Meteorological-based oil temperature prediction in wind turbine gearbox using deep learning

Paul A. Adedeji [a,b] , Obafemi O. Olatunji [a,b], Nkosinathi Madushele [a], Zelda Z. Rasmeni [a,b], Nickey Janse van Rensburg [b]

[a] Department of Mechanical Engineering Science, University of Johannesburg, South Africa
[b] Process, Energy, and Environmental Technology Station, University of Johannesburg, South Africa

BACKGROUND

- About 20% of overall downtime of a wind turbine (WT) across its lifetime is often caused by gearbox failure (Tautz-weinert and Watson, 2016; Bruce et al., 2015).
- Accurate predicting gearbox oil temperature is very relevant to the health monitoring of the WT gearbox.
- While the climatic operational conditions of the WT can influence the temperature of the gearbox oil, such influence is more pronounced in sites where spatio-temporal variation of climatic conditions is highly dynamic.
- Deep learning techniques offers a high prospects but this must be compared to other machine learning models like the Long Short-Term Memory (LSTM), for effectiveness and efficiency

Motivation

Monitoring WT gearbox conditions have been carried out from time series data from the SCADA approach, however,

i. Meteorological variables play significant roles in its influence and this has not been considered
ii. Comparing deep learning models with evolutionary-based soft computing and regression-based models in this space is less explored
iii. Investigating the influence of several model performance evaluation criteria on gear oil temperature prediction to prevent model bias (Adedeji et al., 2022, 2020) is sparse in the literature

Aim and Objectives

The aim of this study is to investigate the effectiveness of the LSTM deep learning model in predicting gearbox oil temperature of a WT gearbox.

Objectives

i. Investigate the effectiveness of gearbox oil temperature prediction from meteorological variables
ii. To compare evolutionary-based machine learning models (ANFIS-GA, ANFIS-PSO) and support vector regression (SVR) with deep learning LSTM model for gearbox oil temperature prediction
iii. To investigate dynamics of several performance metrics in the models developed

METHODOLOGY

Data Collection

- Nojoli wind farm, Eastern Cape, South Africa between 2016 and 2020 – wind farm comprises 44 WTs each with a rated capacity of 2MW
- Wind farm missing data were deleted
- Data Division: 70%- Training ; 30% Testing

Model Architectures

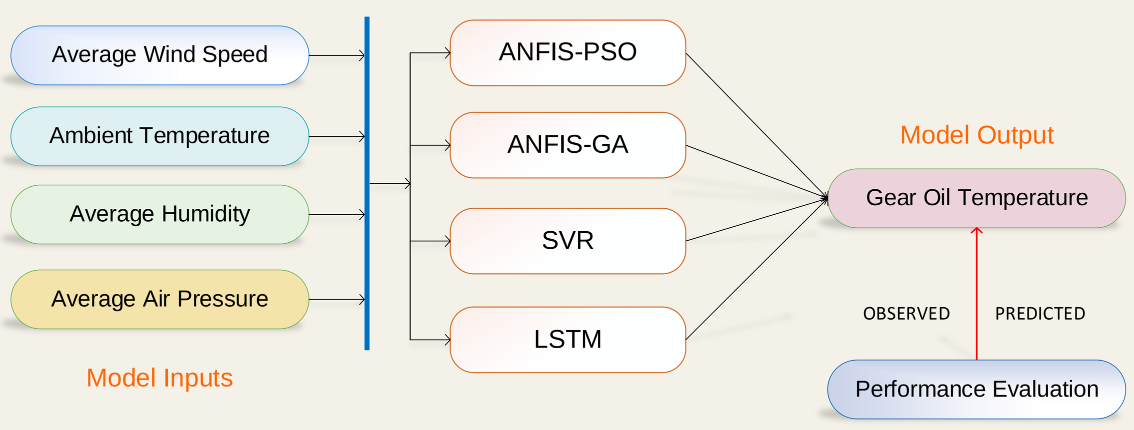

FIGURE 1. Flowchart of Model

PSO & GA Optimized ANFIS Model

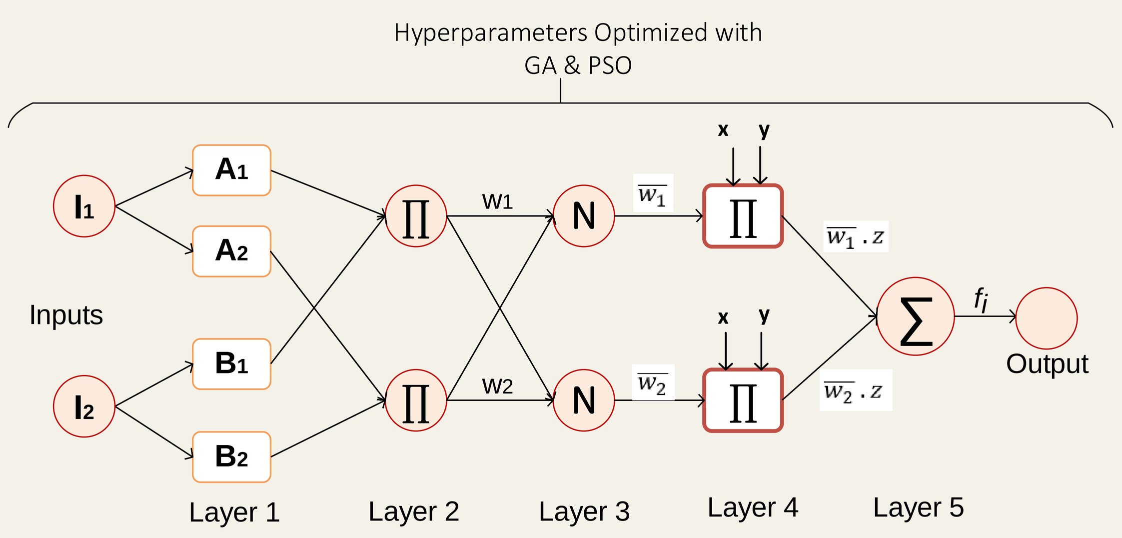

FIGURE 2. Architecture for ANFIS-GA & ANFIS-PSO (Adedeji et al. (2022))

Methodology (Cont'd)

Long Short-term Memory Model

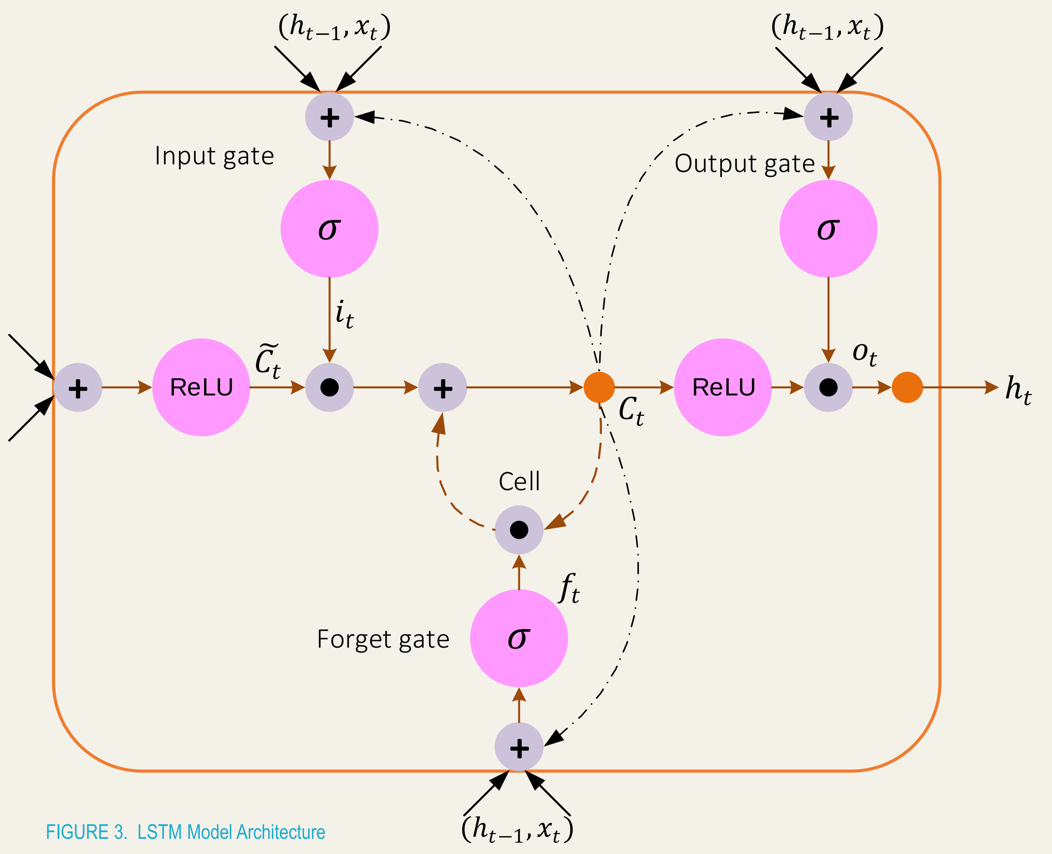

FIGURE 3. LSTM Model Architecture

Support Vector Regression Model

For a given sample space $S = \{(X_p, t_p): 1 \le p \le N\}$ the support vector regression model seeks to fit a linear model $f(X, W, b) = W.X + b$ to minimize the regularized ϵ-intensive loss function, $\mathcal{L}_\epsilon(W, b)$ expressed as (Schölkopf et al., 2002)

$$\mathcal{L}_\epsilon(W, b) = \sum_p \left[t_p - f\left(X_p, W, b\right) \right]_\epsilon + \frac{\lambda}{2} \|W\|^2 \quad (1)$$

such that $[z]_\epsilon = \max(0, |z| - \epsilon)$. An error $t_p - f(X_p, W, b)$ is penalized only if its absolute value is greater than ϵ (Gala et al., 2016). In solving the linear model, the SVR loss function is written as a minimization function over $L(W, b, \xi)$ expressed as a primal problem as follows:

$$L(W, b, \xi) = \frac{1}{2} \|W\|^2 + C \sum_i (\xi_i + \xi_i^*) \quad (2)$$

Subject to

$$\xi_i - \epsilon \le W.X_i + b - y_i \le \xi_i^* + \epsilon, \qquad \xi_i, \xi_i^* \ge 0 \quad (3)$$

where C is the penalty constant, ϵ is the tolerance, ξ distance from loss function.

Model Performance Evaluation Metrics

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N} [y_k - \widehat{y_k}]^2}{N}} \quad (4)$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N} (\widehat{y_k} - y_k)^2}{\sum_{k=1}^{N} (\bar{y} - y_k)^2} \quad (5)$$

$$MAPE = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{y_k - \widehat{y_k}}{y_k} \right| \times 100\,\% \quad (6)$$

$$MAD = \frac{1}{N} \sum_{k=1}^{N} |y_k - \bar{y}| \quad (7)$$

$$rMBE = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{\widehat{y_k} - y_k}{y_k} \right) \quad (8)$$

RESULTS

- Comparison plots of the observed WT oil temperature and predicted were plotted (Figures 4 - 11).
- The values of the performance evaluation metrics applied to each model was also presented (Table 1).

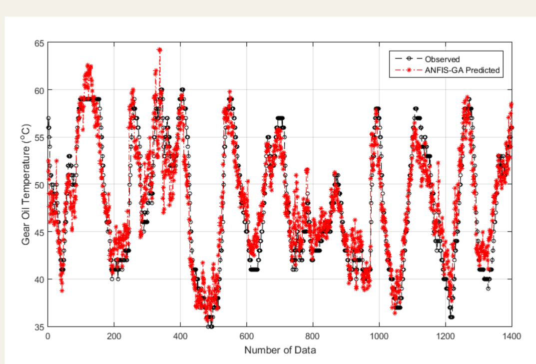

FIGURE 4. Model training with ANFIS-GA

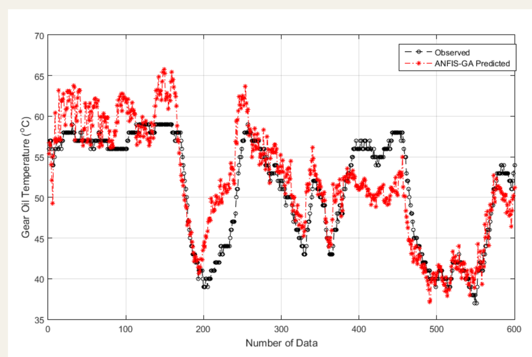

FIGURE 5. Model testing with ANFIS-GA

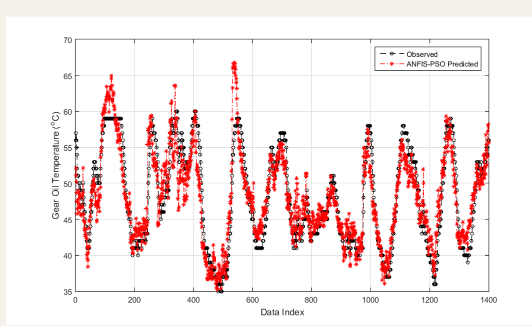

FIGURE 6. Model training with ANFIS-PSO

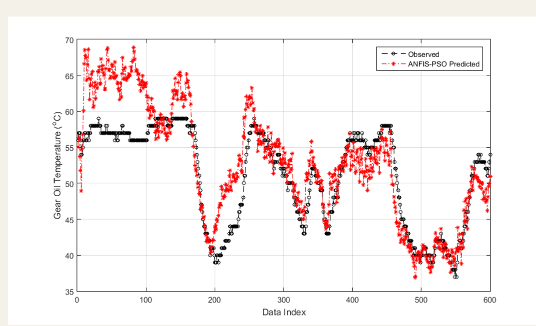

FIGURE 7. Model testing with ANFIS-PSO

NB:

ANFIS- Adaptive Neurofuzzy Inference System
SVR- Support Vector Regression
R^2 – Coefficient of Determination
rMBE- Relative Mean Bias Error
PSO- Particle Swarm Optimization
RMSE- Root Mean Square Error
MAPE- Mean Absolute Percentage Error
LSTM- Long Short-term Memory
GA- Genetic Algorithm
MAD- Mean Absolute Deviation

Results (Cont'd)

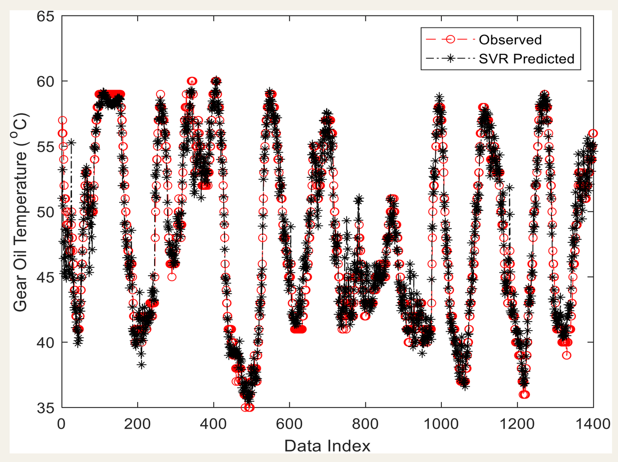

FIGURE 8. Model training with SVR

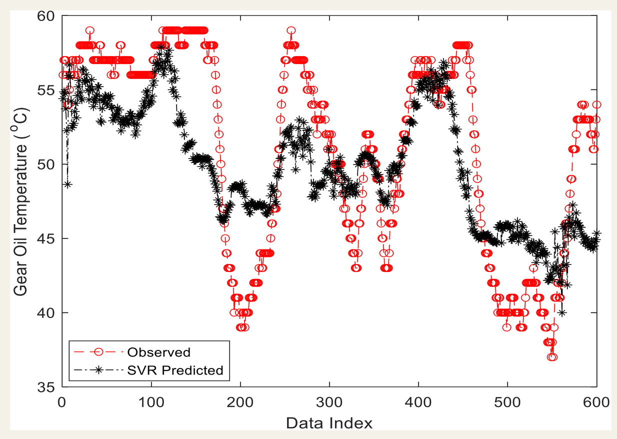

FIGURE 9. Model Testing with SVR

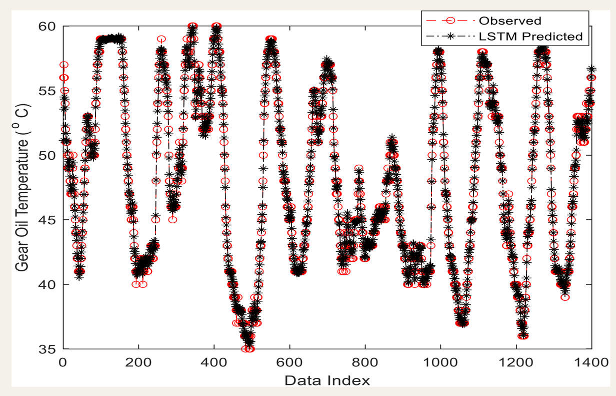

FIGURE 10. Model Training with LSTM

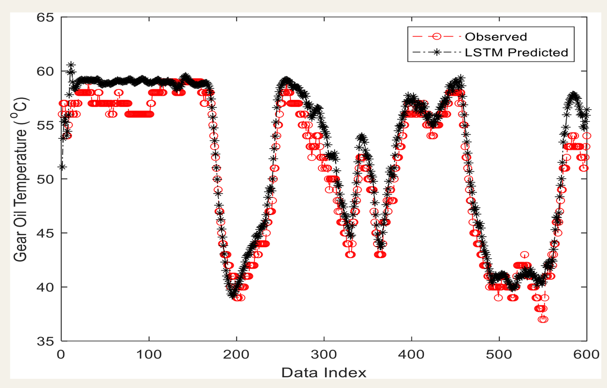

FIGURE 11. Model Testing with LSTM

TABLE 1. Model Performance Evaluation Metrics

	Phase	RMSE	MAD	MAPE	R^2	rMBE	CT (s)
ANFIS-GA	Training	2.61	2.10	4.45	0.850	0.41	118.35
	Testing	3.79	2.93	5.90	0.670	1.74	
ANFIS-PSO	Training	2.81	2.22	4.58	0.830	-5.67E-12	126.79
	Testing	4.48	3.42	6.60	0.550	3.33	
SVR	Training	1.43	1.03	2.20	0.956	-0.041	15.31
	Testing	4.59	3.63	7.70	0.534	-2.34	
LSTM	Training	0.37	0.27	0.58	0.997	0.088	30.00
	Testing	1.82	0.93	3.01	0.926	2.661	

CONCLUSION

- WT gearbox oil temperature can be predicted from meteorological data.
- GA optimized ANFIS performed better than the PSO optimized ANFIS and SVR.
- The LSTM technique, however, outperformed the other models in accuracy (R^2, MAPE), robustness (rMBE) and error deviations (MAD, RMSE), though at a relatively higher computational time compared to SVR.
- LSTM is very useful for the gear oil temperature condition monitoring for the case study wind farm.

REFERENCES

- Adedeji, P.A., Akinlabi, S.A., Madushele, N., Olatunji, O.O., 2022. Evolutionary-based neurofuzzy model with wavelet decomposition for global horizontal irradiance medium-term prediction. J. Ambient Intell. Humaniz. Comput. https://doi.org/10.1007/s12652-021-03639-2
- Adedeji, P.A., Akinlabi, S.A., Madushele, N., Olatunji, O.O., 2020. Neuro-fuzzy resource forecast in site suitability assessment for wind and solar energy: A mini review. J. Clean. Prod. 269, 1–29.
- Bruce, T., Long, H., Dwyer-Joyce, R.S., 2015. Dynamic modelling of wind turbine gearbox bearing loading during transient events. IET Renew. Power Gener. 9, 821–830. https://doi.org/10.1049/iet-rpg.2014.0194
- Gala, Y., Fernández, Á., Díaz, J., Dorronsoro, J.R., 2016. Hybrid machine learning forecasting of solar radiation values. Neurocomputing 176, 48–59. https://doi.org/10.1016/j.neucom.2015.02.078
- Schölkopf, B., Smola, A.J., Bach, F., 2002. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press.
- Tautz-weinert, J., Watson, S.J., 2016. Using SCADA data for wind turbine condition monitoring – a review. IET Renew. Power Gener. 11, 382–394. https://doi.org/10.1049/iet-rpg.2016.0248

ACKNOWLEDGEMENTS

The authors appreciate Nojoli Wind Farm of the Enel Green Power, South Africa for providing the data used for this study and the University of Johannesburg for the support towards carrying out this research.

CONTACT INFORMATION

Corresponding author:

Paul A. ADEDEJI
Tel: +27-655-745-531
Email: padedeji@uj.ac.za