

## Forecasting maize yield at field scale based on high-resolution satellite imagery

Rai A. Schwalbert <sup>a,b,c</sup>, Telmo J.C. Amado <sup>b,c</sup>, Luciana Nieto <sup>c</sup>, Sebastian Varela <sup>c</sup>, Geomar M. Corassa <sup>a,c</sup>, Tiago A.N. Horbe <sup>b</sup>, Charles W. Rice <sup>c</sup>, Nahuel R. Peralta <sup>d</sup>, Ignacio A. Ciampitti <sup>c</sup>.

<sup>a</sup> Agricultural Engineering Department, Federal University of Santa Maria, Rural Science Centre, Santa Maria, RS, Brazil

<sup>b</sup> Soil Department, Federal University of Santa Maria, Rural Science Centre, Santa Maria, RS, Brazil

<sup>c</sup> Department of Agronomy, Kansas State University, 2004 Throckmorton Plant Science Center, Manhattan, KS 66506, USA

<sup>d</sup> Monsanto Argentina, Department of Technology and Development of Corn and Sorghum, Pergamino, Buenos Aires, Argentina

**Keywords:** Yield forecasting models, Maize, Satellite imagery, Yield maps, Model validation, Sentinel-2

**Abstract:** Estimating maize (*Zea mays* L.) yields at the field level is of great interest to farmers, service dealers, and policy-makers. The main objectives of this study were to: i) provide guidelines on data selection for building yield forecasting models using Sentinel-2 imagery; ii) compare different statistical techniques and vegetation indices (VIs) during model building; and iii) perform spatial and temporal validation to see if empirical models could be applied to other regions or when models' coefficients should be updated. Data analysis was divided into four steps: i) data acquisition and preparation; ii) selection of training data; iii) building of forecasting models; and iv) spatial and temporal validation. Analysis was performed using yield data collected from 19 maize fields located in Brazil (2016 and 2017) and in the United States (2016), and normalized vegetation indices (NDVI, green NDVI and red edge NDVI) derived from Sentinel-2. Main outcomes from this study were: i) data selection impacted yield forecast model and fields with narrow yield variability and/or with skewed data distribution should be avoided; ii) models considering spatial correlation of residuals outperformed Ordinary least squares (OLS) regression; iii) red edge NDVI was most frequently retained into the model compared with the other VIs; and iv) model prediction power was more sensitive to yield data frequency distribution than to the geographical distance or years. Thus, this study provided guidelines to build more accurate maize yield forecasting models, but also established limitations for up-scaling, from farm-level to county, district, and state-scales.

**Publicado originalmente en:** Rai A. Schwalbert, Telmo J.C. Amado, Luciana Nieto, Sebastian Varela, Geomar M. Corassa, Tiago A.N. Horbe, Charles W. Rice, Nahuel R. Peralta, Ignacio A. Ciampitti. Forecasting maize yield at field scale based on high-resolution satellite imagery. *Biosystem Engineering*, 171: 179–192

DOI: <https://doi.org/10.1016/j.biosystemseng.2018.04.020>