2024 Tenth Indian Control Conference (ICC)
9 - 11 December, 2024
IISER Bhopal, India

LSTM-EKF Machine Learning State Estimator Validation on Experimental Hybrid Three Tank System

Brajesh Singh Dhakar, Jayaram Valluru

Abstract— A rapid technological advance in data science and push for digitalization in the process industry resulted in the increasing popularity of data-driven-based predictive solutions/decision-making in the process industry in recent years. Data-driven soft sensors or predictive models, which are easier to develop and implement, are proven to be alternative solutions for costly hardware analyzers. The accuracy of the data-driven soft sensors or machine learning-based predictive models highly depends on the data quality. Soft Sensor design for industrial processes often involves handling measurements at different rates and delays. Slow-rate, irregular, and delayed measurements are generally more accurate but suffer from delays due to analysis time, such as laboratory testing. In contrast, fast-rate measurements, such as flow rate/temperature measurements, are available at regular intervals without delay but tend to be less accurate. Further, since these predictive models are developed based on past historical data, developed under normal operating conditions, the predictive performance of these data-driven soft sensors/machine learning-based predictive models deteriorates in the presence of unmodelled faults in the processes. Unknown faults, such as catalyst deactivation, leaks in the process tanks, etc. need to be detected or estimated for accurate estimation of the states of the system.

To improve the prediction performance of data-driven soft sensors/machine learning-based predictive models in the presence of noisy data, this work proposes the implementation of a machine learning-based state estimator. Here, a machine learning-based state estimator is developed by integration of dynamic Deep Neural Networks with the Bayesian State Estimators, such as nonlinear Kalman filters. Performance of the proposed algorithm is validated on experimental hybrid three tank system. Analysis of results indicate that integration of LSTMs with Bayesian estimators can produce accurate state estimates in the presence of noisy input/output data.

Key words: LSTMs, Machine Learning, Bayesian State Estimators.

*Research supported by TIH-IoT Technology Development Program.

B.S. Dhakar is with the Indian Institute of Technology Ropar.

J. Valluru, is with Indian Institute of Technology Ropar (Corresponding author e-mail: vjayaram@iitrpr.ac.in).

I. Design and Development of the Test Bed

The hybrid three tank system shown in Fig. 1 is from Process Control and Data Analytics Lab, Dept. Of Chemical Engineering IIT Ropar. It is designed for validation of model identification techniques and real-time advanced process control algorithms. This setup features three interconnected cylindrical tanks, allowing operation in interactive mode. It monitors water levels and inflow using level and flow transmitters, with I/P converters controlling the valves. The system can switch to hybrid mode through solenoid valves, enabling tank interaction. Load disturbances are introduced by draining or changing tank inflows. In addition to the load disturbances, leaks can be introduced in each cylindrical tank through solenoid valves. These advanced features of hybrid three tank experimental setup can be effectively used to validate the proposed adaptive machine learning based state estimation algorithms.

Inputs and outputs are handled via ACE 2007 wireless DAQ modules, integrated with a router. The system supports MATLAB, LabVIEW, Simulink, Python, Visual Basic, and SCADA software through a TCP library, enabling real-time testing of custom control algorithms.

Key Specifications of the experimental set up include:

- **Product Name:** Hybrid Tank System
- **Control Type:** Direct Digital Control (DDC)
- **DAQ Units:**
  - ACE 2007: 4 inputs, 2 outputs (24VDC)



Figure 1. Experimental set up of Hybrid Three Tank System, Process Control and Data Analytics Lab, IIT Ropar.

- ACE 2011: 8 digital outputs (24VDC)

- **Router:** D-Link R15 AX1500, Wi-Fi 6 (1500 Mbps)

II. Detailed Experiment Design

The hybrid tank system was tested to validate the performance of the proposed machine learning based state estimator. Below is the procedure followed:

Experimental Protocol:

1. Under normal operating conditions (i.e., no leaks in the system), a multistep input signal is given to the three-tank experimental setup, and the output data i.e., level information of each tank, is recorded at regular sampling intervals.
2. Tanks were reset after each experiment.
3. Obtained data sets are divided into 70% training and 30% validation data sets. The training data sets are used to train the machine learning model, here we have considered to use Long Short-Term Memory (LSTM) Networks for developing a predictive model. The developed LSTM based predictive model is further integrated with Bayesian state estimator sensor fusion algorithm to address noisy data issues, multi-rate/delayed measurements and to detect unknown faults, i.e., in the process leaks.

Data Collection:

- Sensor Integration: Wireless DAQ units (ACE 2007, ACE 2011) gathered data from sensors, control valves, and solenoid valves.
- Real-Time Monitoring: MATLAB and LabVIEW were used to log high-resolution data from the system.
- Data Processing: Data was exported to MS Excel, where noise filtering and outlier detection were performed to ensure accuracy.

Model Calibration:

I. **Initial Model Setup:** Using the historical data sets, deep dynamic neural network models are developed.

II. **Parameter adjustment:** Experimental data was compared with the model predictions. To account for the leaks in the process tanks, model is adapted by integrating with the Bayesian state estimators.

III. **Validation:** The calibrated models were validated using new data, with iterative adjustments to improve accuracy.

III. Experimental Evaluation on a Pilot Plant

A laboratory experiment was conducted to assess the performance of the proposed machine learning based state estimator on experimental hybrid three tank system. Figure 1 represents the experimental setup of the Hybrid Three-Tank System at the Process Control and Data Analytics Lab, IIT Ropar. In this setup, water is pumped into the left-hand tank and can be transferred to the middle tank through pipes controlled by on/off valves. Level transmitter sensors installed on each tank measure water levels at a sampling interval of one second.

For this experiment, the input signal is the flow rate of the pump on the left side (refer to Fig. 1), while the water level in the middle tank is monitored. The right-hand tank is not used in this validation. Using this setup, multiple step changes in the input variable (flow rate of the left pump) are introduced and both input and output (i.e., water level in the middle tank) data are collected (refer to Fig. 2 & 3, respectively). Seventy percent of this data is used to train the data-driven model (LSTM), while the remaining 30 percent is reserved for validation.

Figure 4 represents a comparison between the conventional

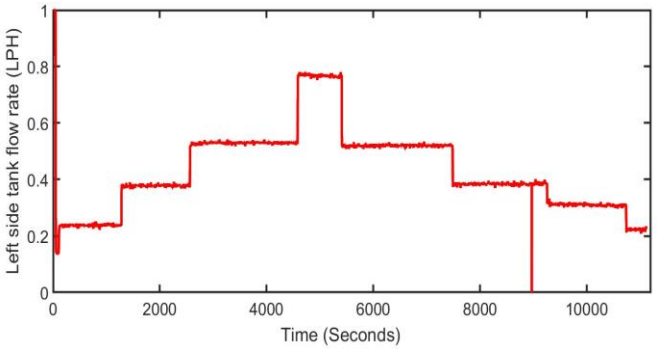


Figure 2: Profile of input signal (Left pump flow rate)

LSTM model predictions and the actual measurement of the middle tank level. In this figure, the solid blue line represents the LSTM model's measurement, while the solid red line indicates the actual measurement. Although the LSTM model follows a similar trend to the actual data, it does not accurately align with the middle tank's true measurements, indicating a limitation in the model's predictive accuracy in this instance.

Figure 5 provides a comparison of the actual middle tank level measurement with predictions from both the conventional LSTM model and the EKF-LSTM (Extended Kalman Filter-LSTM) state estimator model. In this figure,

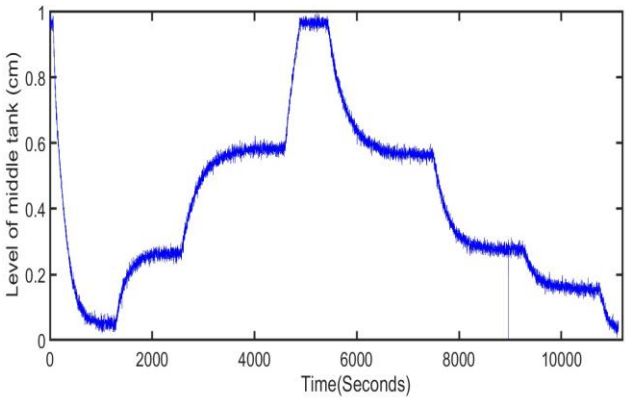


Figure 3. Profile of middle tank level (water level in the middle tank)

the solid blue line represents the EKF-LSTM prediction, the solid green line denotes the conventional LSTM prediction, and the solid red line shows the actual measurement. The results indicate that the EKF-LSTM model closely follows the trend of the actual middle tank measurements, aligning accurately and demonstrating zero bias, making it a feasible approach for this process. Conversely, while the conventional LSTM model also follows the general trend, it does not match the actual measurements or the EKF-LSTM prediction closely, showing a noticeable bias in comparison.

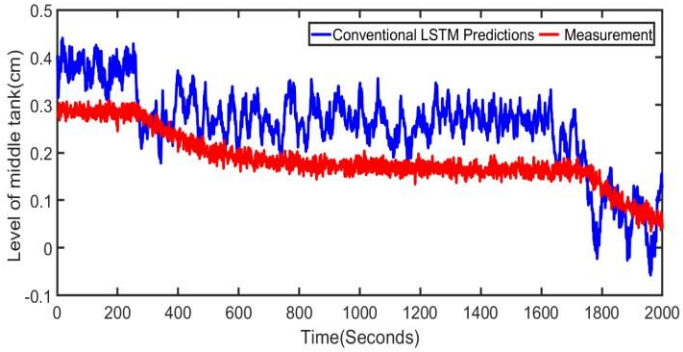


Figure 4. Comparison of profiles between conventional LSTM predictions and actual measurement

Figure 6 presents a comparison of error profiles between the conventional LSTM measurement and the actual measurement, as well as the EKF-LSTM measurement and the actual measurement. The solid blue line represents the error between the conventional LSTM measurement and the actual values, while the solid red line indicates the error between the EKF-LSTM measurement and the actual values. From the figure, it can be concluded that the error associated with the EKF-LSTM measurement exhibits a zero-mean error with no bias, indicating its reliability. In contrast, the error from the conventional LSTM measurement is not zero and shows bias. Furthermore, the RMSE between the EKF-LSTM measurement and the actual measurement is lower compared to that of the conventional LSTM, suggesting a better fit of the EKF-LSTM model to the data and indicating that the predicted values are closer to the actual measurements. Root mean squared error (RMSE) between conventional LSTM predictions and actual measurement is 0.0893, while RMSE between LSTM-EKF predictions and actual measurement is 0.0169.

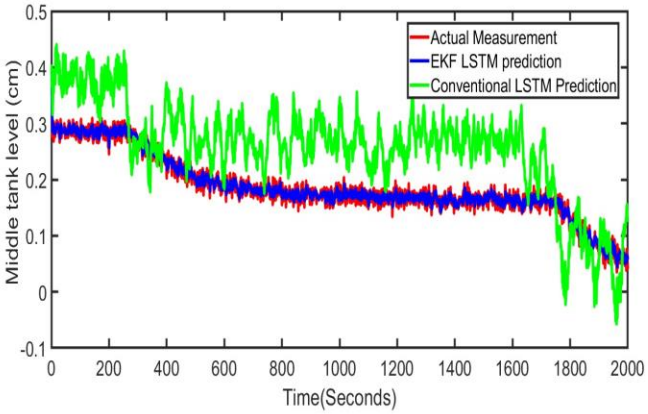


Figure 5. Comparison of profiles between conventional LSTM predictions, actual measurement and EKF- LSTM predictions

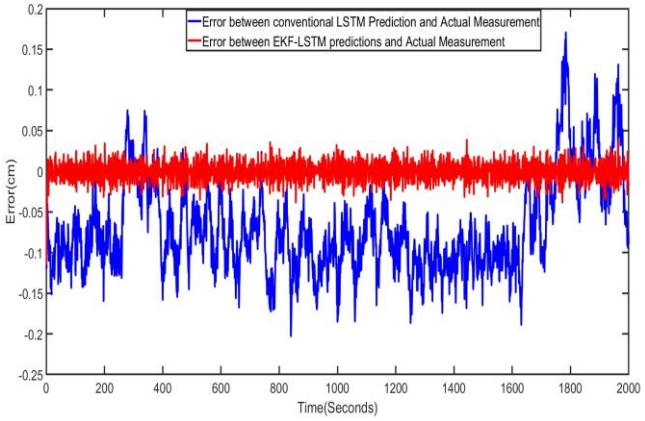


Figure 6. Comparison of error profiles

Working Experiment Video: A video demonstrating the experiment's working setup is available. The video showcases real-time system responses and data collection.

VideoLink:https://youtube.com/@ram_643?si=LW-VWzA7R6EBMvb0

IV. CONCLUSION

In the presence of noisy plant data, predictions obtained through data driven modelling techniques / machine learning models are shown to be less accurate. Further, measurements for most of the processes are sampled at irregular time varying intervals with delay. To address the data quality and measurement sampling uncertainty issues, this work aims in validating the LSTM-EKF based machine learning state estimator on experimental hybrid three tank system. Based on the preliminary results, it is observed that the proposed machine learning based state estimator (LSTM-EKF) has shown improved estimation performance compared to the conventional LSTM for noisy data scenario. Future work is aimed at testing the algorithm performance to multirate and time varying measurement delays.

ACKNOWLEDGEMENT

Authors would like to acknowledge TIH-IoT Technology Development Program for funding this project.

REFERENCES

[1] Fatehi, A, and Huang, B., "Kalman Filtering Approach for Multi-Rate Information Fusion with Irregular Sampling Rates and Variable Measurement Delays." Journal of Process Control 53 (2017): 15-25.

[2] Gopalakrishnan, A., Kaisare, N. S., and Narasimhan, S., "Incorporating Delayed and Infrequent Measurements into Extended Kalman Filter-Based Nonlinear State Estimation." Journal of Process Control 21, Vol. 1 (2011): 119-129.

[3] Fatehi, A. "Modeling a Hybrid Tank Pilot Plant Using First-Principal Models and Experimental Data.", 2015.

[4] Sansana, J., Rendall, R., Wang, Z., Chiang, L.H., Reis, M.S., "Sensor Fusion with Irregular Sampling and Variable Measurement Delays." Industrial & Engineering Chemistry Research 59.6 (2020): 2328-2340.