Erratum - Summary Analysis

Overview

Error

Several EPF pH values were rearranged among individuals between when the raw data was collected and the pH values were scaled for the final publication. This rearrangement only impacted a small number of samples and largely restricted to a single timepoint (Timepoint 1 - 24H). Below is a visualization of this error.

Impact on results and conclusions

The overall impact of the EPF pH correction is very mild and only had a slight impact on the visual patterns of EPF pH response through time, particularly at timepoint 1 (24H). After scaling the pH (total scale), we observed a slight increase in the mean EPF pH within the moderate OA treatment (900 uatm) relative to the control treatment, which resulted in some minor changes in the significant planned comparisons, specifically at timepoint 1. In particular, the comparison between the delta EPF pH of the control and moderate OA treatments went from not significant to significant (pvalue$<$0.05), while the comparison between the delta EPF pH in the moderate OA treatment is no longer significantly lower than zero. Despite some minor shifts in the significance of certain planned comparisons, these new results do not impact our previous conclusions or interpretations.

Importantly, errors that emerged from this rearrangement were restricted to timepoints and treatments that were not used for any downstream molecular analysis (timepoints 9 and 80 at treatments 400 and 2800 uatm) and did not feature in any analysis regarding the longterm trends between calcification and EPF pH.

Resolution

An updated version of the scaled EPF pH datasheet has been produced and is available in the `AE17_Cvirginica_MolecularResponse` Github repo (**URL**) with in a folder titled `/erratum`. In this folder revised versions of any related downstream datasheets using the revised EPF pH data has been generated and stored in the folder `/revised_data` and the revised version of the timeseries analysis (Figure 1) is located in `/revised_figures`.

Visualizing Correction

Packages and Data

Packages

```
library(dplyr,quietly = T,warn.conflicts = F)
library(reshape2,quietly = T)
library(ggplot2,quietly = T)
library(cowplot,quietly = T)
```

Data

```
# Raw EPF pH values
raw_epf <- read.csv("~/Github/AE17_Cvirginica_MolecularResponse/erratum/initial_data/AE17_EPF_raw.csv")
# Scaled EPF pH (total) from original published dataset
scaled_epf <- read.csv("~/Github/AE17_Cvirginica_MolecularResponse/data/Phenotype/AE17_EPFpHData_Experi
# Updated EPF pH (total) values
revised_scaled_epf <- read.csv("~/Github/AE17_Cvirginica_MolecularResponse/erratum/revised_data/AE17_EPI
```

Summary of differences between EPF pH in published datasheet compared to original datasheet

Below we highlight how EPF pH (unscaled) differs in our published data compared to the values on the original collection sheet and how our updated datasheet corrects this error.

Figure 1: EPF pH (measured EPF pH from published datasheet vs. collection sheet)

Here we can see that several points deviate from the 1:1 line, indicating that several EPF pH measurements were rearranged when the EPF pH data was being processed and scaled. Most points that deviate from this line correspond to Timepoint 1, which corresponds to oysters collected 24H after the start of the experimental exposure.

Note : Minor deviations from this 1:1 line reflect small rounding errors between the raw datasheet and the revised datasheet.

Datasheets

X-axis - Measured EPF pH (unscaled) from initial collection datasheet

`AE17_Cvirginica_MolecularResponse/erratum/initial_data/AE17_EPF_raw.csv`

Y-axis - Measured EPH pH (unscaled) from the processed datasheet used for downstream analysis

`AE17_Cvirginica_MolecularResponse/data/Phenotype/AE17_EPFpHData_ExperimentalExposureSamples_Simple.csv`

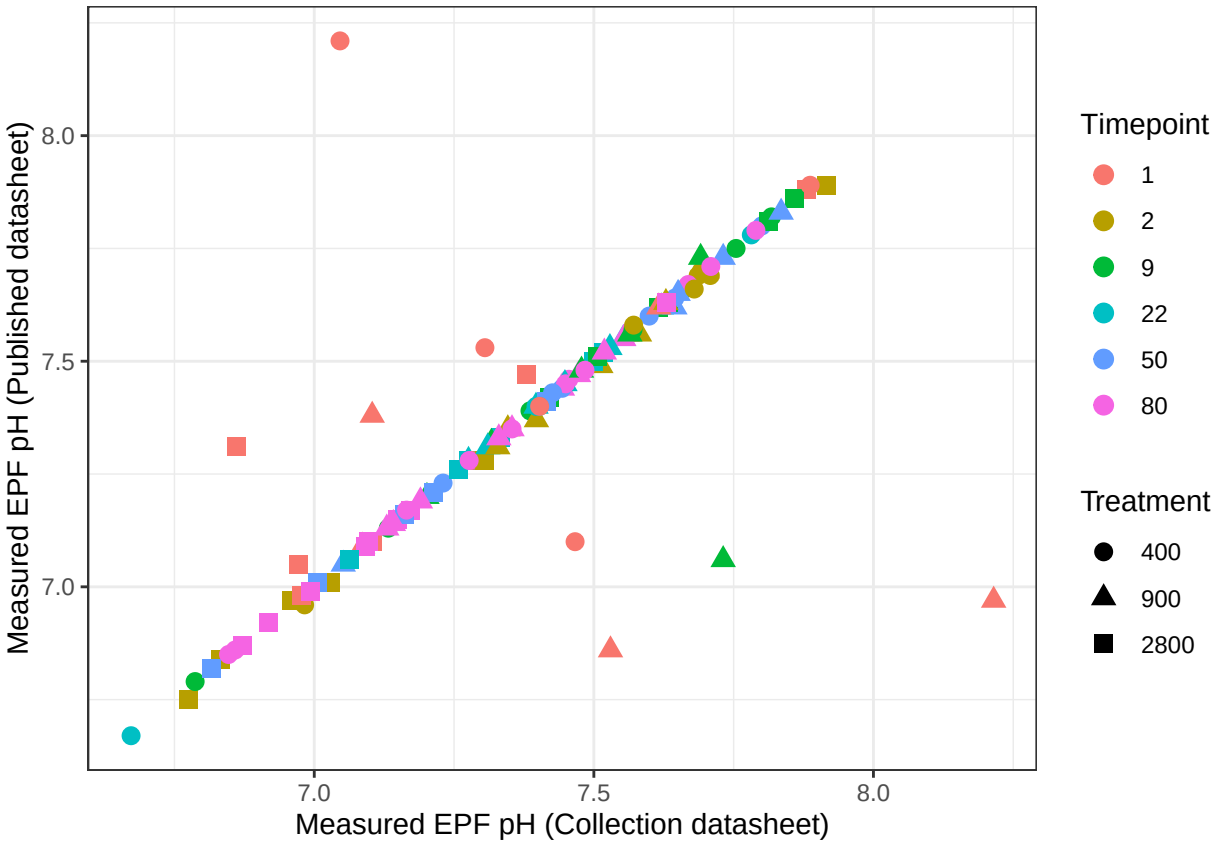


Figure 2: EPF pH (measured EPF pH from the revised published datasheet vs. collection sheet)

Here all points fall along 1:1 line indicating the measured pH in the revised EPF pH datasheet matches the measured EPF pH from the original datasheet.

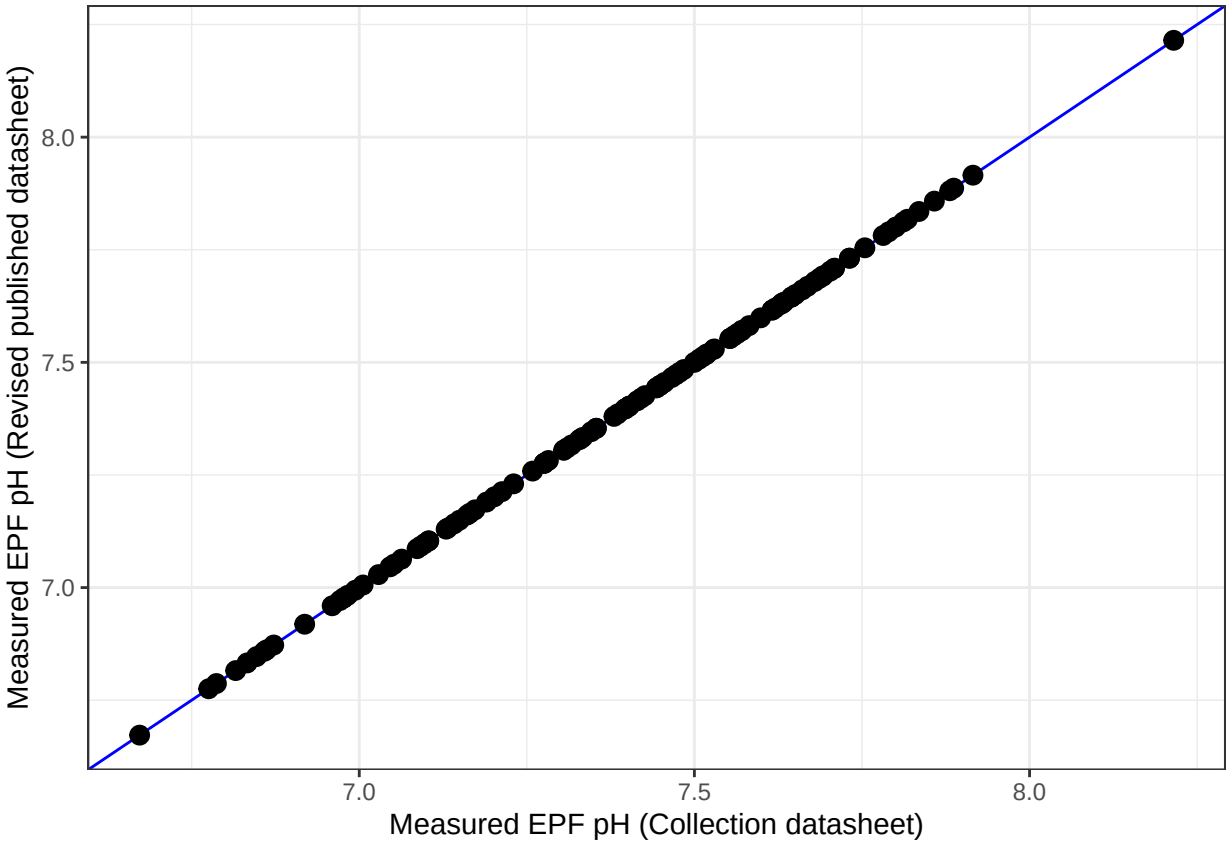
Datasheets

X-axis - Measured EPF pH (unscaled) from initial collection datasheet

`AE17_Cvirginica_MolecularResponse/erratum/initial_data/AE17_EPF_raw.csv`

Y-axis - Measured EPH pH (unscaled) from the revised, processed datasheet used for downstream analysis

`AE17_Cvirginica_MolecularResponse/erratum/revised_data/AE17_EPFpHData_revised.csv`



Impact of error on the timeseries analysis of EPF pH

In the revised version of the figure slight differences in timepoint 1 can be seen. This had a mild impact of significance tests, and primarily resulted in a slightly larger (and significant) difference between the control and moderate OA treatments at that timepoint.

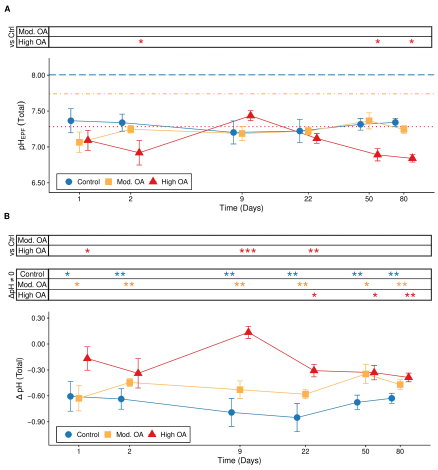


Figure 1: Original published figure 1

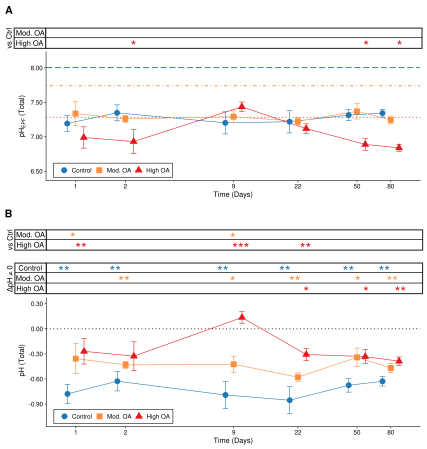


Figure 2: Revised figure 1