Towards field specific phosphate application norms with machine learning

August 28th, 2018 – EAAP 2018, Dubrovnik

Erwin Mollenhorst, Michel de Haan, Jouke Oenema, Rita Hoving, Roel Veerkamp, Claudia Kamphuis

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Nutrient cycle

Animal

Crops Balance Manure

Field

Current situation

Fixed phosphate application norms for crops / grassland

- 3 classes, based on P status of field
- For crops: 50 / 60 / 75 kg P_2O_5 (app. 22 / 26 / 33 kg P)

However, differences in P yield dependent on, e.g.:

- Field
- Crop
- Weather
- .....

Goal

To predict future maize yields

based on farm data and

open source weather data

Dataset from “KTC De Marke”

162 records of maize yields

24 different fields

Years 1996 – 2014

On average 7 times maize

Information on:

- N and P input and output
- Irrigation, P status of field
- Weather data (own weather station and open source)

Predicted variable

Maize yield, expressed in kg P per ha per year

Average yield: 22 kg P (13 - 36)

Generalized boosted regression models

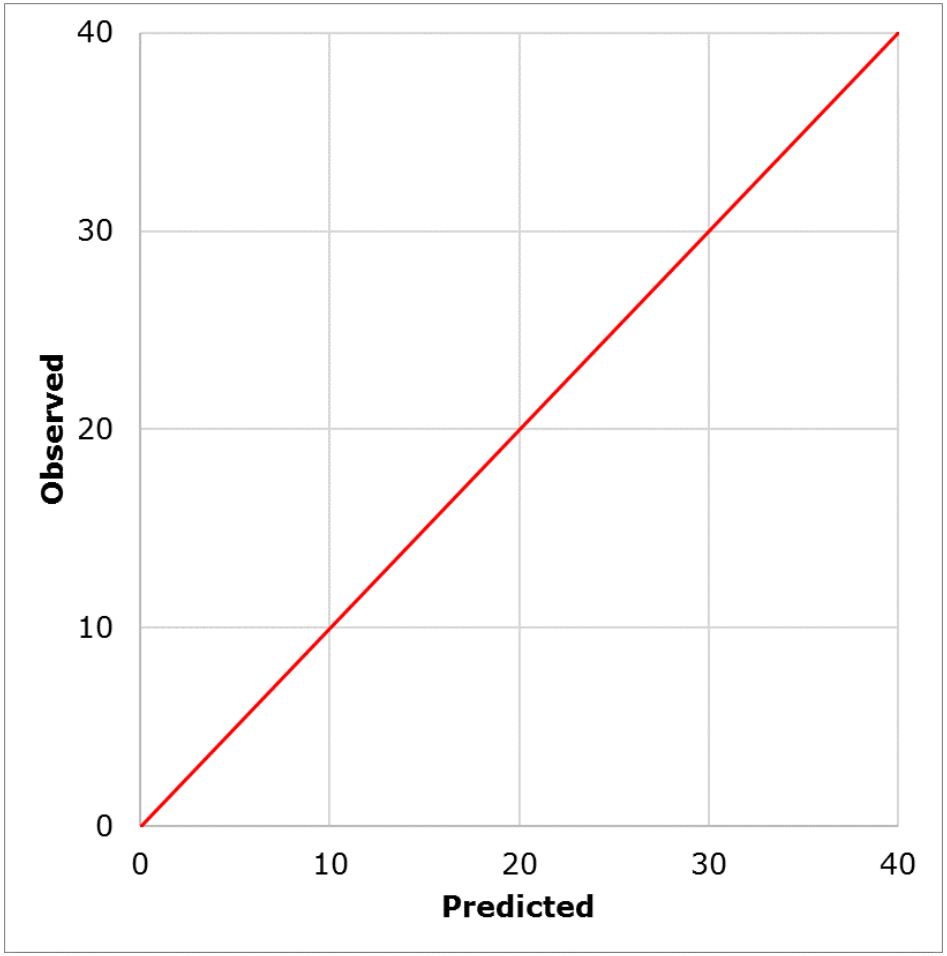
gbm package in R

Validation

70% train, 30% test, 1 year validation

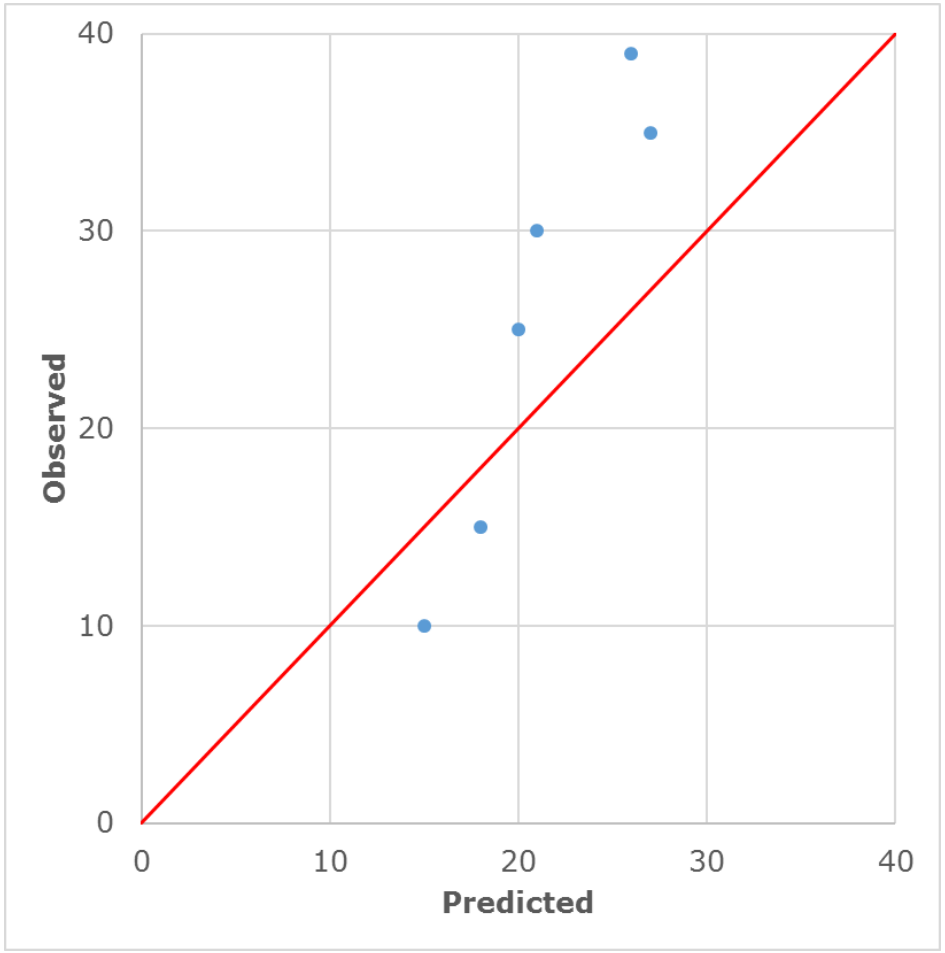
Final performance: 5 validation years combined

Performance criteria



Ideal situation: y = x

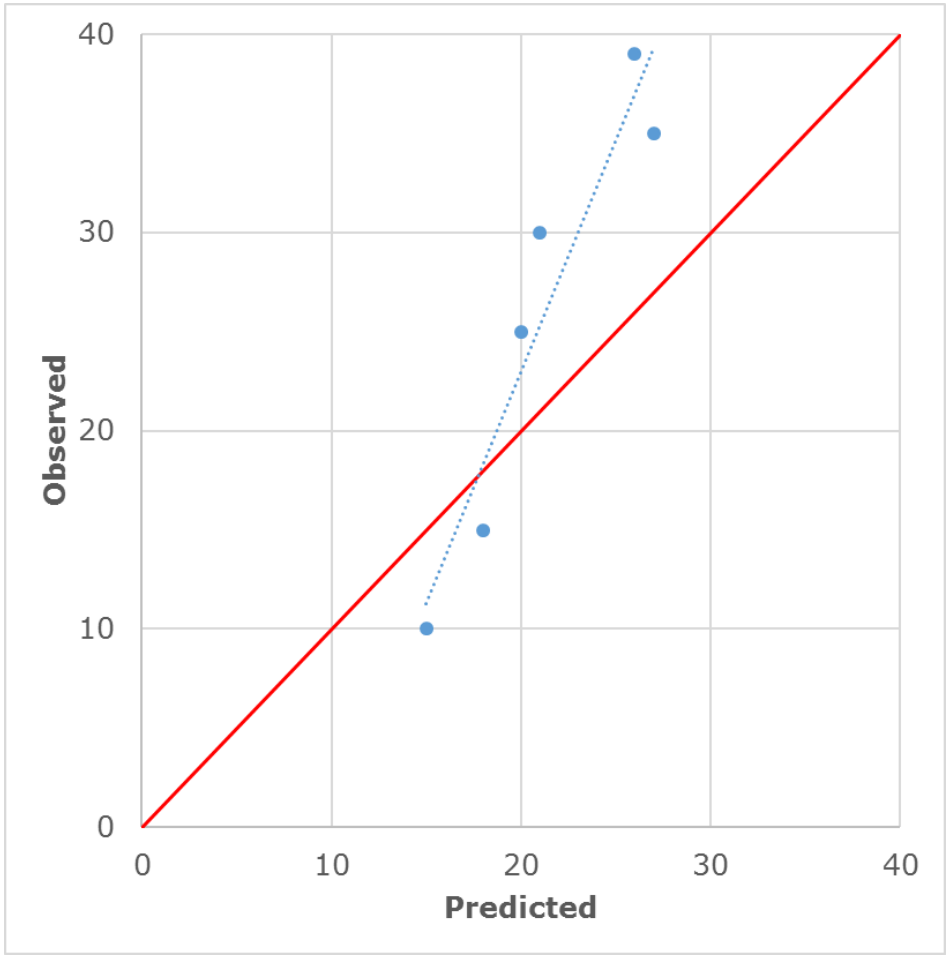
Performance criteria



RMSE - root mean squared error

Deviation from y=x

Performance criteria



RMSE root mean squared error

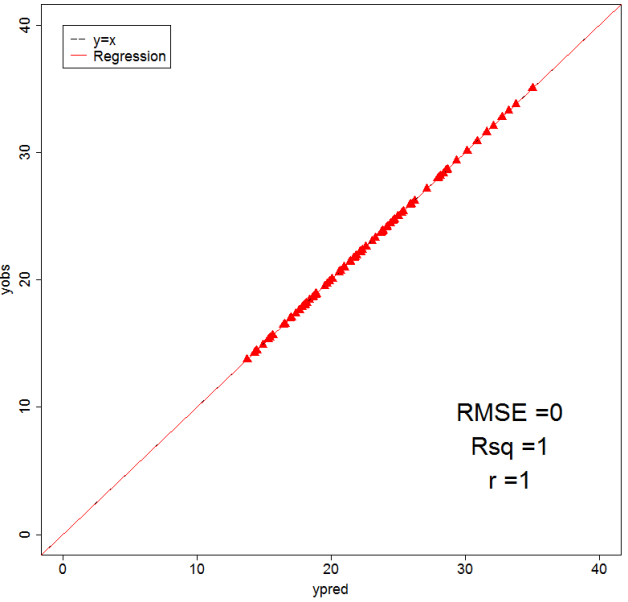
Deviation from y=x

r relative to linear fit

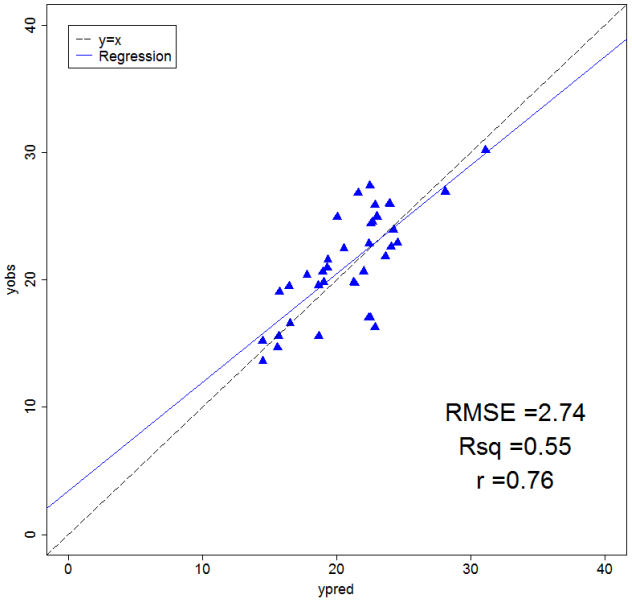
How much variation is explained (trend)

Pyield 2010 – Observed vs predicted

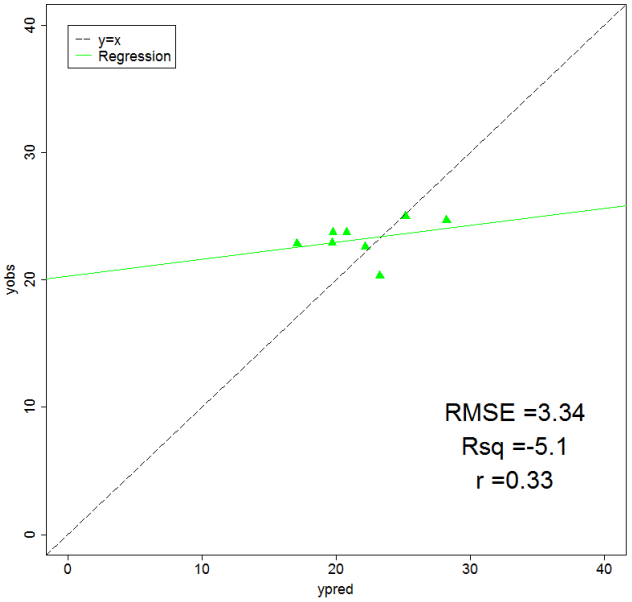
Train



Test

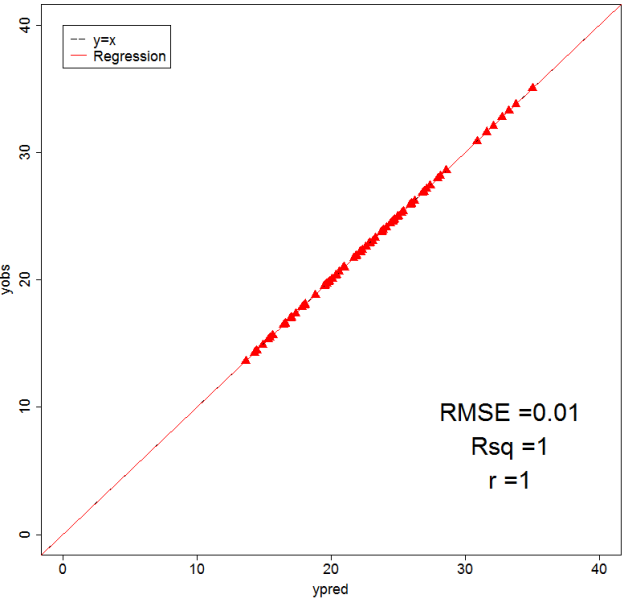


Validation

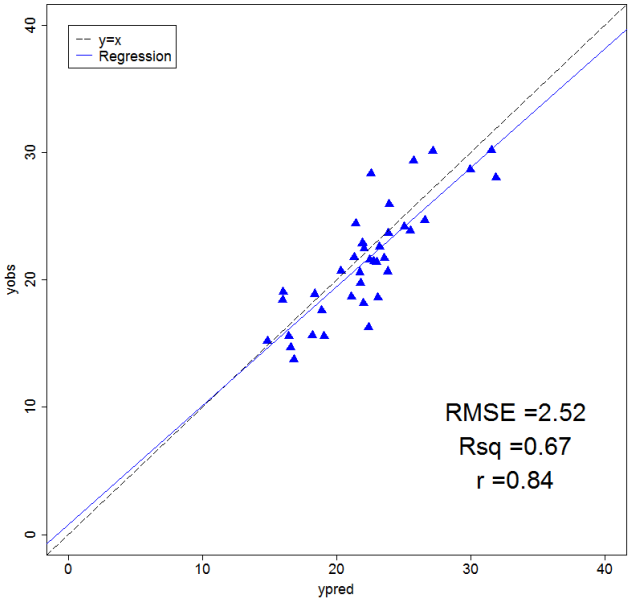


Pyield 2011 – Observed vs predicted

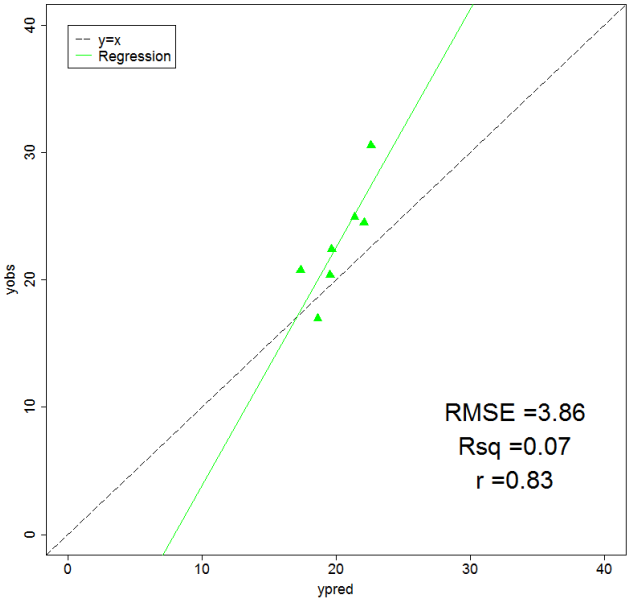
Train



Test

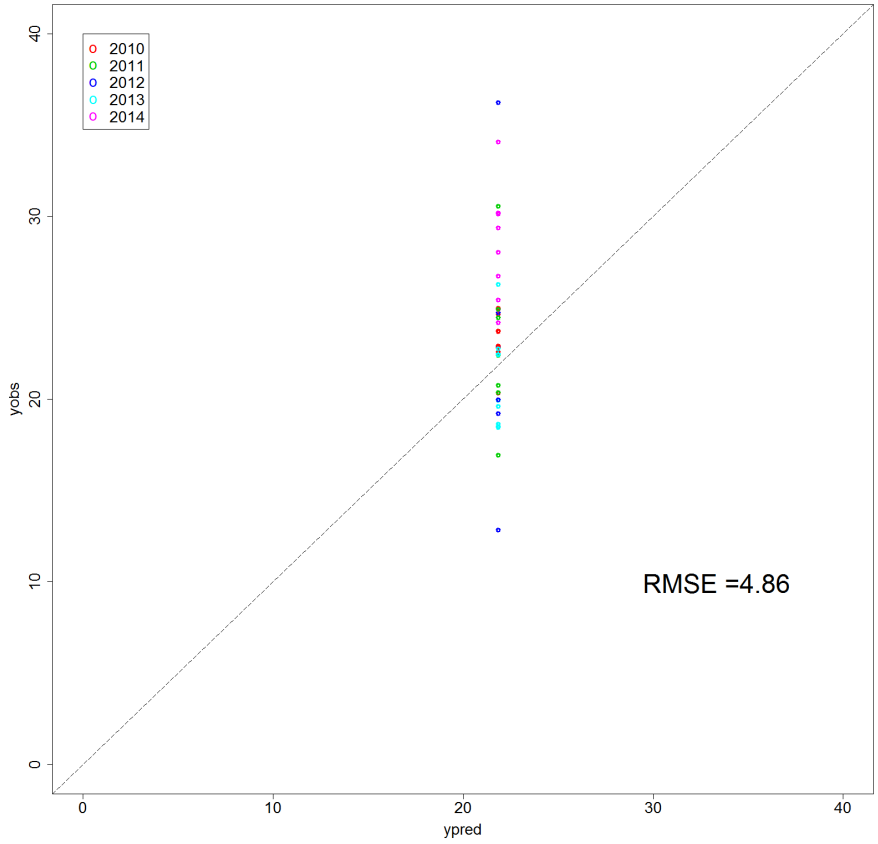


Validation

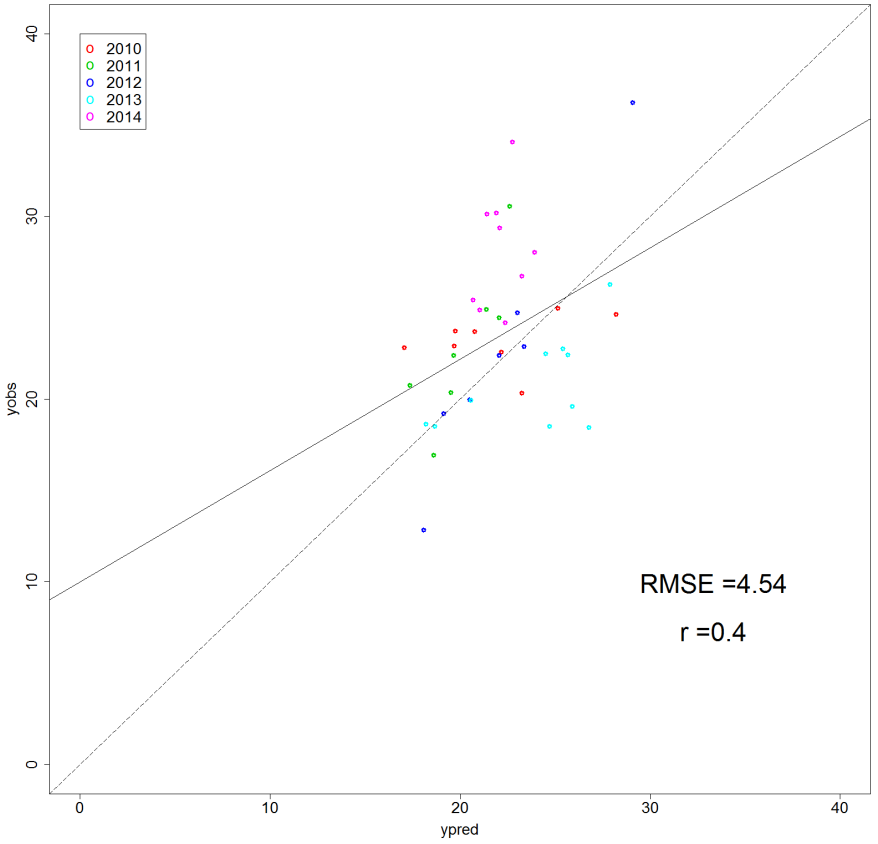


Norm vs model

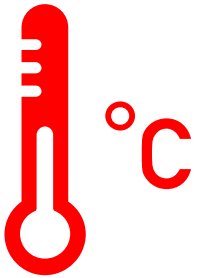
Norm (50 kg P_2O_5 = 22 kg P)



Predicted (validation sets)



Most important variables

Cropping scheme	Soil status	Weather	Yield history
			
Crop in previous year (grass/maize)	Phosphate status field	Maximum temperature in July	Average Pyield maize same field past 7 yrs

Conclusions

Machine learning is marginally better in predicting P yield than a generic norm (similar RMSE)

Furthermore, a trend could be shown in P yield ($r = 0.40$)

Multiple data sources are utilized

To be further explored, e.g., by including grassland

Acknowledgements

KTC De Marke

Gerjan Hilhorst

This project was part of

KB-27-01-0013

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018