

A Data-Driven Framework for Predicting National Vehicular Emissions Using Regulatory Emissions Test Databases of Sri Lanka

Sankha Jayawardhana, Saman Bandara, Varuna Adikariwattage, Thusitha Sugathapala

Abstract

This study presents a data-driven framework for forecasting national vehicular emissions in Sri Lanka using emissions testing records maintained by the Department of Motor Traffic (DMT). These databases, originally collected for compliance verification, remain largely underutilized despite their potential for informing long-term environmental policy. Focusing on passenger cars with internal combustion engines as a case study, a Tier 2 bottom-up emissions estimation methodology is implemented, integrating annual mileage of tested vehicles on a monthly basis and vehicle vintage-size composition. The developed model forecasts pollutant loads including carbon monoxide (CO), non-methane volatile organic compounds (NMVOC), nitrogen oxides (NO_x), nitrous oxide (N₂O), and ammonia (NH₃), in an annual spanning monthly basis. The methodology begins with the derivation of average annual mileage values per vehicle by cross-referencing test entries from the same vehicle across consecutive years. A structured data pre-processing pipeline was developed to extract relevant vehicle characteristics and to clean and align data across multiple testing providers. To assign appropriate emission factors, a hybrid classification scheme was implemented based on both engine horsepower (sourced using a custom web-scraping and fuzzy matching algorithm) and vehicle manufacturing year, which determines its Euro emissions standard category. These parameters were then matched to Tier 2 emission factors from the EMEP/EEA guidelines. To predict future mileage, a Long Short-Term Memory (LSTM) neural network was trained using annual spanning monthly average mileage as the target variable. Input features included real GDP, fleet size of passenger cars, and the proportional distribution of Euro vehicle categories. Additional enhancements such as cyclical encoding of months, moving averages, and MinMax scaling were applied to improve model performance and generalizability. The LSTM model achieved a high predictive accuracy with an R² score of 0.899, confirming its capability to capture temporal trends in vehicular mileage patterns. Future emissions were forecasted by integrating the predicted mileage values with projected fleet compositions and matched emission factors. The fleet composition proportions were extrapolated using Holt-Winters triple exponential smoothing. The model predicts that CO will remain the dominant pollutant, with emissions ranging between 2,104.96 and 2,604.00 metric tonnes on an annual spanning monthly basis. NMVOC emissions are expected to range from 630.07 to 705.53 metric tonnes, and NO_x from 289.38 to 359.99 metric tonnes. N₂O and NH₃ emissions, though lower in volume, remain significant contributors, with projected ranges of 11.56–19.41 and 36.39–46.60 metric tonnes respectively. The framework's modularity allows for straightforward adaptation to other vehicle categories such as motorcycles, three-wheelers, and commercial vehicles, with adjustments to input parameters and emission factor sets. Although hybrid and electric vehicles were excluded due to the lack of centralized testing data, their future inclusion would require only minimal modifications. This research highlights how institutional data, when carefully restructured and modelled, can

yield high-utility tools for national-scale emissions forecasting. The proposed framework not only enhances the value of existing regulatory data but also supports the development of evidence-based transport and environmental policies, particularly in developing countries.

Keywords: *vehicular emissions, emission factors, LSTM prediction*

Author details;

1. Graduate Research Assistant, Department of Civil Engineering, University of Moratuwa.
jayawardhanass.24@uom.lk
2. Emeritus Professor, Department of Civil Engineering, University of Moratuwa.
bandara@uom.lk
3. Senior Lecturer, Department of Civil Engineering, University of Moratuwa.
varunaa@uom.lk
4. Senior Lecturer, Department of Mechanical Engineering, University of Moratuwa.
agtsugathapala@gmail.com