Public School Financial Forecasting with Gaussian Process Regression

Presented at the 38th New England Statistics Symposium, New Haven CT

Eugene Quinn Ph.D.[1], Maggie Baker MBA[2]

[1]East Greenwich RI School Committee, [2]East Greenwich RI School Department

June 3rd, 2025

Statutory Requirements

Public schools in Rhode Island are required by law to produce 3-year forecasts of expenditures and revenues. Rhode Island General Law 16-2-21.2 [2] states that:

On or before November 1, the school committee shall submit to the city or town council an estimate of its expenditures by major program and school revenues for the three (3) years...

The forecasts are collected by the Rhode Island Division of Municipal Finance[3].

School Financial Data Sources

In 2010 the Rhode Island Department of Education (RIDE) standardized the accounting practice for public schools when it implemented the Uniform Chart of Accounts (UCOA) [1].

The UCOA data includes budget and actual values for revenues and expenses with six levels of categorization.

RIDE provides documentation including detailed instructions for coding each UCOA data element, and curates submitted data before making it publicly available on their website.

Granularity of the Data

In any given year, the UCOA data for a district will contain about 2,000 distinct line items, so some degree of summarization is needed.

There are many ways to use the six categorical variables to do this, but the most straightforward way is by a data element called 'Object Code'.

This produces about 150 budget-actual pairs, and aligns with the district's internal financial reporting structure.

UCOA provides a description for each Object Code, and these are generally self-explanatory:

- Object Code 51110: Regular Salaries
- Object Code 41321: Tuition from Other Districts

Required Granularity of the Forecast

While the statute requires only forecasts for 'major categories', governing bodies with oversight responsibility will expect more granularity.

Their oversight usually takes the form of questioning assumptions and criticizing deviations from prior forecasts at the budget line item level.

So in practical terms we require not only a high level forecast, but enough detail to make overseeing bodies comfortable with the results.

Ideally, we want both forecast totals for major budget categories and consistent forecasts at the individual line item level.

The Modelling Challenge

A number of practical considerations make these forecasts challenging.

- There are about 150 distinct Object Codes.
- The historical data contains a budget-actual pair for each Object Code for most years from 2010 through 2024.
- Governing bodies with oversight responsibility expect at least Object Code granularity and will not hesitate to criticize variances.
- Many Object Codes will have data for every year in the range, but in some cases, there may be only one or two years of data.
- The scale of the Object Code totals varies widely, from tens of millions of dollars in one case to hundreds of dollars in others.

The Gaussian Process Regression Approach

Gaussian Process Regression provides a mathematically disciplined way to meet these challenges.

- GP Regression provides a feasible way to satisfy the granularity requirement and provide consistent higher level summaries.
- Having separate models mitigates the problems with widely different scales.
- The models provides a 'sanity check' on budget assumptions, and a way of defending those assumptions.
- The required computations are easily automated with Python.
- Prior information can be incorporated if it is available.

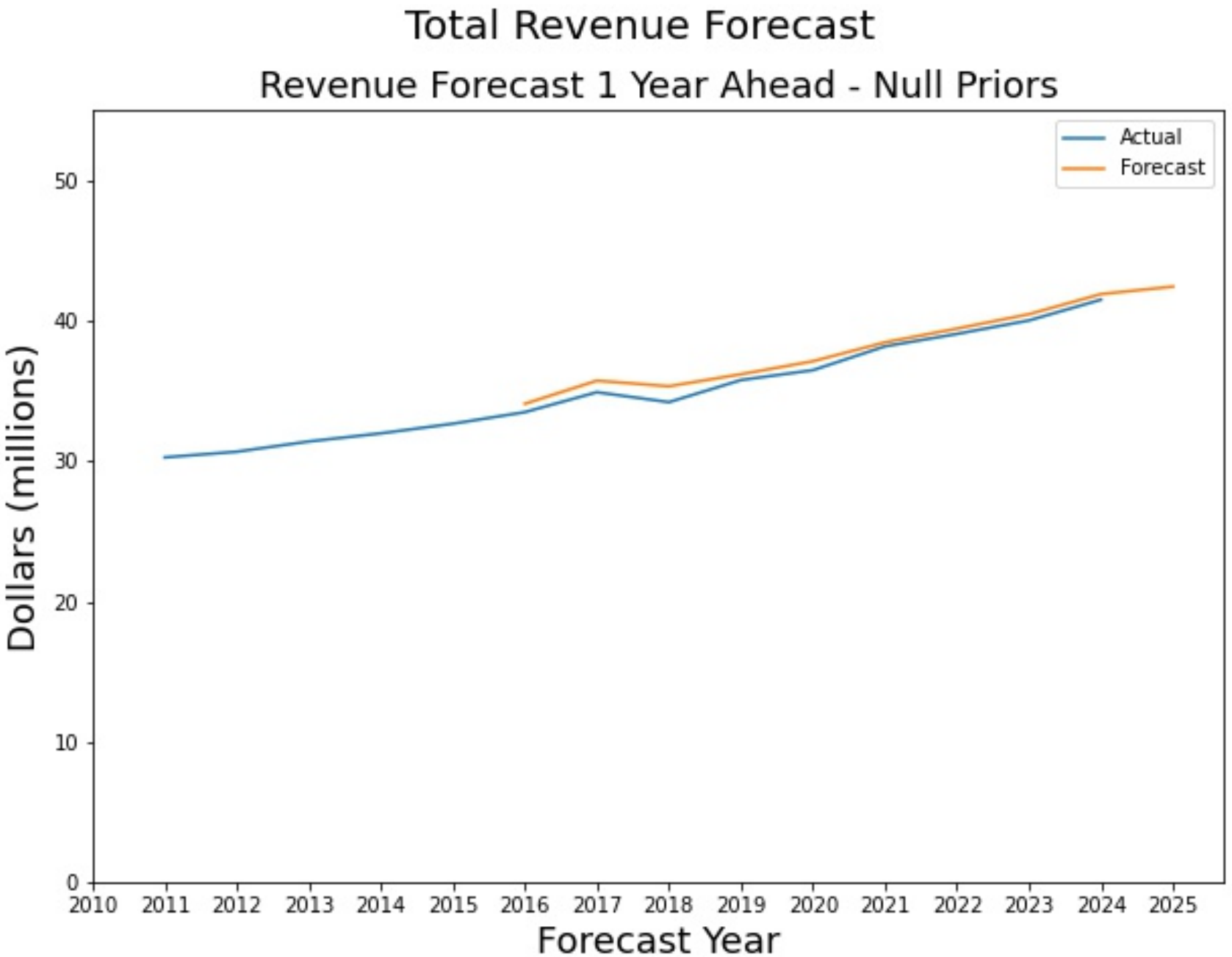
Status of the Project

While our approach looks promising, the implementation is in the early stages and is a work in progress.

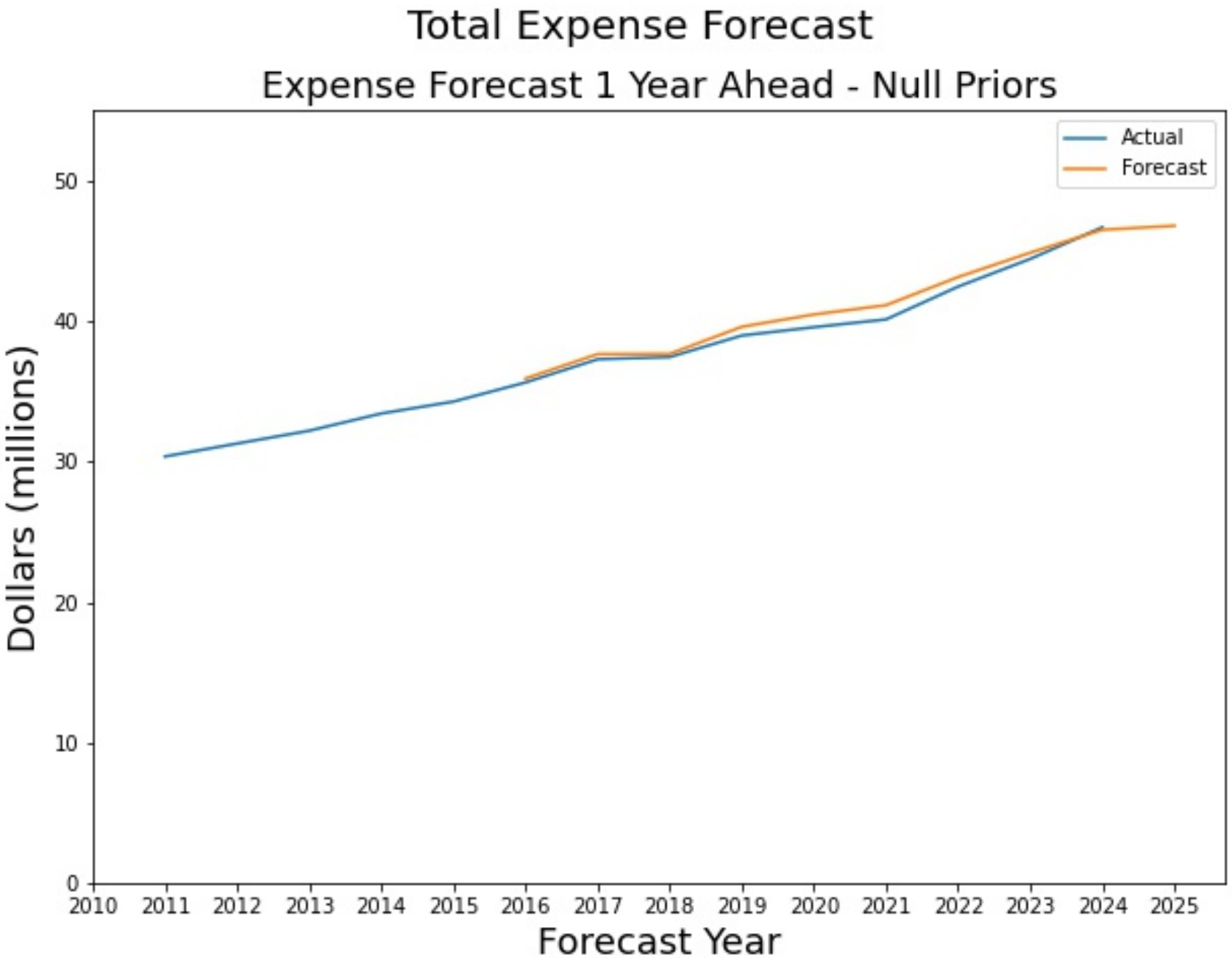
In many cases the details to be worked out are practical considerations rather than mathematical issues. How do we make this work in the environment we find ourselves?

- What kernels work best with this data?
- Stationary or non-stationary models? Or a mixture?
- How and when to use informative priors.
- How to explain the results to oversight bodies.

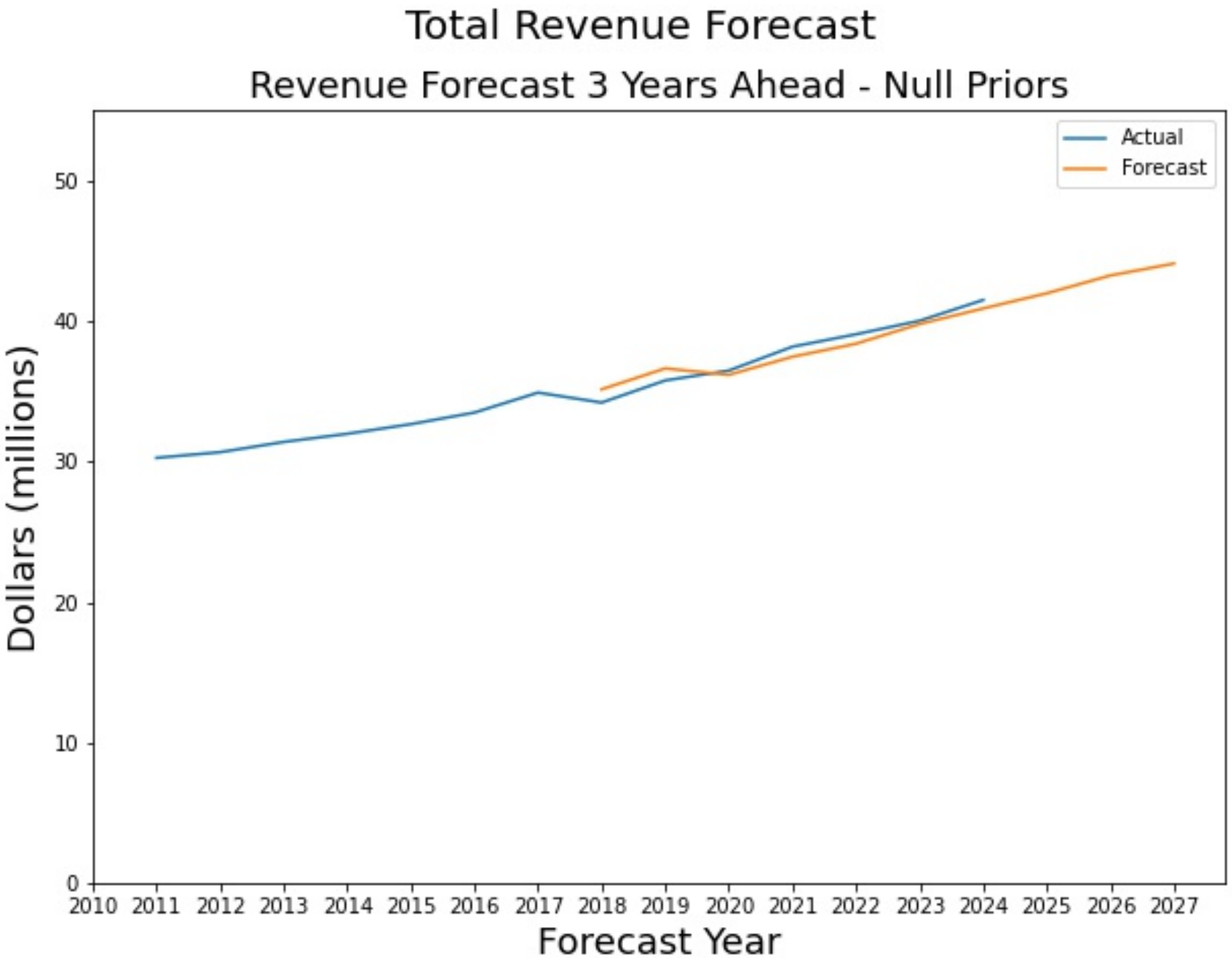
Validation - Total Revenue Forecast - One Year Ahead



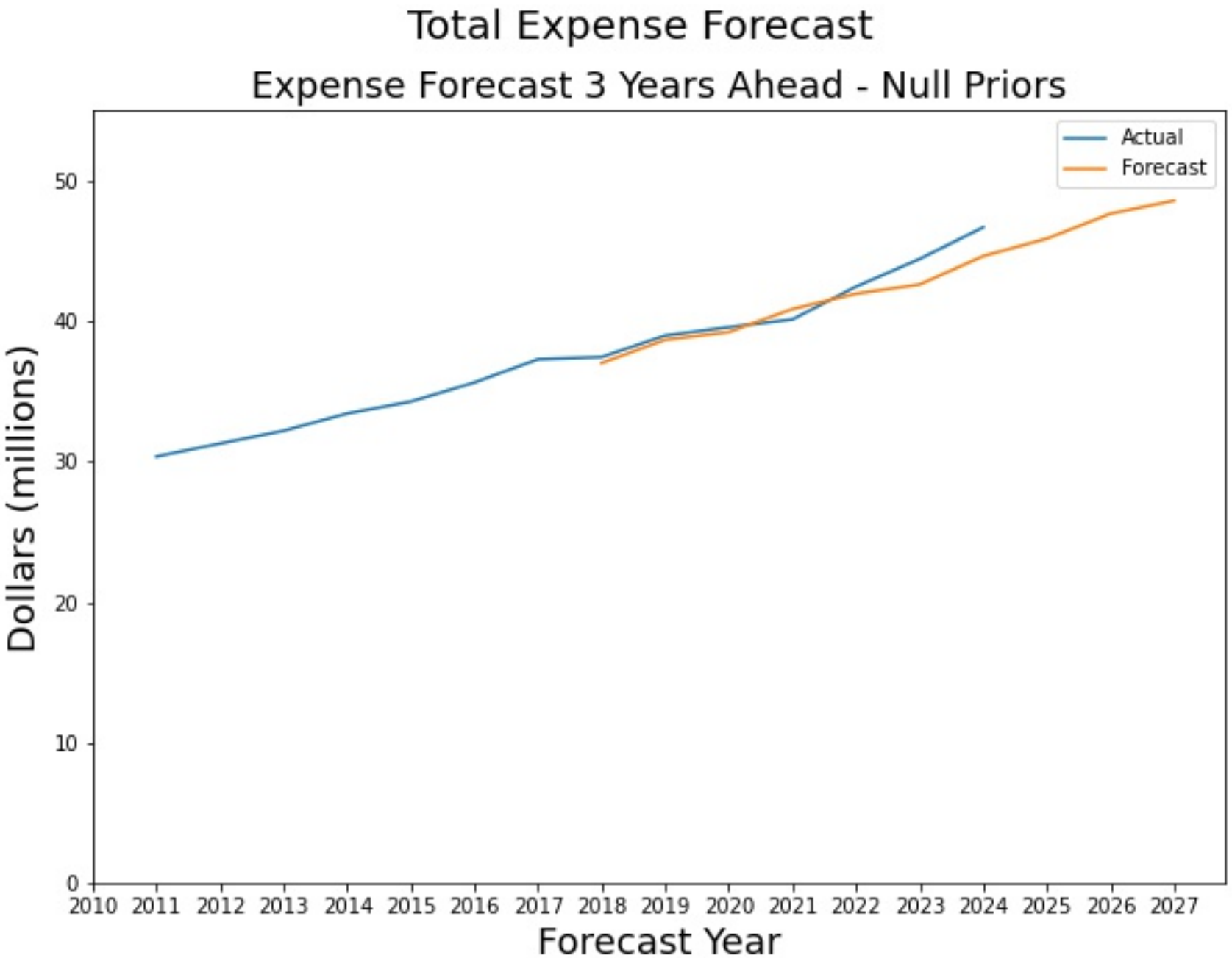
Validation - Total Expense Forecast - One Year Ahead



Validation - Total Revenue Forecast - Three Years Ahead



Validation - Total Expense Forecast - Three Years Ahead



References

[1] Rhode Island Department of Education. *Uniform Chart of Accounts*. URL: https://ride.ri.gov/funding-finance/school-district-financial-data/uniform-chart-accounts.

[2] JUSTIA U.S. Law. *RIGL 16-2-21.2*. URL: https://law.justia.com/codes/rhode-island/title-16/chapter-16-2/section-16-2-21-2/.

[3] Rhode Island Division of Municipal Finance. *What We Do*. URL: https://municpialfinance.ri.gov/about.

[4] Eugene Quinn and Maggie Baker. *Public School Financial Forecasting with Gaussian Process Regression*. URL: https://eugenequinnforegschools.org/NESS_2025_talk.pdf.

[5] C. E. Rasmussen and C. K.I. Williams. *Gaussian processes for machine learning*. MIT Press, 2005.