Columbia River Flow Prediction

Shaun Harrington

2023-07-21

Abstract

River flow forecasts are crucial for efficient hydroelectric operations and maintenance scheduling, particularly in the context of the Columbia River, which is heavily utilized for power generation. This study examined the efficacy of various forecasting methods, including a linear model, Exponential Smoothing model (ETS), Autoregressive Integrated Moving Average model (ARIMA), and an ensemble model in predicting monthly river flows over a year. Using historical river flow data, each model was evaluated based on Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Continuous Ranked Probability Score (CRPS). Surprisingly, the simpler linear model, incorporating a seasonal component and 1-month lag, outperformed the more complex ETS and ARIMA models on all metrics. However, the robustness of these results to long-term changes in climate and weather patterns warrants further investigation. The study concludes by advocating for a tailored approach to forecasting, which takes into account the specific characteristics of river flows and possibly includes other predictive factors, such as climate and weather data.

Introduction

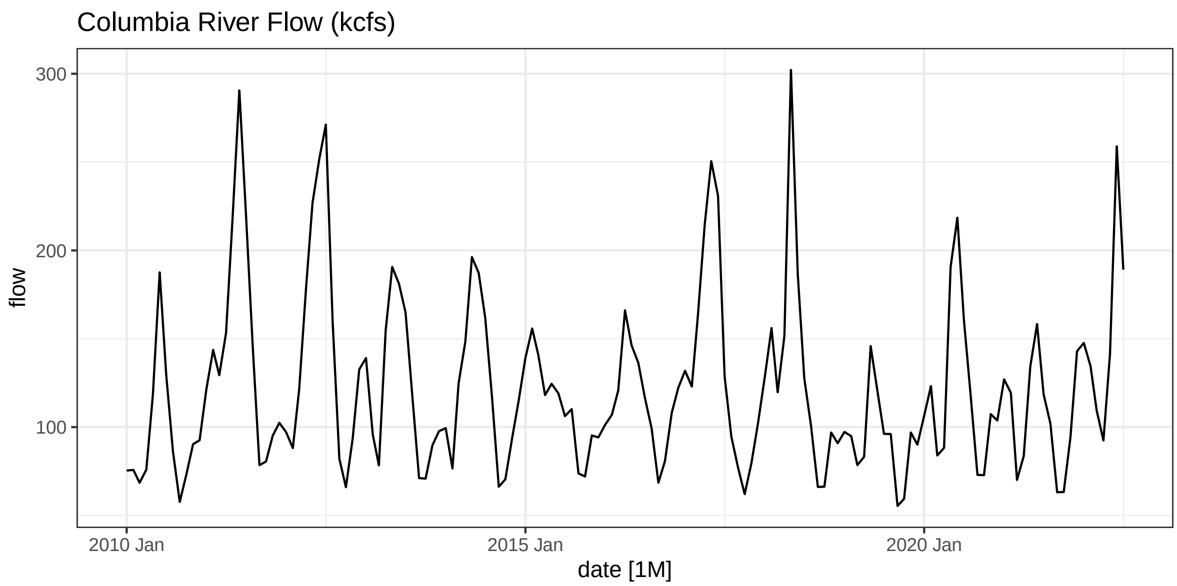
The Columbia River begins in British Columbia, Canada running through Washington state and then along the Oregon border where it empties into the Pacific Ocean. Fourteen hydroelectric dams-11 American and 3 Canadian- are found along it. Their ability to generate power relies heavily on the river's water volume, much of which is sold to California.

While rains can affect the amount of water in the river, the winter snow pack and heat progression throughout Spring and Summer largely determine how much water will flow in the river. The electric utilities that operate these hydroelectric plants must have an understanding of this river to maximize their operations. Short-term river forecasts are essential for daily operations. Hydroelectric plants must maintain their forebay—the body of water just before the dam— within a certain range mandated by the Federal Energy Regulatory Commission. To achieve this, they can manipulate two dials: the flow of water through their turbines and the use of spill gates.

Generally, water is sent through spill gates only when turbines are operating at their full capacity. In the early summer, when river flow is often at its peak, these dams are unable to fully utilize the excess water for power generation. This necessitates the opening of spill gates. Many of these are manually operated, highlighting the need for reliable river flow forecasts to ensure adequate personnel are on hand to manage the operation. Beyond just meeting regulatory guidelines, operators also aim to maintain the highest forebay level possible, as this allows the system to generate electricity more efficiently.

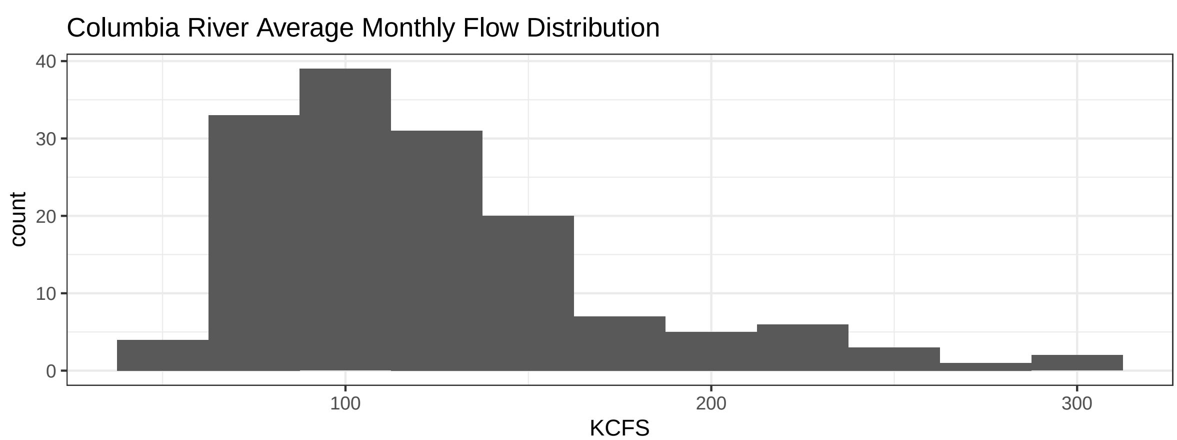
Long-term, these forecasts are necessary in scheduling maintenance of the turbines. During low flow periods of the year, plants have ample capacity to handle the water volume. This allows turbines to be shut down for routine maintenance without forsaking any significant amount of potential revenue.

```r
train %>%
    autoplot(flow) +
    ggtitle("Columbia River Flow (kcfs)")
```



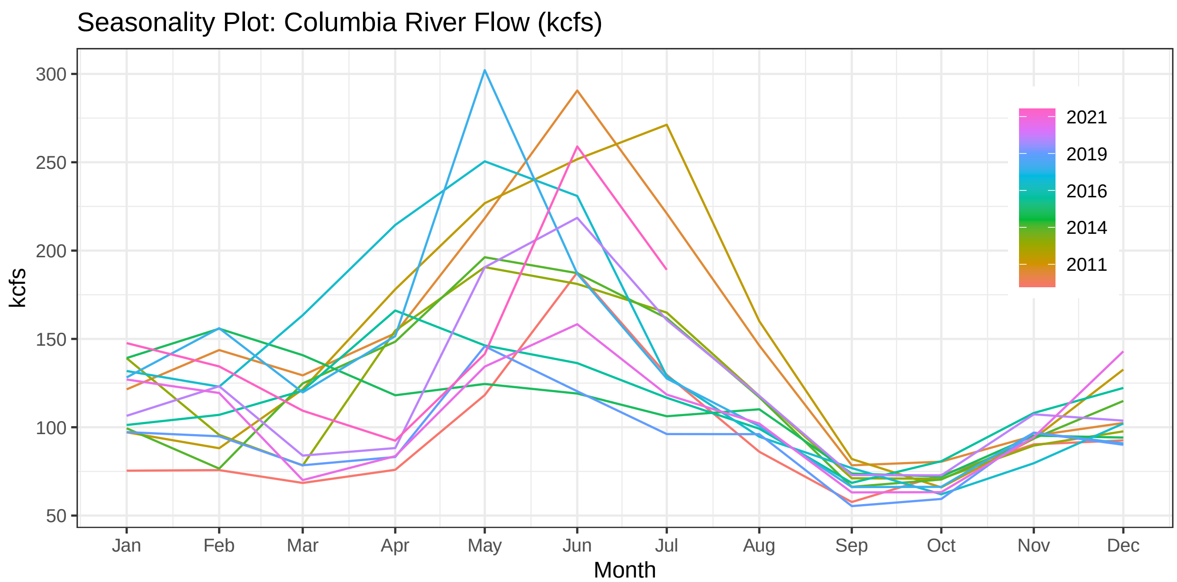
Between 2010 and 2022, the Columbia river has averaged 122.3 kcfs. The minimum in this period was 55.4 kcfs with a maximum of 302.2 kcfs.

```
train %>%
  ggplot(aes(x = flow)) +
  geom_histogram(binwidth = 25) +
  ggtitle("Columbia River Average Monthly Flow Distribution") +
  xlab("KCFS")
```



The seasonality plot below demonstrates how these flows change throughout a year. A complicating matter is the prediction of peak flows. The winter's snowpack determines how much water will flow through the water, but the speed at which Spring warms determines when the flow begins to pick up.

```
train %>%
  gg_season(flow) +
  ggtitle("Seasonality Plot: Columbia River Flow (kcfs)") +
  xlab("Month") +
  scale_y_continuous("kcfs", breaks = seq(0, 400, 50)) +
  theme(legend.position = c(.9, .7))
```



Because of the consequential nature of river flows, the literature abounds in various methods. The National Oceanic and Atmospheric Administration (NOAA) holds the Northwest River Forecast Center (NOAA 2023) that provides the public with 10-day and 120-day river flow forecasts.

Garen (1992) has found an improvement with linear models by implementing three methods: employing principal components regression, using cross validation, and systematic searching for optimal combinations of variables.

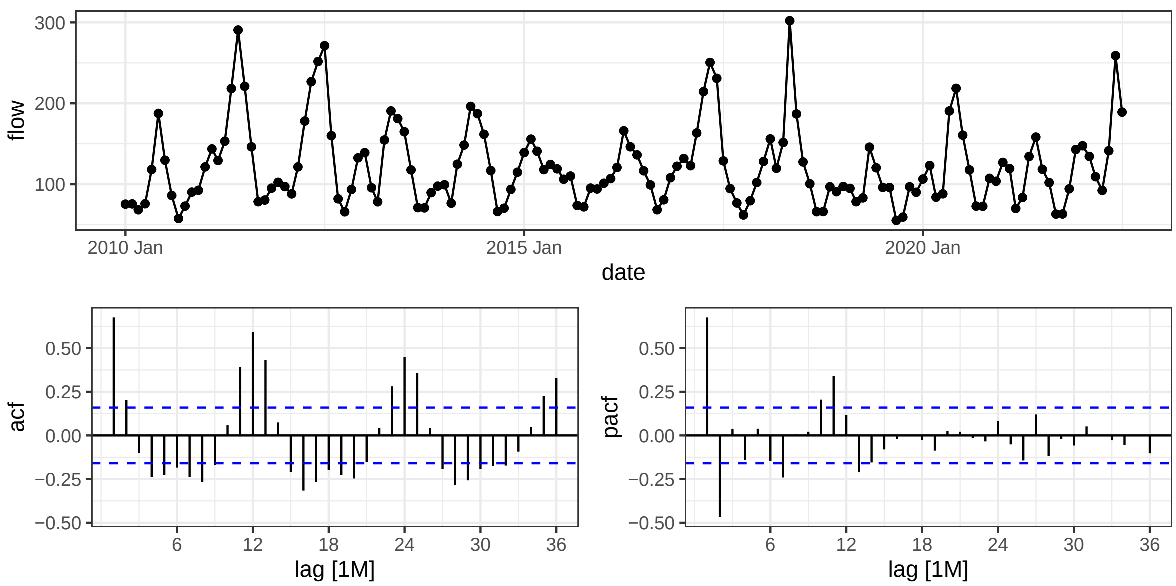
Danandeh Mehr et al. (2022) compares river forecast accuracy using genetic programming (GP), seasonal autoregressive integrated moving average (SARIMA), and an ensemble dubbed GP-SARIMA. The GP and SARIMA were found to both execute well on a day-ahead forecast, though struggled in the longer-term. The GP-SARIMA model was able to reduce the forecast RMSE by 1/4 in the longer-term.

Ilhan (2022) compares 102 machine learning models on day-ahead forecasts. While the models produced a spectrum of results, they were able to decrease the MAE and RMSE by considering an ensemble approach that averages their top performing models.

The approach in this analysis uses classical time series models for predicting one year of monthly river flows. A linear model, ETS, ARIMA, and an ensemble model will be estimated and evaluated on an out-of-sample test set. The river flow data used is provided through the United States Geological Survey (2023) API and aggregated to monthly averages.

Methods

```
train %>%
  gg_tsdisplay(
    #difference(flow, 12) %>% difference(),
    flow,
    plot_type = "partial", lag_max = 36
  )
```

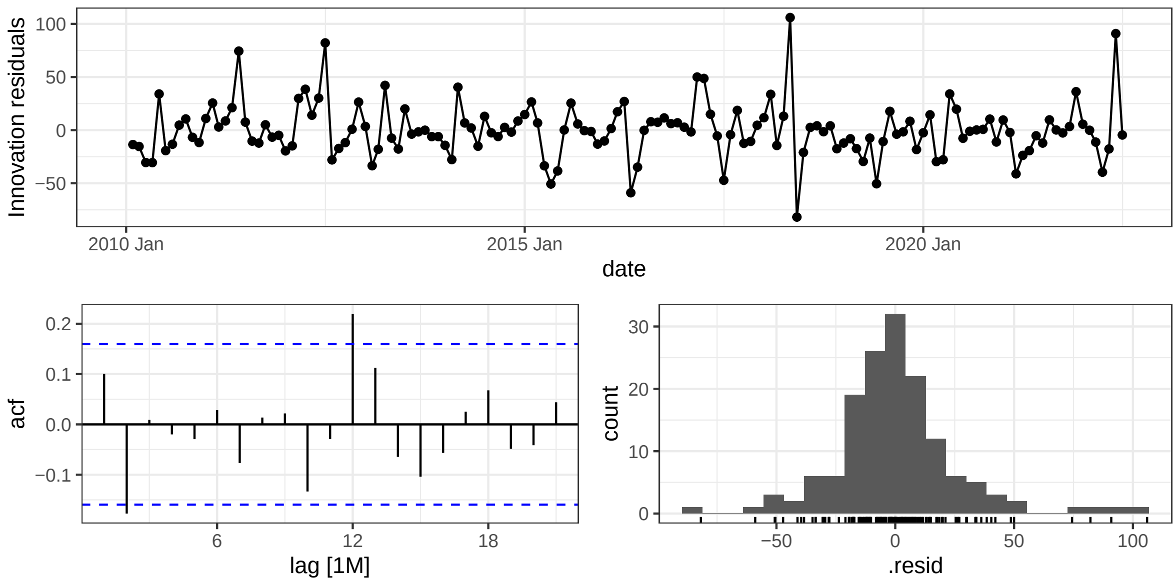


Linear Model

The linear model is a tried and true model used in various situations. For flow prediction, the seasonality, or month, will be used as an exogenous variable due to the highly seasonal nature of river flows. While climate change may introduce a trend to river flows in the very long run, a trend will not be included since we are not making long-term predictions. As such, a model with only a seasonal component will predict the monthly average within the training set for each month in the test set. However, a model with only a seasonal component displays strong autocorrelation within the residuals, suggesting some predictive power is being unused. A 1-month lag will also be included in this model.

```
fit.lm <- train %>%
  model(TSLM(flow ~ season() + lag(flow, 1)))

fit.lm %>%
  gg_tsresiduals()
```

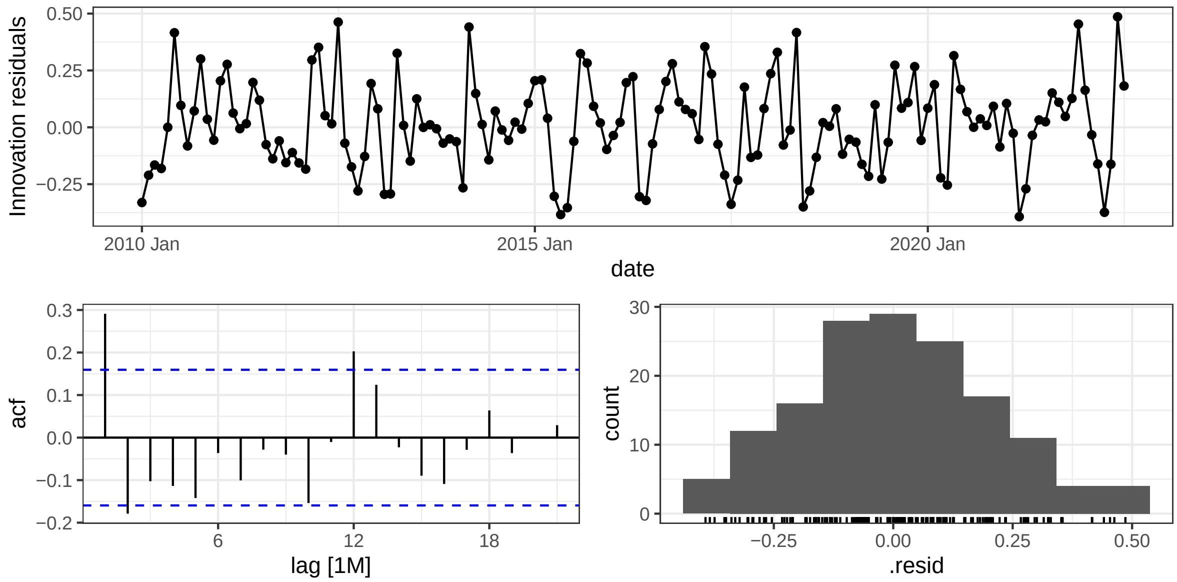


ETS

The ETS model will be estimated for the level and seasonality. Once again, and for the same reasons, a trend will not be included. ETS models do not incorporate exogenous regressors.

```
fit.ets <- train %>%
  model(ETS(flow ~ error() + trend("N") + season()))

fit.ets %>%
  gg_tsresiduals()
```



Though the residuals appear to be mostly white noise, the ACF displays autocorrelation at the 1-month and 12-month lags. This is confirmed with the Ljung-Box test.

```
fit.ets %>%
  augment() %>%
  features(.innov, ljung_box, lag = 12)
```

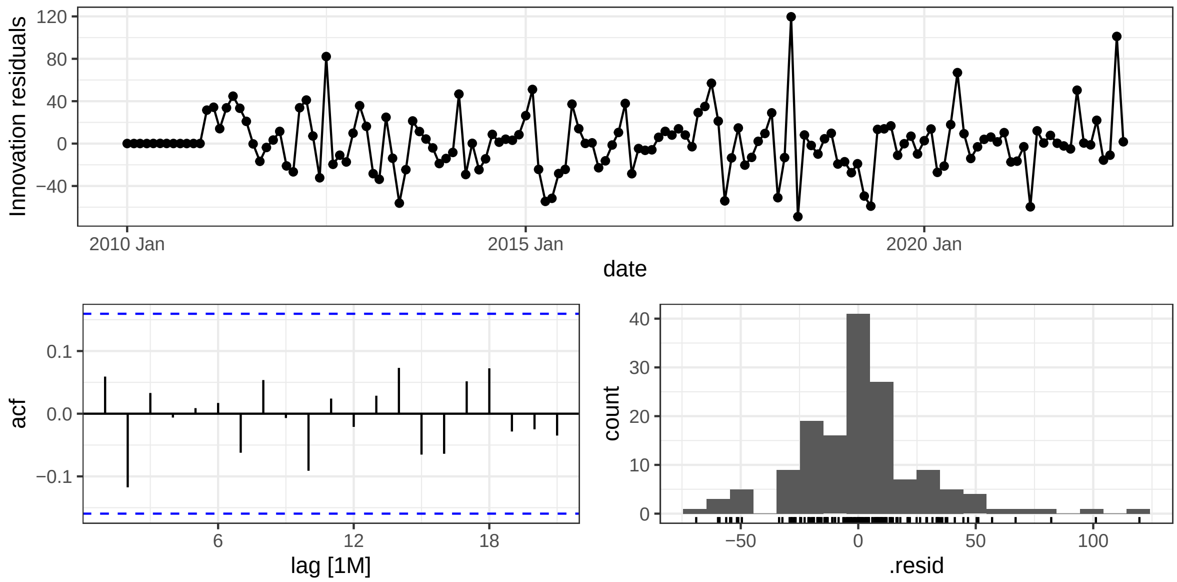
```
# A tibble: 1 x 3
  .model                                          lb_stat lb_pvalue
  <chr>                                             <dbl>     <dbl>
1 "ETS(flow ~ error() + trend(\"N\") + season())"    37.8  0.000165
```

ARIMA

The ARIMA model will be determined through the Hyndman-Khandakar algorithm as specified in the *fable* package.

```
fit.arima <- train %>%
  model(ARIMA((flow)))

fit.arima %>%
  gg_tsresiduals()
```



A visual inspection suggests the residuals are uncorrelated and random, which is also supported by the Ljung-Box test.

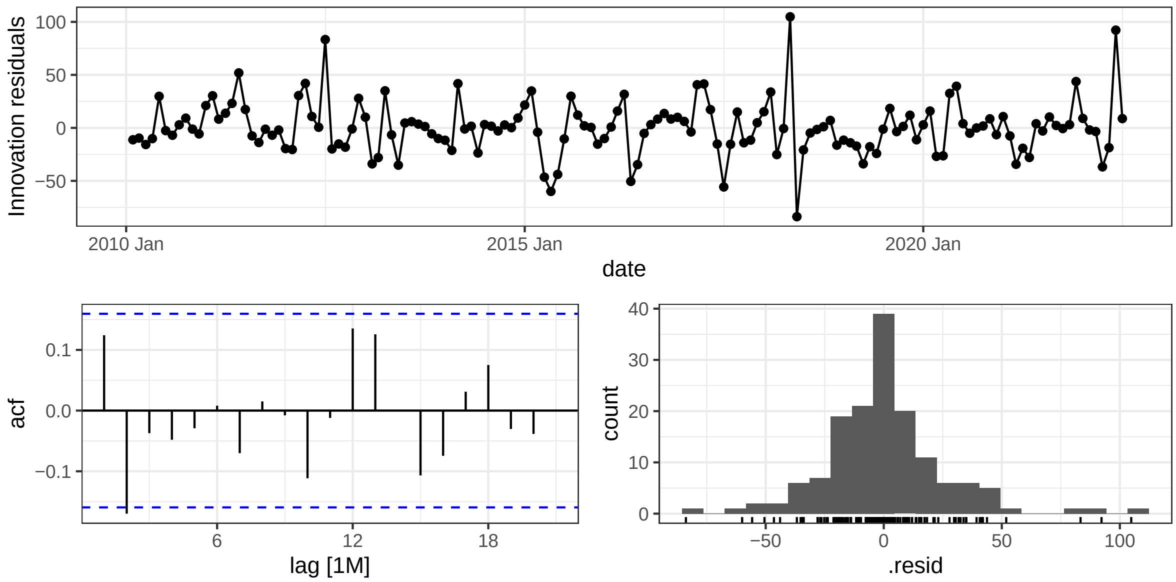
```
fit.arima %>%
  augment() %>%
  features(.innov, ljung_box, lag = 12)
```

```
# A tibble: 1 x 3
  .model         lb_stat lb_pvalue
  <chr>            <dbl>     <dbl>
1 ARIMA((flow))     5.54     0.937
```

Ensemble

The ensemble is a simple average of the prior three models. No egregious problems can be visually inspected in the residuals, though the ACF plot could indicate some autocorrelation is still present.

```
fit %>%
  select(ensemble) %>%
  gg_tsresiduals()
```

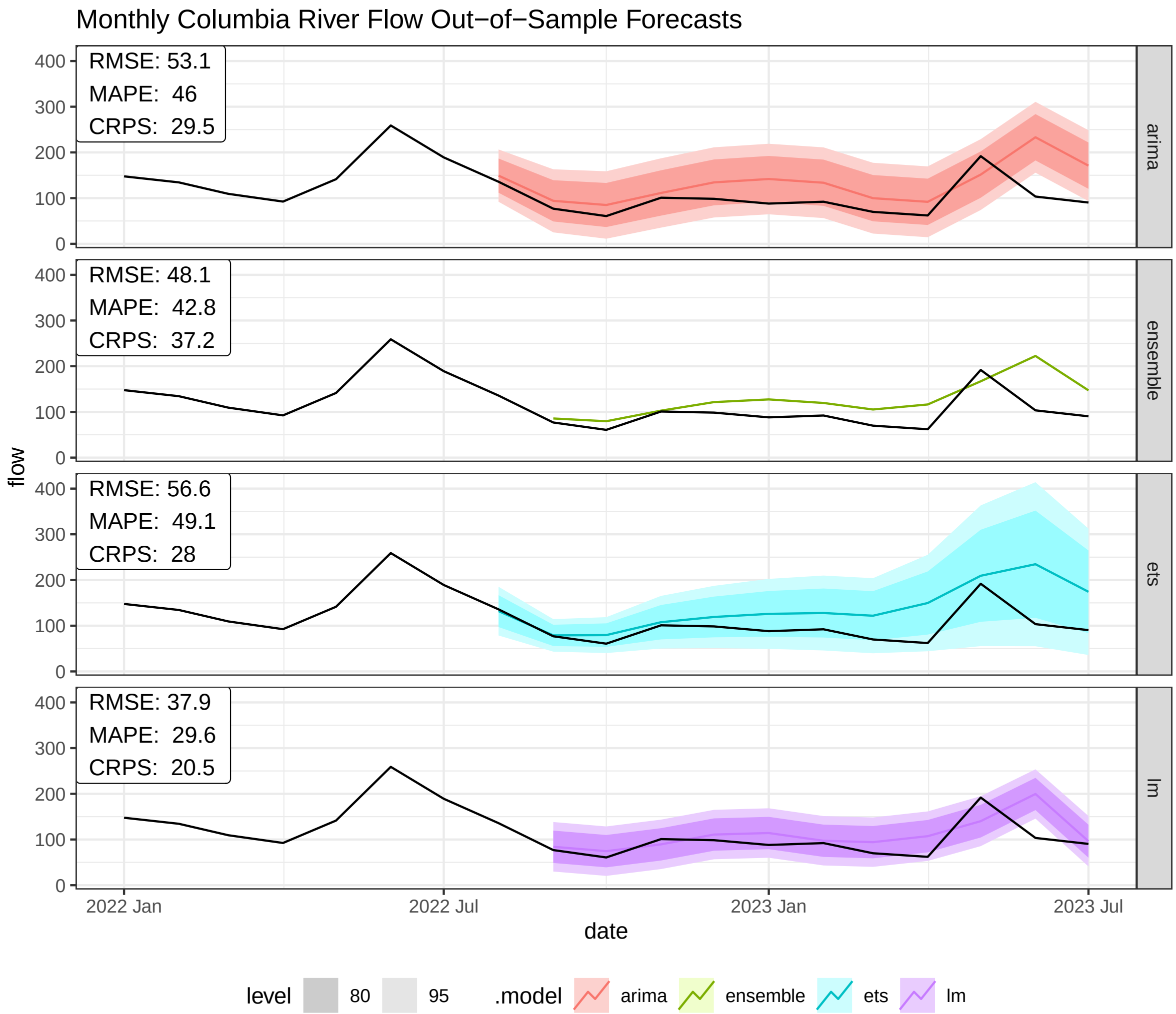


Results

The estimated models produced the following forecasts on the test set:

```
fx %>%
  autoplot(
    data %>% filter(year(date)>2021),
    # level = NULL
  ) +
  ggtitle("Monthly Columbia River Flow Out-of-Sample Forecasts") +
  facet_grid(.model ~ .) +
  geom_label(
    data = fx %>%
      accuracy(
        test,
        measures = list(point_accuracy_measures, distribution_accuracy_measures)
      ),
    aes(x = -Inf, y = Inf, label = paste(
      " RMSE:", round(RMSE, 1),
      "\n", "MAPE: ", round(MAPE, 1),
      "\n", "CRPS: ", round(CRPS, 1)
    )),
    hjust = 0, vjust = 1
  ) +
```

```r
  theme(
    legend.position = "bottom"
  )
```



The linear model outperforms the other models on all metrics (RMSE, MAPE, and CRPS).

The worst model on the point estimate forecast is the ETS model with a *root mean squared error* (RMSE) of 53 kcfs. While it scores higher than the ARIMA on the *continuous ranked probability score* (CRPS), the confidence interval is much greater than that of the other models. However, the linear model and ARIMA have one and two predictions, respectively, that fall outside of their 95% confidence interval. This suggests the models are not properly specified and some improvements could be made by adjusting the models.

ARIMA Statistics:

```r
fit %>%
  select(arima) %>%
  report()
```

```
Series: flow
Model: ARIMA(1,0,0)(2,1,0)[12]
Transformation: (flow)

Coefficients:
         ar1     sar1     sar2
      0.6751  -0.3875  -0.2046
s.e.  0.0636   0.0882   0.0887

sigma^2 estimated as 851.6:  log likelihood=-666.11
AIC=1340.23   AICc=1340.52   BIC=1351.96
```

ETS Statistics:

```r
fit %>%
  select(ets) %>%
  report()
```

```
Series: flow
Model: ETS(M,N,M)
  Smoothing parameters:
    alpha = 0.4867502
    gamma = 0.0001012484

  Initial states:
     l[0]      s[0]     s[-1]     s[-2]     s[-3]     s[-4]    s[-5]    s[-6]
 123.8028 0.8609289 0.7784877 0.5741357 0.5696651 0.9538401 1.258872 1.695363
    s[-7]    s[-8]     s[-9]    s[-10]   s[-11]
 1.512702 1.081726 0.8805814 0.9235965 0.910102

  sigma^2:  0.0425

     AIC     AICc      BIC
1728.484 1732.040 1773.743
```

Discussion

The LM does minimize the RMSE, MAPE, and CRPS. A complicating factor is that while strong seasonality does exist, the amount of water in a given year varies by a good margin as do the springs. A cold spring that suddenly turns hot will cause snow to melt faster causing large peaks. In contrast, a warm spring will cause the snow to begin melting earlier, but not necessarily at a fast rate. The linear model predicts each month to be the historical average with an adjustment considering the prior month, producing a decent forecast.

The ETS and ARIMA models both overestimate the actual flow, likely a result of the high water year for the last period of the training set. The positive ma1 term and negative ma2 term indicate that when there is this large change in the flow, the following month is usually the same direction, but the month after that results in a larger contraction in the opposite direction. This lines up with the warm/cold spring theory postulated in the previous paragraph.

A gamma value near zero in the ETS model indicates that river flows do not respond very much to changes in the prior season. The alpha equal to almost .5 indicates some response to recent months, but not overly so. Together these help explain why the linear model outperformed the other two, more complicated models: the prior years' values are not a good indicator for the current year, and recent months can only go so far. When no exogenous variables are considered, this is nearly a random walk, for which the historical mean is the best approximate.

Conclusion

This research has analyzed the efficacy of different forecasting models in predicting the flows of the Columbia River. Efficient operations of the hydroelectric dams requires reliable estimates of future river flows. This paper compares the predictive power of the linear model, ETS, ARIMA, and an ensemble approach to forecast 12 months of river flows.

The linear model with a seasonality component and 1-month lag outperformed the more sophisticated ETS and ARIMA models on all metrics. This performance is likely due to the difficult nature of predicting river flows. While the linear model performed best on this sample, abnormal years will result in a worse performance. Further study should consider a longer time span and use cross validation to adequately compare the different methods.

Though the ETS and ARIMA overestimated river flows, each offer valuable insights into the nature of the river. The ma terms from the ARIMA model indicate how flows might react to sudden changes in temperature, which could help mitigate the damage from extreme weather events. The ETS model's parameters indicate that the river is not overly responsive to recent periods, suggesting a near random walk nature of the time series.

This comparison demonstrates the benefits of using multiple models for predictive analysis and also show the need for careful inspection of the story that each tells. Exogenous regressors should be included in future studies. An estimate of snow pack, total precipitation, and temperature or heating degree days may be beneficial. As snow pack is a stock measure, accumulated heating degree days could prove useful to measure the amount of spring run off. Rainfall could aid short-term forecasts to see the impact it has to flows.

References

Danandeh Mehr, Ali, Sahand Ghadimi, Hannu Marttila, and Ali Torabi Haghighi. 2022. "A New Evolutionary Time Series Model for Streamflow Forecasting in Boreal Lake-River Systems." *Theoretical and Applied Climatology* 148 (1): 255–68. https://doi.org/10.1007/s00704-022-03939-3.

Garen, David C. 1992. "Improved Techniques in Regression-based Streamflow Volume Forecasting." *Journal of Water Resources Planning and Management* 118 (6): 654–70. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:6(654).

Hyndman, Rob. 2023. *Fpp3: Data for "Forecasting: Principles and Practice" (3rd Edition)*. https://CRAN.R-project.org/package=fpp3.

Ilhan, Akin. 2022. "Forecasting of River Water Flow Rate with Machine Learning." *Neural Computing and Applications* 34 (22). https://doi.org/10.1007/s00521-022-07576-9.

NOAA. 2023. "Northwest River Forecast Center." https://www.nwrfc.noaa.gov/rfc/.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

United States Geological Survey. 2023. "National Water Information System API." https://nwis.waterservices.usgs.gov/nwis/iv/?format=rdb&sites=12472800&startDT=2010-01-01&endDT=2023-07-02&parameterCd=00060,00065&siteStatus=all.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2023. *Tidyverse: Easily Install and Load the Tidyverse*. https://CRAN.R-project.org/package=tidyverse.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2023. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. https://CRAN.R-project.org/package=ggplot2.