



R exercises – Model validation

Marina Peris-Llopis

Research Methodology in Forest Sciences

(3.5 ECTS) 3513148

marinp@uef.fi

Model validation

You have built a model with data from one location and you apply it to a different location.

How can you know if the model is good for the other location? Can you apply it?

Root Mean Squared Error

Describes how much scatter there is between the measured and modeled values.

Standard deviation of the residuals.

How spread are the residuals.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}},$$

Diagram illustrating the RMSE formula components:

- Observation: Points to y_i in the formula.
- Prediction: Points to \hat{y}_i in the formula.

$$\text{RMSE}\% = \frac{\text{RMSE}}{\bar{y}} \times 100,$$

Observations
Mean

Diagram illustrating the RMSE% formula component:

- Observations Mean: Points to \bar{y} in the formula.

BIAS

Describes how much the average level of the modeled values differ from the measured ones.

$$\text{BIAS} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$$

How well the model matches the data.

$$\text{BIAS}\% = \frac{\text{BIAS}}{\bar{y}} \times 100,$$

Tendency to systematically over or underestimate the values.

Model bias

- You can apply a t-test to check if the bias is significant. R code:

```
t.test(y1, y2, paired = TRUE)
```

