IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 20 January 2023, accepted 10 February 2023, date of publication 22 February 2023, date of current version 6 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3247430

RESEARCH ARTICLE

Time-Series Forecast-Based Endpoint Thickness Compensation for Thinning of Sapphire Wafer

YU-KUN LIN, (Student Member, IEEE), AND BING-FEI WU, (Fellow, IEEE)
Institute of Electrical and Control Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan

Corresponding author: Bing-Fei Wu (bwu@nycu.edu.tw)

This work was supported in part by the National Science and Technology Council under Grant MOST 111-2218-E-A49-032-MY4, and in part by the Mechatronics System Research Laboratory of the Industrial Technology Research Institute under Grant M353CH1210.

ABSTRACT The grinder consults measurement modules to control the endpoint thickness of the sapphire wafer thinning process, but these modules have shortcomings. We propose a forecast-based endpoint thickness and online error compensation approach for grinding hard, brittle material. We leverage forecasts to improve grinding efficiency, and address the shortcomings of conventional probe gauges, which can be used only when grinding is paused, as well as those of contact gauges, which require correction from time to time due to wear. We construct a multi-signal time-series forecast model and compare the prediction performance using features from various signals. We implement the forecast model in a vertical grinder for actual grinding to evaluate the effectiveness of the proposed approach. The results show that online error compensation reduces contact times for the contact gauge and maintains precise thickness, greatly facilitating the wafer thinning process. The proposed measurement approach applies not only to wafer thinning but can also be extended to other grinding processes.

INDEX TERMS Wafer thinning process, acoustic emission (AE), tri-axial vibration, wafer endpoint thickness, time-series forecasting.

I. INTRODUCTION

Wafer thinning is a crucial front-end process for manufacturing wafers, and is considered shape processing rather than surface processing. As such, the whole wafer is directly scrapped if the processing fails to meet the quality requirements. Sapphire is a hard, brittle material used to make LED substrates. In recent years, wafer demand has dramatically increased in the industry due to the proliferation of mini-LED and micro-LED products, and increased production capacity is an urgent problem for the industry. Product quality requirements for the wafer thinning process are endpoint thickness, flatness, and surface roughness [1].

Wafer thinning is divided into free abrasive lapping and the fixed abrasive grinding. Although polishing is also free-abrasive, it is not classified as thinning. The free abrasive for lapping is a diamond slurry, which is not fixed on the carrier and produces excellent surface roughness but poor machining efficiency [2]. However, its grinding quality is gradually surpassing that of lapping because the precision and manufacturing technology of grinding wheels has improved, resulting in considerably faster processing times as well as the greatly improved wafer surface roughness and flatness [3].

Since a diamond grinding wheel is sintered with diamond grains and metal binders, the sintered surface is distributed with many abrasive grains similar to small cutting tools and pores that load chips [4]. Furthermore, to grind hard, brittle materials, a grinding wheel must be chosen with good self-sharpening ability and suitable grinding parameters; nevertheless, good sharpening ability implies a high wear rate for the abrasives. Nevertheless, as diamond wheels are expensive, it is not desirable for the abrasive to be consumed too quickly. Strong abrasive self-sharpening ability results in increased manufacturing costs, whereas inadequate self-sharpening ability results in wafer cracks or broken wheel abrasives. The grinding wheel's tilt angle with respect to the workpiece is another significant factor in the specific energy of grinding and wafer flatness. The configuration and

The associate editor coordinating the review of this manuscript and approving it for publication was Ehab Elsayed Elattar.

coordinates between the wheel and workpiece have been defined to investigate the effect of the angular deviation of grinding on the surface profile of the ground wafer [5]. The wheel is more likely to shed abrasives under immense grinding energy.

Grinding is a complex machining process in which many variables cannot be controlled or measured. In addition, disturbances in the grinding environment such as the splashing of cooling water, chips, abrasive particles, noise, and vibration from the spindle operation and grinding. Such noises increase the difficulty of online monitoring [6]. Thus, indirect signals such as acoustic emission (AE), vibration, current, force, and power are used to monitor and evaluate status conditions that characterize the grinding wheel and workpiece [7].

In general, previous studies monitor the workpiece and grinding wheel during the grinding process to improve the workpiece quality by analyzing the status of the workpiece. However, there is a paucity of research on improving the endpoint thickness accuracy when thinning. For sapphire wafers, endpoint thickness error affects the processing time of fine polishing. The more accurate the endpoint thickness, the faster fine polishing is. In general, grinders use a digital probe gauge (PG) to measure wafer thickness, yet during such measurement grinding must be halted; moreover, the last feed must proceed at a slow speed to minimize thickness error. Alternatively, the grinder can use a contact gauge (CG) module to measure wafer thickness during grinding. Since sapphire is harder than the gauge head, long-term use wears down the gauge and reduces the measurement accuracy; moreover, the sapphire surface can be scratched and marked if the gauge does not properly break away from the sapphire.

We propose using the CG module with a forecast model for endpoint thickness compensation for sapphire wafer thinning to improve on the traditional PG's need to stop grinding for measurement and to address the CG's problem with wear and scratches. We implement this approach on the grinder, and compare the results using the PG and CG modules.

We propose a thickness compensation forecast model for wafer thinning. We extract three kinds of time-series signal features and compare the performance of these features using six deep forecast models. We use windowed time series data with different periods to verify the forecast ability. We implement the proposed forecast model into the grinder and use it for sapphire wafer thinning. Finally, we compare the wafer endpoint thickness using the proposed approach with that of a conventional grinder. Our main contributions are as follows:

1) We propose an forecast-based endpoint thickness and online error compensation approach for grinding hard, brittle material. The forecasts are leveraged to improve the grinding efficiency and correct the conventional PG module's disruptive measurements; it also reduces the frequency of corrections due to CG wear. Furthermore, we design and validate an endpoint thickness compensation method to facilitate practical real-time applications during grinding.
2) We construct a multi-signal time-series forecast model and compare its prediction performance with the features of each signal. During the grinding process, preprocessing and feature extraction methods for each signal—AE, vibration, and grinder current—are based on real-time feature analysis. In addition, we compare the performance between each characteristic signal and the forecast model using six different time-series deep forecast models. The Res-LSTM model exhibits superior performance in predicting endpoint thickness compensation values.
3) We implement the forecast model in a vertical grinder for actual grinding to verify the effectiveness of the proposed approach, and compare the endpoint thickness accuracy between this approach and that of traditional grinders. The results attest the effectiveness of the approach.

The following sections are arranged as follows. Section II reviews monitoring technologies for the grinding process, Section III describes the proposed endpoint thickness compensation for the CG module with the forecast model, and Section IV describes forecast models based on characteristic signals and state data acquired in the grinding experiments and compares the performance using different features and models. Section V explains how the compensation approach is implemented in the grinder and compares the results of thinning experiments. Section VI concludes.

II. GRINDING PROCESS MONITORING

As grinding quality affects the success or failure of subsequent processes and product values, it is almost always performed by experienced personnel. Therefore, grinding process monitoring has long been a popular topic in machining optimization, whether for metals or hard, brittle materials. Scholars have identified various signals by which to monitor the grinding process, seeking to identify the significant characteristics at the moment of state change for each signal [8].

AE signals are commonly used to monitor the grinding process. Since the surface of the grinding wheel contains a distribution of tiny, randomly oriented grits and pores, an AE signal with superior sensitivity and a wide frequency band is suitable. Therefore, AE is used to visualize the surface topography of the grinding wheel during the dressing process; this approach facilitates evaluation of the wheel condition when dressing and reduces abrasive waste when overdressing [9]. Likewise, AE can be used to determine whether the wheel surface condition is sharp or dull, which improves the workpiece quality and reduces defects. The wheel surface is dulled when cutting chips fill the surface pores or the grits are worn down, thus losing the ability to grind [10]. Next, the bonder grade of the abrasive can be distinguished by the AE signal; in general, for diamond wheels the bond grade is distinguished by calculating the wear ratio after grinding [11], [12], [13]. In addition, the AE signal is often used to monitor workpiece conditions by identifying defects such as workpiece surface

burning, chatter vibration marks, and roughness [14], [15], [16]. Moreover, AE signal is often mentioned in other related studies, such as non-destructive testing of material broken and damage mechanisms with loading.

During grinding, a worn wheel is exhibits vibration as well as increased grinding force and specific energy of grinding; to identify grinding conditions, many studies use multi-sensors such as accelerometers, dynamometers, current transducers, power meters, and microphones [17], [18], [19].

To extract features, the time-domain signal power of AE and the grinding force are proportional, and frequency-domain features are more effective than time-domain features for evaluating surface conditions [20]. Thus, feature extraction involves time-frequency analytical methods such as the short-time Fourier transform (STFT) [21] and the discrete-time wavelet transform (DWT) [22]. Although novel analytical methods such as ensemble empirical mode decomposition (EEMD) and the Hilbert transform [23] are also used, such transformation algorithms are characterized by massive calculations that preclude their use in real-time systems. Hence, for real-time system applications, the performance of feature extraction depends not only on how precisely the features represent the grinding state but also on the time needed to calculate the features.

Note that the CNN model's convolutional layer cannot be used to simultaneously extract multiple frequency signals, in particular AE and vibration signals [24]. Also, the sampling frequency of the raw AE signal is higher than that for other signals; it is difficult to integrate features with such quantitative differences and ensure that the resultant model converges [25]. Therefore, related approaches generally do not attempt to learn AE signals in combination with other signals. Moreover, convolution for such a large number of raw AE signal data points takes a long time for 1-D as well as 2-D convolution kernels, which makes it challenging to apply this to real-time systems.

III. METHODOLOGY

Here we describe the configuration, the grinding wheel, the parameters, the external control method, the running state data of the vertical grinder, and the experiment for extracting time-series data, as well as the time-series data acquisition module, the sensors, and the feature extraction algorithms for the characteristic signals: AE, tri-axial vibration, and tri-axial motor current.

A. EXPERIMENTAL SETUP

The vertical grinder grinds the workpiece to produce excellent flatness via a stable and precise feed system. A pair of upper and lower non-coaxial rotating shafts rotate simultaneously to ensure a flat surface. The upper shaft is equipped with a cup grinding wheel, which is called the grinding wheel axis (S1); installed on S1, the feeding axis (Z) stabilizes the grinding wheel and supplies micro-feeding. The lower shaft has a porous ceramic vacuum chuck table that holds the ceramic block that fixes the workpiece, called the workpiece axis (S2). During grinding, the speed of S1 is typically faster than that of S2 during grinding, and the two axes rotate in different directions. To grind all of the workpiece, including the center, the abrasive distributed on the outer circumference of the grinding wheel must traverse the center of the workpiece. When grinding, a large amount of cooling water must be sprayed to remove the workpiece and wheel chips.

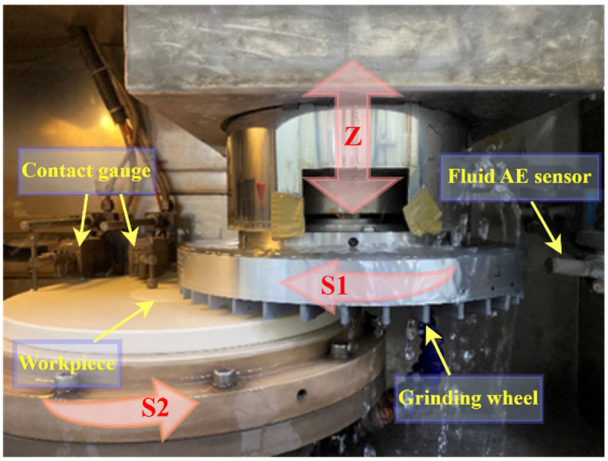

FIGURE 1. Vertical grinder and sensors.

The grinder has a contact gauge (CG) module to measure the wafer thickness during the grinding process: two sensors measure the height of the reference plane (ceramic block) and the workpiece (wafer) simultaneously, and the difference between the two is taken as the thickness of the wafer to be measured.

In this study, the grinding wheel uses a 400 grain diamond abrasive with a metal bonder that ensures a grinding ratio of 0.05:1 between the wheel and the sapphire wafer, the most commonly used configuration in the sapphire wafer thinning process. The workpiece is a 4-inch sapphire wafer bonded to the center of the ceramic block with wax. The grinding path coordinates mirror those defined previously [5]; the tilting angle of the grinding wheel $+\alpha$ is 0.01 degrees, and $+\beta$ is 0.01 degrees. The grinding wheel tilting angle not only determines the workpiece topography but also affects the hard material's grinding force.

The experimental test configuration of the vertical grinder and sensors is shown in Figure 1. The sensors cannot be installed in the work area because the vertical grinder has two rotation axes; therefore, we use noncontact or indirect measurements. As the fluid AE sensor receives the signal through cooling water, it is not affected by the configuration of the working area; however, it must be installed where the cooling water sprays onto the grinding area. The headstock is closest to the grinding wheel, so it is the best place to receive tri-axial vibration signals from the grinding wheel. The accelerators are fixed to the headstock using magnets. The current transducers are mounted on motor drivers on the S1, S2, and Z axes to measure changes in current during grinding.

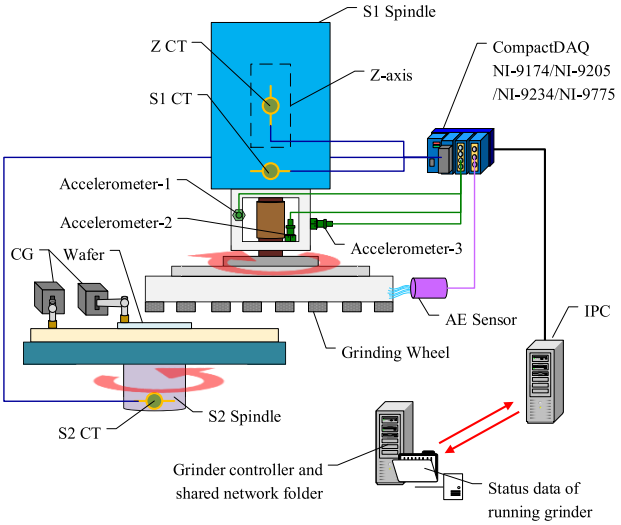

FIGURE 2. Architecture of signal and grinder state data transmission.

TABLE 1. Experimental grinder parameters.

Parameter	Specification
Spindle (S1) speed	1,000 rpm
Chuck (S2) speed	60 rpm
Feed rate (Z)	0.06 mm/min
Endpoint thickness	0.2 mm
Cooling flow rate	15 L/min

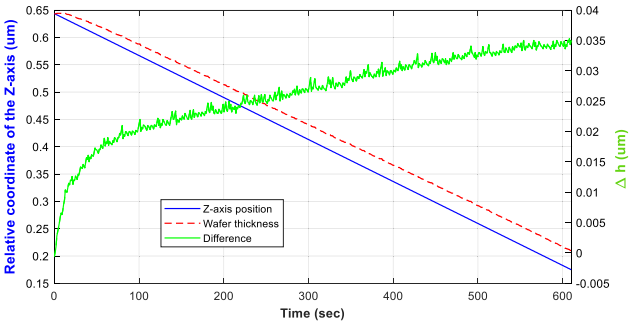

FIGURE 3. Wheel feed path (blue line) and wafer thickness (red dotted line) during grinding. The Δh value is the difference value between the two (green line).

The drivers are installed in the control cabinet behind the grinder.

The architecture of the signal and grinder state data transmission is shown in Figure 2. The data acquisition (DAQ) module is the National Instruments (NI) CompactDAQ (NI-9174), the AE signal is captured by the NI-9775 module, the tri-axial vibration signals are captured by the NI-9234 modile, and the tri-axial current signals are captured by the NI-9205 module. The grinder's state data can be accessed by other devices through a shared network folder.

Additionally, we extracted and processed the characteristic signals and trained the time-series models using Python 3.6.8 with TensorFlow 2.5.0 on a Ubuntu 18.0.4 system, and used Matlab for signal analysis. We used a NVIDIA GeForce GTX3090 GPU with CUDA 11.3 and cuDNN 8.1.0.

B. GRINDING EXPERIMENT TO EXTRACT TIME-SERIES DATA

To obtain the characteristic signal when grinding the sapphire wafer, we conducted 14 grinding experiments; we collected AE, tri-axial vibration, and tri-axial current signals as well as the state data of the running grinder. The grinding workpiece was a 4-inch sapphire wafer fixed to a ceramic block center using wax. The total thickness variance (TTV) after waxing the wafers ranged from 660 to 680 μm because the initial wafer thickness and waxing technique differed.

Table 1 lists the grinder's experimental parameters. The spindle speed (S1), chuck speed (S2), and feed rate (Z) are production-line process parameters for the #400 grinding wheel, and the endpoint thickness is the wafer thickness after grinding.

The wheel exhibits good grinding performance when the S1 (motor driver) current loading is 37% during processing; in contrast, the S1 loading is 40% when the surface condition of the wheel is loaded with chips or the abrasive is dull. Therefore, the grinding wheel must be dressed when the wheel surface is loaded with workpiece chips to ensure good grinding ability. Poor grinding ability causes wafer edges to crack or wheel abrasive to break. For this reason, the grinding wheel was dressed six times during the course of the experiment with a dressing depth of 0.2 um each time.

Figure 3 shows that the grinding wheel feed follows the parameter settings (blue line) and the changing thickness of the wafer being thinned (red dotted line) during grinding; the green line shows the difference between the two. The data in the figure was collected from the 14th experiment, and the X-axis represents the time series of the grinding process. The relative coordinate system of the Z-axis on the left axis of the figure is based on the machining coordinate system, which is based on the plane of the ceramic block. As this plane also served as the CG reference plane, when the wheel position is equal to the thickness of the wafer, this is taken to mean that the grinding wheel is ground to the wafer. The difference Δh, as expressed by (1), between the Z-axis position P_Z and the wafer thickness T_W (DBZPT) is caused by wheel frame deformation or abrasive fallout or dullness of the grinding wheel. The right axis (the green line) represents the difference between the blue and red lines, and increases over time. These changes in difference values varied for each experiment; the grinding factors that determine this change are many and complex, in particular the grinding wheel, which is one reason for the unpredictable grinding feed.

$$\Delta h = T_W - P_Z \tag{1}$$

DBZPT is the feed compensated by the grinding wheel, which means that to thin the wafer to a given thickness, the grinding wheel must feed a corresponding amount at the

specified thickness. Because the feed slope is known, the forecast DBZPT predicts how much further the Z-axis feed must move to thin the wafer to the specified thickness after a few seconds: this is thus the ground truth of this study's time-series forecasting models. We use median filtering (n=7) to preprocess the data before model training to reduce data noise.

C. ACOUSTIC EMISSION SIGNALS

The high frequency and high sensitivity of the AE signal make it suited for monitoring the grinding state; it eliminates noise in the grinding environment and provides information about damage to the wheel and workpiece. Many studies use the AE signal to identify workpiece surface burn and wafer edge cracks during grinding, as well as the grinding wheel's abrasive grade and surface condition. The AE signal of the grinding process contains much information about the state of the wheel [16].

Here we used a fluid AE sensor produced by the Balance Systems Company, which enables noncontact transmission of signals through water and operates at frequencies from 10k to 1MHz. Thus, according to the Nyquist–Shannon sampling theorem [14], the AE signal is captured by NI-9775 at a frequency of 2.5 M/s.

$$Y_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |x_n|^2} \tag{2}$$

$$\sigma_{STD} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (x_n - \mu)^2} \tag{3}$$

$$Y_{variance} = \frac{1}{n-1}\sum_{n=1}^{N} (x_n - \mu)^2 \tag{4}$$

The root mean square (RMS), as expressed by (2), is a statistical method for time-domain signal processing. Analysis of the probability distribution of continuous variables mainly concerns the data's central tendency or degree of dispersion; RMS is a measure of the former [11]. RMS is thus a representative feature of AE signals and is often used to assess the magnitude of the signal power. Many studies confirm that the RMS of the AE signal is positively correlated with the grinding force, and is thus often used to detect chip loading or dulled abrasives in the grinding wheel. Standard deviation (STD) and variance, in turn, as expressed by (3) and (4), are used to extract features that characterize statistical dispersion [9].

Figure 4 shows the characteristic frequency band distribution of the power spectrum (determined by the Welch method) of the AE signal for various grinding durations, where the characteristic bands at which the condition of the wheel surface is observed to change during grinding are the 312.5 to 560 kHz segments and 600 to 900 kHz segments. The ratio of power (ROP), expressed as (5), is used to calculate the power ratio at a particular frequency band in the spectrum. It is known that grinding wheels with different bonding grades exhibit different characteristic frequency band distributions in the spectrum; thus ROP yields excellent performance in identifying grinding wheel grades [12]. After observing the

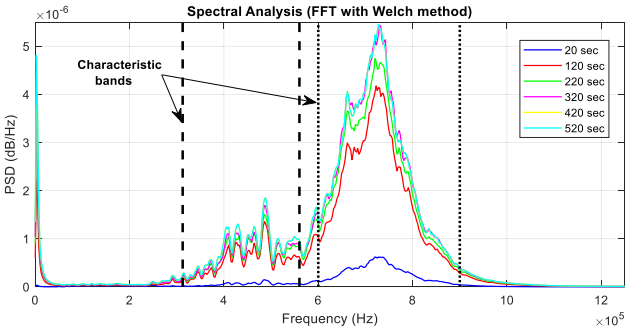

FIGURE 4. Power spectrum density of AE signals at various grinding durations.

spectrum in the experiments, the frequency band was divided into ten segments from 100 kHz to 1 MHz, which were then used as features.

$$X_{ROP} = \frac{\sum_{k=n_1}^{n_2} |x_k|^2}{\sum_{k=0}^{N-1} |x_k|^2} \tag{5}$$

The high sampling rate of the AE signal makes it easy to miss details as the wide frequency distributes characteristics during frequency domain analysis. The discrete wavelet transform (DWT) decomposes feature bands to denoise them or extract features from them [26]. Here we use Daubechies 10 (Db10) as the mother wavelet ψ [13].

Given the characteristic frequency bands, we selected variables (a, b) of wavelet packet decomposition as (1, 2) and (2, 2), where the packet (1, 2) coefficients characterized the 600 to 900 kHz band and the packet (2, 2) coefficients from 312.5 to 560 kHz. The statistical features listed in Table 2 were extracted from packet (1, 2) and (2, 2) coefficients such as STD, mean absolute deviation (MAD), 2-norm, skewness, kurtosis, peak factor, impulse factor, and so on [27]. These features contain much characteristic information, and are easy to use and calculate.

There are thus 3 time-domain features, 9 frequency-domain features, and 22 features from two wavelet packets; the above calculation features a total of 34 items, which are the time-series features of the AE signal for a single second.

D. VIBRATION SIGNALS

The accelerometer converts the vibration of the mass block—its displacement—into an electronic signal; accelerometers are small and low-cost, and can be used for most vibration measurement applications. The vibration signal is often used to monitor grinding conditions such as wheel flutter, chip loading, and dullness, and to predict workpiece surface roughness [28]. Here we used the Benstone 786A accelerometer with a sensitivity of 100 mV/g produced by Wilcoxon Research; the sensors were magnetically fixed to the headstock. Tri-axial grinding wheel vibration signals were captured by NI-9234 modules at a frequency of 25,600 Hz.

When extracting the time-domain features, we also calculated the autocorrelation function, as expressed by (6). The

TABLE 2. Statistical features.

Parameter	Formula
STD	$\sigma = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y_k - \mu)^2}$
MAD	$Y_{MAD} = \frac{1}{N}\sum_{k=1}^{N}\lvert y_k - \mu\rvert$
L2-norm	$\lVert Y\rVert_{L_2-norm} = \sqrt{\sum_{k=1}^{N}\lvert y_k\rvert^2}$
Skewness	$Y_{Skewness} = \frac{1}{N}\sum_{k=1}^{N}(y(k) - \mu)^3$
Kurtosis	$Y_{Kurtosis} = \frac{1}{N}\sum_{k=1}^{N}(y(k) - \mu)^4$
Peak factor	$Y_{peak\ factor} = 0.5 \times (\max(y(k)) - \min(y(k))$
Impulse Factor	$Y_{Impulse\ factor} = Y_{peak\ factor}/mean(abs(y(k)))$
Shape Factor	$Y_{shape\ factor} = Y_{rms}/mean(abs(y(k)))$
Crest factor	$Y_{crest\ factor} = Y_{peak\ factor}/Y_{rms}$

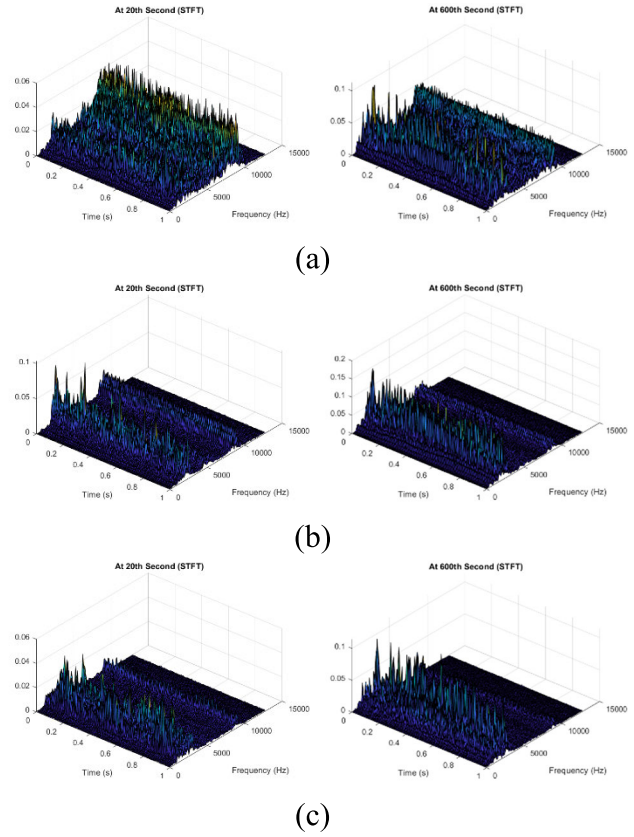

FIGURE 5. Tri-axial vibration spectrograms for 20 and 600 seconds on the (a) x-axis, (b) y-axis, and (c) z-axis, respectively.

autocorrelation function compares a random variable series with itself to compute the signal correlation between different time points. This is often used in signal processing to detect periodicity and is a significant feature describing random variables in the time domain.

$$\mathrm{Y(k)} = \frac{E\left[(X_i - \mu_i)(X_{i+k} - \mu_{i+k})\right]}{\sigma^2} \tag{6}$$

The short-time frequency transform (STFT) is utilized to observe how the frequency of a nonstationary signal changes over time [29]. The STFT of the signal was determined by sliding a window of length N over a signal (here N=256) and calculating the discrete Fourier transform (DFT) of the window data. Following most studies, the window moves over the signal at intervals of 1/2 samples. The DFT of the window data is merged, and the one-dimensional time series signal is transformed into a two-dimensional matrix representing time and power spectral density, which is the spectrogram [30]. Figure 5 shows the tri-axial vibration spectrogram during grinding when starting (20 seconds) and when almost finished (600 seconds). In the tri-axial spectrograms, the power spectral density of the x-axis is distributed the widest, with the maximum power in the high-frequency regions, whereas the power of the z-axis is concentrated in only a few bands, with a power density generally lower than that of the y-axis; this reflects the rigidity of the S1 spindle in the tri-axial direction. We also compared the spectrograms at the beginning of grinding (a) and at the end (b). After a period of grinding, the tri-axial power density in high-frequency regions reduces, which shows the usefulness of the vibration signal's time-frequency characteristics as the wheel surface goes from sharp to dull.

The spectrogram was used to calculate the full-time mean value of the segmented spectra [31]. After selecting the frequency, all values along the time axis were summed and averaged to obtain the full-time mean for the specific frequency band as

$$\mu_{ss}(\mathrm{k}) = \frac{1}{N}\sum_{m=f_{lk}}^{f_{uk}} \lvert S(m, n)\rvert, \tag{7}$$

where $\mu_{ss}(\mathrm{k})$ is the full-time mean of the k-th band, N is the number of matrix elements in the specific band, and f_{lk} and f_{uk} are the upper and lower frequencies of the k-th band, respectively. The STFT matrix is represented by S, in which m is a frequency vector and n is a time vector. When the full-time mean value is calculated, n is the number of time axis columns. The tri-axial spectrograms here were segmented into 13 bands to calculate the features of the full-time mean. Hence, the vibration signal's features include 12 statistical parameters and 13 full-time mean values. As the above calculation features 25 items for each axis, the tri-axial features comprise a total of 75 items.

E. MOTOR CURRENT SIGNALS

In the industry, the motor loading is used to reveal wheel conditions and poor grinding performance. We thus installed a current transducer (CT) on the frequency converter of the tri-axial motor to extract the current signal [25]. We used LEM's HTR 100-SB current transducer and measured the motor loading changes during grinding. The current signals were captured via a NI-9205 module at a frequency of 25,600 Hz, and we used the absolute mean value: the equivalent current value from the raw data of CT. Figure 6 shows the changes in motor current for the S1 and S2 axes during grinding: the motor loads of the S1 and S2 axes increase as the grinding wheel loses its ability to grind. Therefore,

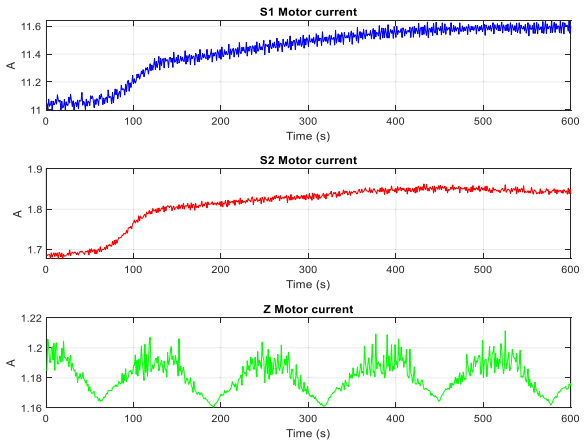

FIGURE 6. Current variation of tri-axial motor during grinding.

the tri-axial motor current's features are 3 equivalent current values.

F. REAL-TIME GRINDER DATA AND EXTERNAL CONTROL

The vertical grinder in this study outputs the running status parameters during processing. Every second, the real-time running status parameters are output to a CSV file, including the S1 rotation speed, S1 current loading, S2 rotation speed, S2 current loading, Z feeding speed, Z current loading, Z relative coordinate position, and workpiece thickness, for a total of 8 parameters. When the grinder commences grinding, a file containing the running status parameters is created in a shared network folder so that other devices can access the parameters for analysis.

Likewise, the vertical grinder can be externally controlled via a CSV parameter file during grinding. The command parameters in this file include motions and motion parameters. Motion control parameters are used to advance, stop, and retract the Z-axis motion; to change the rotation speed of the S1-axis and the S2-axis; and to change the Z-axis feed rate, feed direction, and retracting distance. An external device can send a command file to control the grinding process via the shared network folder when the grinder is operating.

IV. TIME-SERIES FORECASTING

A. TIME-SERIES DATA WINDOWING

We use a sliding window so that data with time sequence relationships can be entered into the forecasting model and predicted in the time sequence, where the window slides over the time sequence data and splits it into past and future time steps.

The window architecture used in this study for DBZPT compensation is illustrated in Figure 7. Inputs represent past data, which is the data of each characteristic signal and running state in the past p seconds, and the DBZPT at each time point is the ground truth (learning target) for model training, which is the labels. The predictions represent future data, predicting the upcoming p seconds of the DBZPT sequence [32].

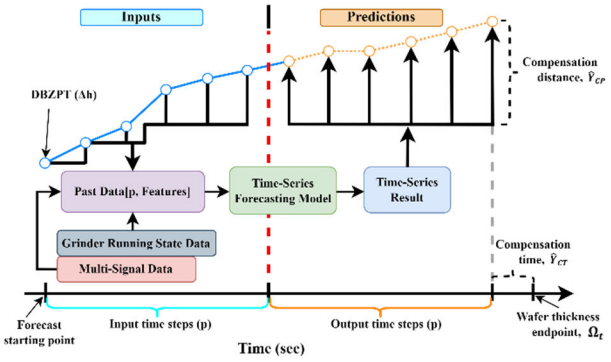

FIGURE 7. Time-series data windowing.

For multiple time-step prediction, the model must learn one period of past data to predict an equal-length period of future data. Therefore, the inputs and predictions are composed of a period of length p: input time-series data of length p, and the model predicts future time-series data of length p.

We forecast the DBZPT sequence to predict the grinding wheel compensation feeding distance in p seconds. As the Z-axis feed rate is constant during grinding, the compensation distance is converted to a compensation time to finish thinning the wafer to the endpoint thickness. The forecasting series data focuses on the p-second results since compensation is performed for p seconds in the future.

B. MULTI-STEP FORECAST MODELS

To forecast the DBZPT sequence data, we investigate six different styles of multi-step neural network models. We extract the characteristic signals for each second and combine these using extraction algorithms. We arrange the features according to the time sequence and experimental order as a dataset, window the data, and use it for network training. We thus train six models on the same dataset with mean square error (MSE) loss, the Adam optimizer, and the mean absolute error (MAE) training metric [33].

1) LINEAR MODEL

The linear model uses a single hidden layer between the input and output layers, and the hidden layer size is set to p, the length of a sliding window. No activation is used for the hidden layer [34].

2) MULTILAYER PERCEPTRON (MLP) MODEL

After evaluating the model architectures with the collected data, the model architecture of the fully connected (FC) and the remaining layers is $128 \times 128 \times 128 \times 64 \times 64$, using the ReLU activation function. No activation function is used for the regression MLP in the output layer, and dropout neurons are inserted after FC layers 2 and 3 for regularization, which is 0.2 [35].

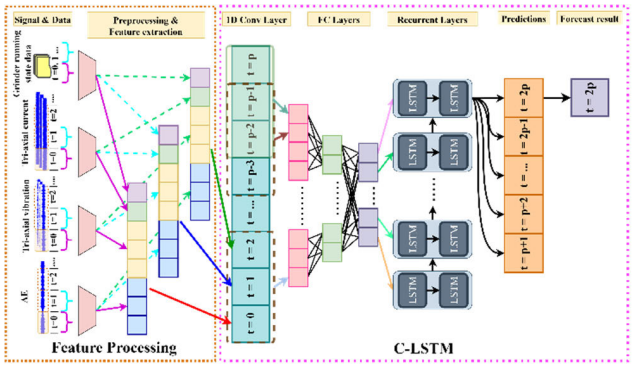

FIGURE 8. Multiple input C-LSTM network.

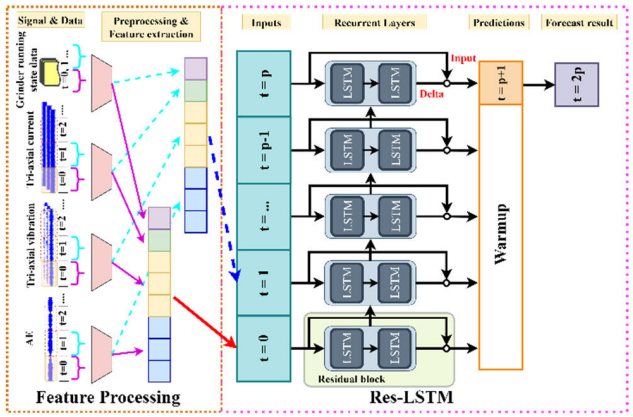

FIGURE 9. Multiple input Res-LSTM network.

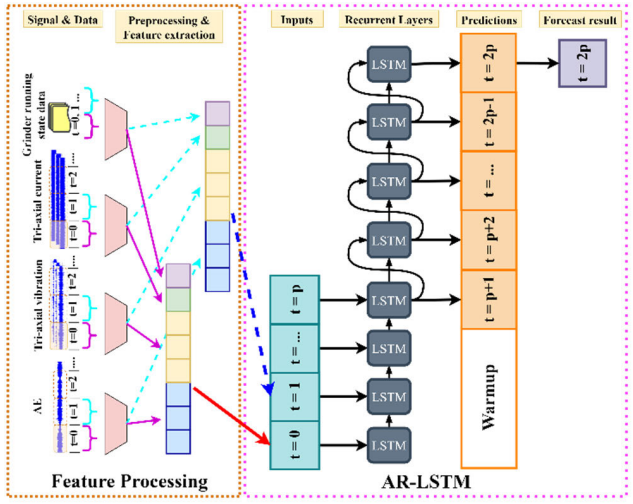

FIGURE 10. Multiple input AR-LSTM network.

3) LSTM MODEL

For the RNN models, two LSTM are stacked, each with 128 LSTM units. The first LSTM produces output for each input, whereas the second LSTM returns output only at the last step [36].

4) C-LSTM MODEL

A 1D convolution layer combines the FC and recurrent layers (C-LSTM) to build a neural network, as shown in Figure 8. The 1D convolution layer has multiple kernels for scanning sequence data for feature extraction because the kernels traverse numerous time steps to obtain context information. In this way we extract the correlation between sequences and efficiently process large amounts of time-series data. In the C-LSTM architecture using 1D convolution operations with 32 different filter kernels of 1×4 size and 1 stride, all network activation functions were ReLU. The FC layer sizes were both 64, and two LSTM layers of size 64 were stacked in the recurrent layers. Because this was a single-shot network, the recurrent layers returned the output at the last step [37], [38].

5) RES-LSTM MODEL

The residual LSTM (Res-LSTM) model uses shortcut connections to build the RNN model, as shown in Figure 9. A residual block with a shortcut connection adds its input to its output, which is then the input plus the residual (Delta). Termed residual learning, this approach allows blocks to avoid degradation during feature transfer, even if it does not learn new features. Moreover, during initial network training, its weights are all zero; residual learning causes the network to output the identity function, which helps the network converge faster. In the Res-LSTM architecture, two LSTM layers of size 64 are stacked in the residual block. Again, as this is a single-step network, the network returns an output for each input. After the warmups, the last output is used as the prediction result for the $2p$-th second [39].

6) AR-LSTM MODEL

In the autoregressive model, the prediction output of each time step is fed back to itself to predict the next time step, as shown in Figure 10. Therefore, each output makes a prediction based on the previous one, as in the classic RNN architecture. We constructed an autoregressive LSTM (AR-LSTM) model and used single-step output because the network decomposed prediction into individual time steps for the recurrent next step, using 128 LSTM units. The final prediction output was taken as the forecast for the $2p$-th second [40].

C. PERFORMANCE COMPARISON BETWEEN CHARACTERISTIC SIGNALS AND MODELS

For the various forecasting models and features, we evaluated the performance using grinding state data. We compared features of different signals to determine which signals were the best predictors of grinding thickness compensation. The features included 34 features from the AE signal, 78 items from the tri-axial vibration signals and motor current values, and 8 items from the grinder running state, where 6 were training model features such as the S1 speed, S1 motor loading, S2 speed, S2 motor loading, Z motor loading, and Z relative

TABLE 3. Results for different signals and models.

Model	Signal					
	AE		Vibration & current		Combined	
	Vali.	Test	Vali.	Test	Vali.	Test
Linear	0.071	0.044	0.090	0.084	0.098	0.088
MLP	0.093	0.054	0.166	0.099	0.163	0.113
LSTM	0.081	0.057	0.104	0.084	0.094	0.076
C-LSTM	0.081	0.059	0.097	0.082	0.093	0.074
Res-LSTM	0.059	0.042	0.071	0.054	0.048	0.042
AR-LSTM	0.097	0.071	0.139	0.125	0.093	0.076

coordinate position. The final DBZPT was the training data label. In this way, we trained models using features from the various signals to represent the grinding state and compare which signal features yielded better prediction performance.

Table 3 shows the cross-validation results separated by signal and model in terms of MAE between the sequence forecasted by the model and the ground-truth sequence. In the dataset from 14 complete grinding experiments, the time-series data from the first to the twelfth experiments was concatenated as the training data, and the last two parts were used for validation and the testing set, respectively. Res-LSTM outperforms the other models.

We also evaluated the results using the combined signals. The AE signal and the combined signal both yield good prediction results for the Res-LSTM model: they perform equally well on the test set but the combined signal is better on the validation set. This suggests that the combined signal provides more stable prediction results. Therefore, Res-LSTM and the combined signal were used to implement the thickness compensation system.

D. FORECAST PERIOD ASSESSMENT

We evaluated the prediction performance using different period lengths for the forecast model. Previous training models used a 60-second sliding window and required architecture design, hyperparameter fine-tuning, and feature comparison. Each experiment took about 600 seconds to complete one grinding run, as the feed rate was set to 0.06 mm/sec. Thus, we considered the input and prediction periods when deciding the p value for the sliding window.

Hence in this section we evaluate the model for different period lengths to determine whether the forecast error was within the allowable range (<7um), that is, the grinding accuracy error of a traditional grinder with the PG module. Figure 11 shows the MAE of six forecast models for 6 different period lengths. To prevent multiplier effects and ensure representative training results, we ensured that the period lengths were not multiples of each other. The results show that the Res-LSTM performance is still best; also, the prediction errors for different periods are nearly almost all within the design range, except for $p = 300$. Table 4 shows the prediction error of the Res-LSTM model on the test dataset: the forecast period should be set according to the process requirements, as long as the period is less than 180 seconds.

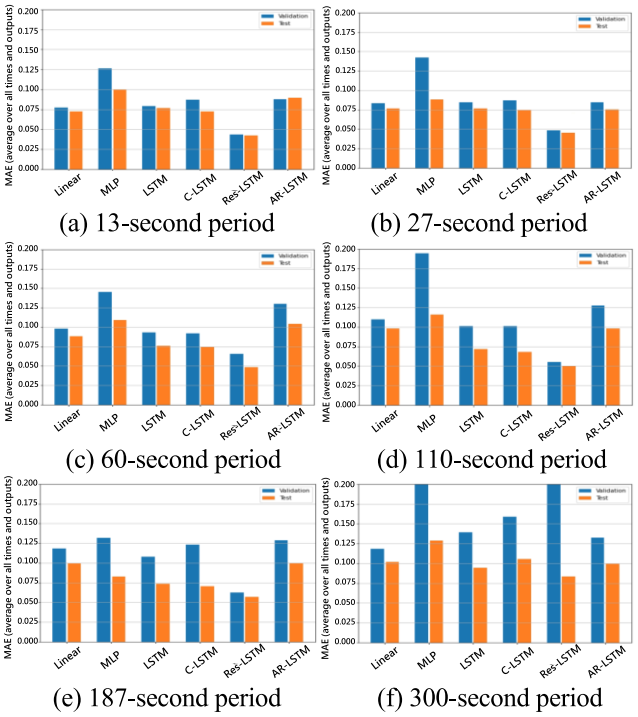

FIGURE 11. Results of 6 models at different period lengths: (a) 13 sec. (b) 27 sec. (c) 60 sec. (d) 110 sec. (e) 187 sec. (f) 300 sec.

TABLE 4. Prediction results for different period lengths.

Res-LSTM						
Period length (sec)	13	27	60	110	187	300
MAE	0.0421	0.0451	0.0484	0.0505	0.0572	0.0833

The forecast error (MAE) does not represent the actual thickness error value.

We thus sought to better understand the stability and accuracy of different period lengths with the forecast model. The endpoint compensation approach included an input period and a prediction period covered by a sliding window, as illustrated in Figure 7. We calculated the compensation distance ($\hat{Y}_{CP}$) according to the feed speed to estimate the time when endpoint thickness would be reached ($\hat{Y}_{CT}$) given the input and prediction periods, and then compensated for the thickness endpoint (Ω_t) using $\hat{Y}_{CT}$.

V. METHOD VERIFICATION RESULTS AND DISCUSSION

A. FORECAST MODEL FOR DBZPT COMPENSATION

We use characteristic signals and a forecast model for online DBZPT compensation. Algorithm 1 of the program builds on the IPC, which collects characteristic signal data and also transmits the state data to the grinder via a shared network folder, as shown in Figure 2. Therefore, the program contains functions for extracting the characteristic signals, collecting the grinder state data, forecasting the DBZPT, and sending commands to the grinder. Since two computers exchange the

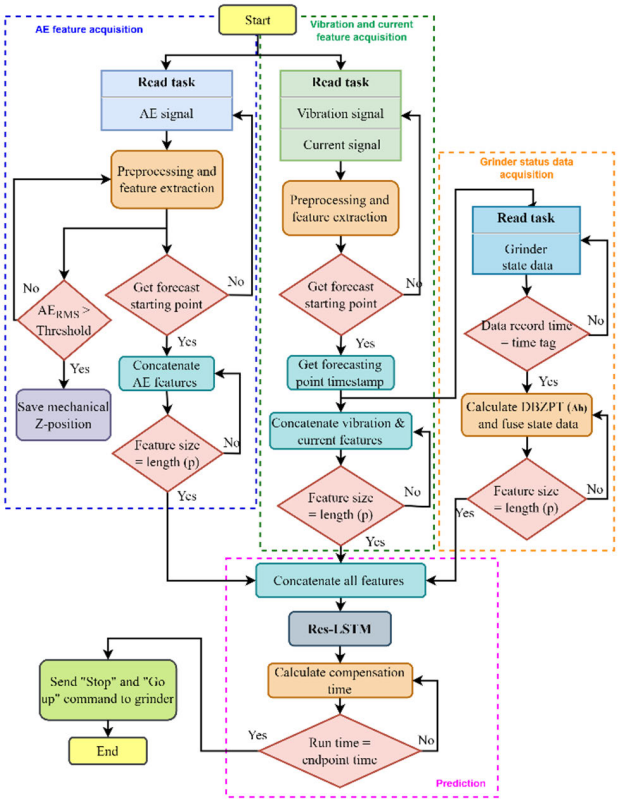

FIGURE 12. Proposed approach.

data, the system times of the computers must be synchronized before executing the program.

A flowchart of the proposed method is illustrated in Figure 12. The signal and task status data is divided into two characteristic signals and one set of grinding state data. As the state data was exported as a file, the state data was delayed, which necessitated further data alignment.

The forecast starting point in the program is not derived from the grinder state data, although the command is issued by the CG module and received by the DAQ NI-9205 module. When the CG module detects that the wafer thickness has reached the forecast starting point, the prediction module begins receiving each signal's features. When the input length has reached p (that is, the length of the input period, see Figure 7), the features are input to the time-series forecast model according to their timestamps. Thus we use the timestamps to find the file with the grinder signal data from the network folder and concatenate features with the same timestamp. The features for period p are then input to the time-series forecast model.

After Res-LSTM outputs the prediction, the resulting DBZPT compensation distance $\hat{Y}_{CP}$ is transformed to the compensation time $\hat{Y}_{CT}$. Thus, at time Ω_t, when the wafer thickness endpoint has been reached, the command file is used to stop feeding to the grinder. There is a 0.5-sec delay after the grinder receives the command to stop, but the effect of this delay on the resultant endpoint thickness is insignificant.

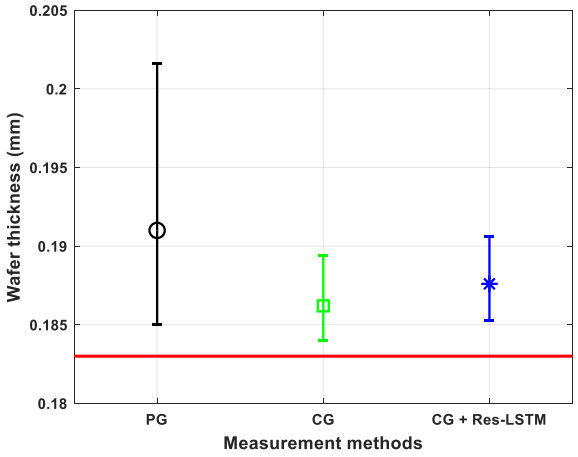

FIGURE 13. Sapphire wafer thinning results.

In the AE acquisition module, a function that determines the AERMS value was added to record the Z-axis position when the grinding wheel contacts the wafer to warn the user that the grinding wheel has touched the wafer.

B. EXPERIMENTAL RESULTS AND DISCUSSION

To verify the feasibility of the proposed approach, we used the proposed forecast compensation method in real-world wafer thinning experiments. Sapphire wafer thinning was executed to compare the endpoint thickness after grinding between the PG, CG, and CG+Res-LSTM grinder. Table 1 shows the grinder parameters.

Figure 13 shows the grinding results with three measurement approaches as the average, minimum, and maximum thickness distributions of PB (black line), CG (green line), and CG+Res-LSTM (blue line), respectively. The thick red line is the target thickness of 0.183 mm; the process must not fall below the target thickness. The CG module, the whole contact measurement, has the smallest error in endpoint thickness. The average thickness of the proposed approach is only 1 μm greater than that of the CG module, and the maximum and minimum distributions are similar. The PG module yields the worst performance.

For the PG experiment, the grinder used two feeding stages: coarse grinding with fast feeding at the beginning and then fine grinding with slow feeding. After the coarse grinding stage, grinding was paused to measure the thickness using the PG, and then the second stage of slow fine grinding proceeded to remove the remaining thickness. During this process, the wheel wear ratio was used to estimate the grinding feed compensation, which yielded a significant error in terms of workpiece accuracy. The grinder manufacturer provided 80 records of PG thickness data for sapphire wafer grinding. The wafer thickness endpoint (Ω) was 0.183 mm, the average thickness was 0.191 mm, and the upper and lower error limits were 0.2015 mm and 0.185 mm, respectively.

For the CG experiment, the grinder was set for a single infeed. The CG module has a measurement error of ± 3 μm [41]. Thus, although the target wafer thickness

was 0.183 mm, Ω was set to 0.186 mm to avoid excessive grinding. In addition, to prevent the CG from scratching the wafer, it was withdrawn before reaching the target thickness. The grinder manufacturer provided 64 records of CG thickness data for sapphire wafer grinding. The wafer thickness endpoint (Ω) was 0.186 mm, the average thickness was 0.1862 mm, and the upper and lower error limits were 0.1894 mm and 0.1840 mm, respectively.

For the proposed CG+Res-LSTM approach, we conducted 10 runs of validation experiments. The wafer thickness endpoint (Ω) was 0.186 mm, as for the CG approach. The sliding window's period (p) was set to 60 seconds, and we used the Res-LSTM forecast model. The threshold value of AERMS was 0.5. After grinding was completed, we waited 2 seconds for the grinding wheel to rise. The average thickness of the experiment was 0.1872 mm, and the upper and lower error limits were 0.1902 mm and 0.1852 mm, respectively.

As this study's forecast model and training data were all from the CG module, the experimental results are similar to the module results, as expected. However, as this model exhibits forecast error as well as delay during communication or file I/O, the average thickness is 0.001 mm more than the CG module. Since this loss is within the CG module's error range, it does not affect the backend polishing process. This result verifies that the DBZPT compensation approach accurately predicts changes in grinding thickness and compensates for this, mitigating CG wear. Also, the fact that the forecast error for a 180-second period is within the allowable range suggests that using a shorter period would yield similar performance. Then, we note that the stability of the sensor determines the quality of the wafer endpoint thickness, so more thinning and gap measurements for grinding can be further investigated.

VI. CONCLUSION

We propose a forecast-based endpoint thickness and online error compensation approach for the grinding of hard, brittle material, where AE, vibration, current signals, and grinding state data are utilized as time-series input features. We analyze the time-frequency characteristics of the AE, tri-axial vibration, and current signals in the grinding process to develop preprocessing and feature extraction methods for each signal. We construct six different multi-signal time-series forecast models and evaluate their performance through verification and testing, and show that Res-LSTM yields the best performance. We also investigate prediction using different period lengths for the forecast model, showing good performance for periods shorter than 180 seconds. We use the resultant compensation approach with the forecast model in wafer thinning experiments and compare the results of the PG and CG modules. The results attest the effectiveness and performance of the proposed CG + Res-LSTM method, which can be extended to other grinding types and materials.

REFERENCES

[1] A. T. Cheung, "Dicing die attach films for high volume stacked die application," in *Proc. 56th Electron. Compon. Technol. Conf.*, May 2006, pp. 1–5.
[2] C.-W. Liu, H.-C. Chen, and S.-C. Lin, "Acoustic emission monitoring system for hard polishing of sapphire wafer," *Sensors Mater.*, vol. 31, no. 9, pp. 2681–2689, 2019.
[3] D. Axinte, N. Gindy, K. Fox, and I. Unanue, "Process monitoring to assist the workpiece surface quality in machining," *Int. J. Mach.*, vol. 44, no. 10, pp. 1091–1108, Aug. 2004.
[4] Y. D. Filatov, "Polishing of precision surfaces of optoelectronic device elements made of glass, sitall, and optical and semiconductor crystals: A review," *J. Superhard Mater.*, vol. 42, no. 1, pp. 30–48, Jan. 2020.
[5] C.-C.-A. Chen and L.-S. Hsu, "A process model of wafer thinning by diamond grinding," *J. Mater. Process. Technol.*, vol. 201, nos. 1–3, pp. 606–611, May 2008.
[6] H. Jamshidi and E. Budak, "On the prediction of surface burn and its thickness in grinding processes," *CIRP Ann.*, vol. 70, no. 1, pp. 285–288, 2021.
[7] L. Lv, Z. Deng, T. Liu, Z. Li, and W. Liu, "Intelligent technology in grinding process driven by data: A review," *J. Manuf. Processes*, vol. 58, pp. 1039–1051, Oct. 2020.
[8] R. Babel, P. Koshy, and M. Weiss, "Acoustic emission spikes at workpiece edges in grinding: Origin and applications," *Int. J. Mach. Tools Manuf.*, vol. 64, pp. 96–101, Apr. 2013.
[9] D. E. Lee, I. Hwang, C. M. O. Valente, J. F. G. Oliveira, and D. A. Dornfeld, "Precision manufacturing process monitoring with acoustic emission," *Int. J. Mach. Tools Manuf.*, vol. 46, no. 2, pp. 176–188, Feb. 2006.
[10] T. W. Liao, C.-F. Ting, J. Qu, and P. J. Blau, "A wavelet-based methodology for grinding wheel condition monitoring," *Int. J. Mach. Tools Manuf.*, vol. 47, nos. 3–4, pp. 580–592, Mar. 2007.
[11] A. Nasiri, J. Bao, D. Mccleeary, S.-Y.-M. Louis, X. Huang, and J. Hu, "Online damage monitoring of SiC_f-SiC_m composite materials using acoustic emission and deep learning," *IEEE Access*, vol. 7, pp. 140534–140541, 2019.
[12] Y.-K. Lin, B.-F. Wu, and C.-M. Chen, "Characterization of grinding wheel condition by acoustic emission signals," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jun. 2018, pp. 1–6.
[13] Y.-K. Lin and B.-F. Wu, "Machine learning-based wheel monitoring for sapphire wafers," *IEEE Access*, vol. 9, pp. 46348–46363, 2021.
[14] P. R. Aguiar, P. J. A. Serni, E. C. Bianchi, and F. R. L. Dotto, "In-process grinding monitoring by acoustic emission," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2004, pp. 1–4.
[15] C. H. R. Martins, P. R. Aguiar, A. Frech, and E. C. Bianchi, "Tool condition monitoring of single-point dresser using acoustic emission and neural networks models," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 3, pp. 667–679, Mar. 2014.
[16] M. S. H. Bhuiyan, I. A. Choudhury, M. Dahari, Y. Nukman, and S. Z. Dawal, "Application of acoustic emission sensor to investigate the frequency of tool wear and plastic deformation in tool condition monitoring," *Measurement*, vol. 92, pp. 208–217, Oct. 2016.
[17] D. He, R. Li, J. Zhu, and M. Zade, "Data mining based full ceramic bearing fault diagnostic system using AE sensors," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2022–2031, Dec. 2011.
[18] J. Badger, S. Murphy, and G. E. O'Donnell, "Acoustic emission in dressing of grinding wheels: AE intensity, dressing energy, and quantification of dressing sharpness and increase in diamond wear-flat size," *Int. J. Mach. Tools Manuf.*, vol. 125, pp. 11–19, Feb. 2018.
[19] C.-H. Lee, J.-S. Jwo, H.-Y. Hsieh, and C.-S. Lin, "An intelligent system for grinding wheel condition monitoring based on machining sound and deep learning," *IEEE Access*, vol. 8, pp. 58279–58289, 2020.
[20] W. Yin-Ling and L. Hua-Cong, "Research on acoustic emission source localization of carbon fiber composite plate based on wavelet neural network," in *Proc. IEEE 10th Int. Conf. Mech. Aerosp. Eng. (ICMAE)*, Brussels, Belgium, Jul. 2019, pp. 302–305.
[21] P. Shi, L. Hong, and D. He, "Using long short term memory based approaches for carbon steel fatigue remaining useful life prediction," in *Proc. Prognostics Syst. Health Manage. Conf.*, Chongqing, China, Oct. 2018, pp. 1055–1060.
[22] A. Marec, J.-H. Thomas, and R. E. Guerjouma, "Damage characterization of polymer-based composite materials: Multivariable analysis and wavelet transform for clustering acoustic emission data," *Mech. Syst. Signal Process.*, vol. 22, no. 6, pp. 1441–1464, Aug. 2008.
[23] Z. Yang, Z. Yu, C. Xie, and Y. Huang, "Application of Hilbert–Huang transform to acoustic emission signal for burn feature extraction in surface grinding process," *Measurement*, vol. 47, pp. 14–21, Jan. 2014.

[24] S. Shao, R. Yan, Y. Lu, P. Wang, and R. Gao, "DCNN-based multi-signal induction motor fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2658–2669, Jun. 2020.
[25] D. M. S. Ribeiro, P. R. Aguiar, L. F. G. Fabiano, D. M. D'Addona, F. G. Baptista, and E. C. Bianchi, "Spectra measurements using piezoelectric diaphragms to detect burn in grinding process," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 3052–3063, Nov. 2017.
[26] H. V. Dang, H. Tran-Ngoc, T. V. Nguyen, T. Bui-Tien, G. De Roeck, and H. X. Nguyen, "Data-driven structural health monitoring using feature fusion and hybrid deep learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 4, pp. 2087–2103, Oct. 2021.
[27] R. Zhao, R. Yan, J. Wang, and K. Mao, "Learning to monitor machine health with convolutional bi-directional LSTM networks," *Sensors*, vol. 17, no. 2, p. 273, Jan. 2017.
[28] M. Miao, Y. Sun, and J. Yu, "Sparse representation convolutional autoencoder for feature learning of vibration signals and its applications in machinery fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 69, no. 12, pp. 13565–13575, Dec. 2022.
[29] P. S. Huang, M. Kim, M. Hasegawa-Johnson, and P. Smaragdis, "Joint optimization of masks and deep recurrent neural networks for monaural source separation," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 23, no. 12, pp. 2136–2147, Dec. 2015.
[30] H. Wang, J. Xu, C. Sun, R. Yan, and X. Chen, "Intelligent fault diagnosis for planetary gearbox using time-frequency representation and deep reinforcement learning," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 2, pp. 985–998, Apr. 2022.
[31] J. S. Rad, Y. Zhang, F. Aghazadeh, and Z. C. Chen, "A study on tool wear monitoring using time-frequency transformation techniques," in *Proc. Int. Conf. Innov. Design Manuf. (ICIDM)*, Aug. 2014, pp. 342–347.
[32] S. Mirifar, M. Kadivar, and B. Azarhoushang, "First steps through intelligent grinding using machine learning via integrated acoustic emission sensors," *J. Manuf. Mater. Process.*, vol. 4, no. 2, p. 35, Apr. 2020.
[33] Z. Zhao, Q. Zhang, and X. Chen, "Applications of unsupervised deep transfer learning to intelligent fault diagonosis: A survey and comparative study," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–28, 2021.
[34] D. M. D'Addona, S. Conte, W. N. Lopes, P. R. de Aguiar, E. C. Bianchi, and R. Teti, "Tool condition monitoring of single-point dressing operation by digital signal processing of AE and AI," *Proc. CIRP*, vol. 67, pp. 307–312, Jan. 2018.
[35] R. Liu, B. Yang, and X. Chen, "Artifical intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.
[36] L. Yao and Z. Ge, "Dynamic features incorporated locally weighted deep learning model for soft sensor development," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
[37] X. Li, J. Wu, Z. Li, J. Zuo, and P. Wang, "Robot ground classification and recognition based on CNN-LSTM model," in *Proc. IEEE 2nd Int. Conf. Big Data, Artif. Intell. Internet Things Eng. (ICBAIE)*, Mar. 2021, pp. 26–28.
[38] R. Huang, Y. Liao, S. Zhang, and W. Li, "Deep decoupling convolutional neural network for intelligent compound fault diagnosis," *IEEE Access*, vol. 7, pp. 1848–1858, 2019.
[39] H. Wang, Z. Liu, D. Peng, M. Yang, and Y. Qin, "Feature-level attention-guided multitask CNN for fault diagnosis and working conditions identification of rolling bearing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4757–4769, Sep. 2022.
[40] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*.
[41] *BSG-V Operation Manual*, SpeedFam Inc., 2015.

YU-KUN LIN (Student Member, IEEE) is currently pursuing the Ph.D. degree with the Institute of Electrical and Computer Engineering, National Yang Ming Chiao Tung University (NYCU), Hsinchu, Taiwan. He has worked as an Electrical Control Engineer at the Industry Technology Research Institute (ITRI) for more than ten years. His current research interests include intelligent grinding processing for hard and brittle materials, signal processing, data mining, and machine learning.

BING-FEI WU (Fellow, IEEE) received the B.S. and M.S. degrees in control engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1992.

Since 1992, he has been with the Department of Electrical and Computer Engineering, where he was promoted to a Professor, in 1998, and a Chair Professor, in 2020. His research interests include image recognition, physiological informatics, vehicle driving safety and control, intelligent robotic systems, and intelligent transportation systems. Since 2019, he has been the President of the Taiwan Association of System Science and Engineering and the Director of the Control Engineering Program, Ministry of Science and Technology, Taiwan. He was the Director of the Institute of Electrical and Control Engineering, NCTU, in 2011. He has received numerous research honors and awards, including the FutureTech Breakthrough Award, from 2017 to 2019, and the Outstanding Research Award, in 2015 and 2019, all from the Ministry of Science and Technology, Taiwan; the Technology Invention Award from the Y. Z. Hsu Foundation, in 2014; the National Invention and Creation Award from the Ministry of Economic Affairs, Taiwan, in 2012 and 2013; the Outstanding Research Award from the Pan Wen Yuan Foundation, in 2012; the Best Paper Award from the 12th International Conference on ITS Telecommunications, in 2012; the Best Technology Transfer Contribution Award from the National Science Council, Taiwan, in 2012; and the Outstanding Automatic Control Engineering Award from the Chinese Automatic Control Society, in 2007. He founded and was the first Chair of the Taipei Chapter of the IEEE Systems, Man and Cybernetics Society (SMCS), in 2003, and was also the Chair of the Technical Committee on Intelligent Transportation Systems of IEEE SMCS in 2011. Currently, he is an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS.