Predict the airway pressure in the respiratory circuit during the breath, given the time series of control inputs

Introduction:

When a patient has trouble breathing, the doctors use a ventilator to pump oxygen into a sedated patient's lungs via a tube in the windpipe. But mechanical ventilation is a clinician-intensive procedure, a limitation that was prominently on display during the early days of the COVID-19 pandemic. At the same time, developing new methods for controlling mechanical ventilators is prohibitively expensive, even before reaching clinical trials. High-quality simulators could reduce this barrier.

Current simulators are trained as an ensemble, where each model simulates a single lung setting. However, lungs and their attributes form a continuous space, so a parametric approach must be explored that would consider the differences in patient lungs.

I found data from Kaggle on a competition that Google Brain and Princeton University collaborated on to simulate a ventilator connected to a sedated patient's lung. The data consists train data and test data.

Below are the variables in the train data:

- id - globally-unique time step identifier across an entire file
- breath_id - globally-unique time step for breaths
- R - lung attribute indicating how restricted the airway is (in cmH2O/L/S). Physically, this is the change in pressure per change in flow (air volume per time). Intuitively, one can imagine blowing up a balloon through a straw. We can change R by changing the diameter of the straw, with higher R being harder to blow.
- C - lung attribute indicating how compliant the lung is (in mL/cmH2O). Physically, this is the change in volume per change in pressure. Intuitively, one can imagine the same balloon example. We can change C by changing the thickness of the balloon's latex, with higher C having thinner latex and easier to blow.
- time_step - the actual time stamp.
- u_in - the control input for the inspiratory solenoid valve. Ranges from 0 to 100.
- u_out - the control input for the exploratory solenoid valve. Either 0 or 1.
- pressure - the airway pressure measured in the respiratory circuit, measured in cmH2O.

The "pressure" is the one that need to be predicted in test data.

I will conduct some basic descriptions on the train data and finally use machine learning on train data in order to predict the airway pressure in test data. All the analysis will be conducted using R.