# Prediction Performance with: : metrica

ShinyApp

bit.ly/3UwM40J

## Basics

**metrica** is a compilation of more than 80 functions designed to quantitatively and visually evaluate the prediction performance of regression (continuous) and classification (categorical) point-forecast models (e.g., APSIM, DSSAT, DNDC, Supervised Machine Learning).

## Using the functions

There are two basic arguments common to all metrica functions: (i) **obs** (Oi; observed, a.k.a. actual, measured, truth, target, label), and (ii) **pred** (Pi; predicted, a.k.a. simulated, fitted, modeled, estimate) values.
Optional arguments include data that allows to call an existing data frame containing both observed and predicted vectors, and tidy, which controls the type of output as a list (tidy = FALSE) or as a data.frame (tidy = TRUE).

## Installation

```
install.packages("metrica")
```

You can install the development version from GitHub with:

```
#install.packages("devtools")
devtools::install_github("adriancorrendo/metrica")
```

## Native datasets

The **metrica package** comes with four example datasets of continuous variables (regression) from the APSIM software:

- **Wheat**: 137 data-points of wheat grain N
- **Barley**: 69 data-points of barley grain number
- **Sorghum**: 36 data-points of sorghum grain number
- **Chickpea**: 39 data-points of chickpea aboveground dry mass

In addition, **metrica** also provides two native examples for categorical variables (classification):

- **land_cover**: binary dataset of land cover using satellite images. Values: 1=vegetation, 0 =other type of land cover.
- **maize_phenology**: data set of maize (*Zea mays* L.) phenology (16 crop development stages).

Check the metrics documentation to find all the performance metrics and their details: metrica

## Regression

```
R2(data = wheat, obs = obs, pred=pred, tidy = TRUE)
#> R2
#> 1 0.8455538
```

```
RMSE(data = wheat, obs = obs, pred = pred)
#> $RMSE
#> [1] 1.666441
```

```
KGE(data = wheat, obs = obs, pred = pred)
#> $KGE
#> [1] 0.9106471
```

Users can also calculate **all (default) or a selected list of metrics** at once using the function **metrics_summary()**:

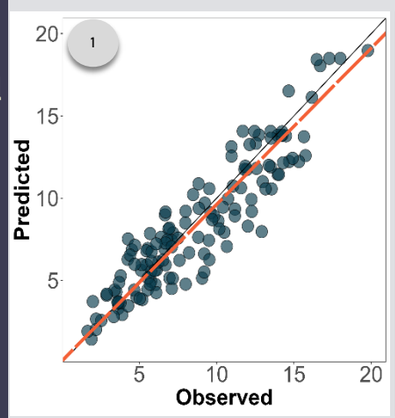
```
sel_r_metrics <- c("R2","MBE","RMSE", "RSR", "NSE",
"KGE", "CCC")

metrics_summary(data = wheat,
                obs = obs,
                pred = pred,
                type = "regression",
                metrics_list = sel_r_metrics)
```

## Plots

```
1. scatter_plot(data = wheat,
                obs = obs, pred = pred)
```

```
2. bland_altman_plot(data = wheat,
                     obs = obs, pred = pred)
```




## Classification

```
accuracy(data=maize_phenology, obs=actual, pred=predicted)
#> $accuracy
#> [1] 0.8834951
```

```
precision(data=maize_phenology, obs=actual, pred=predicted)
#> $precision
#> [1] 0.8335108
```

```
recall(data = maize_phenology, obs=actual, pred=predicted)
#> $recall
#> 1 0.8405168
```

For classification, users can also apply the **metrics_summary()** function to obtain multiple metrics at once:

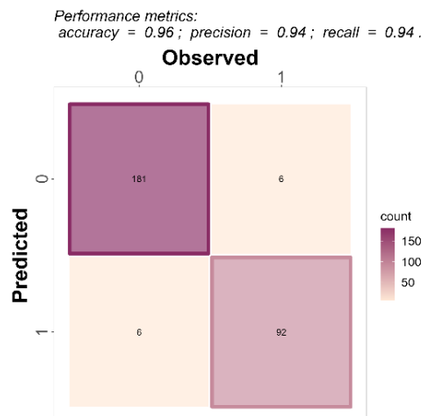
```
sel_c_metrics <- c("accuracy", "precision", "recall",
"fscore")

metrics_summary(data = landcover,
                obs = actual, pred = predicted,
                type = "classification",
                metrics_list = sel_c_metrics,
                pos_level = 1)
```

## Confusion matrix

```
confusion_matrix(data = .,
                 obs = labels, pred = predictions,
                 plot = TRUE,
                 unit="count")
```

### Binomial case




### Multinomial case