

# PSP Forecast 2021 Report

Johnathan Evanilla

2/9/2022

## Read in predictions made during the 2021 season

```
predictions <- pspforecast::read_forecast()

summary(predictions)
```

```
##      version      location      date      name
## Length:533      Length:533      Min.   :2021-04-12      Length:533
## Class :character Class :character 1st Qu.:2021-05-17      Class :character
## Mode  :character Mode  :character Median :2021-06-02      Mode  :character
##                                     Mean  :2021-06-08
##                                     3rd Qu.:2021-06-27
##                                     Max.   :2021-08-23
##
##      lat      lon      class_bins      forecast_start_date
## Min.   :40.71 Min.   : -70.77      Length:533      Min.   :2021-04-18
## 1st Qu.:43.75 1st Qu.: -69.95      Class :character 1st Qu.:2021-05-21
## Median :43.82 Median : -69.74      Mode  :character Median :2021-06-06
## Mean   :43.99 Mean   : -69.21      Mean   :2021-06-12
## 3rd Qu.:44.24 3rd Qu.: -68.35      3rd Qu.:2021-07-01
## Max.   :44.97 Max.   : -66.98      Max.   :2021-08-27
## NA's   :3      NA's   :3
## forecast_end_date predicted_class prob_0      prob_1
## Min.   :2021-04-22 Min.   :0.0000 Min.   : 0.00 Min.   : 0.000
## 1st Qu.:2021-05-27 1st Qu.:0.0000 1st Qu.:88.00 1st Qu.: 2.000
## Median :2021-06-12 Median :0.0000 Median :97.00 Median : 3.000
## Mean   :2021-06-18 Mean   :0.0863 Mean   :87.72 Mean   : 8.002
## 3rd Qu.:2021-07-07 3rd Qu.:0.0000 3rd Qu.:98.00 3rd Qu.: 9.000
## Max.   :2021-09-02 Max.   :3.0000 Max.   :99.75 Max.   :56.000
##
##      prob_2      prob_3
## Min.   : 0.0000 Min.   : 0.00000
## 1st Qu.: 0.0000 1st Qu.: 0.00000
## Median : 0.0181 Median : 0.00038
## Mean   : 2.7384 Mean   : 1.50701
## 3rd Qu.: 2.0000 3rd Qu.: 0.74098
## Max.   :60.0000 Max.   :95.00000
##
```

## Correct Class 3 Predictions

```
closures <- pred_w_results %>%  
  filter(actual_class == 3)
```

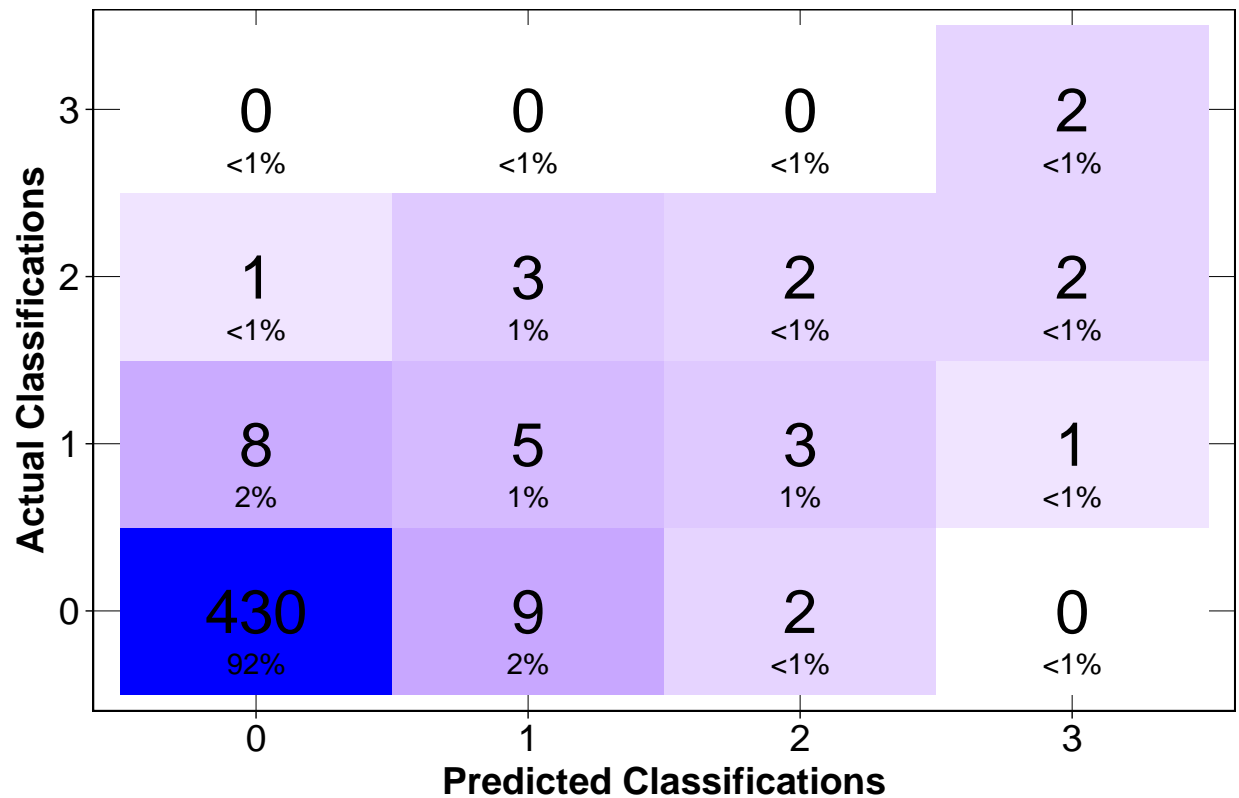
```
closures
```

```
## # A tibble: 2 x 17  
##   version location date      name      lat   lon class_bins forecast_start_d~  
##   <chr>   <chr>   <date>   <chr>   <dbl> <dbl> <chr>   <date>  
## 1 v0.1.3 PSP12.28 2021-05-03 Bear Isl~ 43.8 -69.9 0,10,30,80 2021-05-07  
## 2 v0.1.3 PSP12.28 2021-05-11 Bear Isl~ 43.8 -69.9 0,10,30,80 2021-05-15  
## # ... with 9 more variables: forecast_end_date <date>, predicted_class <dbl>,  
## #   prob_0 <dbl>, prob_1 <dbl>, prob_2 <dbl>, prob_3 <dbl>,  
## #   measurement_date <date>, toxicity <dbl>, actual_class <dbl>
```

## Class 3 Predictions (correct and wrong)

```
## # A tibble: 5 x 14  
##   version location date      name      forecast_start_date forecast_end_date  
##   <chr>   <chr>   <date>   <chr>   <date>   <date>  
## 1 v0.1.3 PSP12.28 2021-05-03 Bear Island 2021-05-07 2021-05-13  
## 2 v0.1.3 PSP12.28 2021-05-11 Bear Island 2021-05-15 2021-05-21  
## 3 v0.1.3 PSP12.28 2021-05-16 Bear Island 2021-05-20 2021-05-26  
## 4 v0.1.3 PSP12.15 2021-05-17 Gurnet      2021-05-21 2021-05-27  
## 5 v0.1.4 PSP12.15 2021-05-24 Gurnet      2021-05-28 2021-06-03  
## # ... with 8 more variables: predicted_class <dbl>, prob_0 <dbl>, prob_1 <dbl>,  
## #   prob_2 <dbl>, prob_3 <dbl>, measurement_date <date>, toxicity <dbl>,  
## #   actual_class <dbl>
```

## Overall Model Performance



2022-02-11

## Overall Accuracy

```
accuracy(pred_w_results, truth=as.factor(actual_class), estimate=as.factor(predicted_class))
```

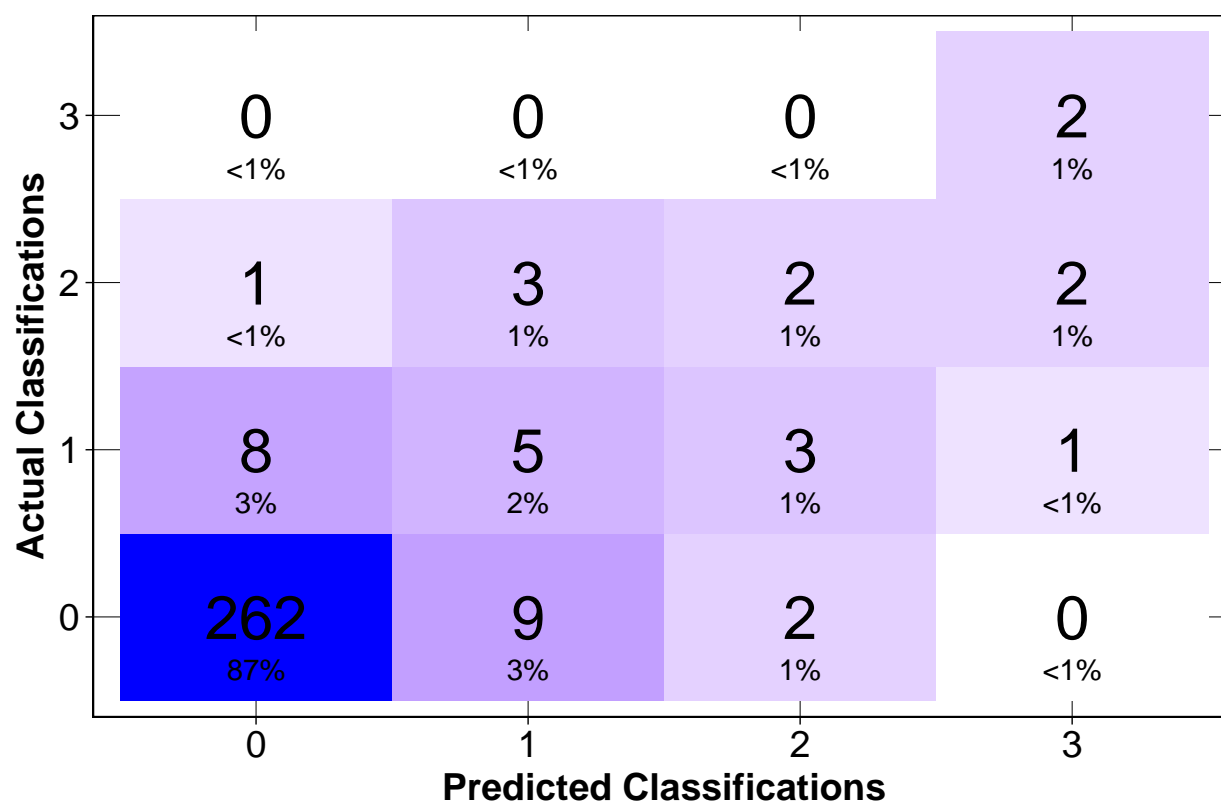
```
## # A tibble: 1 x 3
##   .metric .estimator .estimate
##   <chr>   <chr>      <dbl>
## 1 accuracy multiclass 0.938
```

## Precision

```
precision(pred_w_results, as.factor(actual_class), as.factor(predicted_class))
```

```
## # A tibble: 1 x 3
##   .metric .estimator .estimate
##   <chr>   <chr>      <dbl>
## 1 precision macro    0.490
```

## Western Maine Performance



2022-02-11

## Western Maine Accuracy

```
accuracy(western, truth=as.factor(actual_class), estimate=as.factor(predicted_class))
```

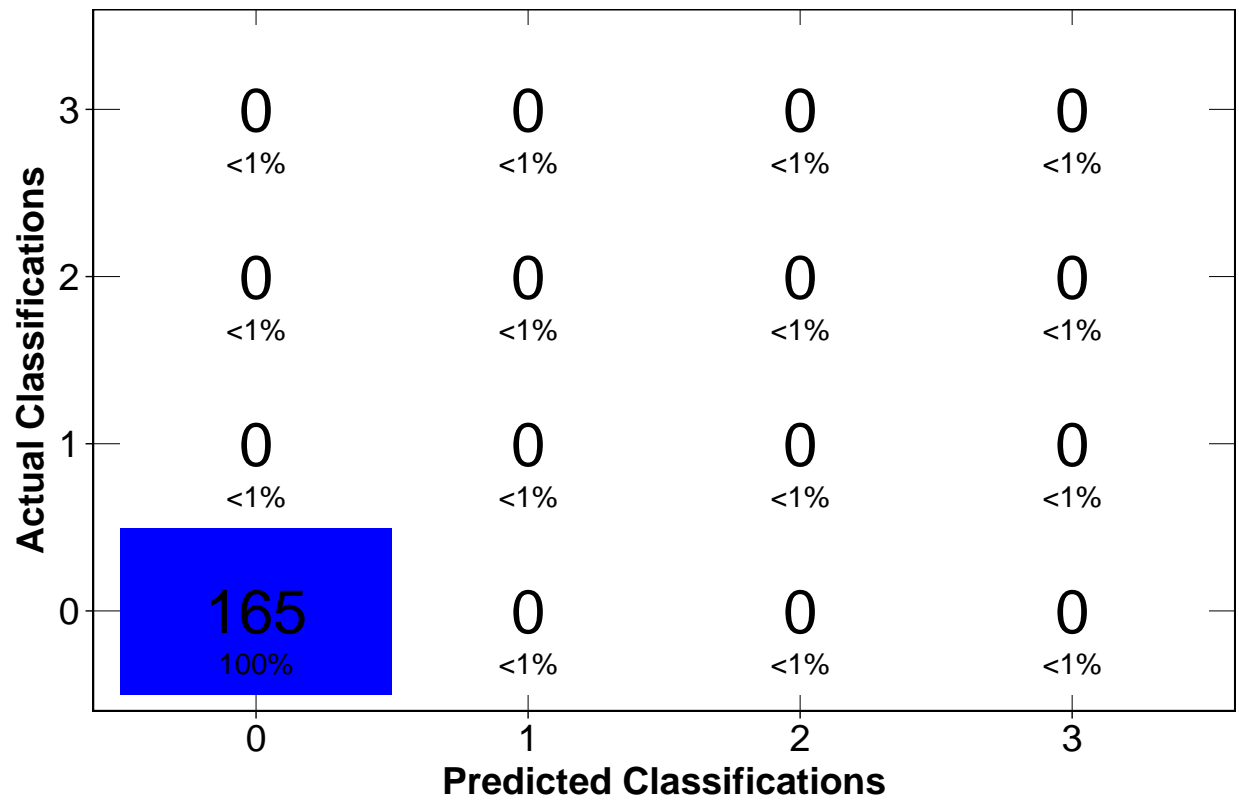
```
## # A tibble: 1 x 3
##   .metric .estimator .estimate
##   <chr>   <chr>      <dbl>
## 1 accuracy multiclass 0.903
```

## Precision

```
precision(western, as.factor(actual_class), as.factor(predicted_class))
```

```
## # A tibble: 1 x 3
##   .metric .estimator .estimate
##   <chr>   <chr>      <dbl>
## 1 precision macro    0.487
```

## Eastern Maine Performance



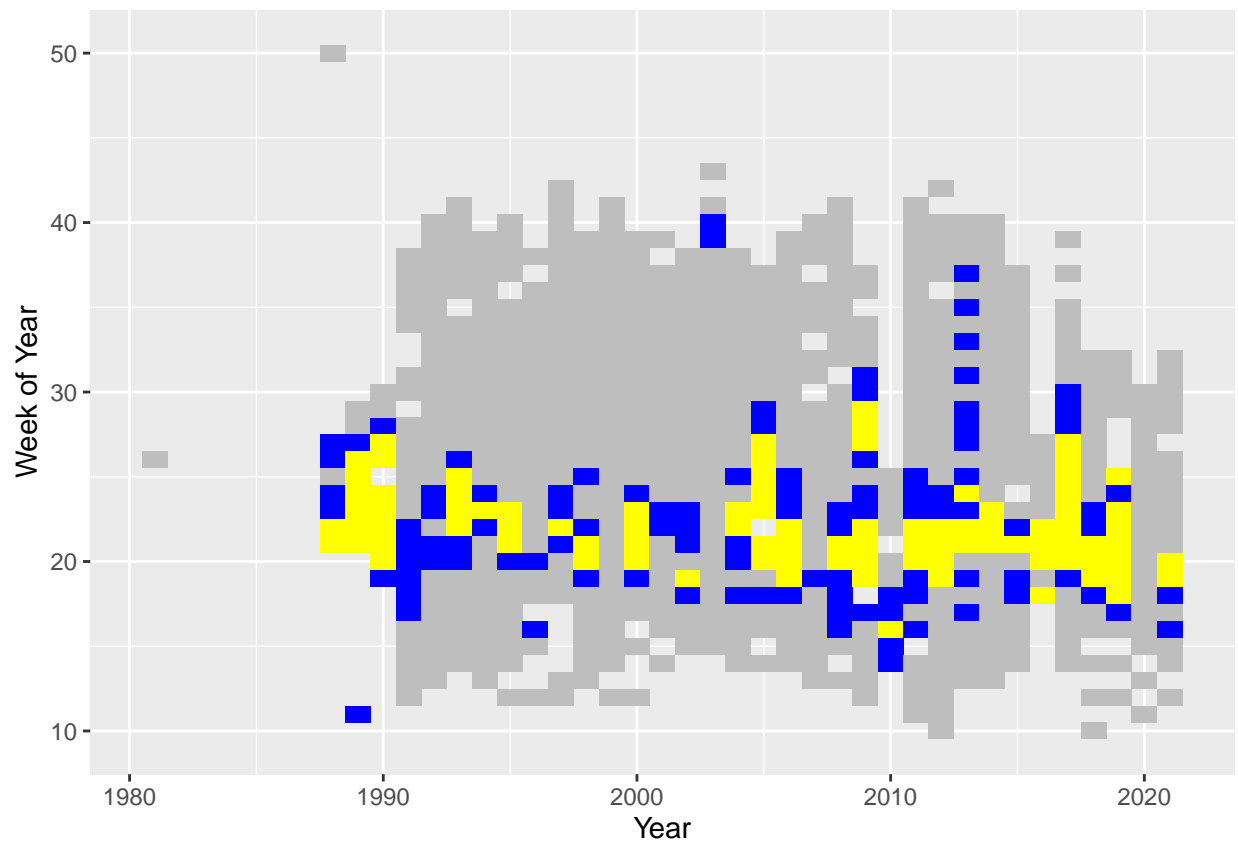
2022-02-11

## Model Versions

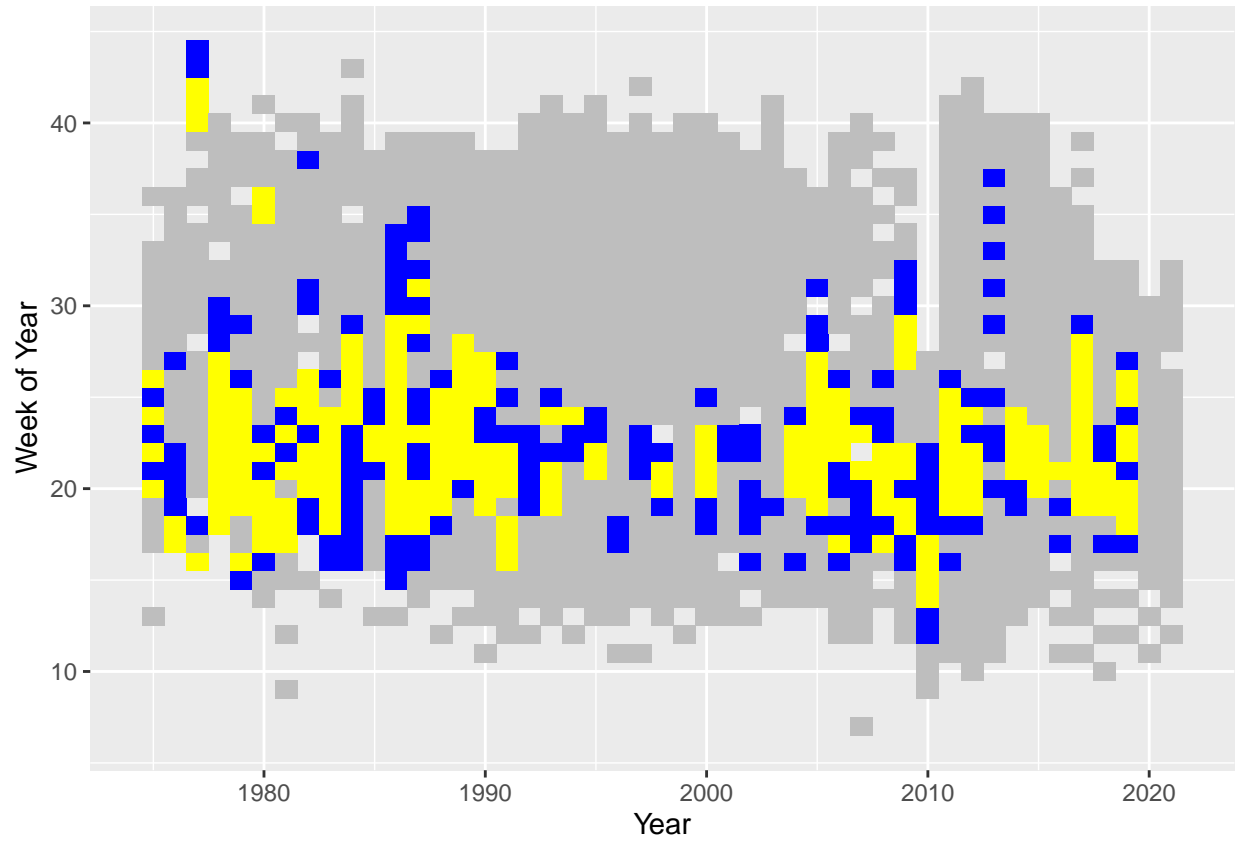
configuration	First Implemented	Description	Reason
v0.1.0	12 April	2 weeks of toxin measurements, 6-10 day step, all 12 toxins	Best performing model going into season.
v0.1.1	26 April	2 weeks of toxin measurements, 6-10 day step, all 12 toxins + sst_cum, doubled size of layers (16 -> 32 nodes)	Adding cumulative sst tuned into bloom seasonality.
v0.1.3	3 May	2 weeks of toxin measurements, 4-10 day step, all 12 toxins + sst_cum	Reduced minimum gap to 4 days in order to include all sites being sampled roughly weekly.
v0.1.4	25 May	2 weeks of toxin measurements, 4-10 day step, all 12 toxins + sst_cum, increased first dropout (0.3 -> 0.4), increased first layer size (32 -> 64 nodes), weighted classes	Weighting classes took away bias toward lowest class and balanced probability distribution.

## Toxicity at Sites over time

### Bear Island

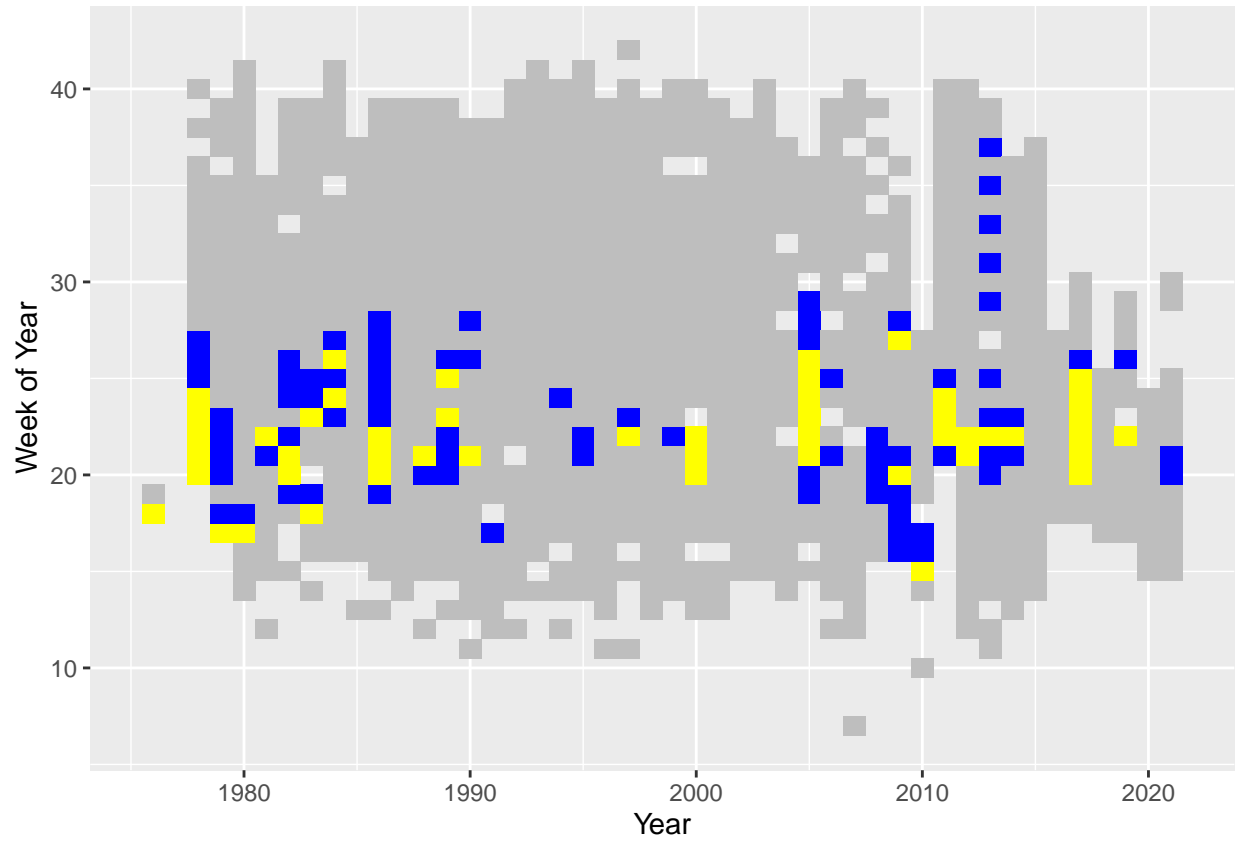


## Lumbo's Hole





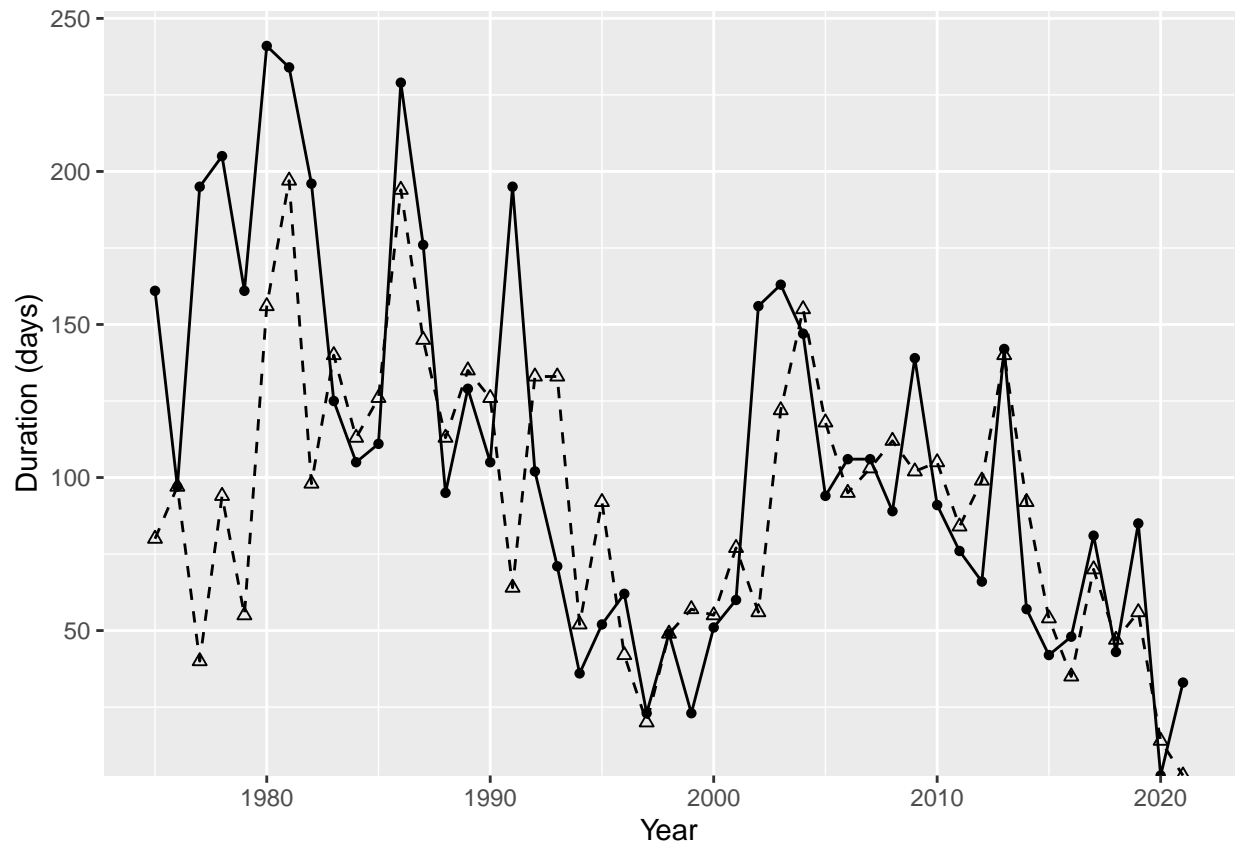
## Gurnet



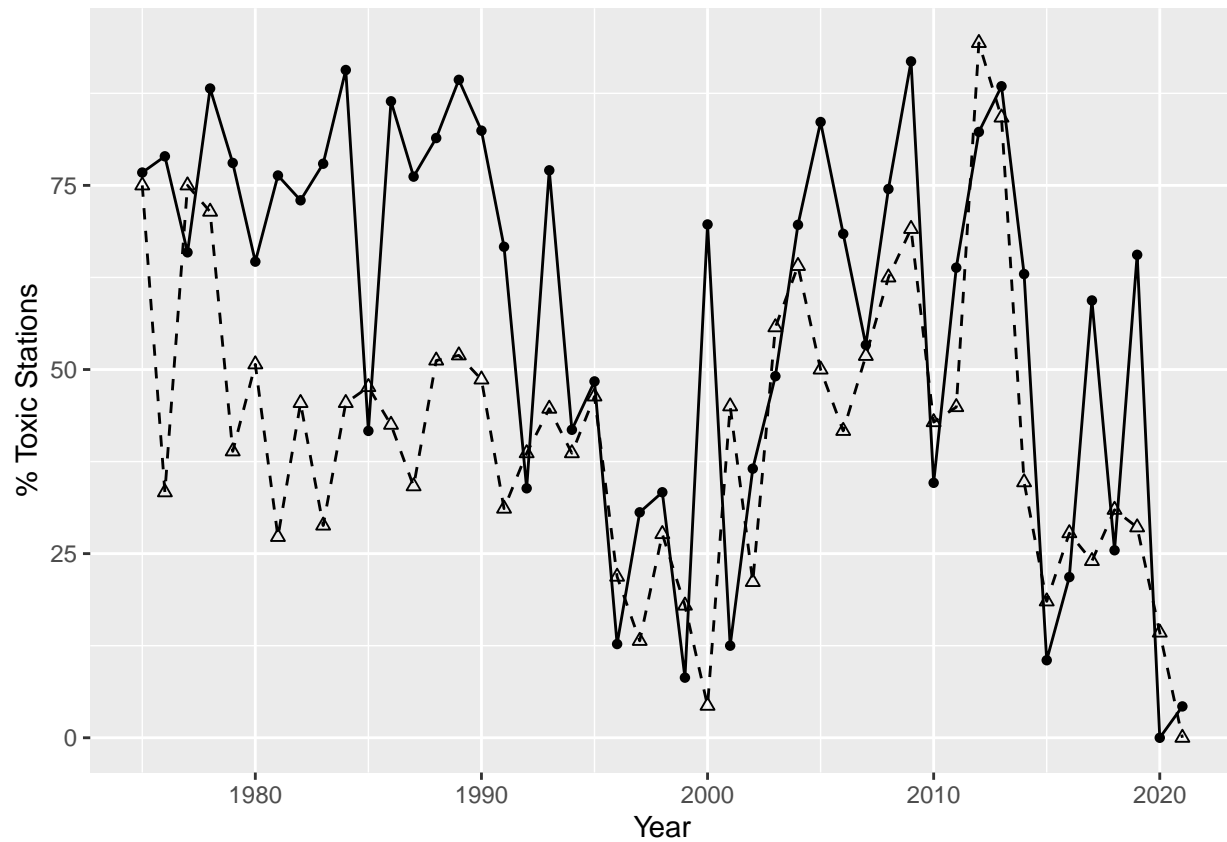
## HAB Index Metric Plots

### PSP Season Duration

## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.

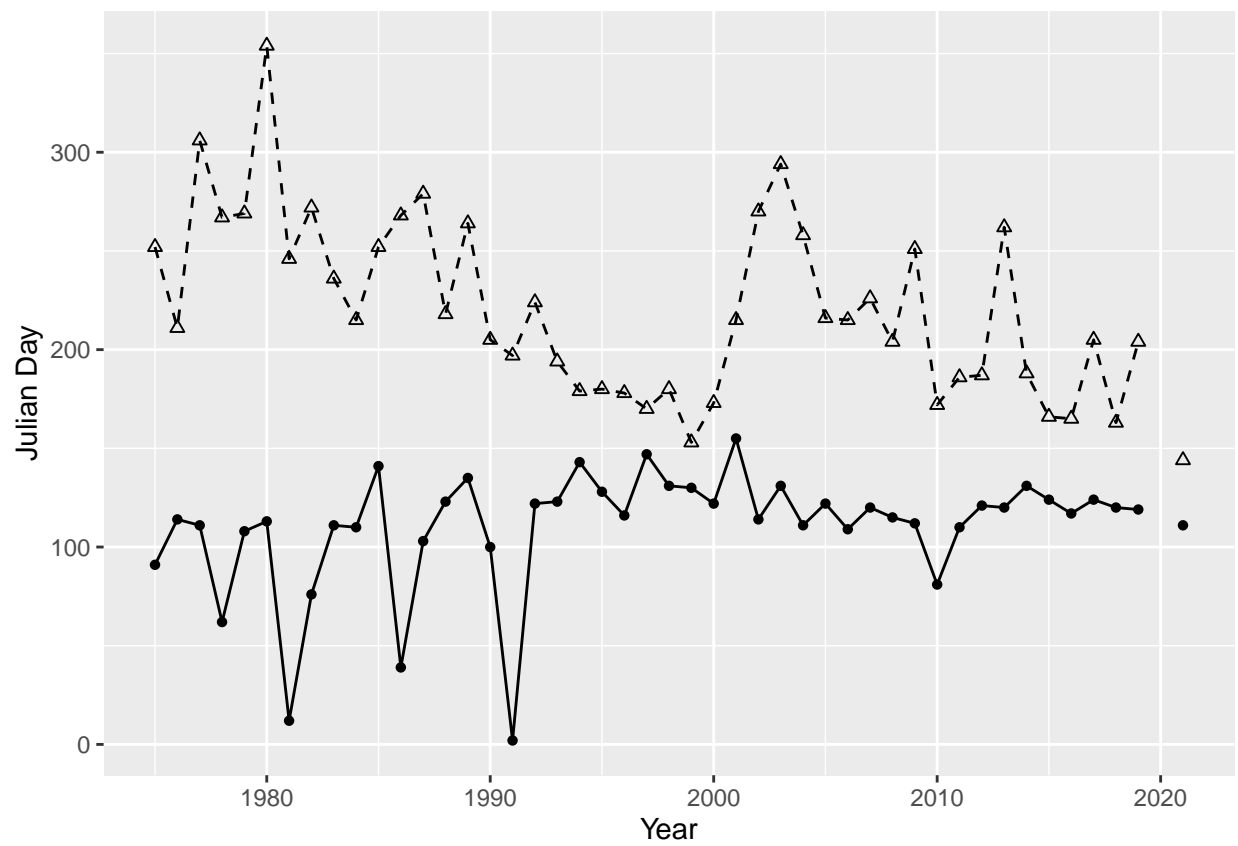


Percent stations with measurement >40



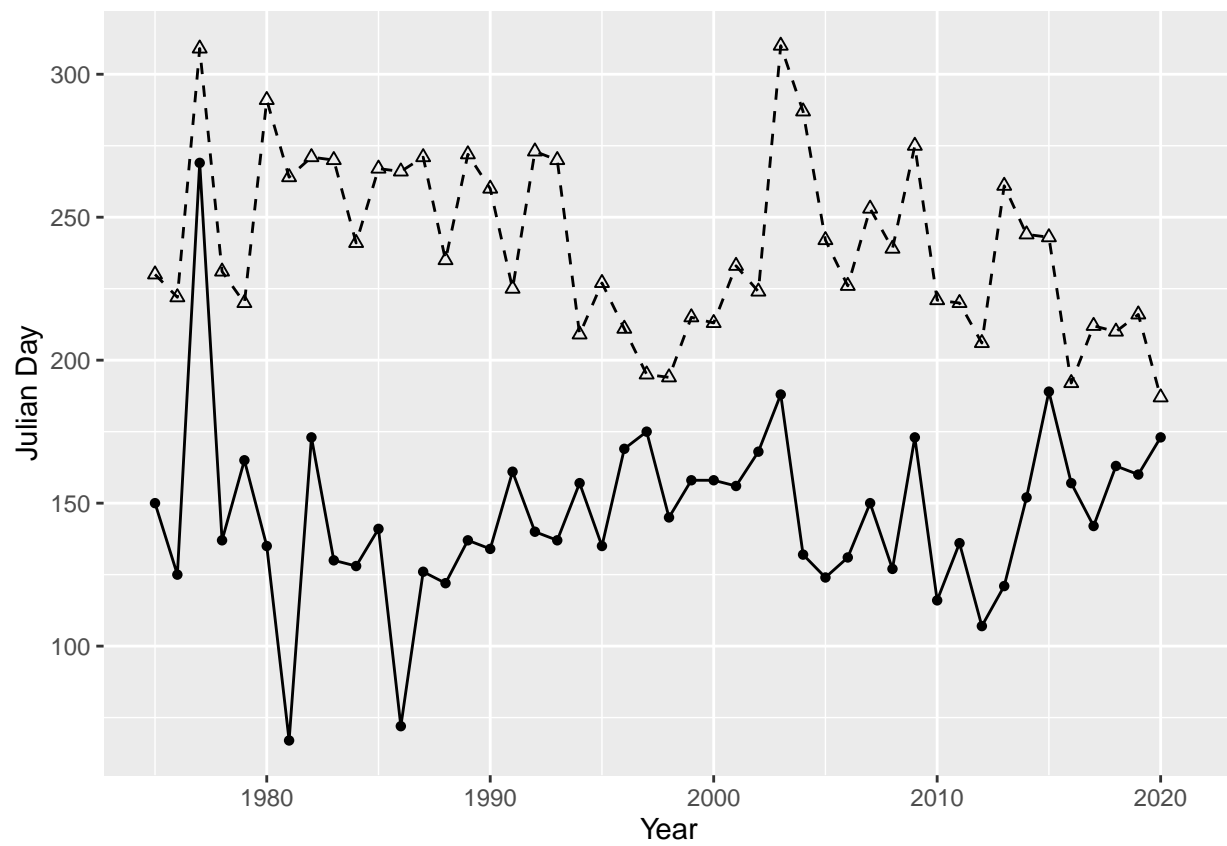
## Western Maine Season Onset and End

Western Maine saw no toxin measurements >40 during 2020.



## Eastern Maine Season Onset and End

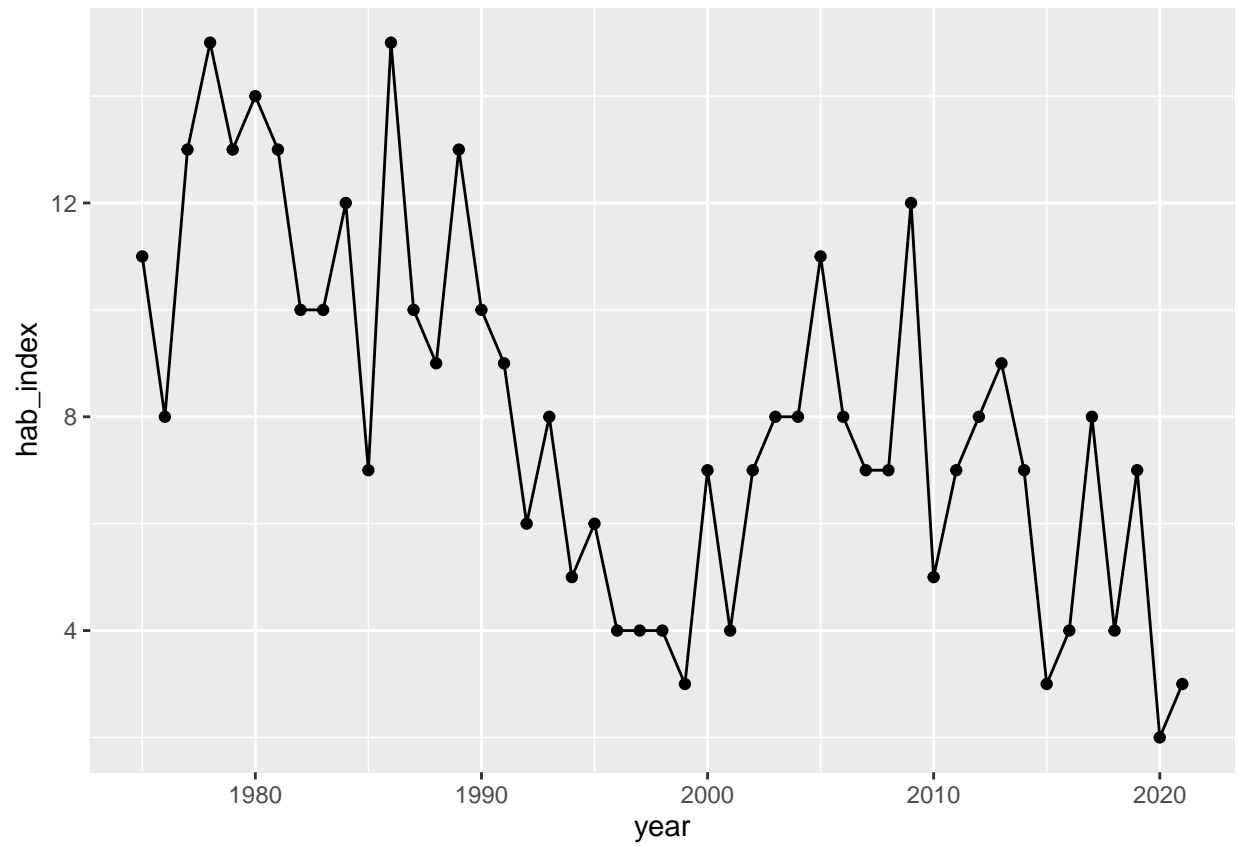
Eastern Maine saw no toxin measurements  $>40$  during 2021.



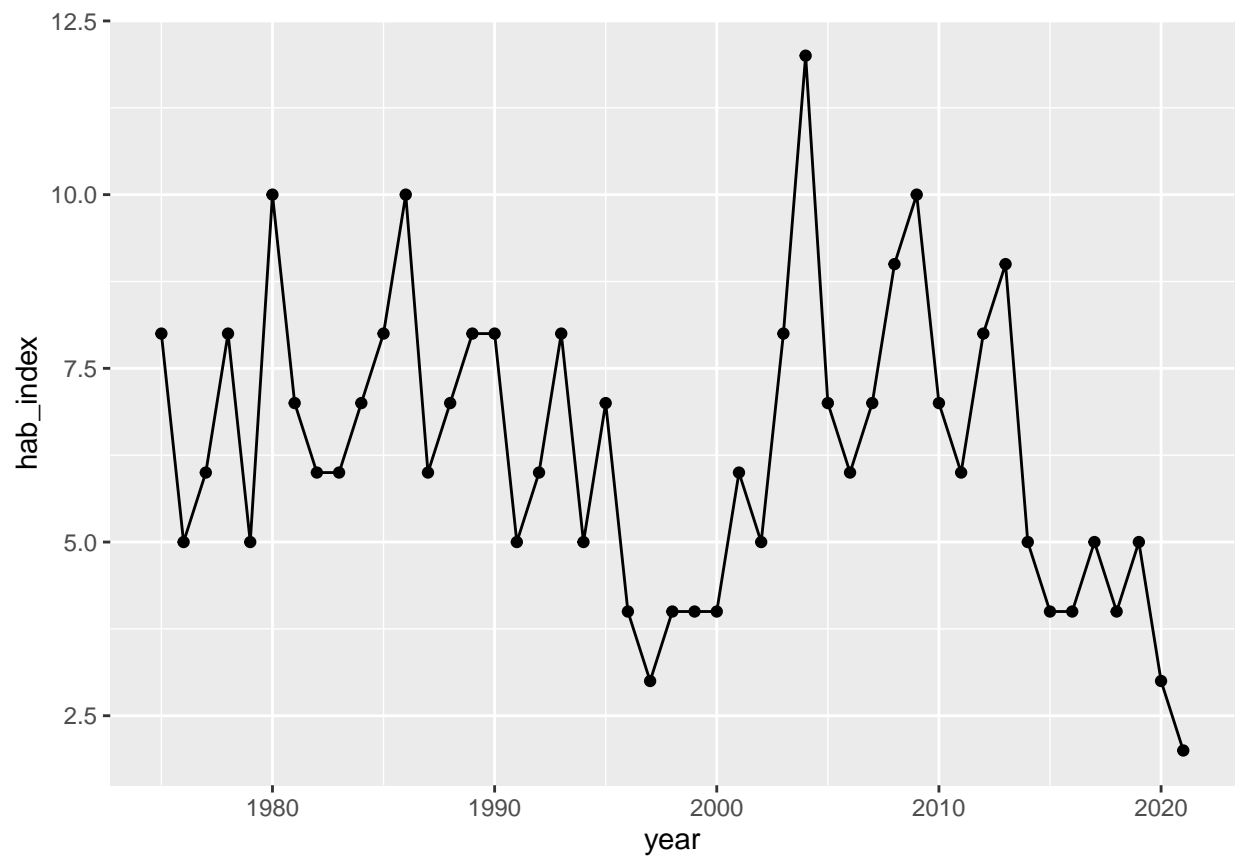
## HAB Index for each region

Calculated per methods found in Anderson, 2014

### Western Maine HAB Index



Eastern Maine HAB Index



```
nrow(predictions)
```

**Our forecast was able to make predictions for 71% of the toxin measurements made during the 2021 season**

```
## [1] 533
```

```
nrow(all_psp %>% filter(year==2021))
```

```
## [1] 741
```

```
nrow(predictions)/nrow(all_psp %>% filter(year==2021))*100
```

```
## [1] 71.92982
```