ARTIFICIAL INTELLIGENCE AND CLEAN AIR: DEVELOPMENT OF NOVEL ALGORITHMS WITH MACHINE LEARNING AND DEEP LEARNING

A Thesis submitted by

EKTA SHARMA

BSc (Mathematical Sciences)

MSc, MPhil (Operations Research)

For the award of

Doctor of Philosophy

2022

This Thesis is affectionately dedicated to

Mrs. Suneeta Sharma (1953-2009), my cherished mother,

who was the wind beneath my wings, and who is missed

beyond measure in the milestones of my life.

~ Ekta Sharma

ABSTRACT

Air pollution has detrimental impacts on the people, the environment, and the global economy; however, this issue is somewhat under-recognised in many nations, despite having an advanced healthcare infrastructure. Accurate, reliable, and real-time forecasting of air pollutants is critical to minimise adverse health outcomes, and improving the quality of human life. In comparison with other developed nations, Australia has a relatively higher standard of air quality. Despite this, approximately 5,000 annual deaths are attributable to air-borne issues (AAS 2021). The Environment Protection Agency (EPA-VIC 2016) forecasts that population growth, increased urbanisation, and uncontrolled fossil-fuel emissions may result in summer smog, getting significantly pronounced after 2030. Australia's arid nature with frequent dust-storms and bushfires could further exacerbate the particle-borne air pollutants. Therefore, scientifically verified, and robust artificial intelligence methods are essential for efficient air quality forecasting that can support the Australian Government's health and environment protection policies.

This doctoral research presents a novel study based on the development of computationally efficient artificial intelligence models that can provide early warnings through an air quality forecasting system. In the first objective, a novel hybrid artificial intelligence framework based on machine learning or an online sequential-extreme learning machine algorithm is developed to emulate hourly air quality variables *i.e.*, fine particulates ($PM_{2.5}$), coarse particulates (PM_{10}), and the visibility reducing particles. These are associated with increased respiratory-induced mortality resulting in recurrent healthcare costs. The second objective further advances the first objective by introducing deep learning algorithms to develop a forecasting system for suspended particulate matter. As remote sensing data are also an important predictor for meteorological variables, the third objective has constructed a one-dimensional convolutional neural network (*CNN*) integrated with a one-directional fully gated recurrent unit (*GRU*) model using satellite and ground-based observations producing reliable estimates of $PM_{10.}$ The study areas considered are all major Australian air pollution hotspots, posing a growing hazard to public health. The results reflect that meteorological data, geographical information, and important statistical metrics have further improved the prediction accuracy in all the objectives.

By providing a high-resolution spatiotemporal forecasting system, the doctoral research results may provide scientific innovation and knowledge contributions facilitating further studies, enhancing the understanding and dynamic evolution of airborne pollutants. These findings could also facilitate the necessary pre-emptive actions during bushfires seasons, frequent dust storms in real-time warnings so our citizens can plan and minimise exposure risks. Another significance of the doctoral study is protecting vulnerable population groups (e.g., frail elderly, people with sickness, expectant mothers, and children) so the proposed models can be updated when new, higher-resolution satellites are launched.

In synopsis, the novel state-of-the-art artificial intelligence models in this doctoral research have the potential to provide significant economic and advisory benefits for the government, particularly in air quality and health policy development. The predictive system can also be of global importance, where air quality poses a serious health risk. Therefore, the outcomes of this doctoral research can be adopted in constructing a real-time forecasting system tailored for environment protection industries, public health, governments, and other stakeholders.

CERTIFICATION OF THESIS

This Thesis is the work of ***Ekta Sharma*** except where otherwise acknowledged. The majority of the authorship in the research papers presented as a Thesis by Publication has been undertaken by the student. The work is original and has not previously been submitted for any other award, except where acknowledged.

Ekta Sharma

PhD Candidate

Endorsements

Professor Ravinesh Deo

Principal Supervisor

Professor Jeffrey Soar

Associate Supervisor

Dr Nawin Raj

Associate Supervisor

Honorary Professor Alfio Parisi

Associate Supervisor

Dr Ramendra Prasad

External Supervisor

Student and supervisors' signatures of endorsement are held at USQ.

STATEMENT OF CONTRIBUTION

The doctoral research thesis has produced five quartile 1 (*Q*1) publications, including a research paper that was generated through the support of the Australian Postgraduate Research (APR.Intern) internship program, and an additional reflective piece of the research work completed during the PhD candidature.

Field of Research (FOR): The focus of this doctoral thesis is on the national priority area of: 'Artificial Intelligence and Image Processing FOR-08', Code 0801, 'Atmospheric Sciences FOR-08', Code 0401, 'Environment, SE0-08', Code 96, 'Air Quality, SE0-08', Code 9601, and 'Public Health and Health Services FOR-08', Code 1117.

Articles 1, 2, and 3 are primary (core) parts of this thesis, and Articles 4, 5, and 6 are the secondary contributions placed in the Appendix section as additional research output completed during the PhD candidature.

The following presents the student contributions and the contributions of the co-authors of the publications.

Article 1: Chapter 3

Sharma, E, Deo, RC, Prasad, R & Parisi, AV 2020, 'A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms', *Science of The Total Environment,* Vol. 709, p. 135934-57 **(Scopus Ranked *Q1, Impact Factor: 7.96 and SNIP*: 2.015; *96th percentile in Environmental Science*).**

The percentage contributions for this paper are: Ekta Sharma 80%, Ravinesh C. Deo 10%, Ramendra Prasad 5%, and Alfio V. Parisi 5%.

Author	Task Performed
Ekta Sharma (*Candidate*)	Exploring the methodology in literature, data collection, and analysis, programming, preparation of tables and figures, writing and revising of the manuscript.

Ravinesh C. Deo (*Principal Supervisor*)	Supervised and assisted in model concepts, provided detailed and beneficial comments on the manuscript, edited, and prepared for the submission. Guidance for the choice of Journal and co-authorship of the manuscript.
Ramendra Prasad (*External Supervisor*)	Editing, proofreading, advised in method, interpretation of results, and co-authorship of the manuscript.
Alfio V. Parisi (*Associate Supervisor*)	Editing, proofreading, and co-authorship of the manuscript.

Article 2: Chapter 4

Sharma, E, Deo, RC, Prasad, R, Parisi, AV & Raj, N 2020, 'Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Networks', *IEEE Access*, Vol. 8, pp. 209503-16. **(Scopus Ranked *Q1, Impact Factor: 3.367, and* SNIP: *1.421, 87th percentile in Engineering*).**

The percentage contributions for this paper are: Ekta Sharma 75%, Ravinesh C. Deo 10%, Ramendra Prasad 5%, Alfio V. Parisi 5%, and Nawin Raj 5%.

Author	Task Performed
Ekta Sharma (*Candidate*)	Exploring the methodology in literature, data collection, and analysis, programming, preparation of tables and figures, and writing and revising of the manuscript.
Ravinesh C. Deo (*Principal Supervisor*)	Supervised and assisted in model concepts, provided detailed comments on the manuscript, edited, and prepared for the submission. Guidance for the choice of Journal and co-authorship of the manuscript.
Ramendra Prasad (*External Supervisor*)	Editing, proofreading, advised in method, and co-authorship of the manuscript.

Alfio V. Parisi (*Associate Supervisor*)	Editing, proofreading, and co-authorship of the manuscript.
Nawin Raj (*Associate Supervisor*)	Advised in method, Editing, proofreading, and co-authorship of the manuscript.

Article 3: Chapter 5

Sharma, E, Deo, RC, Soar J, Prasad, R, Parisi, AV & Raj, N 2022, 'Novel hybrid deep learning model for satellite-based PM10 forecasting in most polluted Australian hotspots', *Atmospheric Environment*, Vol. 279, pp. 119111-24. **(Scopus Ranked *Q1, Impact Factor: 4.79 and SNIP*: 1.4; *95th percentile in Environmental Science*).**

The percentage contributions for this paper are: Ekta Sharma 70%, Ravinesh C. Deo 10%, Jeffrey Soar 5%, Ramendra Prasad 5%, Alfio V. Parisi 5%, and Nawin Raj 5%.

Author	Task Performed
Ekta Sharma (*Candidate*)	Exploring the methodology in literature, data analysis, programming, preparation of tables and figures, writing and revising of the manuscript
Ravinesh C. Deo (*Principal Supervisor*)	Supervised and assisted in model concepts, provided detailed comments on the manuscript, edited, and prepared for the submission. Guidance for the choice of Journal and co-authorship of the manuscript.
Jeffrey Soar (*Associate Supervisor*)	Editing, proofreading, and co-authorship of the manuscript
Ramendra Prasad (*External Supervisor*)	Editing, proofreading, interpretation of results, and co-authorship of the manuscript
Alfio V. Parisi (*Associate Supervisor*)	Editing, proofreading, and co-authorship of the manuscript
Nawin Raj (*Associate Supervisor*)	Editing, proofreading, advised in method, and co-authorship of the manuscript

Article 4: Appendix A

Sharma, E, Deo, RC, Soar J, & Yaseen, ZM 2022, 'Association of Air Pollutants and Novel Coronavirus with Deep Learning: A systematic literature review of the case studies in Asia and Oceania', *Sustainable Cities and Society*, Under review **(Scopus Ranked *Q1, Impact Factor: 7.59 and SNIP*: 2.35; *97th percentile in Civil and Structural Engineering*).**

The percentage contributions for this paper are: Ekta Sharma 75%, Ravinesh C. Deo 10%, Jeffrey Soar 5%, and Zaher Mundher Yaseen 10%

Author	Task Performed
Ekta Sharma (*Candidate*)	Exploring the methodology in literature, data analysis, programming, preparation of tables and figures, writing and revising of the manuscript
Ravinesh C. Deo (*Principal Supervisor*)	Supervised and assisted in model concepts, provided detailed comments on the manuscript, edited, and co-authorship of the manuscript.
Jeffrey Soar (*Associate Supervisor*)	Editing, proofreading, and co-authorship of the manuscript
Zaher Mundher Yaseen (*External Collaborator*)	Editing, proofreading, co-authorship of the manuscript, and preparing for submission.

Article 5 (APR.Intern PhD Internship Outcome) - Appendix B

Sharma, E, Shakeel I, Sancho SS, & Deo, RC 2022, 'Artificial Intelligence-based Bit-wise Decoding of Error-correction Codes for Radio Communications', *IEEE Access*, Under review. **(Scopus Ranked *Q1, Impact Factor: 3.367* and *SNIP*: *1.421; 87th percentile in Engineering*).**

The percentage contributions for this paper are Ekta Sharma 70%, Ismail Shakeel 20%, and Ravinesh C. Deo 10%.

Author	Task Performed
Ekta Sharma (*Candidate and APR Intern*)	Literature review and designing *AI* and *DL* algorithms for radio communications, exploring latest technologies, improving the performance of models, attended workshops and summer school, writing, and revising of the manuscript.
Ismail Shakeel (*Industry Supervisor, DSTG Chief Defence Scientist Fellow*)	Supervised and assisted in model concepts, provided timely suggestions, weekly meetings, edited and prepared for the submission. Guidance for the choice of Journal and co-authorship of the manuscript.
Ravinesh C. Deo (*USQ Academic Mentor*)	Advice on modelling approach, weekly meetings, Guidance on choice of Journal, editing, proofreading, and co-authorship of the manuscript

LIST OF PUBLICATIONS

Refereed Journal Articles

1. **Sharma, E**, Deo, RC, Prasad, R & Parisi, AV 2020, 'A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms', *Science of The Total Environment,* Vol. 709, p. 135934-57 **(Scopus Ranked *Q1, Impact Factor: 7.96 and SNIP*: 2.015; *96th percentile in Environmental Science*).**

2. **Sharma, E**, Deo, RC, Prasad, R, Parisi, AV & Raj, N 2020, 'Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Networks', *IEEE Access,* Vol. 8, pp. 209503-16. **(Scopus Ranked *Q1, Impact Factor: 3.367, and* SNIP: *1.421, 87th percentile in Engineering*).**

3. **Sharma, E,** Deo, RC, Soar J, Prasad, R, Parisi, AV & Raj, N 2022, 'Novel hybrid deep learning model for satellite-based PM10 forecasting in most polluted Australian hotspots', *Atmospheric Environment,* Vol. 279, pp. 119111-24. **(Scopus Ranked *Q1, Impact Factor: 4.79 and SNIP*: 1.4; *95th percentile in Environmental Science*).**

4. **Sharma, E,** Deo, RC, Soar, J, & Yaseen, ZM 2022, 'Association of Air Pollutants and Novel Coronavirus with Deep Learning: A systematic literature review of the case studies in Asia and Oceania', *Sustainable Cities and Society,* Under review **(Scopus Ranked *Q1, Impact Factor: 7.59 and SNIP*: 2.35; *97th percentile in Civil and Structural Engineering*).**

5. **Sharma, E**, Shakeel, I, Sancho, SS, & Deo, RC 2022, 'Artificial Intelligence-based Bit-wise Decoding of Error-correction Codes for Radio Communications', *IEEE Access*, Under review. **(Scopus Ranked *Q1, Impact Factor: 3.367* and *SNIP*: *1.421; 87th percentile in Engineering*).**

6. Galligan, L, King, R, Langlands, T, **Sharma, E** & Pickstone, L 2022, 'Connecting community online and through partnership: a reflective piece', *The International Journal for Students as Partners,* Accepted subject to

minor revision. ***(Open Access, Peer-reviewed journal, McMaster University, Canada, New Journal - under indexing).***

7. Ghimire, S, Deo, RC, Wang, H, Perez, DC, Sanz, SS, Ali, M, & **Sharma, E** 2022, 'A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction', *Expert Systems with Applications,* Under-review. **(Scopus Ranked *Q1, Impact Factor: 6.95 and SNIP*: 3.07; *94th percentile in Artificial Intelligence*).**

8. Ward, A, Soar, J, Singh, K, Taafe, K, **Sharma, E**, Alam, K, Kolbe-Alexander, TL, & Biddle, S 2022, *Journal of Sport & Health Science,* To be submitted. **(Scopus Ranked *Q1, Impact Factor: 7.17 and SNIP*: 2.33; *88th percentile in Sports Science*).**

9. Gharineiat, Z, **Sharma, E**, Musaylh, MA, Kumar, D, Samui, P & Deo, RC 2022, 'Near real-time hourly tidal height forecasting in Australia's wave energy belt zone', *Renewable & Sustainable Energy Reviews,* To be submitted. **(Scopus Ranked *Q1, Impact Factor: 14.98 and SNIP*: 4.68; 97*th* *percentile in Energy*).**

10. Jui, SJJ, Ahmed, AAM, Bose, A, Raj, N, **Sharma, E** & Chowdhury, MWI 2022, 'Spatio-temporal Hybrid Random Forest Model for Tea Yield Prediction using Satellite Derived Variables', *Remote Sensing,* Vol. 14(3), pp. 805. **(Scopus Ranked *Q1, Impact Factor: 5.33 and SNIP*: 1.71; *90th percentile in Earth and Planetary Sciences*).**

11. Ahmed, AAM, **Sharma, E,** Ahmed, MH, Ahmed, O, Sutradhar, A & Farzana, SZ 2021, 'Remote sensing for flood index framework: real-time forecasting with hybrid machine', *Remote Sensing of Environment,* Under-review. **(Scopus Ranked *Q1, Impact Factor: 10.16 and SNIP*: 3.35; *99th percentile in Earth and Planetary Sciences*).**

12. Ahmed, AAM, **Sharma E,** Deo, RC, Nguyen-Huy, Thong, Ali, Mumtaz & Jui, JJ 2021, 'Kernel Ridge Regression hybrid method for wheat yield prediction using satellite-derived predictors', *Remote Sensing,* Under-review. **(Scopus Ranked *Q1, Impact Factor: 5.33 and SNIP*: 1.71; *90th percentile in Earth and Planetary Sciences*).**

ACKNOWLEDGMENTS

This research has been supported by an Australian Government Research Training Program Scholarship and is a joint effort of an entire team of supportive educators and wonderful individuals that have set an outstanding benchmark. The learning curve of my doctoral journey was made easier with their timely responses, excellent research suggestions, and continued support to enhance the quality of each paper/project accomplished so far. My sincere gratitude and appreciation to:

- Principal Supervisor, **Prof. Ravinesh C. Deo** for being instrumental in offering me the doctoral opportunity, significant leadership contributions, and insightful research journey. Prof. Deo mentored me to grow, spend my bandwidth, and energy on what is important, research sincerely, and focus on the bigger picture when things got harder. A sincere thanks for being an unparallel educator, important research discussions, providing the best advice, having faith in me, and unending support in completing the special milestone of my life. It is, and always be an honour to have been mentored by you and be a part of the advanced data analytics lab.

- Associate Supervisor, **Prof. Jeffrey Soar** for being generous in sharing his expertise, providing industry experience, and timely valuable suggestions. Thank you for your generous support, constant motivation, and valuable feedback in direction of future research.

- Associate Supervisor, **Honorary Professor Alfio V. Parisi** for his mentorship, proofreading, research contribution, and comments on the research and journal articles.

- Associate Supervisor, **Dr. Nawin Raj** for his timely and useful suggestions in building up my research foundation, proofreading, and feedback. My sincere gratitude for your assistance with deep learning expertise, which played a valuable role in improving the performances of the algorithms involved throughout the PhD journey.

- External Supervisor, **Dr. Ramendra Prasad** for beneficial scientific advice, expertise in statistical analysis, and interpretation of results. I am very grateful

for your timely feedback and research directions, which helped me immensely in improving the quality of my work throughout the PhD journey.

- Adjunct Professor **Ismail Shakeel** from Defence Science and Technology group, for a great internship opportunity, over 5 months of valuable industry experience, and extending support when needed.
- **Australian Postgraduate Research group** (**APR**), for a successful collaboration with the Australian Defence Department: Defence, Science & Technology Group.
- **Australian Mathematical Sciences Institute (AMSI)** for Optimise funding initiative under the **CHOOSE-MATHS** grant, a conference talk, an interview, a feature AMSI annual research report, and a poster presentation.
- **University of Southern Queensland** for the Research Training Program (RTP) Scholarship for supporting my PhD project, **Graduate Research School, School of Mathematics, Physics and Computing** for guidance and extending support when needed, **Advanced Data Analytics: Environmental Modelling and Simulation Group** led by the Principal Supervisor for immense contributions and my colleagues in brainstorming research ideas.
- **Dr. Aruna Devi** for being instrumental in helping me to connect with my principal supervisor and maybe without her direction, my journey would not have started. A heartfelt thanks and gratitude.
- On a personal note, my deepest gratitude to my husband **Anshul,** who redefined the definition of care and support throughout my doctorate journey, and parallelly worked equally hard to support my dream of expanding research horizons. I truly share this milestone with him. My son **Viraj** and daughter **Eva** for providing understanding, support, and unconditional love. My extended family- my father, my in-laws, and my sister selflessly encouraged me to explore new directions in life and grow. I am indebted and genuinely grateful to all these warm individuals. Thank you.

Finally, to everyone I have connected with, whether it was work, study, or research discussions over my involvement with the **Advanced Data Analytics lab** - your support is overwhelming, and I look forward to staying connected.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 4

A Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Networks Algorithms

Chapter 5

Novel hybrid deep learning model for satellite-based PM10 forecasting in most polluted Australian hotspots

(D) Flowchart detailing the methodology in the model designing phase, as adopted for the construction of the proposed deep learning hybrid *CNN-GRU* model forecast PM_{10} (μg/m^3) at an hourly time horizon.

Figure 2 (A) Cross-correlations (r_{cross}) showing co-variance between the air quality *vs.* its predictor variables (acronyms described in Table 4) used for designing the objective air forecasting tool with a hybrid *CNN-GRU* model for Queensland state.

(B) Partial autocorrelation of PM_{10} (μg/m^3) showing the most significant lag for (a) Queensland, (b) New South Wales, (c) South Australia, and (d) Victoria.

Figure 3 (A) Hotspots chosen in the research to forecast hourly PM_{10} (μg/m^3).

(B) Testing performance of *CNN-GRU vs.* the six others competing models evaluated using the Willmott index of agreement (*WI*).

(C) A *3D*-Bar graph of root mean square error (*RMSE*, μg/m^3).

Figure 4 (A) A combo-clustered column and a line chart showing the Kling-Gupta efficiency (*KGE*; R_{KG}) and mean bias error (*MBE*; %), comparison of the states of *Qld*, *NSW*, *SA*, and *VIC*.

(B) Training and Validation loss comparison of *CNN-GRU* for Queensland site in the testing phase.

LIST OF TABLES

forecasted data. Measurement units are *($\mu g/m^3$)* for $PM_{2.5}$, PM_{10}, & *(Mm^{-1})* for Visibility which is the same for *RMSE*. ✖ = no model developed. Activation functions-*hardlim* (hard-limit), *tribas* (triangular basis), *rbf* (radial basis), *logsig* (log-sigmoid), sigmoid), *tansig* (hyperbolic tangent sigmoid), *sig* (sigmoidal) and *sin* (sine). The optimal performance is boldfaced, presented in **red.**

Table 5 Testing performance of the hybrid *ICEEMDAN-OS-ELM* predictive model *vs.* the hybrid versions of *ICEEMDAN-MLR*, *ICEEMDAN-M5* model tree, and the respective standalone models (*M5* model tree, *MLR*, *OS-ELM*) designed to forecast (a) Particulate matter measuring 2.5 μm or less in diameter ($PM_{2.5}$), (b) Particulate matter measuring between 2.5 μm – 10 μm in diameter (PM_{10}), and (c) Visibility-reducing particles (*VIS*). Note: E_{LM} = Legates and Mc-Cabes Index, *WI* = Willmott's Index, E_{NS} = Nash–Sutcliffe Efficiency, *RMSE* = root mean square error, *MAE* = mean absolute error and *r* = correlation coefficient. The most accurate model is boldfaced, presented in **red.**

Chapter 4

A Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Networks Algorithms

Table 1 (a) Geographic description (b) Data segregation of study sites.

Table 2 Descriptive statistics of *TSP* (μg/m^3) for each study site.

Table 3 Model input variables. ✖ = Data is not monitored.

Table 4 Optimal architecture (in **red** for *CLSTM*) for *TSP* (μg/m^3). The Architecture of the Backpropagation Algorithm.

Table 5 Testing performance of *CLSTM vs.* other competing models. (a) Brisbane, (b) Townsville, (c) Hopeland, (d) Miles airport.

Chapter 5

Novel hybrid deep learning model for satellite based PM10 forecasting in most polluted Australian hotspots

Table 1	(a) The geographic description of the study sites, (b) meteorological descriptive statistics of PM_{10} (μg/m^3) and missing data details for all study sites, and (c) data segregation for each station for the study.
Table 2	Potential model input variables used in the study. After data filtering for missing values, selected variables are shown in red colour.
Table 3	(a) The most optimal model architecture, and training performance of the proposed deep learning hybrid *CNN-GRU* for PM_{10} (μg/m^3). The optimal performance is boldfaced, presented in **red.** (b) The architecture of the backpropagation algorithm.
Table 4	Testing performance of the deep learning hybrid *CNN-GRU vs.* the hybrid versions of *CNN-LSTM*, *CNN-BiLSTM* and the respective standalone models *GRU*, *CNN*, *LSTM*, and *BiLSTM* models designed to forecast PM_{10} for (a) Qld, (b) NSW, (c) SA, and (d) VIC. The most accurate model is boldfaced, presented in **red.**

LIST OF ACRONYMS

AQ	Air Quality
APF	Air Pollutant Forecasting
AI	Artificial Intelligence
DL	Deep Learning
D	Original hourly air quality data
$\mu g/m^3$	Micrograms per cubic metre
μm	Micrometres
PM	Particulate Matter
$PM_{2.5}$	Fine particles with a diameter of 2.5 μm or less
PM_{10}	Coarse particles with a diameter between 2.5 and 10 μm
PM_{10}^{MIN}	The minimum value of PM_{10} (μg/m^3)
PM_{10}^{MAX}	The maximum value of PM_{10} (μg/m^3)
$PM_{2.5}^{MIN}$	The minimum value of $PM_{2.5}$ (μg/m^3)
$PM_{2.5}^{MAX}$	The maximum value of $PM_{2.5}$ (μg/m^3)
$PM_{2.5}^{NORM}$	The normalised value of $PM_{2.5}$ (μg/m^3)
$\overline{PM_{2.5,i}^{FOR}}$	Mean of forecasted PM values
$\overline{PM_{2.5,i}^{OBS}}$	Mean of observed PM values

$PM_{2.5,i}^{FOR}$	Forecasted particulate matter for i^{th} observation less than 2.5 micrometre in diameter
$PM_{2.5,i}^{OBS}$	Observed particulate matter for i^{th} observation less than 2.5 micrometre in diameter
VIS	Visibility-reducing particles
TSP	Total Suspended Particulate Matter up to about 100 µm in Diameter.
TSP_i^{FOR}	Forecasted *TSP* for i^{th} observation.
TSP_i^{OBS}	Observed *TSP* for i^{th} observation.
NEPH	Nephelometer
DERM	Department of Environment and Resource Management
BAM	Beta Attenuation Monitor
DoE	Department of Environment and Science
TEOM	Tapered element oscillating micro-balance
D-TEOM	Dichotomous tapered element oscillating balance
r	Correlation Coefficient
r^2	Coefficient of Determination
SD	Standard Deviation
MSE	Mean Square Error
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error, %

WI	Willmott's Index of Agreement
E_{NS}	Nash–Sutcliffe Coefficient
E_{LM} *or* LM	Legates and McCabe Index
ANN	Artificial Neural Network
ACF	Auto-Correlation Function
$\|FE\|$	Forecasted Error
$\beta_n(.)$	n^{th} mode of the signal $x(i)$
$\bar{m}_{2.5}(i)$	local mean of $x(i)$ for $PM_{2.5}$
μ^t	Constant to control Gaussian white noise amplitude
F	Feature map
$*$	An operator of the convolutional process
F_n	Forget gate
I_n	New input
b_m	Bias vector
i_n	Input gate
δ_n	Candidate cell state
W_m	Weight matrices
W^f	Kernel's weight with feature map
IMFs	Intrinsic Mode Functions
EMD	Empirical Mode Decomposition
EEMD	Ensemble EMD

Adam	Adaptive Moment Estimation
a	Activation Function
Tansig	Tangent Sigmoid
Tribas	Triangular basis
Logsig	Log Sigmoid
Purelin	Positive Linear
sig	Sigmoidal
rbf	Radial Basis Function
ReLU	Rectified Linear Unit
DTR	Diurnal Temperature Range
MLR	Multiple Linear Regression
BMA	Bayesian Model Averaging
MRA	Multiresolution Analysis
ELM	Extreme Learning Machine
VOL	Volterra model
OS-ELM	Online Sequential Extreme Learning Machine
GRU	Gated Recurrent Unit
RF	Random Forest
COPD	Chronic Obstructive Lung Disease
CNN	Convolutional Neural Network
CON	Convolutional layer

POOL	Pooling layer
FC	Fully connected layer
LSTM	Long Short-Term Memory
CLSTM	Convolutional Long Short-term Memory Neural Network
BiLSTM	Bidirectional *LSTM*
EEMD	Ensemble EMD
VMD	Variational Mode Decomposition
CEEMDAN	Complete EEMD with adaptive noise
ICEEMDAN	Improved Version of Empirical Mode Decomposition with Adaptive Noise
FDMS	filter dynamics measurement system
Ppm	parts per million
$FTSP^{RF}$	Forecasted total suspended particles for random forest
$FTSP^{VOL}$	Forecasted *TSP* for Volterra
$FTSP^{MLR}$	Forecasted *TSP* for MLR
$FTSP^{M5}$	Forecasted *TSP* for M5 model tree
$FTSP^{LSTM}$	Forecasted *TSP* for *LSTM*
SO_2	Sulfur Dioxide
O_3	Ozone
CO	Carbon Monoxide
NO_2	Nitrogen Dioxide

Qld	Queensland
NSW	New South Wales
SA	South Australia
VIC	Victoria
WHO	World Health Organisation
NEPM	National Environment Protection Measure
CHAG	Clean & Healthy Air for Gladstone
EPA-V	Environment Protection Authority Victoria
EPA-SA	Environment Protection Authority South Australia
Qld-DoES	Queensland Government Department of Environment and Science
NSW-DoPIE	New South Wales Government Department of Planning, Industry, and Environment
MODIS	Moderate Resolution Imaging Spectroradiometer
GIOVANNI	Goddard online interactive visualization and analysis infrastructure
ECMWF	European Centre for Medium-Range Weather Forecasts
NASA	National Aeronautics and Space Administration

CHAPTER 1 INTRODUCTION

1.1 Background

Air quality (*AQ*) affects our health, the liveability of our cities, towns, and our natural environment. Air quality is one of the major issues within an urban area that affects people's lives while existing observations of *AQ* are not adequate to provide comprehensive information to help vulnerable populations to a better plan. Australia is a developed nation and compared with other nations around the world, it has a relatively higher air quality standard. Despite this, challenges to maintaining clean air still exist that include population growth, increased urbanisation, surging demand for transportation (especially air transportation), and energy consumption. Considering some recent statistics, from October 2019 to February 2020, as Australia grappled with mega-fires with unprecedented intensity, more than 3 billion animals and around 437,000 people were exposed to air with a PM 2.5 concentration of least 25 micrograms per cubic meter of air, which is substantially more than the 15 micrograms per cubic meter of air that the World Health Organization considers an acceptable level for short-term exposure. At times, PM2.5 concentrations increased by more than 3.5-fold because of the fires, as indicated by the authors (Graham et al. 2021) estimate. Every year in New South Wales, at least 279 people die prematurely because of toxic air pollution from the state’s five coal-fired power stations. The health impacts of poor air quality also include 233 babies born with reduced birthweight, 361 people developing type 2 diabetes and 2,614 years of life lost each year due to uncontrolled air pollution from the power stations.

A significant challenge is the insufficient spatial measurement of *AQ* parameters. The ground-based weather stations continuously collect data but are only available in a limited number of locations. On the other hand, the remotely sensed satellite-based data are more suited for spatial studies however, satellites may pass a given location at the same time each day and miss out on the emissions variations at different hours over the whole day. This, combined with the arid nature of Australia with frequent wind-blown dust storms and bushfires, can exacerbate the particulate air pollution both temporally and spatially. Therefore, a robust forecasting method, used to constantly monitor the quality of air, which can consider these factors, is needed to forecast the *AQ* more accurately and efficiently whilst supporting the Federal and Local Government's policy development.

This doctoral thesis aims to investigate the feasibility of AI models that can predict 'particulate matter (*PM* or particle pollution)', as a critical indicator for measuring and controlling the degree of air pollution (measured in μg/m^3) and referring to extremely small solid particles and liquid droplets found in the air (NSW 2021a).

Particulate matter, in the context of this doctoral thesis, includes the following:

- **Fine particles or $PM_{2.5}$:** Extremely fine inhalable particles, with diameters that are generally 2.5 micrometres and smaller (Squizzato et al. 2017). The main sources of $PM_{2.5}$ are road dust, emissions from combustion of gasoline, oil, diesel fuel, wood products or coal- or natural gas-fired power plants, and fireplaces to name a few (Broome et al. 2020). This has been the subject of several research studies that have indicated that long-term *PM* exposure has adverse effects on mortality from cardiovascular and respiratory diseases (O'Neill, Zanobetti & Schwartz 2003; Bateson & Schwartz 2004; Glinianaia et al. 2004; Dominici et al. 2006; Zhou et al. 2010; Leitte et al. 2011; Langrish et al. 2012; Zhou et al. 2014). Owing to its severity and to address rising health issues arising from it, precise and real-time prediction of $PM_{2.5}$ is of vital importance.

- **Coarse particles or PM_{10}:** Inhalable but coarser particles, with diameters that are generally between 2.5 μm and 10 μm (Andrews 2008). The main sources of PM_{10} are dust from construction sites, landfills and agriculture, wildfires and brush/waste burning, industrial sources, wind-blown dust from open lands, pollen, and fragments of bacteria (Chan et al. 1999). The severity of ill health impacts arising from high exposure to PM_{10} range from asthma to bronchitis to strokes and even premature death (Krzyzanowski et al. 2005; Walsh 2014; Kim, Kabir & Kabir 2015).

- **Total Suspended particles or *TSP* or *SPM:*** Airborne particles up to about *100 μm* in diameter are referred to as *TSP* (total suspended particles or suspended particulate matter) (Qld 2021). The main source of these particles is combustion and non-combustion processes, including windblown dust, sea salt, earthworks, mining activities, industrial processes, motor vehicle engines, and fires (Chaulya 2004). This pertinent air pollutant is mainly responsible for the increased nuisance, aspect of soiling of property and materials.

- **Visibility-reducing particles (aerosols)** are also monitored to assess loss of visual amenity (the distance one can see clearly) (Qld 2021). The agricultural, industrial activity remnants, and wildfires are some of its common sources with safety implications in areas such as the aviation industry (Hyslop 2009).

Fundamentally, *PM* comes under one of the most challenging optimisation problems in present times, both for air quality and for climate change policies (Valavanidis, Fiotakis & Vlachogianni 2008). Forecasting air quality and designing a practical air quality system is a very difficult problem. This is due to the chaotic behaviour of air pollutants (Lee & Lin 2008). It also creates further difficulties in tracking their three-dimensional movement over several temporal domains. Their high importance for environmental policy is also one of the reasons for the growing scientific interest in published literature. When we forecast spatially correlated time series data, this becomes more challenging due to linear and non-linear dependencies in the temporal and spatial dimensions (Lin et al. 2018). To alleviate the urgency of an early warning system, accurate and robust air quality models for forecasting are needed which will assist in decision making and policy formulation. (Shaban, Kadri & Rezk 2016).

Recent decades have seen the forecasting of air pollutants undertaken largely by three main categories of models: Physically driven (Kukkonen et al. 2012), data-driven or statistical or machine learning model (Karatzas & Kaltsatos 2007), and deep learning models in air quality (Sharma et al. 2020). Physically driven models, also addressed by dynamic models offer good accuracy and generally focus on the scientific evaluations and are considered the gold standard in the industry such as the Community Multiscale Air Quality (*CMAQ*) model that provides good accuracy (Byun & Schere 2006). An important issue leading to their fading importance in forecasting air quality is higher run-times due to the non-stationarity of predictor variables and sometimes overfitting (Pan et al. 2017). Such limitations can lead to substandard results, providing another pathway for an improved version of data-driven or machine learning models (Ghimire et al. 2019b, 2019a). These statistical models have limited forecasting accuracy due to their inability to handle multivariate and nonlinear time series datasets (Tao et al. 2019). Other factors affecting air pollution forecasting are various meteorological factors, traffic, *etc.* When considering these, sometimes statistical and machine learning models fail to consider parameter correlations and substantial multivariate air quality datasets (Zhu, Liang & Chen 2021). In recent years, with the rapid growth of artificial intelligence and big data techniques, deep learning (*DL*) has attracted considerable attention. Becoming an active research field for air quality forecasting, several studies show that the following characteristics help deep learning in attaining a forecasting and competitive edge in comparison to the earlier counterparts (Zhang et al. 2015):

- Greater precision (Zhang et al. 2015)
- Superior capability to address complex data problems and relatively complex function approximations (Ahmed et al. 2021).
- Nonlinear mapping capability (Saxe, McClelland & Ganguli 2013).
- Analysis of convoluted (*i.e.* stochastic) dataset (Reznikov et al. 2020).
- Ability to use in multiple application areas such as pain intensity estimation (Bargshady et al. 2020), solar radiation forecasting (Ghimire et al. 2019b, 2019a), and seizure diagnosis (Al-Hadeethi et al. 2020) to name a few.

Using deep learning methods, this PhD thesis aims to explore the feasibility of existing machine learning methods to create improved hybrid machine learning models that offer an efficient performance of air quality forecasting. The work aims to also recommend new and advanced models that may, therefore, be of potential use in the health environment and useful to governments globally in the planning and policy-making aspects related to air quality. The following subsections of this chapter present the air pollution problem in more depth, providing the research motivation and objectives, the expected contribution, and finally, the structure of the rest of the thesis.

1.2 Research Problem

Air quality monitoring and modeling have become equally important to ensure that the public gets real-time health alerts and advice based on their areas' pollution levels. This need has advanced air quality forecasting as the major research area in environmental data engineering. Timely forecasting is also beneficial for supporting environmental management decisions as well as averting serious accidents caused by air pollution. Furthermore, several time horizons in the future can also have an immense impact on air pollution management, consequently improving social, economic growth, and public health.

Australia is a developed nation with generally good air quality measurements; however, the current air pollution regulatory standards are not strong enough to protect human health. National air pollution limits currently exceed the World Health Organisation's recommended thresholds and by international comparison, lag significantly. Much stricter standards have been adopted in many other countries, including the United States of America, China, and the European Union (EJA 2021). Each year, more than 5000 Australians die from exposure to air pollution – that's four times the national road toll and could be prevented (AAS 2021). Most of the air pollution in Australia comes from aging coal-fired power stations and motor vehicles. Unlike most other countries, Australia's laws, regulations, and standards for these two sources of toxic pollution are incredibly weak (Dean & Green 2017).

The detrimental health risks of unprecedented Australian bushfire smoke levels have further raised alarming concerns regarding air quality monitoring. The 'Black summer fires' of 2019-20 resulted in the worst-ever air emission footprints to date with critical environmental pollutants, particularly particulate matter (PM) reaching nearly $400\mu g/m^3$; described as hazardous by the World Health Organisation (Davey & Sarre 2020). A massive 80% of Australians were impacted, 27.2 million acres burnt, with over $10 billion in national financial impacts (Ryan, Silver & Schofield 2021). There seem to be discrepancies in the presentation of existing air quality metrics used by states, territories, and the Australian government departments such as the composite Air Quality Index, stratifying health messages into colour-coded bands (very good, good, fair, poor). This air quality information and related health advice across jurisdictions are confusing for the public. Further government investment is needed in air quality forecasting models, public health messaging, and exposure reduction for futuristic measures to protect Australians from the hazards of bushfires and dust storms.

Analysing Australia National Pollutant Inventory (*NPI*) data to estimate air pollutants, coal mines accounted for 42.1% of national PM_{10} air emissions from *NPI* sites. $PM_{2.5}$ from coal mines accounted for 19.5% of the national total, metals for 12.1%, and nitrogen oxides for 10.1% (Hendryx et al. 2020). Air pollution in the Upper Hunter, one of the most polluted regions of Australia was reviewed (Higginbotham et al. 2010) to highlight the inaction of Australian States in addressing residents' health concerns. However, (Broome et al. 2015) studied ambient $PM_{2.5}$ and ozone in metropolitan Sydney and the public health impacts. Mount Isa, Queensland, Port Pirie Smelter in South Australia, La Trobe Valley in Victoria are some of the top particulate producing hotspots in Australia (Price et al. 2012).

Motivated by the critical reasons mentioned above, this doctoral thesis focuses on forecasting the air quality, especially all forms of particulates at short and medium temporal horizons for Australia using the machine and deep learning approach. The thesis also assesses the impact of several potential satellite-extracted and ground-based meteorological variables on forecasting air quality. The study regions are the major Australian hotspots where particulate matter concentration is high such as parts of Queensland, New South Wales, Victoria, and South Australia.

1.3 Research Aims and Objectives

Within the scope of the research gaps already identified in this doctoral thesis, the primary purpose of this research is to develop a robust and computationally efficient hybrid model based on the artificial intelligence applied to the problem of air quality forecasting. The thesis reports on the research which examined a set of case studies in Australia across short and long-term horizons. This is presented as a collection of three high-impact, Scopus *Q*1 publications as a core contribution of this PhD thesis by Publications.

To achieve the aim of this study, the following objectives are presented:

i. Develop and evaluate a novel hybrid air quality early-warning framework *ICEEMDAN-OS-ELM* (**I**mproved **C**omplete **E**nsemble **E**mpirical **M**ode **D**ecomposition with **A**daptive **N**oise (*ICEEMDAN*) with the **O**nline **S**equential-**E**xtreme **L**earning **M**achine (*OS-ELM*)) that emulated hourly particulates (*PM*2.5 and *PM*10) with atmospheric visibility reducing particles

(*VIS*). It was benchmarked with *ICEEMDAN*-multiple-linear regression (*MLR*), *ICEEMDAN-M5* model tree, and standalone: *OS-ELM, MLR, M5* model tree using Queensland and New South Wales data. The article has been published in ***Science of the Total Environment*** (Vol: 709 (2020): Page(s): 135934 – 135957, *ISSN:* 0048-9697).

ii. Design deep learning techniques for hourly suspended particulate matter forecasting with **C**onvolutional **N**eural **N**etwork (*CNN*) and **L**ong **S**hort-**T**erm **M**emory (*LSTM*) networks to create a novel *CNN-LSTM* framework for Australia. The article has been published in ***IEEE ACCESS*** (Vol: 8 (2020): Page(s): 209503 – 209516, *ISSN*: 2169-3536)

iii. Investigate the association of meteorological parameters in forecasting hourly air quality for the most polluted Australian hotspots through hybrid deep learning of **C**onvolutional **N**eural **N**etwork (*CNN*) and **G**ated **R**ecurrent **N**etwork (*GRU*) to create *CNN-GRU* architecture. The article has been published in ***Atmospheric Environment.*** **(**Vol. 279 (2022), Pages(s): 119111-24.**)**

Other than the three primary objectives identified as a core part of the PhD program, the doctoral research has also focussed on a further **two research** objectives covering:

(1) a systematic review paper on air quality and COVID-19, and

(2) an original research paper on Artificial Intelligence-based Bit-wise Decoding of Error-correction Codes for Radio Communications for future defence applications. For this aspect, an APR. Intern PhD internship program (over 5 months) was conducted with external funding from the Australian Department of Defence. A media article by APR.Intern is here.

The details of these additional articles are given in section ***1.5 thesis layout*** and ***Appendix A and B***.

1.4 Significance of the Research study

Using an artificial intelligence approach this thesis reports the potential utility of an air pollution forecasting system that was developed and evaluated at short and medium-scale timesteps with case studies in Australia. The models developed in the study carry significant merits, as summarised below:

i. A significant literature gap has been bridged in this work where novel frameworks have been modelled in all three prime objectives for pollutant modelling, particularly for Australia, and therefore, can be used as a pragmatic tool in monitoring the atmospheric environment.

ii. The atmospheric modelling effort assists in the evaluation of the impact of pollutants on health, and the environment, and contributes to research on air quality and general risk assessments. It can be used as an early warning decision support system for the government of any country.

iii. This research work is a paramount step regarding the construction of an effective air pollutant forecasting technology making a vital impact on the environment and health risk alleviation.

iv. The developed hybrid artificial intelligence and deep learning models are vital for health policymakers and state governments for better future planning and policy development. These modelling strategies can provide timely information for rapid decision-making during extreme events such as bushfires and dust storms.

v. In modelling the air pollution forecasting frameworks, efficient and practical architectures have been designed. This exemplified the proposed methods as a superior tool for forecasting air quality magnitudes both at short and medium scales. This has positive implications for providing advice on air quality and addressing the current challenges to the public health sector.

vi. The limitations of traditional statistical models in modelling the proposed architecture have been identified leading to an understanding of the hybrid frameworks, their advantages, and how they can overcome the challenges experienced in the conventional statistical methods.

The programming algorithms implemented in this research were developed on a Windows 10 platform Intel® core™ *i*7 Generation 9 @ 3.7 gigahertz processing unit, 16 GB memory, *MATLAB,* and Python programming language (Sanner 1999) and freely available open-source libraries (*i.e*., Keras (Ketkar 2017), Tensor Flow (Abadi et al. 2016), and Scikit-learn (Pedregosa et al. 2011)).

1.5 Thesis Layout

This thesis is organised into six chapters, comprising three major studies, presented as a PhD Thesis by Publications. They are followed by the conclusion that summarises the challenges, findings, significance, and scientific contributions of this study and recommendations for future works. Furthermore, the details of the one article under review and one submitted are given in *Appendix A and B.*

The thesis schematic and the six chapters are as follows:

Chapter 1 describes the introductory background and the statement of the problem of the research and presents the objectives of this study.

Chapter 2 presents the area of study, data, and general methodology used in this study and lays the background for the forthcoming chapters. This chapter provides general viewpoints while the specific study area, data, and methods are presented in the respective chapters.

Chapter 3 presents a published journal article in ***Science of the Total Environment*** (Vol: 709 (2020): Page(s): 135934 – 135957, *ISSN:* 0048-9697). This chapter formulates a novel hybrid air quality early-warning framework for Australia, that was modelled at an hourly temporal horizon with online sequential extreme learning machines and improved complete ensemble empirical mode decomposition with adaptive noise for particulate matter 2.5, 10, and visibility reducing particles.

Publication Type: Journal
Scopus Rated: Q1
Percentile: 96%, Category: Environmental Science
Impact Factor: 7.96
SCImago Journal Rank (*SJR*): 1.795
SNIP: 2.015
***DOI*:** https://doi.org/10.1016/j.scitotenv.2019.135934

Chapter 4 presents a published journal article in ***IEEE ACCESS*** (Vol: 8 (2020): Page(s): 209503 – 209516, *ISSN:* 2169-3536). This chapter discusses deep learning techniques for hourly suspended particulate matter forecasting using convolutional neural and long short-term memory networks to create a novel *CNN-LSTM* framework for Australia.

Publication Type: Journal
Scopus Rated: Q1
Percentile: 87%, Category: Engineering
Impact Factor: 3.367
SCImago Journal Rank (*SJR*): 0.587
SNIP: 1.421
***DOI*:** https://doi.org/10.1109/ACCESS.2020.3039002/

Chapter 5 presents a journal article accepted subject to minor revision in ***Atmospheric Environment.*** This chapter presents the association of meteorological parameters in forecasting hourly air quality for the most polluted Australian hotspots through hybrid deep learning of convolutional neural network and gated recurrent unit to create *CNN-GRU* architecture.

Publication Type: Journal
Scopus Rated: Q1
Percentile: 95%, Category: Pollution
Impact Factor: 4.79
SCImago Journal Rank (*SJR*): 1.4
SNIP: 1.4
***DOI*:** https://doi.org/10.1016/j.atmosenv.2022.119111

Chapter 6 presents the summary of the thesis with concluding remarks, limitations, and recommendations for future works.

Appendix A:

This chapter blends our main thesis chapters by discussing the association of air pollutants and novel Coronavirus with deep learning as a systematic literature review with the case studies in Asia and Oceania. The findings of this work are under review in a high-impact research journal of ***Sustainable Cities and Society.***

Publication Type: Journal

Scopus Rated: Q1

Percentile: 97%, Category: Civil and Structural Engineering.

Impact Factor: 7.59

SCImago Journal Rank (*SJR*): 1.65

SNIP: **2.35**

DOI**:** TBA

Appendix B:

This chapter discusses the work done during the student internship with Defence Science and Technology Group (*DST*) and Australian Postgraduate Research (*APR*). The journal article was submitted to ***IEEE ACCESS.*** Artificial Intelligence-based Bit-wise Decoding of Error-correction Codes for Radio Communications was discussed.

Publication Type: Journal

Scopus Rated: Q1

Percentile: 87%, Category: Engineering

Impact Factor: 3.367

SCImago Journal Rank (*SJR*): 0.587

SNIP: 1.421

DOI**:** TBA

For a better understanding of the connection among the studies and articles, the flow story of the thesis is graphically represented in Fig. 1.1.

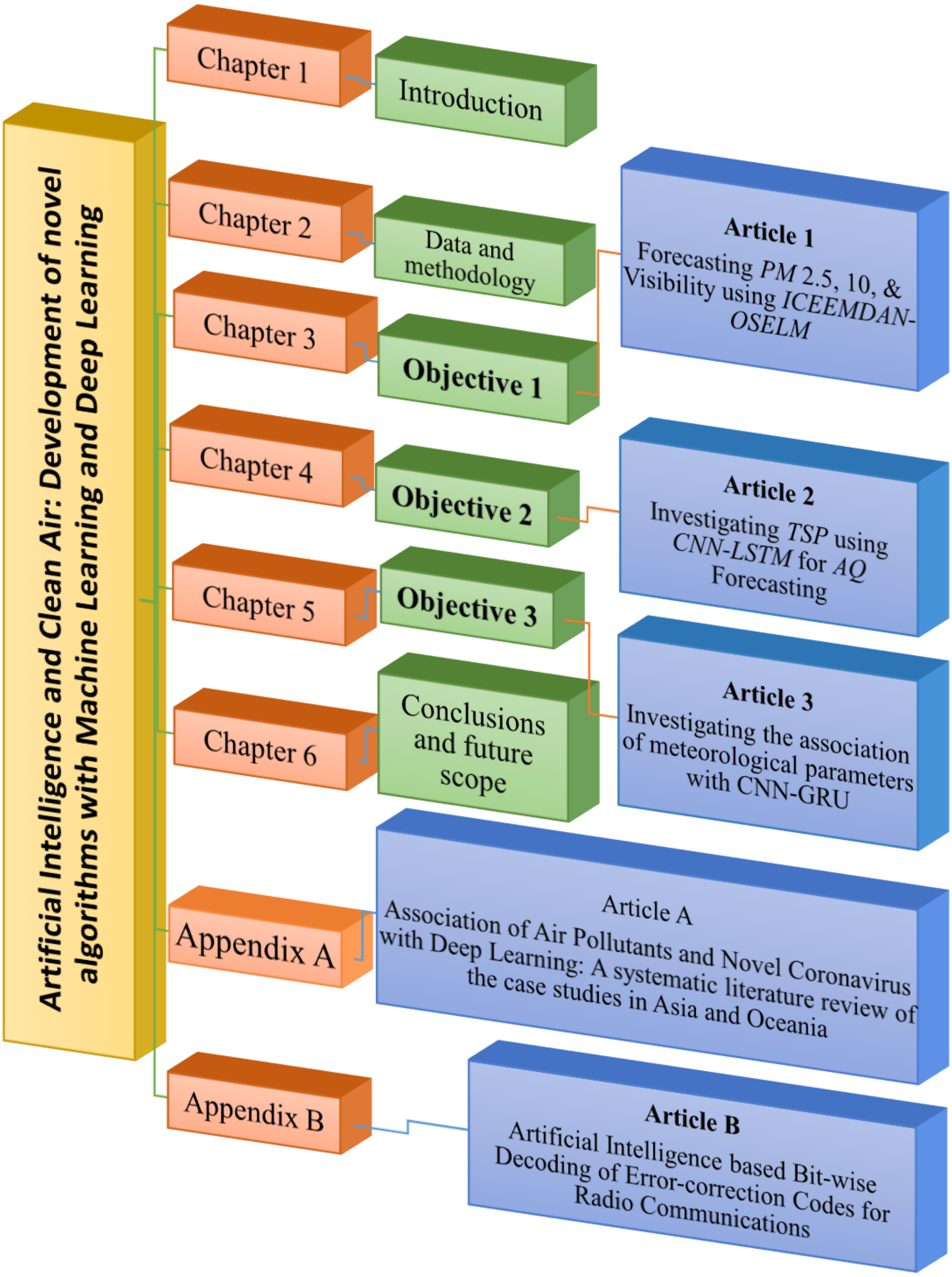


Figure 1.1: Thesis flow chart.

1.6 Chapter Summary

Over time, the patterns of air pollution are changing in Australia, but the presence of air pollution is an increasing challenge for Australian cities and often for rural areas. This is because of changes in the growing population, climate, industry, and technology. The co-occurrence of multiple factors has a significant impact on weather variables and air quality. State and Territory governments in Australia would benefit from a powerful tool to better understand and quantify atmosphere-related risks. Although several physical models have been developed worldwide, they often fail to measure a complex dependence structure. The deep learning models provide a better way to describe the joint behaviour of compound events. The main contribution of this research is to attract considerable attention to a potential and novel integration of deep learning in emulating air quality at several temporal resolutions in regions of Australia where air pollution is becoming a significant problem to public health. The models developed in this work can also be used globally to develop policies and strategies to control air pollution, now and in the future.

CHAPTER 2 DATA AND METHODOLOGY

2.1 Foreword

This chapter provides an overview of all the study locations used in this thesis in developing the early warning hybrid deep learning models for air quality forecasting. These study areas are of paramount importance to Australia where air quality presents a health issue while the proposed models are of global significance. The chosen hotspots have a high concentration of particulate matter, and their importance is described in detail in each of the chapters. The availability, description of data used, duration and limitations if any, are also discussed. Although detailed model development techniques have been presented in the respective chapters, this chapter also presents a brief account of the methodology that was used in developing the deep learning and artificial intelligence models developed for air quality forecasting.

2.2 Study Area

Figure 2.1 shows the study region showing geographical sites considered in Australia for the research.

2.2.1 Queensland (*Qld*)

Queensland located between 22.57° S (latitude) and 144.08° E (longitude) is the second largest state of Australia and one of the primary study areas considered in this thesis. Being a sunshine state, as the temperatures surges, it becomes important to control and reduce the emission of air particles, particularly when the temperature surges (Ren & Tong, 2006). *Qld* records Australia's highest greenhouse gas emissions per capita (CWA 2018) and is also home to nine out of the worst 10 mines for causing air pollution in Australia. These mines are known for generating PM_{10} which is the cause of rising respiratory problems, cancer, and chemicals in the bloodstream (NPE 2016). Queensland stations considered in the study are-

- **Gladstone** (Boyne Island)- A project named Clean & Healthy Air for Gladstone (*CHAG*) was undertaken by *Qld* Health and the Department of Environment and Resource Management to measure the cumulative impacts of hazardous pollutants.

- **Mackay Region** (West Mackay)- Located in the West Mackay, Mackay Region provides important information about air pollutants such as PM_{10} *and VIS*, from the industrial site that faces increasing community concerns from the industrial air emissions on human health (DERM 2011).
- **Brisbane** (Cannon Hill, South-East (*SE*) *Qld*)- Capital of Queensland, this station is situated next to the metropolitan railway and transports coal to the Port of Brisbane. *PM* is measured to evaluate the progress of ongoing measures focussed on reducing coal dust emissions from rail wagons. The station also measures most air pollutants *i.e*., PM_{10}, $PM_{2.5}$, *TSP*, and *VIS*, considered in the study, and therefore, is a critical site to assess the efficiency and feasibility of the proposed *DL* hybrid *AI* models.
- **Townsville** (Townsville Coast Guard)- With the 'Townsville Dust Monitoring Program' *AQ* monitoring started in Townsville in 2007. This site registers community concerns about dust impacts from the Port of Townsville operations.
- **Miles Airport** (Western Downs region)- The station started its monitoring in 2015 and is based in Southwest (*SW) Qld*. This is important to consider evaluating deep learning (*DL*) hybrid models as it measures 12 air pollutants.
- **Hopeland** (Darling Downs region)- Hopeland is a rural station in *SW Qld*. It assesses *AQ* near an area of intensive coal seam gas (*CSG*) production run by Origin Energy. Hopeland station is managed by CSIRO's Gas Industry Social and Environmental Research Alliance (*GISERA*) as part of the Surat Basin Air Quality research.
- **Mount Isa mines**- The Mount Isa Mine has been in operation since the mid-1920s with smelting commencing in the 1930s. Mount Isa has been declared the most polluted postcode in Queensland (ABC 2018) with 12 polluting facilities (NPI 2021). However, it is also being

challenged as the most polluted postcode in Australia (NWSN 2021), with a claim from the National Pollutant Inventory (NPI 2021) that Mount Isa's mining facilities are the highest polluting facility, with almost double the air emissions of the next highest facility in Australia.

2.2.2 New South Wales (*NSW*)

NSW located between 31.25° S (latitude) and 146.92° E (longitude) is a south-eastern Australian state and the most populous state of Australia. *NSW* stations considered in the study are-

- **Sydney** (Campbelltown)- Sydney, the capital of *NSW* is the economic and financial centre of Australia and registers a high concentration of $PM_{2.5}$. It also measures most air pollutants ***i.e***., PM_{10}, $PM_{2.5}$, and *VIS*, considered in the study. Many researchers have taken Sydney as an area of interest (Mannes et al. 2005; Gupta et al. 2007; Morgan et al. 2010) but there is an absence of air quality forecast models based on techniques considered in the thesis.
- **Newcastle** in (Lower Hunter Region)- This station has the world's largest black coal exporting port with three power stations and 34 mines (HVRA 2009) while the Hunter region residents are exposed to industrial air pollution concentrations rivalling any region in Australia (Higginbotham et al. 2010).
- **Hunter Valley –** It has the highest concentration of coal mines and coal train operations in Australia. The national standard of coarse particulates pollution was exceeded 171 times as compared to earlier years in the Hunter region, making it the highest air-polluting region of *NSW (NPI 2021)* and a major contributor to exacerbated heart and lung diseases in the population.

2.2.3 Victoria (*VIC*)

VIC located between 36.98° S (latitude) and 143.39° E (longitude) is a south-eastern Australian state. *VIC* stations considered in the study are-

- **La Trobe Valley-** It is the location of large coal-fired power stations in Victoria, along with the paper mill, and open-cut coal mines. This makes it the third-largest source of particulates along with the emission of 30 toxic substances. The (EPA-VIC 2016) air monitor readings show $PM_{2.5}$ readings of up to 67.4 micrograms per cubic meter in the valley – nearly three times the ambient air goal for $PM_{2.5}$ making it a very important hotspot for the study.

2.2.4 South Australia (*SA*)

SA located between 31.25° S (latitude) and 146.92° E (longitude) is a south-central part of the Australian state. The station that has been considered in the study is-

- **Port Pirie Smelter-** Port Pirie hosts one of the world's biggest lead and sulphur dioxide smelters. This facility is the reason that at least 95% of children in Port Pirie had elevated lead levels in the blood over the last decade. Other pollutants levels that are also substantially high are one-hour sulphur dioxide levels. The standard of which is 0.2 parts per million and > 1000 times since 2003. It has been the topic of research in several studies due to its importance (Baghurst et al. 1999; Taylor & Isley 2014; Taylor, Isley & Glover 2019)

Figure 2.1: The study region shows geographical sites in Australia for the research

2.3 Data Description

Two main data sources were utilised in developing artificial intelligence and hybrid deep learning forecasting models for air quality. Table 2.1 concisely describes the data used with respective sources and other relevant details in achieving each objective.

Table 2.1 Details of all data used in this research.

		Tested Air Quality Parameter	Source	Study period (dd-mm-yyyy)	Proposed model	Other models developed
OBJECTIVE 1	Paper 1 (Chapter 3)	Predictor and Target: *$PM_{2.5}$ ($\mu g/m^3$), PM_{10}* (*$\mu g/m^3$*), *VIS (Mm^{-1})*	*NSW DoPIE (OEH)* and the *Qld DoES*	01-01- 2015 to 31-12-2017	*ICEEMDAN-OS-ELM*	*OS-ELM, ICEEMDAN*-M5 model tree, *M5* model tree, *ICEEMDAN-MLR, MLR*
OBJECTIVE 2	Paper 2 (Chapter 4)	Predictors *: *WD (°TN), WS (°), WST (m/s), WSSD (°C), AT (%), RH (hPa), BP*(μg/m^3), *PM_{10}* (μg/m^3), *$PM_{2.5}$*(μg/m^3), *O_3 (ppm), NO_2 (ppm), CO (ppm), RAD (W/m^2)* Target: *TSP* (μg/m^3)	*Qld* Government Department of Environment and Science	01-01- 2015 to 31-12-2018	*CLSTM*	*LSTM*, Random Forest, M5 model tree, Volterra, *MLR*
OBJECTIVE 3	Paper 3 (Chapter 5)	Predictors †: *DSA (-), TAE (Nm), TAS (Nm), TCZ (Dobsons) CTP (hPa), CTT (K) TCA (-), CAFL (-) CAFM (-), CAFH (-), TPW (kg/m^2), DSA (m/s)* Target: *PM_{10} ($\mu g/m^3$)*	*NASA GIOVANNI*, *Qld DoES*, *EPA VIC* Government of *SA*, and *NSW DoPIE*	01-01- 2018 to 31-12-2020	*CNN-GRU*	*CNN-BiLSTM, CNN-LSTM, GRU, BiLSTM, CNN*

Note: *OEH*- Office of Environment and Heritage; *DoES* - Department of Environment and Science, *EPA*- Environment Protection Authority, *DoPIE*- Department of Planning, Industry, and Environment.
* = *WD*- Wind Direction, *WS*- Wind Speed, *WST*- Wind Sigma Theta, *WSSD*- Wind Speed Standard Deviation, *AT*- Air Temperature, *RH*- Relative Humidity, *BP*- Barometric Pressure, PM_{10}-Coarse particulate matter, $PM_{2.5}$-Fine Particulate matter, O_3 - Ozone, NO_2 – Nitrogen Dioxide, *CO* – Carbon Monoxide, *RAD* - Solar radiation.
† = *DSA*- Dust Scattering AOT, *TAE*- Total Aerosol Extinction, *TAS*- Total Aerosol Scattering, *TCZ*- Total column ozone, *CTP*- Cloud top pressure, *TCA*- Total cloud area fraction, *CAFL*- Cloud area fraction for low clouds, *CAFM*- Cloud area fraction for middle clouds, *CAFH*- Cloud area fraction for high clouds, *TPW*- Total precipitable water vapour, and *SWS*- The surface wind speed.

For objective 1, the aggregated data from the entire state of New South Wales (*NSW*), and Queensland (*Qld*), Australia were used and acquired from the *NSW* Office of Environment and Heritage and the *Qld* Government Department of Environment and Science repositories. The data providers regularly assess the reliability of air quality monitoring stations and adopt preliminary quality checks to ensure an accurate, effective, and cost-efficient mechanism is in place. The data are for the hourly temporal horizon and the period of 01-January-2015 to 31-December-2017. Chapter 3 provides more details regarding data processing and its usage.

For objective 2, the data from the same source as objective 1 were adopted. However, the study focussed on all the hotspots in Australia where suspended particulate matter data were available. These locations were found to be in Queensland (four stations). Hourly data was extracted to forecast *TSP* from 01-01–2015 to 31–12–2018. The data is validated before use for modelling purposes. Chapter 4 provides more details of these data.

For objective 3, the study adopts a wide range of datasets extracted from five sources: 30 meteorological variables or Satellite data are extracted from (NASA-GIOVANNI 2021) for the input variables. *GIOVANNI* is the repository 'Goddard Online Interactive Visualization and Analysis Infra-structure' belonging to the National Aeronautics and Space Administration (*NASA*). This enables feeding the model with critical variables like aerosol optical depth, tropopause pressure, air temperature to name a few that are highly likely to influence the variability of air pollution. The ozone monitoring instrument *(OMI)* spectrometer, the Atmospheric Infrared Sounder (*AIRS)* satellite, and the Modern-Era Retrospective analysis for Research and Applications (*MERRA)* satellite variables were adopted in the study since historical data related to the target variable plays the main role in evaluating air pollutants. The ground-based PM_{10} values were sourced from four government sources for each state.

For Queensland, the data were sourced from the Queensland Government Department of Environment and Science (DoES 2021). For New South Wales (*NSW)* the repository was considered by the New South Wales Government Department of Planning, Industry, and Environment (NSW 2021b). South Australian repository was sourced from The Environment Protection Authority South Australia (EPA-SA 2021), and finally, the data for Victoria was sourced from The Environment Protection Authority, Victoria (EPA-VIC 2021). The data sourced were from 01-Jan-2018 to 31-Dec-2018 for the hourly horizon. A complete list of all variables, including their details is provided in Chapter 5.

2.4 General Methodology

Before the development of hybrid deep learning and artificial intelligence models, it is essential to do a quality check on the data. The missing data imputation was done through a calendar averaging technique during this phase. The data used in the study naturally display stochastic behaviour. In addition, we normalise/scale the input data to avoid dominance and to bring the data to a common scale. This also avoids large numeric ranges from the values of the predictor variables, which in turn may undermine the effects of lower range values. The data are normalized between [0, 1] (Hsu, Chang & Lin 2003). The following formula is used for normalisation which in turn helps in handling large variations in the data (Hsu, Chang & Lin 2003). The formula used in objective 1 is shown below as an example:

$$PM_{2.5}^{NORM} = \frac{PM_{2.5} - PM_{2.5}^{MIN}}{PM_{2.5}^{MAX} - PM_{2.5}^{MIN}} \tag{1}$$

In Eq. 1. $PM_{2.5}^{MIN}$ = The minimum values of $PM_{2.5}$,

$PM_{2.5}^{MAX}$ = The maximum values of $PM_{2.5}$

$PM_{2.5}^{NORM}$ = The normalised $PM_{2.5}$ and similarly for the other variables in subsequent equations used for modelling.

Thereafter to complete the objective of this thesis, the inputs, and the target data were converted into their respective forecast horizons by using the partial autocorrelation function (*PACF*). This step selects most of the best statistically significant lag from the target air quality variable (different *PM* in this study). Furthermore, to address the non-stationarity issues associated with the data, the data were decomposed using an improved version of empirical mode decomposition with adaptive noise (*ICEEMDAN*). For the further objectives, as we had more input variables, we also employed the statistic of the cross-correlation (*CCR*) between the target (*PM*) and the inputs (State governments data or satellite extracted data) to choose the best variables for input for the third objective.

Finally, the grid search was employed to finalise the boundary conditions and best parameters of the models developed in the respective objectives of this study. The optimisation and trial-and-error methods are discussed comprehensively in their respective chapters 3, 4, and 5.

In this thesis, various data intelligent air quality forecasting models and competitive benchmarking approaches were considered and developed, since a robust modelling approach is necessary. The models ranged from the well-known artificial intelligence models such as online sequential extreme learning machine (*OSELM*), Volterra, extreme learning machine (*ELM*), random forest (*RF*), M5 model tree, multiple linear regression (*MLR*), and deep learning models such as *ICEEMDAN*-multiple-linear regression (*MLR*), *ICEEMDAN-M5* model tree, Convolutional long short-term memory (*CLSTM),* long short-term memory *(LSTM),* Bidirectional long short-term memory *(BiLSTM), CNN-BiLSTM, CNN-LSTM,* gated recurrent unit *(GRU),* and *CNN-GRU.*

This evaluated their abilities to forecast air quality data and compare the preciseness, robustness, and limitations of different approaches over the hourly forecast horizon. Pre-processing approaches and data filtering techniques involved *ICEEMDAN*, grid search, and *CNN* that were used to handle the non-stationarity features or extract the spatial features from the inputs and create a hybrid model. These filtering techniques are necessary to enhance the forecast accuracy by decomposing the data into low and high pass filters and selecting the best input parameters for the models used as a hybrid or later in modelling. A trial-and-error method was utilized in this research to select the appropriate parameters of all the models used in the study. Finally, the bootstrap (*B*) technique was incorporated in this research to create an ensemble model that has a good ability to estimate the forecast uncertainty.

In the literature, the *MLR* model is a statistical technique that examines the cause and effect relationship between objective *i.e.* and predictor variables (Deo & Şahin 2017).

Random Forest is an ensemble of decision tree algorithms that can be used for classification and regression predictive modeling (Rajarethinam, Aik & Tian 2020). The main advantage of the M5 model tree is that it provides the results in a simple and comprehensible form of regression equations (Kisi et al. 2017) which can be easily used. In air forecasting and other applications, the *OS-ELM* model is considered a fast and powerful data intelligent approach that can offer better performance compared to other algorithms (Liang et al. 2006; Lan, Soh & Huang 2009). On the other hand, *ICEEMDAN* and B are filtering robust techniques. These pre-processing techniques have been applied in many forecasting studies to improve traditional models (Xiao et al. 2021). This helps in the selection of the best parameters for models and generating prediction bands through an ensemble model (Tiwari & Adamowski 2013; Li & Li 2016; Quilty & Adamowski 2018).

For objective 1 of Chapter 3, *i.e.*, hourly (1.0 hour) forecasting horizon, standalone *OS-ELM*, *MLR*, *M5* model tree was used, and two hybrid approaches were constructed namely *ICEEMDAN-MLR*, *ICEEMDAN-M5* model tree. They were all benchmarked against *ICEEMDAN-OSELM* which was found to outperform the approaches in forecasting $PM_{2.5}$, PM_{10}, and *VIS*. In Chapter 4, which was our objective 2, a hybrid predictive model (*i.e.*, *CLSTM*) was designed to forecast the *TSP*. *CLSTM* was integrated between *CNN* with *LSTM* to attain optimal performance. It was compared against the ensemble of five other competing models namely Random Forest, Volterra, *MLR*, *M5* model tree, and *LSTM* that lagged in their capability to generate satisfactory *TSP* forecasts to the deep learning *CLSTM* hybrid model. In the next objective 3 *i.e.*, Chapter 5, there is an integration of *CNN* and *GRU* model to make the *CNN-GRU* model. The objective model was competitively benchmarked with competitive *CNN*, *LSTM*, *BiLSTM*, *GRU*, and two hybrids *CNN-LSTM* and *CNN-BiLSTM*.

Figure 2.2 illustrates the classification of the models discussed in this section into their respective chapters. Also, the theoretical backgrounds and techniques used to develop the objective models are shown in detail in the related chapter. In summary, the novel and main hybrid models developed in this study were:

i. The novel early warning *AI*-based framework *ICEEMDAN-OSELM* was proposed to forecast all air quality variables (*i.e.*, $PM_{2.5}$, PM_{10} & *VIS*). The *ICEEMDAN* was adopted to decompose the target data into intrinsic mode functions (*IMF*) and a residual component while the *OSELM* robustly extracted predictive patterns by fine-tuning the model generalisation to a near-optimal global solution, representing modelled *AQ* at hourly forecast horizons. Chapter 3 illustrates the complete model development.

ii. A practical and novel approach for Australia was generated through a computationally efficient architecture with a deep learning hybrid *'CLSTM'* model. This addressed gaps in modelling the hourly *TSP* that can help design a user-friendly air pollution forecasting system. A three-layered *CNN* model robustly extracted the data features from 'four' benchmarked models. The last layer analyses all features to forecast the next hour's *TSP* (μg/m^3). The details of the model development steps are clearly shown in Chapter 4.

iii. The *CNN* and *GRU* models are integrated with an extensive list of meteorological parameters as predictor variables to forecast PM_{10} (μg/m^3) for an hourly forecasting horizon. A new and novel approach for air quality forecasting was presented in Chapter 5 where six other *DL* models were developed.

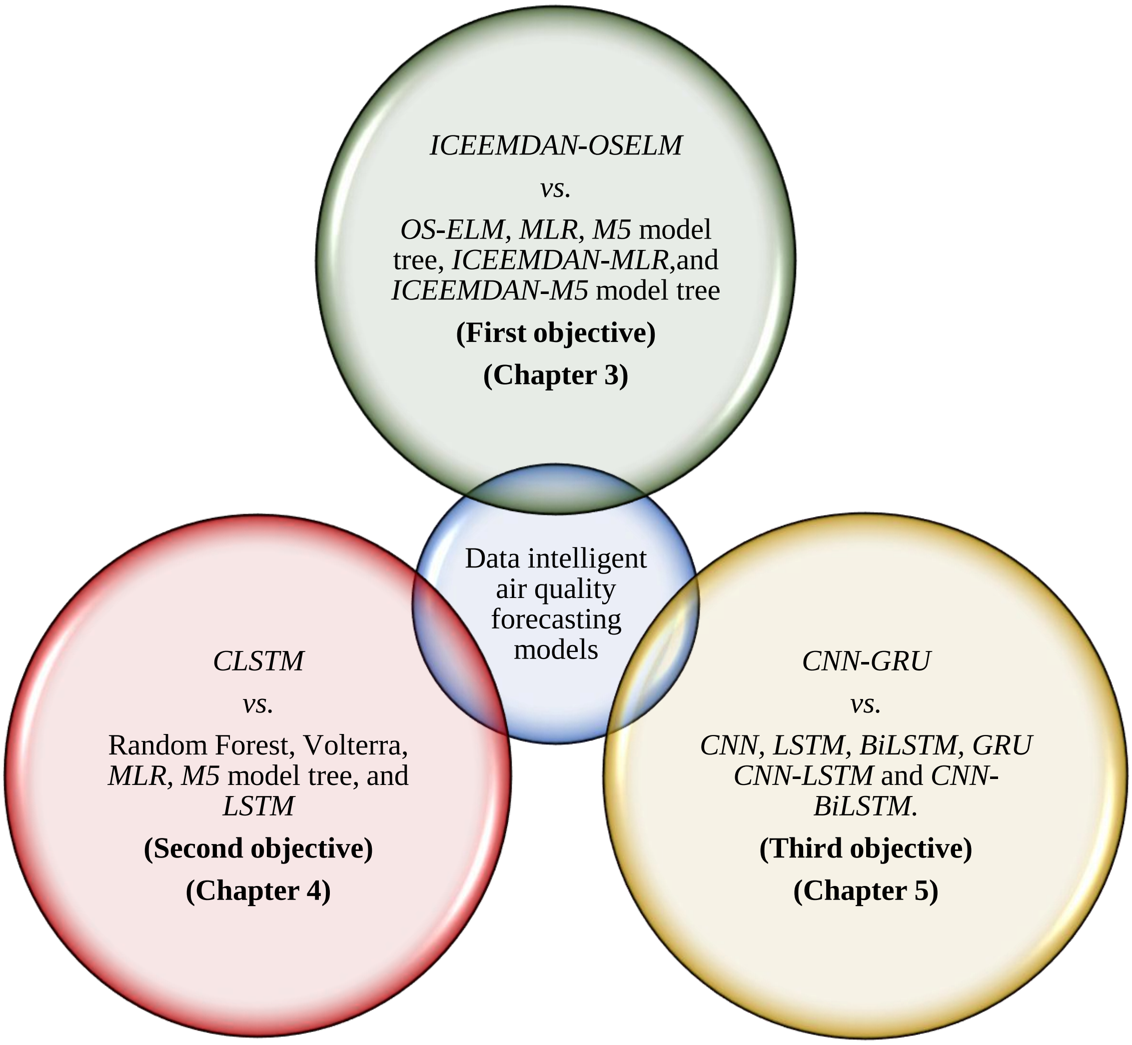


Figure 2.2: Types of data intelligent models developed in each chapter (objective) in this doctoral thesis.

In this doctoral thesis, a wide range of statistical criteria have been used as performance criteria. They include- correlation coefficient (*r*), root-mean-square error (*RMSE; μ*g/m^3), mean absolute error (*MAE; μ*g/m^3), and relative mean square error (*RMSE; μ*g/m^3), mean absolute percentage error (*MAPE*; %), mean bias error (%), Willmott's Index (*WI*), Nash–Sutcliffe efficiency coefficient (E_{NS}) and Legates and McCabe's Index (*L*) were employed to evaluate the performance of the models to forecast air quality data (*PM* in this thesis). The details and mathematical equations for these statistical indices are shown in each chapter of this thesis. Additionally, several plots including a colour-coded heatmap, polar plots, combination plots of column and line charts, box plots, scatter diagrams, time series plots, relative error analysis plot, bar graphs, Taylor plots, and Python training and validation loss comparisons, as well as performance evaluation plots, were also introduced to assess the ability of the models developed in this study for air quality forecasting.

CHAPTER 3 A HYBRID AIR QUALITY EARLY-WARNING FRAMEWORK: AN HOURLY FORECASTING MODEL WITH ONLINE SEQUENTIAL EXTREME LEARNING MACHINES AND EMPIRICAL MODE DECOMPOSITION ALGORITHMS

3.1 Foreword

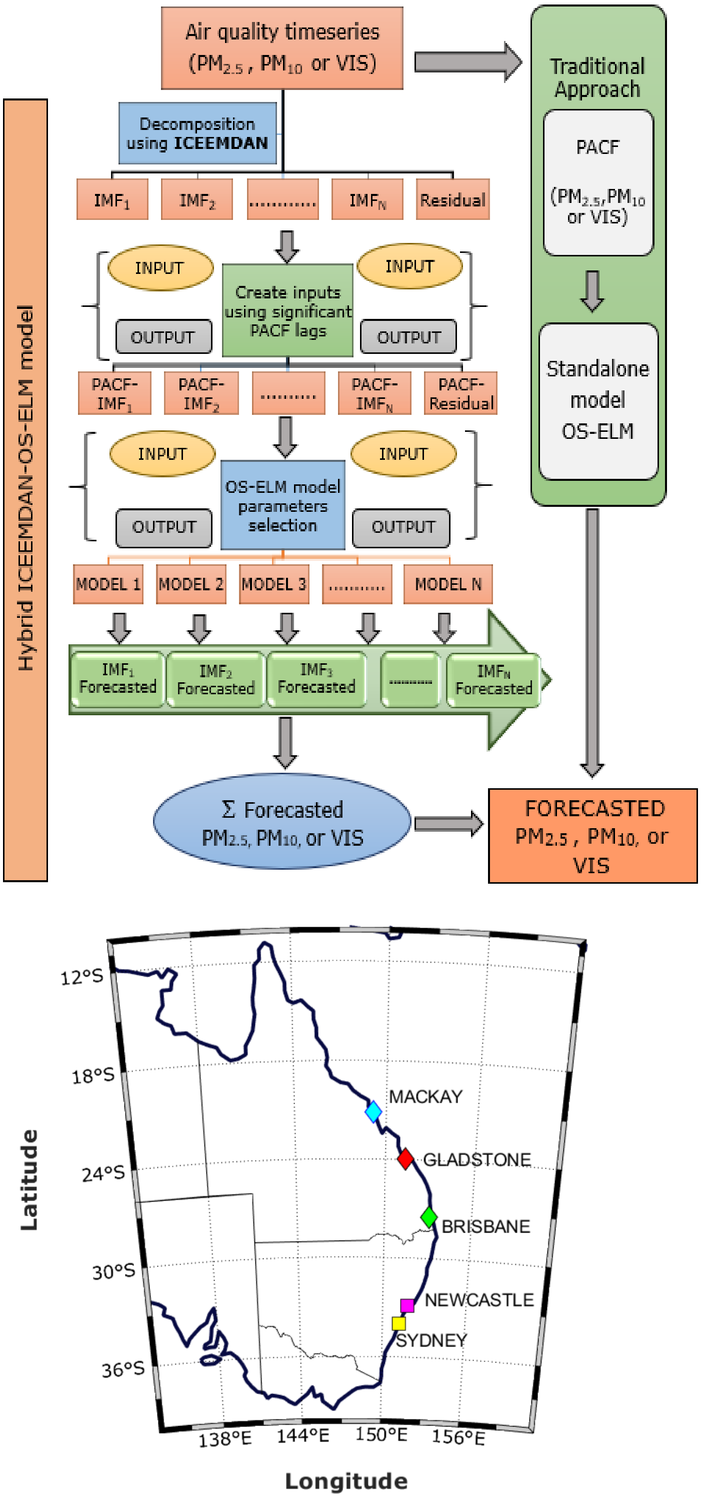
This chapter presents an exact copy of the published article in ***Science of the Total Environment*** journal (Vol: 709 (2020): Page(s): 135934 – 135957, *ISSN:* 0048-9697).

This work focussed on significant challenges of air quality such as the chaotic, non-linear, and high dimensional nature of predictor variables. A practical tool was formulated that produces real-time forecasts to mitigate public health risks. The novelty of this research is the development of the proposed hybrid early warning artificial intelligence (*AI*) framework that can emulate hourly *AQ* variables (*i.e*., Particulate Matter 2.5, $PM_{2.5}$; Particulate Matter 10, PM_{10} and lower atmospheric visibility, *VIS*), associated with increased respiratory induced mortality and recurrent health-care cost. Firstly, hourly *AQ* data series (January-2015 to December-2017) are demarcated into their respective intrinsic mode functions (*IMFs*) and a residual sub-series that reveal patterns and resolve data complexity characteristics, followed by a partial autocorrelation function applied to each *IMF* and residual sub-series to unveil historical changes in *AQ*. To design the prescribed hybrid model, the data are partitioned into training (70%), validation (15%), and testing (15%) sub-sets. The online sequential-extreme learning machine (*OS-ELM*) algorithm integrated with improved complete ensemble empirical mode decomposition with adaptive noise (*ICEEMDAN*) is designed to robustly extract predictive patterns by fine-tuning the model generalisation to a near-optimal global solution, representing modelled *AQ* at hourly forecast horizons. The resulting early warning *AI*-based framework denoted as *ICEEMDAN-OS-ELM* is individually constructed by aggregating the sum of forecasted *IMFs* and residual sub-series. The results are benchmarked with several competing approaches- *e.g*., *ICEEMDAN*-multiple-linear regression (*MLR*), *ICEEMDAN-M5* model tree, and standalone: *OS-ELM, MLR, M5* model tree. Statistical metrics including the root-mean-square error *(RMSE)*, mean absolute error *(MAE)*, Willmott's Index *(WI)*, Legates McCabe's Index *(E_{LM})*, and Nash–Sutcliffe coefficients *(E_{NS})* also evaluated the model's accuracy. Both visual and statistical results register superior results for the *ICEEMDAN-OS-ELM* model, outperforming alternative approaches for all *AQ* variables (*i.e*., $PM_{2.5}$, PM_{10} & *VIS*). For instance, for $PM_{2.5}$, E_{LM} values ranged from *0.65–0.82 vs. 0.59-0.77* (*ICEEMDAN-M5* tree), *0.59-0.74* (*ICEEMDAN-MLR*), *0.28-0.54* (*OS-ELM*), *0.27-0.54* (*M5* tree) and *0.25-0.53* (*MLR*). Furthermore, *ICEEMDAN-OS-ELM* registered the lowest errors, ranging from *0.7-1.03 µg/m³ (MAE), 1.01-1.47 µg/m³ (RMSE)* for $PM_{2.5}$ whereas for PM_{10}, these metrics value were *1.29-3.84 µg/m³ (MAE), 3.01-7.04 µg/m³ (RMSE)* and for Visibility, they were *0.01-3.72 µg/m³* (*MAE* (*Mm^{-1}*)), *0.04-5.98 µg/m³* (*RMSE* (*Mm^{-1}*)). Visual analysis of forecasted and observed *AQ* through a Taylor diagram illustrated the objective model's preciseness, confirming the versatility of the *AI* model in generating *AQ* forecasts. The excellent performance ascertains the *AI* model's potential utility for *AQ* monitoring and subsequent public health risk mitigation.

3.2 Research Highlights

- An artificial intelligence predictive framework devised for air quality prediction.
- Efficient forecasting of air quality with five modelling approaches was recorded.
- *OS-ELM* coupled with *ICEEMDAN* outperformed the other models.
- *AI* models show potential in health informatics and Australia's environment sector.
- *AI* models can empower public health risk mitigation to create a liveable society.

3.3 Graphical Abstract



3.4 Published Article I

This article cannot be displayed due to copyright restrictions. See the article link in the Related Outputs field on the item record for possible access.

CHAPTER 4 A DEEP AIR QUALITY FORECASTS: SUSPENDED PARTICULATE MATTER MODELING WITH CONVOLUTIONAL NEURAL AND LONG SHORT-TERM MEMORY NETWORKS ALGORITHMS

4.1 Foreword

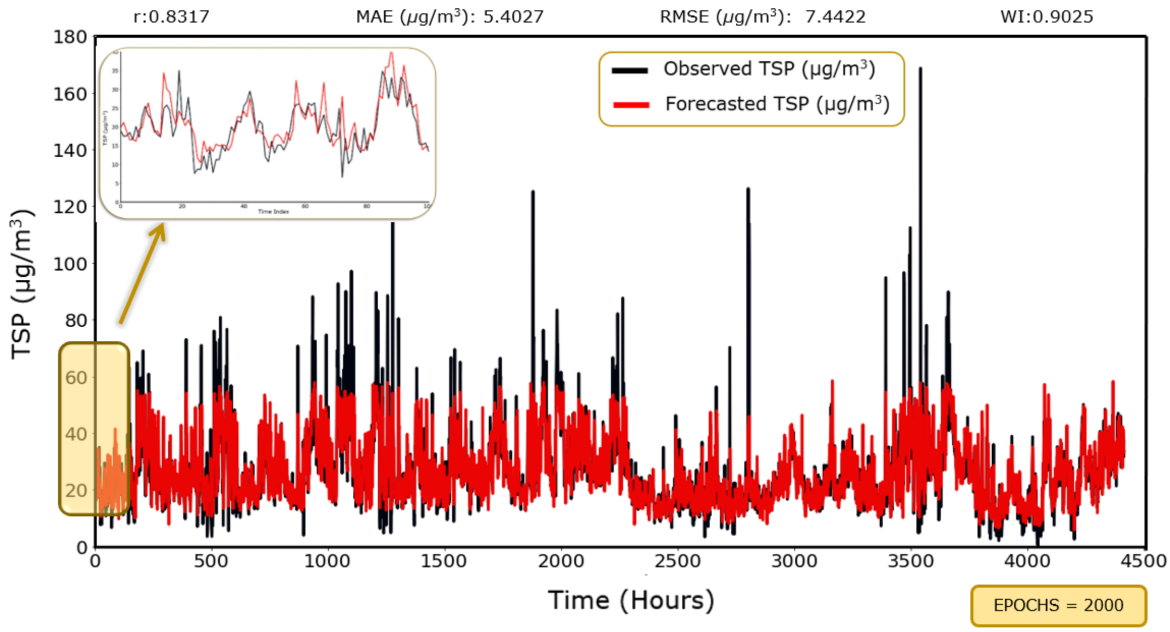
This chapter presents an exact copy of the published article in ***IEEE ACCESS*** (Vol: 8 (2020): Page(s): 209503 – 209516, *ISSN:* 2169-3536).

Total Suspended Particulate (*TSP*) was the subject of this paper, built on Chapter 3. Models were developed for forecasting visibility reducing particles and *PM* using conventional artificial intelligence models, albeit without considering *TSP*. This study built a deep learning hybrid *CLSTM* model where a convolutional neural network (*CNN*) is amalgamed with the long short-term memory (*LSTM*) network to forecast hourly *TSP*. The *CNN* model entailed a data processer including feature extractors that draw upon statistically significant antecedent lagged predictor variables, whereas the *LSTM* model encapsulated a new feature mapping scheme to predict the next hourly *TSP* value. The hybrid *CLSTM* model was comprehensively benchmarked and was seen to outperform an ensemble of five machine learning models. The efficacy of the *CLSTM* model was elucidated in the model testing phase at study sites in Queensland, Australia. Using performance metrics, visual analysis of *TSP* simulations relative to observations, and detailed error analysis, this study ascertained the *CLSTM* model's practical utility for air pollutant forecasting systems in health risk mitigation. This study captured a feasible opportunity to emulate air quality at relatively high temporal resolutions in global regions where air pollution is a considerable threat to public health.

4.2 Research Highlights

- A hybrid *CLSTM* model was generated by integrating with *CNN* and *LSTM* for air quality forecasting.
- *CLSTM* is evaluated against an ensemble of five other competing models.

- A comprehensive evaluation with statistical metrics, charts, and plots of tested data demonstrated that the hybrid *CLSTM* model was able to yield relatively better forecasts and has potential applications in air quality systems.
- *CLSTM* architecture was considerably robust in forecasting hourly data, and therefore indicated its possible application in the air pollution model and future public health and air quality studies.



4.3 Graphical Abstract

4.4 Published Article II

Received November 11, 2020, accepted November 16, 2020, date of publication November 18, 2020, date of current version December 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3039002

Deep Air Quality Forecasts: Suspended Particulate Matter Modeling With Convolutional Neural and Long Short-Term Memory Networks

EKTA SHARMA[1], (Member, IEEE), RAVINESH C. DEO[1], (Senior Member, IEEE), RAMENDRA PRASAD[2], ALFIO V. PARISI[3], AND NAWIN RAJ[1]

[1]School of Sciences, University of Southern Queensland, Springfield Central, QLD 4300, Australia
[2]Department of Science, School of Science and Technology, The University of Fiji, Lautoka, Fiji
[3]Centre for Applied Climate Sciences, School of Sciences, University of Southern Queensland, Toowoomba, QLD 4350, Australia

Corresponding authors: Ekta Sharma (ekta.sharma@usq.edu.au) and Ravinesh C. Deo (ravinesh.deo@usq.edu.au)

The work of Ekta Sharma was supported by the Research Training Scheme from the Australian Government and the School of Sciences, University of Southern Queensland.

ABSTRACT Public health risks arising from airborne pollutants, *e.g.*, Total Suspended Particulate (*TSP*) matter, can significantly elevate ongoing and future healthcare costs. The chaotic behaviour of air pollutants posing major difficulties in tracking their three-dimensional movements over diverse temporal domains is a significant challenge in designing practical air quality systems. This research paper builds a deep learning hybrid *CLSTM* model where convolutional neural network (*CNN*) is amalgamed with the long short-term memory (*LSTM*) network to forecast hourly *TSP*. The *CNN* model entails a data processer including feature extractors that draw upon statistically significant antecedent lagged predictor variables, whereas the *LSTM* model encapsulates a new feature mapping scheme to predict the next hourly *TSP* value. The hybrid *CLSTM* model is comprehensively benchmarked and is seen to outperform an ensemble of five machine learning models. The efficacy of the *CLSTM* model is elucidated in model testing phase at study sites in Queensland, Australia. Using performance metrics, visual analysis of *TSP* simulations relative to observations, and detailed error analysis, this study ascertains the *CLSTM* model's practical utility for air pollutant forecasting systems in health risk mitigation. This study captures a feasible opportunity to emulate air quality at relatively high temporal resolutions in global regions where air pollution is a considerable threat to public health.

INDEX TERMS Air quality forecasting, convolutional neural networks, deep learning, long short-term memory networks.

NOMENCLATURE

APF	Air Pollutant Forecasting.
AQ	Air Quality.
DL	Deep Learning.
$\mu g/m^3$	Micrograms per cubic metre.
$PM_{2.5}$	Fine particles with size 2.5 mm or less.
TSP	Total Suspended Particulate Matter up to about 100 μm In Diameter.
TSP_i^{FOR}	Forecasted *TSP* for i^{th} observation.
RELU	Rectified Linear Units.
WI	Willmott Index of agreement (*WI*)
ADAM	Adaptive Moment Estimation.
MAPE	Mean Absolute Percentage Error (%).
E_{NS}	Nash–Sutcliffe Efficiency.
RELU	Rectified Linear Units.
PM_{10}	Coarse particles between 2.5 And 10 mm.
CLSTM	Deep Learning hybrid Convolutional Long Short-term Memory Neural Network.
TSP_i^{OBS}	Observed *TSP* for i^{th} observation.
L	Legates and Mccabe Index.
r	Pearson's correlation coefficient.

The associate editor coordinating the review of this manuscript and approving it for publication was Asad Waqar Malik.

I. INTRODUCTION

As urbanisation progresses, air pollution is becoming an alarming environmental and societal concern. Besides regular monitoring efforts, there is a rising demand for short-term air pollutant forecasting (*APF*) system. This system can benefit

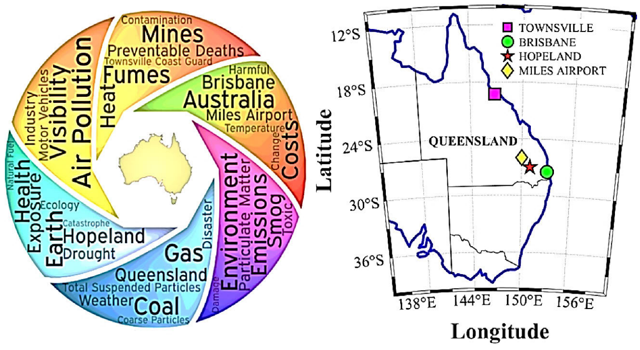
This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

governments in its health policymaking, traffic control in times of heavy pollution, and protecting vulnerable factions of the society (*e.g.* senior citizens, people with health ailments, pregnant ladies, and children).

One critical environmental pollutant that has an indiscriminate emission footprint is particulate matter (*PM*). The diversity and complexity of *PM* movements make the ongoing analysis and forecasting of this health hazard a critically challenging task. This atmospheric property has been extensively studied (*e.g.* [1], [2]). *PM* constitutes $PM_{2.5}$ (Particulate Matter *2.5*, or fine pollutants $< 2.5\ \mu m$ (micrometres)), PM_{10} (Particulate Matter *10,* or coarse pollutants $(2.5 - 10)\ \mu m$), and total suspended particulate matter (*TSP or SPM*) *i.e.* all airborne particles up to $100\ \mu m$ in diameter [3]. *TSP*, measured in μg/m^3, is the subject of this paper, built on our earlier study [4] where we modelled visibility reducing particles and *PM* using conventional artificial intelligence models, albeit without considering *TSP*. This research, therefore, aims to design an *APF* system that predicts *TSP* responsible for recurrent healthcare costs and increased nuisance through the soiling of property and materials. The primary source of *TSP*, being combustion (*e.g*., engines, bushfires, dust, mining, and industrial processes), is currently rising in many nations. There are very few studies globally that studied the acute effects of *TSP* and mortality [3], [5]. Therefore, new research is crucial for modelling the behavior of this health hazard.

In recent years, the forecasting of air pollutants has been accomplished through two methods: first, dynamic or physical models [6], and second, data-based statistical or artificial intelligence models [7] are considered. Physical models provide good accuracy of the physical processes, however, some studies report a significant model-error and a need for long run-times that make the model difficult to implement over a short-term horizon [8]. Fortunately, neural network models can address such issues as they only require proper datasets to speed up the learning and model convergence. Owing to the limitations of physics-based models, an *APF* system based on artificial intelligence can be a viable option to produce better results [9]. In atmospheric predictions, deep learning (*DL*) has attracted considerable attention. Many studies [10] are showing its ability to attain high forecasting precision than its earlier counterpart, or non-*DL* models. Numerous works have implemented *DL* in a diverse range of applications such as solar radiation [11], [12], pain intensity estimation [13], and seizure diagnosis [14]. In these studies, and the others, *DL* was commended for its superior capability to handle complex data (e.g., *TSP*) and approximation through stochastic variables analysis with a nonlinear feature mapping capability.

To develop an *APF* system, this research adopts a convolutional neural network (*CNN*). This is a renowned *DL* algorithm employing efficient multistage architecture through convolution, pooling, and fully connected layers for effective task-dependent and non-handcrafted data attribute representation [15]. Further improvements can be achieved through a secondary *DL* architecture based on long short-term memory (*LSTM*) network [16]. *LSTM's* key merits are that it can



(a) (b)

FIGURE 1. (a) A word crunch showing the problems caused by air pollution in Australia. (b) Study sites in Queensland Australia where the proposed *CLSTM* was implemented.

resolve to vanish gradient issues and explore sequential data relationships through unique (input, forget, output) gates. The amalgamation of *CNN* with *LSTM* is fast becoming a popular area of research in *AQ* and there are some important global studies such as [17]–[19]. However, the construction of the hourly *APF* system for *TSP* and especially for Australia is yet to be explored. Following the aforesaid, there appear to be gaps in the literature. *Firstly,* there is rather fragmentarily available literature involving *DL* algorithms for *TSP* prediction, and *Secondly*, there appear to be gaps in *TSP* prediction application in Australia despite a handful of studies performed elsewhere, *e.g. TSP* concentration models in China sea with back-propagation neural network (*BPNN*) [20], multi-layer filtration system [21], modelling *TSP* with light scattering [22], and with coulter sensors [23]. However, none of these studies have forecasted *TSP w.r.t* air quality. To address such issues, this research aims to build an *APF* system for short-term (hourly) *TSP* forecasts using the *LSTM* algorithm as a versatile model integrated with *CNN* as a feature extraction framework, as with $PM_{2.5}$ [24] and other studies [9], [25]. Despite air pollution causing 3000 premature deaths in Australia with a staggering annual health expense of *AUD* 24.3 billion, there is a dearth of research [26]. Recent spine-tingling Black Summer bushfires across South-eastern Australia (July 2019–February 2020) [27], [28] and ancient dust-storms registered severe air pollution episodes. Australia is an arid continent where the rising global temperature and bushfires, health, and environmental challenges illustrated graphically (Fig. 1 a) are considered serious [29]. An optimal *AQ* target is a critical issue but this is hindered by ineffective policies that warrant the development of a real-time and robust *APF* methodology [4]. Considering this, the novelty of this paper is as follows:

i) A practical study is presented to generate a computationally efficient architecture with the *DL* hybrid *'CLSTM'* model. Our approach aims to address gaps in hourly *TSP* modelling for a user-friendly *APF* system.
ii) *Firstly,* our model design phase employs an ensemble of competing machine learning models: Random Forest (*RF*), Volterra, M5 model tree, and Multiple

Linear Regression (*MLR*) to forecast hourly *TSP. Secondly,* the *LSTM* algorithm is applied to an ensemble of forecasted *TSP* to compare its results with the final output. Thereafter, a three-layered *CNN* model robustly extracts remaining data features *i.e.*, statistically significant antecedent inputs from 'four' (excluding *LSTM*) benchmarked models. The last layer analyses all features involving independent *LSTM* fusion to finally forecast the next hour *TSP* (μg/m^3). This multi-modeling overcomes the inherent deficiencies in single models [30].

iii) The *CLSTM* necessitates the data fusion through a refining process by *CNN* and features mapping by *LSTM*. Extensive evaluation through careful selection of last hour's meteorological variables over hotspots in Queensland, Australia, shows reliable forecasts.

iv) *CLSTM's* efficacy is explored by statistical metrics, visual analysis of forecasted and observed *TSP*, and comprehensively benchmark against an ensemble of five machine learning models including *LSTM*.

v) The hourly predictions of *TSP*, which is near real-time, can help combat public health issues through proactive advisory, planning, and implementing a regulatory *AQ* system, and making this research significantly unique for global applications for human health benefits.

II. THEORETICAL OVERVIEWS

We now provide a brief overview of the objective model (*i.e.*, *CNN* & *LSTM*). The theoretical explanation of standalone models *i.e. MLR* [31], M5 model tree [32], Volterra [30], and Random Forest [33] are presented elsewhere as these are well-known methodologies.

A. FEATURE EXTRACTION: CONVOLUTIONAL NEURAL NETWORK

In this research, *CNN* constructs the proposed *APF* system, denoted as *CLSTM*, which is applied for hourly *TSP* prediction. *CNN*'s are relatively successful Feed Forward Neural Networks [34]. However, few studies *e.g.* [35] have applied *CNN* for *AQ* research. A typical CNN architecture consists of Convolutional layers (*CON*) discovering the data patterns. This transforms local relationships in input features or images using a kernel. A Pooling layer (*POOL*) reduces the target variable dimension while a Fully connected layer (*FC*) generates the probability of each category implied in initial input data [36]. If a = activation function, W^f = kernel's weight with feature map f^{th}, and $*$ = an operator of convolutional process, each convolutional layer extracts *TSP* pattern with a lagged matrix, expressed mathematically as:

$$h_{ij}{}^{k} = a((W^f * x)_i j + b_k) \tag{1}$$

It should be noted that *CNN*'s are computationally intensive in extracting hidden nonlinear features. These features can help a *CNN* model to create filters representing the data patterns [37]. This paper aims to simplify the modelling technique, mainly to satisfy real-time usage for hourly *TSP* architecture. A one-dimensional (1-D) *CON* operator directly forecasts 1-D *TSP* data with a grid search being used to select the channels through three *CON* layers. *Adam* is an optimisation algorithm and *ReLU* is an optimisation algorithm such that *ReLU* is:

$$f(x) = \max(0, x) \tag{2}$$

B. LONG SHORT-TERM MEMORY NETWORK: TIME SERIES PREDICTION

LSTM is a Recurrent Neural Network (*RNN*) [38] with the ability to learn long-range dependence between target and input variables. Hence *LSTM's* are becoming useful in forecasting time series variables, especially in *AQ* research as the objective is to consider dependence on a forecast horizon suited for sequential prediction whilst also evading the gradient decay issues. The memory block of *LSTM* has input, output, and forget gates, assisting the update of information flow, thus making it an excellent choice for *TSP* modelling. This can continuously update the next forecasted value [39]. Successful usage of *LSTMs* are language modelling [40], speech recognition [41], and *AQ* research [42]. The calculations are as follows [9]:

I. If S_{n-1} = Last hidden state, I_n = new input, F_n = forget gate, W_m = Weight matrices, b_m = bias vector, i_n = input gate, $\sigma(\ldots)$, $\tanh(\ldots)$ = Activation functions (Logistic, sigmoid, hyperbolic), the cell state is represented by:

$$F_n = \sigma\left(W_m.\left(S_{n-1}, I_n\right) + b_m\right) \tag{3}$$

A candidate cell state δ_n decides what information will be stored, scaled by i_n –

$$\tanh\left(W_c.\left(S_{n-1}, I_n\right) + b_c\right) \tag{4}$$

$$i_n = \sigma\left(W_i.\left(S_{n-1}, I_n\right) + b_i\right) \tag{5}$$

II. Cell state δ_n combines the earlier state and the present state (δ_{n-1}, δ_n). Here δ_{n-1} is scaled by i_n and δ_{n-1} by F_n:

$$\delta_n = (F_n * \delta_{n-1} + I_n * \delta_n) \tag{6}$$

III. Finally, for the output process, ("output gate" θ_n decide the output state $\mathcal{C}_n$. Here θ_n is filter for output S_n :

$$\theta_n = \sigma(W_\theta.(S_{n-1}), I_n) + b_\theta) \tag{7}$$

$$S_n = \theta_n.tanh(\delta_n) \tag{8}$$

III. MATERIALS AND METHOD

A. RESEARCH AREA

To appraise CLSTM, we utilize hourly air pollutants or TSP (μg/m^3) for Queensland (Qld). Table 1 (a-b) describes the study sites and all relevant data. *Qld* is the second largest state in Australia, exhibiting a soaring rate of greenhouse gas emissions per capita [43]. It has nine (out of ten) worst mines that generate *PM*, causing major respiratory issues with cancer cases [44]. A spatial picture of the study sites is illustrated in Fig. 1 (b). The *Qld*-based

TABLE 1. **(a) Geographic description (b) Data segregation of study sites.**

Station Name	*Location*		
	Longitude (°E)	*Latitude (°S)*	*Elevation (m)*
Brisbane	153.08	27.46	9.0
Townsville	146.81	19.25	9.0
Hopeland	151.04	27.04	316.0
Miles Airport	150.16	26.81	302.0

(a)

(b) *Station Name*	*Data points*	*Training*			*Validation*			*Testing*		
		Period	*Points*	*%*	*Period*	*Points*	*%*	*Period*	*Points*	*%*
Brisbane	35064	01-Jan-2015 to 31-Dec-2017	26304	75.01	01-Jan-2018 to 30-June-2018	4344	12.38	01-July-2018 to 31-Dec-2018	4416	12.59
Townsville	35064	01-Jan-2015 to 31-Dec-2017	26304	75.01	01-Jan-2018 to 30-June-2018	4344	12.38	01-July-2018 to 31-Dec-2018	4416	12.59
Hopeland	34752	14-Jan-2015 to 31-Dec-2017	25992	74.79	01-Jan-2018 to 30-June-2018	4344	12.50	01-July-2018 to 31-Dec-2018	4416	12.70
Miles Airport	30678	02-Jul-2015 to 31-Dec-2017	21918	71.45	01-Jan-2018 to 30-June-2018	4344	14.16	01-July-2018 to 1-Dec-2018	4416	14.39

TABLE 2. **Descriptive statistics of *TSP* (μg/m^3) for each study site.**

Objective Variable	*Study Site*	*Maximum*	*Minimum*	*Mean*	*Median*	*Variance*	*Skewness*	*Kurtosis*	*Standard-Deviation*
***TSP* (μg/m^3)**	Brisbane	2568.4	0.10	24.83	21.4	556.48	41.12	3941.85	23.59
	Townsville	439.9	0.10	25.92	23.2	221.71	3.19	36.46	14.89
	Hopeland	992.3	0.01	17.87	11.6	990.36	12.75	246.42	31.47
	Miles Airport	3709.1	0.40	36.56	20.5	6310.71	13.23	325.94	79.44

air monitoring hotspots are Brisbane, Townsville, Hopeland, and Miles Airport. The Brisbane station monitors *PM* (and *TSP*) which is potentially relevant to emissions from rail wagons transferring coal to the Port of Brisbane. Townsville station began in 2007 in the form of a dust monitoring program by considering community concerns regarding dust impacts from the Port of Townsville. Rural sites: Hopeland and Miles Airport are responsible for assessing *AQ* near an area of intensive coal seam gas productions. Considering the need to develop forecast models for *TSP* especially for these important hotspots, the study designed the proposed *CLSTM* model. Table 2 displays air pollutant *TSP's* inferential statistics. Data were acquired from the *Qld* Department of Environment and Science. Data are continuously evaluated for quality through preliminary analysis. Table 1 (b) shows missing data ('**x**') caused by equipment servicing and other instrumental failures. Following statistical procedures, the hourly mean calendar values were used to replace missing data [45].

B. DEEP HYBRID CLSTM MODEL ARCHITECTURE

Hourly TSP data were used to develop six forecast models shown in the overall schematic (Figure 2) of the proposed deep learning (DL) hybrid CLSTM model. The objective model is compared to the DL model LSTM, including non-DL versions: Volterra, Random Forest, MLR, and M5 model tree. All models were developed on Windows 10 platform Intel®i7 Generation 9 @ 3.7 gigahertz processing unit, 16 GB memory, Python programming language, and freely available open-source libraries (i.e., Keras [46], Tensor Flow [47], and Scikit-learn [48]). To develop a hybrid CLSTM model, data were firstly partitioned; and although there is no consensus over data partitioning ratios [49]; this issue is a critical consideration as it affects the feasibility and capability of the model. In absence of any designated rule [50], the TSP data were initially divided into training sets (01-January-2015 to 31-December-2017) for Brisbane and Townsville (i.e. 75%). For Hopeland, and Miles Airport the period considered was 14-January-2015 to 31-December-2017 (i.e. 75%), and 02-July-2015 to 31-December-2017 (i.e. 72% or 30,678) when these stations started their AQ monitoring. This data division ensured a universally representative hourly-step horizon. Table 1(b) shows that a six-monthly period for validation (01-January-2018 to 30-June-2018) and testing (01-July-2018 to 31-December-2018) purposes for all stations. This is consistent with published literature that emphasises the criticality of partitioning into subsets before model construction to avoid leakage of validation and training data over the upcoming testing subset, thereby introducing a testing bias [51]. The sample data from the training set provided model estimation through hyper-parameters hence the validation set became an essential component of modelling. Table 2 enumerates the descriptive statistics of measured AQ, and Table 3 discusses the model input variables used for hourly TSP (μg/m^3) forecasting. Following earlier literature,

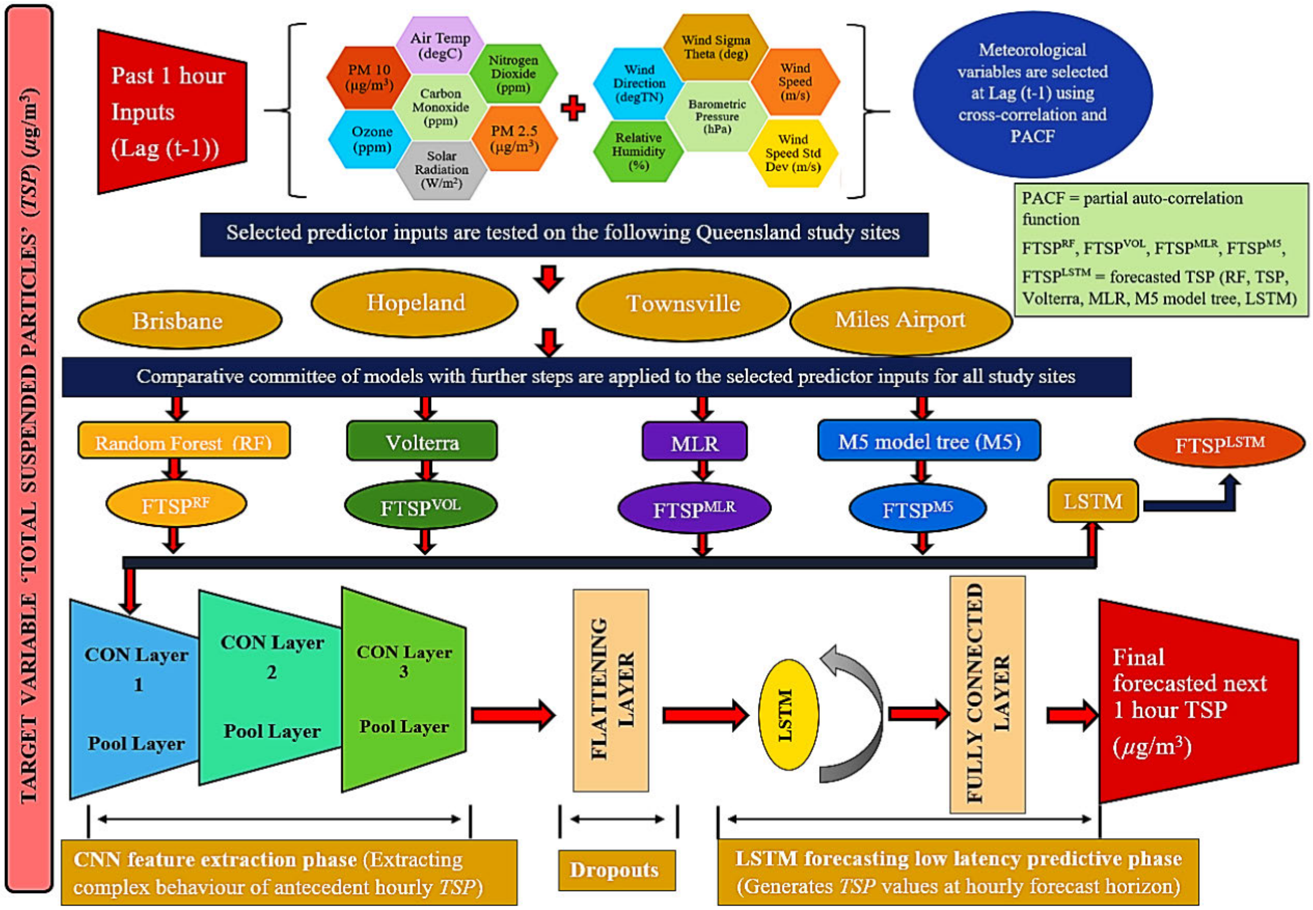


FIGURE 2. Schematic of *APF* system using deep learning hybrid model (*CLSTM*).

the chaotic nature of AQ requires a normalisation process before modelling to conform all data to be within [0, 1] [52]:

$$TSP_{NORM} = (TSP - TSP_{MIN})/(TSP_{MAX} - TSP_{MIN}) \quad (9)$$

In (9), TSP_{MIN}, and TSP_{MAX} = the minimum and maximum values of *TSP*, respectively, while TSP_{NORM} = the normalised *TSP*. In Table 3, Wind (*W*), Direction (*D*), Speed (S), Air Temperature (*AT*), Relative Humidity (*RH*), Barometric pressure (*BP*), (NO_2), Carbon monoxide (*CO*), and Solar Radiation (*RAD*). The next step in model design is illustrated in Fig. 3 (a)-(d). It identifies cross-correlation coefficients (r_{cross}) investigating the co-variance between hourly *AQ* data *vs.* respective predictor variables used to build a hybrid *CLSTM* for all study sites. The blue line indicates a 95% significance boundary. Fig. 4 (a)–(d) utilises partial autocorrelation function (*PACF*) to deduce the correlation of *TSP* series with its own historical lagged values, regressing the shifted series to determine these correlations. Successive time-shifted predictors created by this method denoted as *TSP* ($t - n$) where n is the value of the lag (*e.g.*, $n = 1$ in Fig 4) reveals good correlation with past *TSP* at all sites. Through this method, the significant lagged *TSP* series is then

TABLE 3. Model input variables. × = data is not monitored.

Variable	*Units*	*Brisbane*	*Townsville*	*Hopeland*	*Miles*
WD	(°*TN*)	✓	✓	✓	✓
WS	(*m/s*)	✓	✓	✓	✓
WST	(°)	✓	✓	×	×
WSSD	(*m/s*)	✓	✓	×	×
AT	(°*C*)	✓	✓	✓	✓
RH	(*%*)	✓	×	✓	✓
BP	(*hPa*)	✓	×	×	×
PM_{10}	(μg/m³)	×	✓	×	✓
$PM_{2.5}$	(μg/m³)	×	×	×	✓
O_3	(*ppm*)	×	×	✓	✓
NO_2	(*ppm*)	×	×	✓	✓
CO	(*ppm*)	✓	✓	✓	✓
RAD	(*W/m²*)	✓	✓	✓	✓

employed to build the proposed *CLSTM* model. In Table 4, the optimal framework of *CLSTM* (boldfaced in **red**) *vs.* the comparative models is illustrated. Notably, a grid search was adopted to select the hyperparameters and corresponding

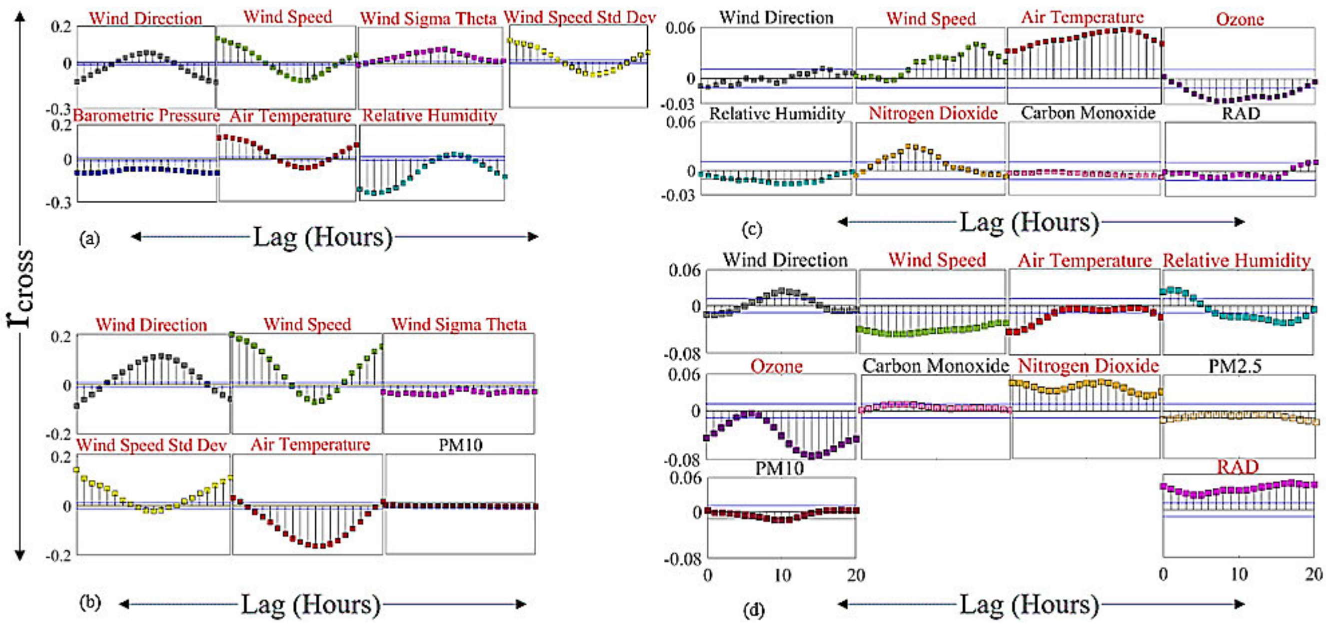


FIGURE 3. Cross-correlations (r_{cross}) investigating co-variance between the *AQ vs.* predictor variables used to build the proposed *APF* system using a hybrid *CLSTM* model. Note: The selected variables are captioned in red (a) Brisbane (b) Townsville (c) Hopeland, (d) Miles Airport.

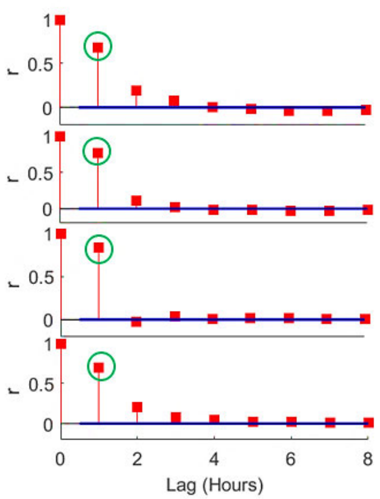


FIGURE 4. Partial autocorrelation of TSP time-series. The green symbol shows the most significant lagged TSP used to develop the proposed APF system using the hybrid CLSTM model. (a) Brisbane (b) Townsville (c) Hopeland, (d) Miles Airport.

optimal value range to finally reach an optimum framework for best feature extraction. The starting numbers are determined as per published studies conducted for *DL* [9], [12]. The model's improvement was monitored by the successive addition of numbers until we reach an optimal set. For instance, as shown in Table 4, when *CLSTM* was tested for a good batch size (from 400 as it is big data) and it continued till the performance became optimal (*i.e.* 2000 epochs and 500 batch size). It is noteworthy that the optimal number of hidden neurons plays a significant role in determining the most feasible model architecture. In this study, it was ensured that the grid search enabled the model to correctly learn the predictor data patterns, and thus, avoided issues like overfitting that can occur in cases when a significant model framework does not converge in a given run time or a small framework might lack the appropriate degrees of freedom to fully extract data patterns [53]. The lowest training *MAE* or *RMSE*, which was obtained by a sequence of hidden neurons in gradual steps was adopted to determine the optimal model architecture [54]. After trialling SoftMax, tangent hyperbolic, and sigmoid functions, '*ReLU'* was found to be the optimal activation function attaining the best *CLSTM* performance. Hyper-parameters of all tested models were chosen through a grid search procedure, which could impose a huge time factor and significant computing costs associated with the machine learning process. Each model search generally takes 10∼11 hours with the computation training and testing time to get reduced to $<$ 20 min after deducing the optimal parameters [9]. Notwithstanding this, the training data size is known to affect the cost and hyperparameter selection [55]. The deep learning hybrid model *CLSTM* utilises a pooling layer to overcome overfitting issues in its training phase, helping minimise and control the parameters and computations involved. It should be clarified that the inputs of *CLSTM* are the hourly *TSP*'s lagged matrix; however, the output is the forecasted *TSP*. The next *i.e.* the fourth layer in the *CLSTM* framework is the *LSTM* architecture itself positioned to analyse the features and later pass those in forecasting *TSP* (μg/m^3) values for the next hour. As enunciated in grid search after hyper-parameter selection, the study utilises the following properties [9]:

- **Least Square Error and Absolute Deviation**(L1, L2-regularisation): The sum of absolute differences and square of differences between observed and forecasted

TABLE 4. Optimal architecture (in red for CLSTM) for $TSP(\mu g/m^3)$.

Model	Hyper-parameters	Grid Search for Optimal set
CLSTM	Epochs	[300, 500, 1000, **2000**]
	Activation function	[SoftMax, tanh, **ReLU**, sig]
	Batch size	[400, **500**, 800, 750, 1000]
	Pooling size, padding	[**2**], [same]
	Convolution Layer 1	[250, 200, 150, 100, **80**]
	Convolution Layer 2	[**20**, 40, 50, 60, 70, 80]
	Convolution Layer 3	[30, 10, 15, 20, **5**]
	LSTM layer (L1)	[30, 40, **50**, 60, 100, 150]
	Optimiser, Dropout	[**Adam**], [**0.1**]
LSTM	Epochs	[50, 100, 500, 1000, 2000]
	Activation function	[SoftMax, tanh, ReLU, sig]
	Batch Size	[300, 500, 800, 750, 1000]
	Optimiser, Drop rate	[Adam], [0.1, 0.2]
	LSTM filter	[50, 60, 100, 200]
Random Forest	Number of Trees	[50, 100, 200, 400, 500, 800]
	Leaf, Foot, Surrogate	5, 1, 'on'
Volterra	Threshold, Regressor	[0,1]
	A, n-start	[m X n] matrix [m>n], [>=1]
MLR	Intercept	[0,1]
	c	[0.0238, 0.758,0.183,0.157]
	α_1	[1.207,0.005,0.008, -0.021]
	α_2	[0.820,0.102, -0.087,0.267]
	α_3	[0.004, -0.012, -0.655, -0.345]
M5 model tree	Threshold, Smoothing	0.05, 15
	Minimum cases, rules	5, 20

The Architecture of the Backpropagation Algorithm	
Beta, β_1, β_2	0.990
Alpha, α	0.001
Epsilon, ε	0.0000001
β_1, β_2 = 1st, 2nd moment estimates exponential decay rate	
α = Learning rate, ε = Small number to prevent zero division	

TSP is minimised through penalisation parameters. Generally, L1, L2 parameters penalises the huge parameter values, thus reducing the models' nonlinearity.

- **Dropout:** is a regularisation adopted to improve the training performance and reduce overfitting. For all iterations, dropout picks up a neuron fraction between [0,1] (hyper-parameter). In this study, dropout = 0.1.
- **Activation Function:** SoftMax, tanh, sigmoidal were tested with Rectified Linear Unit (*ReLU*) as most optimal.
- **Early Stopping:** Kera's *DL* library eliminates overfitting [56] by setting the mode to "minimum" and patience to "45". Here, *ES* typically halts training when validation loss stops, decreasing the number of epochs, specified by the patience term.

Summarising these, the hybrid *CLSTM* was obtained in three distinct stages: **Selection:** Predictors based on meteorological variables at a lag of (*t–1* or using features of the past hour) were selected using r_{cross} and *PACF*. **Validation & Testing:** Validation of the selected inputs and testing on *Qld* sites was carried out through these models: Volterra, Random Forest, M5 model tree, and *MLR*. *LSTM* was applied to the forecasted *TSP* ($\mu g/m^3$) result. **Feature extraction & low latency predictive phase:** The three-layer *CNN* feature extraction phase was employed on forecasted *TSP* (four models). The fourth layer was the low latency predictive fusion phase with an independent *LSTM* on the flattened layer forming a single connected layer, giving the final forecasted next hour *TSP*, with results evaluated *via* equations (10) – (15).

C. MODEL PERFORMANCE CRITERIA

This subsection provides a rigorous evaluation of the hybrid CLSTM relative to its counterpart comparative models adopting a wide range of statistical criteria based on Pearson's correlation (r), Root-Mean-Square-Error (RMSE; $\mu g/m^3$), Mean Absolute Percentage Error (MAPE; %), Mean Absolute Error (MAE; $\mu g/m^3$), Willmott's Index of agreement (WI), Nash–Sutcliffe Efficiency (E_{NS}), and Legates and McCabe Index (L) [50], [57], [58] whose mathematical equations are:

I) Mean absolute error

$$(MAE) = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \left|\left(\boldsymbol{TSP^{FOR}} - \boldsymbol{TSP^{OBS}}\right)\right| \quad (10)$$

II) Root mean square error

$$(RMAE) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(TSP_i^{FOR} - TSP_i^{OBS}\right)^2} \quad (11)$$

III) Pearson's correlation coefficient (r) (12), as shown at the bottom of the next page

IV) Willmott Index of agreement (*WI*)

$$1 - \left[\frac{\sum_{i=1}^{N} \left(TSP_i^{FOR} - TSP_i^{OBS}\right)^2}{\sum_{i=1}^{N} \left(\left|TSP_i^{FOR} - \overline{TSP_i^{OBS}}\right| + \left|TSP_i^{OBS} - \overline{TSP_i^{OBS}}\right|\right)^2}\right] \quad (0 \leq WI \leq 1) \quad (13)$$

V) Nash–Sutcliffe Efficiency (E_{NS})

$$1 - \left[\frac{\sum_{i=1}^{N} \left(TSP_i^{FOR} - TSP_i^{OBS}\right)^2}{\sum_{i=1}^{N} \left(TSP_i^{OBS} - \overline{TSP_i^{OBS}}\right)^2}\right] \quad (\infty \leq E_{NS} \leq 1) \quad (14)$$

VI) Legates and McCabe Index (L)

$$1 - \left[\frac{\sum_{i=1}^{N} \left|TSP_i^{OBS} - TSP_i^{FOR}\right|}{\sum_{i=1}^{N} \left|TSP_i^{OBS} - \overline{TSP_i^{OBS}}\right|}\right] \quad (\infty \leq L \leq 1) \quad (15)$$

where TSP_i^{FOR}, TSP_i^{OBS} = forecasted and observed TSP for i^{th} observation, N = Total number, $\overline{TSP_i^{FOR}}$, $\overline{TSP_i^{OBS}}$ = mean forecasted and observed *TSP*. The study considered that a Gaussian error distribution is likely to imply *RMSE* to be a more appropriate measure of model accuracy compared to the *MAE* [59]. Note that a better performance is attained for *WI* and E_{NS} close to unity [50], [57], [58]. However, the value of *L* (a refined WI) is used to penalise model error more strictly, and so, it is a considerably robust metric, especially for large and relatively complex datasets (*e.g.*, *TSP*).

TABLE 5. Testing performance of *CLSTM vs.* Other competing models. (a) Brisbane, (b) Townsville, (c) Hopeland, (d) Miles airport.

	Models	*r*	*MAE* $\mu g/m^3$	*RMSE* $\mu g/m^3$	*MAPE* %	*WI*	E_{NS}	*L*
(a)	**CLSTM**	**0.93**	**3.54**	**13.97**	**8.43**	**0.96**	**0.86**	**0.80**
	LSTM	0.82	4.69	23.31	10.20	0.83	0.63	0.73
	RF	0.90	15.51	40.74	22.01	0.93	0.81	0.64
	Volterra	0.86	7.91	19.24	25.12	0.95	0.74	0.54
	MLR	0.86	8.13	20.27	27.08	0.89	0.71	0.53
	M5	0.38	11.66	46.71	36.07	0.55	0.52	0.33
(b)	**CLSTM**	**0.83**	5.40	**7.44**	**17.37**	0.91	**0.69**	**0.51**
	LSTM	0.81	5.12	7.94	19.99	0.89	0.64	0.45
	RF	0.79	5.65	8.73	26.74	0.77	0.61	0.44
	Volterra	0.76	**4.63**	9.05	21.56	**0.97**	0.54	0.42
	MLR	0.76	5.34	8.63	27.51	0.84	0.54	0.43
	M5	0.46	9.02	15.69	43.93	0.65	0.37	0.36
(c)	**CLSTM**	**0.67**	**7.98**	**21.97**	49.84	**0.79**	**0.43**	**0.35**
	LSTM	0.55	8.04	24.23	**38.44**	0.66	0.29	0.34
	RF	0.45	8.74	26.11	57.87	0.33	0.18	0.29
	Volterra	0.44	9.96	27.45	43.26	0.70	0.10	0.19
	MLR	0.44	8.91	25.97	59.78	0.51	0.19	0.27
	M5	0.37	9.80	29.62	45.80	0.54	0.10	0.20
(d)	**CLSTM**	**0.85**	**6.60**	**23.11**	**32.52**	**0.87**	**0.69**	**0.61**
	LSTM	0.77	7.76	26.66	53.16	0.81	0.58	0.53
	RF	0.68	9.04	30.27	37.18	0.71	0.45	0.46
	Volterra	0.65	9.31	31.94	35.03	0.82	0.39	0.44
	MLR	0.45	12.01	49.35	113.57	0.62	0.44	0.28
	M5	0.64	13.81	32.03	55.55	0.76	0.39	0.17

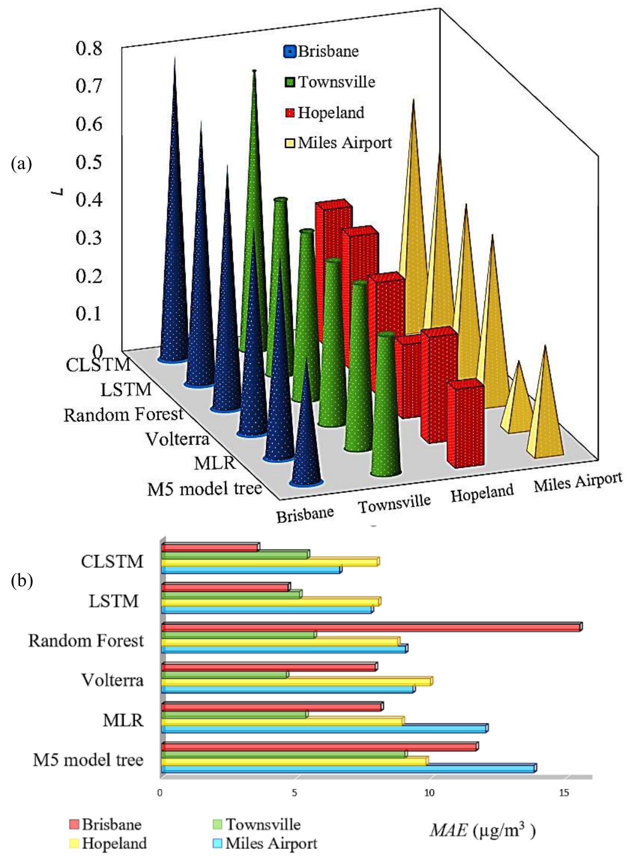


FIGURE 5. (a) Testing performance of *CLSTM vs.* the five other competing models evaluated using Legates & McCabe Index (*L*). (b) 3D-Bar graph of the mean absolute error *(MAE).*

IV. RESULTS AND DISCUSSIONS

This section provides empirical results to assess model performance, demonstrating the effectiveness of the newly designed *APF* system using a hybrid *CLSTM* approach. The results of hybrid *CLSTM* are benchmarked with the five other competing approaches; *e.g.*, *LSTM*, *RF*, Volterra, M5 model tree, and *MLR*. Further model appraisal has been achieved through several model evaluation metrics as described by (10)–(15). Table 5 (a)–(d) shows the predictive performance for all models evaluated in the testing phase. Comparing the results of the study site Brisbane plotted in Fig. (5–9) for *TSP*, this work attained the most precise forecasts for the case of the hybrid *CLSTM w.r.t* all six statistical metrics *(highest* $r \approx 0.93$*, lowest* $RMSE \approx 13.97\mu$g/m^3, $MAE \approx 3.54\mu$g/m^3, *highest* $WI \approx 0.96$, $L \approx 0.80$, and $E_{NS} \approx 0.86$*)* in comparison with statistical metrics of the other models. For instance, $r \approx$[*0.38-0.82*], $MAE \approx$ [*4.69-15.51*]μg/m^3, $MAPE \approx$ [*10.20-36.07*]%, $RMSE \approx$ [*19.24-46.71*] μg/m^3, *WI* $\approx$ [*0.55-0.95*], $E_{NS} \approx$ [*0.52-0.81*], and $L \approx$ [0.33-*0.73*] for the comparative ensemble of models, where the value before [-] is the lower bound and the value after [-] is the upper bound of a metrics. For a complete understanding of *CLSTM,* we evaluate the test performance with all benchmark models (*i.e.*, *LSTM*, *RF*, M5 Model Tree, and *MLR*) through the Legates and McCabe Index (*L*) and a 3D-bar graph of the mean absolute error (*MAE*) of hourly *TSP* forecasts, *i.e.*, Fig. 5(a), (b). The results show the highest '*L*′ and lowest errors, highlighted by the performance metrics. That is, we get the following results: Brisbane ($MAE \approx 3.54$ μg/m^3), Townsville ($MAE \approx 5.40$ μg/m^3), Hopeland ($MAE \approx 7.98$ μg/m^3), and Miles Airport ($MAE \approx 6.60$ μg/m^3), all of which accord with excellent performance of the model in terms of current literature [60]. Our deep learning hybrid model: *CLSTM* shows the highest *L* ($\approx$ *0.80, 0.51, 0.35, 0.61*) for all study sites when compared with the rest of the model ensembles. It should be noted that this is the most stringent performance measure and thus, is an indicator of the superior performance of the proposed *CLSTM* model [61]. Furthermore, the '*p*' value of the *CLSTM* model has been tested for all stations at significance level 0.05, 0.01, and 0.1. The *p*-value is < 0.00001 for all study sites, which indicates the result is significant at $p < 05$, 0.1, and 0.01. Testing performance of *CLSTM* model An alternative account of

$$\frac{\sum_{i=1}^{N}\left(TSP_i^{OBS}-\overline{TSP^{OBS}}\right)\left(TSP_i^{FOR}-\overline{TSP^{FOR}}\right)}{\sqrt{\sum_{i=1}^{N}\left(TSP_i^{OBS}-\overline{TSP^{OBS}}\right)^2}\sqrt{\sum_{i=1}^{N}\left(TSP_i^{FOR}-\overline{TSP^{OBS}}\right)^2}} \quad (-1 \le r \le 1) \tag{12}$$

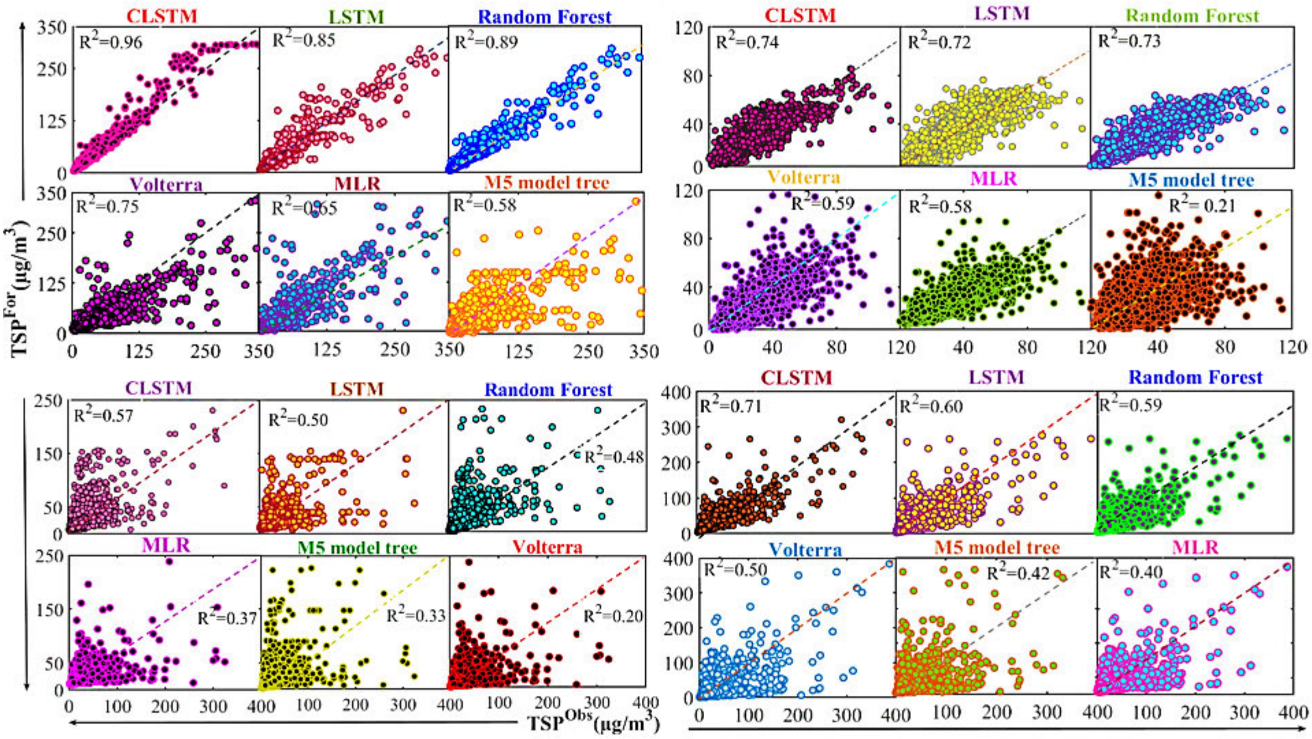


FIGURE 6. Scatterplot of forecasted ('For') *vs.* observed ('Obs') *TSP* in the model test phase. (a) Brisbane, (b) Townsville, (c) Hopeland, (d) Miles Airport. Note that the coefficient of determination (R^2) is indicated in each subplot.

the *CLSTM* model's accuracy is achieved through Pearson's correlation, 'r' comparing the hourly *TSP* (forecasted ('*For*') *vs.* observed ('*Obs*')) and the goodness-of-fit displayed *via* scatterplots (Fig. 6 (a)-(d)). Here, scatterplots are used to authenticate the consensus between regression, coefficient of determination, and predictors variables with a linear fit including the r^2 value are used to outline the model's accuracy [50]. Notably, for short-term forecasts, *CLSTM* gives the best performance to further ascertain its suitability. This deduction is supported by a high r^2-value for Brisbane (*CLSTM* ≈ *0.96*) *vs.* (*LSTM* ≈ *0.85*, Volterra ≈ *0.75*, *RF* ≈ *0.89*, M5 model tree ≈ *0.58*, and *MLR* ≈ *0.75*). Similarly, for Townsville and Miles Airport, the r^2–value for hybrid *CLSTM* is ≈ (*0.74, 0.71*) *vs. LSTM* ≈ (*0.72, 0.60*), *RF* ≈ (*0.73, 0.59*), Volterra ≈ (*0.59, 0.50*), *MLR* ≈ (*0.58, 0.40*), and M5 model tree ≈ (*0.21, 0.42*). It is noticed that, although all models performed relatively poorly for the case of Hopeland, the performance of *CLSTM* was better (≈ *0.57*) than the rest of the models. Importantly, the standalone models produced inferior results in forecasting hourly *TSP* with r^2 values of *LSTM* (≈ *0.50*), *M5* (≈ *0.33*), *RF* (≈ *0.48*), *MLR* (≈ *0.37*), and Volterra (≈ *0.20*). These results concur with the high values of *MAPE* and *RMSE* for this station as elaborated in further discussion. The proposed hybrid *CLSTM* model also showed superior performance *w.r.t* mean absolute error (*MAE*). For instance, in the case of Brisbane, *MAE* ≈ *3.54*μg/m^3 as compared to *MAE* ≈ *4.69*μg/m^3 (*LSTM*), *MAE* (*RF* ≈ *15.51*μg/m^3, Volterra ≈ *7.91*μg/m^3, *MLR* ≈ *8.13*μg/m^3, M5 model tree ≈ *11.66*μg/m^3). Importantly, *MAE* is significantly lower for all stations as compared to the remaining ensemble of models in the research.

It concurs from Table 5(a)–(d), that as the magnitude of *MAE* for the proposed hybrid *CLSTM* model is low (< 10%) for all the study sites, this generally falls in the category of an excellent model, see Ref [60]. However, there also appears to be a subtle variation in model accuracy when the root mean square error (*RMSE*) and the mean absolute percentage error (*MAPE*) is computed. For instance, the hybrid *CLSTM* model is seen to generate an *RMSE* ≈ *13.97*μg/m^3, *7.44*μg/m^3, and *MAPE* ≈ *8.43%*, *MAPE* ≈ *17.37%* for Brisbane and Townsville; (*i.e.*, a good model category with an error < 20%), and fair results with *RMSE* ≈ *21.97*μg/m^3, *23.11*μg/m^3 for Hopeland and Miles Airport (*i.e.*, a fair model category with error > 20%). Notably, the two stations (Hopeland & Miles Airport) generally incur high *RMSE* and *MAPE*, however, the lowest *RMSE* is registered by the hybrid *CLSTM* model in such a way that the error for the other models occurs in a higher range *i.e.*, [*24.23– 29.62*] μg/m^3 for Hopeland. This is also noted for the case of Miles Airport that has the lowest error for the hybrid *CLSTM* model, while the other models are seen to obtain a higher error in the range [*26.66– 49.35*] μg/m^3. Following our results presented to far, we can infer that *CLSTM* accedes to much greater accuracy than the five other models, and thus appears to be a versatile predictive tool in modelling *TSP* at hourly steps. To further confirm this, we revert to radial (or

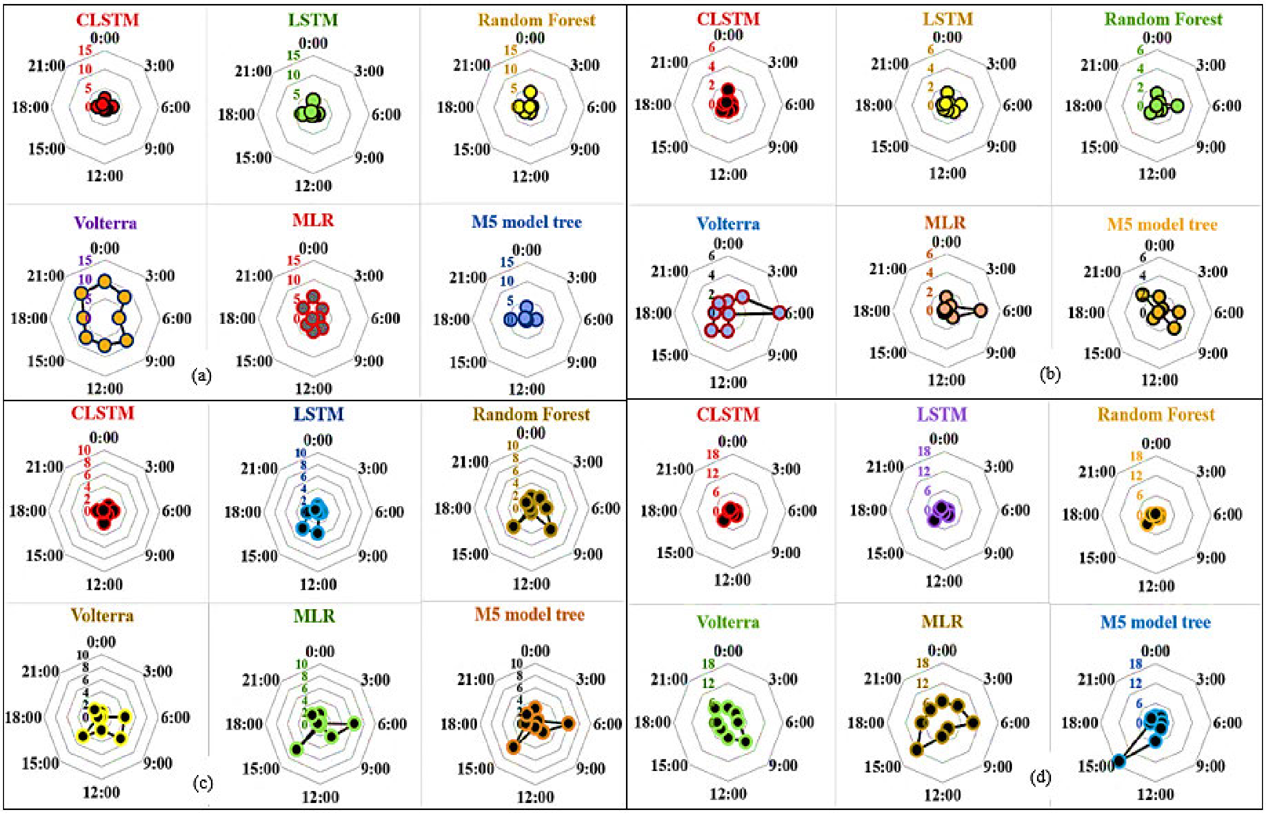


FIGURE 7. Relative forecasted error (%) in *TSP* generated by *CLSTM* vs. five other competing models. (a) Brisbane, (b) Townsville, (c) Hopeland, (d) Miles Airport sites. Note that the axis from origin denotes the proportion of model error per hour.

polar) plots showing the relative (or percentage) forecast error in Fig. 7 (a)-(d). In principle, the performance can be denoted as good when forecasted error (*FE*) is close to trivial. Therefore, the hybrid *CLSTM* model's results are seen to excel to a much greater degree in comparison to its counterpart models. For all study sites, these show an *FE* (%) value from midnight (0.00–23.00) in such a way that the *CLSTM* model elucidates a high degree of accuracy. This performance evaluation also shows that *CLSTM* was succeeded by *RF*, Volterra, *MLR,* and M5 tree with the latter showing high relative *FE* computed at an hourly time-step. Fig. 8 (a)-(d) illustrates a comparison of forecasted *vs.* observed *TSP* where data are averaged for the entire test phase. We have used red to show *CLSTM* simulations and black for observed data. The plots illustrate the preciseness of the objective model against competing approaches. In closing, Fig. 9 (a) illustrates the forecasted the testing ('*For*') *vs.* observed ('*Obs*') *TSP* at the hourly interval in phase for one of the study sites (Townsville). The hybrid *CLSTM* model is seen to yield comparatively high accuracy ($r \approx 0.83$, $MAE \approx 5.40\mu$g/m^3, $RMSE \approx 7.44\mu$g/m^3, $WI \approx 0.91$) between TSP^{Obs} and TSP^{For}. Importantly, the forecasted and observed *TSP* are quite adjacent to each other, suggesting a potential benefit of using *CLSTM* for reliable *AQ* forecasting systems. Fig. 9 (b-d) further shows the efficacy of the hybrid *CLSTM* model illustrating the model *MAE* loss and loss comparison in the training and validation phase from January 2015 – June 2018 for the Townsville study site. To summarise, we aver that the newly proposed deep learning *CLSTM* model has performed very well for the study sites *w.r.t* stringent error evaluation such as, but not limited to the Willmott's Index of agreement, Nash–Sutcliffe Efficiency and Legates and McCabe Index. This was supported by a high correlation of forecasted and observed *TSP* as well as the lowest errors. Following our results, we state that the use of a non-*DL* approach (*e.g.*, *RF*, *M5 Tree*, or *MLR*) is likely to result in inferior performance compared with a hybrid *DL* model (*e.g.*, *CLSTM*) especially for data such as *TSP* that can be highly chaotic in their behaviours over short-term (*e.g.*, hourly) scale. Having said that, we note that the only study site where the performance of the *DL* models was not excellent; but still better than its counterpart models, was Hopeland. On closer observation, we see that this site has a high elevation (316m), as shown in Table 1, and while the rest of the pollutants ($PM_{2.5}$ and PM_{10}) are not considered in its model building, greenhouse gases and solar radiation were considered apart from the other model input variables. This is shown in Fig. 3 and Table 3, which might be a factor affecting the performance of the *DL* model. Similarly, the site where the performance of the *DL* model was best, compared to all others, was Brisbane. On closer observation, we note

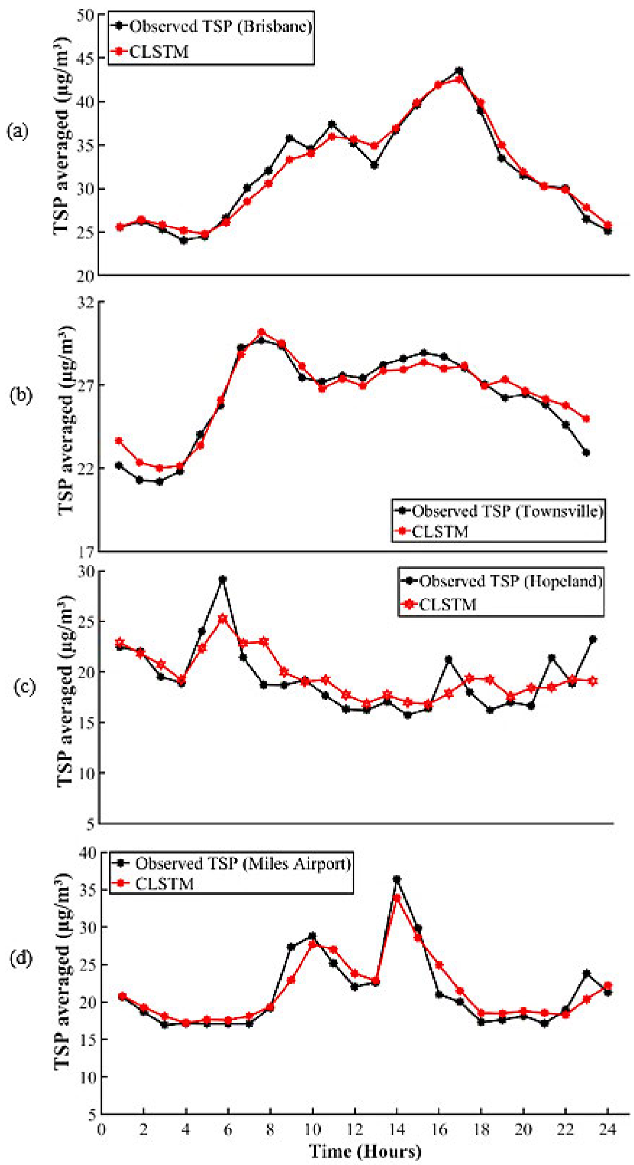


FIGURE 8. **Forecasted *vs.* observed *TSP* in the test phase. (a) Brisbane, (b) Townsville, (c) Hopeland, (d) Miles Airport.**

this site has the lowest elevation (9m; Table 1, Table 3) where none of the gases and solar radiation were considered apart from other selected model input variables. The difference in the availability of model creation data might therefore be a pertinent factor causing a range of differences in the performance of these models.

In summary, the newly proposed deep learning hybrid (*i.e.*, *CLSTM*) model provided a relatively precise performance *i.e.* highest (r), highest (E_{NS}), including a large magnitude of the most stringent model performance metric *i.e.* L, including the lowest *RMSE, MAPE,* and *MAE* compared with the five other benchmark models. Consequently, these results provided compelling evidence, establishing the deep learning hybrid *CLSTM* model as a credible and relatively practical framework for *TSP* predictions. The intelligent framework for modelling air pollutants can also have a major application in the academic and industrial settings, as this particulate matter poses a major health issue. We thus conclude that *CLSTM* can be considered as a novel framework for pollutant modelling, and therefore, can be used as a pragmatic tool in monitoring the atmospheric environment.

V. CONCLUSION

Using deep learning approach, this study reports the potential utility of an air pollution forecasting system developed and evaluated at hourly timesteps. The research work has designed a hybrid predictive model (*i.e.*, *CLSTM*) adopted to forecast the total suspended particulate matter. The newly proposed *CLSTM* model amalgamated convolutional neural network with long short-term memory network to attain optimal performance. For enhanced performance utility, our *CLSTM* firstly employed the *CNN* algorithm to automatically detect and extract important features from predictor variables while *LSTM* collated these features to generate a time series for the next modelling phase. Elucidated by charts, plots, and statistical metrics of forecasted and observed *TSP*, the findings reveal a superior performance of *CLSTM* relative to an ensemble of five competing models. Our findings, including the contributions of this study, are as follows:

i) The paper has bridged significant gaps in modelling hourly *TSP* for an *APF* system, especially for Australia by presenting a constructive research methodology to generate a computationally efficient architecture denoted as a hybrid *'CLSTM'* model.
ii) The ensemble of five other competing models: Random Forest, Volterra, *MLR*, M5 model tree, and *LSTM* appeared to lag in their capability to generate satisfactory forecasts in comparison to the deep learning *CLSTM* hybrid model. Mean absolute and root mean square error for all competing models were significantly higher than the deep learning hybrid model. For argument sake, our simulations registered *MAE* (μg/m^3) = (3.54 against *4.69, 15.51, 7.91, 8.13, 11.66*), and *RMSE* value (μg/m^3) = (13.97 *vs. 23.31, 40.74, 19.24, 20.27, 46.71*) for *CLSTM vs. LSTM*, *RF*, Volterra, *MLR* or M5 model tree algorithms, respectively.
iii) A comprehensive evaluation with statistical metrics, charts, and plots of tested data demonstrated that the hybrid *CLSTM* model was able to yield relatively better forecasts in comparison with the five comparative benchmark models. The performance metrics illustrated the *CLSTM* model to attain efficient performance for short-term *TSP* forecasts. For example, the most rigorous statistical metrics, L, was superior for all four stations with values of (*0.80, 0.51, 0.35, 0.61*) for *CLSTM vs.* (*0.73, 0.45, 0.35, 0.53*) for *LSTM,* (*0.64, 0.44,0.29, 0.46*) for *RF*, (*0.54, 0.42, 0.19, 0.44*) for Volterra, (*0.53, 0.43, 0.27, 0.28*) for *MLR,* and (*0.33, 0.36, 0.20, 0.17*) for M5 model tree. These metrics are providing undisputed evidence that the *CLSTM*

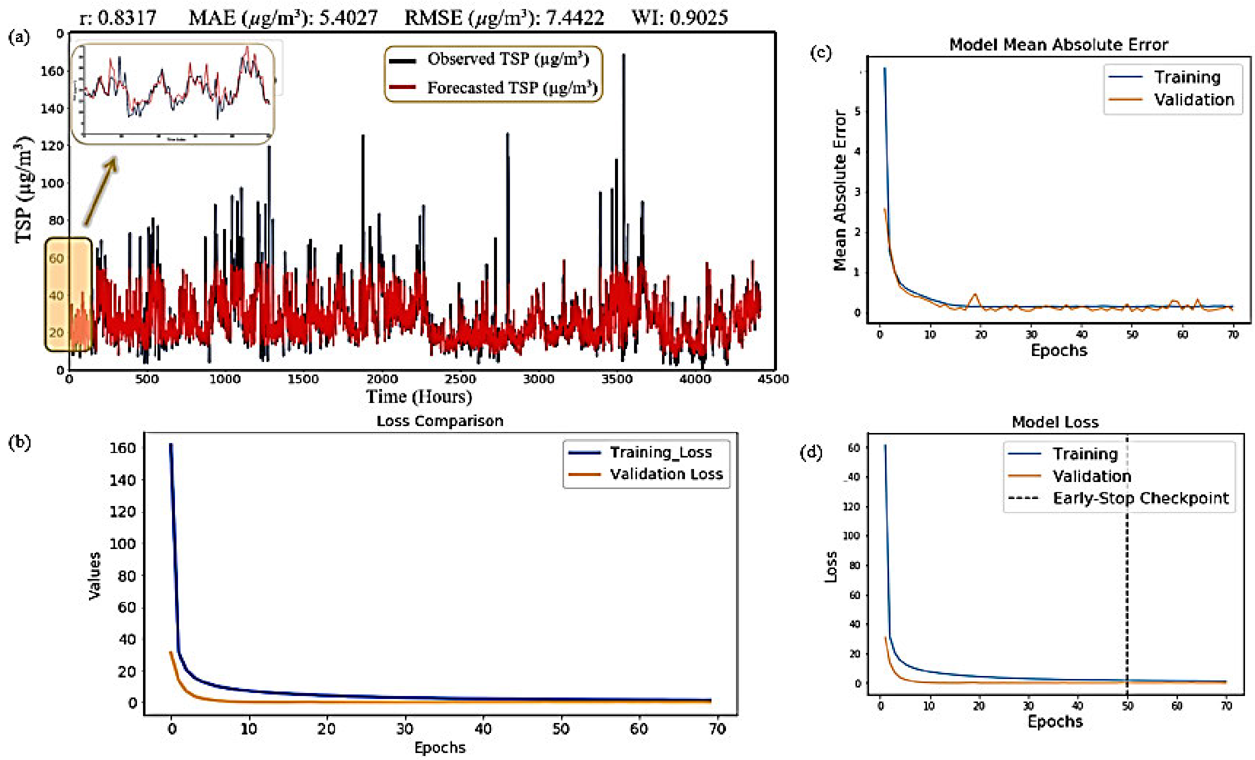


FIGURE 9. (a-d).Performance evaluation with epoch = 2000 at Townsville in the test phase (July-December 2018).

model was very successful in forecasting the hourly air pollutant datasets for all sites.

iv) The percentage error for *CLSTM* also suggested that the fusion of *CNN* and *LSTM* frameworks led to better predictive outcomes, as verified by *MAPE* (*8.43% w.r.t CLSTM vs. 10.20%* (*LSTM*), *22.01%* (*RF*), *25.12%* (Volterra), *27.08%* (*MLR*), and *36.07%* (M5 model tree). In both the training and the testing phase the proposed *CLSTM* hybrid framework was in the 'excellent model category, see Ref [60], as exemplified by *MAE* $< 10\%$ (at all study sites), with low *RMSE* $< 10\%$ for Townsville or $< 15\%$ for Brisbane. This reveals that *CLSTM* architecture was considerably robust in forecasting hourly data, and therefore indicated its possible application in the air pollution model and future public health and air quality studies.

VI. LIMITATIONS AND FUTURE RESEARCH

Despite the efficacy of *CLSTM*, there remain some limitations that can be used to recommend future research.

- Independent testing of the *CLSTM* model at different time steps such as at short-term (*e.g.*, 1–, 5– or 10–minute) resolution may enhance its practical adoption from an end-user perspective. The very short-term forecasting can also lead to a more proactive application of *CLSTM*, in terms of public advisory roles, and especially, the vulnerable populations *e.g.*, children, people with medical issues, senior citizens, and expectant women and planning accordingly.
- In future research, one can improve the *CLSTM* model using ancillary tools such as bidirectional *LSTM* algorithms to generate better outcomes. Also, considering the chaotic nature of air pollutants, the use of an improved complete ensemble empirical mode decomposition with an adaptive noise algorithm to extract temporal information of air pollutant movements could further add value to the proposed *APF* system.
- It is also recommended that future research could incorporate the Internet of Things (*IoT*) to convert the hybrid *CLSTM* model into an application (*app*) for the Internet and mobile phone users. This could be similar to a recently proposed system for forecasting solar ultraviolet (*UV*) radiation tailored to generate very short-term reactive *UV* forecasts for public health risk mitigation [62]. Similarly, an IoT system for air quality forecasting can increase the practicality of *CLSTM*, particularly, ensuring greater public use of the newly proposed *AQ* system.

In closing, the study avers that the deep learning hybrid model: *CLSTM* model carries significant merits, especially in atmospheric modelling effort, assisting in the evaluation of the impact of pollutants on health and environment and contributing to research on air quality and general risk assessments. Consequently, this research work is a paramount step regarding the construction of an effective air pollutant forecasting technology making a vital impact on the environment and health risk alleviation.

ACKNOWLEDGMENT

The authors are thankful to the reviewers for their helpful suggestions and the Queensland Department of Environment and Science for the data.

REFERENCES

[1] P. Gupta, S. A. Christopher, M. A. Box, and G. P. Box, "Multi year satellite remote sensing of particulate matter air quality over Sydney, Australia," *Int. J. Remote Sens.*, vol. 28, no. 20, pp. 4483–4498, Oct. 2007.

[2] M. Oprea, S. F. Mihalache, and M. Popescu, "A comparative study of computational intelligence techniques applied to PM2.5 air pollution forecasting," in *Proc. 6th Int. Conf. Comput. Commun. Control (ICCCC)*, May 2016, pp. 103–108.

[3] X. Xu, H. Ding, and X. Wang, "Acute effects of total suspended particles and sulfur dioxides on preterm delivery: A community-based cohort study," *Arch. Environ. Health, Int. J.*, vol. 50, no. 6, pp. 407–415, Dec. 1995.

[4] E. Sharma, R. C. Deo, R. Prasad, and A. V. Parisi, "A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms," *Sci. Total Environ.*, vol. 709, Mar. 2020, Art. no. 135934.

[5] M. R. Heal, P. Kumar, and R. M. Harrison, "Particles, air quality, policy and health," *Chem. Soc. Rev.*, vol. 41, no. 19, pp. 6606–6630, 2012.

[6] J. Kukkonen *et al.*, "A review of operational, regional-scale, chemical weather forecasting models in Europe," *Atmos. Chem. Phys.*, vol. 12, no. 1, pp. 1–87, Jan. 2012.

[7] K. D. Karatzas and S. Kaltsatos, "Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece," *Simul. Model. Pract. Theory*, vol. 15, no. 10, pp. 1310–1319, Nov. 2007.

[8] S. Pan, Y. Choi, A. Roy, and W. Jeon, "Allocating emissions to 4 km and 1 km horizontal spatial resolutions and its impact on simulated NO_x and O_3 in Houston, TX," *Atmos. Environ.*, vol. 164, pp. 398–415, Sep. 2017.

[9] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113541.

[10] C.-Y. Zhang, C. L. P. Chen, M. Gan, and L. Chen, "Predictive deep Boltzmann machine for multiperiod wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1416–1425, Oct. 2015.

[11] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113541.

[12] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction," *Energies*, vol. 12, no. 12, p. 2407, Jun. 2019.

[13] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, "Enhanced deep learning algorithm development to detect pain intensity from facial expression images," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113305.

[14] H. Al-Hadeethi, S. Abdulla, M. Diykh, R. C. Deo, and J. H. Green, "Adaptive boost LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications," *Expert Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113676.

[15] E. Eslami, Y. Choi, Y. Lops, and A. Sayeed, "A real-time hourly ozone prediction system using deep convolutional neural network," *Neural Comput. Appl.*, vol. 32, no. 13, pp. 8783–8797, Jul. 2020.

[16] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, and T. Chi, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," *Environ. Pollut.*, vol. 231, pp. 997–1004, Dec. 2017.

[17] Q. Zhang, J. C. K. Lam, V. O. K. Li, and Y. Han, "Deep-AIR: A hybrid CNN-LSTM framework forFine-grained air pollution forecast," 2020, *arXiv:2001.11957*. [Online]. Available: http://arxiv.org/abs/2001.11957

[18] T. Li, M. Hua, and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter (PM2.5)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.

[19] G. P. V. Athira, R. Vinayakumar, and K. Soman, "DeepAirNet: Applying recurrent networks for air quality prediction," in *Proc. Int. Conf. Comput. Intell. Data Sci.* vol. 132, 2018, pp. 1394–1403.

[20] J. Chen, W. Quan, T. Cui, and Q. Song, "Estimation of total suspended matter concentration from MODIS data using a neural network model in the China eastern coastal zone," *Estuarine, Coastal Shelf Sci.*, vol. 155, pp. 104–113, Mar. 2015.

[21] L.-S. Wen, C.-P. Lee, W.-H. Lee, and A. Chuang, "An ultra-clean multilayer apparatus for collecting size fractionated marine plankton and suspended particles," *J. Visualized Exp.*, no. 134, p. e56811, Apr. 2018, doi: 10.3791/56811.

[22] Q. Guo, Z. Zhu, Z. Cheng, S. Xu, X. Wang, and Y. Duan, "Correction of light scattering-based total suspended particulate measurements through machine learning," *Atmosphere*, vol. 11, no. 2, p. 139, Jan. 2020.

[23] N. Wang, R. Liu, N. Asmare, C.-H. Chu, and A. F. Sarioglu, "Processing code-multiplexed coulter signals via deep convolutional neural networks," *Lab Chip*, vol. 19, no. 19, pp. 3292–3304, 2019.

[24] G. Liu and S. Shuo, "Air quality forecasting using convolutional LSTM," Stanford Univ., Stanford, CA, USA, Tech. Rep. 8291197, Spring 2018. [Online]. Available: http://cs230.stanford.edu/projects_spring_2018/reports/8291197.pdf

[25] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, "Enhanced deep learning algorithm development to detect pain intensity from facial expression images," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113305.

[26] C. Mathers, T. Vos, and C. Stevenson, "The burden of disease and injury in Australia," *Austral. Health Rev.*, vol. 23, no. 1, p. 216, 2000.

[27] S. Davey and A. Sarre, *The 2019/20 Black Summer Bushfires*. New York, NY, USA: Taylor & Francis, 2020.

[28] P. Yu, R. Xu, M. J. Abramson, S. Li, and Y. Guo, "Bushfires in Australia: A serious health emergency under climate change," *Lancet Planet. Health*, vol. 4, no. 1, pp. e7–e8, Jan. 2020.

[29] M. Hendryx, N. Higginbotham, B. Ewald, and L. H. Connor, "Air quality in association with rural coal mining and combustion in new south Wales Australia," *J. Rural Health*, vol. 35, no. 4, pp. 518–527, Sep. 2019.

[30] R. Prasad, R. C. Deo, Y. Li, and T. Maraseni, "Ensemble committee-based data intelligent approach for generating soil moisture forecasts with multivariate hydro-meteorological predictors," *Soil Tillage Res.*, vol. 181, pp. 63–81, Sep. 2018.

[31] M. Ali, R. Prasad, Y. Xiang, and R. C. Deo, "Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms," *Renew. Sustain. Energy Rev.*, vol. 132, Oct. 2020, Art. no. 110003.

[32] A. Rahimikhoob, M. Asadi, and M. Mashal, "A comparison between conventional and m5 model tree methods for converting pan evaporation to reference evapotranspiration for semi-arid region," *Water Resour. Manage.*, vol. 27, no. 14, pp. 4815–4826, Nov. 2013.

[33] A. Suleiman, M. R. Tight, and A. D. Quinn, "A comparative study of using random forests (RF), extreme learning machine (ELM) and deep learning (DL) algorithms in modelling roadside particulate matter (PM10 & PM2.5)," in *Proc. IOP Conf. Ser., Earth Environ. Sci.*, 2020, Art. no. 012126.

[34] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[35] A. Alléon, G. Jauvion, B. Quennehen, and D. Lissmyr, "PlumeNet: Large-scale air quality forecasting using a convolutional LSTM network," 2020, *arXiv:2006.09204*. [Online]. Available: http://arxiv.org/abs/2006.09204

[36] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities," *Sensors*, vol. 18, no. 7, p. 2220, Jul. 2018.

[37] S. Oehmcke, O. Zielinski, and O. Kramer, "Input quality aware convolutional LSTM networks for virtual marine sensors," *Neurocomputing*, vol. 275, pp. 2603–2615, Jan. 2018.

[38] S. Du, T. Li, Y. Yang, and S.-J. Horng, "Deep air quality forecasting using hybrid deep learning framework," *IEEE Trans. Knowl. Data Eng.*, early access, Nov. 20, 2019, doi: 10.1109/TKDE.2019.2954510.

[39] J. Chen, G.-Q. Zeng, W. Zhou, W. Du, and K.-D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Convers. Manage.*, vol. 165, pp. 681–695, Jun. 2018.

[40] M. Sundermeyer, R. Schlüter, and H. Ney, "LSTM neural networks for language modeling," Tech. Rep., 2012.

[41] S. Han, Y. Wang, H. Yang, W. J. Dally, J. Kang, H. Mao, Y. Hu, X. Li, Y. Li, D. Xie, H. Luo, and S. Yao, "ESE: Efficient speech recognition engine with sparse LSTM on FPGA," in *Proc. ACM/SIGDA Int. Symp. Field-Program. Gate Arrays (FPGA)*, 2017, pp. 75–84.

[42] Y.-T. Tsai, Y.-R. Zeng, and Y.-S. Chang, "Air pollution forecasting using RNN with LSTM," in *Proc. IEEE 16th Int. Conf. Depend., Autonomic Secure Comput., 16th Int. Conf Pervas. Intell. Comput., 4th Int. Conf Big Data Intell. Comput. Cyber Sci. Technol. Congr. (DASC/PiCom/DataCom/CyberSciTech)*, Aug. 2018, pp. 1074–1079.

[43] *Quarterly Update of Australia's National Greenhouse gas Inventory: Mar. 2018, Incorporating Emissions from the NEM up to June 2018*, CWA, Mar. 2018.

[44] *National Pollution Inventory, Department of Agriculture, Water and the Environment, CSIRO Submission no 48 to Senate Community Affairs References Committee, Parliament of Australia, Impacts on Health of Air Quality in Australia*, p. 8. [Online]. Available: http://www.npi.gov.au/resource/particulate-matter-pm10-and-pm25

[45] W. L. Junger and A. P. de Leon, "Imputation of missing data in time series for air pollutants," *Atmos. Environ.*, vol. 102, pp. 96–104, Feb. 2015.
[46] N. Ketkar, "Introduction to Keras," *Deep Learning With Python*. Cham, Switzerland: Springer, 2017, pp. 97–111.
[47] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," Tech. Rep., 2016, pp. 265–283.
[48] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[49] S. Sun and J.-L. Bertrand-Krajewski, "Input variable selection and calibration data selection for storm water quality regression models," *Water Sci. Technol.*, vol. 68, no. 1, pp. 50–58, Jul. 2013.
[50] R. C. Deo, X. Wen, and F. Qi, "A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset," *Appl. Energy*, vol. 168, pp. 568–593, Apr. 2016.
[51] R. C. Deo, M. K. Tiwari, J. F. Adamowski, and J. M. Quilty, "Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model," *Stochastic Environ. Res. Risk Assessment*, vol. 31, no. 5, pp. 1211–1240, Jul. 2017.
[52] R. C. Deo, M. K. Tiwari, J. F. Adamowski, and J. M. Quilty, "Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model," *Stochastic Environ. Res. Risk Assessment*, vol. 31, no. 5, pp. 1211–1240, Jul. 2017.
[53] N. Karunanithi, W. J. Grenney, D. Whitley, and K. Bovee, "Neural networks for river flow prediction," *J. Comput. Civil Eng.*, vol. 8, no. 2, pp. 201–220, Apr. 1994.
[54] M. Ali and R. Prasad, "Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition," *Renew. Sustain. Energy Rev.*, vol. 104, pp. 281–295, Apr. 2019.
[55] R. Yu, J. Gao, M. Yu, W. Lu, T. Xu, M. Zhao, J. Zhang, R. Zhang, and Z. Zhang, "LSTM-EFG for wind power forecasting based on sequential correlation features," *Future Gener. Comput. Syst.*, vol. 93, pp. 33–42, Apr. 2019.
[56] D. Zeng, "Workshop 3: Deep learning with Python," in *Proc. South Dakota State Univ. Symp. Session 8, Workshop 3, United States*, 2019. [Online]. Available: https://openprairie.sdstate.edu/datascience_symposium/2019/sessions/
[57] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—A discussion of principles," *J. Hydrol.*, vol. 10, no. 3, pp. 282–290, Apr. 1970.
[58] D. R. Legates and G. J. McCabe, Jr., "Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, 1999.
[59] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geoscientific Model Develop.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.
[60] K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petković, and S. Ch, "A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation," *Energy Convers. Manage.*, vol. 92, pp. 162–171, Mar. 2015.
[61] D. R. Legates and G. J. McCabe, "A refined index of model performance: A rejoinder," *Int. J. Climatol.*, vol. 33, no. 4, pp. 1053–1056, Mar. 2013.
[62] R. C. Deo, N. Downs, A. V. Parisi, J. F. Adamowski, and J. M. Quilty, "Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle," *Environ. Res.*, vol. 155, pp. 141–166, May 2017.

EKTA SHARMA (Member, IEEE) received the B.Sc. degree in mathematical sciences, the M.Sc. degree in operational research, and the M.Phil. degree from the University of Delhi, India. She is currently pursuing the Ph.D. degree in artificial intelligence with the University of Southern Queensland, Australia. She has served in both academia and industry across varied roles as an Area Manager to a Learning Advisor, Lecturer, and Researcher with universities in Europe, India, and Australia. Her research work has received funding from the Australian Defence Science and Technology Group, The Australian Mathematical Sciences Institute, and the Australian Government. Her research interests include atmospheric sciences, environmental management, data analytics, and deep learning. She is a member of INFORMS and AMSI.

RAVINESH C. DEO (Senior Member, IEEE) is currently an Associate Professor with the University of Southern Queensland, Australia, with research interests in data science, decision systems, and deep learning. He published over 215 articles, including 160 journals (mostly Scopus Quartile 1), seven books under Elsevier, Springer Nature, and IGI Global, and ten book chapters. His work has cumulative citations exceeding 5,400 with H-index of 40. He has supervised over 18 research degrees; he was an Industry Mentor through externally funded grants and a visiting professor internationally and professionally, and is affiliated under scientific bodies. He was awarded internationally competitive fellowships, such as the Queensland *US* Smithsonian, Australia–China Scientist, *JSPS*, Chinese Academy of Science, Australia-India, and Endeavour Fellowship. He serves on Editorial Boards of *Stochastic Environmental Research and Risk Assessment*, *Remote Sensing* and *Energies* journals, including *Journal of Hydrologic Engineering* (ASCE), contributing to teaching, research, and service in data science, mathematics, and modeling.

RAMENDRA PRASAD received the B.Sc. degree in mathematics and physics and the M.Sc. degree in physics from The University of the South Pacific, Fiji, and the Ph.D. degree in modeling and simulations from the University of Southern Queensland, Australia, in 2019. Since 2018, he has been a Lecturer with The University of Fiji. He has authored 15 articles in peer-reviewed journals and reputed conferences. His research interests include advanced machine learning approaches, data analytics, hydrological modeling, energy modeling, energy management, environmental, atmospheric modeling, and ocean wave modeling.

ALFIO V. PARISI the B.Sc. degree (Hons.) in 1981, the M.App.Sc. degree in 1992, and the Ph.D. degree in 1996. He is currently a Professor of Physics with the Faculty of Health, Engineering, and Sciences, University of Southern Queensland, Australia. His research interests include solar UV dosimetry, radiometry, and spectroradiometry, including the development of new techniques on solar measurement, applied to provide improved characterization of solar UV environment, and measurement of solar exposures to plants.

NAWIN RAJ received the B.Sc., B.Ed., PGDMA, and M.Sc. degrees in computational fluid dynamics from The University of the South Pacific and the Ph.D. degree from the University of Southern Queensland (USQ), Australia, in 2015. From 2007 to 2010, he was a Lecturer with Fiji National University. He is currently a Lecturer at USQ. His research interests include artificial intelligence, deep learning, non-linear oscillation, computational fluid dynamics, and oceanography. He is a member of the Australian Mathematical Society and Queensland College of Teachers.

CHAPTER 5 NOVEL HYBRID DEEP LEARNING MODEL FOR SATELLITE-BASED PM10 FORECASTING IN MOST POLLUTED AUSTRALIAN HOTSPOTS

5.1 Foreword

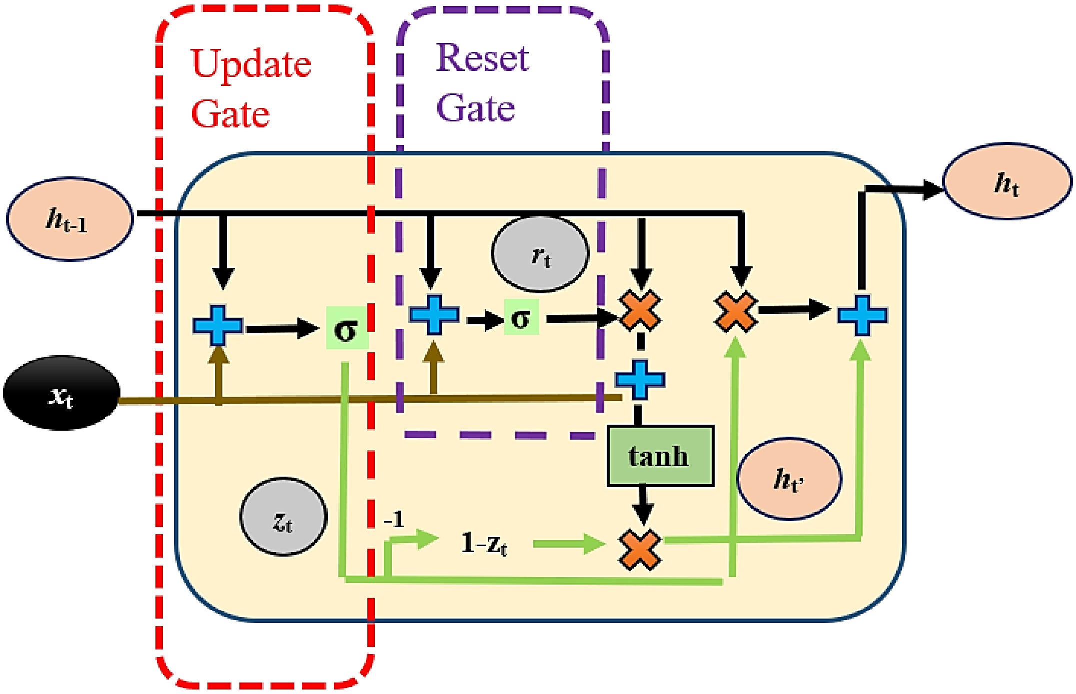
This chapter presents an exact copy of the published article in ***Atmospheric Environment,*** (Vol: 279 (2022): Page(s): 119111-119124).

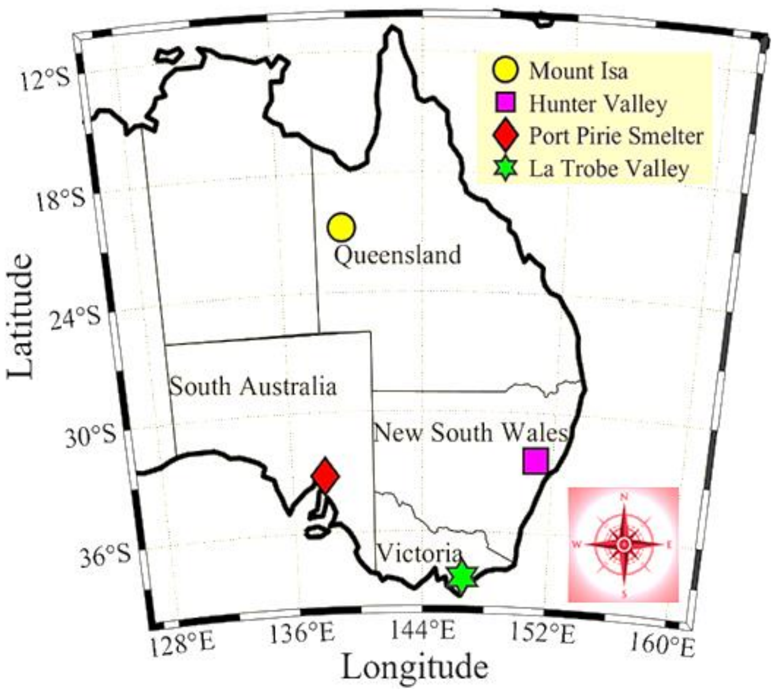
This chapter has built upon earlier chapters 3 and 4 by studying the association of meteorological parameters in forecasting hourly air quality for the most polluted Australian hotspots through hybrid deep learning methods. The early warning tool built for coarse particulates has been done after assessing the impact of the 30 satellite-derived and ground-based environmental pollutants on hourly data from January 2018-December 2020. A one-dimensional convolutional neural network (*CNN*) was integrated with a one-directional fully gated recurrent unit (*GRU*) to forecast the consecutive hour's air quality. The *CNN* model acted as a spatial feature extractor, whereas the new generation *GRU* made it computationally efficient. The hybrid '*CNN-GRU'* was comprehensively benchmarked outperforming an ensemble of six deep learning models. The efficacy of the proposed model was demonstrated in the testing phase at the four most air polluted Australian postcodes. A detailed error analysis with visual and statistical metrics for air quality forecasting ascertains the proposed model's countermeasures to reduce harm and loss. The practical tool can be widely deployed in regions where air pollution poses a significant hazard to public health.

5.2 Research Highlights

- A hybrid *CGRU* model was generated by integrating with *CNN* and *GRU* for air quality forecasting.
- *CNN-GRU* was evaluated against an ensemble of six other competing models.
- A detailed error analysis with visual and statistical metrics for air quality forecasting ascertained the proposed model's countermeasures to reduce harm and loss.

5.3 Graphical Abstract





5.4 Published Article III

This article cannot be displayed due to copyright restrictions. See the article link in the Related Outputs field on the item record for possible access.

CHAPTER 6 CONCLUSIONS AND FUTURE SCOPE

6.1 Synthesis

The findings of this doctoral thesis have advanced new knowledge in artificial intelligence and its applications to the field of atmospheric sciences and acknowledge the vital problem of accurately modeling and forecasting air quality to assist decision-makers. An important contribution to the novel research with new generation early warning artificial intelligence models has been accomplished in the work done. Careful attention has been paid to the choice of models and state-of-the-art procedures for designing final algorithms in all the objectives considered in the study. This makes the resultant hybrids smart, robust, and computationally efficient. The modelling strategy has used advanced machine learning and particularly deep learning approaches for forecasting particulate matter, meteorological parameters, and atmospheric visibility for four important states and regions of Australia.

Utilising the advanced techniques with comprehensive benchmarking, the pollutants were forecasted at an hourly temporal horizon. Given the complexity and 3-dimensional nature of air particles, a diverse range of data intelligence algorithms was utilized to construct a set of hybrid models. Early warning hybrid and novel algorithms developed in this study for Australian air quality are listed as follows-

- *ICEEMDAN-OS-ELM* (Improved complete ensemble empirical mode decomposition (*ICEEMDAN*) was integrated with the online sequential-extreme learning machine (*OS-ELM*) algorithm; wherein *ICEEMDAN* acted as an excellent de-noising algorithm on the chosen noise amplitude, and *OS-ELM* worked as a fast-paced adaptive algorithm that robustly extracted predictive patterns to fine-tune the model generalization to a near-optimal global solution; **objective 1**,
- *CLSTM* (Convolutional neural network (*CNN*) amalgamed with long short-term memory (*LSTM*). Here one of the most popular and effective deep learning algorithms, *CNN* was used to entail a data processer including feature extractors that drew upon statistically significant antecedent lagged predictor variables,

whereas the *LSTM*, another superior algorithm encapsulated a new feature mapping scheme for prediction; **objective 2**,

- *CNN-GRU* (convolutional neural network (*CNN*) was integrated with a one-directional fully gated recurrent unit (*GRU*). Here *CNN* acted as a spatial feature extractor for classification whereas the new generation *GRU* made it computationally efficient; **objective 3.**

The machine learning algorithms and the hybrids that were developed and used as a comparative benchmark in this study are as follows. The main advantage and the reason for using each of them are also stated along with:

- Online sequential-extreme learning machine (*OSELM*) – *OSELM* is one of the popular learning algorithms and has several advantages such as fast learning speed, strong generalization ability, and simple implementation, which make it very suitable for machine learning applications.
- Extreme learning machine (*ELM*) – This machine-learning algorithm is famous for its usage, quicker speed of learning, greater generalisation performance, and appropriateness for several nonlinear kernel functions.
- Random forest (*RF*) – This model can perform both regression and classification tasks. A random forest produces good predictions that can be understood easily. It can handle large datasets efficiently. Moreover, this algorithm provides a higher level of accuracy in predicting outcomes over certain other algorithms such as decision trees.
- M5 model tree - M5 model tree can simulate the phenomena with very high dimensionality and even up to hundreds of attributes. This ability sets M5 apart from regression tree learners at the time (like Multivariate Adaptive Regression Splines (*MARS)*), whose computational cost grows very quickly when the number of features increases.
- Multiple linear regression (*MLR*) - *MLR* algorithm allows the program to account for all the potentially important factors in one model. The advantages of this approach are that this may lead to a more accurate and precise understanding of the association of each factor with the outcome.

- Volterra (*VOL*) - Its main advantage lies in its generality. This means it can represent a wide range of systems. Thus, it is sometimes considered a non-parametric algorithm.

- ICEEMDAN-MLR – The hybrid model captures *IMFs* with less noise and more meaning through *ICEEMDAN* and leads to a more precise understanding *via* the *MLR* model.

- ICEEMDAN-M5 model tree- This hybrid captures high dimensionality through the M5 model tree and extracts several *IMFs* from a single channel through the *ICEEMDAN* algorithm.

The deep learning algorithms and the hybrids (apart from optimal ones) that were developed and used as a comparative benchmark in the study are:

- Long-Short Term Memory (*LSTM*) – One of the most popular deep learning algorithms, LSTM is well-suited for classification, processing, and predicting time series given time lags of unknown duration. Relative insensitivity to gap length provides them with a major advantage over other algorithms.

- Convolutional Neural Networks (*CNN*) - The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision.

- Gated Recurrent Unit (*GRU*) – A type of Recurrent Neural Network (RNN) who have advantages over *LSTMs* in certain cases as they use less memory and are very faster than LSTMs. However, LSTM is preferred when datasets with longer sequences exist, and more accuracy is needed.

- Bidirectional LSTM (*BiLSTM*) – An *RNN* variant, that acts as a discriminative classifier that models the decision boundary between different classes. Therefore, they learn weights associated with data points and perform better in certain cases.

The efficacy of these proposed models was demonstrated in the testing phase at the carefully chosen postcodes of Australia. A detailed error analysis with visual and statistical metrics for air quality forecasting ascertained the objectives' countermeasures to reduce harm and loss. The practical tools designed in the thesis can be widely deployed in regions where air pollution poses a significant hazard to public health. Subsection 6.2 provides challenges in the route of air quality and issues associated with machine learning and deep learning. The important results with novel contributions of this study are summarised in subsection 6.3. Finally, potentially interesting paths for future research are suggested at the end of this chapter.

6.2 Challenges faced in air quality forecasting

Air quality forecasting is a recent development, with most programs initiated only in the last 25 years (Ryan 2016). However, our published literature does contain a discussion section where the difficulties and limitations experienced during the research implementations are discussed in more detail. This section summarises issues encountered while completing the objectives and challenges in both developments and applications of air forecasting models particularly deep learning technology, that could potentially form interesting subjects for further investigation to address these limitations in the future-

- *Complexity*: Some air quality models use complex mathematical techniques to simulate the physical and chemical processes that affect air pollutants as they disperse and react in the atmosphere.

  Recommendation: Artificial intelligence models assist in easier and faster modelling algorithms both through machine learning and deep learning models.

- *Real-time satellite data availability:* There is a need for real-time and preciseness in forecasting air quality models. This warrants robust methodologies for predicting air quality, largely over short-term and real-time scales so the forecasts can be employed effectively by the public for health risk advisory and air pollution prevention.

  Recommendation: Satellite data can be of immense help here, warranting more effective forecasts.

- *The cost involved in downloading data:* The cost involved in subscribing to the available satellite data even with a student subscription is a discouraging factor in the present times. This might be due to the dearth of real-time data availability for satellite extracted air quality variables.

  Recommendation: A nominal or waved-off cost can make future research much more precise and reliable.

- *Underlying uncertainties:* As air quality forecasts must diagnose and predict several pollutants and their precursors in addition to standard meteorological variables, they are, compared with weather forecasts, a higher uncertainty forecast. These uncertainties, or assumptions in the model when estimating the parameters could influence the result. They also include the potential sources of error that are derived from data management, estimating methods, and model structures generated by a purely statistical approach (*i.e.*, goodness-of-fit tests). In addition, choosing the best parameter combination may lead to an underestimation of the uncertainties of the entire model.

  Recommendation: Based on the research findings, practices to facilitate the incorporation of uncertain information should be suggested.

- *Limited length of the data*: Deep learning currently lacks a mechanism for learning abstractions through explicit definitions and works best when we have enough quality training data (Marcus 2018). This factor can plausibly affect the accuracy of the model design by increasing the uncertainties.

  Recommendation: The techniques of data augmentation and synthetic data can come to the rescue here and they do not have the problem of a copyright too.

- *Minimal implementation of data:* Traditionally, the concentrations of air pollutants are measured using air quality monitoring (*AQM*) stations that are highly reliable, and precise, and can measure a wide spectrum of pollutants using standardized analysers. However, these stations have three main drawbacks the significant infrastructure needed for installation due to their bulky size, the complicated operational requirements, e.g., access to grid power, heating/cooling, and secure shelters, and the prohibitive costs of acquiring, setting up, and performing regular maintenance and calibration. These drawbacks reduce the number of installations and result in sparsely distributed *AQM* networks with limited spatial resolution air pollution data.

  Recommendation: An integrated air quality monitoring network may be established that can provide air quality data for various purposes, e.g., to monitor regulatory compliance and to assess the effectiveness of control strategies. This may assist in removing the limitation of sparse *AQM* networks.

Overcoming such challenges and advancing in the construction of intelligent algorithms will be a promising direction for future research.

6.3 Summary of important findings and novelty of the study

This study focuses majorly on the application of deep learning technologies in forecasting air quality. Currently, deep learning is the spearhead of artificial intelligence and perhaps one of the most exciting technologies of the decade. Putting aside future scopes of research, the deep learning research carried out in all the articles has significant merits in ongoing air quality research and health risk mitigation. This is a key step toward ensuring progress and well-being in any nation's meteorology. To summarise, the following are the value propositions of the study-

- This research initially proposed a novel early warning (hourly-step) AI model (denoted as the *ICEEMDAN-OS-ELM*) hybrid system for air quality prediction in several cities and regional locations in Australia.

- There was a comprehensive consideration of the importance and synergistic effects of atmospheric variables to establish a hybrid model for hourly forecasting, to emulate time-series of data on particulate matter 2.5, 10 and atmospheric visibility reducing particles.

- Considering the high predictive utility with an enhanced model testing accuracy of the proposed models, the machine learning models robustly extracted predictive variables necessary to model the air quality target variables. The findings, elucidated by performance metrics and visual analysis of forecasted and observed air quality, clearly captured the improved performance of the early warning *AI* model developed as compared to the other AI models.

- In further articles, the research involved an ensemble of the latest neural network techniques such as the deep learning (*DL*) framework which are known to generate a superior accuracy. These models use multiple feature extraction layers and learn the complex relationships within the data more efficiently.

- To increase testing performance and accuracy, in the second article, a hybrid *DL* predictive algorithm *CLSTM* was developed. The hybrid was produced by integrating the Convolutional Neural Networks (*CNN*) algorithm with Long Short-Term Memory Networks (*LSTM*). The former automatically detect and extracted important features from predictor variables while the latter collated these features to generate a time series for the next modelling phase. The paper bridged the gap in modelling hourly total suspended particles for an air pollution forecasting system, especially for Australia, presenting constructive research.

- In the third article, the paper reported on research to model an early warning tool for coarse particulates when assessing the impact of the 30 satellite-derived and ground-based meteorological pollutants using hourly Australian data.
- A one-dimensional convolutional neural network (*CNN*) was integrated with a one-directional fully gated recurrent unit (*GRU*) to forecast consecutive hours' air quality. The *CNN* model acts as a spatial feature extractor, whereas the new generation *GRU* makes it computationally efficient. The resultant hybrid '*CNN-GRU'* is then comprehensively benchmarked outperforming an ensemble of six other deep learning models.
- The research has also attempted to bridge the gap in modeling hourly coarse particulate matter with a diameter between 2.5 and 10 μm and evaluated the potential impact of ground and important satellite-derived predictors, where PM_{10} levels are high in Australia.
- Appendix Article A paper reported on the research that does a systematic review of the ongoing *COVID-19* pandemic with deep learning and machine learning algorithms, with particular emphasis on air quality in continents of Asia and Oceania. The inter-continent in-depth research based on selective demography carries importance to determine whether the spread of the virus is correlated to atmospheric and meteorological factors.
- Appendix Article B paper harnessed the robustness and adaptivity of deep learning (*DL*) long short-term memory (*LSTM*) architecture to design error-correction decoding algorithms that scale with high-dimensional codes. The paper proposed computationally efficient, binary classification-based, linear architecture '*LSTM*' to assist in solving dimensionality issues. *LSTM* encapsulated a feature mapping scheme to reduce the complexity, number of errors, and computational time.

- The Deep Learning (*DL*) and machine learning (*ML*) methods are proving to be powerful and hence an increasingly popular choice in forecasting air quality, Overall, deep learning technology in the research conducted has proven to be a potential tool to be tested for air quality predictions given the complexity of air quality variables.
- A detailed error analysis in all the studies conducted with visual and statistical metrics for air quality forecasting ascertains the proposed model's countermeasure to reduce harm and loss.
- The comprehensive analysis of systemic weather risk and potential adaptation strategies are expected to support the development of atmospheric models. Therefore, the knowledge achieved from this study will be important for better understanding the influences of the weather on air pollutants, particularly in Australia.
- The novel methods and the practical tools evolved from this study are immensely beneficial and can be widely deployed to the regions of public health concern where air pollution is a significant hazard.
- The tools can also find an important place in health policy or Government policy designing and may be extended to other regions.

Summarising the above points, the novelty of this research is to propose an innovative, robust, self-adaptive, and computationally efficient artificial intelligence-based modelling framework that can be helpful in air quality regulatory planning, especially at an hourly horizon. This research employs intelligent, versatile, and powerful tools, adopted to decompose original air quality data in high and low-frequency patterns. The use of data sub-series, which reveal better clarity of the patterns embedded in air quality predictor variables increases the computational efficacy, as such methods can address issues of non-stationary, repeats, periodic behaviours, and jump type perturbations before such data are utilised. Therefore, the air regulatory plans established through an hourly-based forecast model have potential benefits for human health, especially for the vulnerable factions of the society such as the elderly, children, pregnant women, and people with health problems.

6.4 Recommendations for the future works

The air quality framework based on machine learning and deep learning has been successfully developed in this study for investigating the influence of atmospheric conditions on air quality with special reference to case studies in Australia. However, careful attention needs to be provided to address the limitations, despite the success of the hybrid models developed in this study. Therefore, the following pointers can be a promising step for future research to enhance the understanding of atmospheric and meteorological impact as well as the extent of the impact of the deep learning and feature extraction methods.

- *Testing at several timescales*- The models have been tested at hourly temporal domains. Additionally, the proposed model can be tested at various other temporal horizons such as a real-time (minute), 5-minute, or 10-minute scale. This will be beneficial for the provision of proactive advice at various forecasting horizons however the provision of high-resolution data will affect the success of a real-time system.

- *Better spatially resolved data sets*- A range of long-term and short-term challenges to atmospheric composition such as Volcanic ash, desert dust, natural and anthropogenic gas emissions on aviation and human health with changing atmospheric composition requires considerable additional demands of data timeliness and temporal and spatial resolutions. This near-real-time need for resolutions and observations is indeed a common and an important requirement across a range of impacts, but one that is not always conducive to the significant processing involved in producing fully assured atmospheric composition data.

- *Development of Hybrid model with Bayesian approach*- Construction of an air quality forecast hybrid model with Bayesian Model average (*BMA*) can assist in developing more reliable, and robust (probabilistic) predictions. This can be possible from uncertainty evaluations, ranked models, and multiple competing predictions to further improve the credibility of the hybrid model.

- *Involving Explainable AI (XAI)*- Rapidly evolving state-of-the-art technology of *XAI* produces more explainable models while maintaining a high level of learning performance/prediction accuracy. The resultant framework and tools may further assist in understanding and interpreting predictions made by the deep learning models developed. Such an empirical study and benchmark framework will also help humans understand, trust, and effectively manage the results of artificial intelligence technology.

LIST OF REFERENCES

Note that the references presented here are for Chapters 1, 2, and 6. They do not include the references from the published articles (Chapters 3 to 5). These references are provided in the reference sections of the respective articles.

AAS 2021, Australian Academy of Science, *https://www.science.org.au/curious/people-medicine/air-pollution-whats-situation*

Abadi, M, Barham, P, Chen, J, Chen, Z, Davis, A, Dean, J, Devin, M, Ghemawat, S, Irving, G & Isard, M 2016, 'Tensorflow: a system for large-scale machine learning', *OSDI*, pp. 265-83.

ABC 2018, Australian Broadcasting Corporation, *https://www.abc.net.au/news/2018-11-16/australian-pollution-mapped-by-postcodes/10478620?nw=0*

Ahmed, A, Deo, RC, Raj, N, Ghahramani, A, Feng, Q, Yin, Z & Yang, L 2021, 'Deep Learning Forecasts of Soil Moisture: Convolutional Neural Network and Gated Recurrent Unit Models Coupled with Satellite-Derived MODIS, Observations and Synoptic-Scale Climate Index Data', *Remote Sensing*, vol. 13, no. 4, p. 554.

Al-Hadeethi, H, Abdulla, S, Diykh, M, Deo, RC & Green, JH 2020, 'Adaptive boost LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications', *Expert Systems with Applications*, p. 113676.

Andrews, SQ 2008, 'Inconsistencies in air quality metrics:‘Blue Sky’days and PM10 concentrations in Beijing', *Environmental Research Letters*, vol. 3, no. 3, p. 034009.

Baghurst, PA, Tong, S, Sawyer, MG, Burns, J & McMichael, AJ 1999, 'Sociodemographic and behavioural determinants of blood lead concentrations in children aged 11-13 years: The Port Pirie Cohort Study', *Medical Journal of Australia*, vol. 170, no. 2, pp. 63-7.

Bargshady, G, Zhou, X, Deo, RC, Soar, J, Whittaker, F & Wang, H 2020, 'Enhanced deep learning algorithm development to detect pain intensity from facial expression images', *Expert Systems with Applications*, vol. 149.

Bateson, TF & Schwartz, J 2004, 'Who is sensitive to the effects of particulate air pollution on mortality?: a case-crossover analysis of effect modifiers', *Epidemiology*, vol. 15, no. 2, pp. 143-9.

Broome, RA, Powell, J, Cope, ME & Morgan, GG 2020, 'The mortality effect of PM2. 5 sources in the Greater Metropolitan Region of Sydney, Australia', *Environment international*, vol. 137, p. 105429.

Broome, RA, Fann, N, Cristina, TJN, Fulcher, C, Duc, H & Morgan, GG 2015, 'The health benefits of reducing air pollution in Sydney, Australia', *Environmental Research*, vol. 143, pp. 19-25.

Byun, D & Schere, KL 2006, 'Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system'.

Chan, Y, Simpson, R, Mctainsh, GH, Vowles, P, Cohen, D & Bailey, G 1999, 'Source apportionment of PM2. 5 and PM10 aerosols in Brisbane (Australia) by receptor modelling', *Atmospheric Environment*, vol. 33, no. 19, pp. 3251-68.

Chaulya, S 2004, 'Assessment and management of air quality for an opencast coal mining area', *Journal of Environmental Management*, vol. 70, no. 1, pp. 1-14.

CWA 2018, 'Quarterly Update of Australia’s National Greenhouse Gas Inventory: March 2018 Incorporating emissions from the NEM up to June 2018'.

Davey, S & Sarre, A 2020, *the 2019/20 Black Summer bushfires*, Taylor & Francis, 0004-9158.

Dean, A & Green, D 2017, 'Climate Change, Air Pollution and Health in Australia'.

Deo, RC & Şahin, M 2017, 'Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland', *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 828-48.

DERM 2011, 'Department of Environment and Resource Management '.

DoES 2021, *Queensland Government Department of Environment and Science, https://apps.des.qld.gov.au/air-quality/download/*.

Dominici, F, Peng, RD, Bell, ML, Pham, L, McDermott, A, Zeger, SL & Samet, JM 2006, 'Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases', *Jama*, vol. 295, no. 10, pp. 1127-34.

EJA 2021, *Environmental Justice Australia, https://www.envirojustice.org.au/air-pollution-scorecard/*

EPA-SA 2021, *Environment Protection Authority South Australia, https://data.sa.gov.au/data/dataset/air-quality-monitoring-sites*.

EPA-VIC 2016, 'Final report - Future air quality in Victoria', *Environment Protection Authority, Victoria, https://www.epa.vic.gov.au/for-community/environmental-information/air-quality/clean-air-future-air-quality/future-air-quality-in-victoria*

EPA-VIC 2021, *Environment Protection Authority Victoria, https://www.epa.vic.gov.au/EPAAirWatch*.

Ghimire, S, Deo, RC, Raj, N & Mi, J 2019a, 'Deep Learning Neural Networks Trained with MODIS Satellite-Derived Predictors for Long-Term Global Solar Radiation Prediction', *Energies*, vol. 12, no. 12.

Ghimire, S, Deo, RC, Raj, N & Mi, J 2019a, 'Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms', *Applied Energy*, vol. 253, p. 113541.

Ghimire, S, Deo, RC, Raj, N & Mi, J 2019b, 'Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction', *Energies*, vol. 12, no. 12, p. 2407.

Ghimire, S, Deo, RC, Raj, N & Mi, J 2019b, 'Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms', *Applied Energy*, vol. 253.

Glinianaia, SV, Rankin, J, Bell, R, Pless-Mulloli, T & Howel, D 2004, 'Particulate air pollution and fetal health: a systematic review of the epidemiologic evidence', *Epidemiology*, vol. 15, no. 1, pp. 36-45.

Graham, AM, Pringle, KJ, Pope, RJ, Arnold, SR, Conibear, LA, Burns, H, Rigby, R, Borchers-Arriagada, N, Butt, EW & Kiely, L 2021, 'Impact of the 2019/2020 Australian megafires on air quality and health', *GeoHealth*, vol. 5, no. 10, p. e2021GH000454.

Gupta, P, Christopher, SA, Box, MA & Box, GP 2007, 'Multi year satellite remote sensing of particulate matter air quality over Sydney, Australia', *International Journal of Remote Sensing*, vol. 28, no. 20, pp. 4483-98.

Hendryx, M, Islam, MS, Dong, G-H & Paul, G 2020, 'Air pollution emissions 2008–2018 from australian coal mining: implications for public and occupational health', *International journal of environmental research and public health*, vol. 17, no. 5, p.

1570.

Higginbotham, N, Freeman, S, Connor, L & Albrecht, G 2010, 'Environmental injustice and air pollution in coal affected communities, Hunter Valley, Australia', *Health & Place*, vol. 16, no. 2, pp. 259-66.

Hsu, C-W, Chang, C-C & Lin, C-J 2003, *A practical guide to support vector classification*, Taipei.

HVRA 2009, 'Hunter Valley Research Foundation'.

Hyslop, NP 2009, 'Impaired visibility: the air pollution people see', *Atmospheric Environment*, vol. 43, no. 1, pp. 182-95.

Karatzas, KD & Kaltsatos, S 2007, 'Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece', *Simulation Modelling Practice and Theory*, vol. 15, no. 10, pp. 1310-9.

Ketkar, N 2017, 'Introduction to keras', in *Deep Learning with Python*, Springer, pp. 97-111.

Kim, K-H, Kabir, E & Kabir, S 2015, 'A review on the human health impact of airborne particulate matter', *Environment international*, vol. 74, pp. 136-43.

Kisi, O, Parmar, KS, Soni, K & Demir, V 2017, 'Modeling of air pollutants using least square support vector regression, multivariate adaptive regression spline, and M5 model tree models', *Air Quality, Atmosphere & Health*, vol. 10, no. 7, pp. 873-83.

Krzyzanowski, M, Bundeshaus, G, Negru, ML & Salvi, MC 2005, 'Particulate matter air pollution: how it harms health', *World Health Organization, Fact sheet EURO/04/05, Berlin, Copenhagen, Rome*, vol. 4, p. 14.

Kukkonen, J, Olsson, T, Schultz, D, Baklanov, A, Klein, T, Miranda, A, Monteiro, A, Hirtl, M, Tarvainen, V & Boy, M 2012, 'A review of operational, regional-scale, chemical weather forecasting models in Europe', *Atmospheric Chemistry and Physics*, vol. 12, no. 1, pp. 1-87.

Lan, Y, Soh, YC & Huang, G-B 2009, 'Ensemble of online sequential extreme learning machine', *Neurocomputing*, vol. 72, no. 13-15, pp. 3391-5.

Langrish, JP, Li, X, Wang, S, Lee, MM, Barnes, GD, Miller, MR, Cassee, FR, Boon, NA, Donaldson, K, Li, J, Li, L, Mills, NL, Newby, DE & Jiang, L 2012, 'Reducing personal exposure to particulate air pollution improves cardiovascular health in patients with coronary heart disease', *Environ Health Perspect*, vol. 120, no. 3, pp. 367-72.

Lee, C-K & Lin, S-C 2008, 'Chaos in air pollutant concentration (APC) time series', *Aerosol and Air Quality Research*, vol. 8, no. 4, pp. 381-91.

Leitte, AM, Schlink, U, Herbarth, O, Wiedensohler, A, Pan, XC, Hu, M, Richter, M, Wehner, B, Tuch, T, Wu, Z, Yang, M, Liu, L, Breitner, S, Cyrys, J, Peters, A, Wichmann, HE & Franck, U 2011, 'Size-segregated particle number concentrations and respiratory emergency room visits in Beijing, China', *Environ Health Perspect*, vol. 119, no. 4, pp. 508-13.

Li, X & Li, C 2016, 'Improved CEEMDAN and PSO-SVR Modeling for Near-Infrared Noninvasive Glucose Detection', *Computational and mathematical methods in medicine*, vol. 2016.

Liang, N-Y, Huang, G-B, Saratchandran, P & Sundararajan, N 2006, 'A fast and accurate online sequential learning algorithm for feedforward networks', *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411-23.

Lin, Y, Mago, N, Gao, Y, Li, Y, Chiang, Y-Y, Shahabi, C & Ambite, JL 2018, 'Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning', *Proceedings of the 26th ACM SIGSPATIAL international conference on advances in geographic information systems*, pp. 359-68.

Mannes, T, Jalaludin, B, Morgan, G, Lincoln, D, Sheppeard, V & Corbett, S 2005, 'Impact of ambient air pollution on birth weight in Sydney, Australia', *Occupational and Environmental Medicine*, vol. 62, no. 8, pp. 524-30.

Marcus, G 2018, 'Deep learning: A critical appraisal', *arXiv preprint arXiv:1801.00631*.

Morgan, G, Sheppeard, V, Khalaj, B, Ayyar, A, Lincoln, D, Jalaludin, B, Beard, J, Corbett, S & Lumley, T 2010, 'Effects of bushfire smoke on daily mortality and hospital admissions in Sydney, Australia', *Epidemiology*, pp. 47-55.

NASA-GIOVANNI 2021, *The National Aeronautics and Space Administration, Goddard Online Interactive Visualization and Analysis Infra-structure, https://earthdata.nasa.gov/.*

NPE 2016, 'National Pollutant Energy, Department of the Environment and Energy', vol. 1.

NPI 2021, *National Pollutant Inventory, http://www.npi.gov.au/*

NSW 2021a, *New South Wales Health, https://www.health.nsw.gov.au/environment/air/Pages/particulate-matter.aspx*.

NSW 2021b, *New South Wales Government Department of Planning, Industry, and Environment, https://www.dpie.nsw.gov.au/air-quality/air-quality-data-services*.

NWSN 2021, *The North West Star News, https://www.northweststar.com.au/story/5766470/stark-pollution-report-slammed-by-representatives/*

O’Neill, MS, Zanobetti, A & Schwartz, J 2003, 'Modifiers of the temperature and mortality association in seven US cities', *American journal of epidemiology*, vol. 157, no. 12, pp. 1074-82.

Pan, S, Choi, Y, Roy, A & Jeon, W 2017, 'Allocating emissions to 4 km and 1 km horizontal spatial resolutions and its impact on simulated NOx and O3 in Houston, TX', *Atmospheric Environment*, vol. 164, pp. 398-415.

Pedregosa, F, Varoquaux, G, Gramfort, A, Michel, V, Thirion, B, Grisel, O, Blondel, M, Prettenhofer, P, Weiss, R & Dubourg, V 2011, 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825-30.

Price, OF, Williamson, GJ, Henderson, SB, Johnston, F & Bowman, DM 2012, 'The relationship between particulate pollution levels in Australian cities, meteorology, and landscape fire activity detected from MODIS hotspots', *PloS one*, vol. 7, no. 10, p. e47327.

Qld 2021, *Queensland Government*, *https://www.qld.gov.au/environment/pollution/monitoring/air/air-pollution/pollutants/particles*.

Quilty, J & Adamowski, J 2018, 'Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework', *Journal of Hydrology*.

Rajarethinam, J, Aik, J & Tian, J 2020, 'The Influence of South East Asia Forest Fires on Ambient Particulate Matter Concentrations in Singapore: An Ecological Study Using Random Forest and Vector Autoregressive Models', *International journal of environmental research and public health*, vol. 17, no. 24, p. 9345.

Reznikov, N, Buss, DJ, Provencher, B, McKee, MD & Piché, N 2020, 'Deep learning for 3D imaging and image analysis in biomineralization research', *Journal of*

Structural Biology, vol. 212, no. 1, p. 107598.

Ryan, RG, Silver, JD & Schofield, R 2021, 'Air quality and health impact of 2019–20 Black Summer megafires and COVID-19 lockdown in Melbourne and Sydney, Australia', *Environmental pollution*, vol. 274, p. 116498.

Ryan, WF 2016, 'The air quality forecast rote: Recent changes and future challenges', *Journal of the Air & Waste Management Association*, vol. 66, no. 6, pp. 576-96.

Sanner, MF 1999, 'Python: a programming language for software integration and development', *J Mol Graph Model*, vol. 17, no. 1, pp. 57-61.

Saxe, AM, McClelland, JL & Ganguli, S 2013, 'Exact solutions to the nonlinear dynamics of learning in deep linear neural networks', *arXiv preprint arXiv:1312.6120*.

Shaban, KB, Kadri, A & Rezk, E 2016, 'Urban air pollution monitoring system with forecasting models', *IEEE Sensors Journal*, vol. 16, no. 8, pp. 2598-606.

Sharma, E, Deo, RC, Prasad, R & Parisi, AV 2020, 'A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms', *Science of The Total Environment*, vol. 709, p. 135934.

Squizzato, S, Cazzaro, M, Innocente, E, Visin, F, Hopke, PK & Rampazzo, G 2017, 'Urban air quality in a mid-size city—PM2. 5 composition, sources and identification of impact areas: From local to long range contributions', *Atmospheric Research*, vol. 186, pp. 51-62.

Tao, Q, Liu, F, Li, Y & Sidorov, D 2019, 'Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU', *IEEE Access*, vol. 7, pp. 76690-8.

Taylor, MP & Isley, CF 2014, 'Measuring, monitoring and reporting but not intervening: Air Quality in Australian Mining and Smelting Areas', *Air Quality and Climate Change*, vol. 48, no. 2, p. 35.

Taylor, MP, Isley, CF & Glover, J 2019, 'Prevalence of childhood lead poisoning and respiratory disease associated with lead smelter emissions', *Environment international*, vol. 127, pp. 340-52.

Tiwari, MK & Adamowski, J 2013, 'Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models', *Water Resources Research*, vol. 49, no. 10, pp. 6486-507.

Valavanidis, A, Fiotakis, K & Vlachogianni, T 2008, 'Airborne particulate matter and human health: toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms', *Journal of Environmental Science and Health, Part C*, vol. 26, no. 4, pp. 339-62.

Walsh, MP 2014, 'PM2.5: global progress in controlling the motor vehicle contribution', *Frontiers of Environmental Science & Engineering*, vol. 8, no. 1, pp. 1-17.

Xiao, Y-j, Wang, X-k, Wang, J-q & Zhang, H-y 2021, 'An adaptive decomposition and ensemble model for short-term air pollutant concentration forecast using ICEEMDAN-ICA', *Technological Forecasting and Social Change*, vol. 166, p. 120655.

Zhang, C-Y, Chen, CP, Gan, M & Chen, L 2015, 'Predictive deep Boltzmann machine for multiperiod wind speed forecasting', *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1416-25.

Zhou, Q, Jiang, H, Wang, J & Zhou, J 2014, 'A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network', *Science of The Total Environment*, vol. 496, pp. 264-74.

Zhou, Y, Fu, JS, Zhuang, G & Levy, JI 2010, 'Risk-based prioritization among air pollution control strategies in the Yangtze River Delta, China', *Environ Health Perspect*, vol. 118, no. 9, pp. 1204-10.

Zhu, Y, Liang, Y & Chen, SX 2021, 'Assessing local emission for air pollution via data experiments', *Atmospheric Environment*, vol. 252, p. 118323.

ADDITIONAL ARTICLE I (APPENDIX A) ASSOCIATION OF AIR POLLUTANTS AND NOVEL CORONAVIRUS WITH DEEP LEARNING: A SYSTEMATIC LITERATURE REVIEW OF THE CASE STUDIES IN ASIA AND OCEANIA

Sharma, E, Deo, RC, Soar J, & Yaseen, ZM 2022, 'Association of Air Pollutants and Novel Coronavirus with Deep Learning: A systematic literature review of the case studies in Asia and Oceania', *Sustainable Cities and Society,* Under-review. (*Q1, Impact Factor: 7.59, H Index: 61, SNIP*: *2.35,* and *97th percentile in Civil and Structural Engineering*).

A.1 Foreword

Appendix A presents an exact copy of the article Under-review in ***Sustainable Cities and Society.***

In this work, the candidate has focussed on the severe acute respiratory syndrome coronavirus-2 (*SARS-CoV-2*) that has caused approximately 2 billion confirmed coronavirus disease (*COVID 19*) cases and 4.25 million deaths globally, putting worldwide healthcare systems into crisis. The Deep Learning (*DL*) and machine learning (*ML*) methods are powerful and hence an increasingly popular choice in the challenge to overcome the *COVID-19* pandemic. A novel and an extensive literature review about Asia and Oceania have been conducted to integrate the knowledge and correlation between *COVID-19* and air pollutants. The inter-continent, in-depth research based on selective demography carries importance to determine whether the virus spread is correlated to atmospheric and meteorological factors. From an initial set of 545 articles, a total of 33 articles were finally selected through rigorous rounds.

The study shortlisted the countries in Asia (China, India, Bangladesh, Jordan, Jakarta, Korea) and Oceania (Australia and New Zealand) that published the literature relevant to the criteria considered. It was observed that most of the research belonged to China, the epicenter of the *COVID-19* and most researchers in Asia and Oceania focused on particulate matter *PM* 2.5, followed by meteorology and *PM*10. Random Factor seems to be the most preferred choice in *ML* for modeling and long short-term memory is the most preferred *DL* choice. The study in the end has summed up four future research opportunities in this review.

A.2 Research Highlights

- This systematic review shows that most research papers focused on forecasting or predicting the number of air pollutants.
- All these articles focused on forecasting through ML and DL techniques and prioritized accuracy over interpretability.
- The analysis reveals that the extent of the forecasting accuracy is lesser as compared to that of the models that used estimation.
- The 3-d chaotic nature of air pollutants may demand the application and better suited computationally efficient algorithms such as machine learning or deep learning.
- The preciseness of ML for particulate matter prediction as well as forecasting got to maximum numbers in comparison with the other pollutants. In general, it was observed that the preciseness of peaks with a higher rate of pollution was lesser as compared to the low or medium pollution peaks.
- In the end, from the analysis of the selection, it was noteworthy that harmful and increased contamination concentration of particulate in the southern hemisphere and the Asia pacific is understudied when combined with deep learning or machine learning.

A.3 Under Review Appendix Article I

Association of Air Pollutants and Novel Coronavirus with Deep Learning: A systematic literature review of the case studies in Asia and Oceania

Ekta Sharma [a, *], Ravinesh C. Deo [a], Jeffrey Soar [b], Zaher Mundheer Yaseen [c, *]

[a] Advanced Data Analytics: Environmental Modelling and Simulation Group, School of Sciences, University of Southern Queensland, Springfield QLD 4300, AUSTRALIA

[b] School of Business, University of Southern Queensland, Springfield QLD 4300, AUSTRALIA

[c] New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar 64001, Iraq

E-mail addresses: ekta.sharma@usq.edu.au (E. Sharma), ravinesh.deo@usq.edu.au (R.C. Deo), soar@usq.edu.au (J. Soar), zaheryaseen88@gmail.com (Z M Yaseen)

[*] Corresponding author: E-mail address: ekta.sharma@usq.edu.au & zaheryaseen88@gmail.com

Abstract

The Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has caused approximately 2.4 billion confirmed coronavirus disease (COVID-19) cases and 4.92 million deaths globally. Ongoing COVID-19 pandemic's novel and an extensive literature review along with Deep Learning (DL) and machine learning (ML) algorithms have been studied, with a particular emphasis on air quality in continents of Asia and Oceania. Article aims to integrate the knowledge and correlation on air pollutants and COVID-19 and thereby identifying possible future research. From an initial set of 545 articles, total of 33 articles were selected from the selected region of study through two rigorous rounds of inclusion-exclusion. The authors shortlisted 6 countries in Asia (China, India, Bangladesh, Jordan, Jakarta, Korea) and 2 in Oceania (Australia and New Zealand) that were the only countries to publish the literature relevant to our criteria, and investigated the literature, primarily on the context with keywords, the methodology, algorithms adopted, and importantly, on the air pollutant(s) namely particulate matter (PM) such as 2.5, 10, and the meteorological parameters. From our analysis, we observed that most of the research belonged to China, the epicenter of the COVID-19 and most researchers in Asia and Oceania focused on $PM_{2.5}$, followed by meteorology and PM_{10}.

KEYWORDS: Systematic review; COVID-19; Artificial intelligence; Air quality; coronavirus; deep learning.

List of abbreviations

SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus-2	COVID-19	Coronavirus disease
PM_{10}	Coarse particles with a diameter 2.5 - 10 Mm	WHO	World Health Organization
$PM_{2.5}$	Fine particles with a diameter of 2.5 Mm or less	2019-nCoV	2019 novel coronavirus
SVR	Support Vector Regression	NO**x**	Nitrogen Oxides
DL	Deep Learning	SO_2	Sulfur Dioxide
LSTMB	LSTM with Memory Between Batches	NO_2	Nitrogen Dioxide
LSTM	Long Short-Term Memory	O_3	Ozone
BiLSTM	Bidirectional LSTM	CO	Carbon Monoxide
CNN	Convolutional Neural Networks	DR	Demand Reference
EDLSTM	Encoder-Decoder LSTMs	DT	Decision Trees
LSTMReg	LSTM Network For Regression	SLSTM	Stacked LSTMs
NLP	Natural Language Processing	RF	Random Factor
GCN	Graph Convolutional Network	GRU	Gated Recurrent Unit
CVAE	Conditional Variational Autoencoder	ML	Machine Learning

1. Introduction

Coronavirus disease 2019 (COVID-19) has accelerated the global concern for protecting and improving the health of the public (Fauci, Lane, & Redfield, 2020). The virus that causes COVID-19 is Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2; previously known as '2019 novel coronavirus' (2019- nCoV) (World Health Organization (WHO)). The first outbreak was identified in December 2019 in the People's Republic of China (Hubei Province) (PEREZ, 2020). In March 2020, the outbreak of COVID-19 was declared a pandemic by the World Health Organization (WHO) based on globally growing case notification rates. Unfortunately, there are currently approximately 2.4 billion confirmed COVID-19 cases and 4.92 million deaths globally (World Health Organization (WHO)). The COVID-19 mortality rate is dependent on comorbidities, with severe cardiovascular complications and respiratory failure (Samidurai & Das, 2020). These are similar to those that are influenced by air pollution giving us compelling reasons to be interested in a potential correlation between the two and the consequences for the viral infection (Lelieveld et al., 2020; Xin, Shao, Wang, Xu, & Li, 2021).

It is critical to establish the purpose and the extent to which, particulates such as $PM_{2.5}$ (< 2.5 micrometers (μm)), PM_{10} (diameter 2.5-10 μm) play in the increase, and fatality, of the virus. As particulates have a considerable emissions footprint, (Comunian, Dongo, Milani, & Palestini, 2020; Hashim et al., 2021). The potential role of PM in the spreading of COVID-19 was researched in the first evidence-based research hypothesis by (Setti et al., 2020). Afterward, authors researched the significant correlation between mortality of COVID-19 and air pollution (Pozzer et al., 2020). Their results showed that PM pollutants were the reason for 17% mortality (North America), 27% (East Asia), 19% (Europe), and overall, 15% (worldwide). Since then, several studies have discussed the potential link between PM and COVID-19 and showcased their correlation (Boroujeni, Saberian, & Li, 2021; Chauhan &

Singh, 2020; Comunian et al., 2020; Copat et al., 2020; Frontera, Cianfanelli, Vlachos, Landoni, & Cremona, 2020).

The burning of fuel is directly associated with Greenhouse Gas (GHG) emissions into the environment; hence, all economic activities that require fuel combustion could be considered significant contributors to global pollution (Joung, Kang, Lee, & Ahn, 2020; Mor et al., 2021). Air pollution levels in the world's most densely populated cities have reached dangerously high levels, putting the population's physical health at risk (Anjum, 2020; Bhatti et al., 2021). In today's metropolitan regions, principal pollutants, such as methane (CH_4), nitrogen dioxide, (NO_2) carbon monoxide (CO), sulfur oxides (SOx), as well as secondary air pollutants like ozone (O_3), nitrogen dioxide, (NO_2), and sulfur trioxide (SO_3) have all increased consistently (Magazzino, Mele, & Schneider, 2020). However, experts have recently focused on another critical environmental issue which is the increasing level of small-sized particles (PM_{10} and $PM_{2.5}$) in the big cities (Páez-Osuna, Valencia-Castañeda, & Rebolledo, 2022). These particles are mostly sourced from factories and home heating. In large Asian and Oceanian cities, the principal source of pollution causes specific health problems to the city dwellers (Chung, Zhang, & Zhong, 2011).

Road traffic is the major source of PMs since it runs nearly entirely on petroleum fuels (gasoline & diesel) (Lozhkina, Lozhkin, Nevmerzhitsky, Tarkhov, & Vasilyev, 2016). These particles have been steadily increasing in urbanized regions over the last few decades, posing a threat to human health (Proost & Van Dender, 2012). It is a major cause of traffic congestion, prolonged travel time, and excessive fuel consumption and carbon emissions while preventing efficient travel (Barth & Boriboonsomsin, 2008). As a result, scholars have focused on the planning of urban areas with fewer traffic congestion using specific measures such as public transportation, bicycle incentives, and restriction of private vehicles; low-carbon energy sources are also introduced as energy sources (Liaquat, Kalam, Masjuki, & Jayed, 2010).

Cleaner alternative fuels, on the other hand, contribute less significantly to the energy chain and seem to be unavailable in most urban locations. As a result, the demand for fossil fuel-based transportation has continued to surge over time. This reliance on fossil fuel energy for economic activities remains the major contributor to climate change (Lei, Feng, & Lauvaux, 2020). These concerns have driven the search for cleaner sources of energy for road transportation as such energy sources are believed to be capable of reducing the health impacts associated with the waste gas from transportation systems that rely on fossil fuel.

Regarding PMs, one major issue about them is that they cause serious health problems in the community, problems such as respiratory infections, lung cancer, asthma, chronic obstructive pulmonary disease, etc. (Kim, Chen, Zhou, & Huang, 2018; Zhu et al., 2022). These particulates last longer in the air than the bigger particles due to their small weight and size; they can also easily infiltrate the human lungs and circulatory system (Qiang Wang et al., 2019). As a result, Mayors of most large cities have made improving air quality a top priority. This problem has recently become more prevalent in several places throughout the world. In fact, no other measure comes close to the COVID-19-related lockout. Surprisingly, one unintended consequence of the COVID-19 lockout is a large reduction in both primary and secondary pollution emissions, casting doubt on the long-established link between human activity and air quality.

It is essential to have look at the reported literature on the relationship between COVID-19 lockdown and the air pollution. Based on the survey conducted on Scopus database, 70 research articles were published in this research domain. Using the VOSviewer algorithm, graphical map for the major keywords of those research were presented in Figure 1a and b, 914 and 227 keywords occurrence, respectively. Figure 1a exhibited 12 research clusters and majorly focusing on COVID-19 loackdown, air pollution, forecasting/predictive models, environmental monitoring. Whereas, Figure 1b more relatively focused on two occurrence keywords with total number of keywords 227. Where seven clusters “research domians” were observed maily COVID-19 loackdown, human health, air quality/air pollution. Figure 1c presented the major countries adopted this type of research with total number of 41 countries. In general, Figure 1 revealed the significant of this research topic although all those research established with very short period 2020-2021.

< FIGURE 1 >

Given the significant differences in the extent of COVID-19 transmission globally, it would be worthwhile to evaluate the potential role of air pollution as a contributor to COVID-19 mortality (Contini & Costabile, 2020). As such, several outstanding studies have been published recently on a variety of cases involving several kinds of pollution, including England (Travaglio et al., 2021), USA (Wu, Nethery, Sabath, Braun, & Dominici, 2020), Spain (Briz-Redón, Belenguer-Sapiña, & Serrano-Aroca, 2021), China (K. Chen, Wang, Huang, Kinney, & Anastas, 2020), Vietnam (Ngo et al., 2021), India (Vadrevu et al., 2020), South Africa (Mbandi, 2020), and several other countries (Barbieri et al., 2020). These empirical studies found a substantial link between air pollution and COVID-19 cases or fatality, indicating that poor air quality is a contributing factor in COVID-19 death. This conclusion is consistent with the scientific literature that suggests that air pollution has an impact on the spread of several viral diseases (P. S. Chen et al., 2010; Peng et al., 2020). The basic idea is that a specific particle concentration can promote COVID-19 and make the respiratory system more vulnerable to infection. In reality, infections could be carried by airborne particles, making viral infections even more dangerous.

The literature search did not find papers reporting on comprehensive reviews of Asian and Oceania countries, the gap that the present authors are trying to fill with deep learning (DL). To address the literature gaps, the research novelty is as follows:

- This study reports on research that explores the robustness and potential of deep learning and machine learning for the objectives considered in the study.
- The research investigates the published literature comprehensively with all-inclusive applications of ML and DL in detecting the correlation between COVID-19 and forecasting air quality potentially for Asian and Oceania countries (as shown in Fig. 1).
- The authors suggest recommendations for further studies after reporting on the review in the context of the usage of algorithms in fighting the COVID-19 pandemic.
- Finally, the future research opportunities are summarized at the end of this paper.

The remaining systematic review has been organised as below. Section II explains the strategy with which this review has been conducted. Section III discussed the findings and data analysis. Possible opportunities for further research to fight COVID-19 by targeting air pollution with deep learning and artificial intelligence have been presented in Section IV, with Section V discussing final remarks, shortcomings, and possible suggestions for future research.

< FIGURE 2 >

2. Review Methods

For this study, the authors have adopted a methodology strategy as discussed in (Shi et al., 2021).

2.1 Identification stage – Search approach

Potential databases such as IEEE Xplore, PubMed, Google Scholar, and Springer Link were searched for related articles. Important keywords like ‘COVID-19’, ‘pandemic’ ‘CoV2’ or ‘Coronavirus’ along with ‘air quality, ‘air pollutants’, ‘particulate matter’ and key phrases: ‘machine learning, ‘deep learning, and ‘artificial intelligence’ were searched. Then the combination was searched with the name of continents ‘Asia’ and ‘Oceania’. e.g. (‘pandemic AND Deep learning OR machine learning’ AND ‘COVID-19 OR CoV2 OR coronaviruses AND Asia OR Oceania’).

2.2 Screening round 1 and 2- Criteria for inclusion or exclusion

Primarily the articles were screened (a) based on study region i.e., Asia and Oceania. (b) Duplicate content articles were screened. That means articles that study similar purposes and consider a similar set of data. Then in the second round (c), articles were excluded based on the title or abstract or non-English content. Special care was taken to select only those articles that used machine learning or deep learning techniques and develop algorithms to forecast air pollutants in the ongoing pandemic.

2.3 Eligibility

Around 545 articles were selected from the targeted databases. Fig. 2 illustrates the procedure undertaken with different stages. After the completion of the identification round, 185 articles were selected for the Asian continent and 14 for Oceania. 93 articles were excluded with duplicate content as mentioned in the above Stage. Five articles were excluded with duplicate content for Oceania. From 185 articles we were left with 92 articles after the first round for Asia, and 9 articles for Oceania. Screening round 2 helped in further filtering as 63 articles were removed based on title or irrelevant abstract or if they are not written in the English language.

Similarly, from 9 articles for Oceania, 5 were removed for similar reasons. In the end, 29 full-text articles were assessed for eligibility for Asia and 4 for Oceania.

< FIGURE 3 >

2.4 Extraction of data

After meticulously assessing the eligible articles, the next stage was to extract relevant data involving various DL and ML for our objective of forecasting air pollutants. This was essential to check what relevant algorithms and techniques were used by the researchers in Asia and Oceania. This also helped us to understand the implications due to the pandemic of COVID-19 and forecasting air pollutants.

From the selected articles, the research data was extracted primarily related to the name of the continent/country where the study was conducted, the title of the research, duration of the study, name of ML or DL technique used, and keywords. In the end, the data collected helped in condensing the published literature and in recognizing the future scope. Table I showed the result in the form of a key description of the selected reviewed studies based on Geography.

< TABLE 1 >

3. Analysis of Data and Findings

3.1 Literature considered for Asia and Oceania

Out of 29 full-text eligible articles for Asia, 2 were from Korea (~ 7%), 8 from India (~ 27%), 2 from Bangladesh (~ 7%), 1 from Jakarta (~ 3%), 1 from Jordan (~ 3%), and 15 from China (~ 53%). From Oceania, we had 4 articles, of which 3 were from Australia (75%) and 1 from New Zealand (25%). All these articles were published as original research. The time frame of all research conducted was after Jan 2020. Out of 33 articles combined in the study, 32 (~

97%) articles were published in academic journals, while one (i.e., 3%) article was archived as a pre-print. All the articles (100%) were original research.

3.2 Research Objectives and Context of the study

This article reported on the published literature and discussed the role of air pollutants on the pandemic in 8 countries of Asia and Oceania. Of all the articles from Table I, it was observed that most research was conducted with at least 6 months of data for research, however, some studies such as (Iyer & Dharshini, 2020) from India used 6 years of data from Jan 2015- Dec 2020. The research was on before and after pandemic's effect on AQ in India using machine learning. Research in Korea (Choi & Kim, 2021) used 5 years of data from Jan 2015-Dec 2019 and proposed PCA to DL forecasting models for predicting $PM_{2.5}$.

Seven years is the maximum duration of data from a study in Bangladesh (Shahriar et al., 2021) from Jan 2013-May 2019. The study uses several models such as CatBoost, decision tree amongst others to forecast Atmospheric $PM_{2.5}$. In Australia, the mega-fires of 2019-20 and their impact on air quality were studied with the effect of COVID-19 lockdown in cities of and Sydney (Ward et al., 2020). Over 2 years of data from Jan 2019-Oct to 2020 was used in the study. Overall, 18 studies took more than 1-year data, and 15 studies researched on less than 12 months. This review was further segmented based on the various air pollutants forecasted across different continents such as $PM_{2.5}$, PM_{10}, and Meteorological factors. Out of all studies (n = 33), $PM_{2.5}$ and Meteorological factors (effect of temperature, humidity, greenhouse gases, etc.) seemed to be the most popular choice amongst researchers (n=30 studies or 91%), followed by PM_{10} with (n=16 studies or 49%). At the time of writing this paper, no research was conducted with total suspended particles or visibility reducing particles with Asian or Oceania data.

3.3 Discussion of the techniques used to forecast air pollutants

Table II shows the compilation of all machine and deep learning algorithms considered (in blue) in the study. The optimal algorithm of each article is shown in red. Of all the ML algorithms considered across both continents (m=26), 4 articles (15%) considered the filter, CatBoost, Principal component analysis (PCA), Particle swarm optimization (PSO), and Ridge Classifier (R.Classifier) in the study along with other approaches. Of these only Ridge Classifier was used in Oceanian studies and PCA in Asian studies (twice). Of all these algorithms, RF (Random Forest) seems to be a popular choice for modeling, which occurred 11 times (or 34%) followed by ANN (Artificial Neural Network) (n=8 out of 33 or 24%). The next preference seems to be of DT (Decision Trees), 5 or 15% followed by SVR/SVM (Support vector regression/ Support vector machine) (n=4 or 14%) of which 1 occurrence is in the study based on data in Australia (Oceania). k-NN (k-nearest neighbors) and SMA (Slime mold algorithm) are used thrice or 9%, followed by MLP (Multilayer Perceptron), ARIMA (Autoregressive Integrated Moving Average), Others (Linear regression with gradient, Support Vector Machine with Gradient Descent, InceptionV3, and Resnet50) that were individually used once. ANFIS (adaptive neuro-fuzzy inference system), MLR (Multiple linear regression), MH (Meta-heuristics), Bayesian were also used once. There were hybrids such as MLR-ANN (MLR-Artificial Neural Network), ARIMA-SVM, MLP-Fuzzy Inference, CNB/BNB (Complement Naïve Bayes/ Bernoulli Naïve Bayes), and PSO-SMA-ANFIS are all used only once in the studies.

From (d=14) DL algorithms, the most popular modeling choice across Asia and Oceania was LSTM with 13 occurrences (or 93%), followed by BiLSTM (bidirectional LSTM), and CNN (convolutional neural networks) with 5 occurrences (36%). EDLSTM (Encoder-Decoder LSTMs), CLSTM were used 4 times (29%), GRU (gated recurrent unit) 3 times (21%), and LSTMReg (LSTM Network for Regression) 2 times. The least used algorithms in the study across continents were NLP (Natural language processing), GCN (Graph Convolutional Network), CVAE (conditional variational autoencoder), LSTMB (LSTM with Memory Between Batches), SLSTM (Stacked LSTMs) with hybrids CGRU, and PCA-LSTM being used only once.

< FIGURE 4 >

3.4 Importance of Studies conducted in shortlisted countries

Studies undertaken in Korea (Choi & Kim, 2021; Ryu & Kim, 2020) discussed the application of PCA filter to DL and how the resultant model could overcome the problems related to the current country program emphasizing predicting or forecasting PM especially $PM_{2.5}$. Seven different research works conducted in India focused on varied ML methods and three on DL methods. (Mallik, Soni, Podder, Mishra, & Ahamed, 2020) took an Indian case study that utilized deep learning with remote sensing in the pandemic and noticed changes in $PM_{2.5}$. The results showed that the hybrid method i.e. (MLR-ANN) outperformed with the highest accuracy for the prediction of $PM_{2.5}$. (Balamurali, Pachaivannan, Elamparithi, & Hemamalini, 2021) was a comparative study from Jan 2020 to June 2020 on predicting $PM_{2.5}$ levels using the LSTM model. A variety of LSTMs was considered in the study from Regression to regression with time steps, to memory between batches and stacked LSTMs. A long-term time-series pollution forecast using statistical and DL methods was done (Nath, Saha, Middya, & Roy, 2021). To forecast the future $PM_{2.5}$ and PM_{10} values, historical data and a quantitative approach were carried out. Analysis of AQ Index in India: pre & post-Covid pandemic was considered using efficient ML approaches (Iyer & Dharshini, 2020) through logistic regression and decision tree algorithms. (Nath et al., 2021) found that the major pollutants like $PM_{2.5}$, PM_{10}, NO_2, showed a significant reduction during the social distancing period, compared to the same period in previous years. It highlighted several DL methods such as CNN-LSTM, LSTM, DBN to focus on the prediction of air quality. LSTM was also used to predict AQ in

Delhi with a focus on COVID lockdown (Tiwari, Gupta, & Chandra, 2021). The study showed that there was an unprecedented deterioration of air quality after the full lockdown. The study used a variety of LSTMs for univariate and multivariate modeling. The only Indian research that focused on ML models was (Bera, Bhattacharjee, Sengupta, & Saha, 2021) where $PM_{2.5}$ was predicted in the lockdown during pandemic in Kolkata city with the help of MLR and ANN models. This research signifies the importance of the nonlinear model and its precise prediction of $PM_{2.5}$ over the study area in comparison to the linear model.

Studies in Bangladesh (Shahriar et al., 2021) showed the potential of several hybrids (ANN and SVM with ARIMA) and standalone algorithms such as decision tree and CatBoost filter for the forecasting of $PM_{2.5}$. Also, explanations for the overall prediction, calculations, and meteorology impact of coronavirus using neural networks were shown. The study showed the importance and efficiency of deep learning models for predicting $PM_{2.5}$. Another study (Haque, Hasan Pranto, All Noman, & Mahmood, 2020) discussed the impact of COVID-19 through 5 different DL networks. The researchers noticed that certain atmospheric factors such as humidity and temperature along with sun-hour have a considerable role to play (85.9%) in spreading the coronavirus. They also noticed a >90% effect on mortality with COVID-19 as the humidity had an 8.09 % impact on death. The researchers in Jakarta (Wihayati & Wibowo, 2021) predicted air quality during the COVID-19 outbreak using LSTM. The results obtained show that the Adam optimizer could have brought the results closer to the dataset used. Studies in Jordan (Park & Chang, 2021) assessed and predicted AQ in northern Jordan during the lockdown due to the COVID-19 virus pandemic using ANN. Results of the research indicated a structured and trained artificial neural network could be productive architecture to forecast parameters of AQ adequate preciseness. The concentrations of several pollutants deteriorated in the period of COVID19 with varying results. In the pollutants that are studied, nitrogen dioxide has the most deterioration of 72%, however particulate matter 10 has a minimum deterioration of 29%. It is observed that the maximum number of research papers (n=15 or 53%) have been conducted in China. All the 14 studies focused on particulate matter 2.5 algorithms except (Cole, Elliott, & Liu, 2020). None of the Chinese researchers used filters. Out of 24 ML algorithms used, RF was most widely used in several papers (7 of 15 studies or 47%). ANN was used in 27% and 9 (or 60%) papers using DL algorithms with LSTM being widely used (20%) followed by CNN (20%), BiLSTM, LSTMReg, CVAE, GRU, CLSTM, and CGRU (6%). In Oceania, we have 3 studies from Australia mostly focussing on ML, with RF being the most preferred algorithm (11 times) followed by Others (2 or 50%), DT, CNB/BNB, SVR, ANN (1 or 25%). The only study from New Zealand also used RF as the choice of modeling.

< TABLE II >

3.5 Discussion of the data used and associated problems

Nine studies (28%) used satellite-derived data for calculations and 72% used different types of data, including text and images, to corroborate their findings, as shown in Table II. Many studies did not take much data for the research due to the ongoing pandemic situations and many took multiple weather variables for their research, which may have caused some issues. (Choi & Kim, 2021) discussed the issue of dimensionality problem. They observed the problem with the fewer number of observations. They researched on 5-year data collected at the daily temporal horizon in 8 cities in Korea. Also, a fuzzy inference engine with MLP proposed in another Korean study in Seoul (Ryu & Kim, 2020) alleviated the problems of the current demand reference (DR) program. This study based its findings on the data of meteorology for $PM_{2.5}$ and PM_{10} prediction.

The study conducted in India (Mallik et al., 2020) was based on remote sensing data that have its limitations (Couture, Babin, & Perras, 1993). Another study (Balamurali et al., 2021) collected a variety of real-time data and exploratory variables at 15-minute intervals to apply ANN and LSTM to compare predictions. This involved a lot of pre-processing and cleaning to have a more ordered dataset to avoid any inaccuracies. One of the studies that took the longest duration of data was (Iyer & Dharshini, 2020) to measure the pollution level based on greenhouse gases such as Sulfur dioxide (SO_2), Nitrogen dioxide (NO_2), Carbon monoxide (CO), Ozone (O_3), $PM_{2.5}$ and PM_{10}. (Nath et al., 2021) carried long-term forecasts and compared the efficiency of deep learning models in the pre and post COVID era. The limited data availability may cause less robust techniques such as basic statistical techniques like Holt-Winters to surpass robust machine learning / deep learning models. This can be avoided by increasing the quantity of data or if the formulated model can predict at the daily/monthly time horizon, based on past data. This will result in a better and a reliable model.

Careful observation in our study reveals that the papers on both continents have not considered exogenous variables. This means the absence of all such variables whose value is determined outside the model. The authors see this as a shortcoming, and their inclusion could have increased the preciseness and computational efficiency of the models designed (Tiwari et al., 2021) discussed the need for more data and the importance of Spatio-temporal information while predicting air pollution in the COVID era. The study (Bera et al., 2021) collects spatiotemporal data of $PM_{2.5}$ to obtain remote sensing integrity with the increase of accuracy of data.

There were reports about efforts to overcome limitations of traditional hybrid methods; (Haque et al., 2020) used image data with the DNN as a time and labour saving solution, and (Shahriar et al., 2021) eliminated strong assumptions such as prediction shifts caused by gradient bias. Table III listed prominent COVID-19 datasets used by works considered in this study.

< Table III >

The literature has revealed certain limitations in the observed relationship between COVID-19 and $PM_{2.5}$ pollution; however, this is not enough to establish a causal relationship. Even with powerful machine learning models, disentangling the impacts of strongly correlated parameters is difficult and not always possible. It's especially difficult to distinguish between pollution's direct consequences and the indirect effects of things like economic and racial inequality. Another issue is determining the appropriate proxy (or proxies) for the frequency of interactions between people in a given society, even as the human-to-human form of transmission is undeniably the most prevalent in COVID-19. Considering this, some scholars have correctly emphasized the need for accurate assessment of the movement of the target population (Bontempi, Vergalli, & Squazzoni, 2020; Guo, Yu, Zhang, & Ma, 2021); they have also recommended possible proxies, ranging from specific economic measures and commercial interactions to account for the number of investors or job seekers; analysis of public transportation-related statistics has also been suggested. While identifying and finding additional variables that accurately reflect these factors are difficult in a unified and systematic manner across all the studied regions, there is still room for improving the adopted methodologies in this regard.

The possible disparity between interior and outdoor air pollution is another methodological barrier that is difficult to overcome. This is the case because people are reported to spend an average of 80-90% of their time indoors (Noorimotlagh, Jaafarzadeh, Martínez, & Mirzaee, 2021). As there are no systematic data sources on indoor pollution, we must make the logical assumption that indoor and outdoor pollution is substantially associated in general, as shown in (Harbizadeh et al., 2019). Also, there is an inescapable trade-off between the selection of an analysis scope that exhibits severe levels of air pollution on the one hand and the need for consistency in the other parameters on the other. This study is focused mainly on Asia and Oceania due to the highly polluted environment of the regions which are considered above health-hazard limits. As a result, the COVID-19 context has a significant impact on air pollution, particularly in places like Delhi, Bangkok, Kuala Lumpur, Beijing, Sydney. Owing to this fact, the research progress on the linkage between the air quality indicators and the COVID-19 transmissibility.

It's fascinating to think about random parameters when evaluating air quality. The significance of meteorological parameters such as temperature, humidity, and UV radiation on the transmission of SarS-CoV2 for example, may be less than usually anticipated based on the major predictors. Even though there are compelling arguments that high humidity and temperature levels limits virus transmission (Noorimotlagh, Mirzaee, et al., 2021; Notari, 2021), there are also compelling arguments that they have no impact on the rate of virus transmission. It is worth noting that the literature has not entirely agreed on this hypothesis as there are contrary views about it (Kolluru, Patra, Nazneen, & Shiva Nagendra, 2021; Sangkham, Thongtip, & Vongruang, 2021). Despite the emphasis on the impact of meteorological conditions in this review, some scientists dispute that weather elements have a substantial impact on the transmission of COVID-19 (Qingan Wang et al., 2021). It is important to highlight, however, that fluctuations in meteorological conditions are significantly larger on a worldwide scale, especially where these elements have a greater impact.

4. Future Research

This section presents the scope of further research opportunities on Air pollutants, and ML/DL benefits but is not limited to the COVID-19 pandemic.

4.1 Usage of a Large Dataset and removal of dimensionality

Availability of large and different types of data such as texts, images, Spatio-temporal, remote sensing data will assist in conducting several experiments and applying a range of modeling algorithms. This can result in improved performance. It would also assist in a general, yet ordered, and validated result helping immensely in the pandemic research.

4.2 Enhanced focus on air quality and DL with COVID-19

A small amount of research that considered all the factors of air pollutant forecasting, pandemic, and deep or machine learning algorithms was observed to have been published from the selected continents. Further research can be done particularly focusing on these areas.

4.3 Inaccuracies of greenhouse gas emissions

Limitations of data availability especially data without bias or data with high resolution can create problems in incorrect emissions. This may hold for gases such as CO or O_3. This can create deviation between the model formulated and the results obtained. Future research might take a top-down approach focused on individual pollutants into consideration while designing the models.

4.4 Prior evaluation of the robustness of the models

An upcoming and active research field with innovative algorithms and methods is ML/DL. It keeps on emerging resulting in a better and refined solution. Before formulating a system or approach, the pros and cons of several models need to be considered that may provide better responses to uncertainty quantification in predictions.

4.5 The development of governmental strategy

The following policy recommendations are necessarily based on the findings of the literature review: Government should focus on intervention projects aimed at stimulating the decarbonization of economic activities via the introduction of pollution reduction technologies; this will offer a long-term solution in containing COVID-19, its impacts, and future reoccurrences. A paradigm shift is also necessary acceptance of research, development, & innovations throughout economic sectors. Economic growth can only be decoupled from air pollution by relying on renewable and clean energy sources for economic activities. The promotion of a healthy and clean atmosphere could be a useful factor in reducing the rapid transmission of COVID-19. This study could not reach a consensus result due to the variations in countries across the world, as well as the complexity of COVID-19; however, the adopted empirical procedure in this work laid the groundwork for future research needed to mitigate COVID-19, its effect, and future pandemics.

5. Conclusions and future work

This manuscript examined 545 selected scientific articles focused on Asia and Oceania. These articles narrowed their research on COVID-19 and air pollution and finally, 33 papers were shortlisted that used deep learning or machine learning approaches for the considered objective in forecasting air pollutants. Deep learning and machine learning is an active and evolving research field, particularly in atmospheric sciences. Despite the rapid advancement and popularity over the past years, the usage is noticed to have been limited to the continents of North America and Eurasia. In comparison, fewer studies have been conducted in Asia and particularly the Oceania continent.

We noticed a few important points through this systematic review-

- This systematic review shows that most research papers focused on forecasting or predicting the number of air pollutants. The literature that revolved around primarily the concentration generally took an ensemble of algorithms or simply algorithms based on regression. This may be because this choice helped in an acceptable arrangement between the model performance and analysis of results.
- All these articles focused on forecasting through ML and DL techniques and prioritized accuracy over interpretability.
- Forecasting contaminants such as particulate matter is particularly challenging. The chaotic behaviour of Air pollutants creates major difficulties in tracking their three-

dimensional movement. This might be the reason for the choice of these robust algorithms.

- The analysis by authors reveals that the extent of the forecasting accuracy is lesser as compared to that of the models that used estimation. The 3-d chaotic nature of air pollutants may demand the application and better suited computationally efficient algorithms such as machine learning or deep learning. This may enable us to take care of the complexity in forecasting future contaminants.
- The preciseness of ML for particulate matter prediction as well as forecasting got to maximum numbers in comparison with the other pollutants. In general, it was observed that the preciseness of peaks with a higher rate of pollution was lesser as compared to the low or medium pollution peaks. Furthermore, the forecasting result was constrained to the meteorological contaminants like CO with Nitrogen oxides (NO**x**). Moreover, the models seemed to behave superior for weather that was extreme such as windy, snowy, fall to state a few.
- In the end, from the analysis of the selection, it is noteworthy that harmful and increased contamination concentration of particulate in the southern hemisphere and the Asia pacific is understudied when combined with deep learning or machine learning.

Therefore, the authors suggest that future articles may take care of further challenges of enhancing the models aiming to forecast air pollution along with important pollutants. This article reviewed the study done for prediction and forecasting through several methodologies in the time of the ongoing pandemic. The literature concerning the issues, challenges, methodology, and real advantages was discussed briefly. This review’s strategy is classified into main categories based on the continents: Asia and Oceania. The paper shows that the appropriateness of methods is dependent on the target dataset. Overall, in comparison to the single algorithms, the hybrid algorithms performed better and precisely.

Declaration of Competing Interest

The authors declare no conflict of interest could have appeared to influence the research conducted in this study.

Credit authorship contribution statement

Ekta Sharma: Conceptualisation, Data curation, Formal analysis, Methodology, Software Validation, Writing – Original draft, Review, and Editing. **Ravinesh C. Deo:** Conceptualisation, Methodology, Visualisation, Mentorship, Review, and Editing. **Ramendra Prasad, Alfio V. Parisi, Nawin Raj:** Mentorship, Review, and Editing. **Zaher Mundher Yaseen**; Manuscript revision, discussion, analysis, assessment. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

The first author acknowledges the funding support by Research and Training Scheme (RTS) to USQ from the Australian Government. The authors are grateful to the editor and all reviewers for their beneficial suggestions. The first author acknowledges Dr Ramendra Prasad, Dr Nawin Raj and Honorary Professor Alfio Parisi on their suggestions for an earlier draft of the paper.

References

Anjum, N. . (2020). Good in The Worst: COVID-19 Restrictions and Ease in Global Air Pollution. *Preprints*.

Balamurali, R., Pachaivannan, P., Elamparithi, P. N., & Hemamalini, R. R. (2021). Comparative Study on Predicting Particulate Matter (PM2.5) Levels Using LSTM Models. Research Square Platform LLC. https://doi.org/10.21203/rs.3.rs-436897/v1

Barbieri, D. M., Lou, B., Passavanti, M., Hui, C., Lessa, D. A., Maharaj, B., … Adomako, S. (2020). Survey data regarding perceived air quality in Australia, Brazil, China, Ghana, India, Iran, Italy, Norway, South Africa, United States before and during Covid-19 restrictions. *Data in Brief*. https://doi.org/10.1016/j.dib.2020.106169

Barth, M., & Boriboonsomsin, K. (2008). Real-world carbon dioxide impacts of traffic congestion. *Transportation Research Record*. https://doi.org/10.3141/2058-20

Bera, B., Bhattacharjee, S., Sengupta, N., & Saha, S. (2021). PM2.5 concentration prediction during COVID-19 lockdown over Kolkata metropolitan city, India using MLR and ANN models. *Environmental Challenges*, *4*, 100155. https://doi.org/10.1016/j.envc.2021.100155

Bhatti, U. A., Zeeshan, Z., Nizamani, M. M., Bazai, S., Yu, Z., & Yuan, L. (2021). Assessing the change of ambient air quality patterns in Jiangsu Province of China pre-to post-COVID-19. *Chemosphere*. https://doi.org/10.1016/j.chemosphere.2021.132569

Bontempi, E., Vergalli, S., & Squazzoni, F. (2020). Understanding COVID-19 diffusion requires an interdisciplinary, multi-dimensional approach. *Environmental Research*. https://doi.org/10.1016/j.envres.2020.109814

Boroujeni, M., Saberian, M., & Li, J. (2021). Environmental impacts of COVID-19 on Victoria, Australia, witnessed two waves of Coronavirus. *Environmental Science and Pollution Research International*, *28*(11), 14182–14191. https://doi.org/10.1007/s11356-021-12556-y

Briz-Redón, Á., Belenguer-Sapiña, C., & Serrano-Aroca, Á. (2021). Changes in air pollution during COVID-19 lockdown in Spain: A multi-city study. *Journal of Environmental Sciences (China)*. https://doi.org/10.1016/j.jes.2020.07.029

Chauhan, A., & Singh, R. P. (2020). Decline in PM(2.5) concentrations over major cities around the world associated with COVID-19. *Environmental Research*, *187*, 109634. https://doi.org/10.1016/j.envres.2020.109634

Chen, K., Wang, M., Huang, C., Kinney, P. L., & Anastas, P. T. (2020). Air pollution reduction and mortality benefit during the COVID-19 outbreak in China. *The Lancet Planetary Health*. https://doi.org/10.1016/S2542-5196(20)30107-8

Chen, P. S., Tsai, F. T., Lin, C. K., Yang, C. Y., Chan, C. C., Young, C. Y., & Lee, C. H. (2010). Ambient influenza and avian influenza virus during dust storm days and background days. *Environmental Health Perspectives*. https://doi.org/10.1289/ehp.0901782

Choi, S., & Kim, B. (2021). Applying PCA to Deep Learning Forecasting Models for Predicting PM2.5. *Sustainability*, *13*(7), 3726. https://doi.org/10.3390/su13073726

Chung, K. F., Zhang, J., & Zhong, N. (2011). Outdoor air pollution and respiratory health in Asia. *Respirology*. https://doi.org/10.1111/j.1440-1843.2011.02034.x

Cole, M. A., Elliott, R. J. R., & Liu, B. (2020). The Impact of the Wuhan Covid-19 Lockdown on Air Pollution and Health: A Machine Learning and Augmented Synthetic Control Approach. *Environmental & Resource Economics*, 1–28. https://doi.org/10.1007/s10640-020-00483-4

Comunian, S., Dongo, D., Milani, C., & Palestini, P. (2020). Air Pollution and Covid-19: The Role of Particulate Matter in the Spread and Increase of Covid-19's Morbidity and Mortality. *International Journal of Environmental Research and Public Health*, *17*(12), 4487. https://doi.org/10.3390/ijerph17124487

Contini, D., & Costabile, F. (2020). Does air pollution influence COVID-19 outbreaks? *Atmosphere*. https://doi.org/10.3390/ATMOS11040377

Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., … Ferrante, M. (2020). The role of air pollution (PM and NO(2)) in COVID-19 spread and lethality: A systematic review. *Environmental Research*, *191*, 110129. https://doi.org/10.1016/j.envres.2020.110129

Couture, R., Babin, M., & Perras, S. (1993). Survey of woodcock habitat with Landsat: possibilities and limitations of remote sensing. In *Proceedings of the Eight American woodcock Symposium*.

Fauci, A. S., Lane, H. C., & Redfield, R. R. (2020). Covid-19 - Navigating the Uncharted. *The New England Journal of Medicine*, *382*(13), 1268–1269.

https://doi.org/10.1056/NEJMe2002387
Frontera, A., Cianfanelli, L., Vlachos, K., Landoni, G., & Cremona, G. (2020). Severe air pollution links to higher mortality in COVID-19 patients: The “double-hit” hypothesis. *The Journal of Infection*, *81*(2), 255–259. https://doi.org/10.1016/j.jinf.2020.05.031
Guo, Y., Yu, H., Zhang, G., & Ma, D. T. (2021). Exploring the impacts of travel-implied policy factors on COVID-19 spread within communities based on multi-source data interpretations. *Health and Place*. https://doi.org/10.1016/j.healthplace.2021.102538
Haque, A. K. M. B., Hasan Pranto, T., All Noman, A., & Mahmood, A. (2020). Insight about Detection, Prediction and Weather Impact of Coronavirus (Covid-19) using Neural Network. *International Journal of Artificial Intelligence & Applications*, *11*(4), 67–81. https://doi.org/10.5121/ijaia.2020.11406
Harbizadeh, A., Mirzaee, S. A., Khosravi, A. D., Shoushtari, F. S., Goodarzi, H., Alavi, N., … Goudarzi, G. (2019). Indoor and outdoor airborne bacterial air quality in day-care centers (DCCs) in greater Ahvaz, Iran. *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2019.116927
Hashim, B. M., Al-Naseri, S. K., Al Maliki, A., Sa’adi, Z., Malik, A., & Yaseen, Z. M. (2021). On the investigation of COVID-19 lockdown influence on air pollution concentration: regional investigation over eighteen provinces in Iraq. *Environmental Science and Pollution Research*, 1–19.
Iyer, K. B. P., & Dharshini, V. (2020). ANALYSIS AND PREDICTION OF AIR QUALITY INDEX IN INDIA DURING PRE AND POST COVID PANDEMIC USING MACHINE LEARNING ALGORITHMS. *PalArch’s Journal of Archaeology of Egypt/Egyptology*, *17*(7), 6995–7003.
Joung, T.-H., Kang, S.-G., Lee, J.-K., & Ahn, J. (2020). The IMO initial strategy for reducing Greenhouse Gas (GHG) emissions, and its follow-up actions towards 2050. *Journal of International Maritime Safety, Environmental Affairs, and Shipping*, *4*(1), 1–7.
Kim, D., Chen, Z., Zhou, L.-F., & Huang, S.-X. (2018). Air pollutants and early origins of respiratory diseases. *Chronic Diseases and Translational Medicine*. https://doi.org/10.1016/j.cdtm.2018.03.003
Kolluru, S. S. R., Patra, A. K., Nazneen, & Shiva Nagendra, S. M. (2021). Association of air pollution and meteorological variables with COVID-19 incidence: Evidence from five megacities in India. *Environmental Research*. https://doi.org/10.1016/j.envres.2021.110854
Lei, R., Feng, S., & Lauvaux, T. (2020). Country-scale trends in air pollution and fossil fuel CO2 emissions during 2001–2018: confronting the roles of national policies and economic growth. *Environmental Research Letters*, *16*(1), 14006.
Lelieveld, J., Pozzer, A., Pöschl, U., Fnais, M., Haines, A., & Münzel, T. (2020). Loss of life expectancy from air pollution compared to other risk factors: a worldwide perspective. *Cardiovascular Research*, *116*(11), 1910–1917. https://doi.org/10.1093/cvr/cvaa025
Liaquat, A. M., Kalam, M. A., Masjuki, H. H., & Jayed, M. H. (2010). Potential emissions reduction in road transport sector using biofuel in developing countries. *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2010.07.003
Lozhkina, O., Lozhkin, V., Nevmerzhitsky, N., Tarkhov, D., & Vasilyev, A. (2016). Motor transport related harmful PM2. 5 and PM10: from onroad measurements to the modelling of air pollution by neural network approach on street and urban level. In *Journal of Physics: Conference Series* (Vol. 772, p. 12031). IOP Publishing.
Magazzino, C., Mele, M., & Schneider, N. (2020). The relationship between air pollution and COVID-19-related deaths: An application to three French cities. *Applied Energy*. https://doi.org/10.1016/j.apenergy.2020.115835
Mallik, S., Soni, S., Podder, K., Mishra, U., & Ahamed, M. (2020). Prediction and assessment

of change in PM2.5 during COVID-19 lockdown using remote sensing and deep learning approach: A case study of Kanpur city. Research Square Platform LLC. https://doi.org/10.21203/rs.3.rs-88484/v1

Mbandi, A. M. (2020). Commentary air pollution in Africa in the time of COVID-19: The air we breathe indoors and outdoors. *Clean Air Journal*. https://doi.org/10.17159/CAJ/2020/30/1.8227

Mor, S., Kumar, S., Singh, T., Dogra, S., Pandey, V., & Ravindra, K. (2021). Impact of COVID-19 lockdown on air quality in Chandigarh, India: Understanding the emission sources during controlled anthropogenic activities. *Chemosphere*. https://doi.org/10.1016/j.chemosphere.2020.127978

Nath, P., Saha, P., Middya, A. I., & Roy, S. (2021). Long-term time-series pollution forecast using statistical and deep learning methods. *Neural Computing & Applications*, 1–20. https://doi.org/10.1007/s00521-021-05901-2

Ngo, T. X., Do, N. T. N., Phan, H. D. T., Tran, V. T., Mac, T. T. M., Le, A. H., … Nguyen, T. T. N. (2021). Air pollution in Vietnam during the COVID-19 social isolation, evidence of reduction in human activities. *International Journal of Remote Sensing*, *42*(16), 6126–6152.

Noorimotlagh, Z., Jaafarzadeh, N., Martínez, S. S., & Mirzaee, S. A. (2021). A systematic review of possible airborne transmission of the COVID-19 virus (SARS-CoV-2) in the indoor air environment. *Environmental Research*. https://doi.org/10.1016/j.envres.2020.110612

Noorimotlagh, Z., Mirzaee, S. A., Jaafarzadeh, N., Maleki, M., Kalvandi, G., & Karami, C. (2021). A systematic review of emerging human coronavirus (SARS-CoV-2) outbreak: focus on disinfection methods, environmental survival, and control and prevention strategies. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-020-11060-z

Notari, A. (2021). Temperature dependence of COVID-19 transmission. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.144390

Páez-Osuna, F., Valencia-Castañeda, G., & Rebolledo, U. A. (2022). The link between COVID-19 mortality and PM2.5 emissions in rural and medium-size municipalities considering population density, dust events, and wind speed. *Chemosphere*. https://doi.org/10.1016/j.chemosphere.2021.131634

Park, J., & Chang, S. (2021). A Particulate Matter Concentration Prediction Model Based on Long Short-Term Memory and an Artificial Neural Network. *International Journal of Environmental Research and Public Health*, *18*(13), 6801. https://doi.org/10.3390/ijerph18136801

Peng, L., Zhao, X., Tao, Y., Mi, S., Huang, J., & Zhang, Q. (2020). The effects of air pollution and meteorological factors on measles cases in Lanzhou, China. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-020-07903-4

PEREZ, J.-C. (2020). WUHAN COVID-19 SYNTHETIC ORIGINS AND EVOLUTION. *International Journal of Research -GRANTHAALAYAH*, *8*(2), 285–324. https://doi.org/10.29121/granthaalayah.v8.i2.2020.221

Pozzer, A., Dominici, F., Haines, A., Witt, C., Münzel, T., & Lelieveld, J. (2020). Regional and global contributions of air pollution to risk of death from COVID-19. *Cardiovascular Research*, *116*(14), 2247–2253. https://doi.org/10.1093/cvr/cvaa288

Proost, S., & Van Dender, K. (2012). Energy and environment challenges in the transport sector. *Economics of Transportation*. https://doi.org/10.1016/j.ecotra.2012.11.001

Ryu, J., & Kim, J. (2020). Demand Response Program Expansion in Korea through Particulate Matter Forecasting Based on Deep Learning and Fuzzy Inference. *Energies*, *13*(23), 6393. https://doi.org/10.3390/en13236393

Samidurai, A., & Das, A. (2020). Cardiovascular Complications Associated with COVID-19 and Potential Therapeutic~Strategies. *International Journal of Molecular Sciences, 21*(18), 6790. https://doi.org/10.3390/ijms21186790

Sangkham, S., Thongtip, S., & Vongruang, P. (2021). Influence of air pollution and meteorological factors on the spread of COVID-19 in the Bangkok Metropolitan Region and air quality during the outbreak. *Environmental Research*. https://doi.org/10.1016/j.envres.2021.111104

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Piazzalunga, A., … Miani, A. (2020). The Potential role of Particulate Matter in the Spreading of COVID-19 in Northern Italy: First Evidence-based Research Hypotheses. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2020.04.11.20061713

Shahriar, S. A., Kayes, I., Hasan, K., Hasan, M., Islam, R., Awang, N. R., … Salam, M. A. (2021). Potential of ARIMA-ANN, ARIMA-SVM, DT and CatBoost for Atmospheric PM2.5 Forecasting in Bangladesh. *Atmosphere, 12*(1), 100. https://doi.org/10.3390/atmos12010100

Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., … Shen, D. (2021). Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19. *IEEE Reviews in Biomedical Engineering, 14*, 4–15. https://doi.org/10.1109/rbme.2020.2987975

Tiwari, A., Gupta, R., & Chandra, R. (2021). Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown. *ArXiv Preprint ArXiv:2102.10551*.

Travaglio, M., Yu, Y., Popovic, R., Selley, L., Leal, N. S., & Martins, L. M. (2021). Links between air pollution and COVID-19 in England. *Environmental Pollution*. https://doi.org/10.1016/j.envpol.2020.115859

Vadrevu, K. P., Eaturu, A., Biswas, S., Lasko, K., Sahu, S., Garg, J. K., & Justice, C. (2020). Spatial and temporal variations of air pollution over 41 cities of India during the COVID-19 lockdown period. *Scientific Reports*. https://doi.org/10.1038/s41598-020-72271-5

Wang, Qiang, Kwan, M. P., Zhou, K., Fan, J., Wang, Y., & Zhan, D. (2019). The impacts of urbanization on fine particulate matter (PM2.5) concentrations: Empirical evidence from 135 countries worldwide. *Environmental Pollution*. https://doi.org/10.1016/j.envpol.2019.01.086

Wang, Qingan, Zhao, Y., Zhang, Y., Qiu, J., Li, J., Yan, N., … Zhao, Y. (2021). Could the ambient higher temperature decrease the transmissibility of COVID-19 in China? *Environmental Research*. https://doi.org/10.1016/j.envres.2020.110576

Ward, M., Tulloch, A. I. T., Radford, J. Q., Williams, B. A., Reside, A. E., Macdonald, S. L., … Vine, S. J. (2020). Impact of 2019–2020 mega-fires on Australian fauna habitat. *Nature Ecology & Evolution, 4*(10), 1321–1326.

Wihayati, & Wibowo, F. W. (2021). Prediction of air quality in Jakarta during the COVID-19 outbreak using long short-term memory machine learning. *IOP Conference Series: Earth and Environmental Science, 704*(1), 12046. https://doi.org/10.1088/1755-1315/704/1/012046

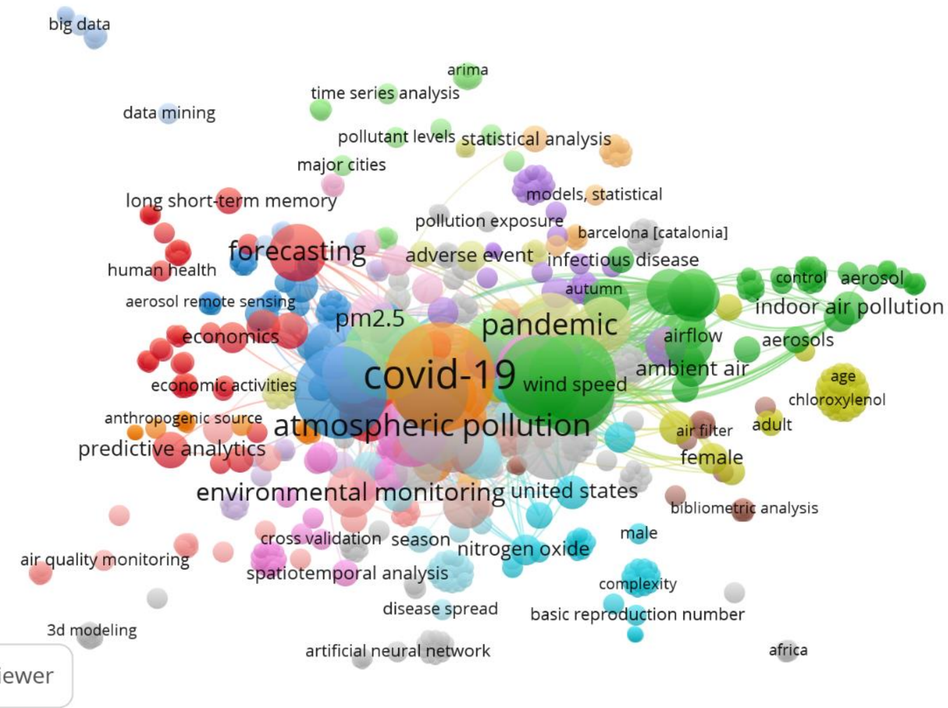
Wu, X., Nethery, R. C., Sabath, M. B., Braun, D., & Dominici, F. (2020). Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *Science Advances*. https://doi.org/10.1101/2020.04.05.20054502

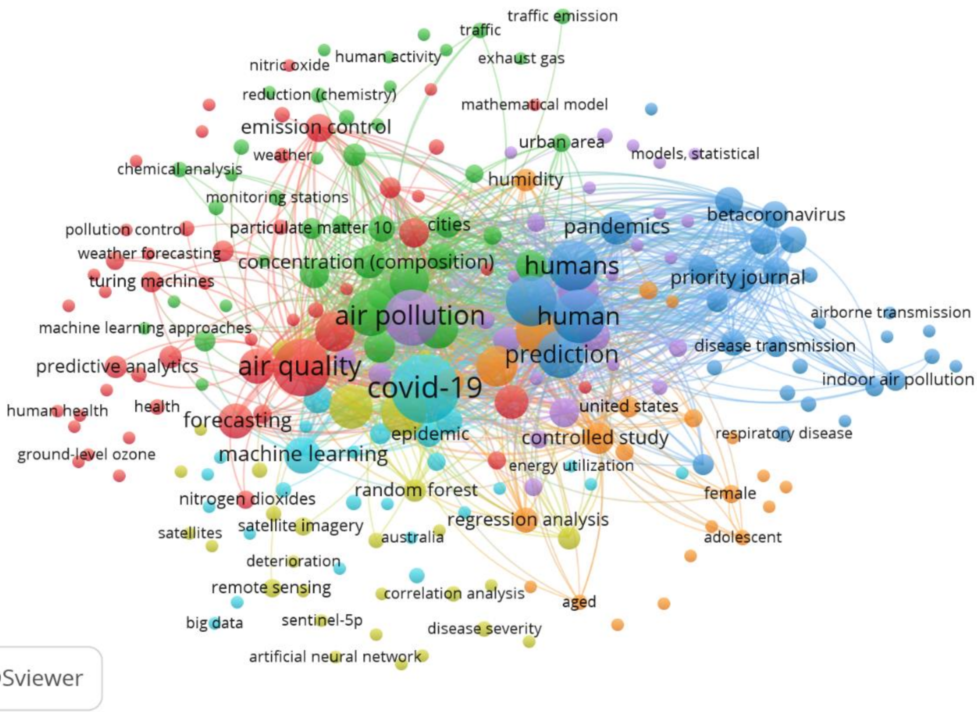
Xin, Y., Shao, S., Wang, Z., Xu, Z., & Li, H. (2021). COVID-2019 lockdown in Beijing: A rare opportunity to analyze the contribution rate of road traffic to air pollutants. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2021.102989

Zhu, W., Zheng, H., Liu, J., Cai, J., Wang, G., Li, Y., … Nie, J. (2022). The correlation between chronic exposure to particulate matter and spontaneous abortion: A meta-analysis. *Chemosphere*. https://doi.org/10.1016/j.chemosphere.2021.131802

Figure Click here to access/download;Figure;List of Figures_APR.docx

List of Figures





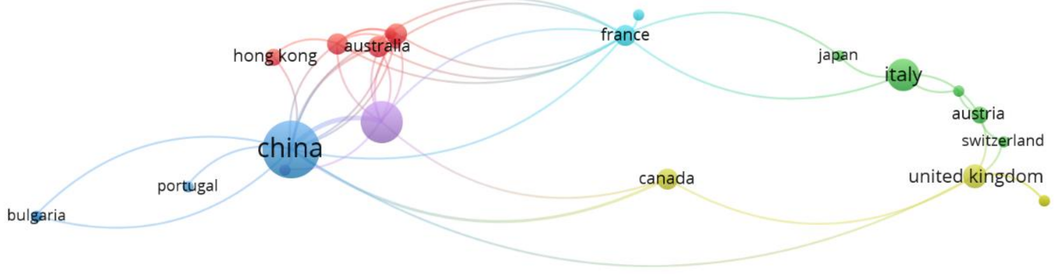


Fig. 1: The major research keywords adopted on the influence of COVID-19 lockdown and the air pollution (a) collection of 914 keywords (b) collection of 227 keywords (c) the major countries established this research domain.

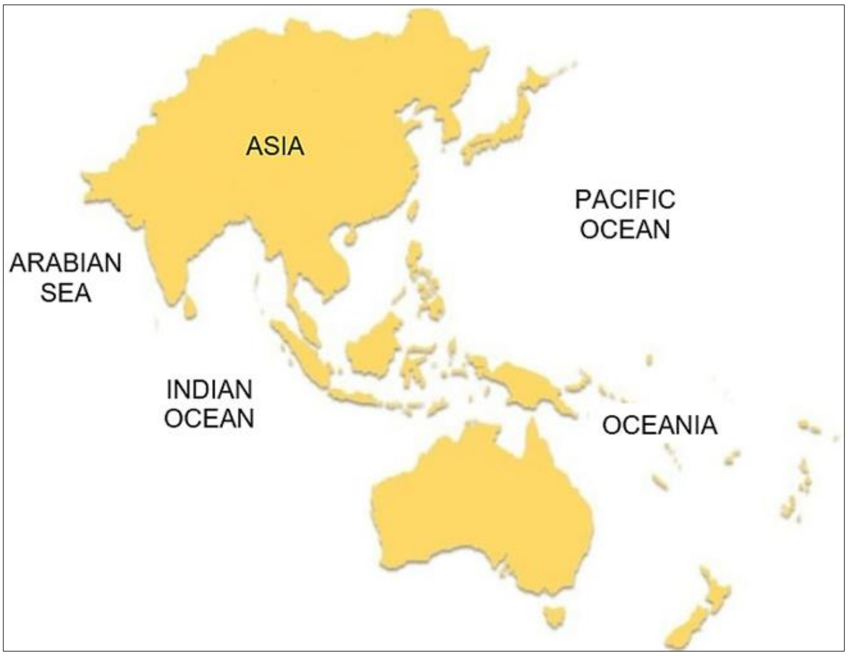


Fig. 2: Illustration of the continents considered for the study.

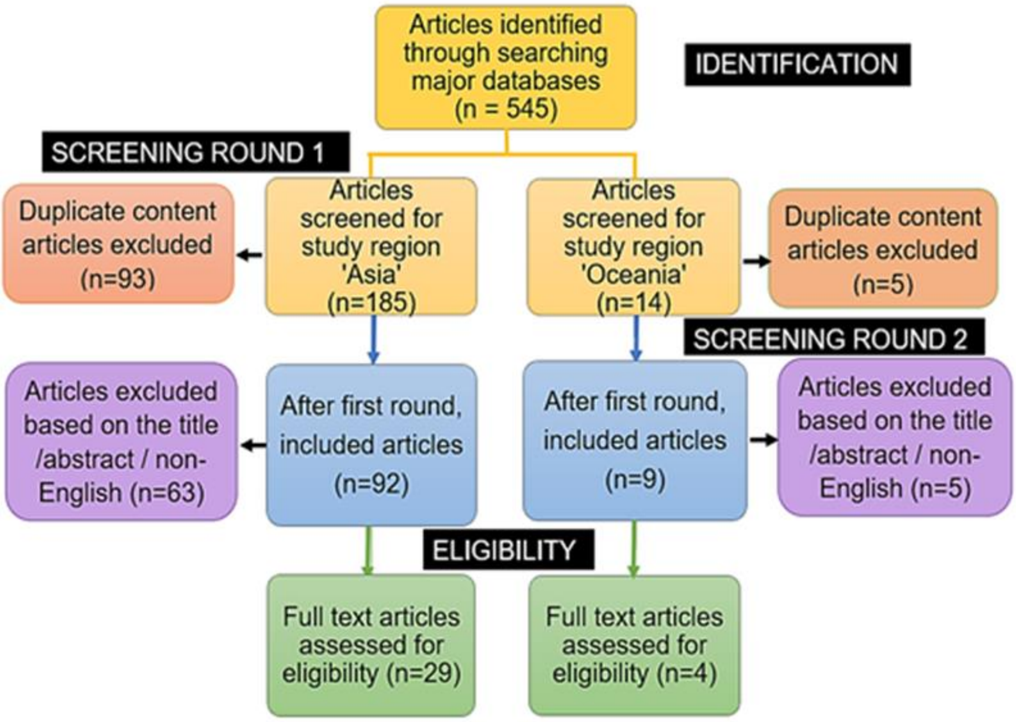


Fig. 3. Systematic selection of studies and flowchart based on PRISMA.

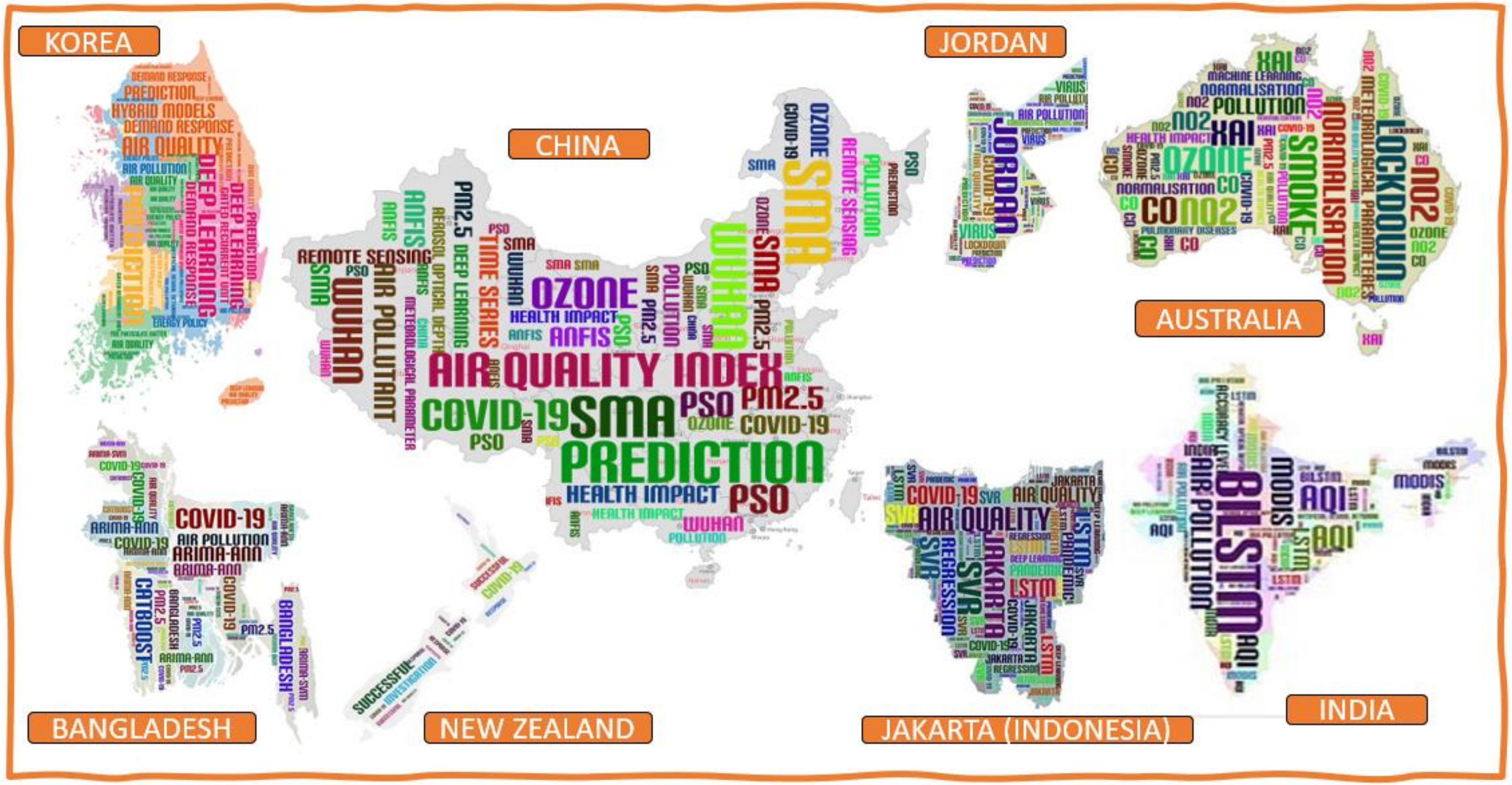


Fig. 4. Word cloud showing the prominent keywords of the specific research undertaken by Asian and Oceania countries shortlisted in this paper.

List of Tables

Table I The description of the Literature review based on Geography. *=Total studies. †=Acronyms.

Country	Title	Based on	Duration	PM2.5	PM10	Metereo--logical	Findings
India (8) *	(Mallik et al., 2020)	A case study focusing on remote sensing and DL in COVID 19. Predicting and estimating changes in PM2.5.	Jan 2019- Apr 2020	✓	✕	✓	Reduction of 26% in PM2.5 in comparison to pre lockdown
	(Balamurali et al., 2021)	Comparative Study on Predicting PM2.5 Levels Using LSTM Models	Jan 2020- Jun 2020	✓	✕	✓	A comparative study to predict PM2.5
	(Iyer and Dharshini, 2020)	Before and after analysis of AQ index through ML in COVID for India	Jan 2015-Dec 2020	✓	✓	✓	Analyses AQ during pre-& post-COVID days
	(Nath et al., 2021)	DL and statistical analysis for the forecast of long-term pollution	Jan 2020- Dec 2020	✓	✓	✓	A comparative study to forecast pollutants
	(Tripathi and Pathak, 2021)	Deep Learning Techniques for Air Pollution	Jan 2020- Dec 2020	✓	✓	✓	CNN-LSTM is proposed as a popular model
	(Tiwari et al., 2021)	Usage of LSTM on AQ of Delhi in the lockdown (COVD 19)	Jan 2019- May2020	✓	✕	✓	BiLSTM model provides the best predictions
	(Bera et al., 2021)	Using ANN and MLR models to predict PM 2.5 in COVID in the city of Kolkata	Mar2020- May2020	✓	✕	✓	Compare the accuracy of models
	(Roy and Chatterjee, 2021)	Increase and spread of AQI and the impact of dual lockdown	Jan 2020-Jun 2020	✓	✕	✓	Rise of air quality level predicted
Bangladesh (2)	(Shahriar et al., 2021)	Forecasting PM2.5 in Bangladesh with CatBoost, DT, Hybrid ARIMA with SVM and ANN	Jan 2013 - May 2019	✓	✕	✕	Suggests using DL for predicting PM2.5
	(Haque et al., 2021)	NN and analysis of correlation of weather and COVID 19	Jan 2020- Aug2020	✕	✕	✓	Weather holds 85.88% impact on COVID19
Korea (2)	(Choi and Kim, 2021)	Forecasting Models for Predicting PM2.5 with DL and PCA	Jan 2015-Dec 2019	✓	✕	✓	PCA in DL can lead to Improvements

	(Cole et al., 2020)	The Impact of the Wuhan lockdown on Air Pollution and Health: ML & Augmented Control Approach	Jan 2014-Feb 2020	✗	✓	✓	NO_2 reduction calculation by 63% from the pre-lockdown level
	(Ryu and Kim, 2020)	Demand response in Korea through PM Forecasting by DL & fuzzy inference	Jan 2020- Feb 2020	✓	✓	✓	The model solves DR programs loopholes
Jakarta (Indonesia) (1)	(Wibowo, 2021)	Prediction of air quality in Jakarta during the COVID-19 outbreak using LSTM	Jan 2020-Jun 2020	✓	✓	✓	Predict air quality with LSTM
Jordan (1)	(Shatnawi and Abu-Qdais, 2021)	Assessing and predicting AQ in northern Jordan during the lockdown due to the COVID-19 virus pandemic using ANN	Jan 2019-Mar 2020	✓	✓	✓	Structured ANN can be a useful tool
China (15)	(Al-Qaness et al., 2021)	ANFIS model for forecasting Wuhan City AQ and COVID-19 lockdown	Jun2016-Jun 2020	✓	✗	✓	Decreases in PM2.5, CO2, SO_2, NO_2.
	(Tan et al., 2020)	Story of environment and people interaction on PM2.5 with COVID 19	Jan 2020-Mar 2020	✓	✗	✗	COVID-19 decreased the PM2.5
	(Han et al., 2021)	Deep-AIR: A Hybrid CNN-LSTM Framework for AQ Modeling	Dec2018-Mar 2020	✓	✗	✓	Air pollutants can model disease transmissibility
	(Zhan et al., 2020)	Prediction of Air Quality in Major Cities of China by Deep Learning	Jan 2015-Dec 2019	✓	✓	✓	Meteorology is the best estimator for NO_2 & PM2.5
	(COSIMO and Marco, 2021)	Hubei area deaths and evidence from Neural Networks: Covid19, economic growth, and air pollution nexus.	Jan 2020-Jul 2020	✓	✓	✓	Strong relation between PM2.5 & COVID-19 deaths.
	(Zhao et al., 2021)	Anomalies generated by COVID 19 with unsupervised PM2.5	Jan 2017-Feb 2020	✓	✗	✗	CVAE detection discerns abrupt changes in PM2.5
	(Lu et al., 2021)	ML in the Yangtze River delta: ambient variations and estimates of PM2.5 during COVID-19 Pandemic	Jan 2019-Feb 2020	✓	✗	✓	Estimating PM2.5 decrease in the area of Yangtze river
	(Han et al., 2021)	Deep-AIR: A Hybrid CNN-LSTM Framework for AQ Modeling	Dec2018-Mar 2020	✓	✗	✓	Air pollutants can model disease transmissibility
	(Li et al., 2021)	ML models and satellite data: estimating the Impact of COVID-19 on the PM2.5 Levels in China.	Nov 2018-Apr 2020	✓	✗	✓	ML estimated spatiotemporally PM2.5 and the level was lowered by 4.8 µg/m3

	(Song et al., 2021)	Air pollution in central and eastern China based on ML in COVID-19 Lockdown.	Jan 2018-Dec 2020	✓	✓	✓	All the measured pollutants of the study were reduced by 16.4%, 24.2%, and 19.8%
	(Hu et al., 2021)	ML insights and covid-19 effect in air pollutants: future control policy	Jan 2015-Dec 2020	✓	✓	✓	PM 2.5, PM 10, SO_2, NO_2, & CO lowered by 39.4%, 50.1%, 51.8%, 43.1%, & 35.1%,
	(Rahman et al., 2021)	Review on ML in COVID-19, AQ, and Human Mobility	Jan 2020-Dec 2020	✓	✓	✓	Significant improvement on AQ in COVID-19 lockdown
	(Wang et al., 2020)	Four-Month Changes in AQ during & after the COVID-19 in China	Jan 2020-Apr 2020	✓	✗	✓	NO_2 lowered by 36−53%. PM2.5 was also reduced.
	(Einsiedler et al.)	Understanding the Impact of transferable models on AQ in lockdown: A technical report	May 2020-Feb 2021	✓	✓	✓	The only study to use transfer learning to fit variables considered.
	(Dai et al., 2021)	Spring Festival and COVID-19 Lockdown: Disentangling PM Sources in Major Chinese Cities	Jan 2015-Dec 2020	✓	✓	✓	−15.4%, −17.0%, −14.5%, −7.6%, −9.7%, & +24.6% changes for NO_2, SO_2, CO, PM10, PM2.5, O_3,
Australia (3)	(Ryan et al., 2021)	Megafires in black summer of 2019-2020 and AQ's health impact with Sydney and Melbourne COVID-19 lockdown.	Jan 2019-Oct 2020	✓	✓	✓	Significant increases of O_3, CO, PM10 & PM2.5.
	(Duc et al., 2021)	COVID-19 lockdown effect in Sydney's AQ.	Apr2020-Jun 2020	✓	✗	✓	NO_2, CO, & PM2.5 decreased, O3 increased
	(Gupta et al., 2021)	Whether the weather will help us weather the COVID-19 pandemic: Using ML to measure Twitter users' perceptions	Jan 2020-Jun 2020	✗	✗	✓	40.4% uncertain about weather's impact, 33.5 % no effect, & 26.1 % some effect.
New Zealand (1)	(Talbot et al., 2021)	Response with the investigation to COVID 19 and weather impact on NZ air quality.	Dec 2018-June 2020	✓	✓	✓	ML found good R for NO_2; Modelling results were weaker for coastal particulate data.

Table II A Compilation of all Machine learning and Deep learning algorithms[#] considered (in BLUE) in this study with their total count. The optimal algorithm for each research work is shown in RED).

ASIA (K=Korea, I=India, B=Bangladesh, JK=Jakarta, J=Jordan, C=China)
OCEANIA (A=Australia, N=New Zealand)

		I								B		K		JK	JO	C															A			N	Count
	Study * →	1	2	3	4	5	6	7	8	1	2	1	2	1	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	1	
Filter	CatBoost									RED																									1
	PCA									BLUE		BLUE																							2
	PSO				BLUE																														1
	RClassifier																																BLUE		1
Machine learning algorithm	ANFIS															BLUE																			1
	SVR/SVM													BLUE									BLUE				BLUE						BLUE		4
	ANN	BLUE							BLUE				BLUE		RED						BLUE		BLUE			BLUE	BLUE								8
	k-NN												BLUE														BLUE						BLUE		1
	ARIMA (AR)									BLUE																									1
	AR-SVM									BLUE																									1
	MLP						BLUE						BLUE																						1
	MLR	BLUE																																	1
	MLR-ANN	RED																																	1
	PSO-ANFIS															RED																			1
	MLP-FI												RED																						1
	DT				BLUE					BLUE	BLUE																BLUE						BLUE		5
	SMA															BLUE								BLUE											2
	RF						BLUE															RED	RED	BLUE	RED		BLUE		RED	BLUE	RED		BLUE	RED	11
	Bayesian																										BLUE								1
	CNB/BNB																																BLUE		1
	Others										BLUE							RED						RED	BLUE	RED						RED	BLUE		7

Deep learning algorithm	CNN	5
	NLP	2
	LSTM	12
	ED-LSTM	4
	Bi-LSTM	5
	LSTM-Reg	2
	LSTMB	1
	SLSTM	1
	CVAE	2
	GRU	1
	PCA-LSTM	1
	CLSTM	3
	CGRU	1
	T. learning	1
Data	Satellite	12

* Reference of all studies are from Table I.

The complete names of models are mentioned in Section III, part C.

Table III A list of prominent COVID-19 Datasets used by works considered in this study

Resources	Link	Details
AitsLab Corona	Aitslab/corona	NLP toolbox for COVID-19 NLP research.
Amazon AWS	aws.amazon.com/covid-19	Public repository for COVID-19 data analysis.
Australia	australia.gov.au/	Official repository of Australian state and territory COVID-19 figures
Bangladesh	dghs.gov.bd/index.php/en/home/5343-covid-19-update	Directorate of Health Services, COVID-19 dashboard of Bangladesh
BSTI Imaging database	bsti.org.uk/training-and-education/covid-19-bsti-imaging-database/	British Society of Thoracic Imaging Covid 19 CT scans data
CORD-19	semanticscholar.org/cord19	Open research database and a free resource for over 52,000 scholarly articles
COVID-19 Graphs	worldometers.info/coronavirus/worldwide-graphs/	This repository gives the tools to visualize the various statistics of COVID-19 using case data
HARVARD Dataverse for China	dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/MR5IJN	Harvard Repository of China's COVID-19 statistics
India	mygov.in/covid-19/	Government of India's repository of COVID-19 cases
Jakarta	corona.jakarta.go.id/en	Jakarta official webpage of COVID-19 cases
Kingdom of Jordan	corona.moh.gov.jo/en	Reports and COVID-19 statistics of Kingdom of Jordan
LitCOVID	ncbi.nlm.nih.gov/research/coronavirus/	Curated literature hub for tracking 2019 novel Coronavirus.
MONTREAL.AI	montrealartificialintelligence.com/covid19/	A resource to use with deep reinforcement learning
NIH NLM LitCovid	ncbi.nlm.nih.gov/research/coronavirus/	Curated literature hub for tracking up-to-date scientific information about COVID-19 with central access to relevant articles in PubMed
Republic of Korea	ncov.mohw.go.kr/en/	Repository of case statistics of Korea
UN Humanitarian data exchange	data.humdata.org/	United Nations OCHA Humanitarian Data Exchange Project
World Health Organisation (WHO)	covid19.who.int/	WHO Coronavirus (COVID-19) Dashboard

References (linked by Endnote. Mentioned already in the manuscript):

Al-Qaness MAA, Fan H, Ewees AA, Yousri D, Abd Elaziz M. Improved ANFIS model for forecasting Wuhan City Air Quality and analysis COVID-19 lockdown impacts on air quality. Environ Res 2021; 194: 110607.

Balamurali R, Pachaivannan P, Elamparithi PN, Hemamalini RR. Comparative Study on Predicting Particulate Matter (PM2. 5) Levels Using LSTM Models. 2021.

Bera B, Bhattacharjee S, Sengupta N, Saha S. PM2. 5 concentration prediction during COVID-19 lockdown over Kolkata metropolitan city, India using MLR and ANN models. Environmental Challenges 2021: 100155.

Choi S, Kim B. Applying PCA to Deep Learning Forecasting Models for Predicting PM2.5. Sustainability 2021; 13.

Cole MA, Elliott RJ, Liu B. The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. Environmental and Resource Economics 2020; 76: 553-580.

COSIMO M, Marco M. A Neural Network Evidence of the Nexus Among Air Pollution, Economic Growth, and COVID-19 Deaths in the Hubei Area. Advances in Environmental and Engineering Research 2021; 2: 1-1.

Dai Q, Hou L, Liu B, Zhang Y, Song C, Shi Z, et al. Spring Festival and COVID-19 lockdown: disentangling PM sources in major Chinese cities. Geophysical research letters 2021: e2021GL093403.

Duc H, Salter D, Azzi M, Jiang N, Warren L, Watt S, et al. The Effect of Lockdown Period during the COVID-19 Pandemic on Air Quality in Sydney Region, Australia. International Journal of Environmental Research and Public Health 2021; 18: 3528.

Einsiedler J, Cheng Y, Papst F, Saukh O. Transferable Models to Understand the Impact of Lockdown Measures on Local Air Quality.

Gupta M, Bansal A, Jain B, Rochelle J, Oak A, Jalali MS. Whether the weather will help us weather the COVID-19 pandemic: Using machine learning to measure twitter users' perceptions. International journal of medical informatics 2021; 145: 104340.

Han Y, Zhang Q, Li VO, Lam JC. Deep-AIR: A Hybrid CNN-LSTM Framework for Air Quality Modeling in Metropolitan Cities. arXiv preprint arXiv:2103.14587 2021.

Haque A, Pranto TH, Noman AA, Mahmood A. Insight about Detection, Prediction and Weather Impact of Coronavirus (COVID-19) using Neural Network. arXiv preprint arXiv:2104.02173 2021.

Hu J, Pan Y, He Y, Chi X, Zhang Q, Song T, et al. Changes in air pollutants during the COVID-19 lockdown in Beijing: Insights from a machine-learning technique and implications for future control policy. Atmospheric and Oceanic Science Letters 2021: 100060.

Iyer KP, Dharshini V. ANALYSIS AND PREDICTION OF AIR QUALITY INDEX IN INDIA DURING PRE AND POST COVID PANDEMIC USING MACHINE LEARNING ALGORITHMS. PalArch's Journal of Archaeology of Egypt/Egyptology 2020; 17: 6995-7003.

Li Q, Zhu Q, Xu M, Zhao Y, Narayan K, Liu Y. Estimating the Impact of COVID-19 on the PM2. 5 Levels in China with a Satellite-Driven Machine Learning Model. Remote Sensing 2021; 13: 1351.

Lu D, Mao W, Zheng L, Xiao W, Zhang L, Wei J. Ambient PM2. 5 Estimates and Variations during COVID-19 Pandemic in the Yangtze River Delta Using Machine Learning and Big Data. Remote Sensing 2021; 13: 1423.

Mallik S, Soni S, Podder K, Mishra U, Ahamed M. Prediction and assessment of change in PM2. 5 during COVID-19 lockdown using remote sensing and deep learning approach: A case study of Kanpur city. 2020.

Nath P, Saha P, Middya AI, Roy S. Long-term time-series pollution forecast using statistical and deep learning methods. Neural Computing and Applications 2021: 1-20.

Rahman MM, Paul KC, Hossain MA, Ali GMN, Rahman MS, Thill J-C. Machine Learning on the COVID-19 Pandemic, Human Mobility and Air Quality: A Review. IEEE Access 2021.

Roy S, Chatterjee A. The Dual Impact of Lockdown on Curbing COVID-19 Spread and Rise of Air Quality Index in India. The Impact of the COVID-19 Pandemic on Green Societies 2021: 113.

Ryan RG, Silver JD, Schofield R. Air quality and health impact of 2019–20 Black Summer megafires and COVID-19 lockdown in Melbourne and Sydney, Australia. Environmental Pollution 2021; 274: 116498.

Ryu J, Kim J. Demand Response Program Expansion in Korea through Particulate Matter Forecasting Based on Deep Learning and Fuzzy Inference. Energies 2020; 13: 6393.

Shahriar SA, Kayes I, Hasan K, Hasan M, Islam R, Awang NR, et al. Potential of arima-ann, arima-svm, dt and catboost for atmospheric pm2. 5 forecasting in bangladesh. Atmosphere 2021; 12: 100.

Shatnawi N, Abu-Qdais H. Assessing and predicting air quality in northern Jordan during the lockdown due to the COVID-19 virus pandemic using artificial neural network. Air Quality, Atmosphere & Health 2021; 14: 643-652.

Song Z, Bai Y, Wang D, Li T, He X. Satellite Retrieval of Air Pollution Changes in Central and Eastern China during COVID-19 Lockdown Based on a Machine Learning Model. Remote Sensing 2021; 13: 2525.

Talbot N, Takada A, Bingham AH, Elder D, Yee SL, Golubiewski NE. An investigation of the impacts of a successful COVID-19 response and meteorology on air quality in New Zealand. Atmospheric Environment 2021; 254: 118322.

Tan Z, Li X, Gao M, Jiang L. The Environmental Story During the COVID-19 Lockdown: How Human Activities Affect PM2. 5 Concentration in China? IEEE Geoscience and Remote Sensing Letters 2020.

Tiwari A, Gupta R, Chandra R. Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown. arXiv preprint arXiv:2102.10551 2021.

Tripathi K, Pathak P. Deep Learning Techniques for Air Pollution. 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). IEEE, 2021, pp. 1013-1020.

Wang Y, Wen Y, Wang Y, Zhang S, Zhang KM, Zheng H, et al. Four-month changes in air quality during and after the COVID-19 lockdown in six megacities in China. Environmental Science & Technology Letters 2020; 7: 802-808.

Wibowo F. Prediction of air quality in Jakarta during the COVID-19 outbreak using long short-term memory machine learning. IOP Conference Series: Earth and Environmental Science. 704. IOP Publishing, 2021, pp. 012046.

Zhan C, Li S, Li J, Guo Y, Wen Q, Wen W. Prediction of Air Quality in Major Cities of China by Deep Learning. 2020 16th International Conference on Computational Intelligence and Security (CIS), 2020, pp. 68-72.

ADDITIONAL ARTICLE II (APPENDIX B) ARTIFICIAL INTELLIGENCE-BASED BIT-WISE DECODING OF ERROR-CORRECTION CODES FOR RADIO COMMUNICATIONS

Sharma, E, Shakeel I, Sancho SS, & Deo, RC 2022, 'Artificial Intelligence-based Bit-wise Decoding of Error-correction Codes for Radio Communications', *IEEE Access*, Under review. (*Q1*, *Impact Factor: 3.367 H Index: 127, SNIP*: *1.421*, and *87th percentile in Engineering*).

B.1 Foreword

Appendix B presents an exact copy of the article under review in ***IEEE ACCESS.***

In this work, the candidate has developed several popular artificial intelligence-driven approaches and then investigated to decode error correction codes (ECC). A systematic and comprehensive comparative strategy has been undertaken to investigate the merits and demerits involving both machine and deep learning (ML/DL) technologies for bit-wise decoding of binary linear block codes. All the technologies considered are known for prediction and recognition problems. The paper considers the complexity of the problem, the error-rate performance, and the overall computational time. The efficacy of the comparison methodologies are assessed using the benchmark codes: the extended binary Golay and the Hamming code. The practicality of the study is advocated through performance evaluation using block error rate (BLER) and power efficiency metrics, including the illustrated error-rate simulations over varying levels of Signal-to-Noise Ratio (SNR) for the designated communication channel. The authors discuss our analysis through the correlation of SNR selections for model training, and the choice of comparative models from an ensemble of ML/DL models. In the proposed model setting, (LSTM) gives the best comparative results (showing BLER ≈ 0.175, ≈ 0.076, ≈ 0.0332, ≈ 0.0129, ≈ 0.0037, and ≈ 0.0003, with the Hamming code), and also reached the theoretical optimum (*BLER* ≈ 0.315, ≈ 0.0656, ≈ 0.0065, ≈ 0.0008, and ≈ 0.0002 for *SNR* 0-5 with the Golay code) compared to the *CLSTM, CNN, OS-ELM,* and *ELM* models. This makes the work a good starting point towards the possibility of future research for longer codes and implementation of the best practices for a robust and resilient architecture. This may offer a powerful toolkit to facilitate quick learning and feature analysis to autonomously generate robust receiver algorithms. The improvements made can enable reliable and uninterrupted radio communications for longer codes that may be challenging to decode efficiently with traditional error correction techniques.

B.2 Research Highlights

- To address the literature gap and propose a feasible, practical, and computationally efficient *ML/DL* architecture after comparing with several approaches to minimize loss function (*e.g.,* block error rate *BLER*).
- To develop several models for error correction that may be standalone such as extreme learning machine (*ELM*), online sequential extreme learning machine (*OS-ELM*), etc., or hybrid such as integrating a convolutional neural network (*CNN*) with *LSTM* generating a hybrid convolutional-*LSTM* (*C-LSTM*) method.
- Generate low error rate performance and secondly, evaluate all the algorithms through *BLER* metrics at various Signal to Noise Ratios (*SNR*) for the benchmark codes: Binary extended Golay Code (24,12) and Hamming Code (7,4).
- Suggest which of the comparative architecture may be considered by future studies and used effectively with longer error correction codes. The developed *ML/DL* architectures are designed by setting 'N' nodes at the output layer to make the decoder implementable for larger and relatively complex codes. The proposed framework after comparison is likely to be competitive with the current state-of-the-art framework in improving the transmission reliability of communication systems.
- To compare best practices and procedures by evaluating all *ML/DL* models' efficacy through visual and statistical metrics on tested data sets.
- To investigate whether any of the comparative architectures can offer a powerful unified toolkit to facilitate the development of computationally efficient and practically implementable *ML/DL*-based decoding algorithms for codes such as Polar codes and low-density parity-check (*LDPC*) codes that may be challenging to handle with the existent techniques in the literature.

B.3 Under Review Appendix Article II

DRAFT. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.DOI

Artificial Intelligence based Bit-wise Decoding of Error-correction Codes for Radio Communications

EKTA SHARMA[1], Member IEEE, ISMAIL SHAKEEL[1,2], Senior Member IEEE, SANCHO SALCEDO-SANZ[1, 3], RAVINESH C. DEO[1], Senior Member IEEE

[1]USQ's Advanced Data Analytics Lab, School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, Queensland 4300, Australia
[2]Contested Communications Branch, Cyber and Electronic Warfare Division, Defence Science and Technology Group, South Australia 5111
[3]Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Spain

Corresponding author: Ismail Shakeel (e-mail: ismail.shakeel@dst.defence.gov.au)

The research was supported by the Defence Science and Technology Group through the Australian Mathematical Sciences Institute Intern Program to the first author Ekta Sharma.

ABSTRACT Emerging demands for designing practical artificial intelligence (*AI*) algorithms for radio communications are the key drivers for the continued technology evolution in wireless communications. Recently, several state-of-the-art advancements have been introduced for error correcting codes *(ECC)* as they are used in practically all cases of message transmission. Growing radio communication literature reports that the complexity of artificial intelligence (*AI*) algorithms grows exponentially with the dimension of the error-correction codes. In this paper, several popular *AI* driven approaches have been developed and then investigated to decode error correction codes *(ECC)*. A systematic and comprehensive comparative strategy has been undertaken investigating merits and demerits involving both machine and deep learning (ML/DL) technologies for bit-wise decoding of binary linear block codes. All the technologies considered are known for prediction and recognition problems. The paper considers the complexity of the problem, the error-rate performance, and the overall computational time. The efficacy of the comparison methodologies is assessed using the benchmark codes: the extended binary Golay and the Hamming code. The practicality of the study is advocated through performance evaluation using block error rate (BLER) and power efficiency metrics, including the illustrated error-rate simulations over varying levels of Signal-to-Noise Ratios *(SNR)* for the designated communication channel. We discuss our analysis through the correlation of *SNR* selections for model training, and the choice of comparative models from an ensemble of *ML/DL* models. In the proposed model setting, *(LSTM)* gives the best comparative results (showing *BLER* ≈ 0.175, ≈ 0.076, ≈ 0.0332, ≈ 0.0129, $\approx 0.0037, and \approx 0.0003$, for *SNR* 0-6 with the Hamming code), and also reached the theoretical optimum *BLER* ≈ 0.315, ≈ 0.0656, ≈ 0.0065, ≈ 0.0008,, *and* ≈ 0.0002 for *SNR* 0-5 with the Golay code) compared to the *vs. CLSTM, CNN, OS-ELM, and ELM* models). This makes our work a good starting point towards the possibility of future researches for longer codes and implement the best practices for a robust and resilient architecture. This may offer a powerful toolkit to facilitate quick learning and feature analysis to autonomously generate robust receiver algorithms. The improvements made can enable reliable and uninterrupted radio communications for longer codes that may be challenging to decode efficiently with traditional error correction techniques.

INDEX TERMS Error-correcting codes; machine learning; AI-enabled wireless communication; high-dimensional codes; Golay code; Hamming code; linear block codes; deep learning.

I. INTRODUCTION

THE phenomenal growth in the number of wireless users and rapid development of wireless communication technologies demands a higher level of information reliability and computational efficiency. These demands have imposed new challenges that will be difficult to address using traditional

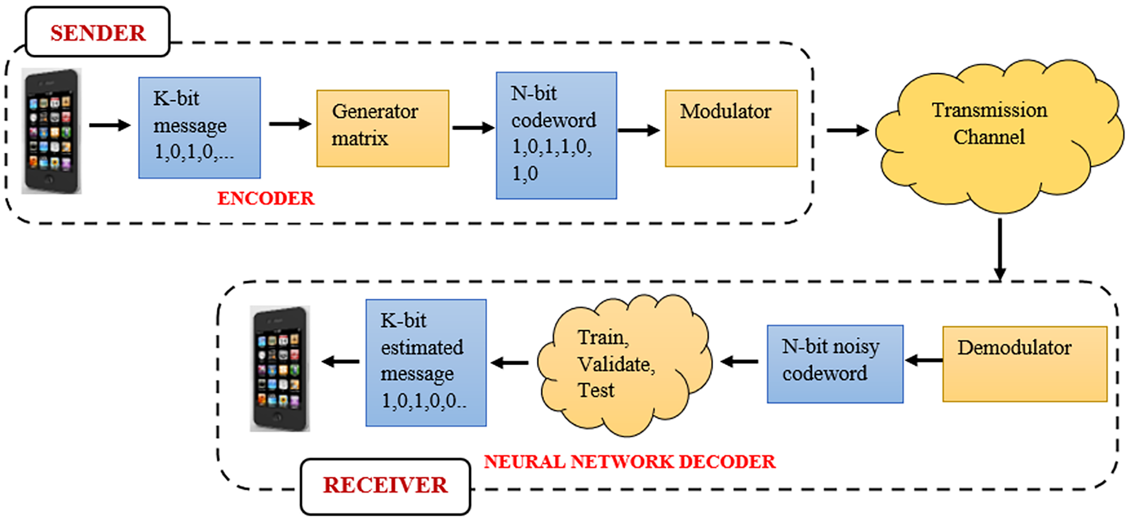


FIGURE 1. Illustrating the Communication system with ML/DL enabled decoder.Redrawn after [1]

methods due to their excessive complexity and sub-optimal performance [2], [3]. Advancements in Machine Learning *(ML)*, and in particular deep learning *(DL)*, have achieved groundbreaking success to stimulate interest in addressing such challenges [4], [5]

There are multiple and successful application areas of deep learning as the models based on *DL* have the potential to attain high performance compared with traditional approaches. They address issues of model-measurement errors, and can speed up the learning and model convergence. However, it is observed that as deep neural networks *(DNNs)* are optimized with limited parameters, passing the inputs through them only requires a limited number of operations, making *DL* computationally efficient [6]. However, as fewer works exist on deep learning in the field of radio communication, the design of efficient algorithms for radio communication systems has become an active area of research. Recent papers such as [7], [8] also emaphasise on the use of networks adopting AI methods for decoding error-correcting codes. The basic functionalities of a radio communication system with an *ML*-based decoder are illustrated in Figure 1.

Error correction codes *(ECCs)* ensure safe and reliable transmission of data over the noisy channel by detecting and correcting bit errors [9]. Error corrections codes can be categorised into two classes, namely, block codes and convolutional codes. Figure 2 shows the different types of error correction codes. Coloured boxes show the two benchmark codes considered in the research.

In this paper, we compared merits and demerits of several smart *ML/DL*-enabled *ECC* decoders for minimising the overall system error rate and improving transmission reliability. All algorithms are assessed using the benchmark codes, the binary extended Golay codes reference [10], and the Hamming code (linear*ECC*) [6].

Many communication applications require longer codes to achieve a target performance for reliable transmission of information [12]. Also, most of the *DL* methods in the literature formulate the problem as a high-dimensional classification problem where each unique word in the code is assigned a label. However, for high-dimensional codes (i.e., longer codes), the number of required labels increases making the implementation of the DL-based algorithms practically impossible. This approach has produced excellent results and appears to meet the expected theoretical performance levels [9]. This paper takes an alternative approach to address the increased decoding complexity for longer codes by formulating the problem as a binary classification problem for each decoded bit and then comparing it with other approaches. The literature [13], [14] shows that DNN methods have been applied for designing decoders, by enabling feature extraction through convolutional neural networks *(CNNs)* and Hamming code performance evaluations using an artificial neural network decoders.

Our motivation is to compare efficient data-driven decoder architectures harnessing the predictive ability of *ML/DL*. The ensemble of algorithms decoding block codes are compared in Section 2, and discussed in detail in the Appendix [15]. Studies such as [16] considered deep learning *(DL)* to improve the decoding of linear codes. To the best of the authors' knowledge, only one study [17] has explored the comparative approach on practically implementable architectures for decoding block error codes. Also, the authors only compared two of deep learning models *(CNN)* and *(LSTM)* in the context of radio communication. The novelty and scientific contributions of this study are, therefore:

i) To address the literature gap and propose a feasible, practical, and computationally efficient *ML/DL* architecture after comparing with several approaches to minimize loss function (*e.g.*, block error rate *BLER*).

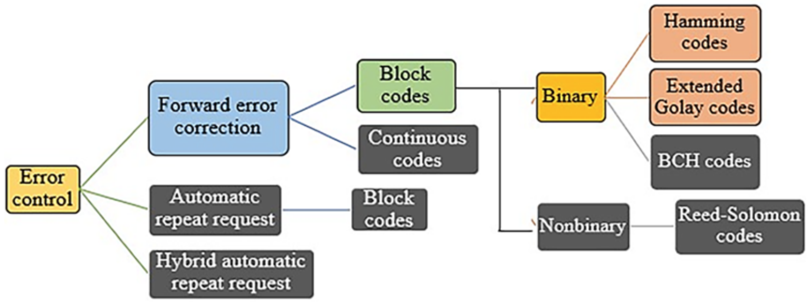


FIGURE 2. Types of Error control codes. The coloured boxes illustrate the benchmark codes considered in the study.Redrawn after [11]

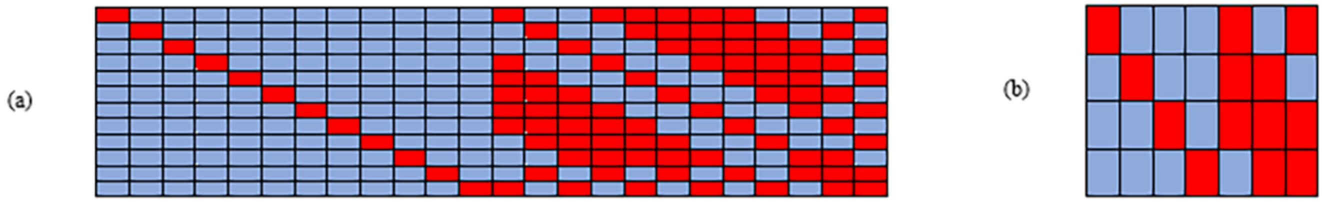


FIGURE 3. Generator Matrix with for 0 and red for 1 for (a) Binary Extended Golay Code, and (b) Hamming Code.

ii) To develop several models for error correction that may be standalone such as extreme learning machine *(ELM)*, online sequential extreme learning machine *(OS-ELM)* etc or hybrid such as integrating a convolutional neural network *(CNN)* with *LSTM* generating a hybrid convolutional-LSTM *(C-LSTM)* method.

iii) Generate low error rate performance and secondly, to evaluate all the algorithms through *BLER* metrics at various Signal to Noise Ratios *(SNR)* for the benchmark codes Binary extended Golay Code (24,12) and Hamming Code (7,4).

iv) Suggest which of the comparative architecture may be considered by future studies and used effectively with longer error correction codes. The developed *ML/DL* architectures are designed by setting 'N' nodes at the output layer to make the decoder implementable for larger and relatively complex codes. The proposed framework after comparison is likely to be competitive with the current state-of-the-art framework in improving the transmission reliability of communication systems.

v) To compare best practices and procedures by evaluating all *ML/DL* models' efficacy through visual and statistical metrics on tested data sets.

vi) To investigate whether any of the comparative architectures can offer a powerful unified toolkit to facilitate the development of computationally efficient and practically implementable*ML/DL*-based decoding algorithms for codes such as Polar codes and low-density parity-check *(LDPC)* codes that may be challenging to handle with the existent techniques in the literature. The rest of this paper is structured as follows:

Section 2 describes the comparative overview of several *ML/DL* models developed in this study. Section 3 introduces the materials and method, describing the detailed architecture of the models developed for comparison. This section ends by describing the evaluation criteria used to evaluate the models.

Next, the experimental details, results, validation process, and discussion of the model are presented and analysed in Section 4. The concluding remarks are presented in Section 5 with limitations and future research.

II. COMPARISON OF ALL METHODOLOGIES: CNN, LSTM, C-LSTM, ELM, OS-ELM

This paper focuses more on the comparison of the well-known methodologies. The details of these models *i.e CNN* [18], *LSTM* [19], *C-LSTM* [20], and traditional *ML* models such as Extreme learning machines *(ELM)* [21] and Online learning extreme learning machine *(OS-ELM)* [22] are presented in Appendix of this paper.

A. DEEP LEARNING METHODOLOGIES: CNN, LSTM, C-LSTM

Convolutional neural network *(CNN)* is a famous and an advanced neural network, that has gained popularity especially in the digital image recognition area. The authors in [23] explores *CNN* to recognize the cognitive radio waveform in an automatic system. Their work concludes that it is an effective model when high noise signals need to be differentiated. Other works such as [24] finds out that *CNN* proves to have strong capability of classification especially in the wavelet denoising technology and needs fewer parameters to learn with reduced chance of overfitting and a better prediction in data based on images. However, the study [25] finds out that *CNN*-based modulation classification performs better in cases when the input size is fixed. In reality this is not possible as the actual radio signal is of variable nature. A model based on *CNN* requires a lot of training data and is less efficient for encoding, suffering from the lack of ability to be invariant to the input data.

Another framework that has gained immense popularity in recent years is Long short-term memory network *(LSTM)*. It is designed to work efficiently and differently as compared to *CNN*, by automatically extracting robust features from noisy radio signals, and using learnt features for modulation. The authors [26] uses this capability implemented on a computational platform that exceeds state-of-the-art accuracy while staying low-cost. The results of this study on over-the-air radio data shows that the *LSTM* model classifies incoming radio signals efficiently, and reliably with superior performance, when compared with other state-of-the-art technologies. However some works with *LSTM* such as [27] concludes that the *LSTM* needs more parameters as compared to *CNN* and often takes more training time. However, its important significance of checking long input sequences without increasing the size of network outshines its shortcoming.

The authors in [28] efficiently explore the feature interaction, and the properties of raw complex temporal signals through another deep learning methodology, named *CNN-LSTM*. It models non-linear patterns in data, where *CNN* extracts complex features from data, and *LSTM* can model temporal information in data. The radio communication research with *C-LSTM* is in developing stages with very few studies to have used it and citeczech2018cnn being one of them.

B. MACHINE LEARNING METHODOLOGIES: ELM, OS-ELM

ELM and OS-ELM are both popular methodologies where the former has attracted several papers in radio communication domain such as [29] due to the benefit of usage, quick learning speed, simplicity, and good generalization performance.

OS-ELM is another fast and accurate online sequential learning algorithm for single hidden layer feed forward networks used by studies such as [22] where the authors ensured the continuous operation of the wireless sensor networks.

C. DEEP LEARNING (DL) VS. MACHINE LEARNING (ML) METHODOLOGIES

Deep learning is part of a broader family of machine learning methods however in comparison, deep learning models are fast becoming a popular choice [15], [30]. With the start of modelling, *DL* models reach their best predictive power making them robust with their hyper parameters. This enables enables them to be a preferred choice for run time predictions. In contrast, *ML* parses the data, learns from it, and makes decisions that are informed. Even though, *DL* methods run the problem of requiring more training data, taking longer time to train (generally on GPU), while *ML* trains on CPU with lesser training data, and hence lesser time to train. Lastly, *DL* can be tuned in several ways, while *ML* has limited tuning capability for hyper parameter tuning. These arguments show that even though *(DL)* is a subset of *(ML)*, it has a wider range of capabilities and can handle more complex tasks than machine learning.

TABLE 1. CODES USED TO ASSESS THE PROPOSED ALGORITHMS.

Name	Distance	Total (N)	Bits Data (K)	Parity (N-K)	Rate (K/N)	Notation
Extended Golay (24,12)	8	24	12	12	0.5	G_{24} OR $[24, 12, 8]_2$-code
Hamming (7,4)	3	7	4	3	0.57	$[7, 4, 3]_2$-code

III. MATERIALS AND METHOD

A. RESEARCH DATA

For the data used, the transmitted message is reconstructed at the receiver side over a physical channel using the *ML/DL*-enabled decoder as shown in Figure1. In this illustration, we note that the transmitter is expected to first encode the randomly generated K-bit message to an N-bit code word, which is then modulated using Binary Phase Shift Keying before transmitting over a noisy channel. An additive white Gaussian noise *(AWGN)* model is considered for the channel. On the receiver end, the soft output N-values from the demodulator go to the decoder where the K-bit message is retrieved. In the proposed approach, the decoder is represented as a Deep Neural Network *(DNN)*.

The *DNN* decoder reconstructs the original message with a low error rate by learning from the received soft values. The entire communication system is then trained to achieve the system performance target in terms of the bit error rate *(BER)* and the block error rate *(BLER)* [15]. The training data comprises the noisy soft-values of the channel output and the expected message sequences generated from the communication system.

It is noteworthy that the present research aims to first develop and then compare several popular artificial intelligence driven approaches and aims to find a receiver-decoding algorithm that has potential for reliable communications over an *AWGN* channel.All the algorithms are evaluated by generating *BLER* metrics at various Signal to Noise Ratios *(SNR)* for the benchmark codes: Binary extended Golay Code (24,12) and Hamming Code (7,4). Table 1 shows more detail on the benchmark codes. The Hamming code is well known for its single-bit error correction and two-bit error detection capability [14]. The (N=7, K=4) linear Hamming block code is based on the principle of mapping 4-bit input data to a 7-bit output codeword using a generator matrix with order G (4x7). The binary extended Golay code (N=24, K=12) has

12 message bits with a code-word length of 24 bits. The Golay code encodes 12 bits of data to a 24-bit word using a generator matrix with order G (12x24). The Golay code can correct up to 3-bit errors and detect up to 7-bit errors. The parameter *'K'* is the dimension of the code and the generator matrix in a standard systematic form is shown in Figure 3. Note that is for 0 and red stands for 1 for both these coding schemes.

B. DEVELOPMENT OF MODEL ARCHITECTURE

An ensemble of five (conventional machine learning and deep learning) models was developed in this paper. The deep and non-deep learning methods are: *ELM, OS-ELM, CNN, LSTM*, and hybrid architecture in this paper is, *C-LSTM*. All of the models were constructed using a Windows 10 Intel®*i*7 Generation 10 platform @3.8 gigahertz unit, 32 GB memory platform using Python programming language, and open-source libraries (*i.e.*, Tensor Flow [31], Keras [32], and Scikit-learn [33]). To develop *LSTM* architecture, randomly generated data were partitioned. Without any standard rule for data partitioning the data for both Hamming and Golay codes were partitioned in 60 % (training), 6.67 % (validation), and 33.3 % (testing) subsets. Data were normalised to lie within the interval [0, 1]:

$$RI_{NORM} = (RI - RI_{MIN})/(RI_{MAX} - RI_{MIN}) \quad (1)$$

In 1 RI_{MIN}, = The minimum values of message inputs, RI_{MAX}, = The maximum values of message inputs received by the Receiver (refer to Figure 1), and RI_{NORM} = The normalised Receiver input.

Table 2 shows an optimal architecture in Red colour with our comparative approaches benchmarked with the model that performed best amongst all. This study adopted a grid search method as an important tool to extract the best features [34] and optimise the resulting model. In Table 2, we show how an optimum framework can be attained by optimising the model hyper-parameters. For a specific *SNR*, testing was continued for a reasonably good batch size starting from 50 and reaching 3000 until an optimal performance was achieved as recommended by the theory w.r.t. each *SNR*. Thereafter each model performance was compared with others. In determining a feasible architecture, the optimal number of hidden neurons plays a significant role [35]. Notably, another important consideration in the model formulation that the authors have considered is over fitting [36]. It is noteworthy that under-fitting can occur when an *ML* framework fails to converge in the desired time and the model framework lacks an appropriate degree of freedom to fully extract the most pertinent data patterns. The grid search adopted in the study aimed to overcome this issue and to enable model learning. The lowest mean absolute error *(MAE)* or root mean square error *(RMSE)* (training), as obtained by a sequence of hidden neurons in gradual steps was adopted to determine the optimal architecture. As activation functions *(AFs)* play a significant role in the robustness of the model being constructed [37] we have used *AFs* based on the *ReLU*, sigmoidal, SoftMax, and hyperbolic tangent equations, which resulted in the best performance for an *LSTM* using the *'ReLU' AF*.

It is imperative to mention that performing a grid search of hyper-parameters can be time-consuming [38] with each model taking ∼ 10 to 15 hours. However, after deducing the optimal parameters, the computation training and testing time for our model was reduced to less than 20 minutes. The following model parameters were modified as part of the optimisation task:

- Absolute Deviation and Least Square Error ($L1$, $L2$-regularisation): Here, we penalised parameters to minimise the sum of the square of differences between the observed.
- Early Stopping *(ES)*: Here, we used this method to avoid overfitting through Kera's *DL* library where the patience setting was '45' and the mode setting was 'minimum'.
- Activation Function *(AF)*: Several activation functions (sigmoidal, SoftMax, leaky_*ReLU*, tanh, and *ReLU* (Rectified Linear Unit) have been tested with the *ReLU AF* as the optimal choice.
- Dropout: Here, we used a regularisation feature together with *ES* which helped our *LSTM* to reduce its overfitting behaviour. Our study used a dropout value of 0.1.

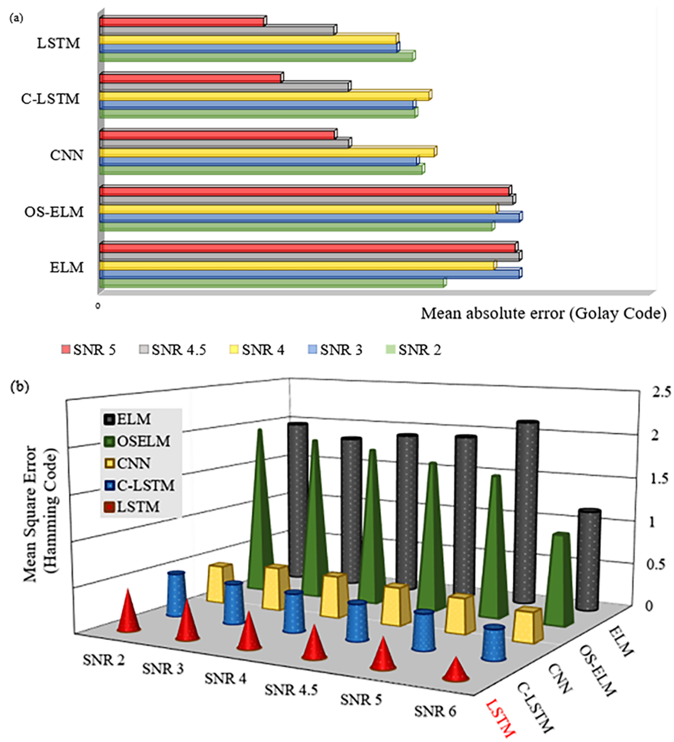


FIGURE 4. (a) Test performance of all comparative models via. mean absolute error (MAE) through Golay Code. (b) 3D- Bar graph of mean square error (MSE) in terms of Hamming Code. Evidently LSTM performance is better.

C. PERFORMANCE BENCHMARKS

All the developed models in comparison were assessed using statistical performance metrics based on the block error rate

TABLE 2. OPTIMAL ARCHITECTURE (IN RED FOR *BLER*.

Model	*Hyper-parameters*	*Optimal set grid search*
LSTM	Epochs	[50, 500, 700, 1000, 3000]
	Activation function	[SoftMax, *Tanh*, *ReLU*, Leaky *ReLU*, *sig*]
	Batch size	[100, 500, 700, 1000, 2000]
	Optimiser, Drop rate	[*Adam*], [0.1, 0.2]
	LSTM filter	[50, 60, 100, 200, 500]
	Dropout	Yes
	SNR	[0,2,3,4,4.5,5,6]
	Platform	Python (version 3.8.5)
C-LSTM	Epochs	[50, 500, 700, 1000, 3000]
	Activation function	[SoftMax, *tanh*, *ReLU*, Leaky_*ReLU*, *sig*]
	Batch size	[400, 500, 800, 750, 1000, 2000]
	Pooling, Padding	[2], [same]
	Layer 1	[250, 200, 150, 100, 80]
	Layer 2	[20, 40, 50, 60, 70, 80]
	Layer 3	[30, 10, 15, 20, 5]
	Activation function	[SoftMax, *tanh*, *ReLU*, Leaky_*ReLU*, *sig*]
	Dropout	Yes
	SNR	[0,2,3,4,4.5,5,6]
	Platform	Python (version 3.8.5)
CNN	Epochs	[50, 500, 700, 1000, 2000]
	Activation function	[SoftMax, *tanh*, *ReLU*, Leaky_*ReLU*, *sig*]
	Batch Size	[100, 500, 700, 1000, 2000]
	Optmiser, Drop rate	[Adam], [0.1, 0.2]
	CNN filter	[50, 60, 100, 200, 500]
	Pooling, Padding	[same]
	Number of Dense layers used	[3 for Input and 1 for Output]
	Dropout	Yes
	SNR	[0,2,3,4,4.5,5,6]
	Platform	Python (version 3.8.5)
OS-ELM	Epochs	[500, 1000, 5000, 10000, 50000]
	Activation function	[tansig, SoftMax, *tanh*, *ReLU*, *sig*]
	Hidden Neurons	[50, 100, 200, 500, 1000]
	Regularisation factor	[50, 100, 500]
	SNR	[0,2,3,4,4.5,5,6]
	Platform	*MATLAB* R2020b
ELM	Epochs	[500, 1000, 5000, 10000, 50000]
	Activation function	[tansig, SoftMax, *tanh*, *ReLU*, *sig*]
	Hidden Neurons	[50, 100, 200, 500, 1000]
	Regularisation factor	[50, 100, 500]
	SNR	[0,2,3,4,4.5,5,6]
	Platform	*MATLAB* R2020b

Backpropagation Algorithm - Standard Architecture	
Beta β_1, β_2	0.990
Alpha, α	0.001
Epsilon, ϵ	0.0000001

$\beta_1, \beta_2 = 1^{st}, 2^{nd}$ Decay rates (exponential moment estimates)
α = Rate of learning, ϵ = Small value to avoid division by zero.

(BLER), mean square error *(MSE)*, root-mean-square-error *(RMSE)*, and mean absolute error *(MAE)*, as follows:

1) Block error rate $(BLER)$ =

$$\text{Number of Block Errors/Number of Test messages} \tag{2}$$

2) Mean square error (MSE) =

$$\frac{1}{N}\sum_{i=1}^{N}\left(msg_i^{SIM} - msg_i^{OBS}\right)^2 \tag{3}$$

3) Mean absolute error (MAE) =

$$\frac{1}{N}\sum_{i=1}^{N}\left|\left(msg_i^{SIM} - msg_i^{OBS}\right)\right| \tag{4}$$

4) Root mean square error $(RMSE)$ =

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(msg_i^{SIM} - msg_i^{OBS}\right)^2} \tag{5}$$

where,
msg_i^{SIM} = simulated message block for the i^{th} message,
msg_i^{OBS} = observed message block for the i^{th} message,
N = Total number of message blocks.

IV. RESULTS AND DISCUSSIONS

This section analyses the obtained results and evaluates the performance of the model, demonstrating the efficiency of the model that performs best after comparing all developed architectures *(i.e. LSTM)* All the competing approaches, that were developed *e.g., CNN, C-LSTM, OS-ELM*, and *ELM* have their results compared and benchmarked. Equations 2-5 suggest further appraisal of the model through various statistical and evaluation metrics.

Tables3 and 4 show the segregation of data in all models when their performance is assessed through a Hamming and Golay code at different Signal to Noise ratios *(SNR)* to calculate the Block Error Rate *(BLER)*. The most accurate performance is shown in red with the authors testing the model performance from *SNR* 0, 2, 3, 4, 4.5, 5, and 6 *dB*, to assess the Hamming code. For the case of the Golay code, the model performance is tested from *SNR* 0, 2, 3, 4, 4.5, and 5 *dB* with the entire data divided into training, validation, and testing subsets. Table3 compares the results of all the *SNR*s. It is observable that out of all models developed, this research work attains the lowest training time (in seconds) for the *LSTM* architecture *w.r.t* all six *SNRs* ≈ 0.0101s (seconds), ≈ 0.0119s, ≈ 0.0468s, ≈ 0.0997s, ≈ 0.1025s, ≈ 0.0359s, and ≈ 0.0420s for Hamming assessment. This is also the case for validation and testing times. For example, the testing time for *LSTM* was ≈ 0.0029s, ≈ 0.0049s, ≈ 0.0089s, ≈ 0.0158s, ≈ 0.0278s, ≈ 0.0238s, and ≈ 0.0436s.

In terms of overall computation time, *LSTM* was most efficient followed by *OS-ELM, ELM, CNN, and C-LSTM* for higher *SNR's i.e.,* $4-6$ *dB*. For lower *SNRs i.e.,* 0, 2, *and* 3 *dB, LSTM's* computation time performance was most efficient followed by other *DL* models *CNN, C-LSTM, OS-ELM, and ELM*. However, in Table5 (algorithm assessment *via* Golay code), we note a subtle variation for lower *SNR's i.e.,* 0- 3 *dB*. The minimum training and testing time has been achieved by *OS-ELM* (≈ 0.2409*s* - ≈ 0.3162*s*, and testing time ≈ 0.1302*s*, ≈ 0.1632*s*), except the validation time, which is the lowest overall for *LSTM* (≈ 0.1298*s*, ≈ 0.1037*s*, ≈ 0.1007*s*, ≈ 0.7051*s*, ≈ 0.1190*s, and* ≈ 0.1920*s*).

All *DL* models performed better in terms of computation times for higher *SNRs*, with *LSTM* followed by *OS-ELM*,

TABLE 3. OPTIMAL ARCHITECTURE (IN RED) FOR *BLER*.

SNR (per information bit) (dB)	Optimum BLER (Theoretical)	Designed Predictive Models	Total Messages	Number of messages			Time to generate messages (seconds)		
				Training	Validation	Testing	Training	Validation	Testing
0	0.1757	LSTM	6,000	3,600	400	2,000	0.0101	0.0039	0.0029
		C-LSTM	6,000	3,600	400	2,000	0.0324	0.0048	0.0065
		CNN	6,000	3,600	400	2,000	0.0207	0.0049	0.0043
		OS-ELM	6,000	3,600	400	2,000	0.1007	0.0135	0.0161
		ELM	6,000	3,600	400	2,000	0.1251	0.0257	0.0263
2	0.0722	LSTM	12,000	7,200	800	4,000	0.0119	0.0049	0.0049
		C-LSTM	12,000	7,200	800	4,000	0.0513	0.0059	0.0079
		CNN	12,000	7,200	800	4,000	0.0309	0.0059	0.0059
		OS-ELM	12,000	7,200	800	4,000	0.1217	0.0263	0.0216
		ELM	12,000	7,200	800	4,000	0.1635	0.0365	0.0326
3	0.0346	LSTM	26,000	15,600	2,400	8,000	0.0468	0.0059	0.0089
		C-LSTM	26,000	15,600	2,400	8,000	0.0498	0.0069	0.0199
		CNN	26,000	15,600	2,400	8,000	0.0279	0.0069	0.0029
		OS-ELM	26,000	15,600	2,400	8,000	0.0910	0.0237	0.0302
		ELM	26,000	15,600	2,400	8,000	0.0975	0.0419	0.0534
4	0.0122	LSTM	96,000	57,600	6,400	32,000	0.0997	0.0169	0.0158
		C-LSTM	96,000	57,600	6,400	32,000	0.1467	0.0179	0.0297
		CNN	96,000	57,600	6,400	32,000	0.1585	0.0249	0.0498
		OS-ELM	96,000	57,600	6,400	32,000	0.0935	0.0308	0.0415
		ELM	96,000	57,600	6,400	32,000	0.2900	0.0673	0.0488
4.5	0.00647	LSTM	132,000	79,200	8,800	44,000	0.1025	0.0249	0.0278
		C-LSTM	132,000	79,200	8,800	44,000	0.4125	0.0877	0.1443
		CNN	132,000	79,200	8,800	44,000	0.4555	0.0608	0.1428
		OS-ELM	132,000	79,200	8,800	44,000	0.1335	0.0269	0.0316
		ELM	132,000	79,200	8,800	44,000	0.1729	0.0586	0.0768
5	0.00349	LSTM	144,000	86,400	9,600	48,000	0.0359	0.0151	0.0238
		C-LSTM	144,000	86,400	9,600	48,000	0.1518	0.0896	0.0544
		CNN	144,000	86,400	9,600	48,000	0.2328	0.0301	0.0720
		OS-ELM	144,000	86,400	9,600	48,000	0.0382	0.0249	0.0373
		ELM	144,000	86,400	9,600	48,000	0.1129	0.0387	0.0792
6	0.000813	LSTM	300,000	180,000	20,000	100,000	0.0420	0.0298	0.0436
		C-LSTM	300,000	180,000	20,000	100,000	0.5479	0.1136	0.2322
		CNN	300,000	180,000	20,000	100,000	0.3317	0.0453	0.1117
		OS-ELM	300,000	180,000	20,000	100,000	0.0733	0.0685	0.1424
		ELM	300,000	180,000	20,000	100,000	0.2408	0.0805	0.0960

ELM, CNN, C-LSTM. Hybrid *C-LSTM* gave better performance than *OS-ELM, ELM, CNN at SNR* 5 *dB (*$\approx$ 3.0131 *seconds)*. For our work to further show the efficacy of our proposed architecture, *LSTM*; Table5,6, and Figure5) are very important. Table6 illustrates results assessed through Golay code, and Figure5) show *BLER vs. SNR* comparison through visual plots. We notice the lowest errors and block error rate *(BLER)* for all *SNRs (*$\approx$ 0.0760, $\approx$ 0.3975, $\approx$ 0.0129, $\approx$ 0.0037, *and* $\approx$ 0.0003 *for SNR* 04-6 *dB)* in the part (a) of Figure5). Similarly, part (b) shows the superiority of *BLER vs. SNR* with Golay code assessment. As *DL* is very powerful, it is not a surprise to see all *DL* models performing well, especially for lower *SNRs i.e.,* 0-3 *dB*, however, with *LSTM* outperforming all comparative approaches for all *SNR's* and reaching closest to the theoretical optimum *(*$\approx$ 0.0656, $\approx$ 0.0181, $\approx$ 0.0065, $\approx$ 0.0008, *and* $\approx$ 0.0002 *for SNR* 0-5 *dB)*. However, for higher *SNRs, LSTM* is followed by *C-LSTM, CNN, OS-ELM, and ELM*. *CNN* showed an impressive performance at *SNR* 4*dB* $\approx$ 0.0086 but fell short when the signal-to-noise ratio increased. It is noteworthy that traditional *ML* was lagging more significantly than *DL* models. The efficacy of the proposed model: *LSTM* is also tested through 3-*D* Bar graphs in Figure4). Here the visual comparison attempts to illustrate the test performance of *LSTM* vs. benchmark models *(i.e., C-LSTM, CNN, OS-ELM, and ELM)* through the mean absolute error *(MAE) via* Golay Code and mean square error *(MSE)* in terms of assessment *via* Hamming code for Figure4. Both illustrations show the superiority of the deep learning models. The performance of *C-LSTM and CNN* for lower *SNRs* is equally strong, however, the hybrid algorithm performed better for *MAE* for higher *SNRs (C-LSTM* $\approx$ 0.4177, $\approx$ 0.3154, $\approx$ 0.2297, *and* $\approx$ 0.4586 *at SNR* 4,4.5, *and* 5 *dB*, as compared to *CNN* $\approx$ 0.4242, $\approx$ 0.3159, *and* $\approx$ 0.2979*)*. The proposed *LSTM* outperforms all the other algorithms for higher *SNRs (MAE* $\approx$ 0.3970, $\approx$ 0.3771, $\approx$ 0.3760, $\approx$ 0.2972, *and* $\approx$ 0.2081 *for SNR* 2,3,4,4.5, *and* 5 *dB* with Golay code assessment).

A subtle variation is observed in model accuracy when we compute the *RMSE* and the *MSE*. For instance, for *SNR* 2 *dB*, we notice, *CNN's MSE and RMSE* were the lowest *(MSE* $\approx$ 0.4291, *RMSE* $\approx$ 0.6550, *as compared to LSTMs (MSE* $\approx$ 0.4714, *RMSE* $\approx$ 0.6865*)*. Despite this observation, the proposed *LSTM* produced the lowest *BLER (*$\approx$ 0.0760, *and MAE* $\approx$ 0.4738*)*, making it the superior model as compared

TABLE 4. TESTING PERFORMANCE OF ALL COMPARATIVE MODELS FOR HAMMING CODE. RED = BEST PERFORMANCE.

SNR (per information bit) (dB)	Optimum BLER (Theoretical)	Designed Predictive Models	Batch Size	Epochs	Errors	Test messages	BLER	MSE	RMSE	MAE
0	0.1757	LSTM	700	1,000	344	2,000	0.1720	0.4804	0.6931	0.4759
		C-LSTM	700	1,000	547	2,000	0.2735	0.4955	0.7039	0.4895
		CNN	700	1,000	807	2,000	0.4035	0.4991	0.7064	0.4990
		OS-ELM	NA*	1,000	547	2,000	0.2735	1.9729	1.4046	0.5377
		ELM	NA	1,000	946	2,000	0.4730	1.9925	1.4115	0.6314
2	0.0722	LSTM	700	1,000	304	4,000	0.0760	0.4714	0.6865	0.4738
		C-LSTM	700	1,000	333	4,000	0.0832	0.4815	0.6939	0.4872
		CNN	700	1,000	306	4,000	0.0765	0.4291	0.6550	0.4975
		OS-ELM	NA*	1,000	382	4,000	0.0955	1.9582	1.3993	0.5346
		ELM	NA	1,000	1099	4,000	0.2747	1.9605	1.4001	0.6291
3	0.0346	LSTM	700	1,000	106	8,000	0.0332	0.4426	0.6652	0.4705
		C-LSTM	700	1,000	292	8,000	0.0365	0.4451	0.6671	0.4734
		CNN	700	1,000	526	8,000	0.0657	0.4862	0.6972	0.4864
		OS-ELM	NA*	1,000	362	8,000	0.0452	1.8658	1.3659	0.5335
		ELM	NA*	1,000	1526	8,000	0.1908	1.8164	1.3477	0.6236
4	0.0122	LSTM	700	1,000	413	32,000	0.0129	0.4003	0.6326	0.4682
		C-LSTM	700	1,000	589	32,000	0.0183	0.4242	0.6513	0.4715
		CNN	700	1,000	1401	32,000	0.0430	0.4667	0.6831	0.4776
		OS-ELM	NA*	1,000	720	32,000	0.0225	1.7946	1.3396	0.5340
		ELM	NA*	1,000	2932	32,000	0.0916	1.8952	1.3766	0.5905
4.5	0.00647	LSTM	700	1,000	368	44,000	0.0083	0.3528	0.5939	0.4602
		C-LSTM	700	1,000	440	44,000	0.0100	0.3990	0.6316	0.4685
		CNN	700	1,000	1540	44,000	0.0350	0.4226	0.6500	0.4799
		OS-ELM	NA*	1,000	455	44,000	0.0153	1.6906	1.3002	0.5322
		ELM	NA*	1,000	2474	44,000	0.0612	1.9191	1.3853	0.5827
5	0.00349	LSTM	700	1,000	179	48,000	0.0037	0.3226	0.5679	0.4587
		C-LSTM	700	1,000	327	48,000	0.0068	0.3873	0.6223	0.4609
		CNN	700	1,000	1344	48,000	0.0280	0.3916	0.6257	0.4704
		OS-ELM	NA*	1,000	432	48,000	0.0090	1.5945	1.2627	0.5241
		ELM	NA*	1,000	2245	48,000	0.0468	2.1281	1.4588	0.5736
6	0.000813	LSTM	700	1,000	70	100,000	0.0007	0.2226	0.4718	0.4495
		C-LSTM	700	1,000	207	100,000	0.0020	0.3214	0.5669	0.4586
		CNN	700	1,000	1096	100,000	0.0109	0.3319	0.5761	0.4684
		OS-ELM	NA*	1,000	290	100,000	0.0029	1.0072	1.0035	0.4991
		ELM	NA*	1,000	1539	100,000	0.0154	1.1504	1.0725	0.5742

to the rest. It also agrees with Table6, using the Golay code assessment, that *LSTMs* performance kept on improving with an increase in *SNRs*. This is followed by the lowest *MSE* *($\approx$ 0.4965, $\approx$ 0.4774, $\approx$ 0.4675, $\approx$ 0.4267, $\approx$ 0.4158, and* $\approx$ 0.3975 *for SNR* 0,2,3,4,4.5, *and* 5 *dB)*, lowest *RMSE* *($\approx$ 0.7045, $\approx$ 0.6887, $\approx$ 0.6837, $\approx$ 0.6532, $\approx$ 0.6448, and* $\approx$ 0.6304 *for SNR* 0,2,3,4,4.5, *and* 5 *dB)*, and lowest *MAE* as has been discussed earlier. We have included four further plots showing the performance evaluation of *LSTM*. Figure6) evaluate *SNR* 6 *dB* when assessed by the Hamming code in the test phase. Figure7) evaluate *SNR* 5 *dB* when assessed by the extended Golay code in the test phase.

Summarising the discussion, we affirm that comparing all the algorithms the deep learning architecture, *LSTM* provided a comparatively precise performance *w.r.t.* all performance metrics considered, *i.e.*, lowest *MAE, MSE, RMSE,* and most importantly, *BLER* at all Signal to noise ratios, when assessed by both the Hamming and binary extended Golay codes. As evident from the results obtained, we conclude that modeling performed through a non- *DL* approach *(e.g., OS-ELM, or ELM)* is likely to result in inferior performance in comparison with the model *(e.g., LSTM)* or other benchmark models *(C-LSTM and CNN)*. Therefore, after a comprehensive comparison we have provided a compelling evidence, establishing the architecture *LSTM* model as a dependable, and computationally efficient framework for decoding error-correction codes at various *SNRs*. The credible and smart framework for modelling error rate calculation might also find a prominent application in the radio communication settings with further complex channel coding schemes. Conclusively, *LSTM* has sufficiently proven an intelligent architecture for developing efficient decoders that can operate at different *SNRs* and henceforth, can be considered as a practical tool for radio communication systems.

V. CONCLUSIONS

Although *DL*, is an emerging inter-disciplinary paradigm, not many papers have so far been published on practically implementable architectures for decoding block error codes for radio communications.

The application of *DL* for decoding error correction codes presents a challenging research field that is still in its early stage. Although previous studies have shown promising results, some extensive challenges are worth exploring further. This research aimed to bridge the gap in the application of deep learning to the physical layer by designing practically

TABLE 5. DATA SEGREGATION AND COMPUTATION TIME ASSESSMENT FOR GOLAY CODE. RED = BEST PERFORMANCE

SNR (per information bit) (dB)	Optimum BLER (Theoretical)	Designed Predictive Models	Total Messages	Number of messages			Time to generate messages (seconds)		
				Training	Validation	Testing	Training	Validation	Testing
0	0.312 or 3.12e-2	LSTM	100,000	60,000	6,670	33,330	0.4327	0.1298	0.1985
		C-LSTM	100,000	60,000	6,670	33,330	4.4567	0.9199	2.0933
		CNN	100,000	60,000	6,670	33,330	0.4583	0.8576	0.1980
		OS-ELM	100,000	60,000	6,670	33,330	0.2409	0.2868	0.1762
		ELM	100,000	60,000	6,670	33,330	0.2679	0.0820	0.1693
2	0.0575 or 5.75e-2	LSTM	133,300	84,000	9,300	40,000	0.4627	0.1037	0.1845
		C-LSTM	133,300	84,000	9,300	40,000	4.4667	0.9171	2.0683
		CNN	133,300	84,000	9,300	40,000	0.4694	0.8577	0.1880
		OS-ELM	133,300	84,000	9,300	40,000	0.2508	0.2788	0.1302
		ELM	133,300	84,000	9,300	40,000	0.2509	0.0630	0.1439
3	1.45e-2 or 0.0145	LSTM	150,000	90,000	10,000	50,000	0.5024	0.1007	0.2243
		C-LSTM	150,000	90,000	10,000	50,000	1.0477	0.2009	0.4408
		CNN	150,000	90,000	10,000	50,000	3.9492	0.6041	1.4497
		OS-ELM	150,000	90,000	10,000	50,000	0.3162	0.5713	0.1632
		ELM	150,000	90,000	10,000	50,000	0.3702	0.0744	0.2530
4	220e-3 or 0.0022	LSTM	198,000	118,000	13,200	66,000	3.0106	0.7051	1.0205
		C-LSTM	198,000	118,000	13,200	66,000	7.7605	1.5699	3.0029
		CNN	198,000	118,000	13,200	66,000	3.9335	0.6383	1.4215
		OS-ELM	198,000	118,000	13,200	66,000	3.3899	0.9939	1.2945
		ELM	198,000	118,000	13,200	66,000	3.3945	0.7947	1.1925
4.5	7.12e-4 or 0.000712	LSTM	2,200,000	1,32,0000	160,000	720,000	3.0276	0.1190	1.0218
		C-LSTM	2,200,000	1,32,0000	160,000	720,000	5.0543	1.4645	2.2012
		CNN	2,200,000	1,32,0000	160,000	720,000	4.8155	1.0287	1.7886
		OS-ELM	2,200,000	1,32,0000	160,000	720,000	3.2473	1.7334	2.8142
		ELM	2,200,000	1,32,0000	160,000	720,000	3.2537	0.1778	2.8173
5	1.97e-4 or 0.000197	LSTM	2,410,000	1,44,0000	170,000	800,000	4.8743	0.1920	0.8440
		C-LSTM	2,410,000	1,44,0000	170,000	800,000	7.3898	1.4354	3.0131
		CNN	2,410,000	1,44,0000	170,000	800,000	16.9530	3.0382	6.1487
		OS-ELM	2,410,000	1,44,0000	170,000	800,000	7.3768	1.8027	3.8440
		ELM	2,410,000	1,44,0000	170,000	800,000	7.4852	1.1928	3.8735

implementable decoders and assessing performance using the binary extended Golay code, which has not been reported in earlier literature. Summarising our comparison notes:

i) The work has focussed on exploring and comparing an ensemble of classification-based and sequential machine learning *(ML)* and deep learning *(DL)* algorithms: *ELM, OS-ELM, CNN, LSTM, and C-LSTM,* and studied the suitability of using these models for developing efficient decoders for radio communication that use longer block codes.

ii) The performance of the developed algorithms were analysed through simulation and compared with theoretical performances.

iii) Of all the models designed and and compared, the study finds out and highlights a linear classification deep model that operates sequentially and gives better results when benchmarked *(i.e., LSTM)*.

iv) The *LSTM* model attained optimal theoretical performance for the benchmark codes. The *LSTM's* simulations registered the lowest *BLER* ($\approx$ 0.175, $\approx$ 0.076, $\approx$ 0.0332, $\approx$ 0.0129, $\approx$ 0.0037, *and* $\approx$ 0.0003 *for SNR* 0-6*)*. *LSTM* outperformed all comparative approaches for all *SNR's* and reached the theoretical optimum ($\approx$ 0.315, $\approx$ 0.0656, $\approx$ 0.0065, $\approx$ 0.0008, *and* $\approx$ 0.0002 *for SNR* 0-5*)* which was the lowest *vs. CLSTM, CNN, OS-ELM, and ELM* models respectively.

v) *LSTM* can be classified as a better comparative model of the ensemble on the basis of the lowest *MAE*, lowest *RMSE* when assessed using the Hamming and Golay codes. Elucidated by various illustrations and tabular calculations, *ELM* and *OS-ELM* models seem to lag considerably in their capability to generate good performance metrics when compared with all the deep learning models (competing for *LSTM, CNN*, and hybrid *C-LSTM)*. However, we noticed that the *OS-ELM* could be a faster algorithm than *CNN, C-LSTM*, and *ELM*, but not *LSTM*.

There is enough affirmation, to set up the architecture *(i.e., LSTM)* as the concrete and hands-on framework for decoding long error correction codes when the decoding problem is formulated as a K-bit binary-classification problem.

A. LIMITATIONS AND FUTURE RESEARCH

LSTM architecture has demonstrated encouraging results comparatively. Nevertheless, there remain some limitations that motivate future research in this field.

- The *ML* algorithms proposed in this paper were optimized and analysed over a *AWGN* channel. The application of reinforcement learning *(RL)* can be explored for decoding over more complex channel scenarios. This may accelerate the power of algorithms through faster convergence rates, and better error rates, making communication systems more resilient.

- Experiments can be conducted in the environment of

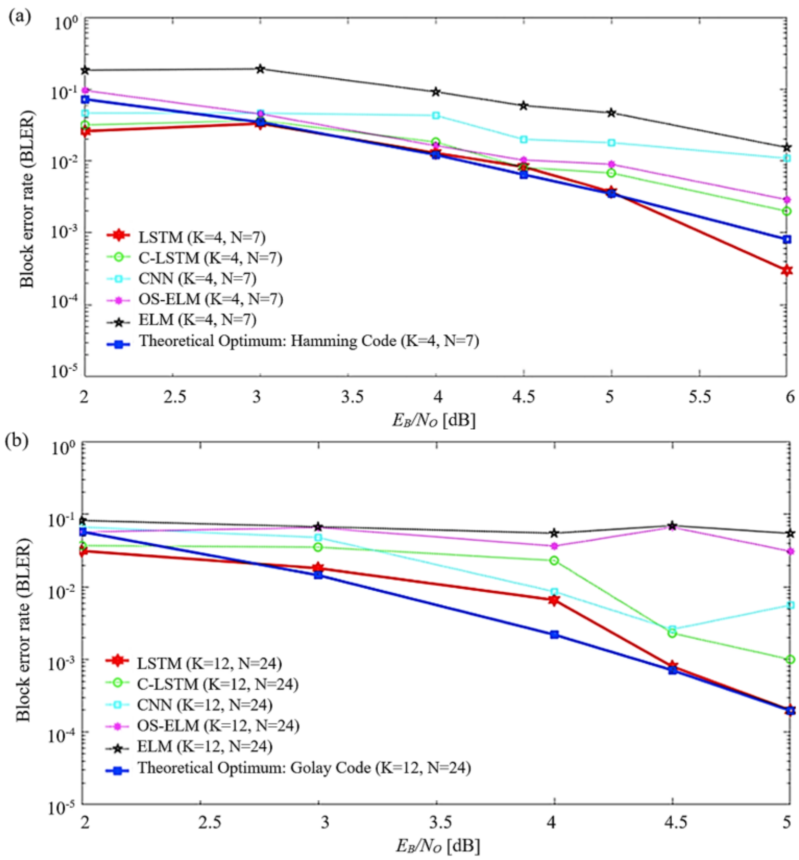


FIGURE 5. (a), (b). BLER versus SNR for all comparative models. (a) Hamming code (b) Golay code- based communication schemes. Evidently, the LSTM model (in red) is closest to the theoretical optimum ().

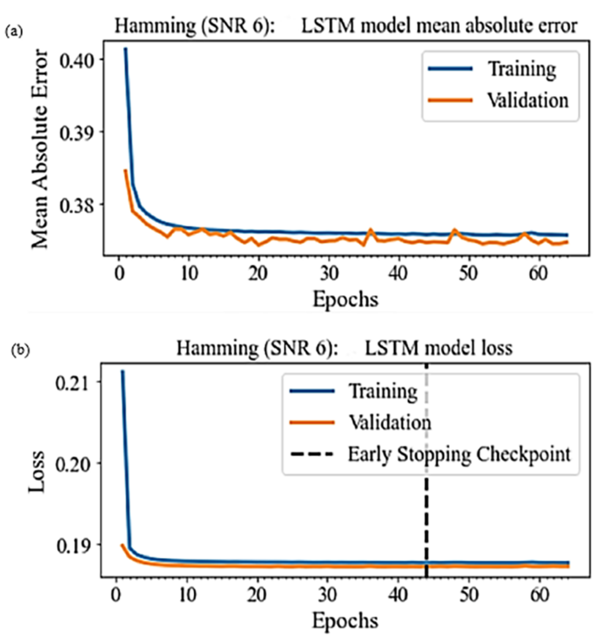


FIGURE 6. (a), (b).Performance evaluation of LSTM at SNR 6 for the Hamming code in the test phase.

deep Q-learning in further studies, exploiting the fact that, during learning, agents can back-propagate error derivatives through (noisy) communication channels. Hence, this approach will use centralised learning, but decentralised execution.

- It will be also interesting to extend these results to much larger families of algebraic block codes, as carried out by the authors in [39]–[41] for diverse families of codes (such as low density parity codes *(LDPC)* and so on. *(LDPC)* codes are important as they lead to lower Bit Error Rate *((BER)* values and reduced decoding complexity with the length of the block.

- The simulations and the performance of the proposed *LSTM* model on image multi-class classification tasks with high number of classes can also be considered in the future researches.

In closing, the authors provide detailed analysis to show that the *DL* architecture proposed in the paper is a useful and insightful way to fundamentally rethink traditional communication systems' design for classification and time-sequential problems. *LSTM* which is one of the comparative model and dominating the results can hold further performance improvements in reducing the error rates. The intelligent framework for modelling decoders can be used as a practical tool for developing intelligent radio communication systems

TABLE 6.
TESTING PERFORMANCE OF ALL COMPARATIVE MODELS FOR GOLAY CODE. RED = BEST PERFORMANCE

SNR (per information bit) (dB)	Optimum BLER (Theoretical)	Designed Predictive Models	Batch Size	Epochs	Errors	Test messages	BLER	MSE	RMSE	MAE
0	0.3120	LSTM	700	1,000	10502	33,330	0.3150	0.4964	0.7045	0.3998
		C-LSTM	700	1,000	10669	33,330	0.3200	0.5797	0.7613	0.4020
		CNN	700	1,000	12902	33,330	0.3871	0.6143	0.7837	0.4093
		OS-ELM	NA*	1,000	14502	33,330	0.4351	2.2745	1.5081	0.5001
		ELM	NA*	1,000	18855	33,330	0.5657	2.6316	1.6222	0.4467
2	0.0575	LSTM	700	1,000	2624	40,000	0.0656	0.4744	0.6887	0.3970
		C-LSTM	700	1,000	3068	40,000	0.0767	0.5577	0.7467	0.4001
		CNN	700	1,000	3864	40,000	0.0966	0.5961	0.7720	0.4089
		OS-ELM	NA*	1,000	11200	40,000	0.2800	2.0475	1.4309	0.4976
		ELM	NA*	1,000	11280	40,000	0.2820	2.3116	1.5203	0.4362
3	0.0145	LSTM	700	1,000	905	50,000	0.0181	0.4675	0.6837	0.3771
		C-LSTM	700	1,000	1765	50,000	0.0353	0.4779	0.6913	0.3977
		CNN	700	1,000	2395	50,000	0.0479	0.5780	0.7602	0.4021
		OS-ELM	NA*	1,000	9275	50,000	0.1855	1.9871	1.4096	0.5324
		ELM	NA*	1,000	8865	50,000	0.1773	2.1913	1.4803	0.5320
4	0.0022	LSTM	2,000	1,000	198	66,000	0.0030	0.4267	0.6532	0.3760
		C-LSTM	2,000	1,000	1518	66,000	0.0230	0.4574	0.6763	0.4177
		CNN	2,000	1,000	561	66,000	0.0085	0.4871	0.6979	0.4242
		OS-ELM	NA*	1,000	3927	66,000	0.0595	1.8678	1.3666	0.5027
		ELM	NA*	1,000	6567	66,000	0.0995	1.9976	1.4133	0.4999
4.5	0.000712	LSTM	2,000	1,000	576	720,000	0.0008	0.4158	0.6448	0.2972
		C-LSTM	2,000	1,000	1656	720,000	0.0023	0.4531	0.6731	0.3154
		CNN	2,000	1,000	1872	720,000	0.0026	0.4650	0.6819	0.3159
		OS-ELM	NA*	1,000	18936	720,000	0.0263	1.2959	1.1383	0.5241
		ELM	NA*	1,000	50184	720,000	0.0697	1.5952	1.2630	0.5320
5	0.000197	LSTM	700	2,000	160	800,000	0.0002	0.3975	0.6304	0.2081
		C-LSTM	700	2,000	800	800,000	0.0010	0.4479	0.6692	0.2297
		CNN	700	2,000	880	800,000	0.0011	0.4567	0.6757	0.2979
		OS-ELM	NA*	2,000	4320	800,000	0.0054	1.1683	1.0808	0.5190
		ELM	NA*	2,000	24720	800,000	0.0309	1.2641	1.1243	0.5270

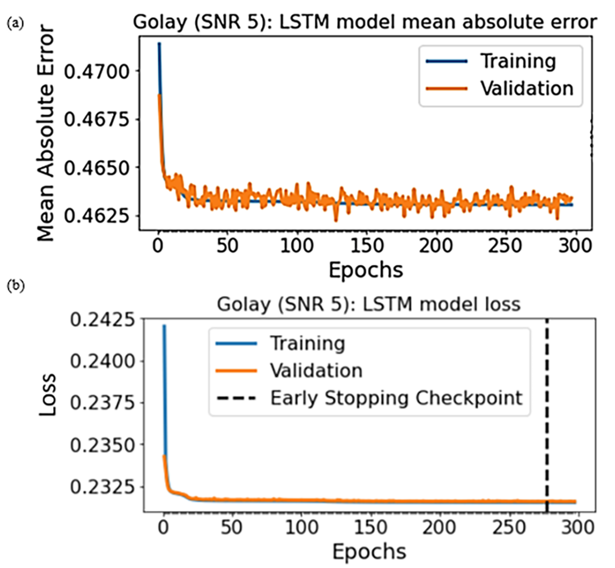


FIGURE 7. (a), (b). Performance evaluation of LSTM at SNR 5 for the binary extended Golay code in the test phase.

VI. APPENDIX. THEORETICAL DETAILS

(Deep learning methodologies)

A. CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN's are an accomplished category of Feed Forward Neural Networks which are superior in finding hidden features and used in studies such as [24], [25]. The general architecture has a layer that finds out the data patterns. This is called Convolutional layer. Next layer, that is the pooling layer brings down the dimension of the target variable, while the last layer creates the probability of initial input data in every category. This is called a Fully connected layer. Mathematically, if Af = function of activation, W = Kernel Weight with feature map 'm', and * = operator that serves the convolutional process. Each convolutional layer *(L)* is represented as:

$$\check{L} = Af((W^m * x_i) + b_k) \tag{6}$$

The structure of a *CNN* model creates data pattern filters. *Adam* and *ReLU* are popularly used as an optimisation algorithm, and mathematically expressed as:

$$f(x) = \check{\max}(0, x) \tag{7}$$

B. LONG SHORT-TERM MEMORY NETWORK (LSTM)

In its general sense, an *LSTM* method is well-known for its capability to resolve the vanishing gradient issues with the

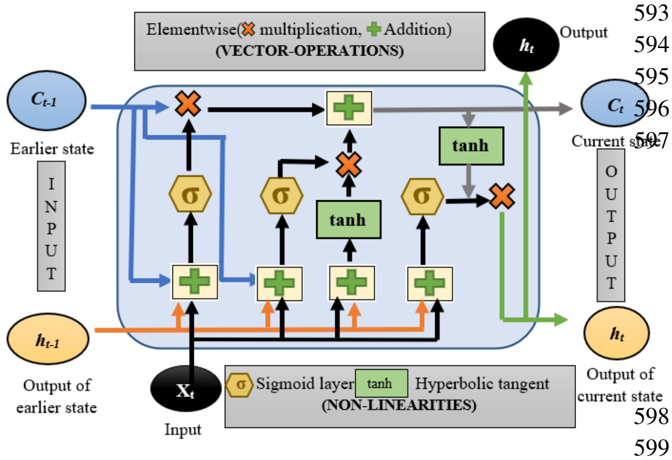


FIGURE 8. Schematic of deep learning LSTM architecture.

help of an input, forget, and output gate [42]. The *LSTM* network branched from recurrent neural networks *(RNN)* [43] emerged as an effective and scalable model for learning problems related to sequential data sets [44]. There are many works accomplished for decoding linear code blocks using *(RNN)* such as [40], [45], [46] The authors in these papers presented novel *(RNN)* based soft decision decoder for block codes, and further demonstrated that the *(RNN)* decoders can be used to improve the performance or alternatively reduce the computational complexity of the algorithms. Contrary to traditional recurrent neural networks, the architecture of an *LSTM* entails a layered connection enabling the system to learn short and long-range temporal dependencies between inputs and target data.

This architecture is not subjected to optimisation hurdles that occur in simple recurrent networks *(SRNs)* as the *LSTM* method can continuously update the next simulated value. This feature has led to its popularity as a sequential algorithm without the vanishing gradient issue [44], making it a potent choice for communication problems in updating information flow in and out of a communication system. Mathematically, an *LSTM* model (shown in Figure 8) can be represented as follows :

1) Assuming h_{t-1} = Last hidden state, x_t = new input, f_t = forget gate, W_f = Weight matrices, b_f = bias vector, i_t = Input, $\sigma(\dots)$, $\tanh(\dots)$ = activation functions (hyperbolic tangent, logistic sigmoid). The state of the cell is shown as follows

$$f_t = \sigma\left(W_f.\left(h_{t-1}, x_t\right) + b_f\right) \tag{8}$$

Information stored in the cell state $\check{C}_t$ is jointly influenced by the forget gate f_t and the input gate i_t:

$$\check{C}_t = \tanh(W_c.\left(h_{t-1}, x_t\right) + b_c) \tag{9}$$

$$i_t = \sigma\left(W_i.\left(h_{t-1}, x_t\right) + b_i\right) \tag{10}$$

2) C_t is the new cell state, which is updated by integrating the earlier state of the cell C_{t-1} and the new candidate state of cell or $\check{C}_t$. Here, the earlier state is affected by f_t or forget gate. The latter gate is scaled by it or the input gate.:

$$C_t = \left(f_t * C_{t-1} + i_t * \check{C}_t\right) \tag{11}$$

$$O_t = \sigma\left(W_O.\left(h_{t-1}, x_t\right) + b_O\right) \tag{12}$$

$$h_t = O_t.\tanh(C_t) \tag{13}$$

3) For the output process, an output gate or a new gate denoted by O_t is built. This decides the state of the cell to be outputted. Cell state C_t activated by tanh function is filtered as a product with O_t. The desired output h_t, obtained after the multiplication result is shown in Figure 8)

Of all models *(LSTM)* seems to have been used widely in radio communication studies. The notable mentions are [26], [27], [34], [42], [44]

C. CONVOLUTIONAL LONG SHORT-TERM MEMORY NETWORK (C-LSTM)

The framework involves the integration of 1-D CNN (One-dimensional convolutional neural network) and long short-term memory network *(LSTM)*. The *(CNN)* model extracts the spatial features, whereas *(LSTM)* which is a kind of recurrent neural network learns order dependence in sequence prediction problems. Both the models have been discussed in detail above. The studies that have used *C-LSTM* in their work on radio communication are [47], [48] (Machine learning methodologies)

D. EXTREME LEARNING MACHINE (ELM)

Extreme learning machine *(ELM)* is a training algorithm for single hidden layer feed-forward neural network. It can be used for feature learning, compression, classification, and regression, clustering preferably with a single layer or multiple layers of hidden nodes. Very few studies such as [49] used *(ELM)* in radio communications, however it is generally used in time series forecasting in other application areas.

E. ONLINE SEQUENTIAL EXTREME LEARNING MACHINE (OS-ELM)

A new and better version of *(ELM)* that learns the data in a sequential manner and allows online learning process either in a block manner or one by one. When the learning process gets completed, the training data gets discarded. The window also has either a fixed or a variable window. It is a fast algorithm and suitable for forecasting as it discards the data if it has already been trained. There are two fundamental learning steps in this algorithm: Firstly, initialization step where the weights and biases are randomly assigned to training data. This computes the output matrix. Secondly

Sequential learning step where updating of weights makes this algorithm accurate and fast. This makes them versatile and more suitable for short term forecasting. There are very few studies that have used OS-ELM in their research for radio communication [22]

VII. ACKNOWLEDGEMENTS

This research was supported by the Defence Science and Technology Group *(DSTG)* under the Australian Postgraduate Research *(APR)* Intern program. We thank Professor Jeffery Soar and Dr. Barbara Harmes *(USQ)* for the editing support. We thank Chris Davey (USQ-Defence PhD Researcher) for his technical support on typesetting the revised version of the document. The authors are thankful to all eight reviewers and the Editor for their constructive suggestions and comments.

REFERENCES

[1] D. Bala, G. Waliullah, N. Islam, I. Abdullah, and M. A. Hossain, “Analysis the probability of bit error performance on different digital modulation techniques over awgn channel using matlab,” Journal of Electrical Engineering, Electronics, Control and Computer Science, vol. 7, pp. 9–18, 2021.

[2] X. Gu, “Intelligent communication systems and networks (mlicom 2017),” Wireless Networks, vol. 25, no. 7, pp. 3701–3702, 2019.

[3] H. Ye, G. Y. Li, and B.-H. Juang, “Power of deep learning for channel estimation and signal detection in ofdm systems,” IEEE Wireless Communications Letters, vol. 7, no. 1, pp. 114–117, 2017.

[4] I. Shakeel, “Machine learning based featureless signalling,” in MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM), pp. 1–9, IEEE, 2018.

[5] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, “Deep learning for wireless communications: An emerging interdisciplinary paradigm,” IEEE Wireless Communications, vol. 27, no. 4, pp. 133–139, 2020.

[6] U. Kumar and B. Umashankar, “Improved hamming code for error detection and correction,” in 2007 2nd International Symposium on Wireless Pervasive Computing, IEEE, 2007.

[7] B. Zhang, B. Tondi, X. Lv, and M. Barni, “Challenging the adversarial robustness of dnns based on error-correcting output codes,” Security and Communication Networks, vol. 2020, 2020.

[8] S. A. A. Ahmed, C. Zor, M. Awais, B. Yanikoglu, and J. Kittler, “Deep convolutional neural network ensembles using ecoc,” IEEE Access, vol. 9, pp. 86083–86095, 2021.

[9] T. O’shea and J. Hoydis, “An introduction to deep learning for the physical layer,” IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563–575, 2017.

[10] F. Bernhardt, P. Landrock, and O. Manz, “The extended golay codes considered as ideals,” Journal of Combinatorial Theory, Series A, vol. 55, no. 2, pp. 235–246, 1990.

[11] R. Kadel, N. Islam, K. Ahmed, and S. J. Halder, “Opportunities and challenges for error correction scheme for wireless body area network—a survey,” Journal of Sensor and Actuator Networks, vol. 8, no. 1, p. 1, 2018.

[12] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: Performance criterion and code construction,” IEEE transactions on information theory, vol. 44, no. 2, pp. 744–765, 1998.

[13] H. Huang, S. Guo, G. Gui, Z. Yang, J. Zhang, H. Sari, and F. Adachi, “Deep learning for physical-layer 5g wireless techniques: Opportunities, challenges and solutions,” IEEE Wireless Communications, vol. 27, no. 1, pp. 214–222, 2019.

[14] A. C. Vaz, C. G. Nayak, and D. Nayak, “Hamming code performance evaluation using artificial neural network decoder,” in 2019 15th International Conference on Engineering of Modern Electric Systems (EMES), pp. 37–40, IEEE, 2019.

[15] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, “Deep learning for wireless physical layer: Opportunities and challenges,” China Communications, vol. 14, no. 11, pp. 92–111, 2017.

[16] E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be’ery, “Deep learning methods for improved decoding of linear codes,” IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 119–131, 2018.

[17] A. Emam, M. Shalaby, M. A. Aboelazm, H. E. Abou Bakr, and H. A. Mansour, “A comparative study between cnn, lstm, and cldnn models in the context of radio modulation classification,” in 2020 12th International Conference on Electrical Engineering (ICEENG), pp. 190–195, IEEE, 2020.

[18] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, “Oracle: Optimized radio classification through convolutional neural networks,” in IEEE INFOCOM 2019-IEEE Conference on Computer Communications, pp. 370–378, IEEE, 2019.

[19] Y. Liu, J. Duan, and J. Meng, “Difference attention based error correction lstm model for time series prediction,” in Journal of Physics: Conference Series, vol. 1550, p. 032121, IOP Publishing, 2020.

[20] S. Pollin, M. Rykunov, A. Bourdoux, H. Sahli, et al., “Convolutional long short-term memory networks for doppler-radar based target classification,” in 2019 IEEE Radar Conference (RadarConf), pp. 1–6, IEEE, 2019.

[21] D. Zabala-Blanco, M. Mora, C. A. Azurdia-Meza, A. Dehghan Firoozabadi, P. Palacios Játiva, and I. Soto, “Relaxation of the radio-frequency linewidth for coherent-optical orthogonal frequency-division multiplexing schemes by employing the improved extreme learning machine,” Symmetry, vol. 12, no. 4, p. 632, 2020.

[22] Y. Bai, L. Cao, S. Wang, H. Ding, and Y. Yue, “Data collection strategy based on oselm and gray wolf optimization algorithm for wireless sensor networks,” Computational Intelligence and Neuroscience, vol. 2022, 2022.

[23] M. Zhang, M. Diao, and L. Guo, “Convolutional neural networks for automatic cognitive radio waveform recognition,” IEEE Access, vol. 5, pp. 11074–11082, 2017.

[24] W. Yongshi, G. Jie, L. Hao, L. Li, W. Zhigang, and W. Houjun, “Cnn-based modulation classification in the complicated communication channel,” in 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), pp. 512–516, IEEE, 2017.

[25] S. Zheng, P. Qi, S. Chen, and X. Yang, “Fusion methods for cnn-based automatic modulation classification,” IEEE Access, vol. 7, pp. 66496–66504, 2019.

[26] Z. Ke and H. Vikalo, “Real-time radio technology and modulation classification via an lstm auto-encoder,” IEEE Transactions on Wireless Communications, 2021.

[27] X. Wang, Y. Zhang, H. Zhang, Y. Li, and X. Wei, “Radio frequency signal identification using transfer learning based on lstm,” Circuits, Systems, and Signal Processing, vol. 39, no. 11, pp. 5514–5528, 2020.

[28] Z. Zhang, H. Luo, C. Wang, C. Gan, and Y. Xiang, “Automatic modulation classification using cnn-lstm based dual-stream structure,” IEEE Transactions on Vehicular Technology, vol. 69, no. 11, pp. 13521–13531, 2020.

[29] N. Wang, T. Jiang, S. Lv, and L. Xiao, “Physical-layer authentication based on extreme learning machine,” IEEE Communications Letters, vol. 21, no. 7, pp. 1557–1560, 2017.

[30] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, “Deep learning convolutional neural networks for radio identification,” IEEE Communications Magazine, vol. 56, no. 9, pp. 146–152, 2018.

[31] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, “TensorFlow: A system for Large-Scale machine learning,” in 12th USENIX symposium on operating systems design and implementation (OSDI 16), pp. 265–283, 2016.

[32] N. Ketkar, Introduction to keras, pp. 97–111. Springer, 2017.

[33] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg, “Scikit-learn: Machine learning in python,” the Journal of machine Learning research, vol. 12, pp. 2825–2830, 2011.

[34] L. Shi, Y. He, B. Li, T. Cheng, Y. Huang, and Y. Sui, “Tilt angle monitoring by using sparse residual lstm network and grid search,” IEEE Sensors Journal, vol. 19, no. 19, pp. 8803–8812, 2019.

[35] J. Hu, J. Zhang, C. Zhang, and J. Wang, “A new deep neural network based on a stack of single-hidden-layer feedforward neural networks with randomly fixed hidden neurons,” Neurocomputing, vol. 171, pp. 63–72, 2016.

[36] D. M. Hawkins, “The problem of overfitting,” Journal of chemical information and computer sciences, vol. 44, no. 1, pp. 1–12, 2004.

[37] P. Ramachandran, B. Zoph, and Q. V. Le, “Searching for activation functions,” arXiv preprint arXiv:1710.05941, 2017.

[38] F. B. Mismar and B. L. Evans, “Deep learning in downlink coordinated multipoint in new radio heterogeneous networks,” IEEE Wireless Communications Letters, vol. 8, no. 4, pp. 1040–1043, 2019.

[39] E. Nachmani and L. Wolf, “Hyper-graph network decoders for algebraic block codes,” Aug. 5 2021. US Patent App. 16/782,919.

[40] E. Nachmani, E. Marciano, D. Burshtein, and Y. Be’ery, “Rnn decoding of linear block codes,” arXiv preprint arXiv:1702.07560, 2017.

[41] T. Thi Bao Nguyen, T. Nguyen Tan, and H. Lee, “Low-complexity high-throughput qc-ldpc decoder for 5g new radio wireless communication,” Electronics, vol. 10, no. 4, p. 516, 2021.

[42] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with lstm,” Neural computation, vol. 12, no. 10, pp. 2451–2471, 2000.

[43] J. Ma, H. Liu, C. Peng, and T. Qiu, “Unauthorized broadcasting identification: A deep lstm recurrent learning approach,” IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 9, pp. 5981–5983, 2020.

[44] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “Lstm: A search space odyssey,” IEEE transactions on neural networks and learning systems, vol. 28, no. 10, pp. 2222–2232, 2016.

[45] Y. Jiang, H. Kim, H. Asnani, S. Kannan, S. Oh, and P. Viswanath, “Learn codes: Inventing low-latency codes via recurrent neural networks,” IEEE Journal on Selected Areas in Information Theory, vol. 1, no. 1, pp. 207–216, 2020.

[46] A. Bennatan, Y. Choukroun, and P. Kisilev, “Deep learning for decoding of linear codes-a syndrome-based approach,” in 2018 IEEE International Symposium on Information Theory (ISIT), pp. 1595–1599, IEEE, 2018.

[47] J. Xie, J. Fang, C. Liu, and X. Li, “Deep learning-based spectrum sensing in cognitive radio: A cnn-lstm approach,” IEEE Communications Letters, vol. 24, no. 10, pp. 2196–2200, 2020.

[48] M. Xu, Z. Yin, M. Wu, Z. Wu, Y. Zhao, and Z. Gao, “Spectrum sensing based on parallel cnn-lstm network,” in 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), pp. 1–5, IEEE, 2020.

[49] C. Qing, W. Yu, B. Cai, J. Wang, and C. Huang, “Elm-based frame synchronization in burst-mode communication systems with nonlinear distortion,” IEEE Wireless Communications Letters, vol. 9, no. 6, pp. 915–919, 2020.

EKTA SHARMA *(IEEE, INFORMS, AMSI* member) is a *Ph.D.* Candidate (Artificial Intelligence) at University of Southern Queensland, Australia. She completed an *M.Phil., M.Sc.* (Operations Research), and *B.Sc.* (Mathematical Sciences), from the University of Delhi, India. She has over a decade of experience in both Academia and Industry across varied roles from Area Manager to Learning Advisor, Lecturer, and Researcher at Universities in Europe, India, and Australia. Her research work has received funding from the Australian Defence Science and Technology Group, The Australian Mathematical Sciences Institute, and the Australian Government. She is working on varied cross-disciplinary research problems in artificial intelligence.

ADJUNCT PROFESSOR ISMAIL SHAKEEL (Senior *IEEE* Member) received a Ph.D. degree in Telecommunications in 2007 and a *BEng (Hons)* degree in Electronic Engineering in 1997 from the University of South Australia. He has completed two master’s degrees at the University of Canterbury *(NZ)* and Monash University in 2001 and 2002 respectively. Ismail joined Defence Science and Technology Group *(DSTG)* in 2011 and is currently with the Cyber Electronic Warfare Division at *DSTG*. Ismail is also an Adjunct Professor at the University of Southern Queensland. Before joining *DSTG*, he has worked in both academia and industry and holds a patent and generated more than 40 technical reports and publications in the field of telecommunications. Ismail’s current research interests include signal detection and classification techniques, artificial intelligence-enabled wireless communication, interference-resistant signaling, chaotic communication, and cooperative wireless communication.

PROFESSOR SANCHO SALCEDO-SANZ was born in Madrid, Spain, in 1974. He received the B.S degree in Physics from Universidad Complutense de Madrid, Spain, in 1998, the Ph.D. degree in Telecommunications Engineering from the Universidad Carlos III de Madrid, Spain, in 2002, and the Ph.D. degree in Physics from Universidad Complutense de Madrid in 2019. He spent one year in the School of Computer Science, The University of Birmingham, U.K, as postdoctoral Research Fellow. Currently, he is a Full Professor at the department of Signal Processing and Communications, Universidad de Alcalá, Spain. He has co-authored more than 220 international journal papers in the field of Machine Learning and Soft-Computing and its applications. His current interests deal with Soft-computing techniques, hybrid algorithms and neural networks in different problems of Science and Technology.

PROFESSOR RAVINESH C DEO (Senior IEEE Member) is a Highly Cited Author, 2021 Clarivate) leads USQ’s Advanced Data Analytics Lab as Professor at the University of Southern Queensland (USQ), Australia. He is a Clarivate Highly Cited Researcher with publications ranking in top 1% by citations for field and publication year in the Web of Science citation index and is among scientists and social scientists who have demonstrated significant broad influence, reflected in the publication of multiple papers frequently cited by their peers. He leads cross-disciplinary research in deep learning and artificial intelligence, supervising 20+ Ph.D./M.Sc. Degrees. He has received Employee Excellence Awards, Elsevier Highly Cited Paper Awards, and Publication Excellence and Teaching Commendations. He has published more than 250 articles, 170 journals, and seven books with a cumulative citation that exceed 9,700 and an H-index of 55.

• • •