

Assessing the Impacts of Environmental and Ecological Variables on the Performance of Fraser
Sockeye Salmon Forecast

Yi Xu¹, Qi Liu², Caihong Fu¹, John Holmes¹

¹ Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo,
BC, Canada

² Fisheries and Oceans Canada, Pacific Region Head Office, 401 Burrard Street, Vancouver, BC,
Canada

Corresponding author:

Yi Xu

Email : xuyiouqd@gmail.com

ORCID: <https://orcid.org/0000-0002-9902-9588>

Abstract

The Canadian Fraser River sockeye salmon (*Oncorhynchus nerka*) is one of the largest stock complexes in North America, supporting major commercial, recreational, and First Nations fisheries. Sockeye fisheries management relies on an annual pre-season forecast of adult returns. In this study, we developed a framework of good visualization with Taylor diagram to evaluate pre-season forecast models on an annual basis and identify external drivers of importance for forecasting sockeye returns. Specifically, we incorporated five new covariates including sea surface temperature in the Gulf of Alaska and other salmon species abundance into the existing forecast models. Results revealed good performances by both Ricker and Power models coupled with the newly included covariates. In addition, models selected more than a decade ago underperformed compared to those selected based on our recent retrospective analysis from 2009-2020. We advocate that forecast models need to be continuously evaluated in the face of environmental change, and that models incorporating non-stationary ecological processes are to be developed with environmental and ecological factors being assessed for their impacts on Fraser sockeye dynamics.

Key words: Fraser sockeye, environmental covariates, *Oncorhynchus nerka*, pre-season forecast, salmon abundance, Taylor diagram

1. Introduction

Sockeye salmon (*Oncorhynchus nerka*) make significant economic, ecological, and social contributions throughout the Northeast Pacific basin. The sockeye salmon complex in the Fraser River, British Columbia, Canada, is one of the largest stock complexes in North America, which spawn in hundreds of distinct locations throughout the 220 000 km² drainage basin (Figure 1; Burgner 1991; DFO 2018). This stock complex has supported major commercial, recreational, and First Nations marine and freshwater fisheries (Ruggerone and Connors 2015), making the sustainable management of these sockeye salmon stocks imperative. The management system of these sockeye salmon consists of pre-season forecast of sockeye returns (a combination of sockeye to be caught by fisheries and those escaping to spawning ground), in-season management and post-season review. Pre-season forecasts are essential and have become mandatory for Fisheries and Oceans Canada (DFO) under a Pacific Salmon Treaty that initiated in 1985 (PSC 2022). Pre-season forecasts of Fraser sockeye returns are used to plan fishery opportunities by fisheries managers and the fishing industry, and feed to in-season assessment modeling (Haeseker et al. 2008; Michielsens and Cave 2019; DFO 2021). Additionally, pre-season forecast assists decision-making processes including international total catch allowance negotiation, allocation of stock assessment resources, and hatchery enhancement planning (Hawkshaw et al. 2020a, b).

Fraser sockeye salmon returns are highly variable not only due to the cyclic nature for some large stocks (with a dominant year preceded and followed by three years of low returns), but also unpredictable productivity at different life stages in both freshwater and marine environments (Ricker 1997; Akenhead et al. 2016a; Huang et al. 2021), which has made pre-season forecasts of sockeye returns challenging.

Despite the tremendous importance and challenges of pre-season Fraser sockeye forecasts, a rigorous evaluation of forecast performance in terms of both accuracy and uncertainty has not been completed. In particular, confidence in pre-season forecasts of Fraser sockeye returns has eroded among fisheries managers and harvesters in recent years because observed returns have frequently diverged outside of the 80% prediction interval (i.e., <10 or >90 of posterior quantiles in Bayesian statistical approaches; DFO 2021). Large differences between pre-season forecasts and observed returns have led to potentially reduced fishery opportunities, missed management targets and escapement goals, as well as increased conservation concerns for Fraser sockeye salmon stocks. Therefore, it is imperative to improve forecasting ability for Fraser sockeye; in particular, it has become necessary to incorporate new information derived from recent scientific research into the forecast process such that concurrent climate change is reflected in the forecast models.

Sockeye returns in any given year are influenced by the abundance of their parental spawners (i.e, escapement or sockeye returns that reach the spawning grounds), the proportions of age classes, and the survival rate of the adult recruits across the entire life cycle from egg to adult (Grant et al. 2010; Hawkshaw et al. 2020a, b). A wide variety of forecast models are typically available at a stock level for Fraser sockeye from non-parametric ‘naïve’ models (i.e., models using recent and historical adult recruits only without considering spawners) to complex stock-recruitment models that incorporate environmental variables into the relationships between spawners and subsequent recruitment (Cass et al. 2006; Grant et al. 2011; DFO 2018; Hawkshaw et al. 2020a, b). More advanced forecast methods have also been developed that use Bayesian statistical approaches to assess uncertainties associated with abundance estimates (Cass et al. 2006; Grant et al. 2011; Akenhead et al. 2016a). Model selection for Fraser sockeye forecasts has

83 been an expert-driven process, primarily adopting forecast models identified based on Jack-knife
84 cross validation (i.e., analysis being carried out by randomly leaving out one year's data at a
85 time) conducted for the period of 1997 to 2004 (Grant et al. 2011).

86 Since the 1990s, Fraser sockeye has generally declined with several populations being
87 assessed as endangered by the Committee on the Status of Endangered Wildlife in Canada
88 (COSEWIC; COSEWIC 2017). Forecast models that were selected by Grant et al. (2011) more
89 than a decade ago have been continuously used in pre-season forecasts despite the fact that their
90 explanatory power was found to be low (Akenhead et al. 2016a; DFO 2021). Moreover,
91 environmental covariates such as coastal sea surface temperature (SST) and the Pacific Decadal
92 Oscillation (PDO) that were shown to be effective in reducing the degree of unexplained stock-
93 recruitment variation in the past two decades (Cass et al. 2006), are no longer applicable for most
94 stocks (Litzow et al. 2020). These time-varying effects of environmental conditions on
95 population and community processes, i.e., non-stationary relationships, are becoming more
96 prevalent in the face of climate change, and have been increasingly recognized (e.g., Litzow et
97 al. 2018, 2020; Ohlberger et al. 2022). Such non-stationarity requires that previously established
98 Fraser sockeye forecast models be continuously evaluated through annual retrospective analysis
99 before carrying out pre-season forecasts. Furthermore, external drivers that are potentially
100 responsible for the interannual and long-term variability of sockeye abundance and productivity
101 continue to be explored and characterized in a changing environment (McKinnell 2008;
102 Ruggerone et al. 2021).

103 In this study, we developed a framework that can be easily used by fisheries managers or
104 other stakeholders to evaluate pre-season forecast models and to identify external drivers of
105 importance for forecasting Fraser sockeye salmon returns on an annual basis. Specially, we

developed computer codes to conduct annual retrospective analysis of multiple Fraser sockeye forecast models and for the first time employ Taylor diagrams (Taylor 2001) with good visualization to display simultaneously three statistics of predictive power, including Pearson's correlation coefficient (denoted as correlation), normalized standard deviation and normalized root-mean-square error (RMSE). Normalization is calculated as dividing both the RMSE difference and the standard deviation of model prediction for a set of years by the standard deviation of observed abundance for the same set of years. The Taylor diagram is a useful visualization tool that has been employed to evaluate the performance of a variety of models such as those for ocean circulation (Lamine et al. 2022), satellite derived chlorophyll a measurements (Lee et al. 2015), biogeochemical dynamics (Salihoglu et al. 2017), and stock assessment and population dynamics of tuna species (Kell et al. 2016; Inna et al. 2019). The framework established through this study can be directly implemented in the process of Fraser sockeye forecast so as to improve the performance of sockeye forecasts and assist stock assessment and fisheries management planning.

2. Materials and methods

2.1 Fraser sockeye spawner and recruit data

In this study, we used the same spawner and recruitment data of Fraser sockeye stocks within the Fraser River watershed (Figure 1) that were approved for the 2022 pre-season forecast (DFO, in press). The spawner data have been collected and maintained by DFO since 1938 (Grant et al. 2011). The recruitment data have been compiled by Pacific Salmon Commission (PSC) since 1952. Similar to Grant et al. 2010, in this study, we treated the recruitment data as observation to represent the true abundance of sockeye returns to the Fraser River mouth. The details on data quality are available in Ogden et al. (2015). Based on shared timing of migration and their entry

into the Fraser River, the stocks are aggregated into four management units: Early Stuart, Early Summer, Summer, and Late Run. In this study, we focused on 18 ‘lake-type’ sockeye stocks that typically spend two years in a nursery lake before migrating seaward and return to their natal rivers at age 4 or 5.

For all stocks, we chose female spawners from the 1948 to 2018 brood years as an indicator of spawner abundance (also called effective female spawners). Effective female spawners (EFS) were defined as “the product of the total female spawners multiplied by female spawner success, where female spawner success was calculated as the proportion of eggs (0%, 50%, or 100%) successfully spawned, based on spawning ground carcass surveys” (Grant et al. 2011; Ogden, et al. 2015). Although 50% in spawning success is arbitrary and covers a wide range of partially spawned females, this category is reasonable for estimating effective female spawners, as long as the partially spawned females are normally distributed and the samples are representative of all partially spawned females (i.e., 50% spawned fish represents half of fully spawned fish and half un-spawned fish, Lingard et al., 2013; Stuart LePage, per. Comm.). For example, if 7000 females are observed to be fully spawned, 1000 females are 50% spawned and 2000 are dead un-spawned, the effective female spawners are 7500 ($7000 + 1000 \times 50\%$), the spawn percentage is 75% ($\frac{7500}{7000+1000+2000}$), and the pre-spawn mortality is 25% ($1 - 75\%$). Recruits included both sexes with the most recent recruitment data (4- and 5-year-old) available up to the 2017 brood year (i.e., the primary brood year returning in 2021). For the Cultus Lake sockeye stock, juvenile abundance estimated during outmigration at a counting fence instead of spawner abundance was used to build the stock-recruitment relationship, because the quality of the adult escapement data is variable and they are heavily affected by enhancement activities (Ackerman et al. 2014).

Similar to Grant et al. 2010, we assumed age composition of Fraser sockeye adults generally followed a similar pattern among brood years around the historical average, which was typically dominated by age 4 year class. For example, to get the total recruits of 2022, we used 2018 brood year EFS to produce a total recruits multiplied by the average stock-specific proportion of age-4 recruits, and used 2017 brood year EFS to produce a total recruits multiplied by the average stock-specific proportion of age-5 recruits. Ten of the eighteen stocks had time series of paired spawner and recruit estimates from brood years going back to 1948, and the remaining eight stocks had shorter time series (Nadina: 1973; Gates: 1968; Scotch: 1980; Fennell: 1967; Weaver: 1966; Portage: 1953), depending on the availability and quality of the data as well as consistency with previous pre-season forecasts (DFO 2018).

2.2 Environmental and ecological covariates

Previous pre-season forecasts have incorporated a few local and large-scale environmental covariates in the Northeast Pacific Ocean, including (1) PDO in the winter preceding outmigration (November-March, Mantua et al. 1997), (2) monthly average SST (April-June) from Entrance Island lighthouse (Ei.SST, Strait of Georgia, near Nanaimo, BC, Canada), (3) monthly average SST (April-July) from Pine Island (Pi.SST, Northeast corner of Vancouver Island) of the year of outmigration, (4) peak Fraser River discharge (FRD.peak), and (5) average Fraser River discharge (FRD.mean) from April to June of the outmigration year, both measured at Hope, BC, Canada (for data sources please refer to DFO 2018; Hawkshaw et al. 2020a, b). All time series of environmental indices were aligned with smolt outmigration year because that is when they are most influential on marine survival of sockeye and other Pacific salmon species (Cass et al. 2006).

Exploration of new biological and environmental covariates that may explain the interannual variability in Fraser sockeye recruitment has long been suggested (Grant et al. 2010; MacDonald and Grant 2012; DFO 2015; Hawkshaw et al. 2020a, b). Recent analyses found that survival of Fraser sockeye exhibited a similar temporal pattern among all stocks, indicating that Fraser sockeye stocks might be related to oceanic environmental conditions and competition among Pacific salmon at a global scale (Akenhead et al. 2016a; Connors et al. 2020; DFO 2021; Ruggerone et al. 2021; Rosengard et al. 2021). Therefore, we included Gulf of Alaska SST (GOA.SST) time series, which was extracted from the COBE SST database (<https://psl.noaa.gov/data/gridded/data.cobe.html>), matching with the Fraser sockeye distribution areas identified based on tagging studies (Myers et al. 1996; Ishii et al. 2005; Supplementary Figure S1). We also included the total abundance (i.e., catch plus escapement) of Pacific salmon returning from the North Pacific Ocean to streams in Asia and North America as covariates in the stock-recruitment models. Specifically, we chose annual abundance of adult pink, chum, and sockeye stocks, respectively, and the sum of all three salmon species abundance (Salmon.Total) from 1950 – 2020 (brood year 1948 – 2018) as covariates, as ocean entry year is believed to be the most important time for sockeye marine survival (Ruggerone and Irvine 2018; Ruggerone et al. 2021). Altogether, ten covariates (six environmental and four ecological) were included in the pre-season forecast models (Figure 2), hypothesizing that inclusion of these covariates can help account for more of the environmental effects on interannual variability in the survival of all Fraser sockeye stocks.

2.3 Forecast models

In this study, we re-evaluated the previously explored models, including eleven naïve and six biological models (Table 1 and Table S1). The biological models (or Stock-Recruit models) of

Ricker, Power, and Larkin were also applied to data of dominant years only (i.e., RickerCyc, PowerCyc, and LarkinCyc) for the purpose of understanding whether or not the dominant year class was driven by different biological and environmental processes. With the Ricker and Power models, we also coupled them with each of the ten covariates to form ten new models. For model descriptions of the 37 pre-season forecast models, please refer to Table 1 and Table S1.

Following previous forecasts (e.g., Cass et al. 2006; Grant et al. 2011; MacDonald and Grant 2012; DFO 2018; DFO in press), we used a Bayesian statistical approach to estimate biological model parameters and Markov Chain Monte Carlo (MCMC) to assess estimation uncertainties. Model specifications are documented in Supplement tables (Priors.csv and mcmc_specs.csv).

2.4 Retrospective analysis and model evaluation

To evaluate the 37 forecast models for the 18 major stocks, we conducted a one-step-ahead retrospective analysis (i.e., iteratively stepping forward through time as each step added a new year to the estimation dataset) for the period between 2009 and 2020. For each stock, all 37 forecast models were ranked using each of the following four performance measures related to forecast error: mean raw error (MRE), mean absolute error (MAE), mean percent error (MPE), and root-mean-square error (RMSE). These four metrics were used in previous studies for Fraser sockeye forecast evaluations (Haeseker et al., 2008; Grant et al., 2010). MRE is the sum of raw errors between forecast (y) and observation (x) divided by the time length T : $MRE = \frac{\sum_{i=1}^T (y_i - x_i)}{T}$. MAE is the sum of absolute errors divided by the time series length T : $MAE = \frac{\sum_{i=1}^T |y_i - x_i|}{T}$ (Willmott and Matsuura 2009). MPE is the averaged percentage difference between forecast and observation over time T : $MPE = \frac{100\%}{T} \sum_{i=1}^T (\frac{y_i - x_i}{x_i})$. RMSE provides a measure of forecast error

217 variance reflecting both bias and uncertainty: $RMSE = \sqrt{\frac{\sum_{i=1}^T (y_i - x_i)^2}{T}}$ (Hyndman and Koehler,
218 2006). For each model, ranks across the four performance measures were averaged to generate
219 an overall rank. Since MRE and MPE can be negative, we took the absolute value before ranking
220 them, and they are noted as Abs(MRE) and Abs(MPE). We then calculated relative rank based
221 on the overall rank to make comparisons across stocks: *Relative rank* =
222 $\frac{\text{Overall rank}}{\text{Numbers of model evaluated}}$. We also provided individual relative rank for Abs(MRE), MAE,
223 Abs(MPE) and RMSE. This approach has been used by other similar studies (Haeseker et al.
224 2008; Grant et al., 2010). In addition, we used the normalized forecast metric (NFM) to measure
225 the bias: $\sum_{i=1}^T \frac{y_i - x_i}{y_i + x_i}$ ([https://demand-planning.com/2021/08/06/a-critical-look-at-measuring-and-](https://demand-planning.com/2021/08/06/a-critical-look-at-measuring-and-calculating-forecast-bias/)
226 [calculating-forecast-bias/](https://demand-planning.com/2021/08/06/a-critical-look-at-measuring-and-calculating-forecast-bias/)). A positive NFM indicates over-forecast while negative under-forecast with
227 0 indicating the absence of bias.

228 We used Taylor diagrams as a visualization tool to guide the process of model selection both
229 in terms of accuracy and uncertainty. The correlation (with values ranging from -1 to 1) between
230 forecasts hindcasted by each model for the years 2009 – 2020 and the observed returns
231 represents forecast accuracy. Correlation values greater than 0.75 indicate high forecast accuracy
232 while smaller values particularly negative values imply inaccurate forecast. The normalized
233 standard deviation represents the magnitude of the forecast variation relative to the observation
234 variation for the given years. The observed sockeye returns from 2009 – 2020 (3 cycles for some
235 major stocks; Supplementary Figure S2) were normalized as a reference point (i.e., observation).
236 The overall performance of a model was expressed as a relative position on the diagram to the
237 observation. Generally, a model that has relatively high correlation, small normalized RMSE,
238 and similar normalized standard deviation as observations is desired. A perfect model is located

at the reference point with both correlation and normalized standard deviation being 1 and normalized RMSE being 0. All models, analyses and visualization were programmed using R 4.2.0 (R Core Team 2022).

3. Results

3.1 Forecast model comparisons

Relative rank based on overall rank of the 37 forecast models across the 18 stocks showed clear patterns of generally poor performances by the naïve models (with the exceptions of LLY and R1C) but relatively good performances by both Ricker and Power models coupled with the newly included covariates (i.e., GOA.SST, Sockeye, Chum, Pink and Salmon.Total) (Figure 3). Relative rank based on Abs(MRE), MAE, Abs(MPE), and RMSE showed similar patterns to those based on overall rank (Supplementary Figure S3 a-d). It is worth noting that the two naïve models (LLY and R1C) and two Ricker models coupled with new covariates (RickerPink and RickerSalmon.Total) had relatively smaller biases with NFM near zero (Supplementary Figure S3e).

Correlations between model forecasts and observations were generally lower for the naïve models than biological models (Figure 4). Correlations were more homogeneous within each stock (across all models) than across the stocks (using the same model). This finding may indicate that model accuracy can be stock-dependent as a consequence of either data quality or specific stock dynamics. The small stocks of Gates and Raft had particularly low correlations across all models, possibly associated with inaccurate yet more economical survey methods applied to insignificant stocks (DFO, 2021). Similarly, standard deviation revealed the disadvantage of the naïve models, which tend to be farther away from 1 (Supplementary Figure S4). Interestingly, naïve models (in particular, LLY model) perform the best among others for

these stocks with the correlation 0.3-0.5. Standard deviation also showed a high degree of consistency across all the models other than the naïve models, particularly for the three stocks, Nadina, Stellako, and Late Stuart, for which the model forecasts either didn't capture uncertainty or had inflated variability compared to observations.

For each of the 18 Fraser sockeye stocks, we presented the retrospective hindcast results (for the period 2009 – 2020) of the 37 forecast models on a Taylor diagram relative to the time series of observations (shown as a black solid circle; Figure 5a,b). The historical forecast based on previously selected models was also presented (shown as a black solid square; Figure 5a, b) relative to the observation. In general, the historical forecast performed poorly compared to many other models. Correlations between the historical forecast and the observation were less than 0.5 for half of the stocks with some close to or even below zero. For all stocks, there was at least one model that outperformed the historical forecast, indicating that had we chosen a good performing model consistently throughout the years, we would have had more accurate forecasts compared to those obtained from the forecast models historically chosen.

Among biological models, the Power model coupled with new covariates (blue icons in Figure 5) generally outperformed other models with higher correlations, similar standard deviations to those in observations, and smaller RMSEs (Figure 5a, b). The performance of the Power model combined with the previously adopted covariates (green icons in Figure 5) was similar to the best models, yet they were rarely used in previous forecasts (Grant et al. 2011; DFO 2018; Hawkshaw et al. 2020a, b). The Ricker model coupled with new covariates (red icons in Figure 5) also seemed to be better compared to those with old covariates (pink icons in Figure 5) with the latter having been frequently used in the previous forecasts. For most stocks, the best model forecasts based on the three statistics (correlation, standard deviation, and RMSE)

shown in the Taylor diagrams were in good agreement with the observation, reaching correlations between 0.75-0.95. However, for a few stocks such as Gates and Raft, none of the existing models produced an accurate forecast, probably due to data quality issues as aforementioned. Naïve models (purple icons in Figure 5) tended to have larger standard deviations and performed worse than biological models in general (except LLY for some stocks).

Age-specific Taylor diagrams are presented for all 37 forecast models relative to the historical forecast (black solid square in Figures S5-6) and the observation (black solid circle in Figures S5-6) for age 4 and age 5 returns, respectively (Supplementary Figure S5 and S6). The age-4 Taylor diagrams showed similar results as the diagrams produced for total abundance (Figure 5a,b), since age 4 was the dominant age class for most stocks. Power models coupled with new covariates consistently outperformed other models for many stocks. The historical forecast of age 4 showed similarly poor performance for most stocks (Supplementary Figure S5). For age 5, sibling models (yellow squares in Figures S6) performed better for many stocks than the naïve and biological models (Supplementary Figure S6).

3.2 Model selection

In 2021, when sockeye forecast was conducted using previously selected forecast models, the forecast error (defined as $\frac{Observation - Forecast}{Forecast} \times 100\%$) was 92%, being the highest among all the sockeye runs in Northeast Pacific (Supplementary Table S2). In 2022, Taylor diagrams started to be implemented to identify the best performing model for many Fraser stocks, as a result, the forecast error has reduced to 30%, much lower than some of the other sockeye runs (Supplementary Table S2). This reduced forecast error also represents a significant improvement compared to the past few years (Hawkshaw et al. 2020a, b). In 2023, when Taylor diagrams were

further implemented, the preliminary results showed that the forecast error was -3% only (pre-season forecast is 1.564 million, in-season adopted run size is 1.606 million; PSC, unpublished document Sep 19, 2023). Although the evaluation period is not long enough to make a conclusion, it would be interesting to continue monitoring the performance of this implementation in the near future.

To understand how the models that were ranked first based on either overall rank or one of the three statistics (correlation, standard deviation, and RMSE) performed compared to other models, we examined the largest three stocks (Chilko, Late Shuswap, and Quesnel) as examples (Figure 6). For Late Shuswap, where the 2022 forecast was done by the historically selected model (RickerEi.SST), the forecast resulted in a deviation of -57%, equivalent to a difference of 1.9 million sockeye salmon relative to the observed returns. In contrast, the model ranked first based on overall rank (i.e., RickerGOA.SST) resulted in the most accurate forecast with the smallest uncertainty (Figure 6). The RickerPi.SST model also performed relatively well, being best according to either correlation, standard deviation, or RMSE. The RS1 model, ranked third based on standard deviation following RickerPi.SST and RickerGOA.SST models, also performed better than the RickerEi.SST model, which may indicate that standard deviation is an important factor to consider in the model selection process. The importance of standard deviation can also be seen with Chilko and Quesnel, where the top-rank models based on standard deviation (i.e., R1C and RickerSockeye, respectively) also performed the best with the highest accuracy and lowest uncertainty (Figure 6). By contrast, RickerCyc, although highest ranked according to correlation and RMSE, performed poorly for Chilko. Similarly, the Larkin model performed poorly for Quesnel despite being ranked highest according to correlation.

4. Discussion

Forecasting Fraser sockeye salmon returns in the upcoming year is a challenging task given the complex life history and the dynamic freshwater and marine habitats in which they reside. In recent years, Fraser sockeye forecast accuracy has been declining, frequently going below 10 or above 90 posterior quantiles, i.e., at the extremes ends of the 80% prediction interval (DFO 2021). For example, in 2021, the median forecast for the total Fraser River sockeye return was 1.33 million fish while the observed return was 2.549 million, 92% above the forecast (Supplementary Table S2). Despite these kinds of extreme differences, the official forecast is widely used by various groups for important planning decisions, including escapement survey planning, fishery opening and licensing, allocation among commercial, recreational and indigenous fisheries, hatchery enhancement activities, research and development, and international salmon treaty negotiation (Haeseker et al. 2008; Michielsens and Cave 2019; Hawkshaw et al. 2020a, b; DFO in press). Developing a framework to improve and streamline the existing forecast process is highly desired for a broad range of stakeholders. In this paper, we developed a framework that allowed fisheries managers and other stakeholders to evaluate the performances of pre-season forecast models on an annual basis by using Taylor diagrams, and to identify external drivers of importance for forecasting Fraser sockeye returns. Specifically, for the first time we incorporated new environmental and ecological covariates (i.e., GOA.SST, Pink, Chum, Sockeye, Salmon.Total) into the Ricker and Power forecast models, which has resulted in consistently robust forecasts across all 18 Fraser sockeye stocks, regardless of data quality and survey method changes temporally over multiple decades and spatially over the entire Fraser watershed. In particular, the Ricker models with ecological covariates of Pink and Salmon.Total tended to have small and negative biases (only slightly under-forecast) across different stocks, which is very appealing from the perspective of conservative management.

4.1 Moving towards ecosystem-based fisheries management

Sockeye salmon have a complex life history, going through vastly different freshwater and marine environments at different life stages, resulting in very unpredictable impacts on stock productivity (e.g., Ricker 1997; Akenhead et al. 2016b; Huang et al. 2021). It can be helpful to have reliable modelling of their complex life history and particularly of the impacts of climate change, ecosystem phase transitions, and competition for food and other resources when conducting forecast of sockeye returns (e.g., Patterson et al. 2016; Satterthwaite et al. 2020; Connors et al. 2020; Wainwright 2021; Ohlberger et al. 2022; Kaeriyama 2022). Moving pre-season sockeye forecast and management towards ecosystem-based approaches that account for ecosystem processes, including fishing and/or climate variability in conjunction with species interactions (Sissenwine and Murawski 2004; Link 2011) could improve the performance of forecast models and by extension, the management systems they serve. Multiple studies have found that sockeye salmon in BC have been negatively affected by the significant increase of Pink salmon abundance in the north Pacific Ocean in recent decades through competition for limited resources on both broad and localized scales (Ruggerone and Connors 2015; Connors et al. 2020; Ruggerone et al. 2021; Litz et al. 2021). Even though existing Fraser sockeye stock assessment and fisheries management are still heavily single-species based, our study is a step forward towards ecosystem-based fisheries management by showing that incorporating the covariates of pink, chum, other sockeye abundance, and their combined abundance, implicitly accounting for inter-specific and intra-specific competition, will improve forecast performance. Our forecast framework revealed that models combined with the salmon covariates were generally highly ranked retrospectively from 2009-2020, highlighting that Fraser sockeye dynamics were closely related to the abundances of other Pacific salmon stocks in the north

Pacific Ocean. In addition, the current area was based on previous tagging studies on distribution of sockeye salmon (Myers et al. 1996; Ishii et al. 2005; Supplementary Figure S1), however, recent high-seas surveys (2019-2020, 2022) in the Gulf of Alaska suggested that Fraser sockeye distribution could be even wider in the North Pacific Ocean, with one sockeye caught in the eastern hemisphere (176.2°E). These surveys also found that Fraser sockeye shared marine habitat with sockeye salmon from other river systems in cooler waters as well as Pink and Chum salmon, although the spatial overlap with the latter two species were less intensive (<https://www.youtube.com/watch?v=thCmnoPEw6s>). The extended period of residence in the GOA by Pacific salmon, their high degree of spatial overlap in this region, and evident co-variability among salmon species mean that the GOA is an area of significance for salmon abundance.

The GOA is influenced by decadal thermal variability and has experienced three marine heatwave events since 2013, resulting in unprecedented reductions in fishery recruitment and shifts in the biological community of this region (Blaisdell et al. 2021). The inclusion of GOA.SST into the pre-season sockeye forecast was another step of moving forward towards ecosystem-based fisheries management. The RickerGOA.SST model (Ricker coupled with GOA.SST) performed extremely well for the Late Shuswap stock (one of the three major Fraser sockeye stocks) compared to all other models with the exception of RickerPi.SST (Figure 3). Had the RickerGOA.SST model been adopted in 2022 for the Late Shuswap stock, the median forecast for this stock would have been reduced from 3.42 million to 1.48 million (equal to the observation), and the error for Fraser River sockeye would have been reduced from -30 to -10% (Supplementary Table S2). It is also worth noting that the current GOA.SST map only represents part of the northeastern Pacific Ocean (Supplementary Figure S1), which can be expanded when

we understand the Fraser Sockeye distribution in GOA better. Nevertheless, our paper sheds some light on how forecasts can be improved by incorporating biotic and abiotic metrics from ecosystem perspectives. Overall, we showed that the forecast models coupled with newly added covariates not only reduced the differences between forecast and observation but also reduced uncertainties, which would be a great help to fisheries managers who tend to face high pressure from political parties when forecast is inaccurate or has high uncertainty.

4.2 Taylor diagram and future implications

The Taylor diagram allowed us to visually compare time series of forecasts from 37 forecast models along with the historical forecasts against the observations in one figure both in terms of forecast accuracy and uncertainty. This approach can help simplify the model selection process by illustrating performance quantitatively among all forecast models. Based on the Taylor diagrams for all 18 Fraser sockeye stocks, we concluded that historically selected forecast models were not ideal compared to many other forecast models. The framework developed through this study can be adopted for future model selection and forecast processes (DFO, in press). In particular, for the first time, standard deviation was taken into consideration when comparing between model prediction and observation. This information could be useful to resource managers when making management decisions. For example, a model with a normalized standard deviation larger than that of observation could result in larger uncertainties during dominant years, but smaller for non-dominant years. Such information helps managers to become aware of potential forecast pitfalls when they encounter unusually large or small standard deviations compared with those of observations.

Many Fraser sockeye stocks exhibit strong cyclic patterns with a dominant year every four years (Supplementary Figure S2). Forecasting for the dominant years was more challenging and

was usually accomplished with lower accuracy compared to other years (DFO, 2021). Our retrospective analysis only covered 12 years (3 cycles); however, as the time series extends in future, it will be interesting to separate dominant years and plot cycle-specific Taylor diagrams to identify models specifically tailored to dominant years. It will also be interesting to examine how the relative positions of these forecast models evolve over time. Theoretically, informative models will move closer to the observation as more years of data become available, assuming that the observation is a reliable reference. Models with false alarms (bad models but appear to have good performance temporarily) will either demonstrate erratic behavior or deviate further from the observation. Monitoring the moving directions of each model for multiple years may provide insights into their relative forecasting skill or potential for improvement, based on relative differences compared to observations.

Currently, our framework allows us to select a model by visually looking at the models' relative positions in Taylor diagrams, which could be subjective. For example, a best RMSE model may not have best standard deviation or best correlation or vice versa. To be more objective, a single distance metric, similar to Mohn's rho value (Mohn, 1999), can be developed with proper weights among three metrics and integrated to automate the model selection process to come up with best machine-suggested models.

4.3 Forecast models for non-stationary relationships

It has been commonly recognized that correlations established based on past observations can be easily broken down due to the non-stationary nature of environmental changes (e.g., Peterman & Dorner 2012; Malick 2020; Litzow et al. 2020). Such non-stationarity calls for continuous evaluation of previously established forecast models. While annually evaluating the performance of forecast models with environmental variables may facilitate more accurate and precise

forecasts, actually implementing the best performing models every year can be disruptive and may risk chasing noise (Winship et al., 2015). Therefore, caution should be taken when interpreting results based on single-year performance while emphasis may be put on performance stability across multiple years in the case of model overfitting or data quality issues. By using a moving data window or down-weighting more distant data, Winship et al. (2015) found that the performances of various models changed over time and that there were a lot of inter-annual variation in the environmental covariates identified for inclusion in a forecast model for a Chinook salmon stock. There is a need to conduct more research to identify models that perform more consistently over time and more robustly to environmental changes, particularly in the light of the connections between life history characteristics of different taxa and forecast model performances (Ward et al. 2015).

In this study, we also found that sibling models outperformed others for age-5 for most of the Fraser stocks, suggesting that sometimes simple models with no knowledge about exact mechanisms can provide satisfactory forecast. This is consistent with other studies (Scheuerell and Williams, 2005). However, there may be even a greater need to develop management systems that are robust to forecast uncertainties (Wainwright 2021), as improving forecast skills solely not necessarily implies improving fisheries management (Rupp et al. 2012).

For the Chilko and Quesnel sockeye stocks, the RickerCyc and Larkin models performed poorly although each model was ranked first based on correlation. Based on this result, we infer that the correlations established for the period 2009 – 2020 may be deteriorating as environmental conditions have been constantly changing. On the other hand, we found that the top-rank models based on standard deviations performed the best with the highest accuracy and lowest uncertainty for the two major stocks Chilko and Quesnel (Figure 6). With the Late

Shuswap sockeye stock, all three models with the first three ranks based on standard deviation (i.e., RickerPi.SST, RickerGOA.SST, and RIC) performed better than the historical forecast (Figure 6). These findings may show that standard deviations established based on past observations are not as easily broken down as correlations and thus could be considered as an important factor in the model selection process.

In terms of stock-recruitment relationships, parametric models, such as Ricker, Power, and Larkin, have been primarily employed in the past (Ricker 1997; Cass et al. 2006; Grant et al. 2010, 2011; Akenhead, et al., 2016a; DFO 2018; Hawkshaw et al. 2020a, b). However, these parametric models have limited forms to encompass the inherent uncertainties associated with large annual variability in multiple environmental factors that affect salmon dynamics throughout their complex life history stages (e.g., Healey, 2011; Martins et al., 2012; Padilla et al., 2015). They are thus inadequate to incorporate anthropogenic changes (Akenhead et al., 2016b), which can result in unexplained residual patterns in the stock-recruit relationships, resulting in less accurate pre-season forecasts. In addition, these parametric models are also difficult to deal with non-stationary sockeye dynamics (Peterman & Dorner 2012; Malick 2020), which may contribute the Chilko sockeye return to have been persistently over-estimated in the last decade (DFO 2021). Climate-induced non-stationarity in relationships between marine environments and fish communities has been increasingly recognized (e.g., Litzow et al. 2018, 2020; Ohlberger et al. 2022) and it has been advocated that new approaches capable of dealing with non-stationarity in sockeye dynamics should be explored (Peterman & Dorner 2012; Malick 2020).

Alternative approaches such as generalized additive models (Wood 2017), boosted regression trees (Elith et al. 2008), random forest (Breiman 2001), empirical dynamic models (Ye et al.

2015), and artificial neural network (Ripley, 1996), use algorithms to learn the relationships between responses and predictors and are not confined to certain functional forms. These more flexible models can be developed to represent non-stationary stock-recruitment relationships in the dynamics of sockeye stocks. Recently discovered similar temporal patterns of Fraser sockeye survival among all stocks (Akenhead et al. 2016a; Connors et al. 2020; DFO 2021; Ruggerone et al. 2021; Rosengard et al. 2021) also suggest that methods based on shared trends and autocorrelation can be explored. Such methods have proven effective in other situations (e.g. Winship et al. 2015; Johnson et al. 2016; DeFilippo et al. 2021), while environmental relationships can be broken down, resulting in higher forecast errors particularly for highly variable marine fish stocks (e.g. Ward et al. 2014). In addition, an ensemble modeling approach has been developed in recent years (Ovando et al. 2022) which can be useful for taking advantages of multiple models to improve forecast skills.

With all alternative approaches being incorporated, Taylor diagrams can then be applied to study how these models perform compared to the traditional naïve models and parametric models. In the future, we can also incorporate multiple covariates into these alternative models to evaluate their relative importance in forecasting sockeye return. In summary, we advocate for continuously evaluating all potential forecast models in the face of increasing environmental change and the development of new models able to deal with non-stationary relationships between environment and sockeye dynamics. This process of continuous evaluation should lead to the identification of good performing forecast models and to improved understanding of the impacts of environmental and ecological factors on the performance of Fraser sockeye forecast.

Acknowledgements

The authors would like to thank all the staff from Fisheries and Oceans Canada and Pacific Salmon Commission who have collected the Fraser sockeye data and contributed in previous forecast model development and evaluation processes. Special thanks are given to recent contributors, including Mickey Agha, Bruce Baxter, Taren Bell, Tom Bird, Brendan Connors, Mary Beth Fagan, Emily Breiteneder, Kelsey Campbell, Peter Chandler, Tracy Cone, Catharina De Monye, Scott Decker, Travis Desy, Kaitlyn Dionne, Maxine Forest, Nicole Frederickson, Jin Gao, Sue Grant, Merran Hague, Mike Hawkshaw, Stacey Hobson, Roy Hourston, Jim Irvine, Les Jantz, Dennis Klassen, Stu LaPage, Steve Latham, Brian Leaf, Doug Lofthouse, Nancy Louie, Jennifer Lynne, Bronwyn MacDonald, Michael Malick, Nate Mantua, Fiona Martens, Catherine McClean, Catherine Michielsens, Matt Mortimer, Judy Munsell, Chuck Parken, Jason Parsley, Matthew Parslow, David Patterson, Gottfried Pestal, Lucas Pon, Nicole Porteous, Tony Rathbone, Karen Richards, Rice Robert, Loraine Roper, Gordon Rose, Sarah Rosengard, Gregory Ruggerone, Jamie Scroggie, Amy Seiders, Michael Staley, Angus Straight, Kayla Suhan, Madeline Thomson, Matthew Townsend, Strahan Tucker, Maxime Veilleux, Tanya Vivian, Paul Welch, Timber Whitehouse, Serena Wong, and Zhipeng Wu (in alphabetic order by last name). We gratefully acknowledge invaluable contributions of two anonymous reviewers and CJFAS associate editor, who provided constructive feedback on our initial manuscript. Your insights greatly enhanced the quality of our work. This project was initially supported by Sockeye and Pink Analytical Program, Fraser and Interior Area, DFO when the lead author was leading the 2022 Fraser sockeye pre-season forecast process. The remaining work was done while the lead author was supported by the Salmon Data Unit, Fishery and Assessment Data Section, and Biological Resource Management Section, Fraser and Interior Area, DFO. The lead author would like to thank Shelee Hamilton and Brittany Jenewein for their support.

536 **Copyright**

537 © 2023 The Crown. This work is licensed under a Creative Commons Attribution 4.0
538 International License (CC BY 4.0), which permits unrestricted use, distribution, and
539 reproduction in any medium, provided the original author(s) and source are credited.

540 **Data availability**

541 All data and code are available in a Github repository. https://github.com/yi-xu/Sockeye_paper
542 Access can be granted when contacting corresponding author.

543 **Competing interests**

544 The authors declare there are no competing interests.

545 **Supplementary material**

546 Supplementary data are available with the article at [https:](https://doi.org/xxxx/cjfas-xxxxx)
547 [//doi.org/xxxx/cjfas-xxxxx](https://doi.org/xxxx/cjfas-xxxxx).

548

References

- Ackerman, P.A., Barnetson, S., Lofthouse, D., McClean, C., Stobbart, A., and Withler, R.E. 2014. Back from the Brink: The Cultus Lake Sockeye Salmon Enhancement Program from 2000 - 2014. Can. Manuscr. Rep. Fish. Aquat. Sci. 3032: vii + 63p.
- Akenhead, S.A., Irvine, J.R., Hyatt, K.D., Johnson, S.C. and Grant, S.C.H. 2016a. Stock-recruit analyses of Fraser River sockeye salmon. N. Pac. Anadr. Fish Comm. Bull 6: 363-390.
- Akenhead, S.A., Irvine, J.R., Hyatt, K.D., Johnson, S.C., Michielsens, C.G.J. and Grant, S.C.H. 2016b. Habitat manipulations confound the interpretation of sockeye salmon recruitment patterns at Chilko Lake, British Columbia. N. Pac. Anadr. Fish Comm. Bull. 6: 391-414.
- Blaisdell, J., Thalmann, H.L., Klajbor, W., Zhang, Y., Miller, J.A., Laurel, B.J. and Kavanaugh, M.T. 2021. A Dynamic Stress-Scape Framework to Evaluate Potential Effects of Multiple Environmental Stressors on Gulf of Alaska Juvenile Pacific Cod. Front. Mar. Sci. 8:656088. doi: 10.3389/fmars.2021.656088
- Breiman, L. 2001. Random Forests. Machine Learning 45, 5–32.
- Burgner, R. L. 1991. Life history of sockeye salmon (*Oncorhynchus nerka*). In Pacific Salmon Life Histories. Groot, C. and Margolis, L. (eds.). University of British Columbia Press, Vancouver, Canada. 1: pp 1-118.
- Cass, A., Folkes, M., Parken, C., and Wood, C. 2006. Pre-season run size forecasts for Fraser River sockeye in 2006. Can. Sci. Advis. Sec. Res. Doc. 2006/060: pp. iii + 72.
- Connors, B., Malick, M.J., Ruggerone, G.T., Rand, P., Adkison, M., Irvine, J.R., Campbell, R., and Gorman, K., 2020. Climate and competition influence Sockeye salmon population dynamics across the Northeast Pacific Ocean. Can. J. Fish. Aquat. Sci. 77(6): 943–949.

571 COSEWIC, 2017. COSEWIC assessment and status report on the Sockeye Salmon
572 *Oncorhynchus nerka*, 24 Designatable Units in the Fraser River Drainage Basin, in Canada.
573 Committee on the Status of Endangered Wildlife in Canada. Ottawa. xli + 179 pp.

574 DeFilippo, L. B., Buehrens, T. W., Scheuerell, M., Kendall, N. W., & Schindler, D. E., 2021.
575 Improving short-term recruitment forecasts for coho salmon using a spatiotemporal
576 integrated population model. Fisheries Research, 242, 106014.

577 DFO, 2015. Pre-Season Run Size Forecasts for Fraser River Sockeye (*Oncorhynchus nerka*) and
578 Pink (*O. gorbuscha*) Salmon in 2015. Can. Sci. Adv. Sec. Sci. Response 2015/014: pp. 55.

579 DFO, 2018. Pre-season run size forecasts for Fraser River Sockeye (*Oncorhynchus nerka*)
580 salmon in 2018. Can. Sci. Advis. Sec. Sci. Resp. 2018/034: pp. 70.

581 DFO, 2021. Pre-Season run size forecasts for Fraser River Sockeye (*Oncorhynchus nerka*) and
582 Pink (*O. gorbuscha*) Salmon in 2021. Can. Sci. Advis. Sec. Sci. Resp. 2021/038: pp. 105.

583 DFO, in press. Pre-season run size forecasts for Fraser River Sockeye (*Oncorhynchus nerka*) in
584 2022. Can. Tech. Rep. Fish. Aquat. Sci. 2023/3496: pp. 349.

585 Elith, J., Leathwick, J.R. and Hastie, T., 2008. A working guide to boosted regression trees. J.
586 Anim. Ecol., 77: 802-813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>

587 Grant, S.C.H., MacDonald, B.L., Cone, T.E., Holt, C.A., Cass, A., Porszt, E.J., Hume, J.M.B.,
588 and Pon, L.B., 2011. Evaluation of uncertainty in Fraser sockeye (*Oncorhynchus nerka*)
589 wild salmon policy status using abundance and trends in abundance metrics. DFO. Can. Sci.
590 Advis. Sec. Res. Doc. 2011/087. pp.183.

591 Grant, S.C.H., Michielsens, C.G.J., Porszt, E.J., and Cass, A.J. 2010. Pre-season run size
592 forecasts for Fraser River sockeye salmon (*Oncorhynchus nerka*) in 2010. Can. Sci. Advis.

593 Sec. Res. Doc. 2010/042: pp.125.

594 Haeseker, S.L., Peterman, R.M., Su, Z., and Wood, C.C., 2008. Retrospective evaluation of
595 preseason forecasting models for sockeye and chum salmon. N. Am. J. Fish. Manag. 28(1):
596 12–29.

597 Hawkshaw, M., Xu, Y., and Davis, B. 2020a. Pre-season Run Size Forecasts for Fraser River
598 Sockeye (*Oncorhynchus nerka*) and Pink (*Oncorhynchus gorbuscha*) Salmon in 2019. Can.
599 Tech. Rep. Fish. Aquat. Sci. 3391: vi + 52 p.

600 Hawkshaw, M., Xu, Y., and Davis, B. 2020b. Pre-season Run Size Forecasts for Fraser River
601 Sockeye (*Oncorhynchus nerka*) Salmon in 2020. Can. Tech. Rep. Fish. Aquat. Sci. 3392: iv
602 + 56 p.

603 Healey, M. 2011. The cumulative impacts of climate change on Fraser River sockeye salmon
604 (*Oncorhynchus nerka*) and implications for management. Can. J. Fish. Aquat. Sci. 68: 718-
605 737.

606 Huang, A-M., Pestal, G., and Guthrie, I. 2021. Recovery Potential Assessment for Fraser River
607 Sockeye Salmon (*Oncorhynchus nerka*) – Nine Designatable Units: Probability of
608 Achieving Recovery Targets - Elements 12, 13, 15, 19-22. DFO Can. Sci. Advis. Sec. Res.
609 Doc. 2021/043. x + 96.

610 Hyndman, Rob J. and Koehler., 2006. Another look at measures of forecast accuracy. Int. J.
611 Forecast.. 22 (4): 679–688. doi:10.1016/j.ijforecast.2006.03.001.

612 Inna S., Lehodey, P., Sibert, J., and Hampton, J., 2019. Integrating tagging and fisheries data
613 into a spatial population dynamics model to improve its predictive skills. Can. J. Fish.
614 Aquat. Sci. 77(3): 576-593. <https://doi.org/10.1139/cjfas-2018-0470>

615 Ishii, M., Shouji, A., Sugimoto, S., and Matsumoto, T., 2005. Objective Analyses of Sea-Surface
616 Temperature and Marine Meteorological Variables for the 20th Century using ICOADS and
617 the Kobe Collection. *Int. J. Climatol.*, 25: 865-879.

618 Johnson, K. F., Councill, E., Thorson, J. T., Brooks, E., Methot, R. D., & Punt, A. E., 2016. Can
619 autocorrelated recruitment be estimated using integrated assessment models and how does it
620 affect population forecasts? *Fisheries Research*, 183: 222-232.

621 Kaeriyama, M., 2022. Warming climate impacts on production dynamics of southern
622 populations of Pacific salmon in the North Pacific Ocean. *Fish. Oceanogr.*
623 doi:10.1111/fog.12598

624 Kell, L. T., Kimoto, A., and Kitakado T., 2016. Evaluation of the prediction skill of stock
625 assessment using hindcasting. *Fish. Res.*, 183: 119–127.

626 Lamine, B.E., Schickele, A., Goberville, E. Beaugrand, G., Allemand, D., and Raybaud, V.
627 2022. Expected contraction in the distribution ranges of demersal fish of high economic
628 value in the Mediterranean and European Seas. *Sci Rep* 12, 10150.
629 <https://doi.org/10.1038/s41598-022-14151-8>

630 Lee, Y. J., Matrai, P.A. Friedrichs M.A.M., Saba, V.S., Antoine, D., Ardyna, M., Asanuma, I.,
631 Babin, M., Belanger, S., Benoit-Gagne, M., Devred, E., Fernandez-Mendez, M., Gentill, B.,
632 Hirawake, T., Kang, S.-H., Kameda, T., Katlein, C., Lee, S.H., Lee, Z., Melin, F., Scardi,
633 M., Smyth, T.J., Tang, S., Turpie, K.R., Waters, K.J., and Westberry, T.K., 2015. An
634 assessment of phytoplankton primary productivity in the Arctic Ocean from satellite ocean
635 color/in situ chlorophyll-a based models, *J. Geophys. Res. Oceans*, 120: 6508– 6541,
636 doi:10.1002/2015JC011018.

637 Lingard, S., Melville, C., and McCubbing, D. 2013. Gates Creek Salmon Project Adult
 638 Escapement Assessment Project #12.SON.01 (Part2). InStream Fisheries Research Inc.
 639 https://a100.gov.bc.ca/pub/acat/documents/r40299/12.SON.01Gates_final_1386167072858_6166228483.pdf (last assessed Dec 7, 2023)
 640
 641 Link, J., 2011. Ecosystem-based Fisheries Management: Confronting Tradeoffs. Cambridge
 642 University Press, Cambridge.
 643 Litz, M.N.C., Agha, M., Dufault, A.M., Claiborne, A.M., Losee, J.P., and Anderson, A.J. 2021.
 644 Competition with odd-year pink salmon in the ocean affects natural populations of chum
 645 salmon from Washington. Mar. Ecol. Prog. Ser. 663:179-195.
 646 Litzow, M. A., Ciannelli, L., Puerta, P., Wettstein, J. J., Rykaczewski, R. R., and Opiekun, M.,
 647 2018. Non-stationary climate–salmon relationships in the Gulf of Alaska. Proc. Royal Soc.
 648 B., 285 (1890), 20181855. <https://doi.org/10.1098/rspb.2018.1855>
 649 Litzow, M. A., Hunsicker, M. E., Bond, N. A., Burke, B. J., Cunningham, C. J., Gosselin, J. L.,
 650 Norton, E. L., Ward, E. J., & Zador, S. G., 2020. The changing physical and ecological
 651 meanings of North Pacific Ocean climate indices. PNAS, 117(14): 7665–7671.
 652 <https://doi.org/10.1073/pnas.1921266117>
 653 MacDonald, B.L., and Grant, S.C.H. 2012. Pre-season run size forecasts for Fraser River
 654 sockeye salmon (*Oncorhynchus nerka*) in 2012. Can. Sci. Advis. Sec. Res. Doc.
 655 2012/011(April): pp. v + 64 p.
 656 Malick, M.J. 2020. Time-varying relationships between ocean conditions and sockeye salmon
 657 productivity. Fish Oceanogr. 29: 265– 275.
 658 Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and Francis, R.C. 1997. A Pacific

659 Interdecadal Climate Oscillation with Impacts on Salmon, Bull. Am. Meteorol. Soc.,
 660 78(6):1069–1079.

661 Martins, E.G., Hinch, S.G., Cooke, S.J., Patterson, D.A., 2012. Climate effects on growth,
 662 phenology, and survival of sockeye salmon (*Oncorhynchus nerka*): a synthesis of the
 663 current state of knowledge and future research directions. Rev.Fish Biol. Fish. 22: 887–914.

664 McKinnell, 2008 Fraser River sockeye salmon and climate; a re-analysis that avoids an
 665 undesirable property of Ricker’s curve Prog. Oceanogr., 77: pp. 146-154

666 Michielsens, C.G.J., and Cave, J.D., 2019. In-season assessment and management of salmon
 667 stocks using a Bayesian time-density model. Can. J. Fish. Aquat. Sci. 76: 1073–1085.
 668 dx.doi.org/10.1139/cjfas-2018-0213

669 Mohn, 1999. The retrospective problem in sequential population analysis: an investigation using
 670 cod fishery and simulated data. ICES J. Mar. Sci., 56: pp. 473-488

671 Myers, K.W., Aydin, K.Y., Walker, R.V., Fowler, S., and Dahlberg, M.L. 1996. Known Ocean
 672 ranges of stocks of Pacific salmon and steelhead as shown by tagging experiments, 1956-
 673 1995. N. Pac. Anadr. Fish Comm. Doc. 192(4).

674 Ogden, A.D., Irvine, J.R., English, K.K., Grant, S., Hyatt, K.D., Godbout, L., and Holt, C.A.
 675 2015. Productivity (recruits-per-spawner) data for sockeye, pink and chum salmon from
 676 British Columbia. Can. Tech. Rep. Fish. Aquat. Sci. 3130: vi + 57 p.

677 Ohlberger, J., Ward, E. J., Brenner, R. E., Hunsicker, M. E., Haught, S. B., Finnoff, D., Litzow,
 678 M. A., Schwoere, T., Ruggerone, G. T., and Hauri, C., 2022. Non-stationary and interactive
 679 effects of climate and competition on pink salmon productivity. Glob. Chang. Biol., 28(6):
 680 2026–2040. <https://doi.org/10.1111/gcb.16049>

681 Ovando D., Cunningham, C., Kuriyama, P., Boatright, C., and Hilborn, R., 2022. Improving
682 forecasts of sockeye salmon (*Oncorhynchus nerka*) with parametric and nonparametric
683 models. Can. J. Fish. Aquat. Sci. 79(8): 1198-1210. <https://doi.org/10.1139/cjfas-2021-0287>

684 Padilla, A., Rasouli, K., and Déry, S.J., 2015. Impacts of variability and trends in runoff and
685 water temperature on salmon migration in the Fraser River Basin, Canada, Hydrol. Sci. J.,
686 60:3, 523-533, doi: 10.1080/02626667.2014.892602

687 Patterson DA, Cooke SJ, Hinch SG, Robinson KA, Young N, Farrell AP, Miller KM (2016) A
688 perspective on physiological studies supporting the provision of scientific advice for the
689 management of Fraser River sockeye salmon (*Oncorhynchus nerka*). Conserv. Physiol.
690 4(1): cow026.

691 Peterman, R.M. and Dorner, B. 2012. A widespread decrease in productivity of sockeye salmon
692 (*Oncorhynchus nerka*) populations in western North America. Can. J. Fish. Aquat. Sci.
693 69(8): 1255–1260.

694 PSC, 2022. Treaty between the Government of Canada and the Government of the United States
695 of America concerning Pacific Salmon. 2022. Prepared by Pacific Salmon Commission.
696 <http://www.psc.org/publications/pacific-salmon-treaty/>

697 R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for
698 Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

699 Ricker, W.E. 1997. Cycles of abundance among Fraser River sockeye salmon (*Oncorhynchus*
700 *nerka*). Can. J. Fish. Aquat. Sci. 54: 950–968. doi:10.1139/cjfas-54-4-950.

701 Ripley, B. D., 1996. Pattern Recognition and Neural Networks. Cambridge.

702 Rosengard, S.Z., Freshwater, C., McKinnell, S., Xu, Y., and Tortell, P.D., 2021. Covariability of

703 Fraser River sockeye salmon productivity and phytoplankton biomass in the Gulf of Alaska.
 704 Fish. Oceanogr. 30: 666-678.

705 Ruggerone, G.T., and Connors, B.M., 2015. Productivity and life history of sockeye salmon in
 706 relation to competition with pink and sockeye salmon in the North Pacific Ocean. Can. J.
 707 Fish. Aquat. Sci. 72(6): 818–833. doi:10.1139/cjfas-2014-0134.

708 Ruggerone, G.T. and Irvine, J.R., 2018. Numbers and biomass of natural- and hatchery-origin
 709 pink, chum, and sockeye salmon in the North Pacific Ocean, 1925-2015. Mar. Coast. Fish.
 710 Dyn. Manage. Ecosyst. Sci. 10: 152-168.

711 Ruggerone, G.T., Irvine, J.R., and Connors, B., 2021. Did recent marine heatwaves and record
 712 high Pink salmon abundance lead to a tipping point that caused record declines in North
 713 Pacific salmon abundance and harvest in 2020? NPAFC Tech. Rep. 17: 78-82.

714 Rupp, D.E., Wainwright, T.C., and Lawson, P.W., 2012. Effect of forecast skill on management
 715 of the Oregon coast coho salmon (*Oncorhynchus kisutch*) fishery. Canadian Journal of
 716 Fisheries and Aquatic Sciences. 69(6): 1016-1032.

717 Salihoglu B., Arkin, S.S., Akoglu, E., and Fach, B.A., 2017. Evolution of Future Black Sea Fish
 718 Stocks under Changing Environmental and Climatic Conditions. Front. Mar. Sci. 4:339.
 719 doi:10.3389/fmars.2017.00339

720 Satterthwaite, W.H., Andrews, K.S., Burke, B.J., Gosselin, J.L., Greene, C.M., Harvey, C.J.,
 721 Munsch, S.H., O’Farrell, M.R., Samhour, J.F., and Sobocinski, K.L., 2020. Ecological
 722 thresholds in forecast performance for key United States West Coast Chinook salmon
 723 stocks. – ICES J. Mar. Sci., 77: 1503–1515. doi:10.1093/icesjms/fsz189

724 Scheuerell, M. D., and Williams, J. G., 2005. Forecasting climate-induced changes in the

725 survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*).
 726 Fisheries Oceanography, 14(6): 448-457.

727 Sissenwine, M., and Murawski, S., 2004. Moving beyond ‘intelligent thinking’: advancing an
 728 ecosystem approach to fisheries. In: Perspectives on Ecosystem-based Approaches to the
 729 Management of Marine Resources. Mar. Ecol. Prog. Ser., vol. 274: pp. 269–303.

730 Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. J.
 731 Geophys. Res. 106: 7183–7192. doi:10.1029/2000JD900719.

732 Wainwright, T.C. 2021. Ephemeral relationships in salmon forecasting: A cautionary tale. Prog.
 733 Oceanogr., 193, 102522

734 Ward, E.J., Holmes, E.E., Thorson, J.T., and Collen, B. 2014. Complexity is costly: a meta-
 735 analysis of parametric and non-parametric methods for short-term population forecasting.
 736 Oikos, 123: 652–661. doi:10.1111/j.1600-0706.2014. 00916.x.

737 Willmott, C. J. and Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the
 738 root mean square error (RMSE) in assessing average model performance. Clim. Res. 30:
 739 79–82. doi:10.3354/cr030079.

740 Winship, A.J., O’Farrell, M.R., Satterthwaite, W.H., Wells, B.K., Mohr, M.S., 2015. Expected
 741 future performance of salmon abundance forecast models with varying complexity. Can. J.
 742 Fish. Aquat. Sci. 72: 557–569. dx.doi.org/10.1139/cjfas-2014-0247

743 Wood S., 2017. Generalized Additive Models: An Introduction with R, 2 edition. Chapman and
 744 Hall/CRC.

745 Ye H., Beamish, R.J., Glaser, S.M., Grant, S.C., Hsieh, C.H., Richards, L.J., Schnute, J.T., and
 746 Sugihara, G., 2015. Equation-free mechanistic ecosystem forecasting using empirical

747 dynamic modeling. *Proc. Natl. Acad. Sci.* 112(13): E1569–E1576.

Figure captions

Figure 1. Locations of 18 major Fraser Sockeye salmon stocks where spawning data were collected. Color indicates different run type: red for Early Stuart, green for Early Summer run (including Bowron, Fennel (Upper Barriere), Scotch, Nadina, Pitt, Seymour, and Gates), blue for Summer run (including Chilko, Quesnel, Late Stuart, Stellako, and Raft), and purple for Late run (including Late Shuswap, Cultus, Portage, Weaver, and Birkenhead).

Figure 2. Interannual variations of ten environmental and ecological variables from 1950-2020, including sea surface temperature (SST, unit in °C) at Entrance Island and Pine Island, mean and peak Fraser River discharge at Hope (unit m³/s), Pacific Decadal Oscillation (PDO), SST in the Gulf of Alaska, North Pacific Sockeye, Chum, Pink salmon abundance (catch plus escapement, number in million), and combined total of the three salmon species.

Figure 3. Relative rank among all 37 forecast models for all 18 stocks. The relative rank (scale of 0-1) for an individual stock was derived from the overall rank table by dividing the rank of a model by the number of models evaluated for this specific stock. Blank indicates the model was not applicable to the stock.

Figure 4. Correlations between retrospective forecasts for the period of 2009 – 2020 based on all 37 forecast models and observations for all 18 Fraser Sockeye stocks. Correlation between historical forecasts (denoted as Forecast) and observations was also shown.

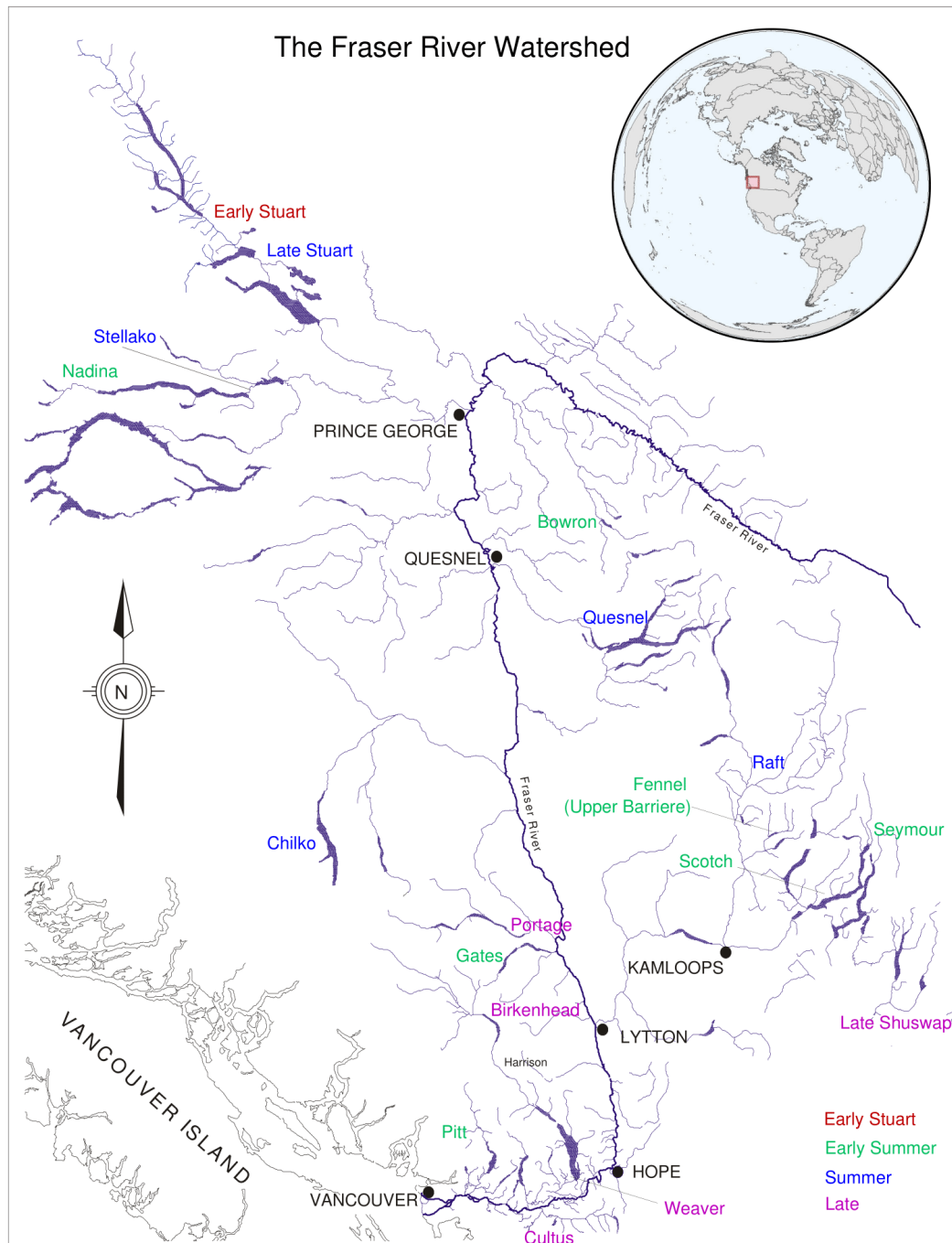
Figure 5. (a) Taylor diagrams for Early Stuart Sockeye stock, seven Early Summer run (Bowron, Fennel (Upper Barriere), Scotch, Nadina, Pitt, Seymour, Gates), and one Summer run (Chilko). (b) Taylor diagrams for four Summer run (Quesnel, Late Stuart, Stellako, Raft) and five Late run (Late Shuswap, Cultus, Portage, Weaver and Birkenhead). Each Taylor diagram compares 37

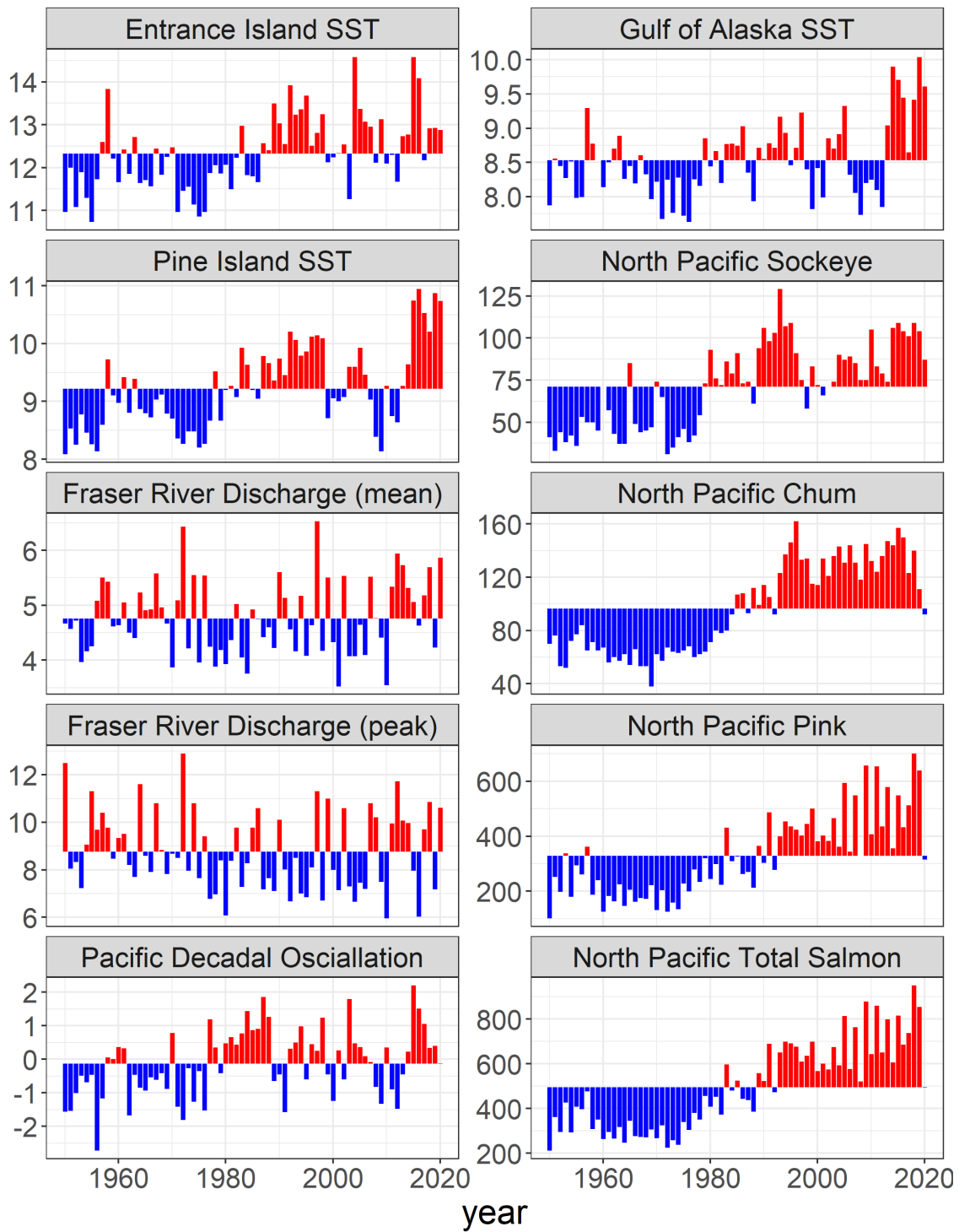
model forecasts and the historical Forecast (black solid square) against the Observation (black solid cycle on the x-axis). The distance from the origin is the normalized standard deviation with the normalized value for observations being 1. The angle describes the correlation (0 to 1) between model forecasts and observations. The dashed arcs around the Observation illustrate the root-mean-square error (RMSE). Models with negative correlations are not shown for each stock. The closer the model is to the Observation, the better predictive power the model has.

Figure 6. Comparisons among the forecast based on the historically selected model (Forecast) for the year 2022 and the forecasts by three top-ranked models using either best overall rank (Rank), standard deviation (SD), correlation (R), or root-mean-square error (RMSE) for three most abundant Fraser sockeye stocks: Chilko, Late Shuswap and Quesnel. The 50 percentile (P50: solid black line) and forecast distribution (box: P25-P75, range: P10-P90) were compared with the observation of the Sockeye return in 2022 (PSC preliminary results). *The Chilko stock used age-specific forecast models with age 4 using RickerEi while age 5 Sibling; Descriptions for the model abbreviations are in Table 1.

Table 1. Descriptions of 37 models under three categories (A: non-parametric/naïve, B: biological, and C: biological models coupled with covariates) along with Sibling model that was used as an additional option for forecasting age 5 specifically, where R is recruitment, S is spawner, a for the intercept measuring density-independent per capita R , and b for the slope measuring density-dependent R .

MODEL CATEGORY	DESCRIPTION
A. Non-Parametric (Naïve) Models	
LLY, R1C, R2C, RAC, TSA, RS1, RS2, RS4yr, RS8yr, MRS, RSC	Models only used returns from selected years, not considering spawners. Detailed model descriptions are listed in Supplementary Table S2.
B. Biological Models (Stock-recruit Models)	
Ricker	Bayesian Ricker model, $\log_e(R_t / S_t) = a - b S_t + \mathcal{E}_t$
RickerCyc	Same as above, using cycle line data only
Power	Bayesian power model, $\log_e(R_t) = a + b \log_e(S_t) + \mathcal{E}_t$
PowerCyc	Same as above, using dominant year data only
Larkin	Bayesian Larkin model, $\log_e(R_t) = a + b_1 \log_e(S_t) + b_2 \log_e(S_t) + b_3 \log_e(S_t) + \mathcal{E}_t$
LarkinCyc	Same as above, using cycle line data only
C. Biological Models (B for Ricker or Power) coupled with Environmental and Ecological Covariates	
B_FRD.mean	B coupled with Mean Fraser discharge flow from April to June
B_FRD.peak	B coupled with Peak Fraser Discharge at a given year
B_Ei.SST	B coupled with Mean Entrance Island sea-surface temperature (SST) from April to July
B_Pi.SST	B coupled with Mean Pine Island SST from May to July
B_PDO	B coupled with Pacific Decadal Oscillation in winter preceding outmigration from November
B_GOA.SST	B coupled with Mean Gulf of Alaska annual SST
B_Pink	B coupled with Abundance of pink salmon in the North Pacific Ocean
B_Chum	B coupled with Abundance of chum salmon in the North Pacific Ocean
B_Sockeye	B coupled with Abundance of sockeye salmon in the North Pacific Ocean
B_Salmon.Total	B coupled with Abundance of pink, chum, sockeye salmon altogether in the North Pacific
Additional model for age 5	
Sibling	Bayesian sibling model, $\log_e(R_{5,t}) = a + b \log_e(R_{4,t-1}) + \mathcal{E}_t$



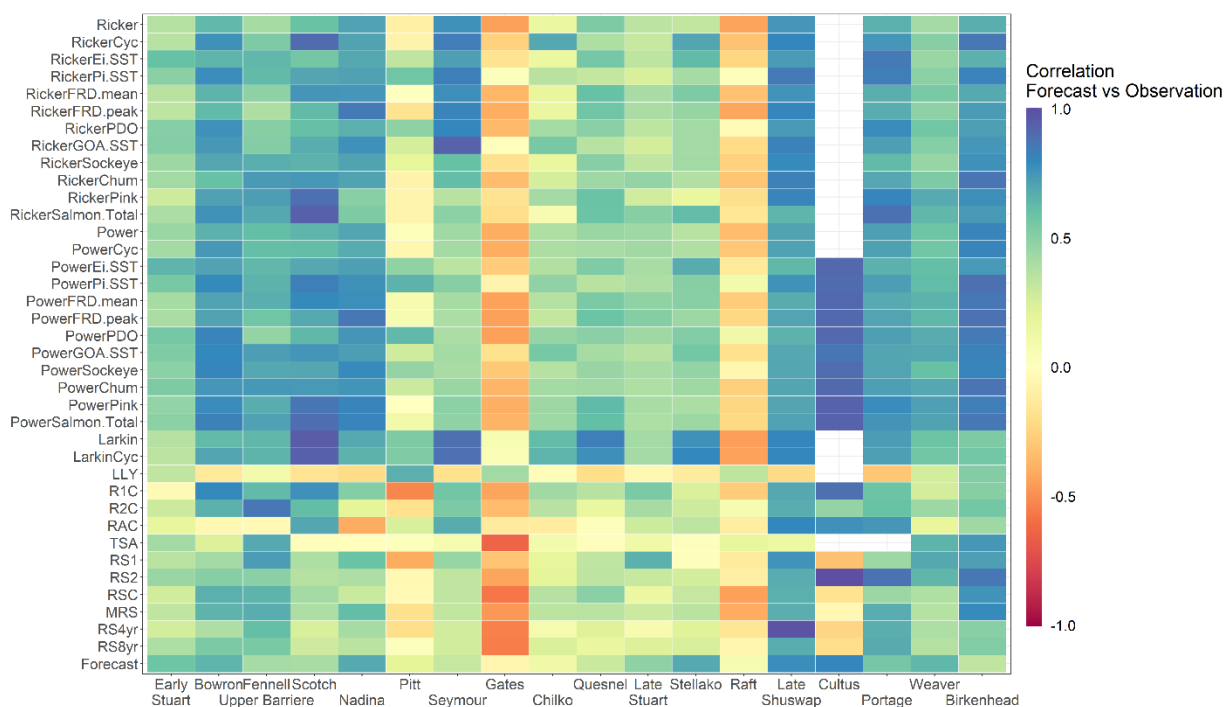


796 Figure 3.



797

798 Figure 4.



799

