Article

Predictive Modeling of Surface Wear in Mechanical Contacts under Lubricated and Non-Lubricated Conditions

Ali Rahman [1,*], Muhammad Khan [2] and Aleem Mushtaq [3]

1   Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, 64283 Darmstadt, Germany
2   School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK; Muhammad.A.Khan@cranfield.ac.uk
3   Department of Electronics and Power Engineering, Pakistan Navy Engineering College, National University of Sciences and Technology, Karachi 75350, Pakistan; aleem@pnec.nust.edu.pk
*   Correspondence: ali.rahman@stud.tu-darmstadt.de

Abstract: The surface wear in mechanical contacts under running conditions is always a challenge to quantify. However, the inevitable relationship between the airborne noise and the surface wear can be used to predict the latter with good accuracy. In this paper, a predictive model has been derived to quantify surface wear by using airborne noise signals collected at a microphone. The noise was generated from a pin on disc setup on different dry and lubricated conditions. The collected signals were analyzed, and spectral features estimated from the measurements and regression models implemented in order to achieve an average wear prediction accuracy of within $1 \, mm^3$.

1. Introduction

Mechanical structures and components generate airborne noise due to their contact mechanics even in normal operational conditions. The amplitude and the frequency of the noise have a direct relation with the surface wear of the contact zone [1]. As the material wears out due to contact, a rapid release of energy from localized sources within a material causes a noise to be generated, hence the two factors; noise and wear seem in relation with each other [2]. Hence, it is not wrong to say that the spectrum of the emitted noise can be evaluated for the health diagnosis of machine components; its wear and friction resistance, as confirmed from the literature [3–5].

The correlation between wear and generated noise is not a new phenomenon for machine health diagnosis. This has been established since the 1970s [6]. Studies indicate the effectiveness of analyzing contact based acoustic emissions in order to calculate the wear occurred in a component for various machine processes like sliding, cutting, milling, and tool manufacturing [1,7–16]. The spectrum of the obtained acoustic emissions is further processed to measure count rates, RMS voltage, and also FFT analysis [10,13]. The signal analysis provides the acoustic parameter, whereas a wear parameter is also measured (wear area, wear volume, degradation, coefficient of friction) in order to correlate the two parameters [17–19]. The tool wear monitoring has been a regular exercise for industrial applications [20].

For this purpose, recently, Kong et al. described the methods of online tool wear monitoring. They used Kernel principal component analysis to fuse the sensitive tool wear signals. They constructed a predictive model for tool wear based on support vector regression and correlated the fused signal and the actual tool wear. The results showed that the prediction accuracy was close to 100% [21]. Another study used Titanium cylindrical flat pins over a Silicon carbide abrasive disc under dry sliding conditions [22]. It used the Taguchi method to predict wear behavior against variable load, speed, sliding duration, and

signal-to-noise ratio. The obtained results showed a good correlation of wear and acoustic emissions under variable operating conditions, and a predictive model based on linear regression equation was also developed. Contact based signals were also studied under the effect of lubrication in the past [23–25]. Recently, researchers showed the possibility of the trend modeling between acoustic emissions and the wear in terms of accumulated signal energy and changes in spectrum amplitude [9,25–29]. However, most of these efforts used a single pin/point contact sensor and analyzed the spectrum of noise with different speed, sliding distance, and temperature values.

An extensive research literature is available on signal processing techniques that are used for analyzing acoustic emissions extracted from a contact based sensor and further correlated with the possible surface wear parameters. Several regression techniques are used in developing these correlations. However, the predictive modeling of wear volume through a non-contact microphone, still requires comprehensive effort. A reliable quantification of surface wear volume would be a major breakthrough in the field of condition based monitoring and early fault detection. In this paper, a predictive model has been derived and used to quantify the surface wear volume in a surface contact by using airborne noise recorded by means of a non-contact microphone. Airborne noise was generated from the operation of a pin on disc setup. The setup was run on different loads with and without lubrication. The noise was analyzed in order to determine features to extract, which were then computed via Non-Parametric Spectral Estimation. These were then considered along with the nature of the data acquisition experiments in order to formulate an appropriate predictive model.

2. Method

2.1. Experimental Setup and Wear-Noise Generation Scheme

Authors have used the experimental rig as described in their recent work on noise and wear correlation [1]. In total, 37 experiments were performed on a pin on disc tribometer as shown in Figure 1 and considered for the statistical model development and validation.
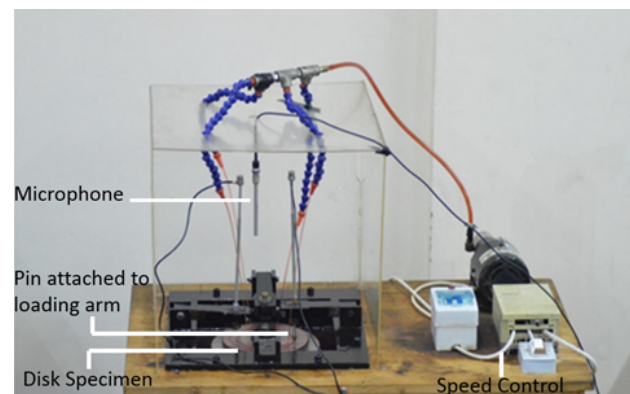

Figure 1. Pin on disc tribometer.

The tribometer was composed of metallic arms to act as pin holders. A circular mild steel disc was mounted on a shaft and motor assembly that allowed the disc to rotate at a constant speed of 250 rpm. A high speed steel material pin was used and assembled with the help of the pin holder in a way so that it made a prefect mechanical contact with the circular disc as shown in Figure 1. The pin had a diameter of 5 mm and a contact diameter of 0.04 mm. Additional details about the pin's material are listed in Table 1.

Table 1. Pin material.

Steel Grade	Hardness (HRC)	Chemical Constitution (wt%)					
		Carbon	Tungsten	Molybdenum	Chromium	Vanadium	Silicon
HS3-3-3	64	1.00	2.98	2.83	4.27	2.30	-

The pin tip at contact was made semi-circular in shape. It was assumed that the pin would not be worn out under sliding conditions at any time of experiments and hence the change in airborne noise would only be caused due the surface wear of the disc as it was made of a softer material. Threads of up to 10 mm in length were present on the other end of the pin, which were used to mount metallic weights and allowed to provide a tangential loading condition on the pin and disc contact during experiments. An oil pump was also used to provide lubrication at the contact and at a rate of 5 mL/s. The selected grade of lubricant was 10W-30 with a density of 877 kg/m^3.

The noise was recorded by a GRAS 40PP free-field microphone (GRAS Sound & Vibration, Holte, Denmark). The frequency range of the selected microphone was 20 kHz, and it was able to record noises with an upper limit of 128 db. It was vertically placed at the center of the disc as shown in Figure 1. The raw noise signals were acquired by an NI 9234 data acquisition module (National Instruments, Austin, TX, USA) in terms of sound pressure (units pascal) and sampling rate of 25.6 kHz. Images of the wear scars on the disc were taken from six different positions, the same as it was performed in the previous work, with a portable microscope (Dino-Lite AM413T, AnMo Electronics Corporation, Hsinchu, Taiwan) at a magnification of 200× as shown in Figure 2 [1].
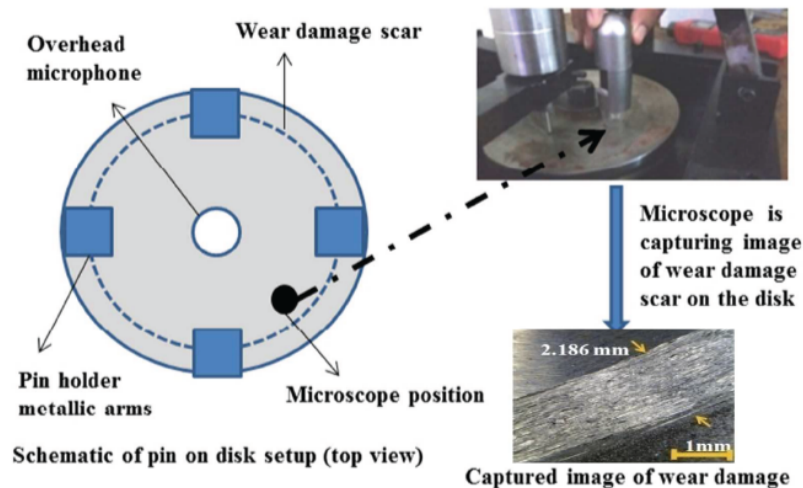

Figure 2. Portable microscope used to capture images of the disc scar (Image reused under STM Guidelines. Content rights are owned by and permission requests for further reuse are handled by SAGE Publishing, CA, USA) [1].

Using ASTM standard G99 (Standard Test Method for Wear Testing with a Pin-on-Disc Apparatus), disk volume loss V due to wear scar was calculated [17]. The wear scar sliding length measurements were measured from the captured images and used in ASTM standard formulation as provided in Equation (1):

$$V = 2\pi R \left[r^2 \arcsin\left(\frac{d}{2r}\right) - \left(\frac{d}{4}\right)\sqrt{4r^2 - d^2} \right] \tag{1}$$

where R = wear track radius, d = wear track width, r = pin end radius, and V is the volumetric wear, assuming no significant pin wear.

The load and lubrication specification breakdown for the 37 experiments is given in Table 2. The duration of each experiment was 6 min, with the wear on a disc measured after 30 s intervals, thus resulting in a total of 444 measurements.

Table 2. Breakdown of experiments and measurements.

| Specification	Number of Experiments	Number of Measurements
No lubrication, no load	9	108
Lubrication, no load	8	96
Lubrication, 4.91 N load	10	120
Lubrication , 9.81 N load	10	120

2.2. Signal Analysis and Feature Extraction

Cutting processes tend to be stochastic [30] with generated noise consequently dependent upon the progression of the tool wear [31]. Hence, the stationarity of the collected data needs to be determined in order to decide the features extraction methodology. If the acquired signals turn out to be non-stationary, then Time-Frequency features would need to be computed; otherwise, frequency domain features would suffice [32]. To help with the analysis for stationarity, the mean and Autocorrelation Sequence (ACS) of one of the 30 s measurements was computed. The ACS of a process X can be defined as:

$$R_{xx}[k] = E\{X[n]X[n+k]\} \tag{2}$$

where n is the time index, k the lag index, and E the expectation operator. Hence, the ACS is a measure of the correlation between two samples of the same stochastic process separated by a lag k.

The evolution of the mean of a 30 s long measurement, computed using a moving average, is plotted against time in Figure 3. Figure 4 shows the computed ACS against the lag indexes for the same measurement. Figure 5 displays the same ACS with the lag index limited to $|100|$. Observing the three figures, we can conclude that:

1. The mean μ is relatively constant.
2. For $k = 0$, $R_{xx}[k]$ is positive.
3. $R_{xx}[k]$ is an even function.
4. $R_{xx}[0]$ is max$\{R_{xx}[k]\}$.
5. $R_{xx}[k]$ approaches μ^2 as k increases.
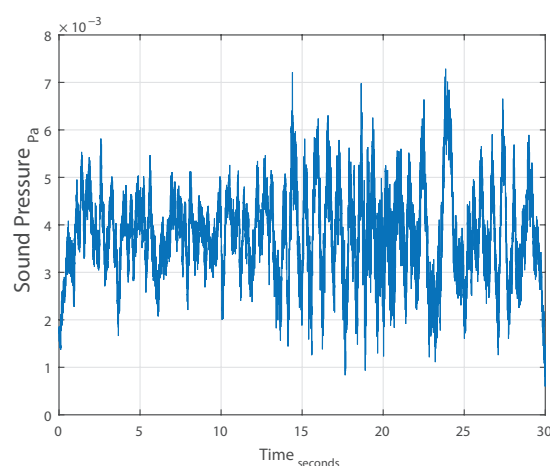
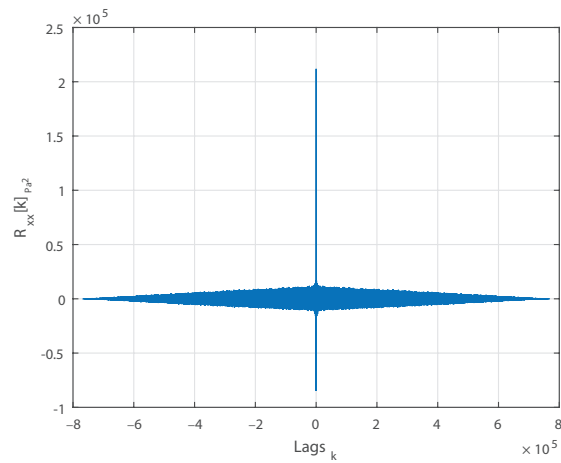**Figure 3.** Moving average of a single measurement.

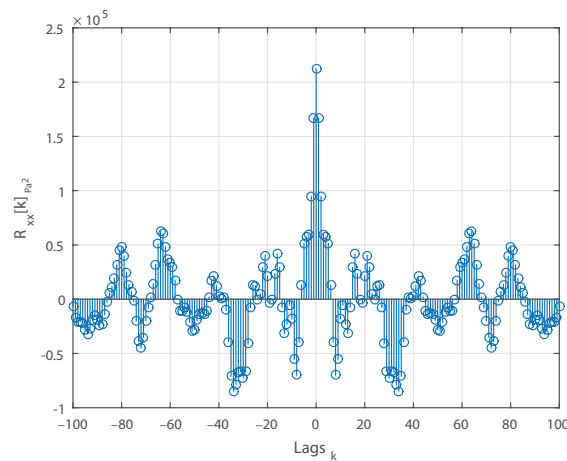**Figure 4.** Autocorrelation sequence of a single 30 s long measurement.

Figure 5. Autocorrelation sequence with lags limited to 100.

Hence, the measurement can be classified as a wide-sense stationary (WSS) process [33]. Moreover, since the measured signal resembles a Gaussian distribution, as seen in Figure 6, this WSS process can be considered strict stationary (SS) as well [34].
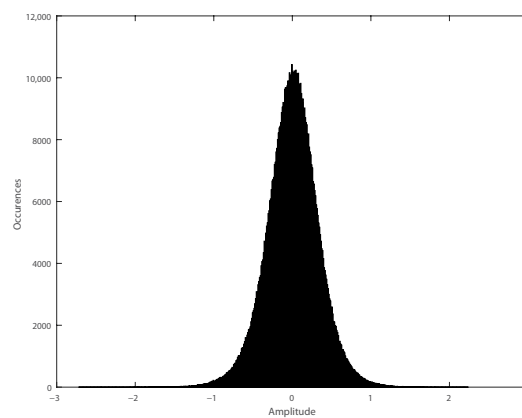

Figure 6. Distribution of measured signals' samples.

Given the result of the stationarity analysis, the power spectral density (PSD) can be estimated. According to the Wiener–Khinchin theorem, the PSD is the squared discrete

Fourier transform (DFT) of the ACS of each measurement signal [35]. The conventional periodogram estimator is considered analogous to this definition [36].

Since using the periodogram would have resulted in large fluctuations around the actual PSD due to the method's inherent asymptotic noise [37], it is imperative to revert to techniques that reduce spectral variance.

Hence, the Welch method [38], which involves the segmentation of the signal being analyzed, and averaging of periodograms can be used. While segmentation reduces spectral resolution, the use of segment overlap mitigates this reduction when compared to the Bartlett method for the same number of segments [39]. The subsequent generation of redundant spectral information due to the use of overlapping is attenuated through the application of a windowing function, such as the Hamming window [40], over each segment. The Welch method for PSD estimation can be summarized in the following steps:

For periodograms $k = 1, ..., K$, a segment k is defined as:

$$X_k(n) = X(n + (k-1)D) \qquad n = 0, ..., M-1 \tag{3}$$

with M being the length of each segment and D controlling the degree of overlap. The number of periodograms involved, K, is dependent upon the previous two parameters and, N, the total length of the signal.

$$K = \frac{N - M}{D} + 1 \tag{4}$$

Each periodogram can be defined as:

$$I_{WX,WX}^{M}(e^{jw}, k) = \frac{1}{MZ} \left| \sum_{n=0}^{M-1} w(n) X_k(n) e^{-jwn} \right|^2 \tag{5}$$

where Z is the normalization factor considered due to the introduction of the Hamming window function $w(n)$. By averaging this ensemble of periodograms, the Welch estimator is defined as:

$$\tilde{C}_{XX}(e^{jw}) = \frac{1}{K} \sum_{k=1}^{K} I_{WX,WX}^{M}(e^{jw}, k) \tag{6}$$

For each measurement of duration 30 s, a Welch estimator with $M = 1280$ was used along with an overlap of 50%. Since sampling frequency used was 25.6 kHz, $N = 768,000$. Consequently, the PSD estimates $\tilde{C}_{XX}(e^{jw})$ have a spectral resolution of 20 Hz. A lower window length wasn't set in order to prevent spectral smearing and higher bias.

For the measurements we recorded, it was observed that the frequency ranges of interest were $0 \leq f \leq 3$ kHz and $4 \leq f \leq 5$ kHz. This can be visualized in Figure 7, which displays $\tilde{C}_{XX}(e^{jw})$ for 108 measurements of 30 s each. For the sake of visual brevity, only measurements corresponding to absence of load and lubrication have been included.

The identification of these spectral regions of interest led to selection of the PSD estimates of 200 discrete frequency bins as features for our eventual surface wear prediction model.
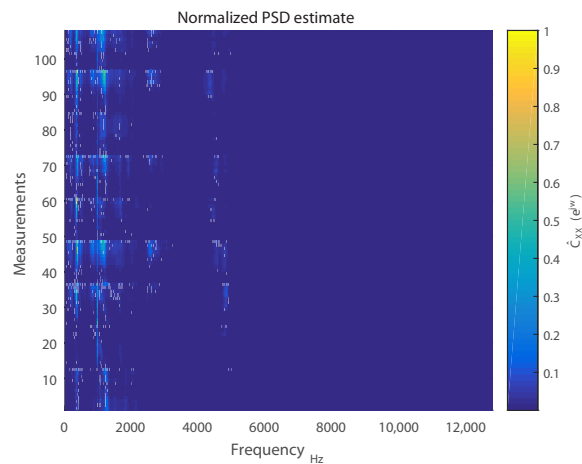
Figure 7. Normalized PSD estimates.

2.3. Prediction Model Formulation

Since volumetric loss of one disc was measured over a span of 6 min in intervals of 30 s, they are cyclically cumulative, and the task of surface wear prediction can be modeled as a time series forecasting problem with a time horizon [41] of depth 12.

For subsequent discussion, let h denote the time horizon, with y_h and X_h corresponding to the surface wear and spectral features of the measurement at horizon h. The model for predicting surface wear estimates g_h can be expressed as f. A common solution for a forecasting problem is the recursive approach, which involves the prediction being fed back to the model, hence the name, as part of the input for prediction of the target belonging to the next time horizon [42]. However, this approach, illustrated in Figure 8, is extremely sensitive to prediction error due to error propagation [43].
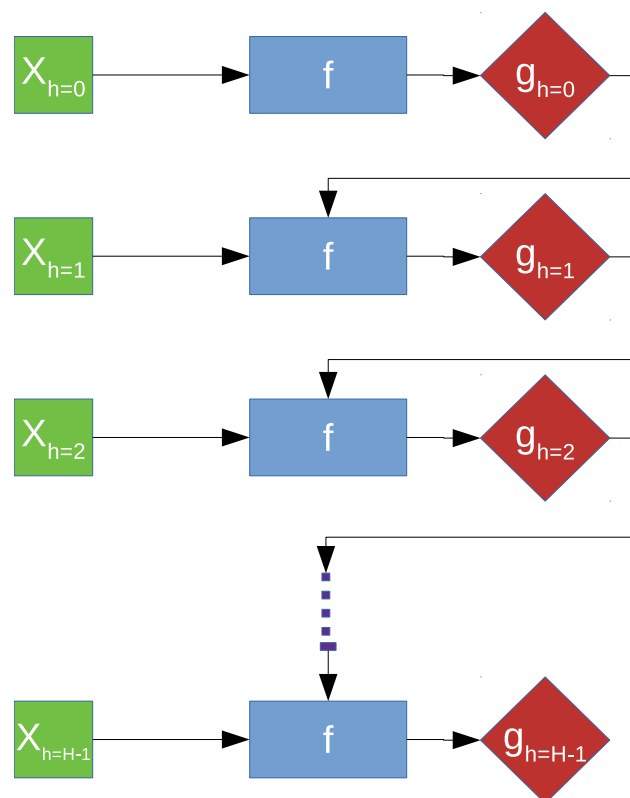

Figure 8. Recursive approach to time series forecasting.

However, the time series aspect of the problem could not be ignored since the quantity to be predicted is the cumulative surface wear, with each 30 s wear process dependent on the previous one. Hence, we implemented time-embeddings using a sliding window approach for our spectral feature s [44]. Consequently, we arrive at a regression problem defined in Figure 9.
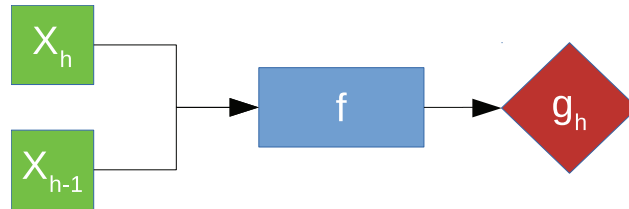

Figure 9. Proposed approach.

A sliding window, consisting of inputs belonging to the current time horizon h and the previous one $h - 1$, is used as the new inputs which will be used to predict the cumulative surface wear at horizon h. In other words, we have set up a high dimensional regression problem with 400 spectral features as input for every cumulative surface wear output.

2.4. Choice of Regression Models

As can be discerned in Figure 7, the spectral features exhibit collinearity due to similar PSD estimates between some neighboring discrete frequency bins. Accordingly, a least-squares (LS) linear regression based model will not be optimal for prediction [45]. An optimal solution will seek to constrict the effects of collinearity between these features by attenuating the weights assigned to them during regression [46]. One method would be to implement a subset selection of features. However, this approach exhibits high variance due to the binary process of picking or dropping a feature entirely. Consequently, this would result in negligible improvement in prediction performance [46].

Ridge regression [47] is the alternate approach that can be adopted. It is a technique involving the shrinking of feature weights by imposing a penalty on their sizes. For a regression model defined by Equation (7), where X is an input matrix of features, y is the vector of output targets, and w the vector of weights assigned to features:

$$y = wX \tag{7}$$

The LS linear regression solution would be the set of weights that minimize the sum of squared errors between the targets and predictions, as defined in Equation (8):

$$\hat{w} = \arg\min_{w} \left(y - w^T X \right)^T \left(y - w^T X \right) \tag{8}$$

Comparatively, the optimization function for ridge regression includes the additional β parameter in order to penalize the size of weights as discussed previously. Hence, converging to the solution for Equation (9), requires reduction in the magnitudes of feature weights.

$$\hat{w} = \arg\min_{w} \left(\left(y - w^T X \right)^T \left(y - w^T X \right) + \beta w^T w \right) \tag{9}$$

While this causes an estimation bias, it decreases prediction error due to a reduction in variance [48]. The superiority of this method over feature subset selection has been proven in the work of Frank and Friedman (1993) [49].

Both LS linear regression and ridge regression suffer from a common drawback, the underlying linearity assumption. Accordingly, they offer poor and highly biased predictions when the relationship between features and the target is more complex. To overcome this issue, we will be moving on from the ridge regression method to Kernel Ridge Regression

(KRR) [50], through which our spectral features will be mapped into a different dimensional space using a nonlinear kernel.

The Chi-Squared (χ^2) kernel [51], defined in Equation (10), is used. The kernel utilizes the χ^2 distance between two vectors a and b in order to generate a new feature vector, with γ controlling the variance of the kernel. This parameter allows us to improve the generalization of our regression model and reduce overfitting:

$$\kappa(a, b) = \exp\left(-\gamma \sum_i \frac{(a[i] - b[i])^2}{a[i] + b[i]}\right) \tag{10}$$

Our choice of kernel function is influenced by the assumption that our spectral features consisting of PSD estimates are similar to sparse histogram features for which the χ^2 kernel performs well [52].

3. Results and Discussion

We used the dataset of our collected measurements to train and validate the following regression models, using Scikit-learn [53]:

- LS Linear Regression
- Ridge Regression
- χ^2 KRR

The dataset was randomly sorted and partitioned such that 404 measurements were used as the training set while the remaining 40 were kept for validation. The training set could not be reduced further; otherwise, the quantity of observations available would be lower than the number of features, which would bar convergence to a unique solution for LS linear regression [54].

Two metrics were used in order to evaluate the performance of the regression models: the R-squared (R^2) and the Mean Absolute Error (MAE). The former determines the model's goodness-of-fit and how well its predictions map towards the target variables [55], while the latter is a statistic denoting the mean of the error magnitudes between the target values and the model's predictions [56]. In our context, the MAE represents the mean disparity, in mm^3, between actual and predicted surface wears.

The performance of the models is summarized in Table 3.

Table 3. Performance of regression models.

Metric Model	R^2		MAE (mm^3)	
	Training Set	**Validation Set**	**Training Set**	**Validation Set**
Least-Squares	0.9759	0.3599	0.720	2.727
Ridge	0.9170	0.9034	1.284	1.173
Kernel Ridge	0.9600	0.9716	0.824	0.635

These results were achieved with the following model parameters:

- β: 0.15
- γ: 0.1

These parameters were set after hyper-parameter tuning. It can be noted that, while LS linear regression provides the lowest error on the training set, it is a result of overfitting since its performance on the validation set is dismal. Ridge regression provides better generalization and avoids the collinearity problem, but its performance isn't ideal. The use of a χ^2 kernel resulted in the validation error of ridge regression being nearly halved and a significantly higher R^2 score. The disparity in performance of the three models can be visually interpreted through the predictions shown in Figures 10–15.

Figure 10. Surface wear prediction on the training set using least squares linear regression.

Figure 11. Surface wear prediction on the validation set using least squares Linear Regression.

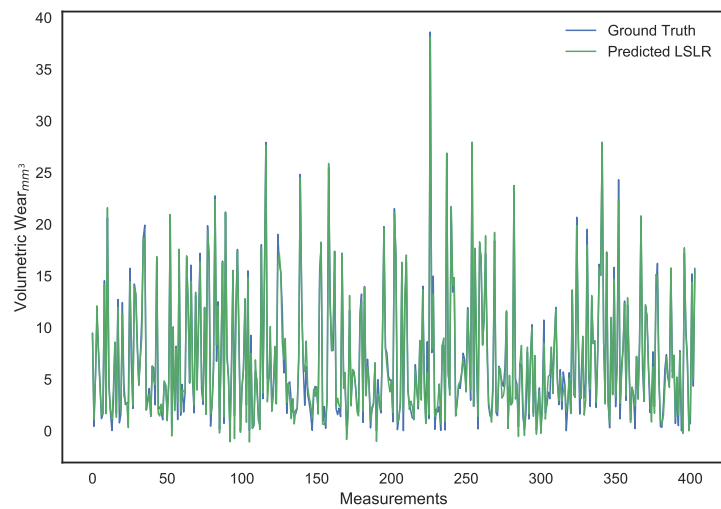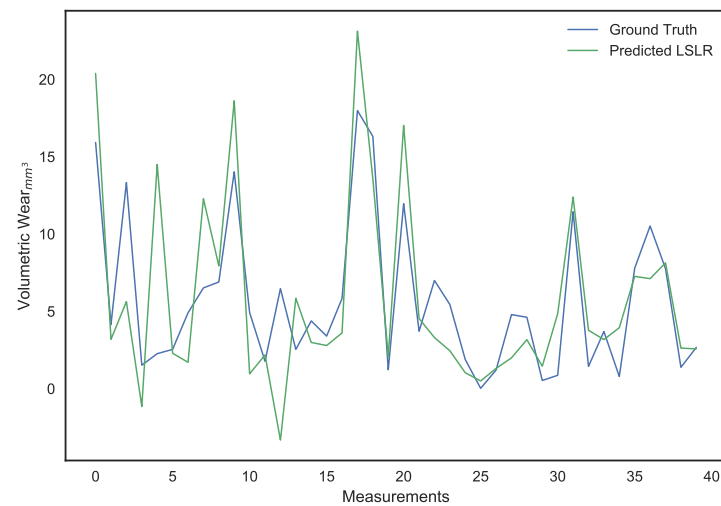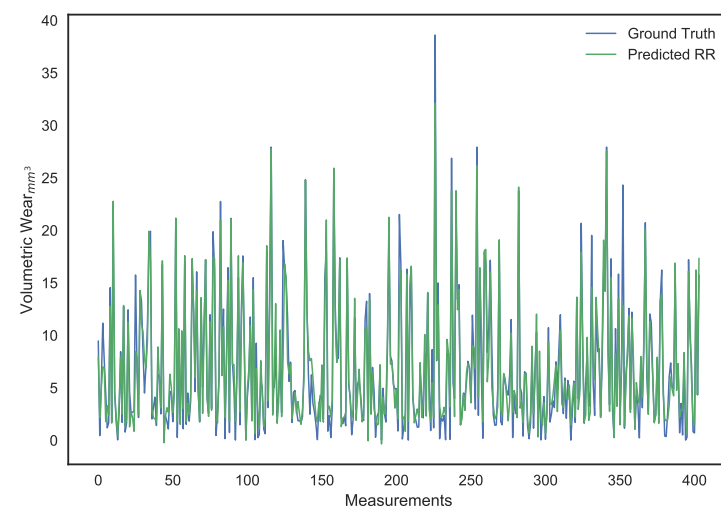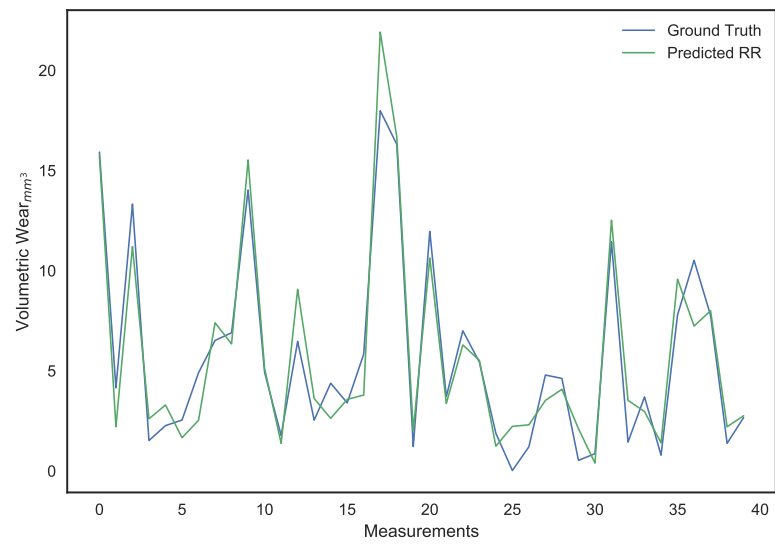**Figure 12.** Surface wear prediction on the training set using ridge regression.

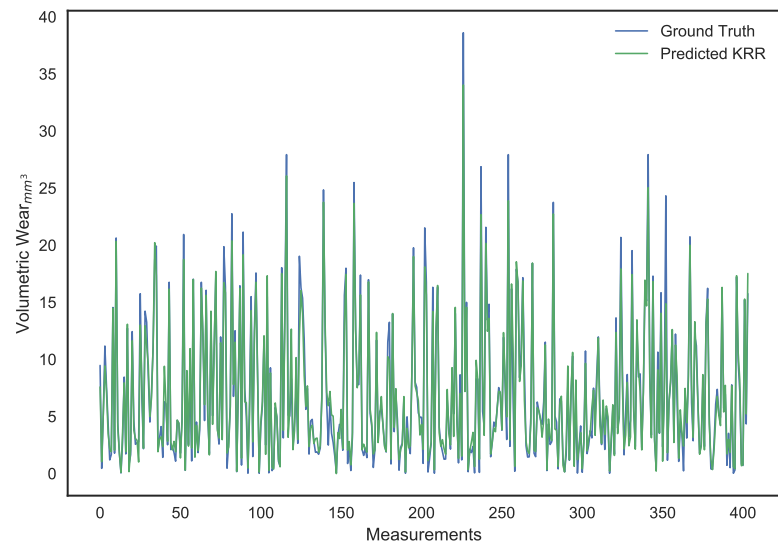**Figure 13.** Surface wear prediction on the validation set using ridge regression.

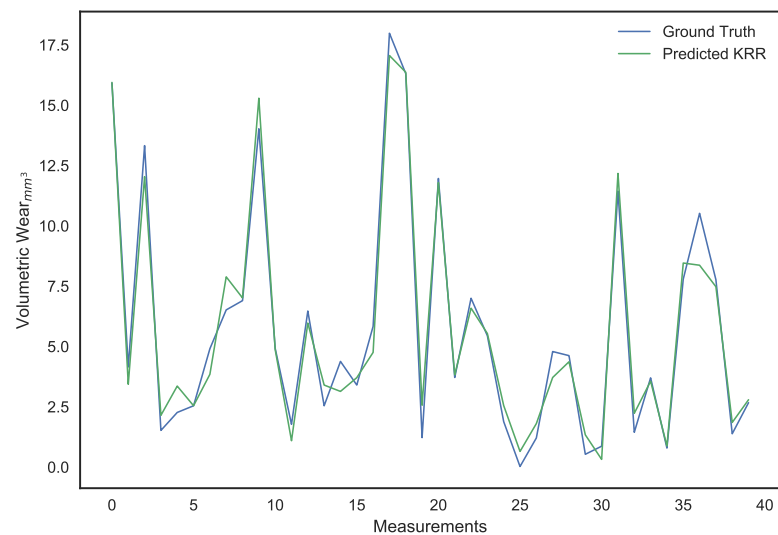**Figure 14.** Surface wear prediction on the training set using χ^2 kernel ridge regression.

Figure 15. Surface wear prediction on the validation set using χ^2 kernel ridge regression.

It can be observed that, apart from χ^2 regression, the other models even predicted negative surface wear losses. The deviations of each of the model's predictions from the actual volumetric loss can be noted in Figure 16, which illustrates how, as a whole, the deviations for the χ^2 kernel ridge regression are smaller than other models. Most significantly, as demonstrated by the validation MAE metric, it predicts surface wear with an average deviation of under $1\,\text{mm}^3$.
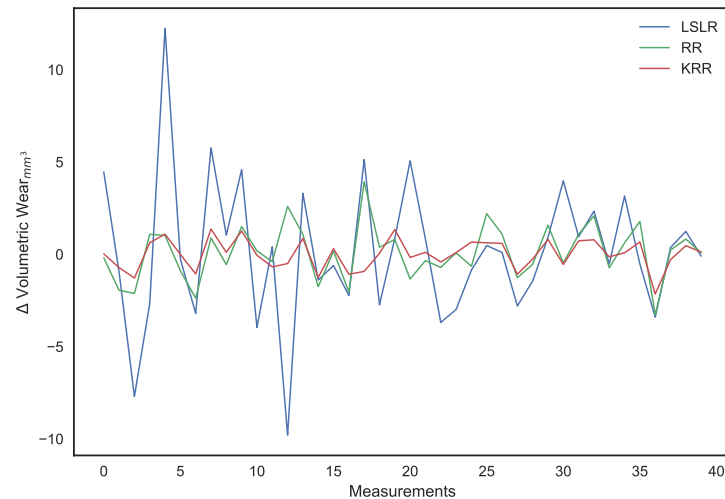

Figure 16. Deviation between model predictions and actual surface wear.

Since the load and lubrication conditions for each measurement were not factored into the input feature space for the prediction model, it is pertinent to ensure that the model is agnostic to these conditions and is not returning favorable performance for a subset of load and lubrication specifications. Additionally, it needs to be ensured that the performance of the χ^2 kernel ridge regression model is generalized and not optimized only for a particular split of training and validation data split.

The dataset of 444 measurements was randomly sorted and split an additional nine times, while ensuring that there were a minimum of five measurements corresponding to each load and lubrication specification in a validation set. The KRR model, with the previously defined hyper-parameters, was trained and validated on these sets separately. The distribution of measurements against their specifications in the validation sets is provided in Table 4, with the breakdown of MAE achieved for different specifications tabulated in Table 5. The first set in the tables corresponds to the validation set used to compare KRR with other regression models.

Table 4. Distribution of measurements in validation sets.

Specifications	Sets									
	1	2	3	4	5	6	7	8	9	10
No lubrication, no load	7	7	9	11	11	8	7	12	9	9
Lubrication, no load	10	14	5	10	11	9	10	8	13	12
Lubrication, 4.91 N load	13	8	13	6	12	11	8	10	9	8
Lubrication , 9.81 N load	10	11	13	13	6	12	15	10	9	11

Table 5. Breakdown of metrics across validation sets.

Specifications	MAE (mm^3)									
	1	2	3	4	5	6	7	8	9	10
No lubrication, no load	0.690	1.007	0.705	0.742	0.772	0.811	0.424	0.991	0.979	0.764
Lubrication, no load	0.534	0.740	0.766	0.676	0.723	0.776	0.793	0.816	0.660	0.978
Lubrication, 4.91 N load	0.740	0.752	1.030	0.762	0.962	0.570	0.908	0.736	0.845	0.863
Lubrication , 9.81 N load	0.561	0.845	0.647	0.586	0.436	0.868	0.869	0.642	0.611	0.441

Given the distribution of measurements against their specifications in Table 4, it can be ascertained that the prediction model is indeed agnostic to the load and lubrication conditions. This conclusion is backed up by the breakdown of metrics present in Table 5, from which it can be observed that the prediction error was below 1 mm^3 for nearly all cases. The tabulated MAE metrics also certify that the χ^2 kernel ridge based prediction model is generalized. The overall averaged R^2 and MAE metrics are listed in Table 6 for the various training and validation sets. It can also be observed from the various validation sets that the surface wear was successfully predicted with an average disparity significantly under 1 mm^3 between the predicted volume and the ground truth.

Table 6. Performance of KRR model on different training and validation sets.

Metric Model	R^2		MAE (mm^3)	
	Training Set	Validation Set	Training Set	Validation Set
1	0.960	0.972	0.824	0.635
2	0.960	0.952	0.814	0.817
3	0.960	0.954	0.821	0.799
4	0.960	0.956	0.827	0.678
5	0.960	0.945	0.824	0.765
6	0.960	0.957	0.819	0.752
7	0.960	0.957	0.820	0.780
8	0.958	0.962	0.834	0.805
9	0.960	0.945	0.820	0.763
10	0.959	0.967	0.828	0.759
Average	0.960	0.957	0.823	0.756

4. Conclusions

In this work, the feasibility of predicting surface wear for a wear process on a pin on disc setup via airborne noise was studied. Qualitative signal analysis was performed on the recorded measurements in order to determine an appropriate methodology for feature extraction. The nature of the experiment and the manner in which wear was recorded were taken into account, and the task was approached from a time series perspective to allow the wear prediction to be modeled as a standard regression problem.

Possible regression models were examined in order to engineer an optimal solution for the input feature space. Consequently, χ^2 kernel ridge regression provided predictions for surface wear with an average accuracy of within 1 mm^3 for both the training and validation sets of measurements, while being agnostic to the load and lubrication specification of each measurement. Performance of the model was cross-validated using multiple training and validation sets. Hence, it can be asserted that a functional quantification between surface wear and airborne noise, via its PSD, has been obtained.

Extensive validation of this trained model on unseen tribopair setups, i.e., new load and lubrication conditions, needs to be ascertained and will be carried out in a future study.

Author Contributions: Conceptualization, A.R., M.K. and A.M.; investigation, A.R. and M.K.; formal analysis, A.R.; writing—original draft preparation, A.R.; writing—review and editing, A.R. and A.M.; supervision, M.K. and A.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

β	penalty term in ridge regression
$\boldsymbol{w}$	weight vector in least-squares linear regression
$\boldsymbol{X}$	matrix of input feature space for the measurements
$\boldsymbol{y}$	vector of target variable in regression problem
γ	kernel variance
κ	kernel function
μ	mean of a signal
$\tilde{C}_{XX}(e^{jw})$	power spectral density estimate
D	number of segments
d	width of wear track
E	expectation operator
h	time horizon
$I^M_{WX,WX}(e^{jw},k)$	periodogram for segment of length M from signal X
K	number of segments
M	segment length of a signal, used in the Welch Method
N	total length of the signal
R	radius of wear track
r	pin end radius
R_{xx}	autocorrelation sequence of signal X
V	volumetric wear of disc in mm^3
$w(n)$	window function
X_h	Spectral features for the noise measurement at horizon h
y_h	surface wear measured at horizon h
Z	normalization factor

References

1. Khan, M.; Basit, K.; Khan, S.; Khan, K.; Starr, A.G. Experimental assessment of multiple contact wear using airborne noise under dry and lubricated conditions. *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.* **2017**, *231*, 1503–1516. [CrossRef]
2. Li, X. A brief review: Acoustic emission method for tool wear monitoring during turning. *Int. J. Mach. Tools Manuf.* **2002**, *42*, 157–165. [CrossRef]
3. Hase, A.; Wada, M.; Mishina, H. The relationship between acoustic emissions and wear particles for repeated dry rubbing. *Wear* **2008**, *265*, 831–839. [CrossRef]
4. Shirakashi, T.; Gong, W.; Obikawa, T. In-process Monitoring of Tool Damage by Active Method-Behavior of Damping Ratio with Tool Wear Development. *J. Jpn. Soc. Precis. Eng.* **1995**, *61*, 1750–1754. [CrossRef]
5. Dybała, J.; Zimroz, R. Rolling bearing diagnosing method based on empirical mode decomposition of machine vibration signal. *Appl. Acoust.* **2014**, *77*, 195–203. [CrossRef]
6. Gandhi, O.; Sharma, J. Signature response for friction and wear. *Asle Trans.* **1979**, *22*, 365–368. [CrossRef]
7. Da Silva, R.H.L.; da Silva, M.B.; Hassui, A. A probabilistic neural network applied in monitoring tool wear in the end milling operation via acoustic emission and cutting power signals. *Mach. Sci. Technol.* **2016**, *20*, 386–405. [CrossRef]
8. Prakash, M.; Kanthababu, M.; Rajurkar, K. Investigations on the effects of tool wear on chip formation mechanism and chip morphology using acoustic emission signal in the microendmilling of aluminum alloy. *Int. J. Adv. Manuf. Technol.* **2015**, *77*, 1499–1511. [CrossRef]
9. Trabelsi, H.; Kannatey-Asibu, E. Tool wear and sound radiation in metal cutting. *Modeling of Machine Tools: Accuracy, Dynamics, and Control*; American Society of Mechanical: New York, NY, USA, 1990; pp. 121–131.
10. Duspara, M.; Sabo, K.; Stoic, A. Acoustic emission as tool wear monitoring. *Teh. Vjesn. Tech. Gaz.* **2014**, *21*, 1097–1101.
11. Kagnaya, T.; Boher, C.; Lambert, L.; Lazard, M.; Cutard, T. Wear mechanisms of WC–Co cutting tools from high-speed tribological tests. *Wear* **2009**, *267*, 890–897. [CrossRef]
12. Hase, A.; Wada, M.; Koga, T.; Mishina, H. The relationship between acoustic emission signals and cutting phenomena in turning process. *Int. J. Adv. Manuf. Technol.* **2014**, *70*, 947–955. [CrossRef]
13. Hisakado, T.; Warashina, T. Relationship between friction and wear properties and acoustic emission characteristics: Iron pin on hardened bearing steel disk. *Wear* **1998**, *216*, 1–7. [CrossRef]

14. Yan, W.; O'Dowd, N.P.; Busso, E.P. Numerical study of sliding wear caused by a loaded pin on a rotating disc. *J. Mech. Phys. Solids* **2002**, *50*, 449–470. [CrossRef]
15. Fan, Y.; Gu, F.; Ball, A. Modelling acoustic emissions generated by sliding friction. *Wear* **2010**, *268*, 811–815. [CrossRef]
16. Ghazali, M.J.; Mansor, N.; Nuawi, M.Z.; Ghani, J.A. Facile wear failure monitoring of commercial bearing alloys using I-kaz method. *Eng. Fail. Anal.* **2013**, *35*, 545–553. [CrossRef]
17. ASTM G99-95a(2000)e1. *Standard Test Method for Wear Testing with a Pin-on-Disk Apparatus*; ASTM International: West Conshohocken, PA, USA, 2000.
18. Jiaa, C.; Dornfeld, D. Experimental studies of sliding friction and wear via acoustic emission signal analysis. *Wear* **1990**, *139*, 403–424. [CrossRef]
19. Dutta, S.; Pal, S.K.; Sen, R. On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression. *Precis. Eng.* **2016**, *43*, 34–42. [CrossRef]
20. Kong, D.; Chen, Y.; Li, N.; Tan, S. Tool wear monitoring based on kernel principal component analysis and v-support vector regression. *Int. J. Adv. Manuf. Technol.* **2017**, *89*, 175–190. [CrossRef]
21. Serrao, P.; Prabhu, R.; Chiranth, B.; Mohammed, Y. Application of Taguchi Method to Predict the Abrasive Wear Behavior of CP Titanium. *J. Mech. Eng. Autom.* **2016**, *6*, 13–17.
22. Edmonds, J.; Resner, M.S.; Shkarlet, K. Detection of precursor wear debris in lubrication systems. In Proceedings of the 2000 IEEE Aerospace Conference Proceedings, Big Sky, MT, USA, 25 March 2000; IEEE: Piscataway, NJ, USA 2000; Volume 6, pp. 73–77.
23. Miettinen, J.; Andersson, P. Acoustic emission of rolling bearings lubricated with contaminated grease. *Tribol. Int.* **2000**, *33*, 777–787. [CrossRef]
24. Filonenko, S.; Nimchenko, T.; Kosmach, A. Model of acoustic emission signal at the prevailing mechanism of composite material mechanical destruction. *Aviation* **2010**, *14*, 95–103. [CrossRef]
25. Ahmaida, A.; Zhen, D.; Gu, F.; Ball, A. Gear wear process monitoring using acoustic signals. In Proceedings of the 21st International Congress on Sound and Vibration 2014, Beijing, China, 13–17 July 2014; International Institute of Acoustics and Vibrations: Auburn, AL, USA, 2004; pp. 2742–2749.
26. Hase, A.; Mishina, H.; Wada, M. Microscopic study on the relationship between AE signal and wear amount. *Wear* **2013**, *308*, 142–147. [CrossRef]
27. Karim, Z.; Nuawi, M.Z.; Ghani, J.A.; Ghazali, M.J.; Abdullah, S.; Mansor, N. Sliding wear evaluation of aluminum alloy (7075-T6) on hardened steel (AISI4340) via non-contact technique by I-kaz™ multi-level analysis. *Wear* **2015**, *334*, 99–104. [CrossRef]
28. Seemuang, N.; McLeay, T.; Slatter, T. Using spindle noise to monitor tool wear in a turning process. *Int. J. Adv. Manuf. Technol.* **2016**, *86*, 2781–2790. [CrossRef]
29. Fekrmandi, H.; Unal, M.; Baghalian, A.; Tashakori, S.; Oyola, K.; Alsenawi, A.; Tansel, I.N. A non-contact method for part-based process performance monitoring in end milling operations. *Int. J. Adv. Manuf. Technol.* **2016**, *83*, 13–20. [CrossRef]
30. Snr, D.E.D. Sensor signals for tool-wear monitoring in metal cutting operations—A review of methods. *Int. J. Mach. Tools Manuf.* **2000**, *40*, 1073–1098.
31. Liang, S.; Dornfeld, D. Tool wear detection using time series analysis of acoustic emission. *J. Eng. Ind.* **1989**, *111*, 199–205. [CrossRef]
32. Sandsten, M. Time-Frequency Analysis of Time-Varying Signals and Non-Stationary Processes. Lecture Notes; Lund University, 2018. Available online: http://www.maths.lu.se/fileadmin/maths/personal_staff/mariasandsten/TFkompver4.pdf (accessed on 31 December 2020).
33. Kay, S. *Intuitive Probability and Random Processes Using MATLAB®*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.
34. Leon-Garcia, A. *Probability, Statistics, and Random Processes for Electrical Engineering*; Pearson Education: London, UK, 2017.
35. Stoica, P. *Introduction to Spectral Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 1997.
36. Stoica, P.; Moses, R.L. *Spectral Analysis of Signals*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2005.
37. Jayaweera, S.K. *Signal Processing for Cognitive Radios*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
38. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **1967**, *15*, 70–73. [CrossRef]
39. Hayes, M.H. *Statistical Digital Signal Processing and Modeling*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
40. Rahi, P.K.; Mehra, R. Änalysis of power spectrum estimation using Welch method for various window techniques. *Int. J. Emerg. Technol. Eng.* **2014**, *2*, 106–109.
41. Weigend, A.S. *Time Series Prediction: Forecasting the Future and Understanding the Past*; Routledge: Abingdon, UK, 2018.
42. Bontempi, G.; Taieb, S.B.; Le Borgne, Y.A. Machine learning strategies for time series forecasting. In *European Business Intelligence Summer School*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 62–77.
43. Taieb, S.B.; Sorjamaa, A.; Bontempi, G. Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomputing* **2010**, *73*, 1950–1957. [CrossRef]
44. Van Vaerenbergh, S.; Via, J.; Santamaría, I. A sliding-window kernel RLS algorithm and its application to nonlinear channel identification. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, Toulouse, France, 14–19 May 2006; IEEE: Piscataway, NJ, USA, 2006; Volume 5, p. V.

45. Wold, S.; Ruhe, A.; Wold, H.; Dunn, W., III. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735–743. [CrossRef]

46. Friedman, J.; Hastie, T.; Tibshirani, R. *The Elements of Statistical Learning*; Series in Statistics; Springer: New York, NY, USA, 2001; Volume 1.

47. Price, B. Ridge regression: Application to nonexperimental data. *Psychol. Bull.* **1977**, *84*, 759. [CrossRef]

48. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]

49. Frank, L.E.; Friedman, J.H. A statistical view of some chemometrics regression tools. *Technometrics* **1993**, *35*, 109–135. [CrossRef]

50. An, S.; Liu, W.; Venkatesh, S. Face recognition using kernel ridge regression. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1–7.

51. Zhang, J.; Marszałek, M.; Lazebnik, S.; Schmid, C. Local features and kernels for classification of texture and object categories: A comprehensive study. *Int. J. Comput. Vis.* **2007**, *73*, 213–238. [CrossRef]

52. Demir, B.; Bruzzone, L. Fast and accurate image classification with histogram based features and additive kernel SVM. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 2350–2353.

53. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

54. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [CrossRef]

55. Cameron, A.C.; Windmeijer, F.A. An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econom.* **1997**, *77*, 329–342. [CrossRef]

56. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]