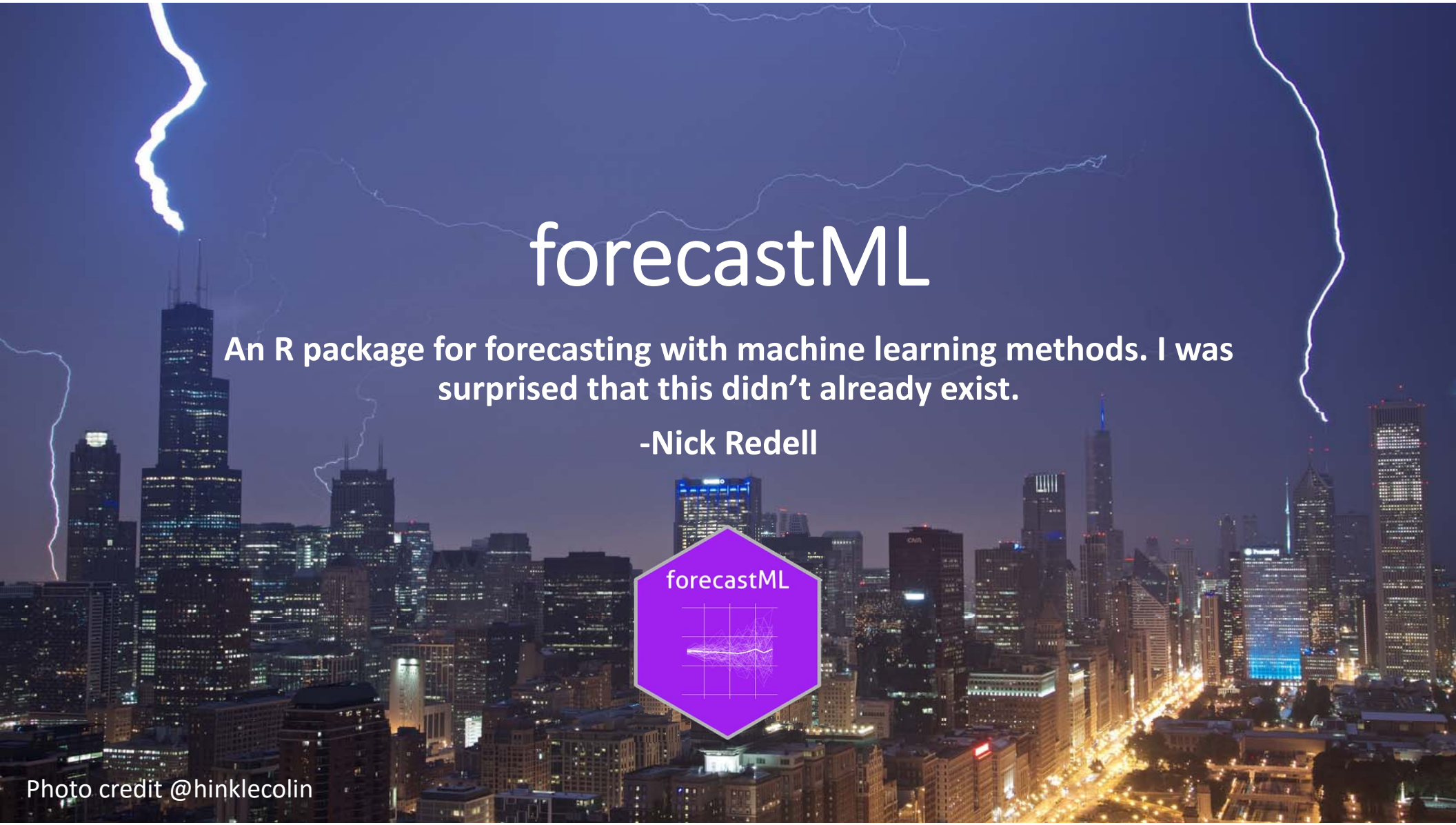
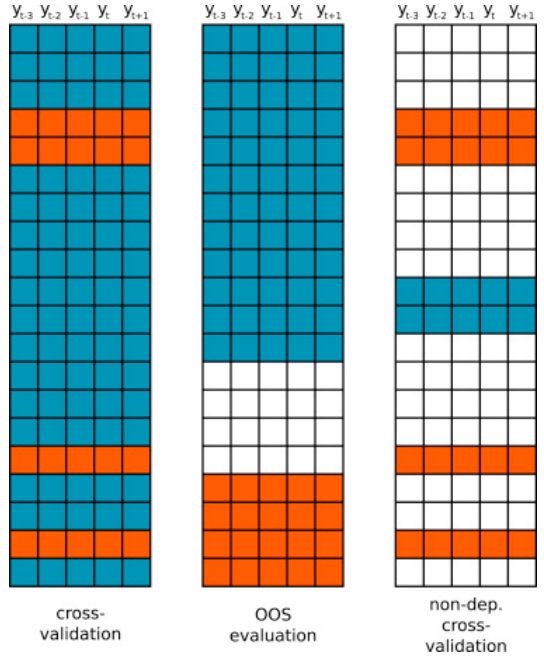

forecastML
An R package for forecasting with machine learning methods. I was surprised that this didn't already exist.
-Nick Redell
forecastML
Photo credit @hinklecolin

Where did the idea come from?

- "A note on the validity of cross-validation for evaluating autoregressive time series prediction" by Bergmeir, Hyndman, & Koo

- Analysis needs that shaped the package
  - Assessing forecast model stability at various historical times—backtesting—was important
  - Needed to work with classes of models that supported newer interpretable machine learning methods (e.g., Shapley values).
  - Quick iteration and ensembling across many models



The evolution of forecastML

1. 1 time series with lagged features and no date support
2. Support custom lags per feature and dates
3. Model many time series with lagged features
4. Things are slow…fix with parallelization
5. Support other types of model features
6. Test, test, test, and input argument type checks

The good, the bad, and the ugly

- **Good**
  - **Documentation**: Helpful error messages, thorough R help docs, vignettes, cheat sheets, even the code has liberal documentation.
  - **Flexibility**: Use any model, just return a data.frame of predictions
- **Bad**
  - **Helper functions:** The code is linear with no explicit helper functions. There's lots of if (x) {fun()} else if (y) {fun()} else if (z) {fun()}…sigh
- **Ugly**
  - **Slow**: group_by(group) %>% mutate("feature_lag_1" = lag(feature, n))

Continuous improvement

- **tsibble**
  - I'm a fan of increased standardization and consolidation in R.
  - This will also support sub-daily time intervals.
- **Speed up**
  - A particular slow spot is creating high dimensional forecasting datasets (help).
- **Unit tests**
  - There are going to be edge cases. Hyperparameter plotting needs more attention.
- **Added functionality**
  - More control over external cross-validation windows (simple-ish).
  - Multiple outcomes (a pretty massive re-write).
  - Setting a seed for future.apply::future.lapply() for reproducibility.

Location

- https://github.com/nredell/forecastML
- https://cran.r-project.org/web/packages/forecastML/index.html