurements, which are the outputs of the model. However, collecting such extensive data through experimentation requires significant time and effort, which has been one key barrier for constructing data-driven energy prediction models. Recent advances in sensing and data management, such as MTConnect, have started to address these barriers by enabling the real-time collection and remote retrieval and processing of manufacturing data. This section discusses how we designed the experiments to collect the data used to construct a data-driven energy prediction model.

2.1 Data acquisition system

Figure 1 shows the overall data acquisition system used to collect energy-consumption data contextualized with machining data, such as process parameters, NC blocks, and tool positions. The machining data were collected from a FANUC controller, and the power time-series data was collected using a High Speed Power Meter (HSPM) from System Insights. The power consumption data represents the power consumed by the entire machine tool, including auxiliary components such as the cooling system and the controller. Both types of data were collected using MTConnect and synchronized and organized using an MTConnect agent. Bhinge et al (2014) and Helu et al. (2014) describe the hardware platform and data acquisition system in greater detail.

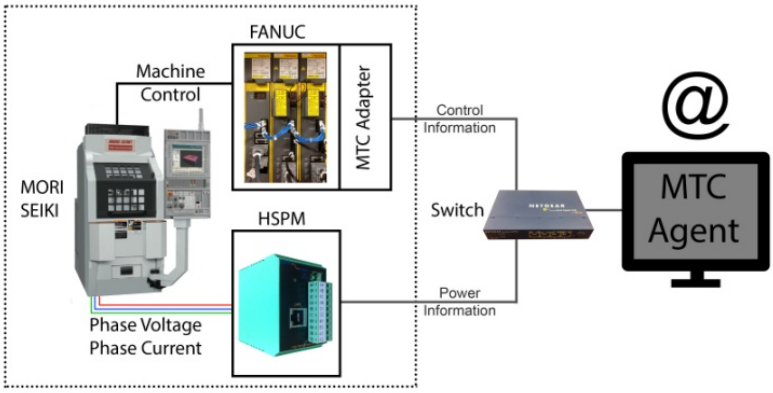


Figure 1. Data acquisition system

2.2 Data processing

To extract insights about a machine tool, the data collected from a target machine needs to be properly processed and contextualized. When constructing our energy prediction model, we used feature extraction to identify the process parameters that possibly influence energy consumption from the raw data. Figure 2 shows how we classified the data into three groups based on the level of post processing applied: direct, derived, and simulated data.

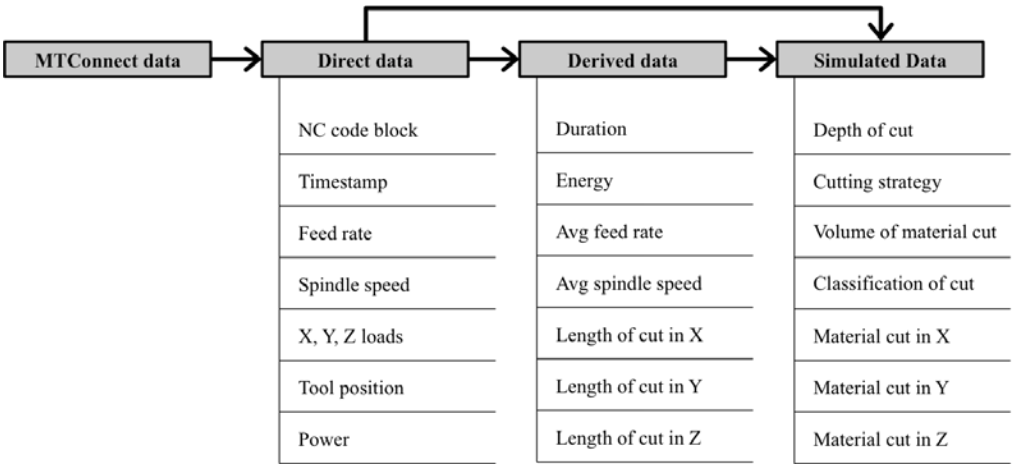


Figure 2. Categorization of types of manufacturing data obtained using MTConnect

Direct data was the raw data collected from the machine tool controller and added sensors using MTConnect. This data included the NC code block, timestamp, instantaneous feed rate, instantaneous spindle speed, instantaneous loads on each axis, instantaneous tool position, and instantaneous power. Instantaneous power was measured using an externally installed power meter. The MTConnect agent synchronized the direct data using a common time stamp.

Derived data, which is data corresponding to cutting operations, was generated by applying simple calculations to sets of direct data. For example, the machining process in a conventional automated machine tool is composed of sequences of cutting operations that can be described using a set of control parameters represented by an NC code block. To construct the energy prediction model, we needed to determine the relationship between the control parameters and the corresponding energy consumption in every NC code bock. Specifically, we computed the total energy, average feed rate, average spindle speed, and length of cut in x- and y-directions over the duration of a block of NC code corresponding to a single

cutting operation. We then used the length of cuts in the x- and y-directions to determine the length and direction of the cut. A sequence of such cuts constituted the complete toolpath.

Simulated data was generated by simulating the tool movements associated with the sequence of cutting operations for the machining process. Such data was needed because the block-averaged data for each NC code block did not provide enough detail to distinguish actual material removal operations from other tool movements (e.g., air cut above workpiece). To determine the actual amount of material removed, we applied a reverse simulation of the entire cutting process using the instantaneous position data retrieved as direct data. This simulation required knowledge of the workpiece dimensions and the tool diameter. To simulate the cutting operation, we constructed a 2-dimensional mesh on the surface of the workpiece, tracked the material removed during each cut, and redefined the elements in the mesh after every block of NC code. From the positional displacement obtained in the derived data, the toolpath of the tool for each NC code block was tracked and the material removed was calculated. The data extracted from this simulation included the depth of cut, volume of material removed, cutting strategy (i.e., climb or conventional milling), and classification of cut (e.g., air cutting, rapid motion without cutting, feed with cutting). The cutting strategy was determined from the cutting simulation by tracking the direction of angular rotation of the tool and the number of elements being cut on either side of the centerline of the tool.

2.3 Experimental design

The training output for our model was energy consumption data corresponding to each block of the NC code and its corresponding machining parameters, all of which can be collected in near-real time and efficiently processed and retrieved remotely. The experimental design and data processing technique used for generating the training data for this study have been described in previous work (Helu, et al., 2014; Bhinge et al., 2014; Park et al., 2015). In this section, we briefly present the basic setup and data processing steps used in the experiments. Figure 3 shows the sample part designed to collect training data for the data-driven energy prediction model. Table 1 shows the specific details of the workpiece, machine tool, and cutting tool used in this experimental study.

As shown in Figure 3, there were five basic cutting operations – face milling, contouring, pocketing, slotting, and plunge – that were involved in machining a part. In addition, there were three non-cutting operations – air cut in the *x*-*y* plane, air cut in the *z* direction, and rapid motion – that were also included in the experiments. Because process parameters, such as feed rate, spindle speed, and depth of cut, could have affected energy consumption, the test parts were produced using different combinations of process parameters to investigate this relationship. For the objective of this paper, a single tool and workpiece were chosen, but an expansion of this experimental setup could also involve variations in cutting tool geometry and workpiece materials. A Taguchi technique (Box et al., 1979) was employed to design the experiments to ensure a fractional-factorial combination for each set of process parameters in each operation. Table 2 shows the levels chosen for the depth of cut, chip load (feed or thickness of chip removed by one cutting edge of the tool), and spindle speed used to machine the parts. The levels were chosen to cover the entire range of prescribed milling parameters for the tool-workpiece material combination. The feed rate f (mm/min) is the product of the spindle speed (RPM), the number of tool teeth, and the chip load (mm/tooth).

As noted in Section 1, 18 parts were machined for this study, which provided a total of 196 face-milling, 108 contouring, 54 slotting and pocketing, and 32 plunge experiments. Each line of NC code corresponded to a cutting operation and tool motion and was combined with the corresponding process parameters and output energy consumption. Unlike traditional data-collection procedures, each line of NC code for a part was treated as a separate experiment. This allowed us to conduct a large number of experiments by machining a modest number of parts. The face milling operations on the first 9 parts were carried out in the y direction, and the remaining 9 parts were milled in the x direction. The separation of milling operations in the *x*- and *y*-directions was necessary to measure the energy consumption accurately for the target machine. The datasets collected from machining all 18 parts were then used to construct the energy prediction model for each (cutting or non-cutting) operation.

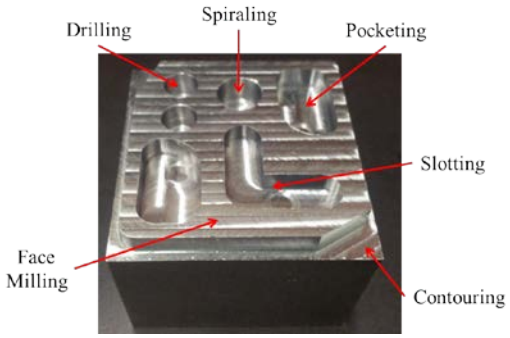


Figure 3. Test part design for experimentation.

Table 1. Experiment details (Bhinge et al., 2014).

Workpiece Material	Cold Finish Mild Steel 1018
Workpiece Dimensions	63.5mm x 63.5mm square cut to a length of 56mm
Machine Make	Mori Seiki NVD 1500
Machine Type	Micro NC Milling Machine
Tool Material	Solid Carbide
Tool Diameter	3/8" (9.525 mm)

Table 2. Experiment levels chosen for different factors.

Level	Spindle Speed (RPM)	Chip Load (mm/tooth)	Depth of Cut (mm)
1	1500	0.0254	1
2	3000	0.0330	1.5
3	4500	0.0432	3
4	6000	0.0508	-

2.4 Data used for energy prediction model

The direct, derived, and simulated data were used to construct the energy prediction model of the milling machine. There were five basic input (predictor) variables based on the fundamental parameters of a milling machine tool that affect energy consumption (see Figure 4): feed rate, spindle speed, depth of cut, cutting direction, and cutting strategy. Three of the input variables – feed rate, spindle speed and depth of cut – were quantifiable measurements defined as follows:

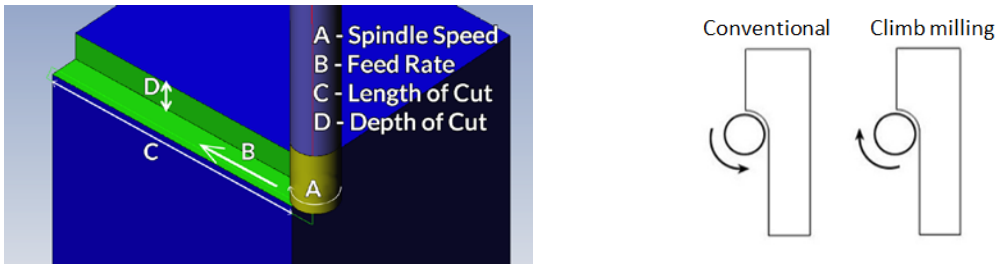
- $x_1 \in \mathbb{R}$ **Feed rate**: The average velocity at which the tool is fed, which can be retrieved from the controller data
- $x_2 \in \mathbb{R}$ **Spindle speed**: The average rotational speed of the tool, which can be retrieved from the controller data
- $x_3 \in \mathbb{R}$ **Depth of cut**: The actual depth of material that the tool is cutting, which can be obtained from the cutting simulation

The remaining input variables – cutting direction and cutting strategy – were qualitatively (or categorically) labeled. Qualitative variables were represented numerically by codes to construct a regression model. For example, a vector of K binary or bits represented a qualitative variable with K independent categorical features; only a single bit was nonzero to indicate the associated category among K possible categories

(Hastie et al. 2009). This approach was used to represent the cutting direction and strategy as coded variables to convert qualitative features into quantitative features:

- $(x_4, x_5, x_6, x_7) \in \{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\}$ **Cutting directions**: x-cut, y-cut, z-cut, or xy-cut, which were represented as coded variables (1,0,0,0), (0,1,0,0), (0,0,1,0) and (0,0,0,1), respectively.
- $(x_8, x_9, x_{10}) \in \{(1,0,0), (0,1,0), (0,0,1)\}$ **Cutting strategies**: Conventional milling, climb milling, or a combination of both (as in slotting), which were represented as coded variables,(1,0,0), (0,1,0) and (0,0,1), respectively.

Using categorical or coded variables, the prediction model was able to represent any combination of cutting direction and cutting strategy. Furthermore, the use of coded variables allowed the prediction model to be constructed using the entire training data. Otherwise, we would have needed to partition the dataset into subsets with few data points according to each combination of cutting direction and cutting strategy if we had to construct an individual prediction function for each combination of features.



(a) Process parameters of a milling process (b) Cutting strategy

Figure 4. Machining process parameters.

The output (response) variable of the energy prediction model was the energy per unit length of cut, which is a quantitative measure. Within each code block i, the power consumption of a phase, $P_k^{(i)}$, $k = 1, \dots, N_P$, over a time duration of t_k, was retrieved as a time series dataset using MTConnect. The total energy consumption, $E^{(i)}$ for NC code block i, was computed as:

$$E^{(i)} = \sum_{k=1}^{N_P} P_k^{(i)} \times t_k. \quad (1)$$

The number of data points, N_P, in each NC code block i depended on the duration of the corresponding operation and the sampling rate for the power measurement, which was 100 Hz in this study. We generalized the energy consumption by using energy density $y^{(i)} = E^{(i)}/l^{(i)}$ (i.e., the energy per unit length of cut) as the output response feature. That is the length of cut scales the predicted energy consumption. Predicting the energy density implicitly included the dependence of the duration of cut on the feed rate and length of cut. It also allowed us to predict the energy consumption of a part with different (unseen) dimensions, which made the model spatially scalable. Note that we modeled the relationship between the averaged values of the process parameters and the average power (or energy density) across the duration of the block.

For the 18 parts machined for the experiments in this study, a total of 12,299 datasets of input feature vector $\boldsymbol{x}$ and output feature y were generated after post-processing and cutting simulation; each dataset corresponded to an individual NC code block. We filtered out those datasets that corresponded to NC code blocks that had duration shorter than 2 seconds except for those blocks corresponding to rapid motion. This process prevented statistically low quality data from biasing the prediction model since data from blocks of longer duration are more stable. The filtered dataset $\boldsymbol{D} = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m\}$, where $m = 3{,}214$ (i.e.,

data from 3,214 NC code blocks remained after filtering) was further categorized into seven different datasets $\{\boldsymbol{D}_1, \dots, \boldsymbol{D}_q, \dots, \boldsymbol{D}_7\}$ that corresponded to the seven cutting operations described in Figure 3 and Section 2.3; each dataset $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$ contained m_q NC code blocks for the cutting operation type q.

3. DATA-DRIVEN APPROACH FOR ENERGY PREDICTION

To construct a data-driven energy prediction model for a machine tool using the data described in Section 2, we can apply Gaussian Process (GP) regression because it can construct a non-linear regression model with high-dimensional input features using a relatively small number of training data. As a non-parametric regression technique, GP regression can model the input and output relationship without using a set of pre-defined basis functions. Instead, it uses bases formed from the training data. Due to this flexibility, GP regression is able to model complex relationships among input variables and a target response with the least number of hyper-parameters. Additional benefits of GP regression are its ability to quantify uncertainties in the predicted values and its ability to update the regression model incrementally. GP regression has been applied to many fields, including modeling robotics (Nguyen-Tuong et al. 2009), human motions (Wang et al., 2008), and traffic flow (Kim et al. 2011). The following sections describe the procedure we applied to construct the energy prediction model using GP regression.

3.1 Gaussian Process

GP regression is employed to approximate the unknown energy prediction function $f(\boldsymbol{x})$ using historical data on the machining process parameters and corresponding energy consumption. A GP is a collection of random variables (stochastic process), any finite set of which has a joint Gaussian distribution (Rasmussen and Williams 2006). By treating the values of the unknown function $f(\cdot) = \text{GP}(m(\cdot), k(\cdot,\cdot))$ as a collection of random variables, GP describes the function probabilistically as a multivariate Gaussian distribution specified by its mean function $m(\cdot)$ and the covariance function $k(\cdot,\cdot)$. The mean function $m(\cdot)$ captures the prior mean of the target function, which is usually assumed to be zero. The covariance function $k(\cdot,\cdot)$ quantifies the correlation between input data in terms of their function values.

In GP regression, we assume that the output $y = f(\boldsymbol{x}) + \epsilon$ is measured with noise $\epsilon \sim N(0, \sigma_\epsilon^2)$, which is Gaussian distributed with zero mean and variance σ_ϵ^2. The values for the unknown function $f(\boldsymbol{x})$ are treated as random variables and modeled by a Gaussian distribution for incorporating prior knowledge captured in the historical data. Suppose the current dataset is denoted by $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$ for the machining operation type q. The measured output $y^{new} = f_q(\boldsymbol{x}^{new}) + \epsilon^{new}$ corresponding to the new input feature $\boldsymbol{x}^{new}$ and the historical outputs $\boldsymbol{y}^{1:m_q} = \{y^1, \dots, y^{m_q}\}^T$ in the training dataset $\boldsymbol{D}_q$ follow a multivariate Gaussian distribution (Rasmussen and Williams 2006):

$$\begin{bmatrix} \boldsymbol{y}^{1:m_q} \\ y^{new} \end{bmatrix} \sim N\left(\boldsymbol{0}, \begin{bmatrix} \mathbf{K} & \boldsymbol{k} \\ \boldsymbol{k}^T & k(\boldsymbol{x}^{new}, \boldsymbol{x}^{new}) \end{bmatrix}\right), \tag{2}$$

where $\boldsymbol{k}^T = \{k(\boldsymbol{x}^1, \boldsymbol{x}^{new}), \dots, k(\boldsymbol{x}^{m_q}, \boldsymbol{x}^{new})\}$ and $\mathbf{K}$ is the covariance matrix (kernel matrix) whose (i,j)th entry is $\mathbf{K}_{ij} = k(\boldsymbol{x}^i, \boldsymbol{x}^j)$. The value of the covariance function $k(\boldsymbol{x}^i, \boldsymbol{x}^j)$ quantifies the amount the two input feature vectors $\boldsymbol{x}^i$ and $\boldsymbol{x}^j$ change together. Note that the more the two vectors $\boldsymbol{x}^i$ and $\boldsymbol{x}^j$ differ, the closer the value of the covariance approaches zero, which implies that the two input vectors are not correlated in terms of their function values. An effective kernel function can be chosen considering the characteristics of the target function. Noting that energy consumption varies smoothly with the changes in the machining parameters (Diaz et al. 2011), we use a squared exponential kernel function that can effectively describe a continuously varying function. The squared exponential kernel function evaluates the covariance between the two input feature vectors $\boldsymbol{x}^i$ and $\boldsymbol{x}^j$ as (Neal, 1996):

$$k(\boldsymbol{x}^i, \boldsymbol{x}^j) = \sigma_s^2 \exp\left[-\frac{1}{2}\sum_{r=1}^{n}\left(\frac{x_r^i - x_r^j}{\lambda_r}\right)^2\right] + \sigma_\epsilon^2 \delta_{ij}. \tag{3}$$

The kernel function is described by the hyper-parameters $\boldsymbol{\theta} = \{\sigma_s, \sigma_\epsilon, \boldsymbol{\lambda}\}$. The term σ_s^2 is referred to as the signal variance, which quantifies the overall magnitude of the covariance value. The term σ_ϵ^2 is referred to as the noise variance, which quantifies the level of noise assumed to exist in the observed output response. The Kronecker delta function δ_{ij} serves to selectively specify the noise variance σ_ϵ^2 to the covariance value $k(\boldsymbol{x}^i, \boldsymbol{x}^j)$; that is, the noise signals added to different measurements are assumed to be independent and the noise correlation is non-zero only when $i = j$. The vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r, \dots \lambda_n)$ is referred to as the characteristic length scales to quantify the relevancy of the input features in $\boldsymbol{x} = (x_1, \dots, x_r, \dots x_n)$ for predicting the response y. Note that we used a total of $n = 10$ input features in this study. A large length scale λ_i indicates weak relevance, while a small length scale λ_i implies strong relevance of the corresponding input feature x_i.

The hyper-parameters $\boldsymbol{\theta} = \{\sigma_s, \sigma_\epsilon, \boldsymbol{\lambda}\}$ are determined by maximizing the log-likelihood of the measurement data. Using the definition of GP in Eq. (2), the log-likelihood function of data $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$ can be expressed as (Rasmussen and Williams 2006):

$$L(\boldsymbol{\theta}; \boldsymbol{D}_q) = P(\boldsymbol{y}^{1:m_q} | \boldsymbol{x}^{1:m_q}; \boldsymbol{\theta}) = N(0, \mathbf{K}) = \frac{1}{\sqrt{(2\pi)^{m_q}|\mathbf{K}|}} \exp\left(-(\boldsymbol{y}^{1:m_q})^T \mathbf{K}^{-1} (\boldsymbol{y}^{1:m_q})\right). \tag{4}$$

Note that $\mathbf{K}$ is the covariance matrix whose (i, j) entries are defined as shown in Eq. (3). The optimum hyper-parameters $\boldsymbol{\theta}^* = \{\sigma_s^*, \sigma_\epsilon^*, \boldsymbol{\lambda}^*\}$ are then determined as that maximize the log-likelihood of the training data $\boldsymbol{D}_q$ as (Rasmussen and Williams 2006):

$$\begin{aligned} \boldsymbol{\theta}^* &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, L(\boldsymbol{\theta}; \boldsymbol{D}_q), \\ &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} -\frac{1}{2}(\boldsymbol{y}^{1:m_q})^T \boldsymbol{K}^{-1} \boldsymbol{y}^{1:m_q} - \frac{1}{2}\log|\boldsymbol{K}| - \frac{m_q}{2}\log 2\pi. \end{aligned} \tag{5}$$

With the gradient $\nabla \log L(\boldsymbol{\theta}; \boldsymbol{D}_q)$ of the log-likelihood function $L(\boldsymbol{\theta}; \boldsymbol{D}_q)$ available, Eq. (5) can be solved using a mathematical optimization algorithm. We use Gaussian Processes for Machine Learning (GPML), a GP package implemented in MATLAB® to optimize the hyper-parameters (Rasmussen and Nickisch 2013).

After the measurement data and the covariance function (hyper-parameters) are updated, GP regression predicts the unknown response y^{new} corresponding to a new input feature vector $\boldsymbol{x}^{new}$ in a probabilistic fashion. Since the distribution conditional on any subset of the data assumed to be Gaussian distributed, the posterior distribution $p(y^{new} | \boldsymbol{D}_q, \boldsymbol{x}^{new})$ on y^{new} given the historical dataset $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$ and the new input feature vector $\boldsymbol{x}^{new}$ can be expressed as a 1-D Gaussian distribution (Rasmussen and Williams 2006):

$$p(y^{new} | \boldsymbol{D}_q, \boldsymbol{x}^{new}) = N\left(y^{new}; \mu(\boldsymbol{x}^{new} | \boldsymbol{D}_q), \sigma^2(\boldsymbol{x}^{new} | \boldsymbol{D}_q)\right). \tag{6}$$

The posterior distribution $p(y^{new} | \boldsymbol{D}_q, \boldsymbol{x}^{new})$ can be described by its mean μ and variance σ^2, which can be expressed, respectively, as (Rasmussen and Williams, 2006):

$$\mu(\boldsymbol{x}^{new} | \boldsymbol{D}_q) = \boldsymbol{k}^T \mathbf{K}^{-1} \boldsymbol{y}^{1:m_q}, \tag{7}$$

$$\sigma(\boldsymbol{x}^{new} | \boldsymbol{D}_q) = \sqrt{k(\boldsymbol{x}^{new}, \boldsymbol{x}^{new}) - \boldsymbol{k}^T \mathbf{K}^{-1} \boldsymbol{k}}\,. \tag{8}$$

That is, we can obtain the mean function $\mu(\boldsymbol{x}^{new} | \boldsymbol{D}_q)$ from the GP regression to predict the most probable energy density $y^{new} = f_q(\boldsymbol{x}^{new}) + \epsilon^{new}$ for a given input feature vector $\boldsymbol{x}^{new}$ and the standard deviation function $\sigma(\boldsymbol{x}^{new} | \boldsymbol{D}_q)$ to quantify the uncertainty in the predicted value of y^{new} at $\boldsymbol{x}^{new}$. The energy

consumption per each machining operation is then aggregated to predict the total energy consumption (with some estimated uncertainty bound) for machining a part.

3.2 Estimating test error

Selecting the type of basis function and choosing the optimum feature sets precedes fitting a prediction model to the training dataset. For GP regression, once the type of kernel function is specified, the optimum feature selection is implicitly carried out by optimizing the hyper-parameters for the kernel function. For example, the optimized length scales $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r, \dots \lambda_n)$ for the exponential squared function automatically weigh the importance of the corresponding features in predicting the output response since a smaller λ_r implies a larger influence of the corresponding input feature x_r on the output response y. This property of feature weighting, generally known as automatic relevance determination (ARD) (Neal 1996), simplifies the construction of the energy prediction model since all features are being included to construct the energy prediction functions without explicitly conducting the feature-selection procedure.

Depending on the machining operation, the parameters in the input feature vector x affected the energy density value y differently. For each machining operation type q with the dataset $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$, we constructed the individual energy-prediction function for that operation using GP regression. We then estimated (generalization) errors for each prediction function using the holdout cross-validation technique (Hastie et al. 2009). Note that here the (generalization) error was estimated to provide insight into how well each individual energy prediction function would perform with unseen test data.

For each machining operation type q with the dataset $\boldsymbol{D}_q = \{(\boldsymbol{x}^i, y^i) | i = 1, \dots, m_q\}$, we trained the model and computed the error rates as follows:

(1) Randomly divide the dataset D_q into the training dataset D_q^{tr} with m_q^{tr} training data points and the test dataset $\boldsymbol{D}_q^{te}$ with m_q^{te} test data points. In this study, we set the ratio $m_q^{tr} : m_q^{te} = 7:3$, which is a common ratio used to estimate the accuracy (i.e., test error) of predictions for supervised learning algorithms (Hastie et al. 2009).

(2) Construct the energy density prediction function $f_q(x)$ by computing $\mu(\boldsymbol{x}|\boldsymbol{D}_q)$ and $\sigma(\boldsymbol{x}|\boldsymbol{D}_q)$ using the training dataset $\boldsymbol{D}_q^{tr}$.

(3) Predict the energy densities corresponding to the input features in the test dataset D_q^{te} and computed the error by comparing them to the true energy densities in the test dataset D_q^{te}. The error was measured in terms of the mean absolute error (MAE), which was more insensitive to outliers than the root mean square error (RMSE) (Willmott and Matsuura 2005):

$$\mathrm{MAE}_q = \frac{1}{m_q^{te}} \sum\nolimits_{\{(\boldsymbol{x}^i, y^i) \in \boldsymbol{D}_q^{te}\}} \left| \mu(\boldsymbol{x}^i | \boldsymbol{D}_q) - y^i \right|. \tag{9}$$

MAE_q quantifies the average deviation between the predicted and measured values. To further quantify how much the predicted values in terms of a ratio to the mean density $\bar{y}_q$ for the machining operation type q, we use the normalized mean absolute error (NMAE) (Gustafson and Shaocai 2012):

$$\mathrm{NMAE}_q = \frac{\sum_{\{(\boldsymbol{x}^i, y^i) \in D_q^{te}\}} \left| \mu(\boldsymbol{x}^i | \boldsymbol{D}_q) - y^i \right|}{\sum_{\{(\boldsymbol{x}^i, y^i) \in \boldsymbol{D}_q^{te}\}} y^i} = \frac{\mathrm{MAE}_q}{\bar{y}_q}. \tag{10}$$

Note that we could have computed the average deviation between the predicted and measured densities, *i.e.*, MAE_q, by simply multiplying NMAE_q with the measured mean density $\bar{y}_q$ (for the machining operation type q).

The value of NMAE_q could have fluctuated depending on the selected training and test datasets. To quantify the test error reliably, 100 values of NMAE_q were computed using the procedure above. The averaged value μ_{NMAE} was then determined and used as an error measure in this study. Note that the number of repetitions was chosen empirically so that a stable, representative mean value could be determined irrespective of the selected training and test datasets.

Table 3 compares the estimated (generalization) errors for the energy density prediction function for each machining operation type. The averages for the normalized mean absolute error μ_{NMAE} (computed using 100 NMAE values from 100 test experiments) for the cutting operations are different due to the different cutting mechanisms and the different numbers of training data used for constructing the models. Overall the μ_{NMAE} values range between 8% and 45%; the smallest values occur for the feed with cut operations. The standard deviation σ_{NMAE} for the average of the normalized mean absolute error quantifies the variability in the estimated μ_{NMAE}.

Table 3: The estimated test error for each energy density prediction function (for rapid motion, the time filtering is not applied since the duration of rapid motions are mostly less than 2 sec).

Operation type		Number of NC blocks	Average duration (sec)	$\mu_{\bar{y}_q}$ (J/mm)	μ_{MAE} (J/mm)	μ_{NMAE} (%)	σ_{NMAE} (%)
Feed with cut	Face milling	1225	18.973	221.963	18.736	8.440	0.472
	Contouring	401	7.278	247.463	30.957	12.488	1.585
	Slotting	119	4.753	240.515	26.148	10.881	1.173
	Pocketing	196	4.213	262.518	33.423	12.735	1.935
	Plunge	115	8.461	1722.108	501.457	29.189	3.212
Non cut	Air cut	384	6.999	841.273	70.157	8.350	0.798
	Rapid motion	110	0.528	100.572	44.069	44.653	12.521

Table 3 also shows that when the size of training data was small, for example for the plunge and rapid motion operations, the estimated values fluctuated significantly. Since the average duration of these operations was extremely small when compared to other operations, the data quality was much lower, which resulted in greater prediction error and variation. Also, the plunge operation was conducted using an end-mill tool without a center cut. This non-standard use of the cutting tool can be one of the causes for the fluctuation in the energy prediction for the plunge operation. To gain better insight into the accuracy of the energy prediction model, the value of μ_{NMAE} can be compared to the coefficient of variation (CV) in the energy density values, which captures the inherent fluctuation of the energy density in the training dataset. The coefficient of variation is computed based on absolute difference between the energy densities and its mean:

$$\mathrm{CV} = \sum_{i=1}^{N} \left|y^i - \bar{y}\right| \Big/ \sum_{i=1}^{N} y^i = 77.165\% \tag{11}$$

The values of μ_{NMAE} obtained by the energy prediction model are much lower than the value of CV, which implies that the energy prediction model captures the variations in the energy density induced by different machine operations and parameters well.

3.3 Uncertainty quantification in the prediction model

Using the energy density prediction model for each machining operation type q represented by the mean energy density function $\mu_q(\boldsymbol{x}|\boldsymbol{D}_q)$ and the associated standard deviation function $\sigma_q(\boldsymbol{x}|\boldsymbol{D}_q)$, the total energy consumption for machining a part can be estimated from the NC codes. First, we can estimate the energy consumption $\hat{E}^i$ and the standard deviation S^i from the input feature $\boldsymbol{x}^i$ of the NC code block i performing the machining operation type q as:

$$\hat{E}^i = \mu_q\left(\boldsymbol{x}^i|\boldsymbol{D}_q\right) \times l^i, \tag{12}$$

$$S^i = \sigma_q\left(\boldsymbol{x}^i\middle|\boldsymbol{D}_q\right) \times l^i, \tag{13}$$

where l^i is the length of cut specified for the operation by the NC code. Aggregating all the NC blocks for the machining operation type q, the predicted total energy consumption $\hat{E}_q$ and the associated standard deviation S_q can be computed for that operation type:

$$\hat{E}_q = \sum\nolimits_{\{(\boldsymbol{x}^i, y^i)\in \boldsymbol{D}_q\}} \mu_q\left(\boldsymbol{x}^i|\boldsymbol{D}_q\right) \times l^i, \tag{14}$$

$$S_q = \sqrt{\sum\nolimits_{\{(\boldsymbol{x}^i, y^i)\in \boldsymbol{D}_q\}} \left(\sigma_q(\boldsymbol{x}^i|\boldsymbol{D}_q) \times l^i\right)^2}. \tag{15}$$

Finally, the estimated total energy consumption $\hat{E}$ for machining a whole part and the standard deviation S associated with the estimation can be computed by summing the mean predicted energy $\hat{E}_q$ and accumulating the standard deviation S_q for all machining operation types, $q = 1, \ldots, Q$, where Q = 7 (including all cutting and non-cutting operations). Because the energy consumed in each machining operation is considered independent of the energy consumed by other machining operations, $\hat{E}$ and S are expressed as:

$$\hat{E} = \sum\nolimits_{q=1}^{Q} \hat{E}_q, \tag{16}$$

$$S = \sqrt{\sum\nolimits_{q=1}^{Q} \left(S_q\right)^2}. \tag{17}$$

Note that the energy density $\hat{y}$ is represented to be a Gaussian random variable in the framework of GP regression. Because a linear combination of Gaussian random variables is also Gaussian, the predicted total energy E, which is computed as a linear combination of the energy densities, is also Gaussian. The probability distribution on the total energy E then can be expressed as $E \sim N\left(\hat{E}, S^2\right)$ with the mean $\hat{E}$ and the standard deviation S given in Eq. (16) and Eq. (17), respectively.

4. VALIDATION TESTS

The energy prediction model constructed based on GP regression was used to predict the energy consumption for machining a generic part. This section discusses the validation of the trained energy prediction function using unseen test data.

4.1 Data collection from a blind test

Figure 5 shows a generic part, the geometry of which is quite different from the part used in the training process (see Figure 3). The cutting and non-cutting operations used to produce the generic test part are face milling, pocketing, plunge, air cut, and rapid motion.

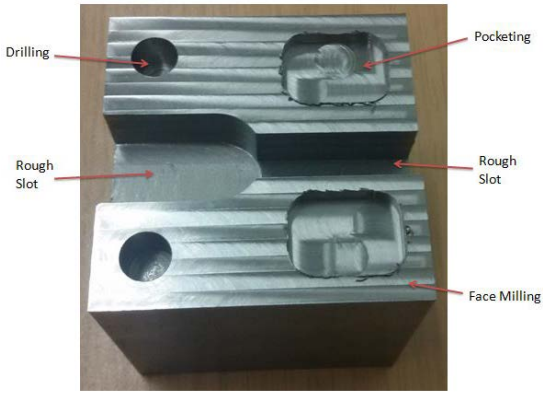


Figure 5. Generic test part used to validate the energy prediction model

The accuracy of the energy prediction model depended on how the machining parameters for a test part were distributed relative to the machining parameters used in the training dataset. If the machining parameters for the test part were completely different from the machining parameters used to collect the training dataset, the accuracy of the prediction fell. To study how the energy prediction model generalized over unobserved test data, we validated the energy prediction model by machining three test parts with the geometry shown in Figure 5, but we intentionally varied the spindle speeds in these experiment as shown in Table 4. Comparing the spindle speeds used to machine the 18 training parts to those in Table 4, the first test part uses the same spindle speed while the second and third use different spindle speeds. We chose these spindle speeds to evaluate the model's capability of predicting the energy density values in incrementally more unexplored parameter space. For all test parts, the depth of cut was set to 1 mm.

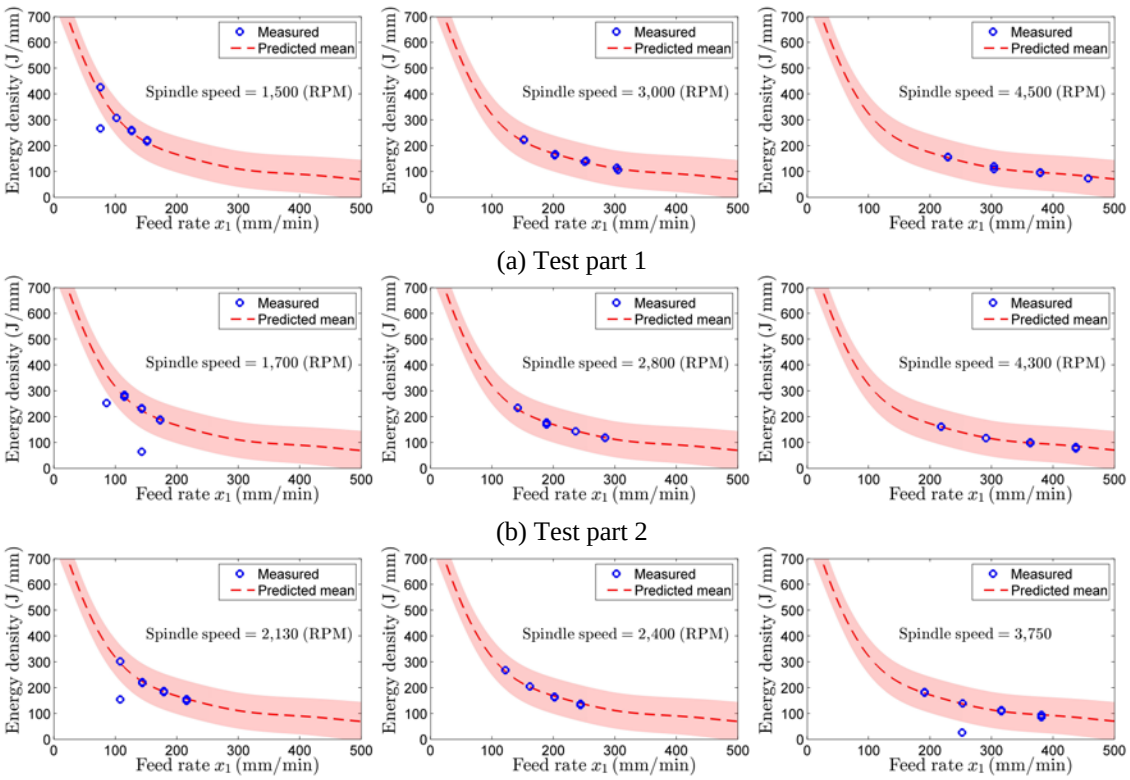
Table 4: Spindle speeds chosen for the blind tests

	Used spindle speeds (in RPM)
Training parts 1~18	$\{1{,}500;\ 3{,}000;\ 4{,}500\}$
Test part 1	$\{1{,}500;\ 3{,}000;\ 4{,}500\}$
Test part 2	$\{1{,}700;\ 2{,}800;\ 4{,}300\}$
Test part 3	$\{2{,}130;\ 2{,}400;\ 3{,}750\}$

4.*2 Prediction result*

Figure 6 shows the measured energy density values y and the predicted energy density function $\hat{y}$ for the face milling operations with different spindle speeds and different feed rates. To visualize the high-dimensional prediction function for the energy density, we fix the other machining parameters for y-direction cut and conventional cutting strategy and set the depth of cut to 1 mm. For each plot, the curve shows how the energy density varies with the feed rate for fixed spindle speed. The influence of the spindle speed on the energy density can be studied by comparing the curves shown in the figure. In each plot, the dash line represents the predicted mean $\mu_1(\boldsymbol{x}|D_1)$ and shaded band represents the 95% confidence bound on the predicted energy density, i.e., $\mu_1(\boldsymbol{x}|D_1) \pm 1.96\sigma_1(\boldsymbol{x}|D_1)$.

As Figure 6 shows, the energy density measurements for the face milling operations in test parts 1, 2, and 3 are well captured by the energy density prediction function for each spindle-speed/feed-rate combination. The overall trend of the energy density is well predicted by the mean function $\mu_1(\boldsymbol{x}|D_1)$. In addition, most measurements are within the 95% confidence bound on the predicted energy density. The width of the confidence bound changes depending on the distribution of the training data used to build the model. In general, the confidence bound for high feed rate is larger because a fewer number of data points were collected in this region to build the model.



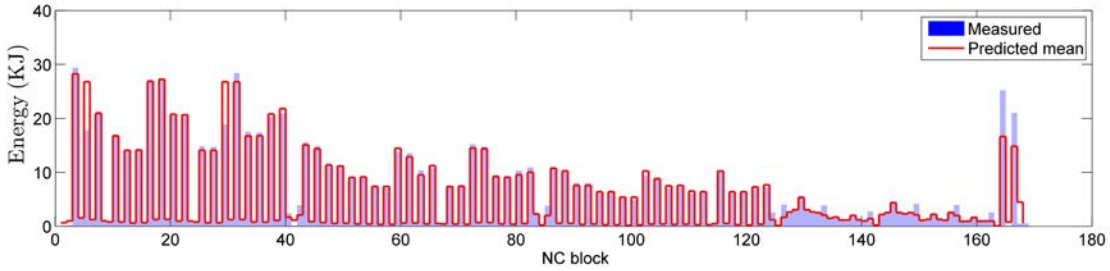
(a) Test part 1

(b) Test part 2

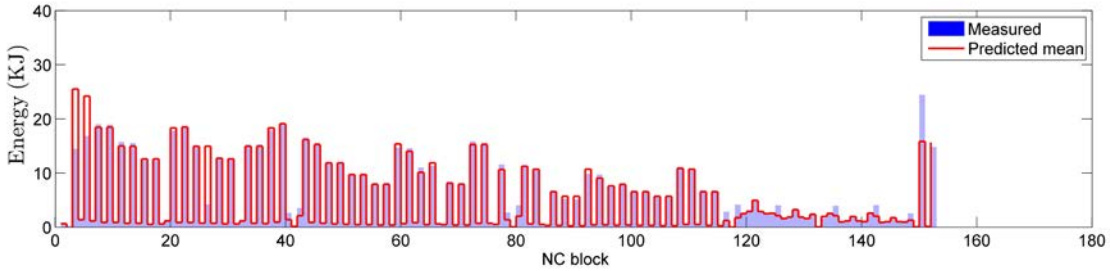
(c) Test part 3

Figure 6. Predication of energy density values for generic test parts (machined using face-milling, y-direction cut, conventional cutting strategy, and depth of cut = 1mm). The band represents $\mu_1(\boldsymbol{x}^i|D_1) \pm 1.96\sigma_1(\boldsymbol{x}^i|D_1)$.

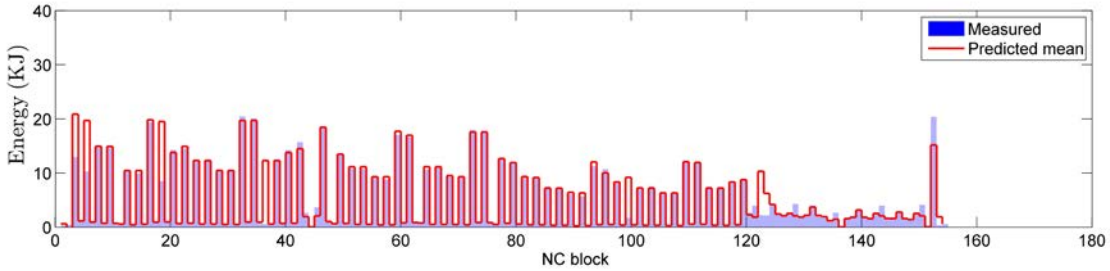
Figure 7 compares the predicted and the measured energy consumption for each individual NC code block. To predict the energy consumption $\hat{E}^i$ for block i, the type of machine operation q is first identified and the energy density prediction function $\mu_q(\boldsymbol{x}^i|\boldsymbol{D}_q)$ corresponding to that operation q is used. The predicted (mean) energy for block i is then computed using Eq. (12). In general, the predicted energy consumption values match well with the measurements. The deviation of the mean energy prediction from the measured energy consumption increases from test part 1 to 3. This is because the machining parameters in test part 3 are the furthest away from the observed values in the training data.



(a) Test Part 1



(b) Test Part 2



(c) Test Part 3

Figure 7. Predication of total mean energy consumptions including all operations

Finally, Table 5 compares the predicted and measured energy consumption using the normalized mean absolute error (NMAE) and relative total error (RTE) defined as:

$$\text{NMAE} = \frac{\sum_{\{i\in \text{ NC blocks}\}} \left|\hat{E}^i - E^i\right|}{\sum_{\{i\in \text{ NC blocks }\}} E^i}, \tag{18}$$

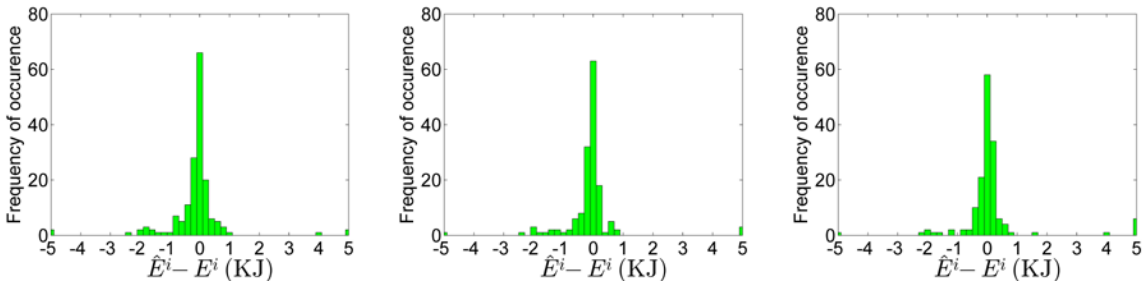
$$\text{RTE} = \frac{\left|\hat{E} - E\right|}{E}. \tag{19}$$

Table 5: Summary of prediction results on the generic test parts

	No. of data	Averaged block duration (sec)	NMAE (%)	Measured Total energy (KJ)	Prediction total energy (KJ)	Standard deviation (KJ)	RTE (%)
Test 1	168	10.014	9.577	909.266	891.022	35.259	-2.007
Test 2	152	9.567	10.359	768.605	766.901	37.696	-0.222
Test 3	154	9.577	13.553	761.272	806.761	34.508	5.976

Note that the NMAE in Eq. (18) is defined using the predicted energy $\hat{E}^i$ and the measured energy E^i for each NC code block i, whereas the NMAE in Eq. (10) is defined using the predicted energy density $\hat{y}^i$ and the measured energy density y^i. Thus, the energy prediction with the longer length of cut l^i will contribute more to the value of NMAE in Eq. (18). In spite of this dependence on the geometry, the measure can still quantify the mean absolute errors of the three test cases in a relative manner. As Table 5 shows, the NMAE for the three test parts are less than 15%, which are consistent with the estimated error using the training dataset based on the hold-out cross-validation method. In other words, the energy prediction model generalizes quite well for the unseen test dataset, which validates the effectiveness of the model in predicting the energy consumed to machine a generic part.

While the NMAE quantifies error in the predicted energy for a single cut, the RTE quantifies the errors in the predicted total energy consumption for producing a whole part. Table 5 shows that for all test cases, the RTE is less than 6%. In addition, the measured total energy falls within the 95% confidence bound $\hat{E} \pm 1.96S$ on the predicted total energy. The RTEs for the energy prediction are small for all three test parts because the errors $\hat{E}^i - E^i$ are distributed centered at the zero-mean with an almost equal chance to over- or underestimate the energy as shown in Figure 8. The overestimations and the underestimations on the block-wise energy consumptions are canceled out when they are summed up to compute the total energy consumption. Therefore, the block-wise energy prediction results in accurate estimation on the total energy consumption for machining a whole part.



(a) Test part 1 (b) Test part 2 (c) Test part 3

Figure 8. Distributions of errors

5. SELECTION OF MACHINING STRATEGY

In addition to predicting the energy consumption, the energy prediction functions can also be used to determine an energy-efficient toolpath to machine a part or to enable novel monitoring strategies by highlighting abnormal behavior. In this section, we discuss the use of energy prediction functions to select the toolpath that uses the least amount of energy to machine a part.

5.*1 Experiments for toolpath planning*

The machine-tool coordinates (x, y, z), with respect to the global reference, represent the location of the cutting tool. The toolpath is then described by the temporal sequences of these coordinates. The tool's sequential moves with respect to the geometry of a workpiece determine the cutting direction and the cutting strategy. Figure 9 shows four different toolpaths that were explored to machine the pocket shown in Figure 10. Table 6 shows the process parameters used to execute these four different toolpaths. Each toolpath is composed of cuts in different directions and with different cutting strategies. The goal is to select the toolpath that minimizes the predicted energy consumption before actually machining the part. This prediction can then be compared to the true energy consumption measured during experiment.

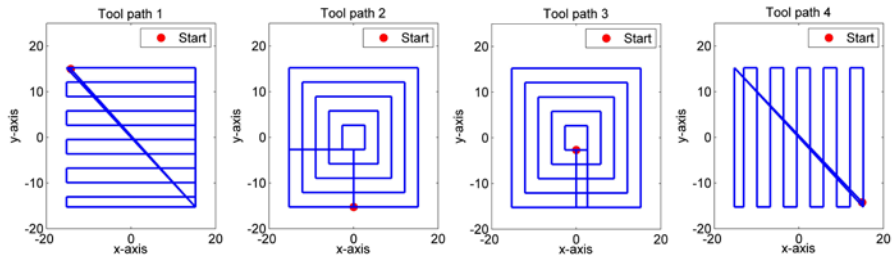


Figure 9. Toolpath comparison

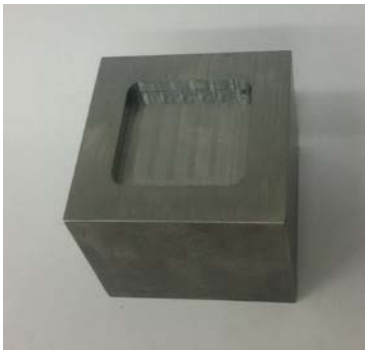


Figure 10. Part geometry for ordering different toolpath strategies

Table 6. Process parameters used in the experiments to order different toolpaths

Levels	Cutting Speed (RPM)	Chip Load (inches)	Depth of Cut (mm)
1	1500	0.001	1.5
2	300	0.002	1.5

5.*2 Energy-efficient toolpath selection*

Table 7 summarizes the results of the experiments to machine the part shown in Figure 10 using the four different toolpaths in Figure 9. The required energy varies depending on the toolpath used, and the energy prediction function predicts the total energy consumption with good accuracy. Figure 11 compares the measured and predicted energies for each toolpath. The error bar on the predicted energy usage represents the 95% interval, i.e., $\mu \pm 1.96\sigma$, for the predicted total energy consumption. Note that the measured energy values all fall within the 95% confidence bound on the predicted total energy. With the predicted energies, the toolpaths can be ordered in terms of their energy consumption, and the toolpath with the minimum energy consumption can be selected accordingly.

Table 7. Summary of prediction results on the toolpath comparison using the generalized energy prediction function.

	No. of data	Averaged block duration (sec)	NMAE (%)	Measured Total energy (kJ)	Prediction total energy (kJ)	Standard deviation (kJ)	RTE (%)
Path 1	73	6.806	11.041	273.389	282.377	9.466	3.288
Path 2	111	5.181	15.436	303.466	320.896	8.995	5.744
Path 3	109	5.415	11.655	327.643	338.277	8.788	3.246
Path 4	84	6.225	16.015	301.907	311.833	11.300	3.288

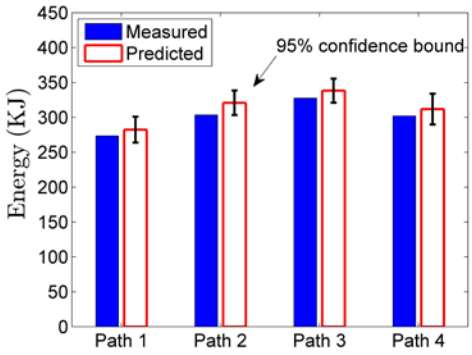


Figure 11. Total energy consumption comparison for different toolpath strategies.

6. CONCLUSIONS

This study demonstrates the use of a non-parametric regression model, namely the Gaussian Process (GP), to predict the energy consumption of a machine tool. The GP models the complex relationships between the input machining parameters and output energy consumption and constructs a prediction function for the energy consumption with confidence bounds. Even though the training datasets in this study include only 18 experimental parts, the models constructed using the machine-learning approach are able to reliably predict the energy consumption for machining a generic test part with the milling machine tool. This is primarily because of the block-wise experimentation and data analysis conducted which rendered each block of NC code as an experiment in itself. In addition, the energy prediction function is used to select the optimum toolpath that uses the least amount of energy to machine the same part.

There are other parameters that can possibly affect the energy consumption pattern of a target machine. For example, the workpiece material or cutting tool geometry and material can affect the energy consumption pattern of the target machine. By including these parameters as input features, the energy prediction model can be further improved and generalized. In the future, we plan to conduct additional experiments to collect datasets that include these features to improve the robustness and generalizability of the energy prediction model.

To effectively establish the energy consumption pattern of a machine tool over time, the energy prediction model would need to be updated continuously with new measurement data to account for the time-varying characteristics of the machine tool (e.g., due to tool wear and machine tool deterioration). Incorporating these characteristics, particularly tool wear, into the modeling approach is one area of future study. Another area of future work is constructing a near-real-time energy prediction model for a machine tool by combining a near-real-time data collection framework with an adaptive GP regression model. We are currently developing a near-real-time data collection framework to retrieve raw data from a milling machine tool and its sensors and convert the data into relevant input features. In addition, we are currently investigating the use of sparse representation of the covariance matrix to reduce the computational and storage demands of GP regression, which can help to update the GP regression model with near-real-time streaming data. We expect that the energy prediction function can be constructed using a fraction of training data points (perhaps as few as 10%), which can reduce the training time without sacrificing accuracy significantly.

As alluded to in Section 5, an energy prediction model that is continually updated can be used to monitor the condition of machine components. One area where we can apply energy prediction for machine tool monitoring is anomaly detection. For example, a sudden, unexpected event (such as tool breakage or machine collision due to incorrect tool offsets) may cause a deviation between the predicted and actual energy consumption or power demand, which can trigger an immediate alarm. Developing monitoring strategies based on deviations between the predicted and measured energy consumption or power demand represents a potentially impactful area of future study.

Finally, in addition to energy consumption, toolpath selection can be based on other criteria, which may include minimum machine operating time, minimum impact to a machine tool, and optimized surface roughness. Given data-driven predictive models for different performance features, such as time, tool wear, or surface roughness, an efficient toolpath can be chosen by considering these impacts individually or simultaneously. This study suggests one possible scenario of integrating the energy consumption prediction function into a CAM system. The integration of the energy prediction model into the CAM system would allow smart and well-informed decision regarding toolpath selection.

In conclusion, this study shows that with advanced data collection and processing techniques, prediction models can be constructed to predict energy consumption of a machine tool with multiple operations and multiple process parameters. The specific energy prediction model that was generated in this study would work for generic parts machined on a Mori Seiki NVD1500. The methodology that was described, though, could be used to create prediction models for other machine tools to enable improved planning and operations in various shop-floor environments.

Acknowledgement and Disclaimer

The authors acknowledge the support by the Smart Manufacturing Systems Design and Analysis Program at the National Institute of Standards and Technology (NIST), Grant Numbers 70NANB12H225 and 70NANB12H273 awarded to University of California, Berkeley, and to Stanford University respectively. In addition, the authors appreciate the support of the Machine Tool Technologies Research Foundation (MTTRF) and System Insights for the equipment used in this research. Certain commercial systems are identified in this paper. Such identification does not imply recommendation or endorsement by NIST; nor does it imply that the products identified are necessarily the best available for the purpose. Further, any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NIST or any other supporting U.S. government or corporate organizations.

REFERENCES

Bhinge, R., Park, J., Biswas, N., Helu, M., and Dornfeld, D., Law, K., and Rachuri, S. (2014). "An Intelligent Machine Monitoring System Using Gaussian Process Regression for Energy Prediction," *IEEE International Conference on Big Data* (IEEE BigData 2014), Washington, DC.

Box, G., Hunter, J.S., and Hunter, W.G. (1979). "Statistics for Experimenters: Design, Inovation, and Discovery," Wiley.

Diaz, N., Helu, M., Jarvis, A., Tonissen, S., Dornfeld, D. and Schlosser, R. (2009). "Strategies for Minimum Energy Operation for Precision Machining," *Proceeding of MTTRF 2009 Annual Meeting*, Shanghai, China.

Diaz, N., Redelsheimer, E. and Dornfeld, D. (2011). "Energy Consumption Characterization and Reduction Strategies for Milling Machine Too Use," *Proceeding of 18th CIRP International Conference on Life Cycle Engineering*, Braunschweig, Germany.

Dietmair, A. and Verl, A. (2009). "Energy Consumption Forecasting and Application for Tool Machines," *Modern Machinery Science Journal*, pp. 62-67.

Draganescu, F., Gheorghe, M. and Doicin, C.V. (2003). "Models of Machine Tool Efficiency and Specific Consumed Energy," *Journal of Material Processing Technology*, 141, pp. 9-15.

Gustafson, W.I., and Shaocai, Y. (2012). "Generalized Approach for Using Unbiased Symmetric Metrics with Negative Values: Normalized Mean Bias Factor and Normalized Mean Absolute Error Factor," *Atmospheric Science Letter*, 13(4), pp. 262-267.

Gutowski, T., Dahmus, J. and Thiriez, A. (2006). "Electrical energy requirements for manufacturing processes," *Proceedings of 13th CIRP International Conference of Life Cycle Engineering*, Lueven, Switzerland.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning, Springer, New York, NY.

Helu, M., Robinson, S., Bhinge, R., Bänziger, T., and Dornfeld, D. (2014). "Development of a Machine Tool Platform to Support Data Mining and Statistical Modeling of Machining Processes," *Proceeding of MTTRF 2014 Annual Meeting*, San Francisco, CA

Kim, K., Lee, D. and Essa, I. (2011). "Gaussian Process Regression Flow for Analysis of Motion Trajectories," *Proceeding of IEEE ICCV*.

MTConnect Institute (2014). MTConnect v. 1.3.0, http://www.mtconnect.org/downloads/standard.aspx.

Neal, R. M. (1996). Bayesian learning for neural networks, Springer-Verlag, New York.

Neugebauer, R., Denkena, B. and Wegener, K. (2007). "Mechatronics Systems for Machine Tools," *Annals of the CIRP*, 56, pp. 657-686.

Nguyen-Tuong, D., Seeger, M. and Peters, J. (2009). "Model learning with local Gaussian process regression," *Advanced Robotics*, 23(15), pp. 2015-2034.

Park, J., Bhinge, R., Biswas, N., Srinivasan, M., Helu, M., Rachuri, S., Dornfeld, D. and Law, K. (2015). "A generalized data-driven energy prediction model with uncertainty for a milling machine tool using Gaussian Process," *ASME 2015 International Manufacturing Science and Engineering* Conference, Charlotte, NC

Rasmussen, E and Nickisch, H. (2013). Gaussian process regression and classification toolbox version 3.2, downloaded from: http://www.gaussianprocess.org/gpml/code/matlab/doc/.

Rasmussen, C., and Williams, C. (2006). "Gaussian Process for machine learning," MIT Press.

Teti, R., Jemielniak, K., O'Donnell, G. and Dornfeld, D. (2010). "Advanced monitoring of machine operations," *CIRP Annals-Manufacturing Technology*, 50, pp. 717-739

U.S. Energy Information Administration (EIA). (2014). AEO2014 Early Release Overview, Report Number: DOE/EIA-0383ER.

Vijayaraghavan, A., Sobel, W., Fox, A., Dornfeld, D. and Warndorf, P. (2008). "Improving Machine Tool Interoperability Using Standard Interface Protocols: MTConnect," *Proceeding of 2008 International Symposium on Flexible Automation*, Atlanta, USA.

Vijayaraghavan, A. and Dornfeld, D. (2010). "Automated Energy Monitoring of Machine Tools," *CIRP Annals – Manufacturing Technology*, 59, pp. 21-24.

Wang, J.M., Fleet, D.J and Hertzmann, A. (2008). "Gaussian Process Dynamical Models for Human Motion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), pp. 283–298.

Willmott, C., and Matsuura, K. (2005). "Advantage of the Mean Absoute Error over the Root Mean Square Error (RMSE) in Assessing Average Model Performane," *Climate Research*, 30, pp. 79-82.