Stock Separation of Chinook salmon at Bonneville Dam for Adult Upstream Migration Model (2008)

White Paper

W. Nicholas Beer
School of Aquatic and Fishery Sciences
University of Washington
Box 358218
Seattle, WA 98195
30 September 2008

Summary

The Adult Upstream Migration (AUM) model (CBR 2008a) makes daily, in-season predictions of the destination of fish that have arrived at BON. It requires stocks to be distinguished at BON by destination on a daily basis. In 2007, we began to use the historic passage patterns of PIT-tagged fish and visual counts to make this real-time prediction. Annual PIT-tag returns at BON, MCN, IHR and PRD were used to compute total returns to upstream zones. First, the difference in visual counts between MCN and BON were used to estimate the LCO returns. Second, the "within-zone" passage distribution (of PIT-tagged fish) was weighted by visual count numbers to determine the daily counts heading to a destination.

Recognizing that Chinook run timing has both genetic and environmental elements (Beer 2007; Anderson and Beer in review) we can refine this process further. In 2008, we began predicting the timing of the spring Chinook run at Bonneville using a genetic and environmental timing model along with in-season corrections based on environmental conditions and observations of passage (CBR 2008b).

For the 2009 passage season, this process is further refined. Adult Chinook timing and abundance is related to Jack-Chinook timing and abundance, therefore the timing model

can be used to predict the annual offset of the run timing compared to the arrival day average. Timing offset are a function of stock compositions and environmental conditions during winter and early spring, so the stock separation fractions can shift as well. This analysis: 1) identifies differences and similarities between the PIT-tagged fish arrivals and the run as a whole, 2) generates a method to predict stock separation based on PIT-tagged returns to verify consistency in timing differences and upstream travel rates, 3) uses visual counts and passage timing to obtain run timing distributions, and 4) recommends the technical methods for refining the AUM stock separation process during in-season runs.

Methods

First, identify returns of PIT-tagged Chinook that can be used for timing information. Adult PIT-tagged Chinook salmon that return to BON in each year have many attributes that can be identified by using their unique identification code. This includes: their release date, location and other pertinent information that identifies their origin, run type, rearing type, release location and other information (CBR 2008c). Aggregations of individual PIT-tag records within a year are necessary because each record is only an individual fish and we seek to understand the relationship of environmental predictors to Chinook returns in terms of their life-history and natal river. I considered aggregations of one or more of the following: HUC, River Region, Release Site and Species/Run/Rearing-Type (SpRRT) of the release group. Aggregations of records using any one of these individually was not appropriate. HUC alone was unsatisfactory because over the last 8 years, mixed or varying runs and rearing types indicate mixed life-histories sharing a HUC and a change in research activities and tagging efforts. For example, PIT-tag recoveries for the Clearwater River (HUC=17060306) were 95% of type “13H” in 2007 but only 43% of type “13H” in 2003. Release site alone was not satisfactory either for similar reasons, and tended to be too specific, i.e. it segregated otherwise similar fish unnecessarily. River region grouped fish according to environmental considerations without the (sometimes arbitrary) HUC boundaries but had mixtures of life-histories as well. SpRRT alone did not have the desired location component. Aggregating PIT-tag records by a combination of River Region, i.e.

destination and SpRRT gave the best signal for timing measures. These became PIT-tagged groups used in this analysis.

Chinook have multiple life-history strategies that blur the boundaries between life-history strategies and they are potentially identified incorrectly at the time of tagging or are given an ambiguous identification. Fish coded as Rearing="U" or "Run"=5 are ambiguous. Identification of adult Chinook is done as follows:

- Records of passage within the same year are ignored. None of these records give information on return timing of fish that have spent a winter in the ocean.
- Recoveries one year after release in spring or early summer are most likely Jacks.
- Recoveries two years after release in spring or early summer are adults.
- Recoveries two years after release in the fall are considered Jacks. Out migration after a fall release is improbable.
- Recoveries three years or more after release are adults.

Result: Jack fish within a group generally arrive later than the group's adults, and are positively correlated with the next year's adult returns. Both these are true for the visual counts. A regression of the adult peak day on the previous year's jack arrival day is highly significant (R^2=0.92 not shown) as has been shown for the spring and summer run at large. However, the relationship within individual groups can not generally be determined due to a lack of power: 1) There are only a few groups that have sufficient numbers in many years to obtain jack and adult mean arrivals. Pit-tag returns are on the order of ~ 3% of the visual counts during the spring and summer runs over the last decade. The fall run has even lower ratios of PIT-tagged fish. In turn, jack counts from the previous year are on the order of ~5-15% of the spring adult run. 2) Some groups are changing, and their life-history is conspicuously different across years. This is partially illustrated with Figure 1 that shows distinctions between the properties of PIT-tag group arrival statistics between 2000 and 2008 (at time of writing, late fall run PITs are not included). For example, "UpperCol.12H" (Summer hatchery fish from the Upper Columbia River) are a very consistent group, the mean time of arrival and the spread of the group is quite consistent. The earlier stocks are more variable by comparison, but this ignores annual differences that shift the entire timing sequence ahead or behind the average timing of the run.

Second, verify sequencing of the stocks. Keefer et al. (2004) observed sequencing for radio-tagged fish. Although there was variability in the sequences between years, the general impression is that arrival timing has a strong genetic component. The PIT-tag groups are sorted by year and plotted according to their mean arrival day (Figure 2. This also shows visual mean arrival days from the arrival distribution (Beer 2007))

Result: Figure 2 demonstrates that the annual run timing signal is most apparent in the early arriving stocks while within-group variability is more dominant for stocks that arrive later in the year. This is consistent with the observation that over 30 years of passage at BON, the standard error on the estimate for the mean spring peak day is 5.5 days while for the fall run peak it is 2.5 days. The fall run is much less variable. The summer run is evolving so that count and timing patterns have been difficult to predict.

Third, normalize the arrival distributions to make arrival timing more predictable. Early season arrivals are known to be influenced by environmental conditions (Beer 2007) which interacts with genetic predispostion of the fish to return at the optimal time (Anderson and Beer, in review). So, I hypothesize that the run is compressed or expanded by the environmental conditions such that the earliest arrivals are most influenced by annual environmental variability and the lastest arrivals are least affected. Using the estimated mean arrival day M_y of the spring run with a year (y), I compute a year-independent arrival date ($\hat{T}_{i,y}$) which is expressed as a real number instead of an integer, for each stock (i) in each year (y) with a tapered correction for arrival timing relative to the entire season:

$$\hat{T}_{i,y} = T_{i,y} - \Delta T_y \cdot \left(1 - \frac{T_{i,y} - M_y}{D_{final} - M_y}\right) \tag{1.1}$$

for $T_{i,y} <= D_{final}$, where D_{final} is the last day to apply the tapered correction, M_y is the visual mean peak day for year y, $T_{i,y}$ is the observed mean arrival day of the PIT-tagged group i. Eqn (1.1) ensures that the largest normalizing correction is applied to stocks arriving earliest and there is no offset for any stock where $T_{i,y} >= D_{final}$.

Result: The effectiveness of this method for normalizing the arrival times and eliminating the within-year environmental factor is shown in Figure 3 which illustrates the relative benefit of adding a tapered correction. The abscissa is the mean arrival day of group *i* in year *y* ($\overline{T}_{i,y}$) and the ordinate is $mean(abs(T_{i,y} - \overline{T}_{i,y})) - mean(abs(\hat{T}_{i,y} - \overline{\hat{T}}_{i,y}))$ so positive values are improvements to the prediction of timing and negative values are not. Some of the summer runs have been very inconsistent with $mean(abs(T_{i,y} - \overline{T}_{i,y}))$ values up to 21 days (MiddleCol.11U). To the extent that their genetic composition has varied, this type of correction for environmental conditions will not improve the run's predictability since genetic predispositions are dominating the run time signal for the group. This is consistent with the observation that early runs are also the most compact, having less spread than the later ones. This can be seen in Figure 1 for the spring and summer stocks arriving before the end of summer near day 220. For the stocks illustrated prior to day 220 with more than 10 fish in each group, a regression of standard deviation of arrival on the mean arrival day is highly significant ($p< 0.0001$). The fall runs are quite different. Most of the points are clustered together, suggesting remarkable consistency and probably, different timing mechanisms.

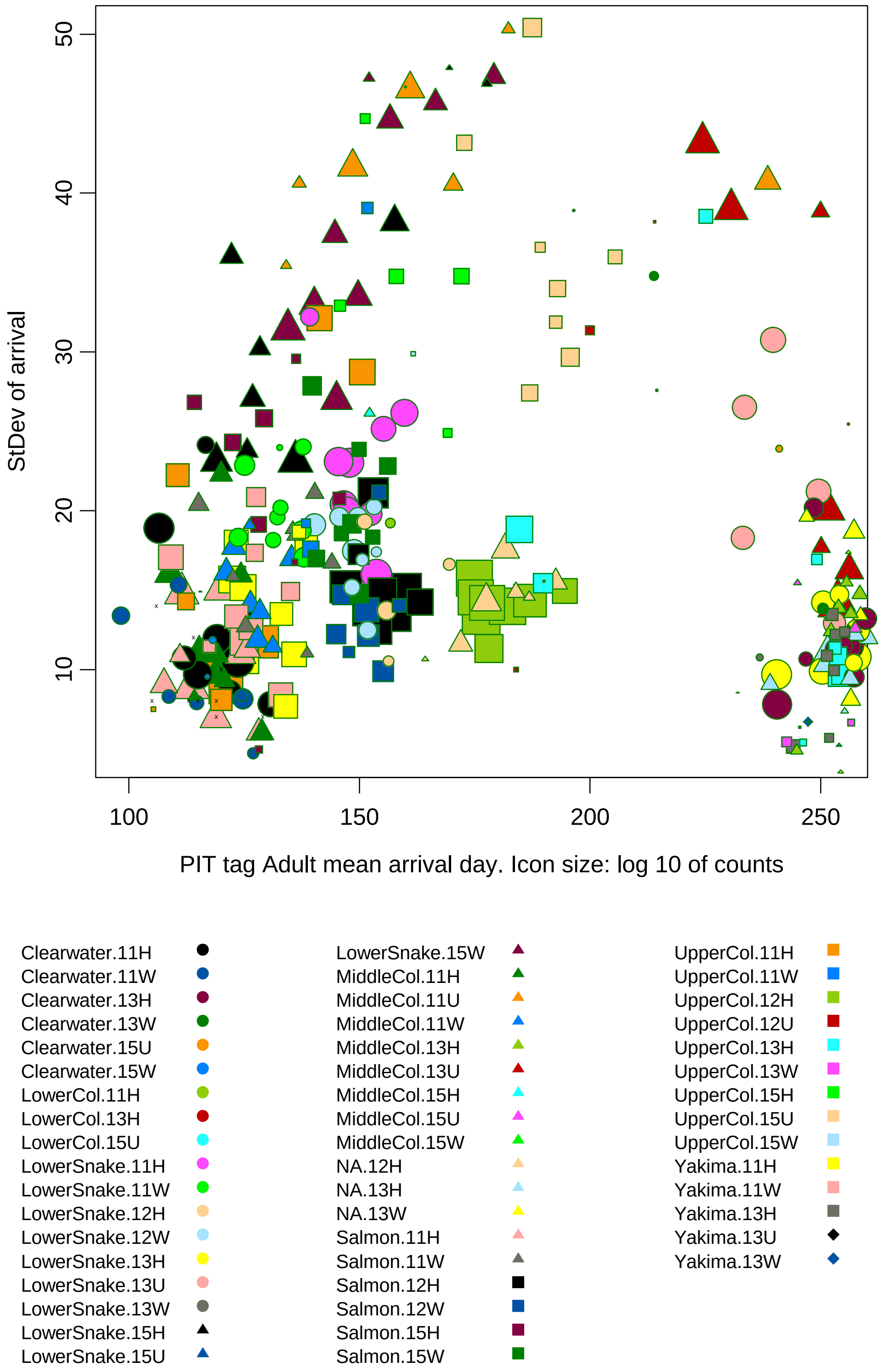

Figure 1 Distributions of PIT-tagged Adult Chinook passing Bonneville Dam 2000 to 2008. Relationship of mean arrival day to spread of arrival. Each icon represents the returns of a single group of PIT-tagged fish in a particular year. UpperCol.12H fish have been very consistent in their arrival timing and distribution. LowerSnake.15W have been much less so. Details of the early stocks are shown below and individual stocks are in the Appendix.

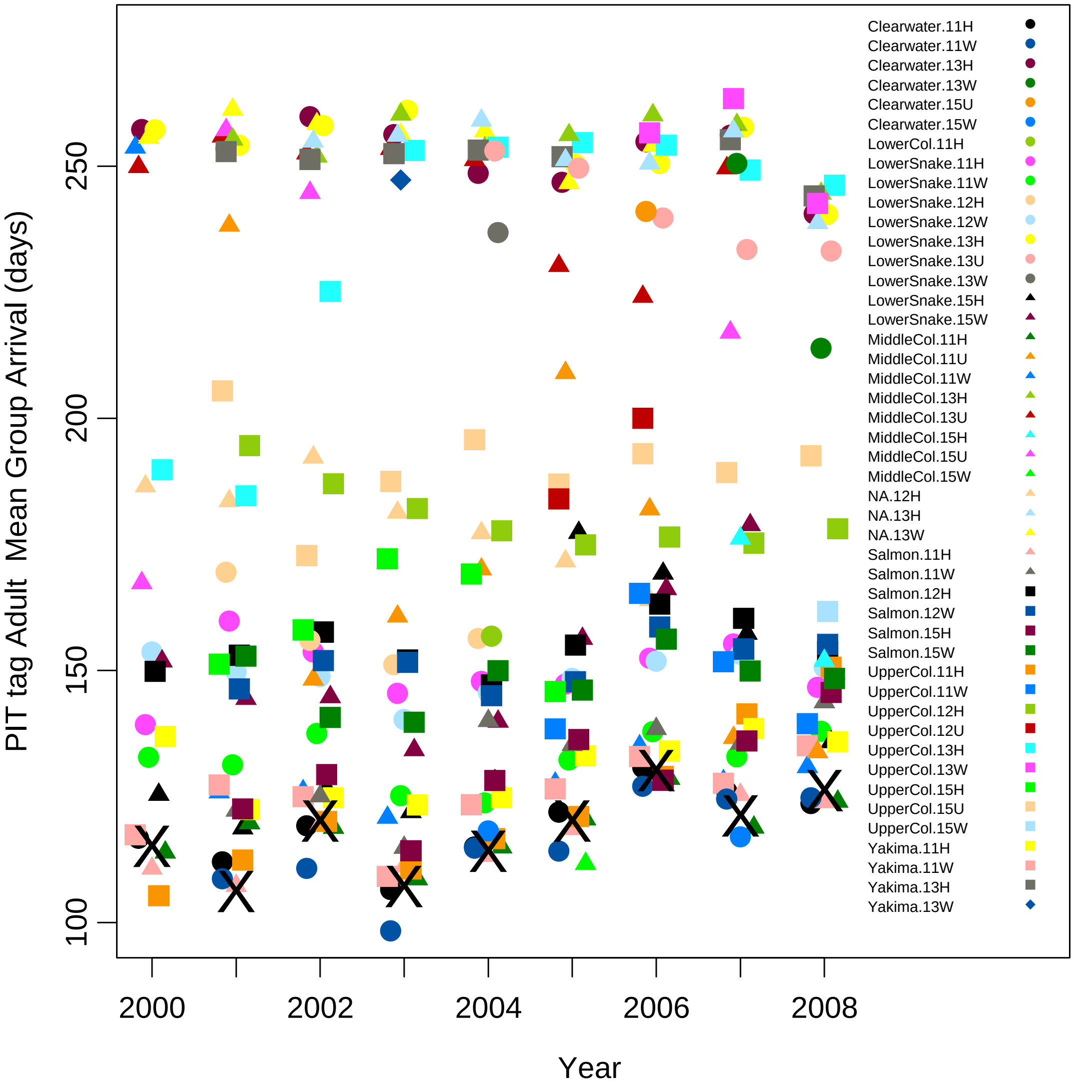

Figure 2 Sequencing in all PIT tag adult return groups atBON based on destination river system and Species Run and Rearing Type (SpRRT) designation. The black X's are the fitted Spring Adult Arrival mean day. Points dithered slightly to improve readability.

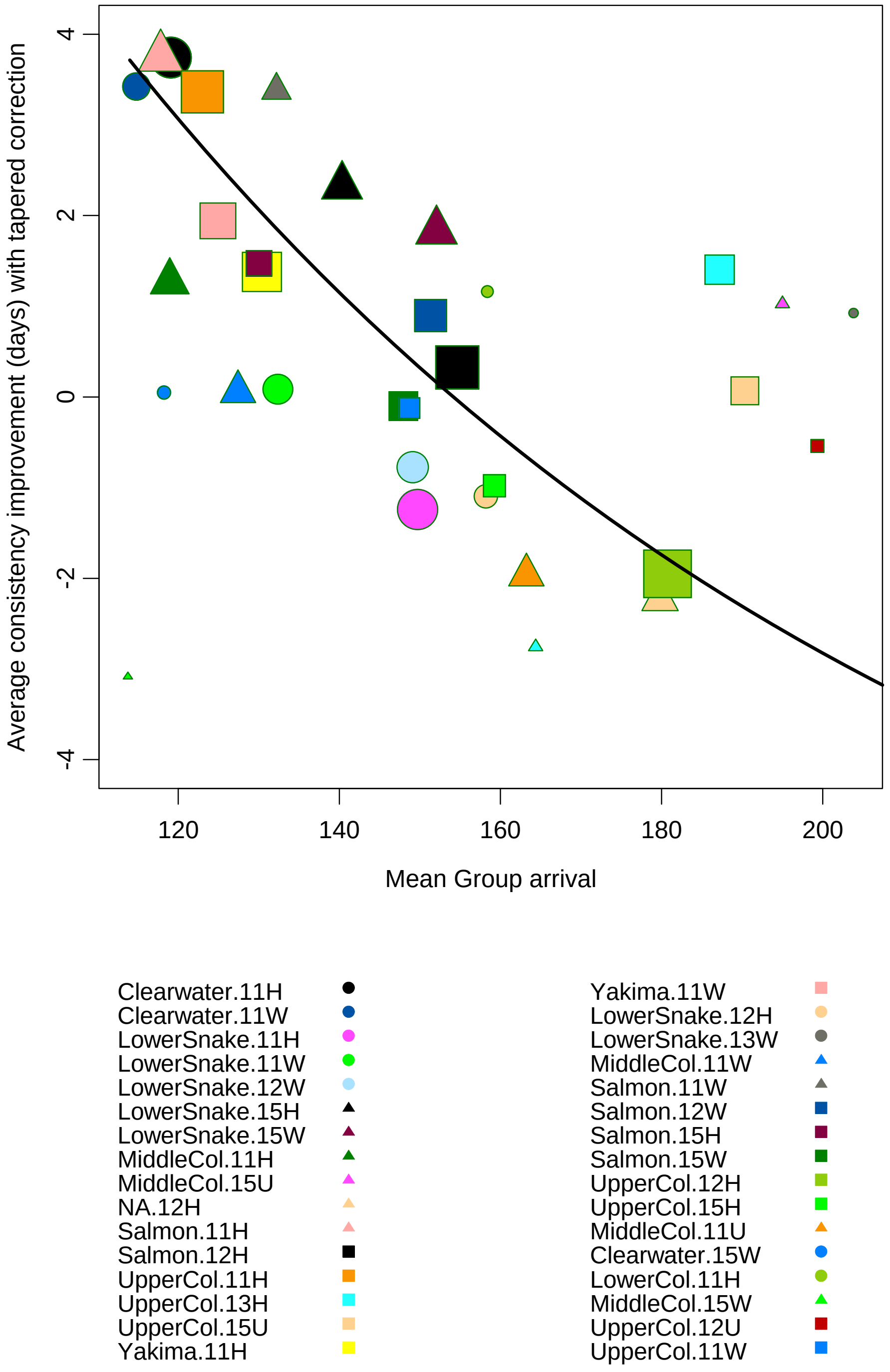

Figure 3 Applying a tapered correction between "mean arrival (calendar) day" of a group and "end of summer season day" to the arrival timing of PIT-tag groups improves their predictability relative to the other groups. The more positive points have the greatest improvments, i.e. the group's timing anomaly corresponds closely to the visual count's timing anomaly. This deteriorates over time. Evidence that annual differences in adult run timing compress or expand the runs, not simply shift

the runs. Arrivals of PIT-tagged fall runs (after day 210) have negative numbers. A taper out to day 145 is certainly beneficial.

Fourth, apply normalizing method to visual counts at all dams to account for annual environmental variability. The impression that the runs are "compressible" is justified since predictability is greatest for the earliest stocks and diminishes smoothly over time once the annual offset is known. Thus, improvments in the predicted destination of fish passing Bonneville can be made by ignoring the calendar time of arrival and examining the relative timing of arrival. The simplest approach is to imagine the season as being comprised of several arbitrary time steps between the beginning and the end of the run. These time steps are stretched or shrunk in a systematic way to normalize the arrivals for comparison. At a dam (*i*), passage can be transformed to the normalized arrival schedule with an algorithm similar to the one used for the PIT group means. Ignoring subscripts for dam *i* and year y, the arrivals for each integer day (*t*) are mapped onto the normalized timeframe $\hat{t}_y$ as:

$$\hat{t} = t - \Delta T \cdot \left(1 - \frac{t - M}{D_{final} - M}\right) \tag{1.2}$$

The summer run at present appears to be a mixture of fish that respond to the offset and others that do not. Note that the method can be applied to any portion of the run and the taper adjusted:

1. Setting D_{final} to a large value will offset all the fish equally.
2. Restricting $t \leq k$ will normalize only the first *k* days of the run.

There is a corresponding count of fish, n_t, on time step $\hat{t}$, which is allocated proportionally to the two integer days (greater and less than $\hat{t}$). First, the normalized integer days τ are identified with the modulo operator: $\tau_a = \hat{t} - \text{mod}(\hat{t},1)$ (the "floor" of $\hat{t}$) and $\tau_{a+1} = \tau_{a+1} + 1$ (the "ceiling" of $\hat{t}$). The counts are then allocated as follows: $n_{\tau_a} = (1 - \text{mod}(\hat{t},1)) \cdot n_t$ and $n_{\tau_{a+1}} = \text{mod}(\hat{t},1) \cdot n_t$. Depending on the normalization details, ther may be more than two values to sum for a day, so the final count of the fish on day τ is:

$$n_{\tau} = \sum_{j=0,1,2,\ldots} n_{\tau_{a+j}} \qquad (1.3)$$

An illustration of this for BON dam is in Figure 4.

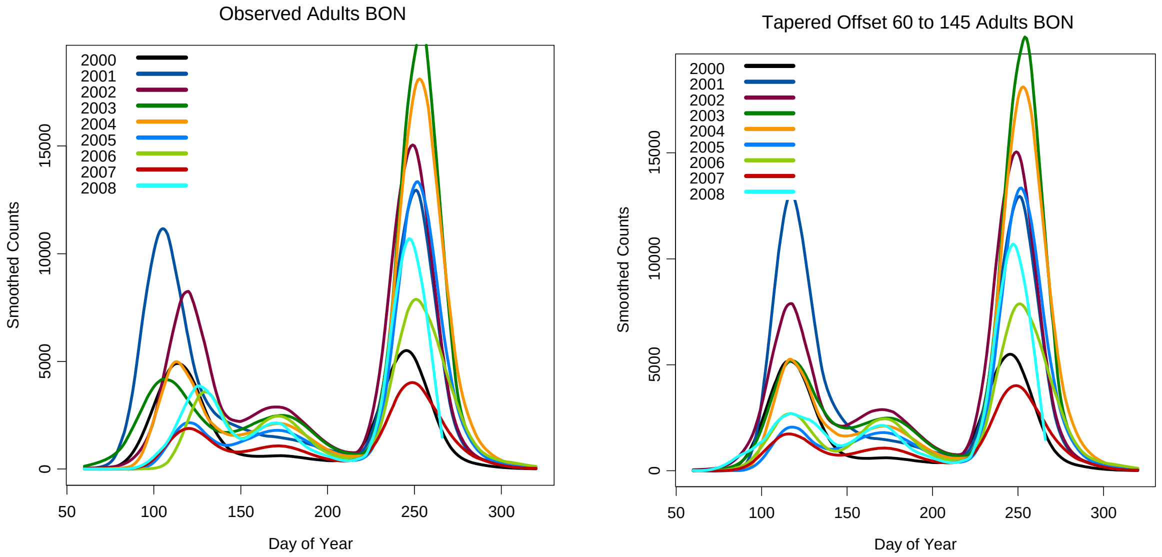

Figure 4 For BON dam passage, the benefits of normalizing the timeframe of the observed passage over recent years is illustrated. Note the alignment of the peaks in the right frame for the first mode of the passage distributions compared to the left frame.

Fifth, use passage patterns at upper dams to back-calculate travel times from one dam to another. The annual pattern of arrivals at BON is a good predictor of the timing at MCN, and this in turn is a prediction of the passage at PRD, IHR and PRO together with the remainder of the fish destined for a river location somewhere between these dams. Although we expect mortality and residency to reduce the number of fish between the lower and upper dams, this is not always the case. At the time of writing there is no known resolution to the problem that the sum of the PRD, IHR and PRO dams is greater than the MCN passage during a significant stretch of the season for every year since 2001 with the exception of 2006 (e.g. see Figure 5). The lag between the BON and MCN locations is determined by selecting the value of *lag* that satisfies:

$$\min\left(\sum_{i=t_0+lag}^{i=t_1} (BON_i - MCN_{i+lag})^2 / (t_1 - t_0)\right) \quad \Big| \quad lag = 1,2,3,\ldots$$

These values vary by year and season (spring-summer t_0=1, t_1=220 and fall t_0=221, t_1=365). From BON to MCN over the last 9 years the average spring lag was 6.9 days for spring/summer and 4.4 for fall. Between MCN and three dams above it, a variation in the method is used to accommodate the fact that PRD and PRO are significantly farther upstream than IHR. Thus assuming that extra time is required for ascent of either the Columbia River to PRD or the Yakima River to PRO, a site specific lag of j = 1, 2 or 3 days was evaluated, as well as choosing a common lag for all fish to account for travel between MCN and the Snake-Columbia confluence, such that it satisfies:

$$\min\left(\sum_{i=t_0+lag}^{i=t_1}\left(MCN_i-\left(PRD_{i+lag}+PRO_{i+lag+j}+IHR_{i+lag+j}\right)\right)^2/(t_1-t_0)\right) \quad \Big| \quad lag=1,2,3,\ldots$$

A site-specific ascent lag to PRO and PRD of 2 days stabilized the common lag to 1.5 days for the spring and summer runs and 1 day for the fall runs (see Table 1).

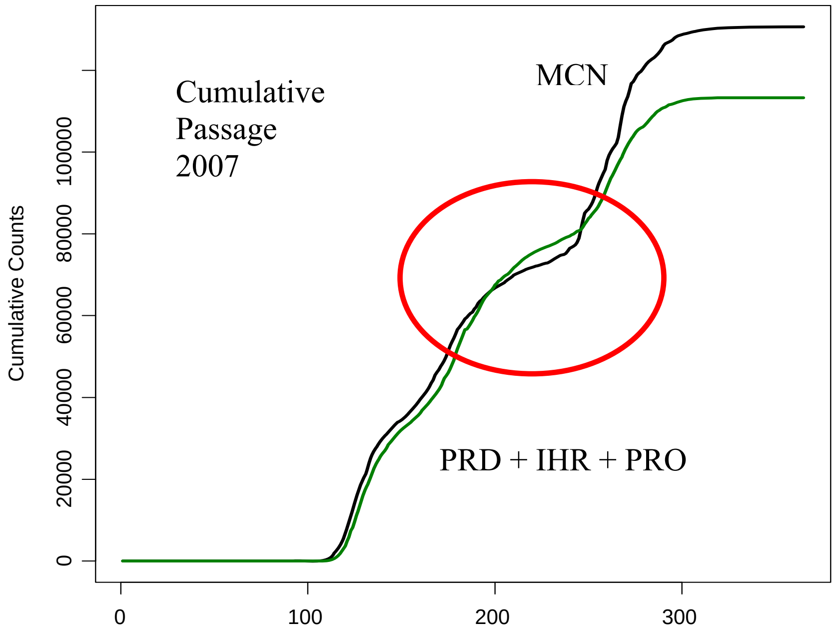

Figure 5 Upper dam counts paradoxically exceed lower dam counts.

Table 1 Estimated lags between dams on the Columbia/Snake Rivers

	Lag1	Lag 2
	BON to MCN	MCN to upper dams + $j = 2$ days to PRD + $j = 2$ days to PRO + $j = 0$ days to IHR

Year	Spr/Sum	Fall	Spr/Sum	Fall
2000	10	2	2	1
2001	7	7	2	1
2002	8	5	2	1
2003	6	5	1	1
2004	5	4	1	1
2005	6	5	2	1
2006	7	5	2	1
2007	7	6	2	1
2008	6	NA	1	NA

Finally, use a passage prediction to obtain a stock separation fraction. Unfortunately, neither observations of PIT tag returns nor any visual counting method at BON is completely reliable for gauging the abundances in the run at large from day to day. The PIT returns in particular are biased toward the summer stocks and the fall run has the lowest proportion of PIT tagged fish (Figure 6). However, some estimate of the expected passage numbers must be made and the previous year's arrivals may be our best hope. Both jack and adult returns (separately enumerated at Columbia/Snake River dams) are auto-correlated from year to year due to age overlaps in the returning stocks and their shared ocean survival experience. Over the last 20 years, jacks have been harbingers of adult returns for spring and summer stocks ($R^2 = 0.78$) but have not been as useful for predicting the fall runs. In the fall, the previous year's adult run is the better predictor ($R^2 = 0.68$) of the adult run.

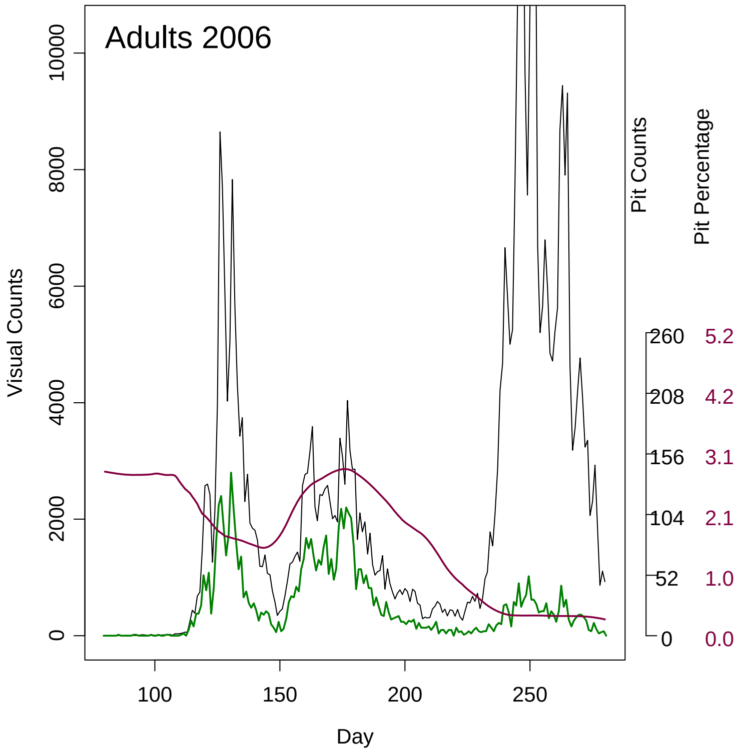

Figure 6 Relationship of PIT passage counts to visual counts at BON for 2006.

A data-based method of making predictions is preferred over more theoretical methods. Over recent years (2006 – 2008), passage timing of the normalized adult returns from the previous year were better predictions of adult passage than the normalized jack returns.

Assuming that a predicted arrival pattern has been identified (e.g. the adult returns for the previous year) the stock separation process can begin. The stock separation process hinges upon:

1. Normalizing the timeframe and counts of upper dam passages according to Eqn(1.2) and Eqn(1.3)
2. Shifting the passage numbers back in time downstream to BON using dam to dam survival. Survival rates are on the order of 98% although hard to determine when more fish are counted upstream than downstream.
3. Shifting the passage distribution back in time downstream to BON using travel time lags.
4. Identifying the destination of the fish as they pass BON on day τ (normalized travel day) and conversion back to day *t*.

The number of fish that passed BON on day τ that will then continue to MCN is based on the observed numbers at MCN and the travel time lag:

$$MCNpass_\tau = BON_\tau - S_{BON-MCN} \times MCN_{\tau+lag1} .$$

and the fraction of fish that do not pass MCN on day τ but stay in the Lower Columbia is $fLCO_\tau$:

$$fLCO_\tau = (BON_\tau - MCNpass_\tau) / BON_\tau .$$

Passage to PRD dam, IHR dam, and PRO dam are computed similarly:

$$(PRDpass + IHRpass + PROpass)_\tau = MCNpass_\tau - S_{MCN-upper} \times (PRD_{\tau+lag1+lag2} + IHR_{\tau+lag1+lag2} + PRO_{\tau+lag1+lag2})$$

From this, fractions of fish destined to the three upper dams (*fPRD*, *fIHR*, and *fPRO*) and the remaining area (Hanford reach of the Columbia River, lower Yakima River, and MCN pool, a.k.a. *fHAN*) are determined.

$$fHAN_\tau = \frac{MCNpass_\tau - (PRDpass + IHRpass + PROpass)_\tau}{MCNpass_\tau}$$

$$fIHR_\tau = \frac{IHRpass_\tau}{BON_\tau} , \text{ with others similar}$$

Under certain circumstances with a late season lag less than an early season lag, a gap in the counts is possible. The value for the gap day(s) is smoothed with a three day window of counts from before and after the gap.

Finally, any fraction or passage count can be converted back to calender days with the reverse of the offset method. For example, using Eqn (1.2), $\hat{t}_y$ and t_y are switched and then n_t is computed by switching n_t and n_τ.

In practice, the counts are smoothed over at least 7 days because the noise in the arrival distributions is a nuisance without significant predictive benefit.

Illustrations of the method are shown below using 2007 passage as a predictor for the 2008 fractions.

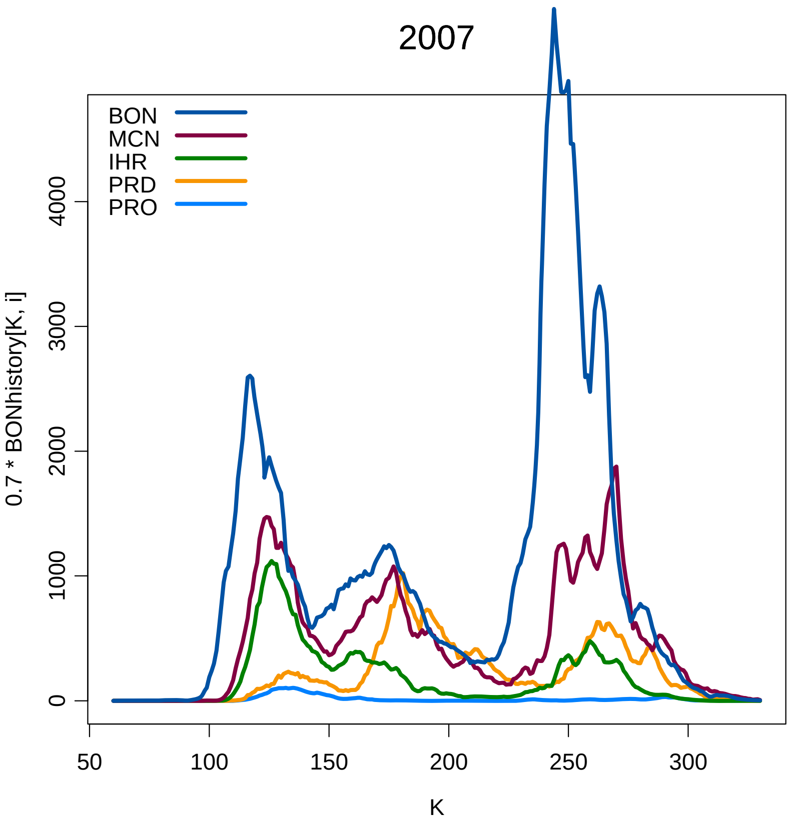

Figure 7 Raw observations of passage at five sites during 2007.

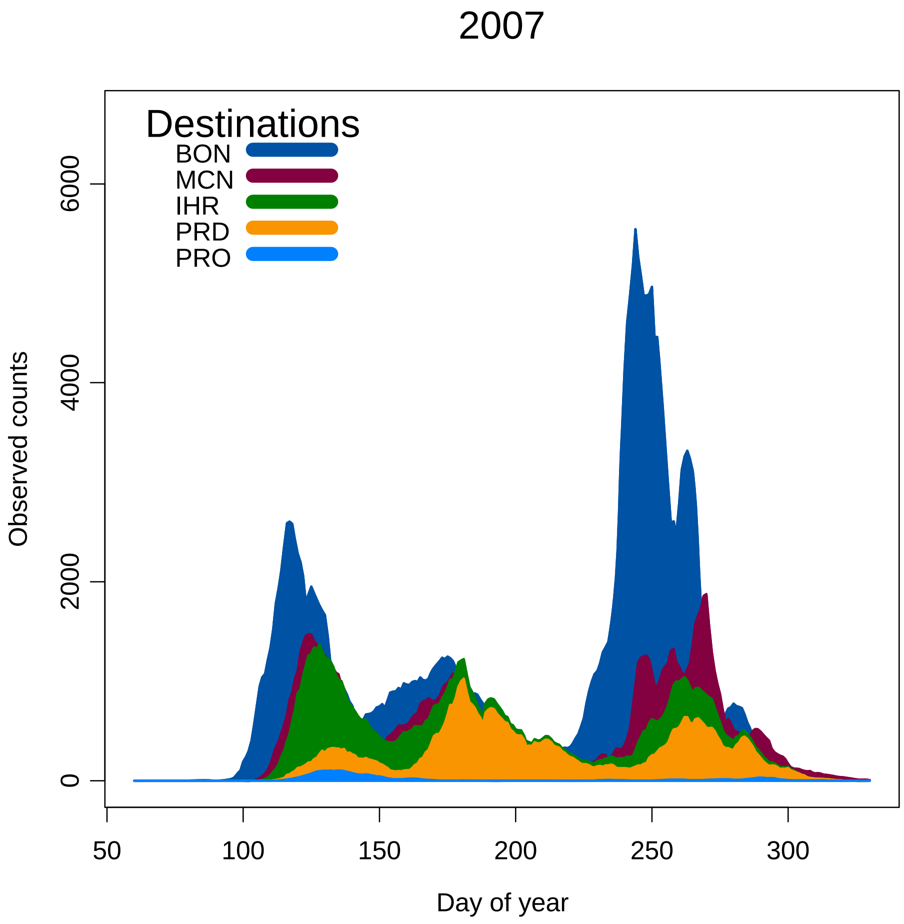

Figure 8 Expected destinations of fish in 2007. (One week smooth)

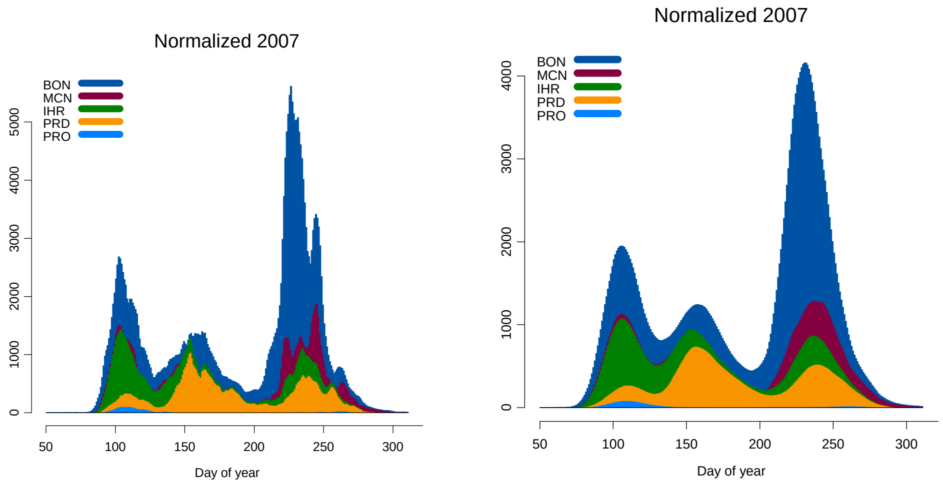

Figure 9 Prediction of passage destinations for stock separation purposes in normalized time frame. Left: one week smoothing. Right: S-Plus 6.1 (Insightful Corp.) “supsmu” cross-validated variable span smoother.

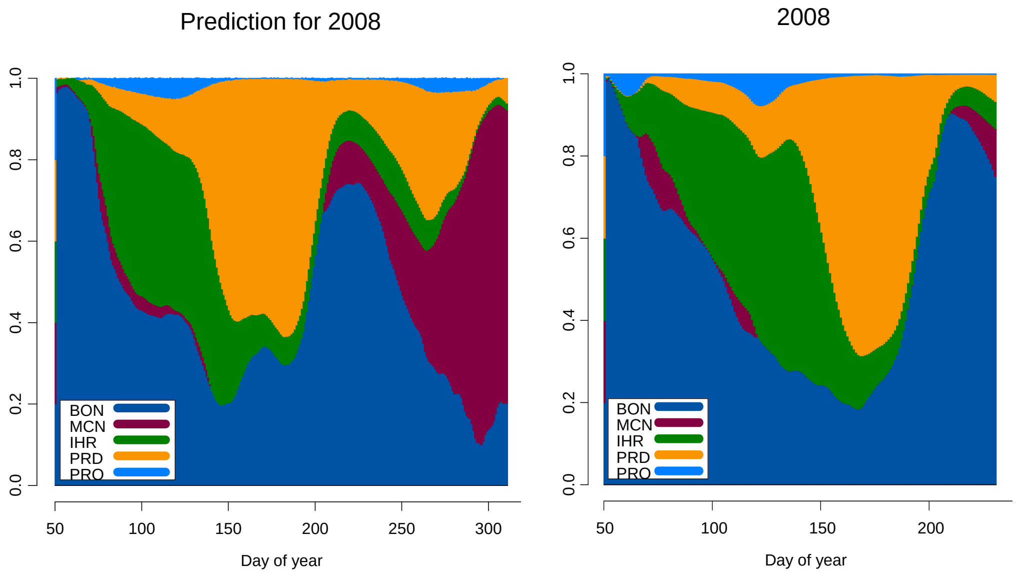

Figure 10 Left: The final partitioning of the Chinook passing BON to five possible destinations based on 2007 passage. Note that the fish destined for BON could go anywhere between BON and MCN including tributaries and hatcheries. Those designated MCN are destined to go to any part of the Yakima, Snake, or Columbia Rivers below PRO, IHR, and PRD respectively. Fish destined for the other dams could go anywhere above them. Right: Best estimate of 2008 passage fractions as of time of writing. Notice that the 2008 observations only go through day 250.

Conclusions

Stock separation fractions at BON are generated with a data processing algorithm using historic and current observations. Inputs required on a daily basis in-season to adjust the stock separation fractions include:

1. Historic, simulated, or idealized passage distributions at five dams: BON, MCN, IHR,PRD, and PRO that will represent their final distribution.
2. The Adult Peak Day prediction for spring Chinook at BON(CBR 2008). This provides a real-time update to the current year's offset which stretches or contracts the expected distribution.
3. Swim time lags between lower and upper dams.
4. Offset time limits

Development in the future could include:

1. Methods to better match an historic passage distribution with a predicted one. This may require more detailed information on the age structure of the populations passing dams, ocean survival metrics, and details on changes in management practices above dams that alter the life histories of stocks from that river.
2. Methods to mix the jack and adult predictions of passage. This may require some stability on the part of the runs which appear to be changing. Most notably, the summer run has been high, but quite variable since 2001.
3. Increased understanding of the mechanisms that create timing of summer and fall runs. The spring run timing has been tied to January upwelling in the Pacific ocean and in-river flows prior to migration (Anderson and Beer in review). Since the freshwater life history of the stocks in the summer and fall runs is different, differences in the response to ocean conditions need to be examined as well.

References

Anderson, J. and Beer, W.N. *in press.* Oceanic, riverine, and genetic influences on spring Chinook salmon migratory timing.

Beer, W.N. 2007. Run timing of adult Chinook salmon passing Bonneville dam on the Columbia River. On-line at: http://www.cbr.washington.edu/papers/chinook.run.timing.2007.pdf (Accessed September 1, 2008)

CBR. 2008a. Adult Passage Predictions based on Visual Counts for Combined Columbia and Snake River Stocks. On-line at: http://www.cbr.washington.edu/crisprt/index_adult.html (Accessed September 1, 2008).

CBR. 2008b. Bonneville Dam Adult Spring Chinook Daily Peak Arrival and Run Size Prediction. On-line at:

http://www.cbr.washington.edu/crt/get?FishType:Stock:Year=fc1:SpringAdult:2008&Config=adultPeak&Segment=Bonneville_Dam&Plot=adultPeak&DataSource=apc&AutoScale=on&DateInput=crispAdultPeak (Accessed September 1, 2008).

CBR. 2008c. DART PIT Tag Adult Returns by Observation Year Detail Report. On-line at: http://www.cbr.washington.edu/dart/pit_obs_adult_obs_de.html. Accessed September 1, 2008.

Keefer, M., Peery, C., Jepson, M., Tolotti, K., Bjornn, T. and Stuehrenberg, L. 2004. Stock-specific migration timing of adult spring-summer Chinook salmon in the Columbia River Basin. North American Journal of Fisheries Management 24:1145-1162.