

Transfer Learning for Pandemic Forecasting

Focus Questions

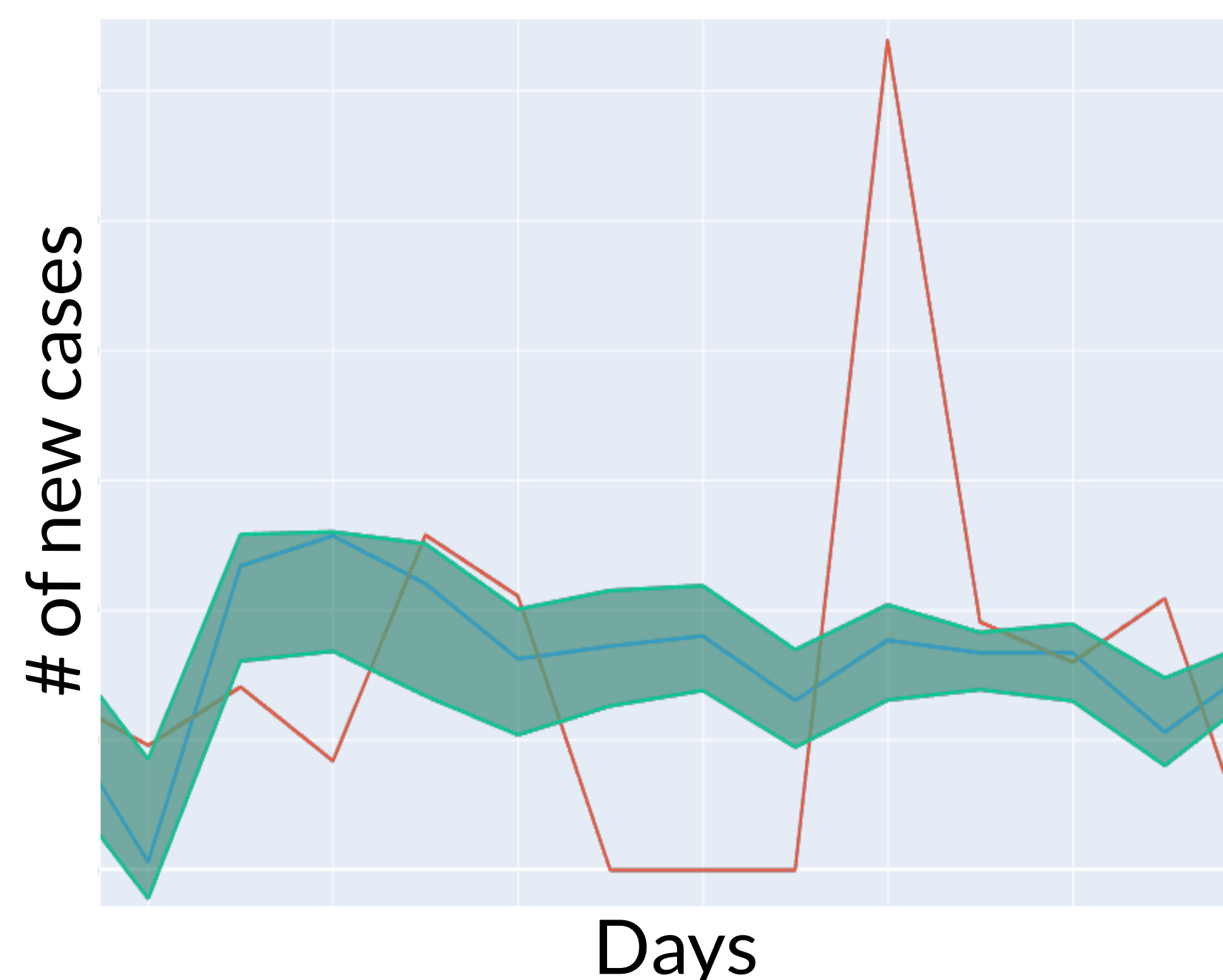
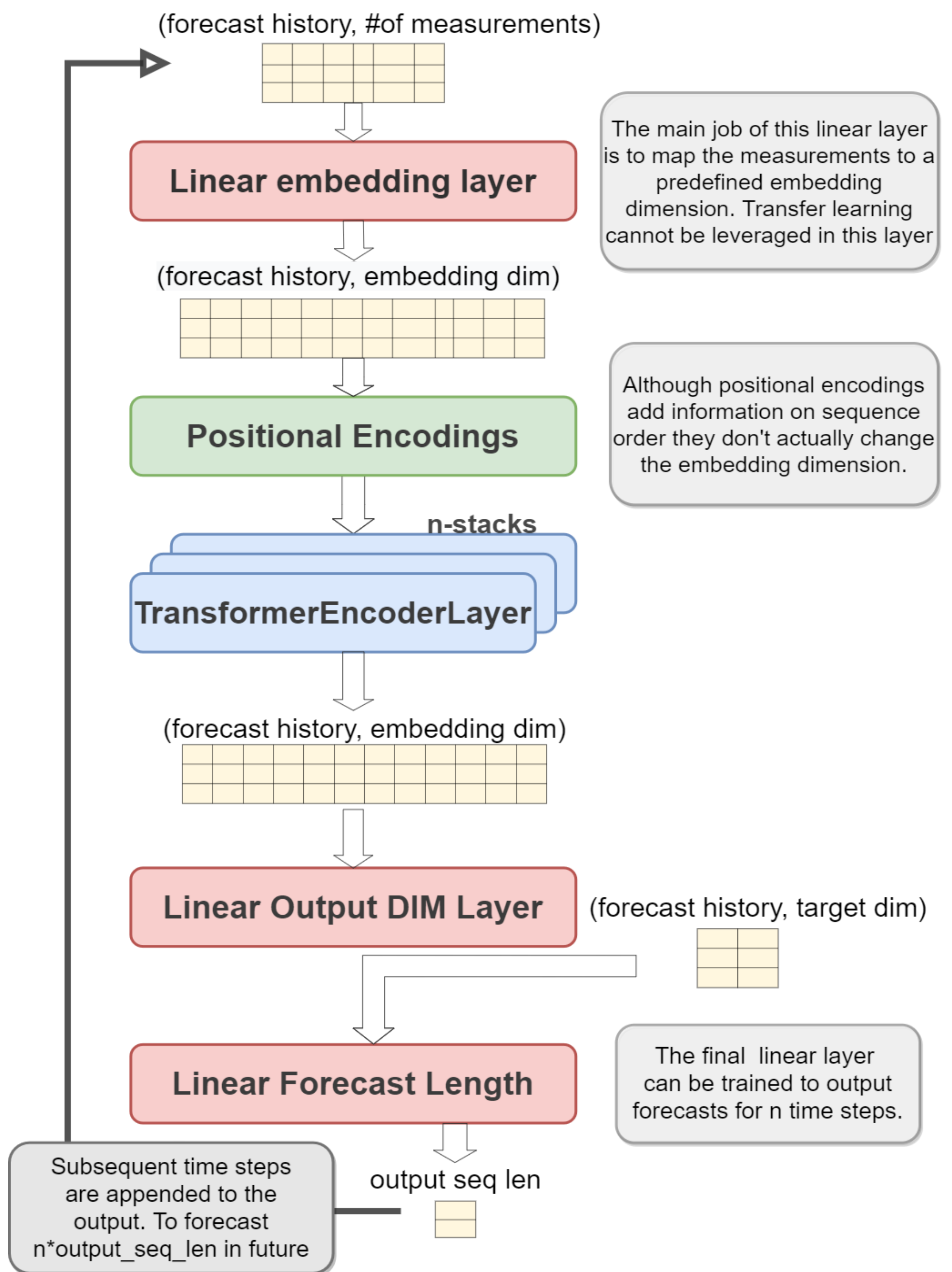
1. Can pre-training a time series model on un-related data improve its performance?
2. What architectures facilitate positive transfer while limiting negative transfer?
3. Do modern Seq2Seq models from NLP work well with respect to COVID-19 forecasting?
4. How do we design models that forecast and explain the links between policy interventions and case numbers?

Preliminary Results

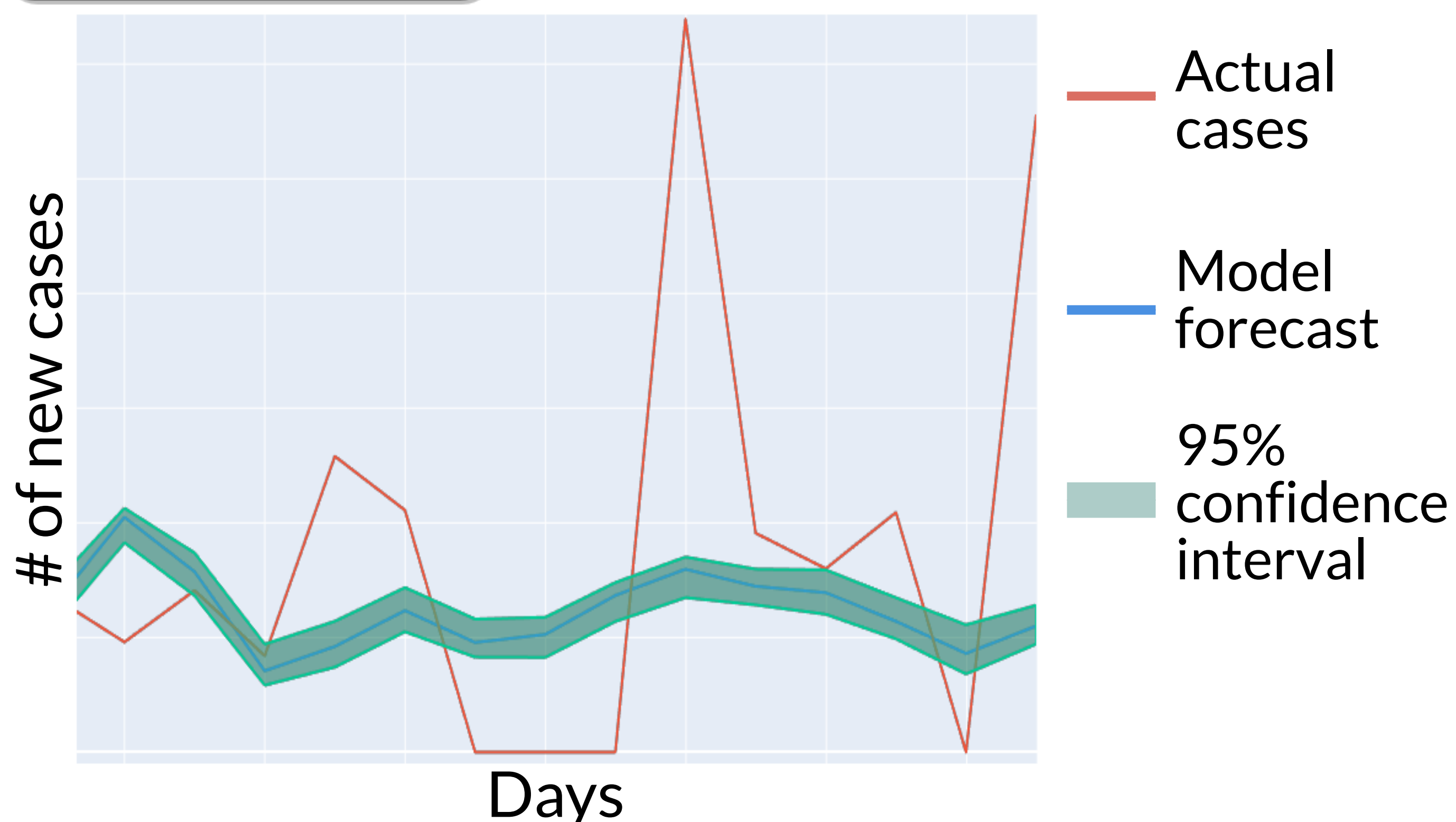
1. Training models with un-related time series data helps to improve performance on the majority of counties we studied.
2. In general, it helps both with respect to validation and test loss MSE.
3. Using transfer learning also seems to widen the model's confidence interval and make it better at gauging its own uncertainty.
4. Counties seem to have their own unique set hyper-parameters that impact performance.

Model Architecture

The main model is a TransformerEncoderLayer modified for transfer learning. We use a simple linear decoder.



Model w/ transfer
Palm Beach County, Florida, USA



Model with **no** Transfer
Palm Beach County, Florida, USA