Predicting the regional onset of the rainy season in West Africa

Motivation
Particularly in regions, where precipitation is limited to a few months per year, the reliable determination of the onset of the rainy season and the start of the sowing time is of crucial importance for sustainable food production. For farmers in the Volta Basin, the onset of the rainy season (ORS) is the most important variable: Planting too early is enhancing the risk of total crop failure, and planting too late is reducing valuable vegetation time (Laux et al., 2008). Due to a very high spatial and temporal variability of precipitation amounts, a non-uniform distribution of the rains during the rainy season and a high variability in the onset dates, farmers have problems to decide when to start with the sowing preparations. In the Volta Basin, the onset seldom occurs abruptly and is often preceded by short isolated showers with intermittent dry spells of various lengths, which are often misinterpreted as the start of the rains (false starts).

Methodologies
- Development and application of two different ORS definitions, for past mode (1961-1999) and predictive mode (current season)
- Regionalization via s-mode Principle Component Analysis (PCA) due to high spatial rainfall variability
- Linear Discriminant Analysis to judge day by day whether the ORS has already begun
- Linear Regression Analysis to identify the temporal shift of the ORS dates between the regions
- Multi Objective Fuzzy Rule-Based Classification (MOFRBC) of large-scale meteorological fields to identify weather patterns which are significantly linked with the ORS (not presented here, see therefore Laux et al., 2007)

Results
Fig. 1 is showing the membership functions of the ORS definition accounting for the most important agricultural aspects. The ORS is calculated as the first day of year (DOY), where:
\[ \Gamma = y_1 \cdot y_2 \cdot y_3 > \text{threshold (past mode)} \]
\[ \Gamma = y_1 \cdot y_2 > \text{threshold (predictive mode)} \]
Positive trends towards delayed regional ORS dates were observed using both definitions (Fig.2, left). PC5 is showing the highest number of false starts in April, and thus, the highest probability of total crop failure (Fig.2, right).

LRA was applied to estimate the regional ORS dates for the ongoing season. Tab. 2 shows the estimated models. The model quality is depending strongly on the mutual \( \Gamma \) combination (Fig. 5) of the definition. The correlation coefficients of the regional ORS dates were calculated to identify hot spots to improve the models.

References