Can we predict the burden of wasting in crisis-affected countries? Findings from Somalia and South Sudan

ADDITIONAL FILE 1

Francesco Checchi1\*

Séverine Frison1

Abdihamid Warsame1

Kiross Tefera Abebe2, UNICEF South Sudan

Jasinta Achen3, UNICEF Somalia

Eric Alain Ategbo2, UNICEF South Sudan

Mohamed Ag Ayoya3, UNICEF Somalia

Ismail Kassim2, UNICEF South Sudan

Biram Ndiaye3 UNICEF Somalia

Mara Nyawo4 UNICEF ESARO

1 Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine

2 United Nations Children’s Fund, South Sudan Country Office

3 United Nations Children’s Fund, Somalia Country Office

4 United Nations Children’s Fund, East and Southern Africa Regional Office

\* Corresponding author: [Francesco.checchi@lshtm.ac.uk](mailto:Francesco.checchi@lshtm.ac.uk)

## Causal framework for wasting



Figure S5. Causal framework for wasting among children, used to identify potential predictors.

## Predictive accuracy of additional Somalia models

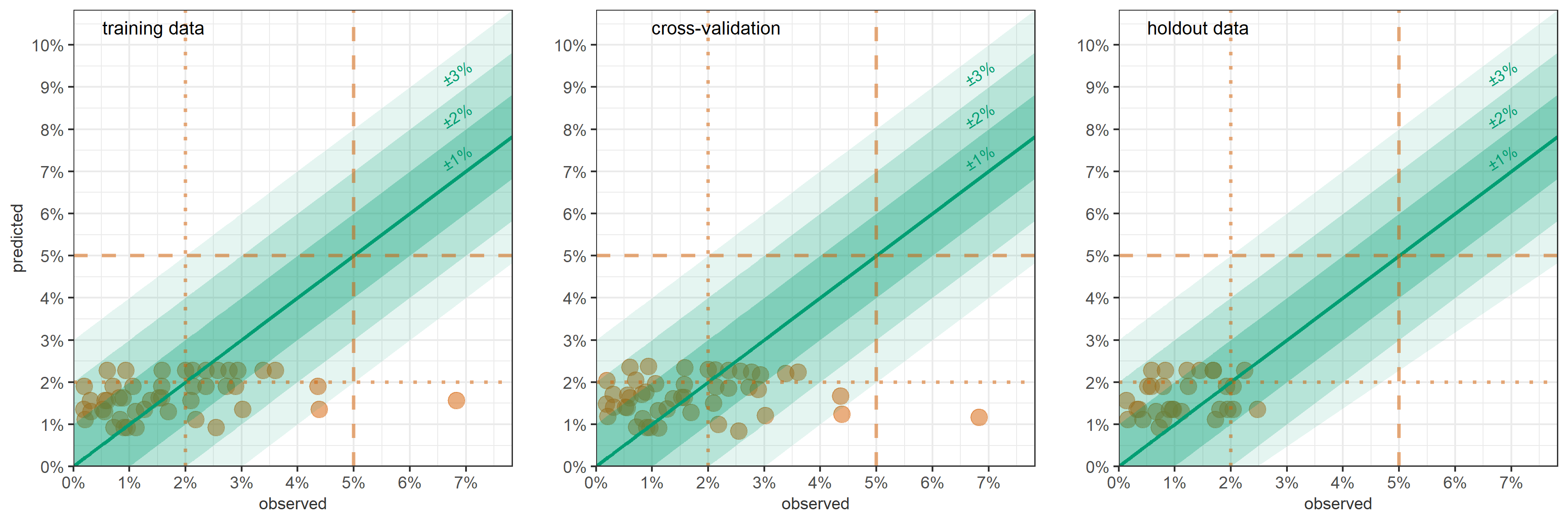


Figure S6. GLM-predicted versus observed severe wasting (MUAC + oedema) prevalence, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used severe wasting prevalence thresholds.

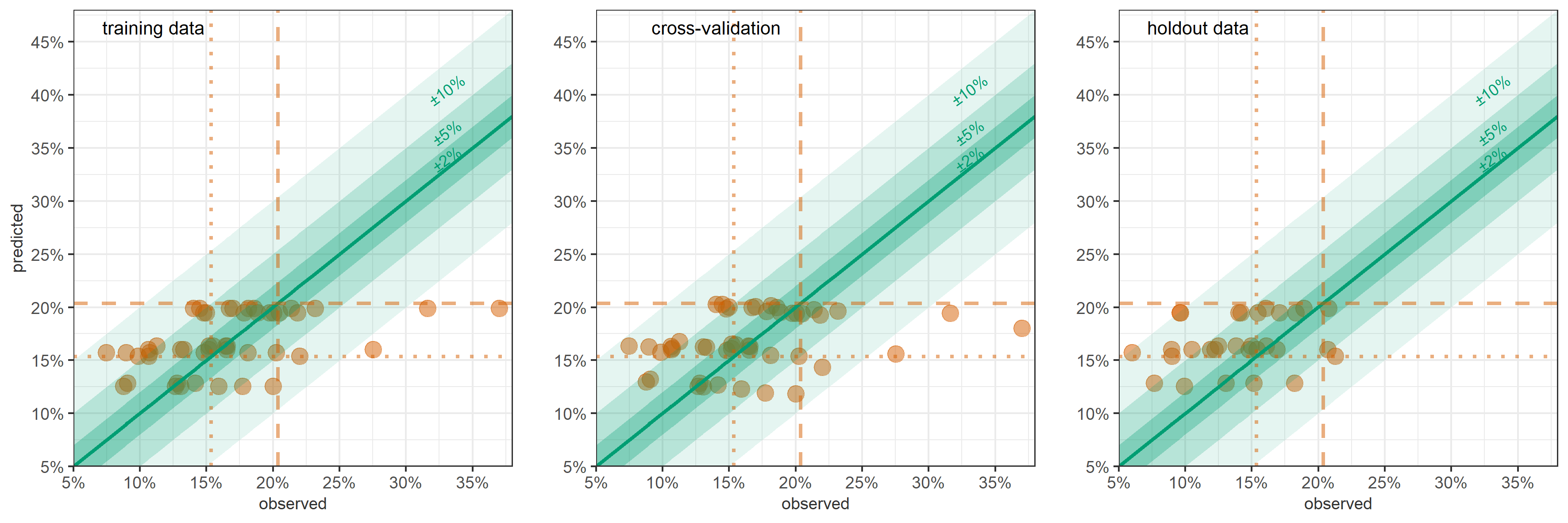


Figure S7. GLM-predicted versus observed wasting (WFH + oedema) prevalence, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

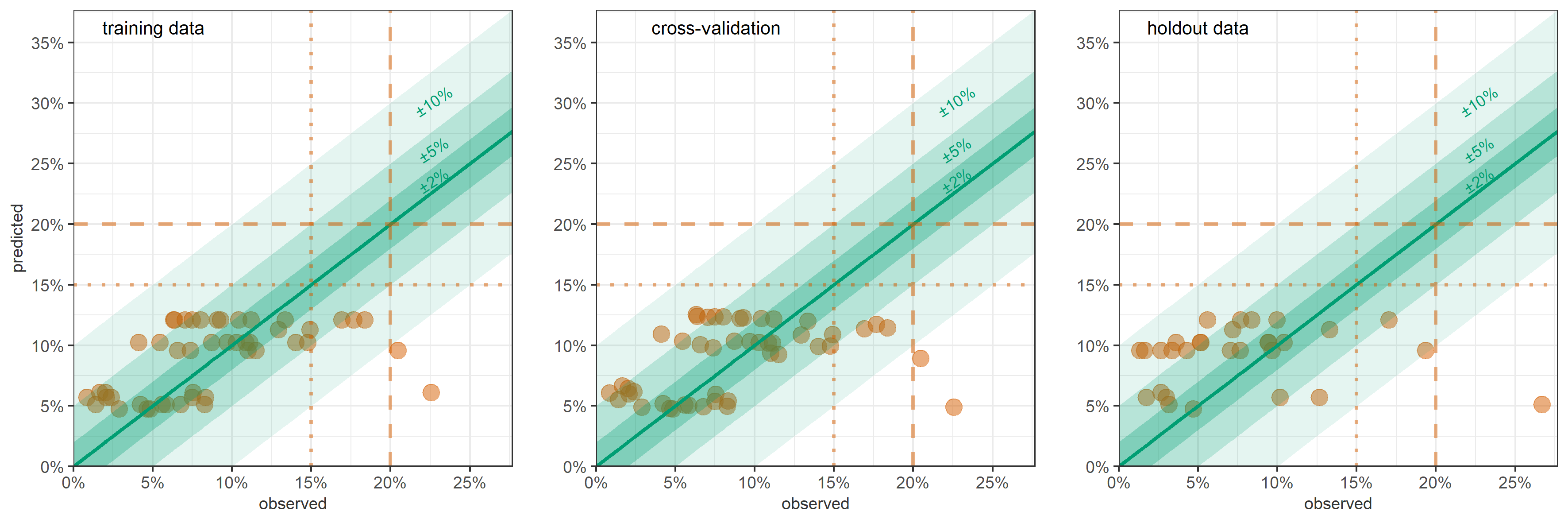


Figure S8. GLM-predicted versus observed wasting (MUAC + oedema) prevalence, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

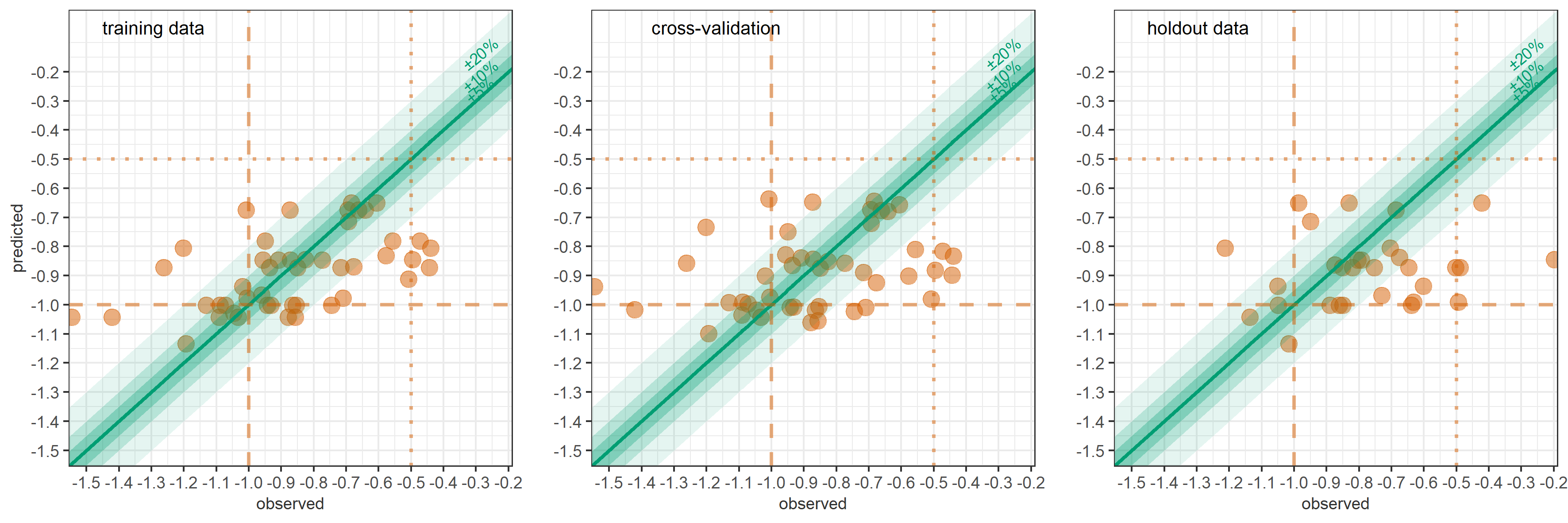


Figure S9. GLM-predicted versus observed mean WFH, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.

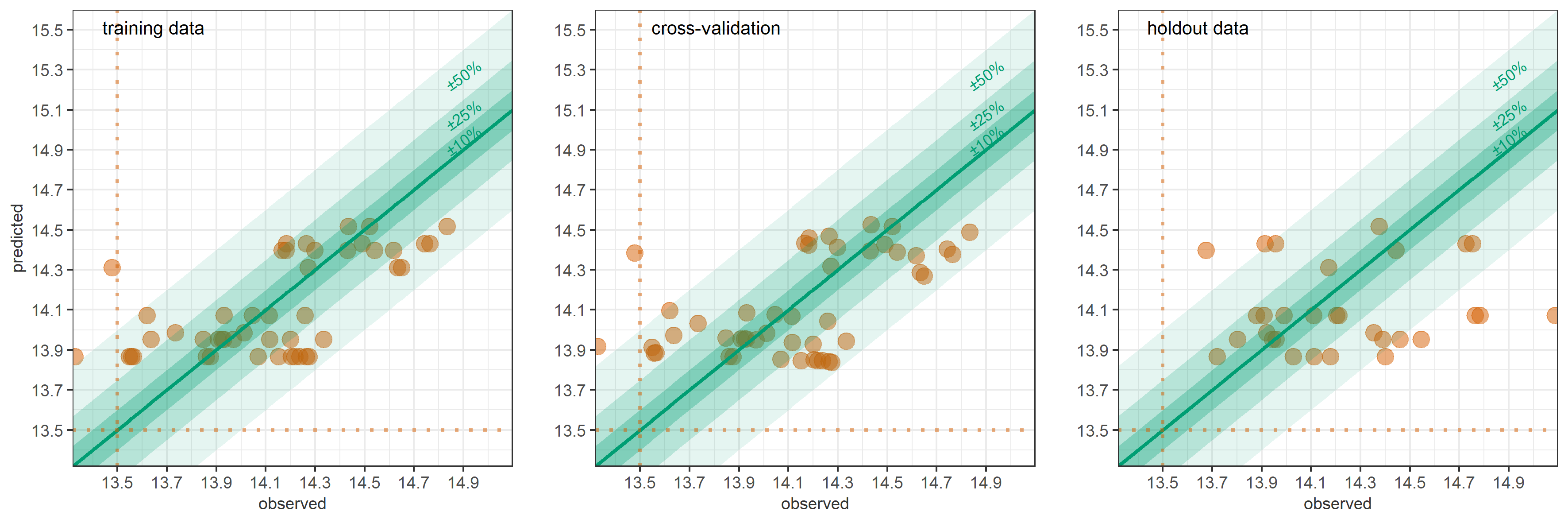


Figure S10. GLM-predicted versus observed mean MUAC, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.

Table S7. Performance of random forest models in Somalia, by wasting outcome.

| Statistic | | Binary:  Wasting (WFH + oedema) | | Continuous:  WFH | |
| --- | --- | --- | --- | --- | --- |
| Estimation performance | | | | | |
| Mean square error | training data | 0.00060 | | 0.01057 | |
| LOOCV | 0.00303 | | 0.05244 | |
| holdout data | 0.00265 | | 0.06680 | |
| Relative bias | LOOCV | -5.7% | | +7.6% | |
| holdout data | 15.6% | | +29.9% | |
| Relative precision of 95%CI | LOOCV | ±21.8% | | ±18.2% | |
| holdout data | ±16.8% | | ±12.4% | |
| Coverage of 95%CI | LOOCV | 48.9% | | 57.4% | |
| holdout data | 63.3% | | 33.3% | |
| Coverage of 80%CI | LOOCV | 25.5% | | 46.8% | |
| holdout data | 53.3% | | 30.0% | |
| Classification performance by wasting prevalence threshold (n = denominator of percentage) | | | | | |
| Sensitivity  (lower threshold) | LOOCV | ≥15% | 66.7% (33) | n/a | |
| holdout data | 72.2% (18) |
| Sensitivity  (upper threshold) | LOOCV | ≥20% | 11.1% (18) |
| holdout data | 42.9% (7) |
| Specificity  (lower threshold) | LOOCV | <15% | 50.0% (14) |
| holdout data | 33.3% (12) |
| Specificity  (upper threshold) | LOOCV | <20% | 93.1% (29) |
| holdout data | 91.3% (23) |
| Most important predictors | lag | Importance | p-value | Importance | p-value |
| Measles incidence rate | previous 0-2mths | 0.0002 | 0.050 | 0.006 | 0.010 |
| previous 1-3mths | 0.0003 | 0.040 | 0.006 | 0.010 |
| Cholera incidence rate | previous 0-2mths | not among most important predictors | | 0.003 | 0.020 |
| previous 1-3mths | 0.001 | 0.099 |
| Water price | previous 3-5mths | 0.0002 | 0.208 | 0.003 | 0.198 |
| previous 1-3mths | 0.0002 | 0.139 | 0.001 | 0.594 |
| previous 0-2mths | 0.0001 | 0.653 | 0.001 | 0.554 |
| Terms of trade | previous 4-6mths | 0.0001 | 0.495 | 0.001 | 0.188 |
| previous 3-5mths | 0.0001 | 0.644 | 0.001 | 0.426 |
| previous 2-4mths | 0.0001 | 0.337 | 0.001 | 0.455 |
| NDVI | previous 0-5mths | 0.0002 | 0.069 | 0.002 | 0.178 |

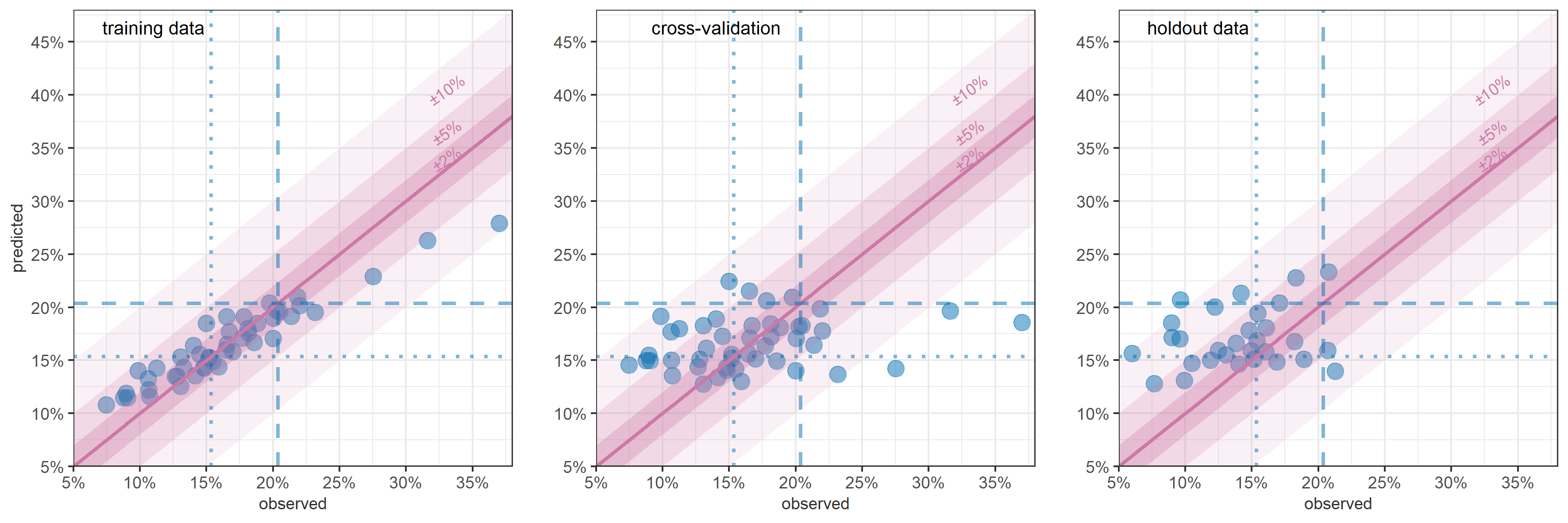


Figure S11. RF-predicted versus observed wasting (WFH + oedema) prevalence, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

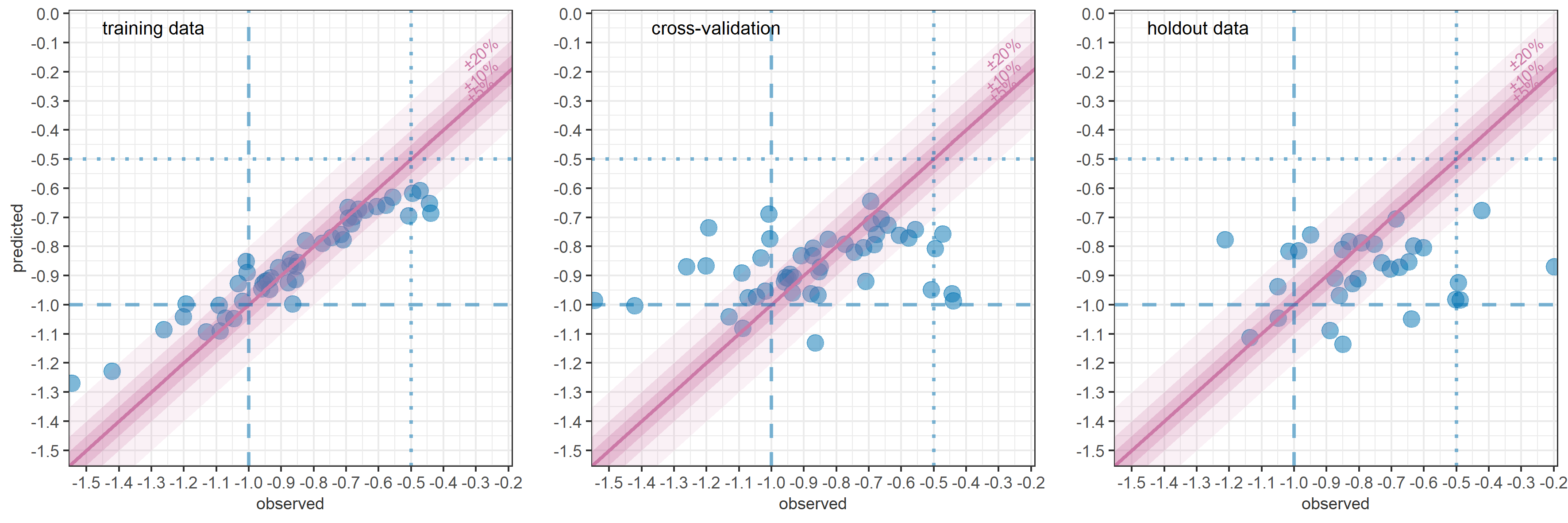


Figure S12. RF-predicted versus observed mean WFH, Somalia, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.

## Predictive accuracy of additional South Sudan models

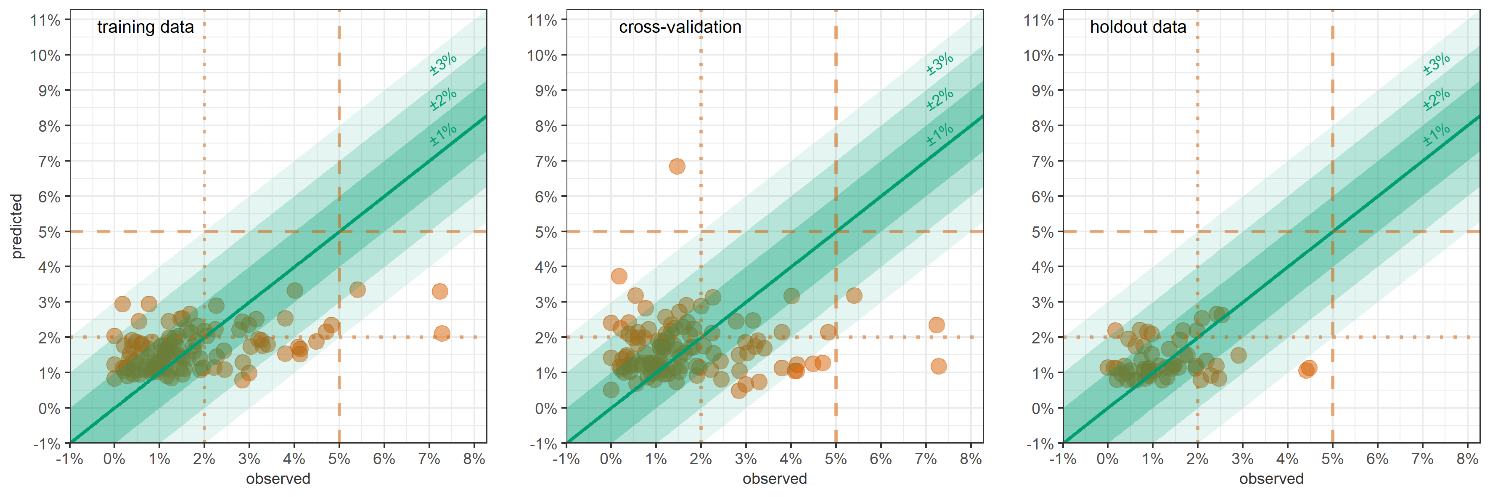


Figure S13. GLM-predicted versus observed severe wasting (MUAC + oedema) prevalence, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used severe wasting prevalence thresholds.

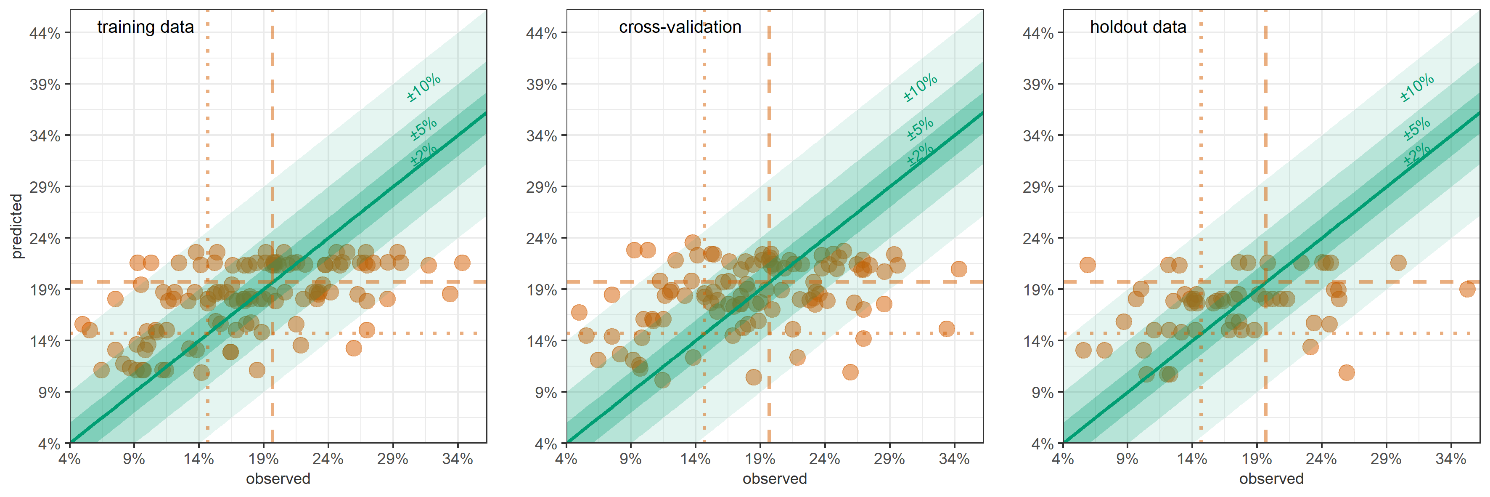


Figure S14. GLM-predicted versus observed wasting (WFH + oedema) prevalence, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

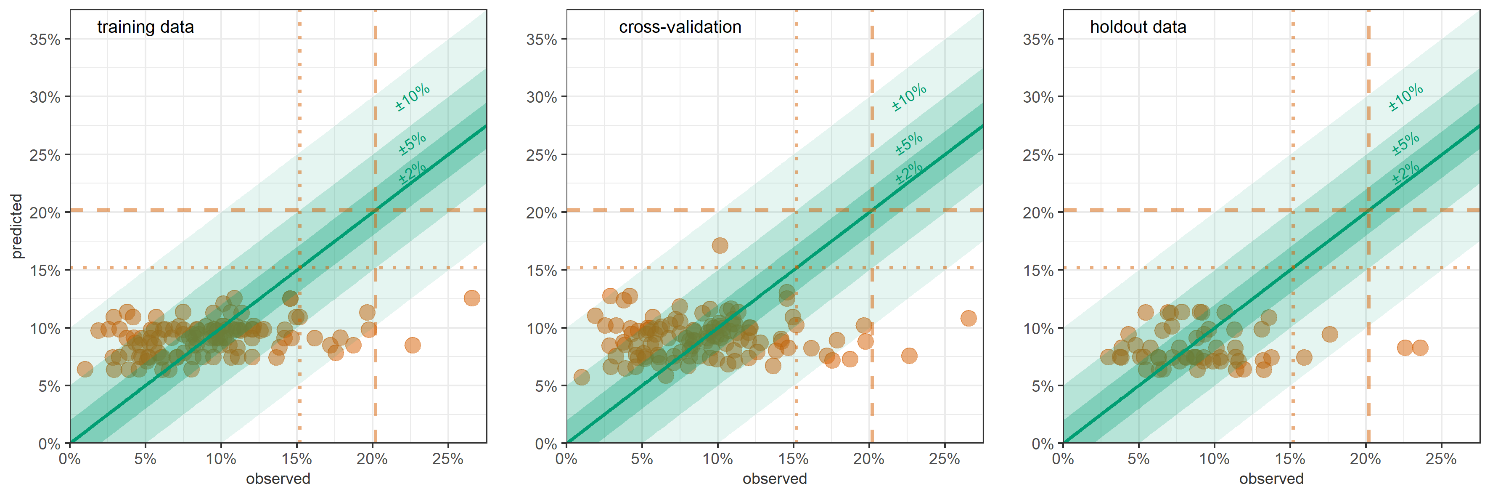


Figure S15. GLM-predicted versus observed wasting (MUAC + oedema) prevalence, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

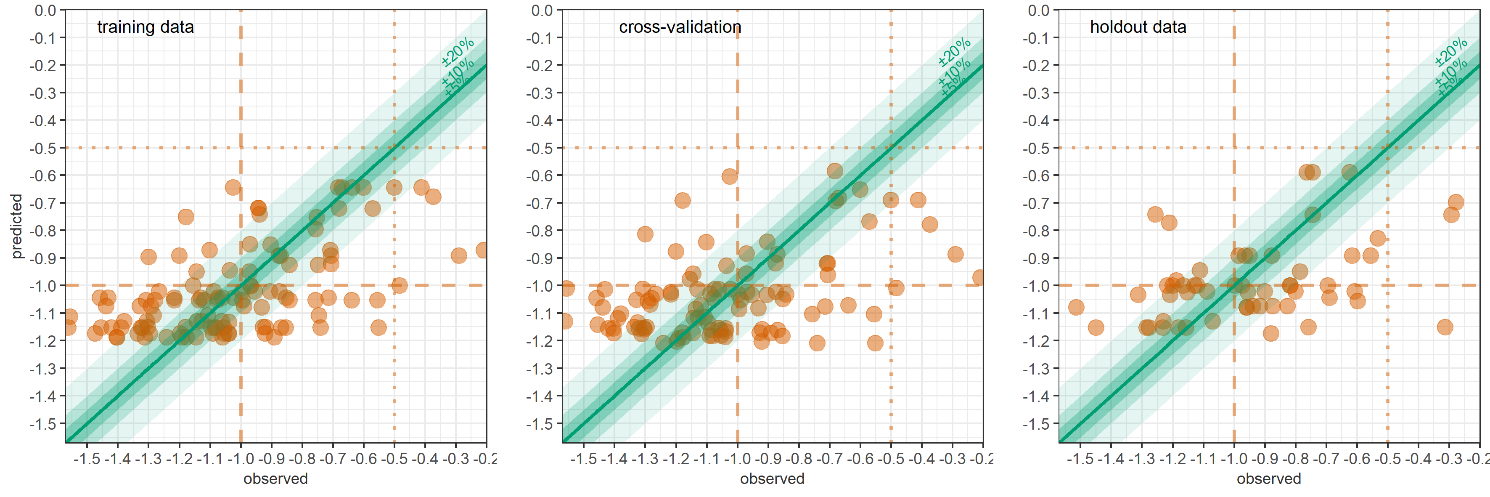


Figure S16. GLM-predicted versus observed mean WFH, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.

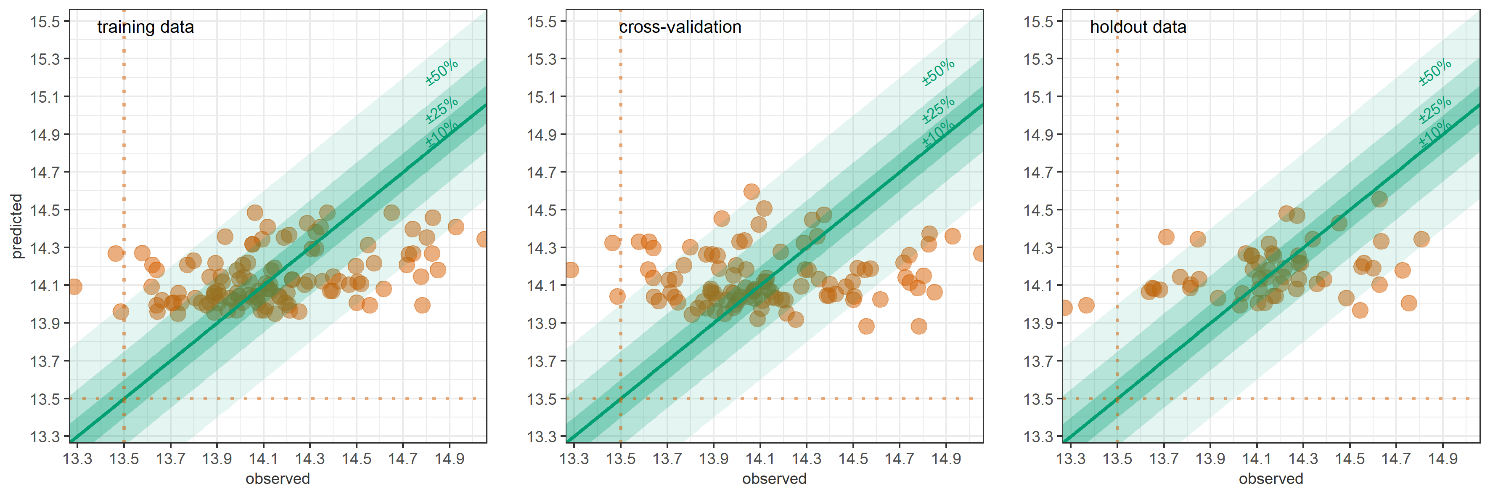


Figure S17. GLM-predicted versus observed mean MUAC, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.

Table S8. Performance of random forest models in South Sudan, by wasting outcome.

| Statistic | | Binary:  Wasting (WFH + oedema) | | | Continuous:  WFH | |
| --- | --- | --- | --- | --- | --- | --- |
| Estimation performance | | | | | | |
| Mean square error | training data | 0.00067 | | | 0.01225 | |
| LOOCV | 0.00367 | | | 0.07329 | |
| holdout data | 0.00348 | | | 0.07153 | |
| Relative bias | LOOCV | -8.9% | | | +2.8% | |
| holdout data | +2.5% | | | +20.9% | |
| Relative precision of 95%CI | LOOCV | ±19.3% | | | ±16.8% | |
| holdout data | ±11.8% | | | ±9.6% | |
| Coverage of 95%CI | LOOCV | 41.7% | | | 35.7% | |
| holdout data | 51.8% | | | 46.4% | |
| Coverage of 80%CI | LOOCV | 27.8% | | | 27.8% | |
| holdout data | 35.7% | | | 32.1% | |
| Classification performance by wasting prevalence threshold (n = denominator of percentage) | | | | | | |
| Sensitivity  (lower threshold) | LOOCV | ≥15% | 94.3% (88) | | n/a | |
| holdout data | 97.6% (42) | |
| Sensitivity  (upper threshold) | LOOCV | ≥20% | 11.3% (71) | |
| holdout data | 35.7% (28) | |
| Specificity  (lower threshold) | LOOCV | <15% | 29.6% (27) | |
| holdout data | 7.1% (14) | |
| Specificity  (upper threshold) | LOOCV | <20% | 88.6% (44) | |
| holdout data | 89.3% (28) | |
| Most important predictors | lag | Importance | p-value | | Importance | p-value |
| Main livelihood type | n/a | 0.0002 | 0.010 | | 0.0083 | 0.010 |
| Terms of trade | previous 4-6mths | 0.0002 | | 0.089 | 0.0037 | 0.129 |
| previous 3-5mths | 0.0001 | | 0.505 | 0.0048 | 0.089 |
| Number of measles vaccine doses administered per population | previous 2-4mths | 0.0003 | 0.069 | | 0.0053 | 0.109 |
| previous 3-5mths | 0.0001 | 0.356 | | 0.0037 | 0.208 |
| previous 0-2mths | 0.0001 | 0.218 | | 0.0038 | 0.178 |
| Rate of insecurity events | previous 4-6mths | 0.0002 | 0.119 | |  |  |
| previous 3-5mths | 0.0001 | 0.446 | |  |  |
| previous 2-4mths | 0.0001 | 0.495 | |  |  |
| Total rainfall | previous 0-5mths | 0.0004 | 0.010 | | 0.0073 | 0.010 |
| Proportion of the population that is internally displaced | n/a |  |  | | 0.0020 | 0.139 |

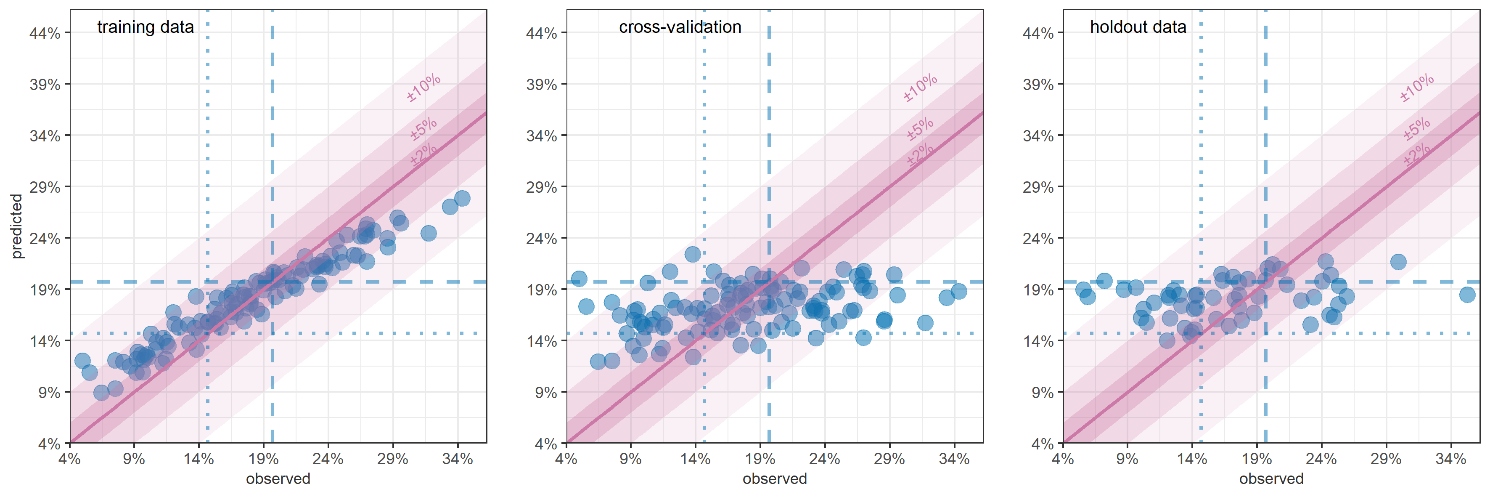


Figure S18. RF-predicted versus observed wasting (WFH + oedema) prevalence, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote commonly used wasting prevalence thresholds.

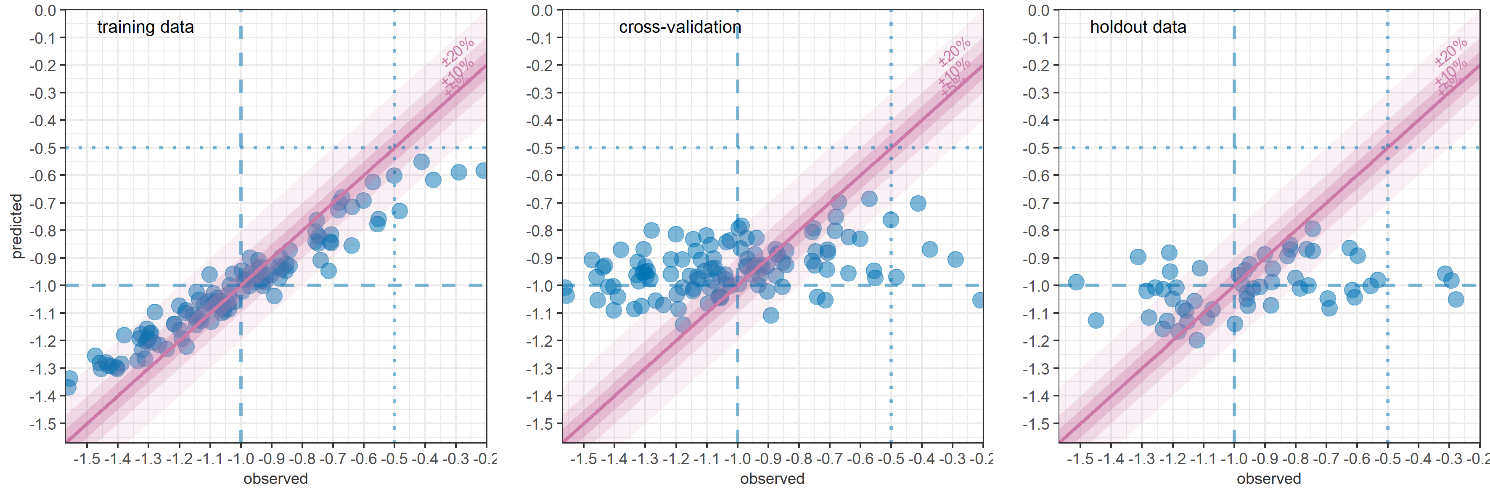


Figure S19. RF-predicted versus observed mean WFH, South Sudan, by district-month, on training data, LOOCV and holdout data. Shaded channels indicate different absolute deviance of predictions. Vertical dotted lines denote potentially useful thresholds.