Western Alaska Salmon Stock Identification Program

Technical Document:[1] 16
Version: 1.0

Title: Examining prior sensitivity using the chum salmon baseline
Authors: C. Habicht, W.D. Templin, J. Jasper
Date: September 26, 2011

Introduction

During the joint Advisory Panel (AP)/Technical Committee (TC) meeting held in Anchorage on September 21 and 22, 2011, Gene Conservation Laboratory (GCL) presented options for establishing priors for both sockeye and chum salmon that are required for analyzing fishery mixtures using Bayesian methods. By the end of the meeting, we had consensus from the AP, pending final TC approval, to use a combination of internally-derived priors based on results from associated fishery strata for the first set of strata and a sequential-prior approach for the remaining priors (Addendum to Technical Document 13, "*Selection of a prior for mixed stock analysis*"). However, during the discussions leading to this decision, the AP requested a sensitivity analysis to examine the effect of different priors on the direction and magnitude of bias and magnitude of error in stock composition estimates. Here we provide the results from this sensitivity analysis.

Methods

Two test sets of chum salmon were used to test the sensitivity of estimates to the choice of priors, a set from coastal western Alaska and a set from Southern District, Alaska Peninsula. For each set, 400 fish were selected from populations assigned to the respective reporting group. These individuals were removed from the baseline and used as mixtures to test sensitivity to the choice of priors. Testing followed the methods used for the 100% proof tests outlined in Technical Document 5, "*Status of the SNP baseline for sockeye salmon*", except for the priors and sample sizes. The prior "sample size" was set to 1 fish.

[1] This document serves as a record of communication between the Alaska Department of Fish and Game Commercial Fisheries Division and the Western Alaska Salmon Stock Identification Program Technical Committee. As such, these documents serve diverse ad hoc information purposes and may contain basic, uninterpreted data. The contents of this document have not been subjected to review and should not be cited or distributed without the permission of the authors or the Commercial Fisheries Division.

In the first set of analyses, the first prior used was the regional uniform prior (described in Technical Document 13), where the prior for each region is set at the same weight (weight of each region equals 1 divided by the number of regions) and the priors within regions are distributed evenly across all the populations within that region (weight of each population equals weight of the region divided by the number of populations within that region). We then used sequential priors, first using the results from the regional uniform prior as the prior for the second analysis, and then using the results from this second analysis for the prior in the third analysis (Figure 1 and 2).

The second set of analyses used the uniform binary method described in the presentation to the AP/TC on September 21, 2011 (Addendum to Technical Document 13) followed by sequential priors. This uniform binary prior is used for the initial prior and is based on expert opinion. We used the expert opinion recommendations from the Department presented in the Addendum to Technical Document 13. In this prior, of the G total number of reporting groups, $G^{(IN)}$ groups are deemed likely to contribute to a mixture and are tagged "IN" while $G^{(OUT)}$ groups are deemed unlikely to contribute significantly and are tagged "OUT". The prior parameter value (α) assigned to the group proportions for each of these sets of reporting groups is: $\alpha_g = 0.01$ for $g \in G^{(OUT)}$ and $\alpha_{g'} = \frac{(1-0.01 \times G^{(OUT)})}{G^{(IN)}}$ for $g' \in G^{(IN)}$. For coastal western Alaska the reporting groups tagged "IN" were: Asia, Kotzebue, coastal western Alaska, and upper Yukon River. For the South Peninsula, all reporting groups were tagged "IN", so the analysis was identical to the regional uniform prior. We then used sequential priors, using the results from the uniform binary prior as the prior for the second analysis, and then using the results from this second analysis as the prior in the third analysis.

Results

Results are shown in Figures 1 and 2. Regional uniform priors provided the most downward discrepancy from the truth for correct allocation and upward discrepancy from the truth for incorrect allocations. Discrepancies from the truth were similar for priors based on the binary uniform (D) and the first sequential originating with the regional uniform (B). The smallest discrepancies from the truth were for priors based on the second sequential originating with the regional uniform (C) or the first and second sequential priors originating with the binary uniform (E and F). Misallocations were more pronounced in the coastal western Alaska tests (Figure 1) than in the Southern Alaska Peninsula tests (Figure 2). In the South Peninsula tests, the results from the two test sets were identical because the uniform regional prior and the binary uniform prior methods provided identical weights to all the reporting groups, simply because all stocks were deemed possibly present by the AP in this fishery.

Discussion

These results support expending effort to develop appropriate priors. Informative priors provide the largest relative decrease in misallocations to reporting groups that are not represented in the mixture. For example, in coastal western Alaska, misallocation to the Northern District, Alaska Peninsula decreased from 2.96% to 0.17%, a 94% relative decrease when an informative prior was used. The higher misallocation was obtained using the regional uniform prior, while the lower value was obtained using methods analogous to those proposed for WASSIP (Addendum to Technical Document 13), where information from other associated strata are used to inform the prior.

The effect of the prior through the sequential-prior process was quickly lost. In the coastal western Alaska tests, the effect of the initial prior was lost after the second sequential analysis (discrepancy < 0.1%), whereas for the southern Alaska Peninsula tests, the effects were gone after the first sequential analysis (discrepancy < 0.1%). By using the approach outlined in the Addendum to Technical Document 13, we anticipate that the effect of the initial priors would be minimal in the initial strata and lost after the first sequential analysis because the associated fishery mixture estimates are likely to be more similar to the mixture under analysis than the binary uniform prior.

The consistent misallocations to the upper Yukon River in tests of coastal western Alaska (Figure 1) were likely due to the artifact that one of the collections in the baseline (Jim River) was mis-assigned to the coastal western Alaska reporting group and should have been assigned to the upper Yukon River group. This population is genetically similar to other upper Yukon

River collections and was the farthest upstream collection in the Yukon River assigned to the coastal western Alaska reporting group. This population has been reassigned into the upper Yukon River reporting group for future analyses, so these apparent misallocations should become smaller if the analysis were to be repeated using the revised reporting groups.

The methods outlined in the Addendum to Technical Document 13 should produce priors that substantially reduce discrepancies from the truth when allocating mixtures to reporting groups for chum salmon compared with the regional uniform prior. We anticipate an improvement for sockeye salmon as well, although it may not be as pronounced because sockeye salmon have deeper genetic differentiation than chum salmon. On the other hand, we have more reporting groups for sockeye salmon and small misallocations to many reporting groups will add to significant numbers of misallocated fish. Therefore, it seems prudent to invest in methods to minimize biases and errors by incorporating the most appropriate prior information.

Questions for Technical Committee

1) Given that this analysis was not designed to provide a comprehensive examination of the sensitivity to priors of estimating of stock composition estimates, do these methods and results provide enough information to conclude that an informative prior is better than an uninformative prior?

2) Are these methods appropriate to test the hypothesis that stock composition estimates are sensitive to the prior, at least in some cases?

Figures

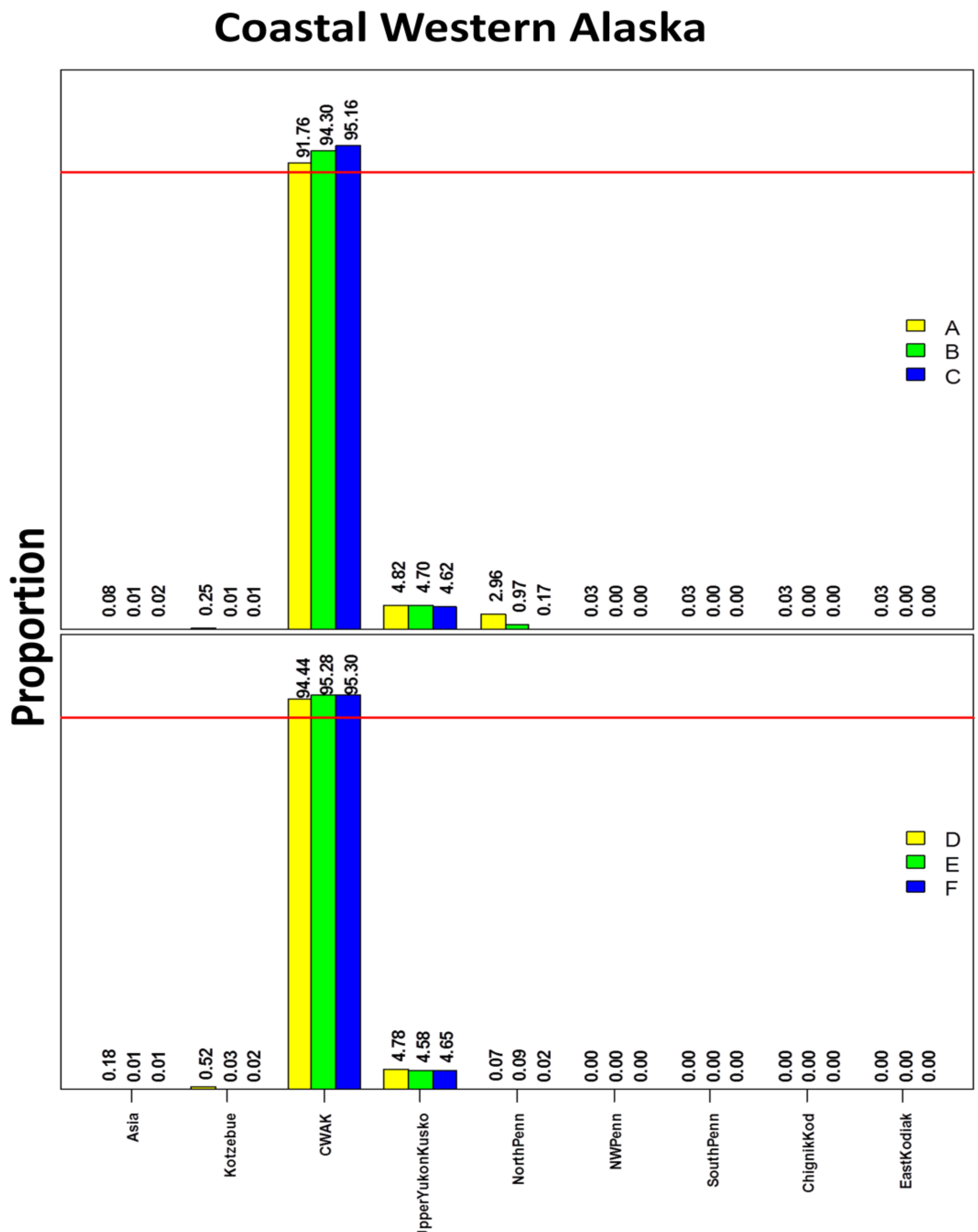

Figure 1. Sensitivity analysis of priors for a mixture of chum salmon from baseline populations of the coastal western Alaska reporting group: A) regional-level prior; B and C) sequential priors following the regional-level prior; D) uniform binary prior; E and F) sequential priors following the uniform binary prior. The red horizontal line is at 90%.

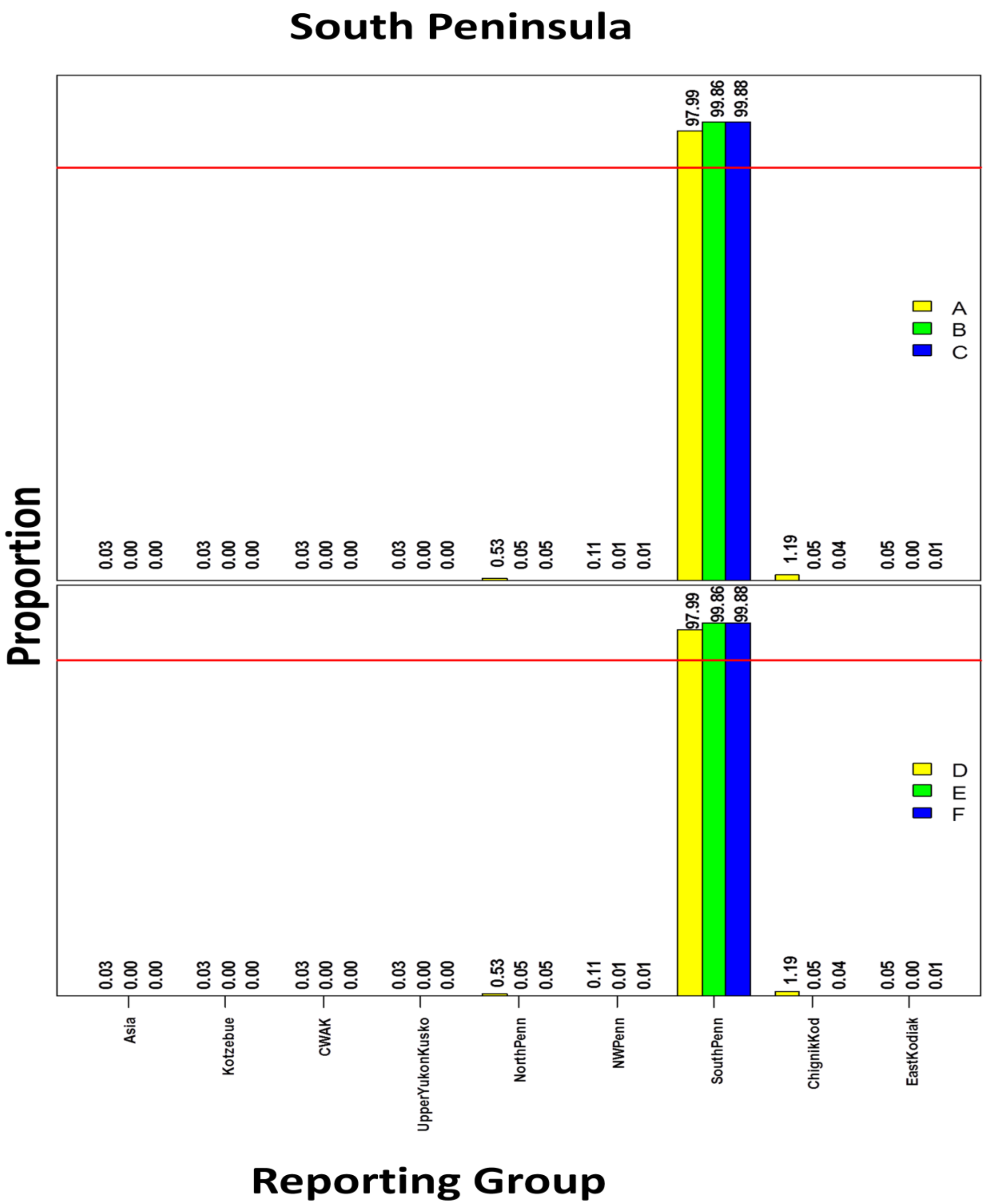

Figure 2. Sensitivity analysis of priors for a mixture of chum salmon from baseline populations of the Southern Alaska Peninsula reporting group: A) regional-level prior; B and C) sequential priors following the regional-level prior; D) uniform binary prior; E and F) sequential priors following the uniform binary prior. The red horizontal line is at 90%.