

## **Erratum for ‘Assessing present and future risk of water damage using building attributes, meteorology and topography’**

Claudio Heinrich-Mertsching

*Norwegian Computing Center, Oslo, Norway.*

E-mail: [claudio@nr.no](mailto:claudio@nr.no)

Jens Christian Wahl

*Norwegian Computing Center, Oslo, Norway.*

Alba Ordoñez

*Norwegian Computing Center, Oslo, Norway.*

Marita Stien

*Gjensidige Forsikring ASA, Oslo, Norway.*

John Elvsborg

*Gjensidige Forsikring ASA, Oslo, Norway.*

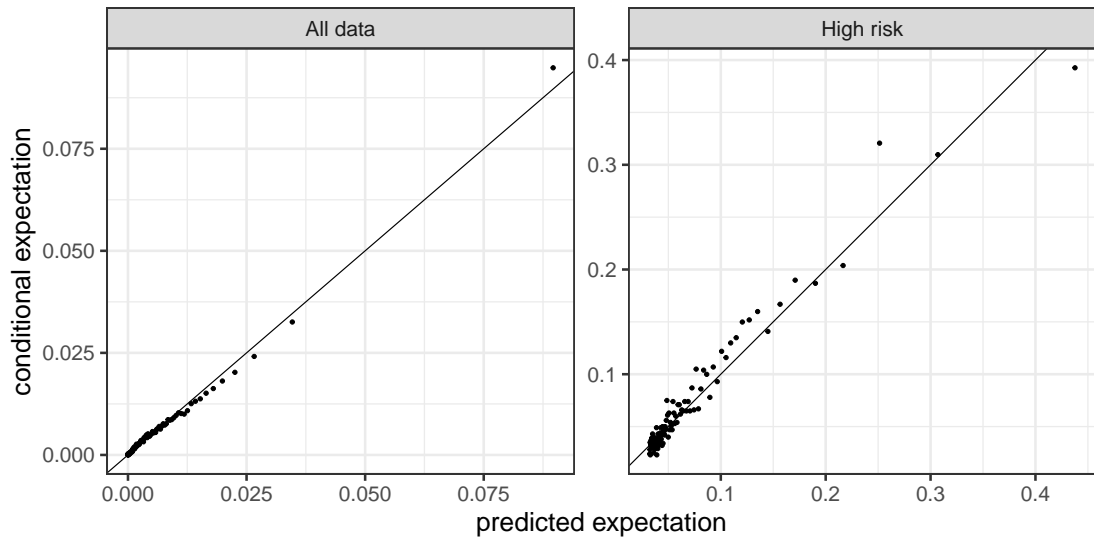
Ola Haug

*Norwegian Computing Center, Oslo, Norway.*

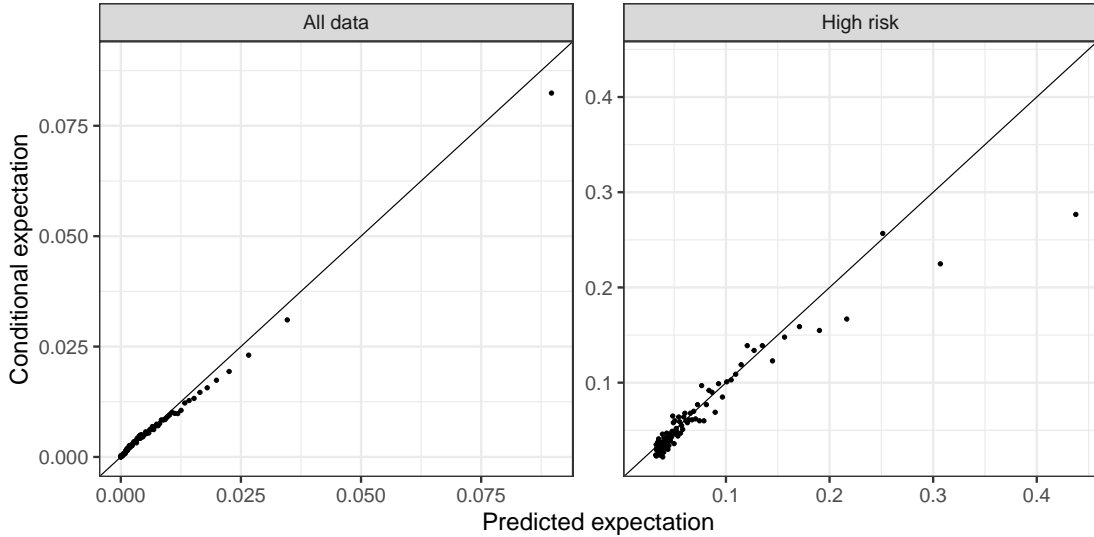
Thordis L. Thorarinsdottir

*Norwegian Computing Center, Oslo, Norway.*

The purpose of this note is to correct an erroneous figure contained in the original document, namely the conditional expectation diagrams (CEDs) shown in Figure 2. Due to an error in our code, the  $y$ -coordinate for each point showed the average probability of containing one or more claims for the category, rather than the expected number of claims for the category. This led to systematically too low  $y$ -values, the difference being caused by contracts with more than one observed claim.



**Fig. 1.** Corrected conditional expectation diagrams for the quarterly negative binomial model using both topographic and climatological covariates. The right hand side diagram is constructed from the 100,000 contracts (1.7% of the total data) with the highest predicted number of claims  $\mu_i$ .



**Fig. 2.** Incorrect conditional expectation diagram shown in the paper.

Figure 1 shows the corrected conditional expectation diagram, the erroneous version in

the preprint is shown in Figure 2, for comparison. The figures show that the error led to a systematic underestimation of the conditional expectation of high-risk customers, which accounts for most of the model bias for high-risk customers that was apparent in the previous diagrams. Consequently, our model shows overall very good calibration, even for high-risk customers. In particular, their risk is not over-estimated, as discussed in the article. It can be observed that the correction has the strongest effect on the rightmost points, both in the overall and the high-risk CED. Since the difference is caused by contracts with more than 1 claim, this indicates that most of these cases occur in the groups with high predicted risk. While this is quite intuitive, it is not obvious since the CEDs are constructed from the crossvalidation data and are, therefore, out-of-sample.