# Results

*Exploratory models*

From our exploratory analysis of multiple modeling structures, we found that the average root mean square error (RMSE) for predictions from 2017 to 2021 was lowest for the random forest model, followed by the GLMM and GAMM models for both the juvenile and spawner life history stages (Fig. 1). The formulas for the models with the lowest RMSE in the exploratory analysis differed between the life-stages. For the models of juvenile density, the GAMM included a three dimensional spline for the spatio-temporal variability (i.e., s(UTM\_E\_km, UTM\_N\_km, yr)), while the GLMM included both an equilibrium spatial random effect describing the average deviates in juvenile densities across all years and spatio-temporal deviations from the equilibrium between years that were assumed to be i.i.d.. Using RMSE, the most predictive random forest model for juvenile densities included all of the environmental covariates as well as the spatial and temporal data, and the structure of the model included 11 variables randomly sampled as candidates at each split (i.e., the mtry argument in the R random forest function) and 1000 decision trees. From the exploratory analysis of spawner densities, the formula for the GAMM model with the lowest RMSE was identical to the juvenile model, while the GLMM model included only spatiotemporal random deviates and not the average random spatial deviates constant across all years. The random forest model for the spawner densities included data for year and not the UTM coordinates for the survey location: the best model also contained 11 variables and 200 trees.

*Temporal and spatial out-of-bag predictability*

We examined the predictive of ability of the three models for different temporal and spatial out of bag samples and compare models by, again, averaging the RMSE for years 2017 to 2021. Three patterns emerged: first, the random forest model continued to have the lowest RMSE for both the predicted spatial and temporal out-of-bag samples for the juvenile and spawner life history stages (Fig. 1); however, the random forest models had larger increases in RMSE relative to its respective exploratory model. Second, out-of-bag predictions from survey designs including “annual” and or “tri-annual” observations had lower RMSE compared to only index observations across all models and life stages. Third, predicting spatial out-of-bag samples based on reduced future survey effort resulted in lower RMSE compared to predicting one or two years into the future for all survey locations (i.e., maintaining some consistent sites each year appears to produce more accurate model predictions than keeping all sites but not sampling every year).

*Choosing the GLMM model to prioritize survey decisions an evaluate future environmental effects*

While the RMSE for the random forest model was lower relative to the GAMM or GLMM in both the exploratory and out-of-bag cross validation analyses, the random forest model with the lowest RMSE had 11 variables sampled at each split, which is quite high relative to the 13 total variables in the model. This results in trees being too independent and the model loses generalizability to unseen data. This can be observed by the larger relative increases in RMSE for out-of-bag sample estimates using the random forest model (Fig. 1). With lower overall and relative RMSE for GLMM compared to the GAMM, we chose the GLMM as the best model for characterizing the effects of the model covariates on Coho salmon densities from 1998 to 2021, and forecasting future densities in 2080 under RCP 8.5 climate scenarios.

*Covariate effects for the GLMM model*

We considered nine environmental covariates without interactions in each GLMM model and found little evidence of multicollinearity for the juvenile and spawner, with the maximum absolute correlation between any two variables equal to 0.44 and 0.41, respectively. For the juvenile Coho, stream slope (StrmSlope) had the largest negative effect on densities, while outmigration distance (OUT\_DIST) had the largest positive effect (Fig. 2). For the spawning Coho, stream slope (StrmSlope) and mean winter precipitation index (MWMT\_Index) had similarly large negative effects on densities, while solar intensity (Sol\_Mean) had a positive effect (Fig. 3). Within our model, only the winter precipitation (W3Dppt) and spring precipitation (SprPpt) can be forecast into the future under various climate scenarios. While the effect of both these variables was negative for juvenile densities and positive for spawner densities (Fig. 2 and 3), the effects were smaller relative to the other covariates.

*Estimated juvenile and spawner abundance indexes for survey locations*

At the population level, the estimated annual abundance of juvenile and spawning Coho in Oregon coastal tributaries was similar to the observations across all years (Supplemental Fig.1 and Supplemental Fig. 2). In general, the precision and accuracy of the juvenile abundance estimates were worse than the spawner abundance estimates, with the Beaver Creek, Sixes River, and Salmon River being the least accurate among the juvenile populations (Supplemental Fig. 1). Similarly, Beaver Creek and the Sixes River, produced consistent biases in the spawner estimates relative to the observations as well (Supplemental Fig. 2).

*Comparing the spatial distribution of juvenile and spawner densities in 2021 and 2080*

We used the RCP 8.5(?) model to forecast winter (W3Dppt) and spring (SprPpt) precipitation under future climate scenarios and held all other environmental covariates to their mean values to predict future juvenile and spawner abundance. *Need to get access to the google drive again.*

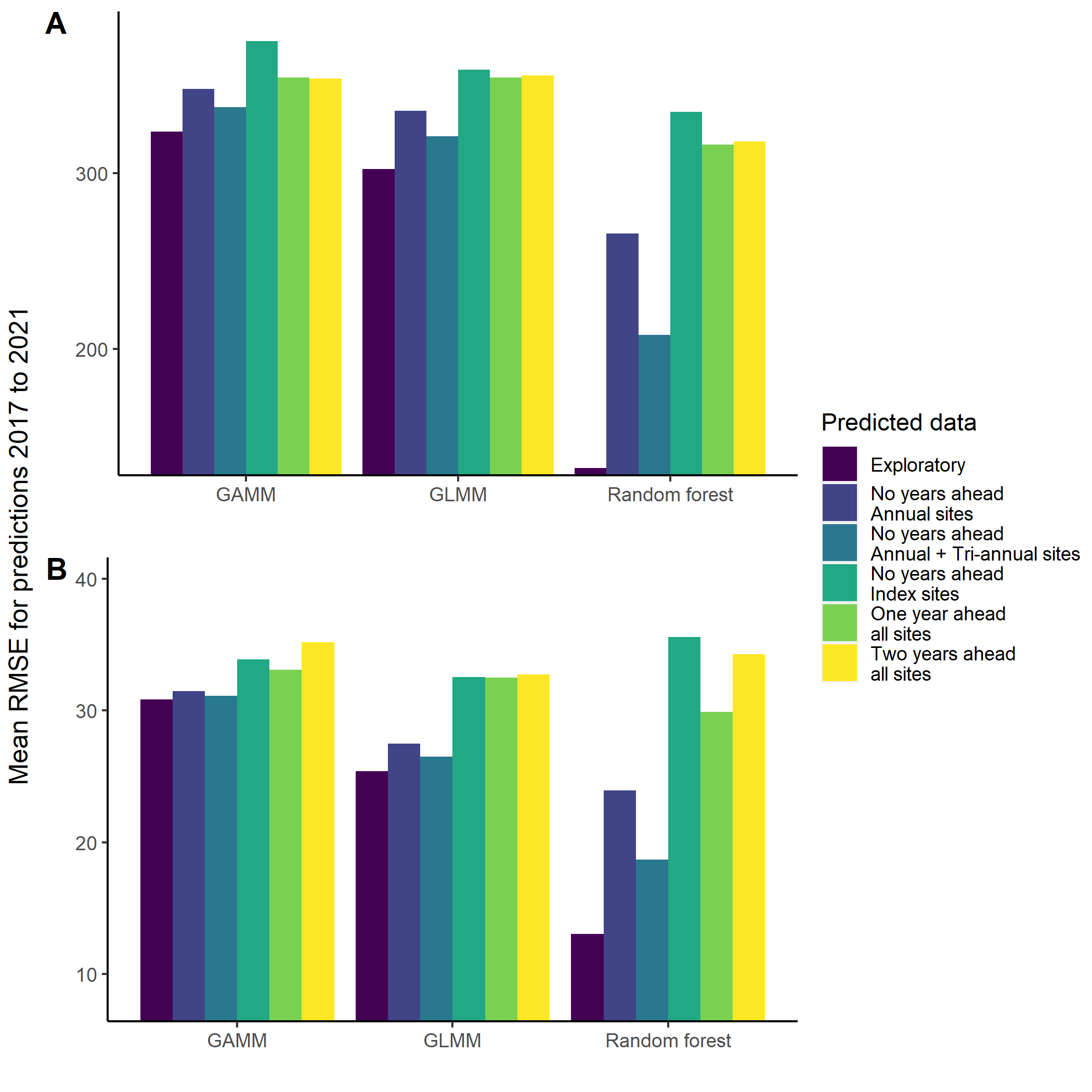


Figure 1. Average root mean square error (RMSE) from 2017 to 2021 for each model with the lowest RMSE in the exploratory analysis, and five out-of-bag samples of juvenile (panel A) and spawner (panel B) densities with different levels of survey effort. (ggplot\_predictive\_RMSE.r)

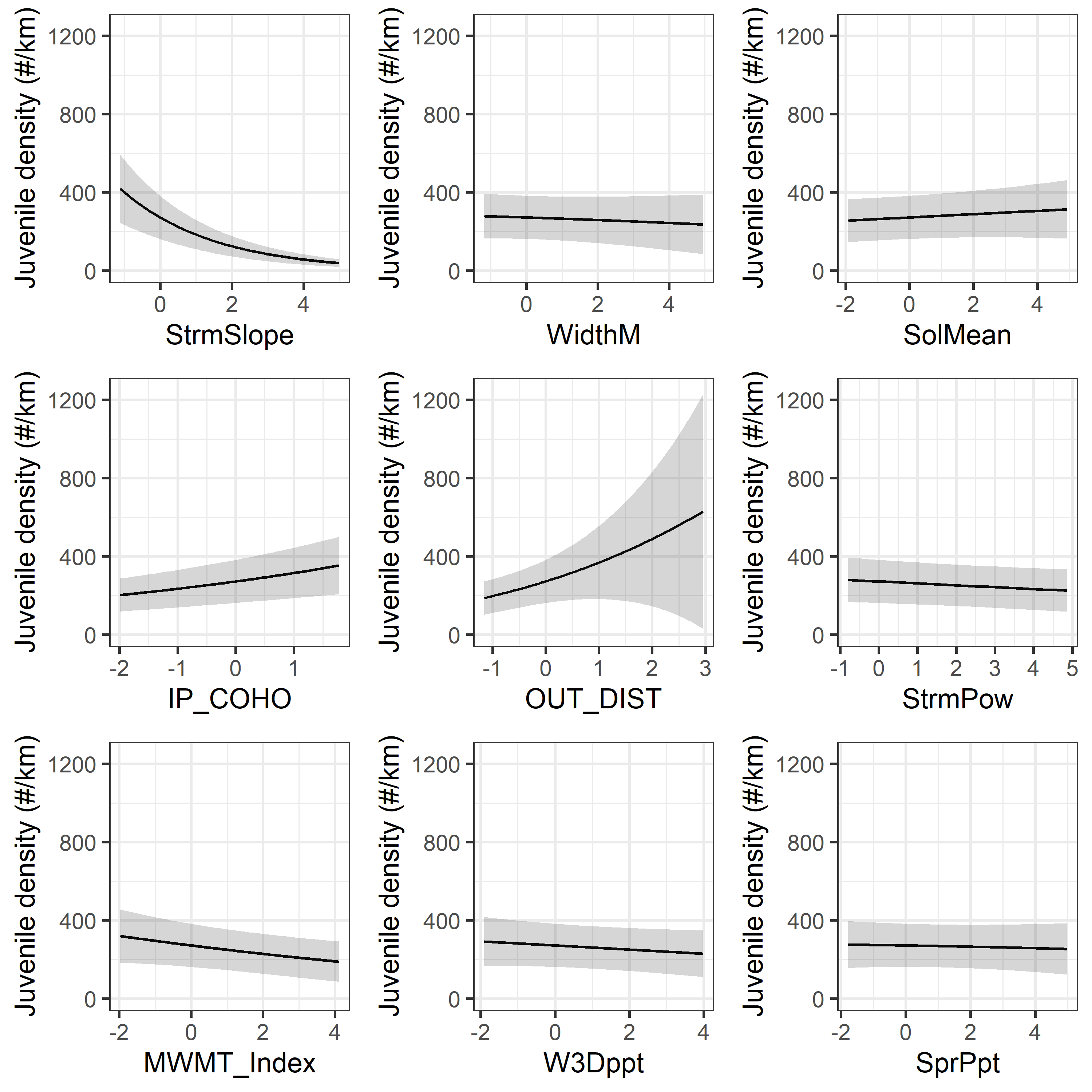


Figure 2. Marginal effect of the nine environmental covariates for the GLMM model predicting juvenile Coho salmon densities in coastal Oregon tributaries from 1998 to 2021. The range of the covariates have been Z-scored to standard the estimated effects. Shaded areas represent the 90% confidence intervals and the solid represents the mean effect. (ggplot\_marginalPlot.r)

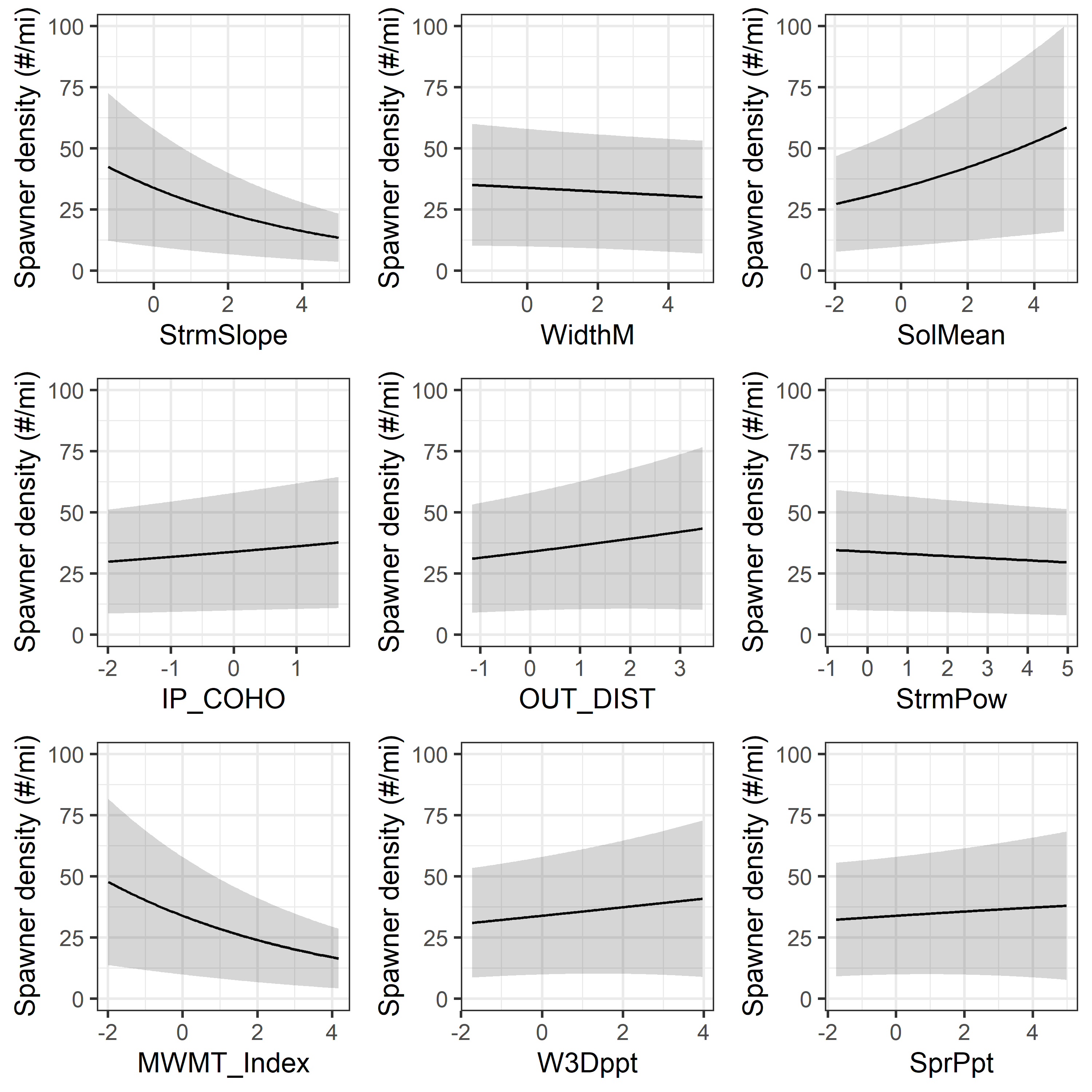
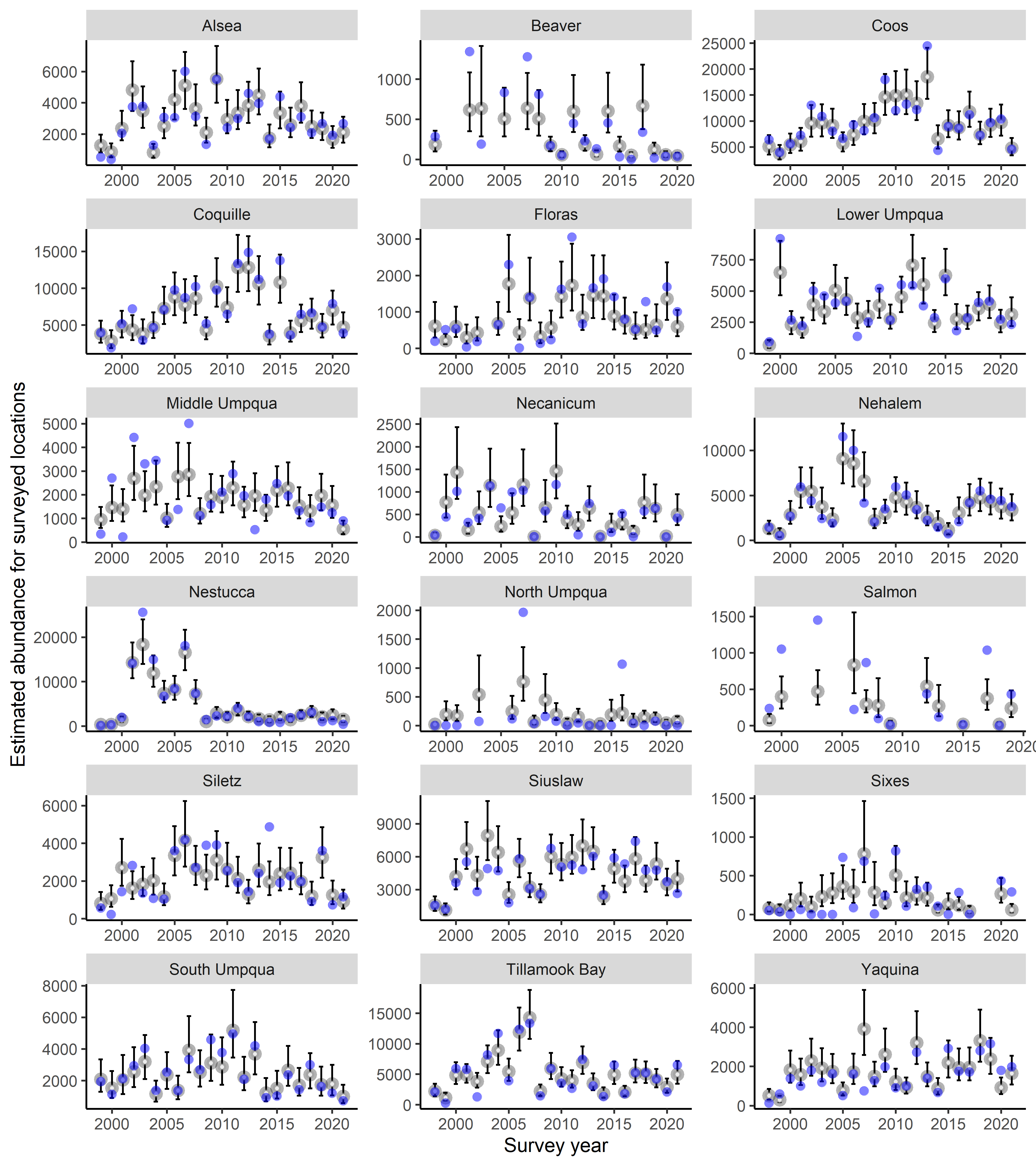
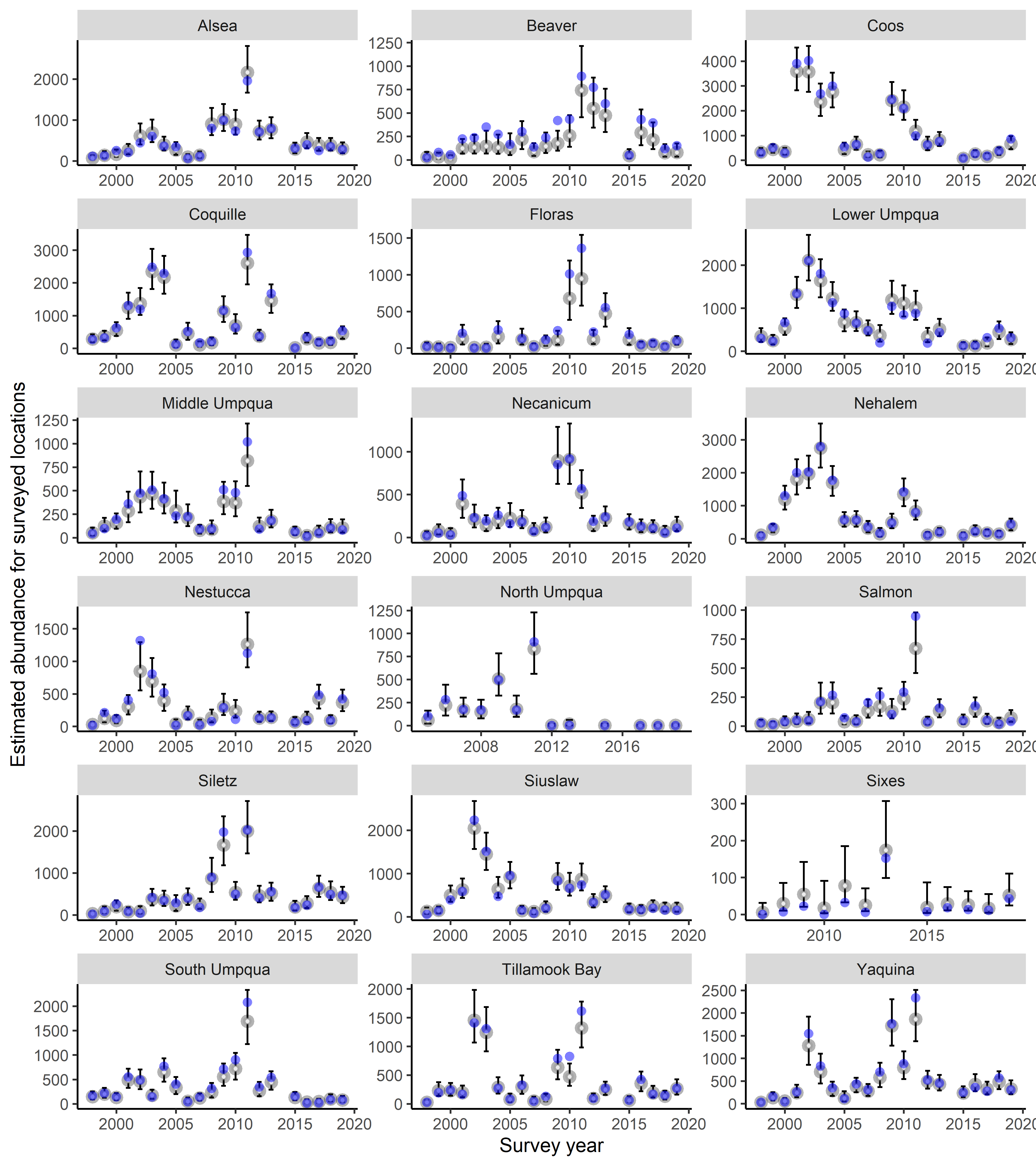


Figure 3. Marginal effect of the nine environmental covariates for the GLMM model predicting spawning Coho salmon densities in coastal Oregon tributaries from 1998 to 2021. The range of the covariates have been Z-scored to standard the estimated effects. Shaded areas represent the 90% confidence intervals and the solid represents the mean effect. (ggplot\_marginalPlot.r)



Supplemental Figure 1. Estimated (grey circles) with 90% confidence intervals (lines) and observed (blue circles) juvenile Coho abundance in 18 Oregon coastal tributaries from 1998 to 2021 based on survey locations in each year and assuming each survey location accounts for one kilometer of tributary.



Supplemental Figure 1. Estimated (grey circles) with 90% confidence intervals (lines) and observed (blue circles) Coho spawner abundance in 18 Oregon coastal tributaries from 1998 to 2021 based on survey locations in each year and assuming each survey location accounts for one mile of tributary.