Evaluation of Acoustic Emission as a Predictor of Laser Power in Laser Welding

IOP Conf. Series: Materials Science and Engineering **1332** (2025) 012041 doi:10.1088/1757-899X/1332/1/012041

Evaluation of Acoustic Emission as a Predictor of Laser Power in Laser Welding

H H Libutti-Núñez[1], L-W Hsu[1], S Parchegani[1], K S B Ribeiro[1], W Moreira Bessa[1], A Salminen[1]

[1]Department of Mechanical and Materials Engineering, University of Turku, Turku, Finland

E-mail: antti.salminen@utu.fi

Abstract. In Laser Welding (LW), multiple sources of data can be used to perform process monitoring. Acoustic Emission (AE) has demonstrated advantages since it does not require severe adaptations into the existing system. Optical microphones, specifically, are capable of sampling signals in the order of MHz, opening a vast possibility for monitoring on high frequency domains. In this work, two methodologies of processing AE are presented, assessing the potential of optical microphones as a robust data source for LW and predictor of the laser power. The experimental setup consisted of 22 bead-on-plate runs on E36 steel, with different laser powers, from 1 kW to 6 kW in 500 W intervals. The experiment was monitored via an optical AE microphone at a sampling rate of 2 MHz, and the acquired signals were split in segments of regular intervals. The first methodology is based on the TSFEL library for feature extraction from the data and the usage of Machine Learning (ML) regressors to predict the laser power. The second is based on computing spectrograms using Short Time Fourier Transform (STFT) and a Convolutional Neural Network (CNN) to predict the laser power. Additionally, each datapoint was then transformed via a 2-dimensional Principal Component Analysis (PCA) reduction for qualitative evaluation. Based on a test set evaluation on unseen data, both methods have achieved a strong prediction performance for the laser power, resulting in a $R^2$of approximately 0.92 and MAE of approximately 0.3kW. The methodology proposed in this work presents an advancement in AE processing, enabling a digital-first, automated LW monitoring system.

1 Introduction

Laser-based manufacturing is becoming a key methodology in several modern industrial sectors. Apart from advanced techniques such as power control and beam shaping, the key factors that make it an attractive technology is the smaller laser beam focal point size, which allows for higher precision applications. In Laser Welding (LW) a direct implication of focal size is the low power input, which avoids severe thermal distortion and rework. However, the broader market adoption of laser welding remains limited. Process defects, which include porosity, cracking, misalignment, undercut, and other defects associated with the highly dynamic nature of the melt pool, degrade product quality and result in wasteful rework and economic losses [1].

Process monitoring in laser welding leverages a combination of sensor technologies and advanced data processing to enable real-time process optimisation and defect detection. Infrared and thermal cameras are routinely employed to capture temperature distributions and melt pool dynamics; however, these systems require constant emissivity calibration and can be adversely affected by splatter or fumes [2]. Simultaneously, previous research has demonstrated that Acoustic Emission (AE) can be a valuable source of information for LW, especially when working in high frequency domains [3]. Currently, there is active research in using AE in the identification of process anomalies and discontinuities, as well as in mapping these irregularities to specific defects [4]. AE monitoring has strongly developed over the past decade, notably with the development of optical microphones that operate based on the interferometry principle: the device consists of a pair of mirrors and a laser source and detector, through which it can detect the difference in air pressure [5]. The main advantage of optical microphones is the extremely fast acquisition rate, currently exceeding the MHz order, while being immune to electromagnetic interference, unlike piezoelectric transducers. A higher sampling rate allows for a broader spectrum analysis. At

the same time, advances in data processing techniques and computing power allow for models that can learn representations and patterns in data, and at the same time, faster computers allow real time process analysis. Recent studies have combined highly successful Machine Learning (ML) techniques and AE data in various engineering fields [4, 6].
Laser power is one of the major parameters in LW, together with the welding speed [7], therefore, externally monitoring the laser power can provide important information, especially when debris on optical components such as fibre ends and lens windows reduce the power delivery, consequently changing process parameters.
This work elaborates on optical microphone AE to predict the in process laser power. Applications of this methodology include online process diagnostics, Non-Destructive Testing (NDT), quality control and feedback control systems.

2 Methods

The present work elaborates on two alternatives for predicting the laser power from acoustic emission signals in LW, which is currently a knowledge gap in process monitoring. The first method is based on data preparation by the TSFEL signal processing library [8], automatically providing a rich set of metrics based on the input data, and the subsequent use of commonly used ML regressors to predict the laser power based on the extracted features. The second methodology uses Short Time Fourier Transform (STFT) to create spectrograms and a Convolutional Neural Network (CNN) to predict laser power based on the spectrograms. Since CNNs typically operate on images, they have the potential of recognising patterns on the 2-dimensional spectrograms of frequency over time.
In this regard, a set of bead-on-plate experiments were performed using different laser powers and monitored with an optical microphone. Our experimental setup is based on an IPG YLS-10000-MM fibre laser mounted on an ABB IRB 4600 6-axis articulated general purpose Robot. The lens windows have been verified prior to every experiment run to ensure the fidelity of the programmed power. The robot was operated under control logic, enabling synchronised activation/deactivation cycles and motion coordination. A constant welding speed of $3\,\mathrm{m\,min^{-1}}$ was maintained. Figure 1 demonstrates the experimental setup, indicating the base plate, sensor position, and other components of the setup.

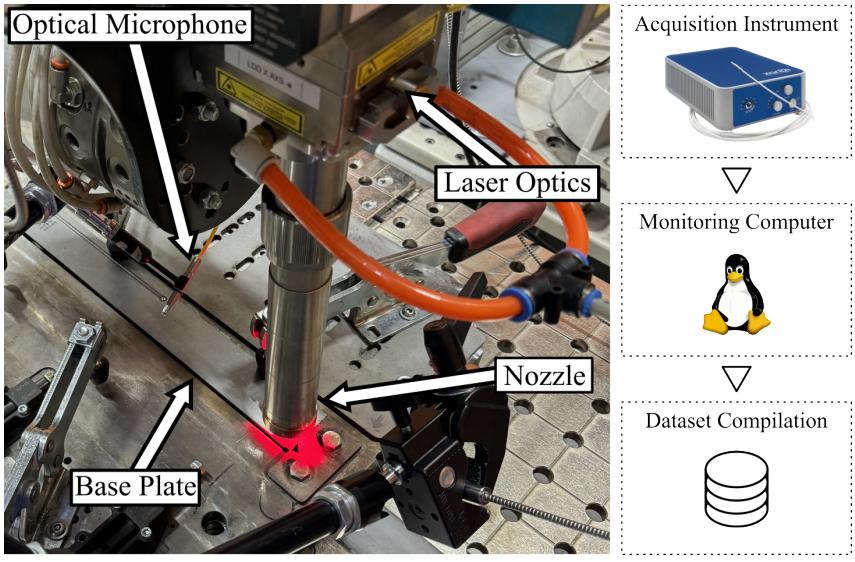

Figure 1: Experimental setup and data collection scheme.

The selected microphone was a Xarion Eta250 Ultra optical microphone, positioned $140\,\mathrm{mm}$ from the process zone, with a set acquisition rate of $2\,\mathrm{MHz}$. The microphone was activated 200 milliseconds prior to laser initiation and deactivated 300 milliseconds afterwards to ensure all valid experiment data is included in the recordings. Data acquisition and processing were performed by a QASS Optimizer4D industrial computer equipped with a preamplifier. In total, 22 bead on plate runs were performed on NV E36 structural steel from SSAB, which were cut with fibre laser into workpieces with dimensions of $400\,\mathrm{mm}$ x $50\,\mathrm{mm}$ x $5\,\mathrm{mm}$. The bead on plate runs were

performed with laser powers ranging from 1 kW to 6 kW in 500 W increments. For each power, two runs have been performed under a steady co-axial argon shielding gas supply, delivered through the nozzle depicted in figure 1.
The data pre-processing consists of selecting the relevant portion of the signals, slicing into segments of equal length. The data selection was based on a manual analysis of the time domain signals, resulting on an interval within the stable laser power that ignores the start and stop curves for the laser. The automatic triggering of the microphone allowed it to capture uniform-length segments, resulting in recordings of approximately 3.54s in length. Figure 2 shows the chosen start and stop ranges on time domain plots, on different recordings. The total collected data, therefore, consisted of 22 recordings of 7 077 888 audio samples each.

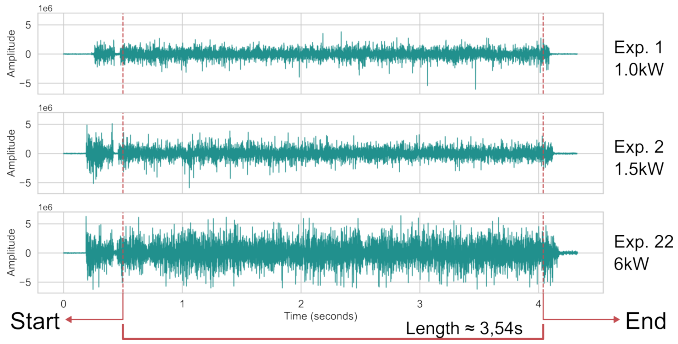

Figure 2: Signal selection interval across recordings.

Following data selection, the signals were processed through the pipeline depicted in Figure 3. A strict 50:50 train-test split (Figure 3b) that isolates data from entire experimental runs within either subset. This prevents data leakage by ensuring that audio signals from individual welding trials remain exclusively in training or testing datasets. For this dataset, two experiments were done at each power setting, resulting on 11 for training and 11 for testing.
Next, each recording is split into 54 equally sized windows (figure 3c), which serve as short observations of the LW process, each with a length of approximately 66 ms. The reasoning behind the windowing was based on the length of the input data (which is computationally heavier), the possibility of creating more data observations (useful for training and evaluation), and creating a prediction methodology capable of being applied in real time. This creates the case for a data processing method that operates on short observation windows. Next, we employed two main methodologies for relating AE signals to laser power. Firstly, we perform feature extraction using a widely adopted library for time-series analysis (figure 3d.1). Secondly, we explore a deep learning approach using spectrogram representations of the signal segments (figure 3d.2). Finally, the predictions are evaluated on the same basis, using widely adopted scoring functions (figure 3e).

2.1 TSFEL

The Time Series Feature Extraction Library (TSFEL) is a Python package that generates an extensive set of metrics for an input time series. The full list of features can be accessed in the official project documentation website [8]. In this study, the full configuration was used, i.e. extracting all the available possible features from the system. For each segment, 312 features were extracted, including spectral skewness and wavelet transformation metrics, resulting in a table of several observations for each experiment and their corresponding features.
The library post-processing functions for removing redundant and low-variance variables, yielding 105 selected features. The feature extraction via TSFEL is followed by the training and evaluation using different ML methods. The selected regressors were Support Vector Regressor (SVR) and Random Forest Regressor (RFR), since they are widely adopted and based on different modelling strategies. The dataset has been standardised prior to the training and evaluation.

2.2 CNN

As a counterpart to the static feature extraction method of the last subsection, a deep learning model was implemented. For this, each window was transformed to a spectrogram, with the use of Scipy signal processing library. The parameters for the STFT spectrogram generation were a Hann window of width 4096, with 75% of

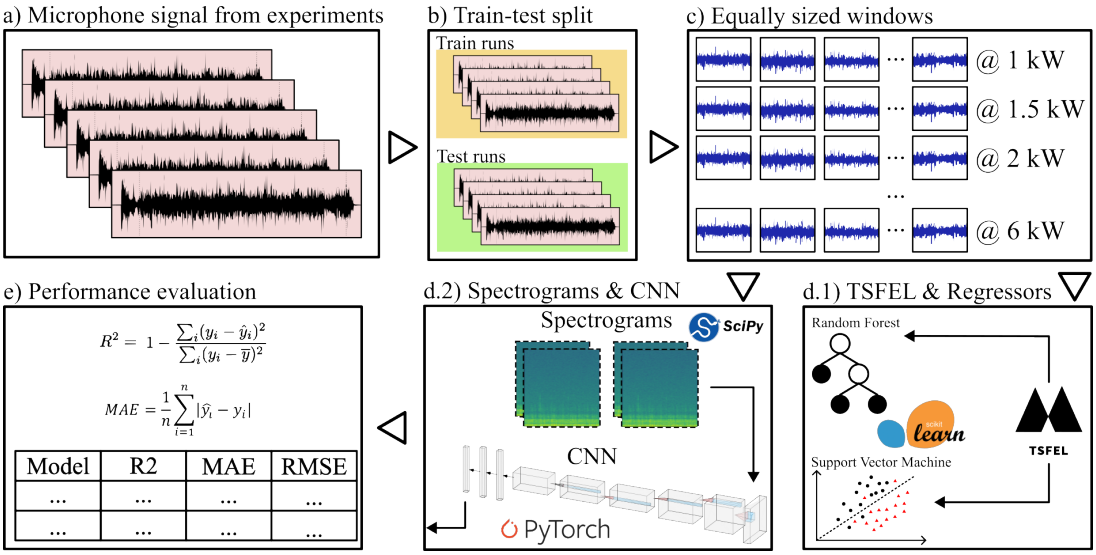

Figure 3: Data processing steps.

overlap between each window. The spectrograms were then converted to the dB scale for range compression. The spectrograms of 2049x131 elements were cropped and padded as a 2048x128 2-D array, since base 2-sized arrays facilitate data transformations. This data layout naturally matches the one of images, which enables the use of CNNs, that have been part of breakthroughs in Artificial Intelligence since the past 2 decades. A CNN was implemented using the Pytorch framework. Together with the CNN, a feature extraction module was also implemented using the encoder step of the network. For that, the feature map resulting from the encoder step was flattened and interpreted as a 1-d vector, to be used as a qualitative comparison for features extracted by TSFEL. The training was carried out using the Adam optimiser, with an initial learning rate of 0.001, and the model was trained for 100 epochs. The used loss function was MSE, which is suited for predicting real-valued variables.

2.3 Model Evaluation

As stated in figure 3, the testing portion of the data come from experiments alien to the model training. Although this results in a smaller train to test ratio, in this case 50:50, this prevents data leakage, contributing to more robust and significant evaluation scores. As with the training, known labels from the experiment were used for the model evaluation, and the comparison between the predicted output and the known target variable results in the score. To compare both proposed methods, the scoring functions R^2, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were implemented, with definitions stated in equations 1, 2, and 3. In all three equations, y_i is the i-th true value in the dataset, $\hat{y}_i$ is the i-th predicted value, and $\bar{y}$ is the mean of the true values.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

3 Results

The dataset preparation resulted in 1188 audio windows, each containing 131 072 audio samples. This dataset was subsequently split into a training set of 594 windows and a test set of 594 windows.

For each signal window, feature extraction and model training steps were performed according to the scheme illustrated in figure 3. The test dataset was then used to evaluate the quality of the features extracted by the TSFEL library (method 2.1) and the representational power of the features learned by the CNN encoder (method 2.2).
Plots in a lower-dimensional space were generated using Principal Component Analysis (PCA) for a qualitative assessment of how well the learned features capture patterns within the data. The PCA plots for features derived from TSFEL and those extracted from the CNN are depicted in figure 4. In these plots, the colour of each point indicates the laser power used during the signal acquisition. It can be observed that, for both sets of features, data points corresponding to similar laser power levels tend to cluster in proximity within the 2D PCA space, suggesting that both feature extraction approaches have successfully captured variance related to the input laser power. The explained variance for the first 2 principal components of TSFEL (PCA1 and PCA2) was approximately 25%. For PC1, the features constituting those principal components were the "absolute energy", and the "mean spectrogram coefficient at 580645Hz", what may indicate that the spectrum surrounding 580kHz is affected by different laser powers, in our configuration. For PC2, the wavelet entropy, median frequency and neighbourhood peaks features were the ones contributing the most. Alternatively, the PC1 and PC2 components for the CNN features are responsible for approximately 73% of the variance.

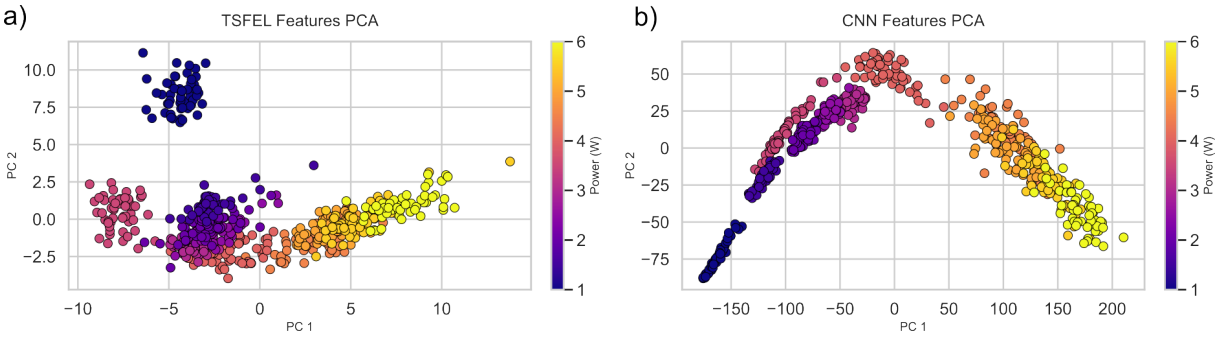

Figure 4: Scatter plots of features reduced to two principal components via PCA. Points are coloured by laser power. (a) TSFEL-derived features; (b) CNN-extracted features.

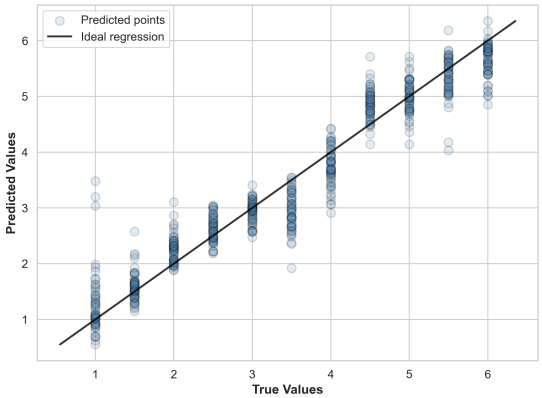

Figure 5: Predicted versus true values for the TSFEL + SVR methodology.

A plot consisting of predicted versus true values is depicted in figure 5, demonstrating the trend and apparently stable error distribution across the entire dataset, indicating that the error magnitude is also approximately constant across values. Next, the predictive performance of the models associated with each feature set was quantitatively evaluated using the R^2, MAE, and RMSE metrics. The test set served as the input for predictions, and the resulting predicted values were compared against the true laser power labels. The evaluated metrics for each methodology are presented in table 1.
According to the evaluation metrics in Table 1, the "TSFEL + SVR" method achieved the highest R^2performance with a score of 0.926, and an MAE of 0.309 kW, and a RMSE of 0.429 kW^2. The "TSFEL + RFR" also presented a

Table 1: Regression performance comparison for laser power prediction on the test set.

Method	R^2	MAE (kW)	RMSE (kW)
TSFEL + RFR	0.918	0.246	0.452
TSFEL + SVR	0.926	0.309	0.429
CNN	0.913	0.339	0.458

lower but comparable performance, with a $R^2$of 0.918. The CNN approach yielded an $R^2$of 0.913. These R^2values, all over 0.9, indicate that all three approaches can capture a significant portion of the variance in the laser power based on AE signals. The performance of the models can be further increased by running more experiments, which will result in more training and testing data, and by exploring different sets of hyperparameters, which in the case of the CNN is the number of layers and architecture.

4 Conclusion

This study addressed process monitoring and data processing in the context of LW. Two methodologies for processing AE signals from LW were developed to predict laser power. Bead on plate experiments were performed using different programmed laser powers, as a data source. A set of experiments was used to train ML models, and another set was used as the evaluation dataset. The first data processing method was based on a time series processing library, followed by the training and evaluation of the SVR and RFR models. In parallel, a deep learning based method was implemented, using a STFT pre-processing and conversion from signal to spectrograms, followed by the training of a CNN to predict a single real-valued variable, the laser power. The evaluation over different sets of metrics demonstrate that both methods demonstrate a strong predictive power. The library-based method achieved a R^2, MAE and RMSE of 0.92, 0.29 and 0.20, respectively. Comparatively, the deep learning based method achieved a R^2, MAE and RMSE of 0.91, 0.36 and 0.22, respectively. The results suggest that AE can be successfully employed for predicting laser power from LW, given the performance of the models.
Future developments of this research include a more detailed investigation of the performance of the regressors by using different models and hyperparameters and collecting a larger dataset.

Acknowledgements
The research team acknowledges for the support from the project CaNeLis - Carbon-neutral lightweight ship structures using advanced design, production, and life-cycle services, which is funded by Meyer Turku Oy, Cavitar Oy, SSAB Europe Oy and Business Finland (3360/31/2022). The project is part of NEcOLEAP leading company of Meyer Turku Oy.

References

[1] Angeloni C, Francioso M, Liverani E, Ascari A, Fortunato A and Tomesani L 2024 *Lasers Manuf. Mater. Process.* **11** 3–24 doi:10.1007/s40516-023-00216-7

[2] Cai W, Wang J, Zhou Q, Yang Y and Jiang P 2019 *Proceedings of the 5th International Conference on Mechatronics and Robotics Engineering* 9–15 doi:10.1145/3314493.3314508

[3] Tomcic L, Ederer A, Grabmann S, Kick M, Kriegler J and Zaeh M F 2022 *Journal of Laser Applications* **34** 042052 doi:10.2351/7.0000796

[4] Schmidt L, Römer F, Böttger D, Leinenbach F, Straß B, Wolter B, Schricker K, Seibold M, Pierre Bergmann J and Del Galdo G 2020 *Procedia CIRP* **94** 763–768 doi:10.1016/j.procir.2020.09.139

[5] Fischer B 2016 *Nature Photon* **10** 356–358 doi:10.1038/nphoton.2016.95

[6] Manthei G and Guckert M 2022 *J Nondestruct Eval* **42** 4 doi:10.1007/s10921-022-00913-x

[7] Maculotti G, Genta G and Galetto M 2024 *Quality and Reliability Engineering International* **40** 202–219 doi:10.1002/qre.3377

[8] Barandas M, Folgado D, Fernandes L, Santos S, Abreu M, Bota P, Liu H, Schultz T and Gamboa H 2020 *SoftwareX* **11** 100456 doi:10.1016/j.softx.2020.100456