Fisheries and Oceans Canada / Pêches et Océans Canada
Ecosystems and Oceans Science / Sciences des écosystèmes et des océans

Pacific Region

Canadian Science Advisory Secretariat
Science Advisory Report 2016/008

RUN TIMING AND DIVERSION RATE MODELS FOR FRASER RIVER SOCKEYE

Fraser River Sockeye Salmon returning to the spawning beds. Photo credit John Wallace.

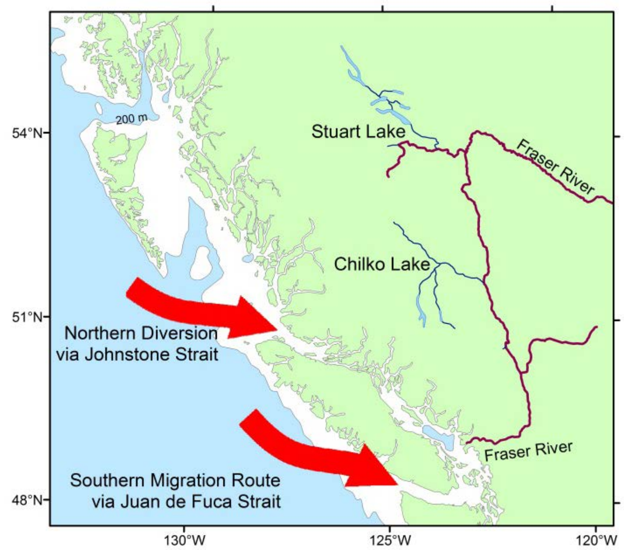

Figure 1. Migration routes available to Fraser River Sockeye returning to spawn.

Context:

Fisheries and Oceans Canada (DFO) Fisheries Management has requested that Science Branch provide advice for pre-season forecasts of adult Fraser Sockeye Salmon (Oncorhynchus nerka) marine run timing and diversion rate (the proportion of fish migrating through Johnstone Strait versus Juan de Fuca and Johnstone Straits combined). These forecasts are essential for planning fisheries and are a Canadian responsibility under the Pacific Salmon Treaty (PST).

The statistical models presently used to generate the run timing and diversion forecasts have degraded, and new models that utilize real-time oceanographic data have been developed and incorporated into new run timing and diversion rate models to improve performance.

The objective of this review is to assess the performance of newly developed models for the run timing of the Early Stuart and Chilko sockeye stocks, and for forecasting the northern diversion rate of Fraser River sockeye. DFO Science will use advice arising from this Canadian Science Advisory Secretariat (CSAS) Regional Peer Review Process to provide Fisheries Management, Pacific Salmon Commission and the Canada-US bilateral Fraser River Panel with annual forecasts of Fraser River Sockeye Salmon run timing and diversion rate.

This Science Advisory Report is from the October 27-28, 2015 regional peer review on the Run Timing and Diversion Rate Models for Fraser River Sockeye. Additional publications from this meeting will be posted on the [Fisheries and Oceans Canada (DFO) Science Advisory Schedule](#) as they become available.

SUMMARY

- Fisheries and Oceans Canada (DFO) Fisheries Management has requested that Science Branch provide advice for pre-season forecasts of adult Fraser Sockeye Salmon (*Oncorhynchus nerka*) marine run timing and diversion rate (the proportion of fish migrating through Johnstone Strait versus Juan de Fuca and Johnstone Straits combined).
- Numerous models to explain the relationship between observed migratory patterns of Fraser River Sockeye Salmon (the run timing and the northern diversion) and environmental correlates - including ocean surface currents, wind stress, and geomagnetic fields - were developed, and measures of performance - including precision and accuracy - were calculated. These data sources are updated frequently enough to allow their use for pre-season forecasts.
- A list of top-ranked models has been identified based on performance metrics. Relative performance of the models is presented as probability-based forecast certainty plots to guide managers and analysts in the selection of models. Multiple variable models perform better than single variable models, and current velocity variables (OSCAR - **O**cean **S**urface **C**urrent **A**nalyses – **R**eal time, and NEPSTAR - **N**orth **E**ast **P**acific **S**almon **T**racking and **R**esearch) dominate top performing models. It is recommended that the ranking process be undertaken annually until its robustness across multiple years is determined.
- The models examined provide statistical relationships to forecast run timing and northern diversions based on data collected from 1983 to 2012, and assume these years represent typical conditions. The forecasting of atypical events requires further model assessment based on management risk tolerance criteria.
- The ranking of the top performing models assumes a neutral weighting for bias and precision. This approach for ranking does not reflect specific objectives provided by fisheries managers, and it is recommended fisheries management consider the merits of this weighting during the pre-season, prior to selecting which models to use for the annual forecasts.
- A separate examination of El Niño and non-El Niño years revealed El Niño conditions during May or June of the return year are associated with a later Early Stuart run timing. Later Chilko run timing occurs when El Niño conditions are observed over any time between December and May of the return year. El Niño conditions in any month between September of the year prior to adult return, and May of the return year, are associated with a significantly higher northern diversion.
- Many of the models developed are dependent on data that are currently publicly available, but sourced from external agencies (e.g. National Oceanic and Atmospheric Administration [NOAA]). Others are dependent on contractors (NEPSTAR) and require an annual financial commitment.
- The evaluation of thousands of statistical relationships between environmental variables and fisheries data to identify an ensemble of models based on performance metrics is an approach that can be applied to other areas of fisheries management. An understanding of the underlying biophysical relationships provides confidence in how to interpret the statistical relationships, and insight into conditions that influence the strength of the relationships.

INTRODUCTION

Management of the Fraser River Sockeye fishery includes a pre-season planning component based, in part, on the forecast of the arrival time of salmon in local waters (when 50% of a stock has passed through Fisheries and Oceans Canada [DFO] statistical area 20, Juan de Fuca Strait), and the migration route around Vancouver Island (the northern diversion rate). Forecast errors can lead to missed catch opportunities, or overfishing. Since 1986, DFO Science has provided the Pacific Salmon Commission (PSC) with pre-season forecasts based on the statistical relationships between migratory patterns and environmental variables.

Every year the forecasts provided by DFO are compared to the post-season estimates of run timing and northern diversion based on observations by the PSC. For many years the return timing forecasts for Early Stuart Sockeye were moderately accurate, but their effectiveness has substantially declined in recent years (Figure 2, a and b). The Chilko timing forecast error has varied greatly throughout its history. The recent increase in forecast errors may be due to changes in environmental dynamics such that the statistical predictors used in previous models are no longer coupled to the variables they are attempting to forecast, or, as recent data are added to the time series, the statistical models themselves may be sensitive to outliers.

Figure 2(c) shows the northern diversion time series; every year prior to 1977 the majority of Fraser River sockeye returned to spawn via the southern route.

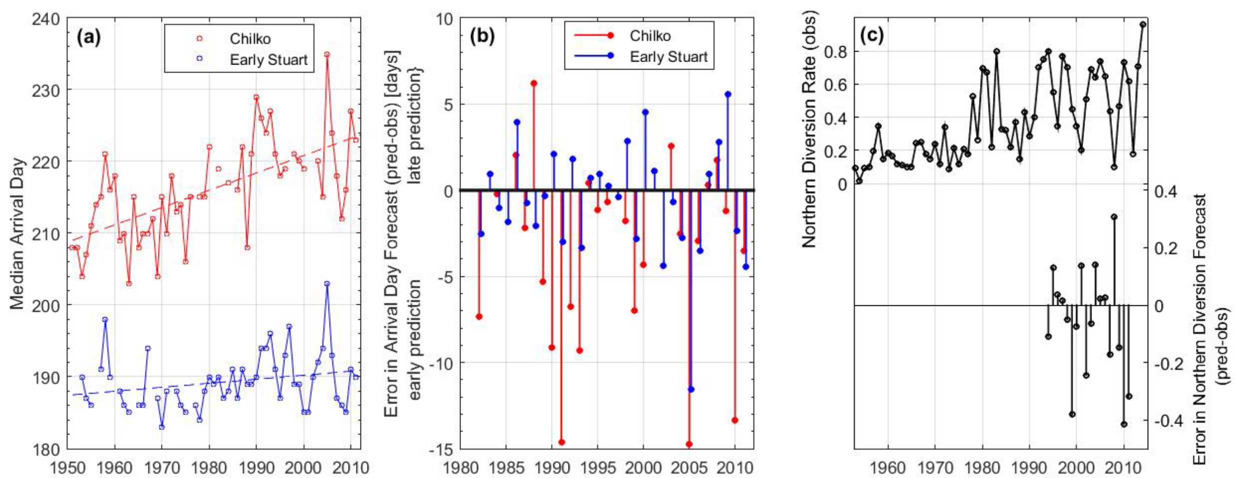

Figure 2. (a) Run timing (median arrival day in DFO statistical area 20) of Early Stuart and Chilko Sockeye from post-season estimates supplied by the Pacific Salmon Commission (PSC). (b)The timing forecast errors for early Stuart and Chilko Sockeye shown as the difference between the run timing predicted using earlier models and observations. (c) Top panel, The Northern Diversion (ND) Rate estimated from observations by the PSC; lower panel, the ND forecast error as the difference between predictions made with previous models and the observed ND.

The statistical relationships between observed migratory patterns of Fraser River Sockeye Salmon (the run timing and the northern diversion) were examined for the time period 1983-2012. The statistical models were evaluated by performance testing to determine forecast precision and accuracy. Statistical models with high performance rankings are good candidates to produce improved annual forecasts of Fraser River Sockeye Salmon migratory patterns.

ANALYSIS

An analytical framework was developed to evaluate statistical models that relate the run timing of Fraser River Sockeye (Early Stuart, Chilko) with environmental predictor factors, including sea surface temperature, sea surface salinity, near-surface ocean currents, and surface wind stress. These environmental variables, plus Fraser River discharge, sea level at Tofino, British Columbia (B.C), and geomagnetic fields, are used as predictors of the northern diversion rate of Fraser River Sockeye Salmon.

Four different types of statistical models were considered: naïve, linear regression with single and multiple variables, generalized additive models (GAM), and shape constrained additive models (SCAM). The environmental factors are assessed as single predictor variables and then combined with other variables in multiple linear regression models to determine which variables, geographical regions, time lags, and time-averaging periods reveal statistically significant relationships.

The influence of El Niño events was specifically examined by analyzing El Niño and non-El Niño years separately.

Data and Methods

The historical data of northern diversion rate and run timing were provided as estimates from observations made by the PSC. These data are the dependent variables in the models, and evaluating their statistical relationship with environmental data is used to forecast the northern diversion rate and run timing.

Eight types of environmental data were considered as predictor variables: sea surface temperatures and salinities, Fraser River discharge, relative sea level, earth magnetic field estimates, wind stress, and ocean current velocity. The latter includes two data sources, one from a numerical model (NEPSTAR - **N**orth **E**ast **P**acific **S**almon **T**racking and **R**esearch), and the other from satellite data (OSCAR - **O**cean **S**urface **C**urrent **A**nalyses – **R**eal time). These data sources are updated frequently enough to allow their use for pre-season forecasts.

The *Bivariate EnSo Timeseries* or the "BEST" index (Smith and Sardeshmukh 2000) was used to represent El Niño events. Based on both equatorial water temperature and atmospheric pressure the BEST index is considered a conservative measure resulting in fewer El Niño events than other indices.

Sources of sea surface temperature (SST) data include the National Ocean and Atmospheric Administration's (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) global grid of in-situ and satellite derived estimates, and daily measured surface temperatures and salinities from lighthouse stations along the B.C. coast.

Sea level data were provided by the Canadian Hydrographic Service for Tofino, B.C. Fraser River discharge was represented by data provided by the Canadian Water Survey for Hope, B.C.

A mathematical representation of the earth's magnetic field, provided by the International Geomagnetic Reference Field model, was used to determine the magnetic field strength and inclination angle at three locations: the mouth of the Fraser River, the seaward entrance of Johnstone Strait, and the seaward entrance to Juan de Fuca Strait.

Wind stress for the NE Pacific, derived from surface marine data gathered by NOAA's National Centers for Environmental Prediction (NCEP), was accessed on a 2^{o} by 2^{o} grid.

The OSCAR data series consists of the monthly mean north-south and east-west components of surface current velocity derived from satellite altimeter and scatterometer data since 1983 (Bonjean and Lagerlof 2002). Calculated on a ⅓^{o} latitude by ⅓^{o} longitude grid for the NE Pacific the data have been smoothed using a 5^{o} by 5^{o} filter to avoid local anomalies.

The NEPSTAR ocean model provides daily mean currents on a 1^{o} latitude by 1^{o} longitude grid for the NE Pacific over the period 1980 to 2013 (Thomson et al. 2013). Velocity fields are available at several depths and the currents at the sea surface and 30 m were used in the NEPSTAR regression analysis.

All the environmental predictor variables, with the exception of NEPSTAR, are available via public access on the internet through DFO or external agencies. NEPSTAR data are acquired through an independent contractor for a fee.

Every statistical model was assessed and those with statistically significant relationships (based on statistical significance [*p*] and coefficient of determination [R^2]) were tested further with performance analysis using retrospective and jackknife methods. The performance measures calculated for each model were used to rank the models.

Results

Significant relationships were found between the run timing of Early Stuart and Chilko stocks and El Niño events. El Niño conditions (as defined by BEST) during May or June of the return year are associated with a later Early Stuart timing, and a later Chilko run if El Niño conditions occur at any time between December and May of the return year. El Niño conditions in any month between September of the year prior to adult return, and May of the return year indicates there is a significantly higher northern diversion rate.

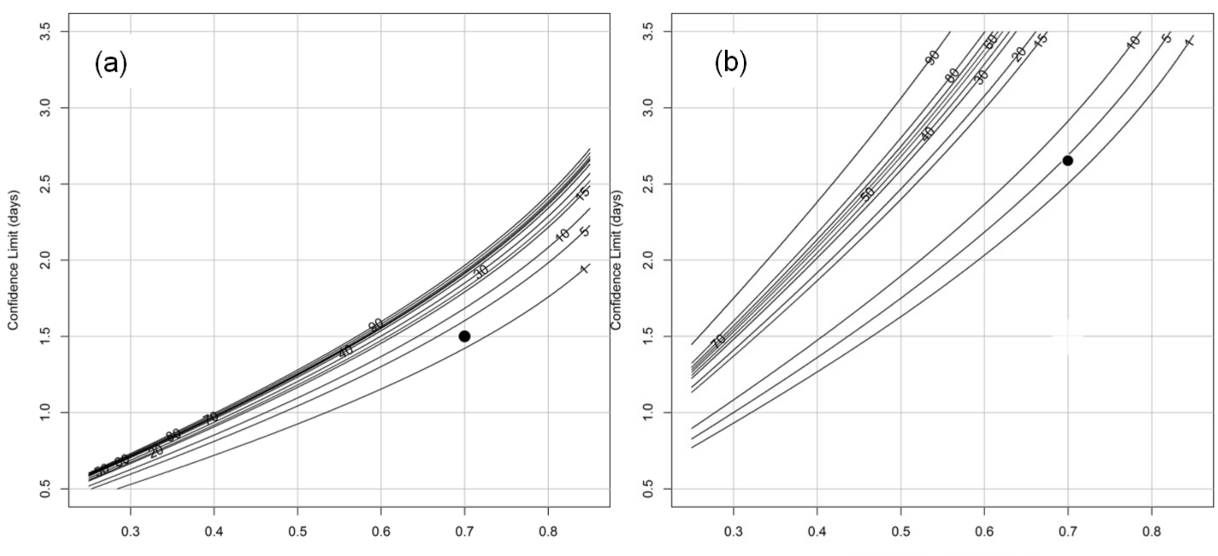

Figure 3. Forecast certainty as represented by confidence limit and confidence level for the models predicting run timing for Early Stuart (a) and Chilko (b) Sockeye Salmon. The number on each contour line defines how many models meet the conditions to satisfy the certainty criteria.

Generalized additive models regularly produced biologically unrealistic results and thus were excluded from performance analysis. Naïve models generally placed below the median value of rank; most were ranked near the bottom. Forecasts produced by the multiple regression models consistently performed better than the single variable models; all the top 50 models are multivariate.

The number of top performing models to consider when making forecasts depends on the user's criteria for accuracy and certainty, with the tradeoff being greater accuracy at the expense of lesser certainty and vice-versa. The analyses of the models show that with 70% confidence the top five timing models can forecast the arrival date of Early Stuart runs within ± 1.5 days and Chilko runs within ± 2.7 days (see Figure 3). The top ten northern diversion models are confident at the 70% level of forecasting the diversion rate within ± 0.13 (figure not shown).

Approximately 160,000 Early Stuart run timing models were evaluated and 381 passed the filter based on statistical significance (*p*) and coefficient of determination (R^2). For the Chilko run timing models 251,000 models were evaluated and 249 remained after filtering, and 290,000 models were evaluated for northern diversion forecasting with 177 selected for further examination. The variables associated with the top 50 run timing and northern diversion models, and their geographic locations, are shown in Figure 4.

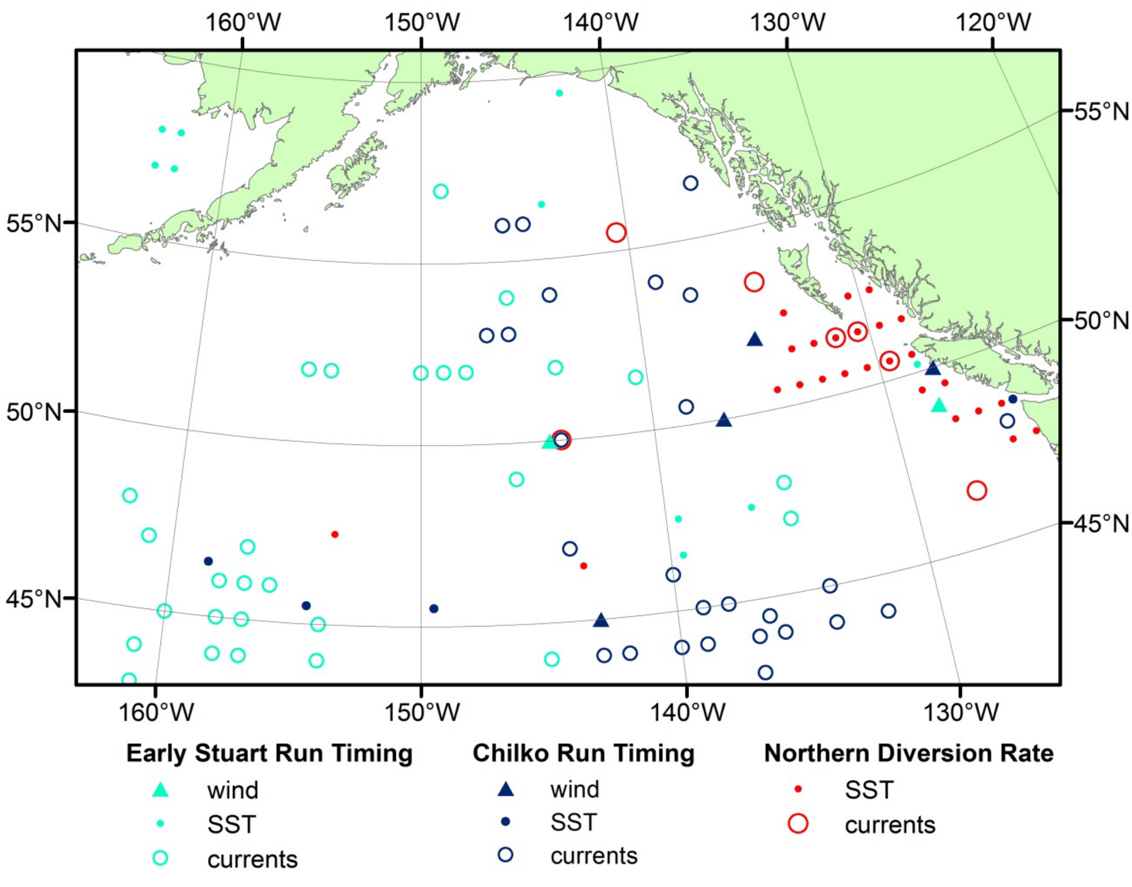

Figure 4. Locations of variables for the top 50 run timing forecast models (Early Stuart and Chilko) and northern diversion forecast models based on retrospective analyses.

A subset of the results of performance analysis for the models is given in Table1. The ranking of the models is based on an average of the many performance measures applied, and could differ if any particular performance measures was considered more suitable than others. It is clear that multiple variable models perform better than single variable models, and that current velocity data (OSCAR and NEPSTAR) dominate top performing models.

Table 1. The data source, year, and month of variables of the top three models, based on the average rank of performance measures, analysed with retrospective and jackknife methods; the year refers to variables in the return year (0) or the previous year (-1).

	Retrospective Analysis			Jackknife Analysis		
PM rank	Data type	Year	Month	Data type	Year	Month
	Early Stuart Timing					
1	OSCAR, OSCAR	-1,1	1,2	OSCAR, OSCAR, OSCAR	-1,0,-1	1,1,5
2	OSCAR, OSCAR, SST	-1,0,-1	1,1,8	OSCAR, OSCAR, SST	-1,0,-1	1,1,8
3	OSCAR, OSCAR SST	-1,0,-1	1,1,7	OSCAR, OSCAR, SST	-1,0,-1	1,1,7
	Chilko Timing					
1	NEPSTAR	-1,0,0,1	7,2,3,5	NEPSTAR	-1,-1,0,-1,-1,0,0	6,5,4,3,5,3,3
2	NEPSTAR	-1,0,0,-1,0	7,3,3,5,3	OSCAR, OSCAR, OSCAR	0,-1,-1	5,12,11
3	NEPSTAR	-1,0,-1	6,3,5	NEPSTAR	-1,0,0,-1	7,2,3,5
	Northern Diversion Rate					
1	NEPSTAR	-1,1,0	6,1,1	NEPSTAR	-1,-1,0	7,1,1
2	NEPSTAR	-1,0,-1,0	1,4,4,1	NEPSTAR	-1,0-1,0	1,4,1,1
3	NEPSTAR	-1,-1,0,0	7,1,4,1	NEPSTAR	-1,-1,-1,0	7,1,4,1

Sources of Uncertainty

The approach used to evaluate models to forecast Fraser River Sockeye return timing and diversion rate is based on the statistical relationship between environmental correlates and the dependent variables: the run timing and the northern diversion. These dependent data, provided by the Pacific Salmon Commission (PSC), are based on observations using methods that have changed over time. Any uncertainties in the dependent dataset have not been examined, and so the statistical relationships identified by the models do not take this source of uncertainty into account. However, all the models are evaluated using the same dependent data and therefore this uncertainty is not expected to bias the inter-comparison of the models.

The ranking of a statistical model will reflect the relative importance of accuracy and precision to its intended use. In the absence of specific direction on the relative importance of accuracy and precision, all the performance metrics were considered equally in the ranking of the models. The ranking of the models by the average of the performance measures can change depending on the weight associated with any particular performance measure.

Environmental variables that do not vary significantly over distance, and are used as input to several statistical models, may cause these models to be ranked similarly. When assessing an ensemble of models it is necessary to review the location of the input data to verify the importance of this effect.

The forecast certainty curves (see Figure 3) show that a more precise forecast window is associated with less confidence in the prediction. A comparison of the Early Stuart (a) and Chilko (b) models shows the greater uncertainty associated with the Chilko run timing models.

The relationships between salmon migratory patterns and historical environmental data are likely to vary with time, but the statistical models assume they do not. A forecast using an ensemble of several models can be used to evaluate the significance of this assumption.

The models examined provide statistical relationships to forecast run timing and northern diversions based on data collected from 1983 to 2012, and assume these years represent typical conditions. The forecasting of atypical events requires further model assessment based on management risk tolerance criteria.

CONCLUSIONS AND ADVICE

An analytical framework was developed to evaluate the performance of statistical models that relate the run timing of Early Stuart and Chilko Sockeye Salmon, and the northern diversion of Fraser River sockeye salmon, with environmental factors, including sea surface temperature, sea surface salinity, geomagnetic variation, ocean currents, and surface wind stress.

A list of top-ranked models has been identified based on multiple performance metrics. Relative performance of the models is presented as probability-based forecast certainty plots to guide managers and analysts in the selection of models. Multiple variable models perform better than single variable models, and current velocity data (OSCAR and NEPSTAR) dominate top performing models. It is recommended that the ranking process be undertaken annually until its robustness across multiple years is determined.

The ranking of the top performing models assumes a neutral weighting for bias and precision. This approach for ranking does not reflect specific objectives provided by fisheries managers and it is recommended fisheries management consider the merits of this weighting during the pre-season, prior to selecting which models to use for the annual forecasts.

Retrospective and jackknife analyses identified the top performing models as those with multiple variables that include current velocity and sea surface temperature. Many of these models are dependent on data that are currently publicly available, but sourced from external agencies (e.g. NOAA). Others are dependent on contractors (NEPSTAR) and require an annual financial commitment. The NEPSTAR current velocity variable is particularly important to the top models forecasting the run timing of Chilko Sockeye and the northern diversion of Fraser River Sockeye. The use of many of the top performing models depends on access to these data.

A separate examination of El Niño and non-El Niño years revealed El Niño conditions during May or June of the return year are associated with a later Early Stuart run timing. Later Chilko run timing occurs when El Niño conditions are observed over any time between December and May of the return year. El Niño conditions in any month between September (of the year prior to adult return) and May of the return year are associated with a significantly higher northern diversion.

The evaluation of thousands of statistical relationships between environmental variables and fisheries data to identify an ensemble of models based on performance metrics is an approach that can be applied to other areas of fisheries management. An understanding of the underlying biophysical relationships provides confidence in how to interpret the statistical relationships, and insight into conditions that influence the strength of the relationships.

SOURCES OF INFORMATION

This Science Advisory Report is from the October 27 – 28, 2015 regional peer review on the Run Timing and Diversion Rate Models for Fraser River Sockeye. Additional publications from this meeting will be posted on the Fisheries and Oceans Canada (DFO) Science Advisory Schedule as they become available.

Bonjean, F., and G.S.E. Lagerloef, 2002: Diagnostic Model and Analysis of the Surface Currents in the Tropical Pacific Ocean, Journal of Physical Oceanography, Vol. 32, No. 10, pages 2938-2954.

Smith, C.A, and P.D. Sardeshmukh, 2000. The effect of ENSO on the intraseasonal variance of surface temperatures in winter. International Journal of Climatology, Vol. 20, No. 13, pages 1543-1557.

Thomson, R., R. Hourston, and S. Tinis. 2013. OSCURS for the 21st Century: Northeast Pacific Salmon Tracking and Research (NEPSTAR) Project, Year 3 Interim Report. Annual report submitted to the Pacific Salmon Commission. 37p.

THIS REPORT IS AVAILABLE FROM THE:

Centre for Science Advice
Pacific Region
Fisheries and Oceans Canada
3190 Hammond Bay Road
Nanaimo, BC V9T 6N7

Telephone: (250) 756-7208
E-Mail: csap@dfo-mpo.gc.ca
Internet address: www.dfo-mpo.gc.ca/csas-sccs/

ISSN 1919-5087

Correct Citation for this Publication:

DFO. 2016. Run Timing and Diversion Rate Models for Fraser River Sockeye. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2016/008.

Aussi disponible en français :

MPO. 2016. Modèles de période de montaison et du taux de déviation pour le saumon rouge du fleuve Fraser. Secr. can. de consult. sci. du MPO, Avis sci. 2016/008.