

# Estimating rice production in the Mekong Delta, Vietnam, utilizing time series of Sentinel-1 SAR data

Kersten Clauss<sup>a,\*</sup>, Marco Ottinger<sup>a</sup>, Patrick Leinenkugel<sup>b</sup>, Claudia Kuenzer<sup>b</sup>

<sup>a</sup>Department of Remote Sensing, Institute of Geography and Geology, University of Wuerzburg, Wuerzburg, Germany

<sup>b</sup>German Remote Sensing Data Center (DFD), Earth Observation Center (EOC), German Aerospace Center (DLR), Oberpfaffenhofen, Wessling, Germany

---

## Abstract

Rice is the most important food crop in Asia and rice exports can significantly contribute to a country's GDP. Vietnam is the third largest exporter and fifth largest producer of rice, the majority of which is grown in the Mekong Delta. The cultivation of rice plants is important, not only in the context of food security, but also contributes to greenhouse gas emissions, provides man-made wetlands as an ecosystem, sustains smallholders in Asia and influences water resource planning and run-off water management. Rice growth can be monitored with Synthetic Aperture Radar (SAR) time series due to the agronomic flooding followed by rapid biomass increase affecting the backscatter signal. With the advent of Sentinel-1 a wealth of free and open SAR data is available to monitor rice on regional or larger scales and limited data availability should not be an issue from 2015 onwards. We used Sentinel-1 SAR time series to estimate rice production in the Mekong Delta, Vietnam, for three rice seasons centered on the year 2015. Rice production for each growing season was estimated by first classifying paddy rice area using superpixel segmentation and a phenology based decision tree, followed by yield estimation using random forest regression models trained on in-situ yield data collected by surveying 357 rice farms. The estimated rice production for the three rice growing seasons 2015 correlates well with data at the district level collected from the province statistics offices with  $R^2$ s of 0.93 for the Winter-Spring, 0.86 for the Summer-Autumn and 0.87 for the Autumn-Winter season.

Keywords: rice, SAR, Sentinel-1, time series, agriculture, food security

---

## 1. Introduction

The Mekong Delta, often referred to as "Vietnam's Rice Bowl", is the biggest rice producing region in Vietnam with the majority of its land area used for cultivating paddy rice. In 2015 Vietnam was the the worlds third biggest exporter and fifth largest producer  
5 of rice (FAOSTAT, 2015). Accurate and timely information about the rice production

---

\*Corresponding author

Email address: kersten.clauss@dlr.de (Kersten Clauss)

is vital in the context of food security, trade policy, land and water management and budgetary planning. Rice production and trade affect people on a global scale, as rice feeds half of the constantly growing population, especially in developing regions in Asia, Latin America and Africa (Kuenzer & Knauer, 2013; Fairhurst & Dobermann, 2002; Alexandratos & Bruinsma, 2012; Khush, 2005; United Nations, 2015). The world market price of rice is highly linked to the export from the 12 biggest rice exporters, which are responsible for 90% of the global rice trade (Muthayya et al., 2014). To alleviate the effects on food security in rice importing countries caused by the volatile rice market, initiatives like the FAO Rice Market Monitor, the GEOGLAM Crop Monitor and Asia-RiCE provide information on crop status, early warning on crop damage and factors influencing expected rice production. In spite the importance of this topic decision and policy makers still heavily rely on data published by national statistics offices, collected through samples, aggregated to administrative units and published six months to a year after the relevant harvest has been performed. The Mekong Delta is a coastal floodplain and subject to natural hazards such as floods, droughts and saltwater intrusion, which destroy rice crops and impact farmers livelihoods and food security. Remote sensing based rice production estimations can aid in the timely delivery of relevant information to decision and policy makers regarding rice production, trade and food security.

Rice is one of the few crops that can be grown under the condition of agronomic flooding. This management practice refers to a constant or periodic water layer covering the soil in which the plants grow, for the purpose of weed and pest control. In the Mekong Delta rice fields are usually flooded prior to transplanting rice seedlings from a nursery into the fields. The growth of rice plants can be divided into the vegetative, reproductive and ripening phase (De Datta, 1981) and each phase coincides with a change in plant morphology. These changes affect the wave-plant-water interaction of electromagnetic waves in the visible light as well as the microwave spectrum, resulting in a unique temporal signal of rice areas when observed with multispectral or SAR sensors. Detecting this temporal signal in remotely sensed time series has been the most frequently applied method to map rice areas with multispectral and microwave sensors (Kuenzer & Knauer, 2013; Mosleh et al., 2015; Dong & Xiao, 2016).

More than 90% of the global rice production is grown in Asia, where rice is often cultivated in areas with high precipitation and frequent cloud cover. Frequent cloud cover has been listed as a challenge for remote sensing of rice with multispectral sensors, which therefore require acquisitions with a high temporal frequency to accurately measure the flooding events of rice fields (Xiao et al., 2006). Active microwave sensors are much less affected by cloud cover and due to their all-weather, day and night imaging capabilities and have been used since the 1990s to map rice areas. Studies using time series data from C- and X-band sensors have shown their potential for rice mapping, the European Remote Sensing satellites (ERS) 1 and 2 (Aschbacher et al., 1995; Kurosu et al., 1995; Patel et al., 1995; Chakraborty et al., 1997; Le Toan et al., 1997; Panigrahy et al., 1997; Liew et al., 1998b; McNairn & Brisco, 2004; Diuk-Wasser et al., 2006), Radarsat (Liew et al., 1998a; Panigrahy et al., 1999; Ribbes, 1999; Shao et al., 2001; Li et al., 2003; Choudhury & Chakraborty, 2006; Yonezawa et al., 2012; Yang et al., 2016; Zhang et al., 2016), Envisat ASAR (Advanced Synthetic Aperture Radar) (Bouvet et al., 2009; Bouvet & Le Toan, 2011; Karila et al., 2014; Nguyen et al.,

2015), TerraSAR-X (Lopez-Sanchez et al., 2011; Pei et al., 2011; Asilo et al., 2014; Nelson et al., 2014), COSMO-SkyMed (Asilo et al., 2014; Nelson et al., 2014; Corcione et al., 2016; Busetto et al., 2017; Boschetti et al., 2017; Phan et al., 2018) and more recently Sentinel-1 (Clauss et al., 2017; Torbick et al., 2017; Onojeghuo et al., 2018; Nguyen et al., 2016; Son et al., 2017). These studies focussed on mapping rice areas, which are flooded prior to transplanting or seeding and achieved accuracies between 78% and 98%.

The backscatter signal over rice crops is sensitive to changes in above-ground plant biomass at the C- and L-band wavelengths, due to the penetration of the leaf canopy by the microwaves (Aschbacher et al., 1995; Le Toan et al., 1997; Inoue et al., 2002). This behavior can be exploited to monitor rice growth and estimate rice yield, which is correlated to the above-ground biomass (De Datta, 1981). The complex interaction of electromagnetic waves with the rice plant in its various growing stages leads to a backscatter signal consisting of multiple backscatter mechanisms, such as volume scattering from the canopy, double and multiple bounces between plants, soil and water (Le Toan et al., 1997). A number of modelling approaches have been studied to use the resulting backscatter signal to calculate biophysical parameters of the rice plant and estimate rice yield. Multivariate regression (Li et al., 2003) and neural networks (Chen & McNairn, 2006; Jia et al., 2013) have been used to predict rice yield and biomass, plant height and age has been correlated to C-band  $\sigma^0$  with polynomial regression (Chakraborty et al., 2005), rice ear weight derived from a rice canopy scattering model has been used to predict rice yield with a linear regression model (Zhang et al., 2017) and SAR backscatter has been used as input parameter to estimate rice yield with the DND (Salas et al., 2007) and ORYZA2000 (Shen et al., 2009) models. While it has been noted that SAR backscatter is influenced by biophysical plant variables besides the weight of the grains (i.e. yield) and linear models might be unsuitable to fully explain the backscatter-yield relation, multiple empirical models have been successfully applied to estimate rice yields (Zhang et al., 2017; Li et al., 2003; Chen et al., 2011).

Estimating rice production, the product of yield and harvested area, from remote sensing data requires the mapping of rice area and estimation of rice yield for the mapped area. In this study we propose to combine a previously published, Sentinel-1 time series based, rice mapping procedure (Clauss et al., 2017) with rice yield predictions based on seasonal random forest regression models to estimate the rice production in the Mekong Delta, Vietnam, for three rice growing seasons centered on the year 2015. The objectives of this study are, to:

- estimate rice production in the Mekong Delta in 2015 using empirical models
- study the potential of Sentinel-1 time series for rice production estimation at the regional scale
- estimate the transferability of seasonal regression models by predicting rice production for years differing from their training data

## 2. Materials and Methods

### 2.1. Study Area

The Mekong Delta (MKD) covers an area of circa 40,000 km<sup>2</sup> and is located between 8.5°–11.5°N and 104.5°–106.8°E where the Mekong River empties into the South China Sea. It is Vietnam's second most populous administrative region, with a population of 17,600,000, and the source for half of Vietnam's yearly rice production (General Statistics Office of Vietnam, 2017b). The Mekong Delta consists of 13 provinces, including the independent municipal city of Can Tho, further sub-divided into 133 urban and rural districts, provincial cities and towns (see figure 1). The Delta is a flat plain with fertile soils originating from alluvial sediments transported by the Mekong River. Large parts of the plain were subject to regular flooding by the single-peak pulse of the Mekong River. The current flooding regime, especially of agricultural areas, is controlled by multiple dams, dykes, sluices and regulatory measures being implemented upstream (Kuenzer et al., 2013b,a). Fields under controlled flooding schemes are flooded every three to four years in-between rice cycles to cover the soil with a new layer of sediment (Nguyen et al., 2012). The Delta as a whole is in an anthropocene state and subject to challenges of urbanization, agricultural intensification, anthropogenic water management, land subsidence salinization, sea-level rise, climate change and natural hazards, such as typhoons, flooding and drought (Renaud et al., 2013; Erban et al., 2014; Van Khanh Triet et al., 2017). In 2016 disasters caused by natural hazards lead to 264 deaths, 5,400 destroyed houses and 527,700 hectare (kha) of damaged rice crops (General Statistics Office of Vietnam, 2017b).



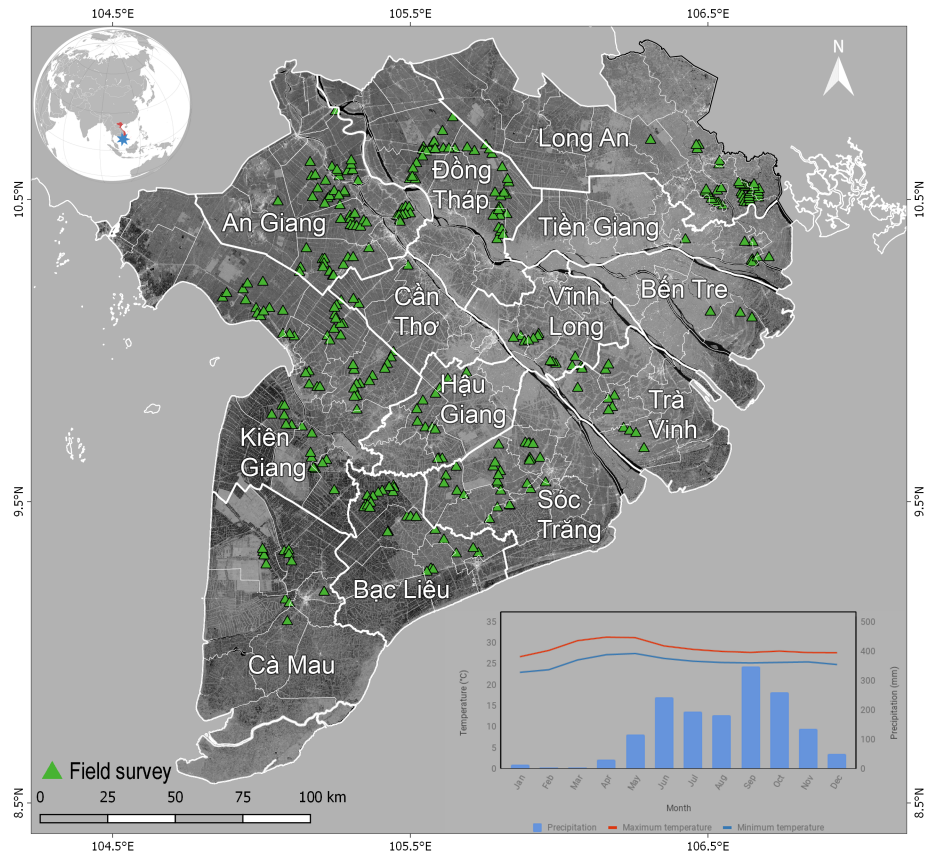


Figure 1: Overview of the Mekong Delta, its administrative units and locations of the field surveys. The inset globe indicates the location of Vietnam (red) and the study area (blue star). The climate chart shows the temperature (from (Saha et al., 2014)) and precipitation (from (Funk et al., 2015)) in Soc Trang in 2015. Background is the median backscatter of all Sentinel-1 acquisitions in 2015 at VH polarization. Administrative boundaries of second and third tier units from (Hijmans et al., 2015).

The Delta is located in the monsoon climate zone, *Am* according to the Köppen-Geiger classification (Peel et al., 2007), with monthly mean minimum temperatures exceeding  $20^{\circ}\text{C}$  throughout the year. Precipitation is on average 1800mm per year and characterized by a rainy season from June–December, caused by the SW monsoon, and a dry season from December–May, caused by the NW monsoon (see figure 1 - inset climate chart). This climate in combination with the water available from the Mekong River for agronomic flooding enables the year-round cultivation of rice plants with up to three harvests per field and year.

The cultivation of rice in the Delta is grouped into three growing seasons, according to sowing and harvesting dates:

- Hè-Thu (Summer-Autumn season)

125 • Thu-Đông (Autumn-Winter season)

• Đông-Xuân (Winter-Spring season)

The specific transplanting and harvest dates are governed by local water management practices and irrigation plans as well as the farmers selection of rice variety and fertilizer input. Availability of irrigation water is dependant on dykes that are usually controlled at the commune level. This results in heterogenous crop calendars with spatial variations. The sowing and harvest dates for rice are not defined in the statistical yearbooks where seasons are grouped by their harvesting dates. Due to inconsistency in existing definitions used for the rice seasons and lack of a definitive source we defined the temporal boundaries for each season. We define rice grown between 1st of November of the previous year and 31st of March to belong to the Winter-Spring season (WS), between 1st of March to 31st of August to the Summer-Autumn season (SA) and between 1st of July and 31st of January the following year to the Autumn-Winter season (AW). Based on existing rice season definitions for the Mekong Delta we defined the seasons with a temporal overlap to ensure the identification of all heading and flooding events when analyzing the time series. For Autumn-Winter this study follows the definition of Mùa (Main Wet Season), which ensures Thu-Đông is captured as well. The season boundaries are comparable to the season definitions used in previous studies on rice seasonality in the Mekong Delta (Nguyen et al., 2015; Sakamoto et al., 2006; Son et al., 2014; Sakamoto et al., 2007; Son et al., 2017; Phan et al., 2018).

The prevalent rice cultivation technique in the Mekong Delta is the transplanting of rice seedlings from seedbeds into the flooded paddy fields. The selection of rice variety by the farmers is largely governed by irrigation water availability, influences of salinization and the cropping intensity of the paddy, i.e. the number of rice harvests per year (Nguyen et al., 2012). The transplanting of seedlings into flooded soils creates a distinct temporal pattern in SAR backscatter time series due to the double bounce effect between water surface and rice tillers (Le Toan et al., 1997; Wang et al., 2005). Paddy fields form relatively homogenous clusters in the Mekong Delta, which lends itself to analysis with SAR satellite data.

## 2.2. Sentinel-1 SAR time series

We created time series from images acquired by the Synthetic Aperture Radar sensors aboard the Sentinel-1A and Sentinel-1B satellites which have been launched on 2014-04-03 and 2016-04-22, respectively. Both satellites carry active microwave sensors and measure backscatter at the C-band wavelength. After an In-Orbit Commission Review regular publication of Sentinel-1A datasets started on 2014-10-03, with the first Mekong Delta image acquired three days later. Each of the polar-orbiting satellites has a revisit time of 12 days at the equator, interleaved by six days, enabling at least one Sentinel-1 acquisition of every point on earth every six days in their current, two satellite constellation with coverage increasing from the equator towards the poles (Sentinel-1 Team, 2013). However, this theoretical maximum coverage is not achieved in large parts of the world due to the capacity of the on-board storage and the data downlink. The actual coverage is constantly increasing, with a large rise in available data after the inclusion of the European Data Relay Satellite EDRS-A to improve data downlink. The

default operation mode over land areas is dual-polarized, vertically transmitted and horizontally received (VH) and vertically transmitted and received (VV), Interferometric Wide Swath (IW) with 250 km swath width. Images in the default, high resolution mode, have a geometric resolution of 5 m by 20 m. The standard products delivered to users via the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) are at Level-1C processing level and are Single-Look Complex (SLC), containing amplitude and phase information, and Ground Range Detected (GRD), without phase information. In the GRD format the data is delivered at 10 m pixel spacing, after applying a 4 by 1 multi-looking pre-processing (Sentinel-1 Team, 2013).

We used all available Sentinel-1 scenes acquired over the Mekong Delta in Interferometric Wide Swath, at high resolution, dual polarization, i.e. the default operation mode, delivered in Ground Range Detected format. For the 2015 rice growing seasons the time series was built from 111 Sentinel-1A acquisitions. Creation of the time series for the 2017 rice growing seasons was performed using 319 acquisitions. The increase in temporal density of the time series is owed to the availability of Sentinel-1B data in 2017 as well as increased imaging and downlink capabilities due to the use of the first European Data Relay Satellite (EDRS-A) (Sentinel-1 Team, 2018).

### 2.3. Reference Data

Three reference datasets have been collected for training the regression models and assessing the accuracy of the estimated rice production. The first reference dataset was collected by performing field surveys and questionnaires at rice fields throughout the Mekong Delta. In total 357 field surveys were conducted between 2017-10-10 and 2018-01-24 covering 48 districts (see table 1). Special consideration was given to the distribution of the survey points, to ensure a minimum distance between each point as well as a distribution covering all provinces and rice cropping schedules in the Mekong Delta. Prior knowledge about rice production per district was used to stratify the survey locations. At each survey point the GPS coordinate of the field edge, a close-up photo of the rice plant, showing the growing stage, and a photo in each cardinal direction, to show the surrounding land-cover, was collected (see figure 2). At each location a questionnaire was conducted with the farmer to gather the transplanting date of the current season, transplanting and harvest dates of the previous two seasons, rice yield of the previous seasons, planted rice variety and which seasons they grow rice in. Additionally information was collected about issues with soil, droughts, floods or salinity intrusion during any of the growing seasons and if they transplant salt tolerant varieties during any of the seasons.

Table 1: Collected field surveys per province.

Province	Number of surveys	Start of survey	End of survey
An Giang	54	2017-10-11	2017-10-18
Bac Lieu	28	2017-10-12	2017-10-21
Ben Tre	3	2017-12-23	2017-12-23
Ca Mau	15	2017-10-12	2017-10-14
Can Tho	16	2017-12-22	2017-12-22
Dong Thap	42	2017-10-11	2017-10-25
Hau Giang	13	2017-10-17	2017-10-17
Kien Giang	57	2017-10-10	2017-10-13
Long An	67	2017-12-23	2018-01-24
Soc Trang	28	2017-10-12	2017-10-15
Tien Giang	9	2017-12-22	2017-12-22
Tra Vinh	12	2017-12-18	2017-12-18
Vinh Long	13	2017-12-19	2017-12-26

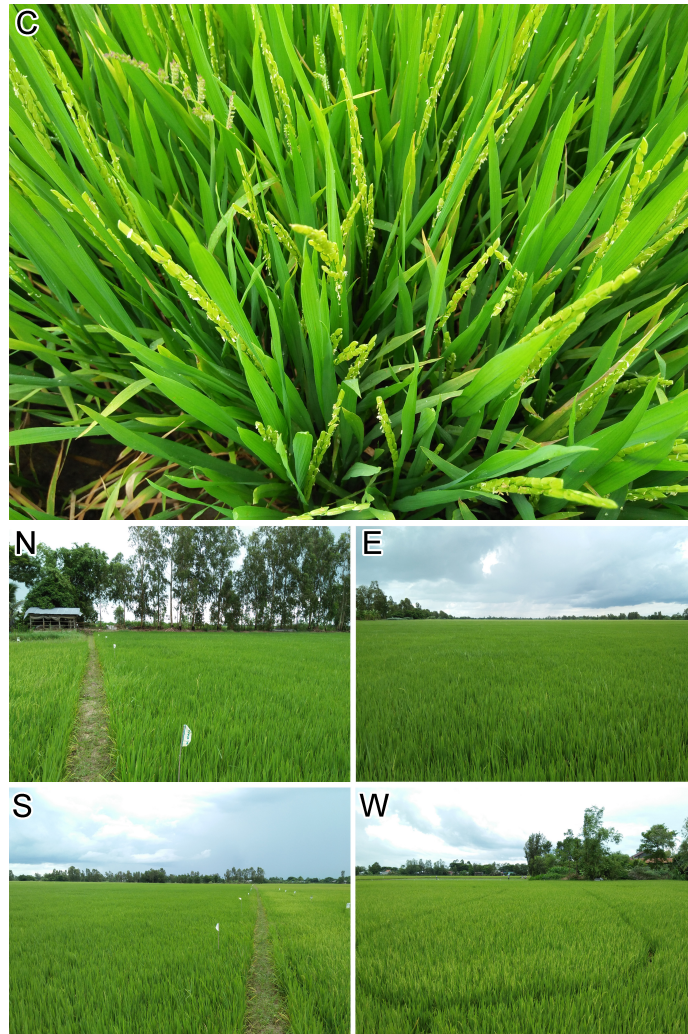


Figure 2: Example of photos taken during the field survey in An Giang province on 2017-10-17 at 10.4117°N 105.3077°E. Picture taken of close-up of the rice plant (C), view North (N), East (E), South (S) and West (W) from the location where the surveyor noted the GPS coordinates.

For assessing the accuracy of the production estimations data about the rice production per growing season and administrative unit was collected from statistics offices in Vietnam. The second reference dataset was collected from the General Statistics Office of Vietnam (GSO) containing rice production volumes at the province level published in the Statistical Yearbook of Vietnam 2015 (General Statistics Office of Vietnam, 2016) and on the GSO website (General Statistics Office of Vietnam, 2017a). The third reference dataset contains district level data and was compiled from the province's statistics offices, which collect and sample rice production volumes and report them to the GSO for aggregation into the Statistical Yearbook of Vietnam. The Statistical Yearbook of

each province has been published as a printed book in Vietnamese language, covering the last 5 years and had to be collected in person at the province's statistics offices. In total 13 yearbooks were collected and the rice production volume at district level transcribed and translated.

#### 2.4. Rice production estimation

We estimated rice production from Sentinel-1 data by applying a superpixel segmentation, calculating the time series at the object level, splitting the time series according to the rice cropping calendar, classifying rice areas and performing a regression analysis for each season. An overview of the processing workflow is given in figure 3.

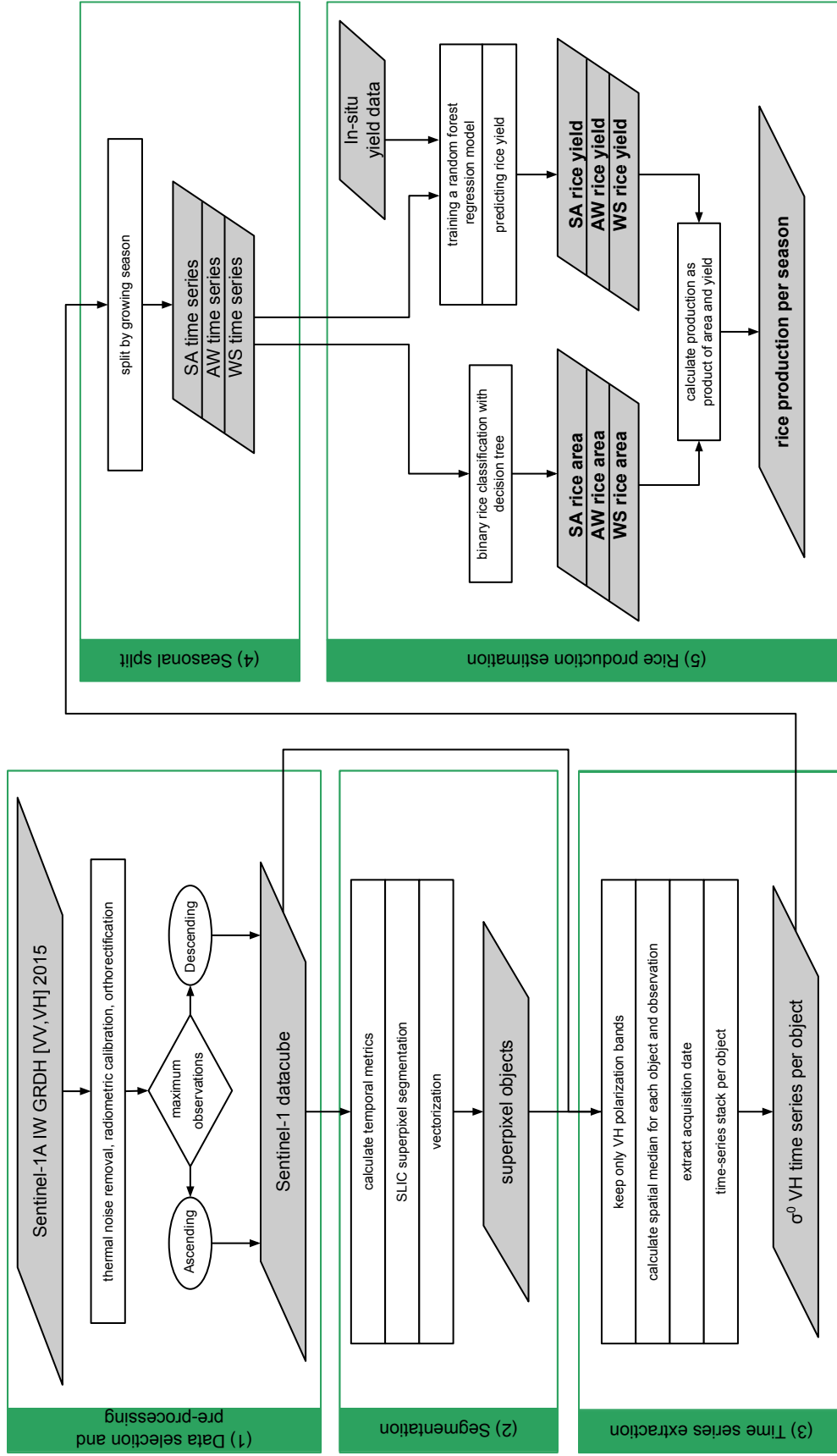


Figure 3: Workflow of the rice production estimation based on Sentinel-1 time series.

Creation of the object based time series with the steps 1) Data selection and pre-processing, 2) Segmentation and 3) Time series extraction is described in detail in a previous publication (Clauss et al., 2017) and is presented here in summarized form. This study utilized Sentinel-1 data provided as an analysis ready data cube by the Google Earth Engine (GEE - <https://earthengine.google.com/>) team, which co-locates a petabyte scale satellite data archive with vast processing capabilities (Gorelick et al., 2017). The Sentinel-1 data cube contains scenes pre-processed to the backscatter coefficient  $\sigma^0$  (sigma0 or sigma nought) after application of an orbit file, thermal noise removal, radiometric calibration and orthorectification. In order to limit the effects of look angle and orbit direction the data was filtered to the orbit direction with most observations, which for the Mekong Delta is the ascending orbit (figure 3 - 1). Overlapping swaths introduce noise into the time series due to differing look angles, which were reduced by applying pixel-wise filtering and masking to the most common look angle at the swath boundaries. The basis for the segmentation procedure, shown in figure 3 - 2, are multi-temporal metrics that have been calculated for all 2015 rice growing seasons using Sentinel-1 scenes from 2014-11-01 to 2016-01-31. We calculated the median and standard deviation of the VH polarized data and the 10th percentile of the VV polarized data, since these provide the necessary contrast between temporarily flooded fields and other land cover (Clauss et al., 2017). The resulting image was then segmented with the Simple Linear Iterative Clustering (SLIC) superpixel algorithm (Achanta et al., 2012) using a compactness value of 75, which balances space proximity and pixel values of the segments, and a mean segment size of 2 ha. The superpixel segments were then used to extract median  $\sigma^0$  values at VH polarization for each segment and Sentinel-1 acquisition in the data cube, arriving at a spatially averaged, per object, instead of per pixel, time series (figure 3 - 3). By changing the temporal extent of the data cube this procedure was used to extract three object based Sentinel-1 time series - one each for the 2014 to 2015 Winter-Spring, the 2015 Summer-Autumn and the 2015 to 2016 Autumn-Winter rice growing season (figure 3 - 4).

The segmentation based,  $\sigma^0$  VH Sentinel-1 time series have then been used to calculate the rice production for each 2015 rice growing season (figure 3 - 5). Estimating the rice production as a product of yield and area further requires the classification of sown rice area and the estimation of the rice yield. The rice area classification procedure used in this study is almost identical to the one published in (Clauss et al., 2017) with the exception that in this study a seasonal time series was used for the classification instead of a whole year and we consequently arrive at three rice area classifications - one for each rice growing season. Classifying rice areas is performed using a phenology based decision tree, which uses key growing stages of the rice plant as basis for the decision nodes. Backscatter time series at C-band wavelength and VH polarization exhibit distinct minima during the flooding of the rice fields and maxima at the heading stage (Le Toan et al., 1997). Minima and maxima of the full time series were used to differentiate rice areas from vegetation that does not experience agronomic flooding as well as from land cover that shows little backscatter variation over time, such as built-up areas, mangrove and bare soil. Local minima and maxima in the temporal signal are calculated within a 90 day window to detect a singular flooding event followed by rapid biomass growth, which helps to discern rice fields from natural wetlands and sporadically flooded areas. The temporal distance was chosen based on the shortest



reported growth duration for a single rice crop in the study area. The result are three classifications exhibiting the rice growing areas per 2015 rice growing season.

270 Rice yield for the 2015 rice growing seasons was estimated using the seasonal time series and a random forest regression model. The regression model was fitted on the yield reference data from the field surveys and time series extracted at these fields. Seasonal models were trained on reference data from the 2017 rice growing seasons. Therefore the time series from the Sentinel-1 data cube was extracted in the same manner as described for the 2015 rice growing seasons adapting the temporal slicing  
275 to the three 2017 rice growing seasons covered in the reference dataset. Instead of the whole Mekong Delta these time series have been extracted only at the locations of the field surveys. To ensure the use of a spatially averaged time series for the model fitting process, the reference data points have been manually moved into the center of the corresponding rice fields and data was extracted using the median inside a 50m circular  
280 buffer around the points for each image. Three groups of Sentinel-1 backscatter time series were derived, each group belonging to a rice growing season and each time series coded to the yield recorded in the farmer survey.

To ensure a compatible regression feature space for the selected years and the temporal transferability of the models we decided to calculate time series features that are able  
285 to describe time series of varying temporal density. Selection of the time series features for the regression models was guided in part by analysis of published regression results, expert knowledge and the usage of the Python package tsfresh for time series feature extraction based on scalable hypothesis (Christ et al., 2016). For each time series we calculated the features: standard deviation, mean, max, 10th, 25th, 50th, 75th and 90th  
290 percentile as well as the difference between the 90th and 10th and the 75th and 25th percentile, respectively. The percentiles have also been calculated after the time series was shifted, or lagged, by 10, 20, 30, 40 and 50 days. After performing a fourier transformation of the original time series the minimum, maximum and mean of the imaginary parts and the phase angle were calculated. The time series and yield reference data was  
295 split into 70% reserved for model training and 30% for testing the performance with cross-validation. This split was performed individually for each season, providing six independent training and test datasets. We trained three random forest regression models with the training data, called  $RF_{WS}$  for the Winter-Spring season model,  $RF_{SA}$  for Summer-Autumn and  $RF_{AW}$  for the Autumn-Winter rice growing season. We used the  
300 random forest module implemented in scikit-learn in version 0.19.1 (Pedregosa et al., 2011), set the number of trees to 500, the parameter of minimum sample size for a node split to 6, used bootstrapping and default implementation values for the remaining parameters. To estimate the model performance we calculated  $R^2$ , root mean squared error (RMSE) and mean absolute error (MAE) for each model comparing the predicted  
305 yield to the surveyed yield of the test split.

Yield for the 2015 rice growing seasons was predicted using the 2015 segment based time series and the random forest regression models  $RF_{WS}$ ,  $RF_{SA}$  and  $RF_{AW}$  for the according season. Rice yields were only predicted for those segments previously classified as rice. With this procedure we arrived at three datasets containing the classified  
310 rice area and predicted yield for each season and the complete Mekong Delta. Rice production was calculated by multiplying the predicted yield with the area of the according segment object. These production values were then summed up at the district level to

arrive at district level rice production estimations for three 2015 rice growing seasons.

315 In addition to testing the regression model performance with withheld reference data  
from the field surveys the accuracy of the rice production estimations has been assessed  
using independent reference data collected by Vietnam's statistics offices. Seasonal  
production estimations were aggregated at the district level to compare them to the  
district level data published by the statistics offices in each province and calculated  
320  $R^2$  and RMSE. After aggregating the production estimations at the province level we  
compared them to the province level data published by the GSO. To test differences  
between the province and district level reference datasets we aggregated the district  
level data from the statistics offices of the provinces to the province level and compared  
it to the province level dataset published by GSO, also calculating  $R^2$  and RMSE.

### 3. Results

325 Results of the rice area classification, yield estimation and production estimation  
are shown in figure 4. We estimated a total rice production of 9,521,757 tonnes for the  
Winter-Spring, 7,266,215 tonnes for the Summer-Autumn and 5,627,803 tonnes for the  
Autumn-Winter season. Throughout all seasons most rice was produced in the Kien  
Giang province. An Giang and Dong Thap are the second and third biggest producers in  
330 the Summer-Autumn and Autumn-Winter seasons, whereas in the Winter-Spring season  
Long An and An Giang are the second and third largest producers, respectively. The  
sown area estimates are similar for the Winter-Spring and Summer-Autumn seasons with  
1,564,234 hectare and 1,575,348 hectare, respectively, which is about twice the area  
classified for the Autumn-Winter seasons at 746,019 hectare. The regression models  
335 indicate the highest yields with the least spatial variation for the Winter-Spring season  
(figure 4 - top left) with a mean of 6.34 tonnes per hectare, the Summer-Autumn season  
exhibits the largest spatial variety in estimated yields with the lowest yields in the central  
areas of the Delta (figure 4 - center left) and an average yield of 5.24 tonnes per hectare.  
At the province level the models reported the highest average yields in Bac Lieu for both  
340 the Winter-Spring and Summer-Autumn seasons, whereas Kien Giang had the highest  
average yield in the Autumn-Winter season.

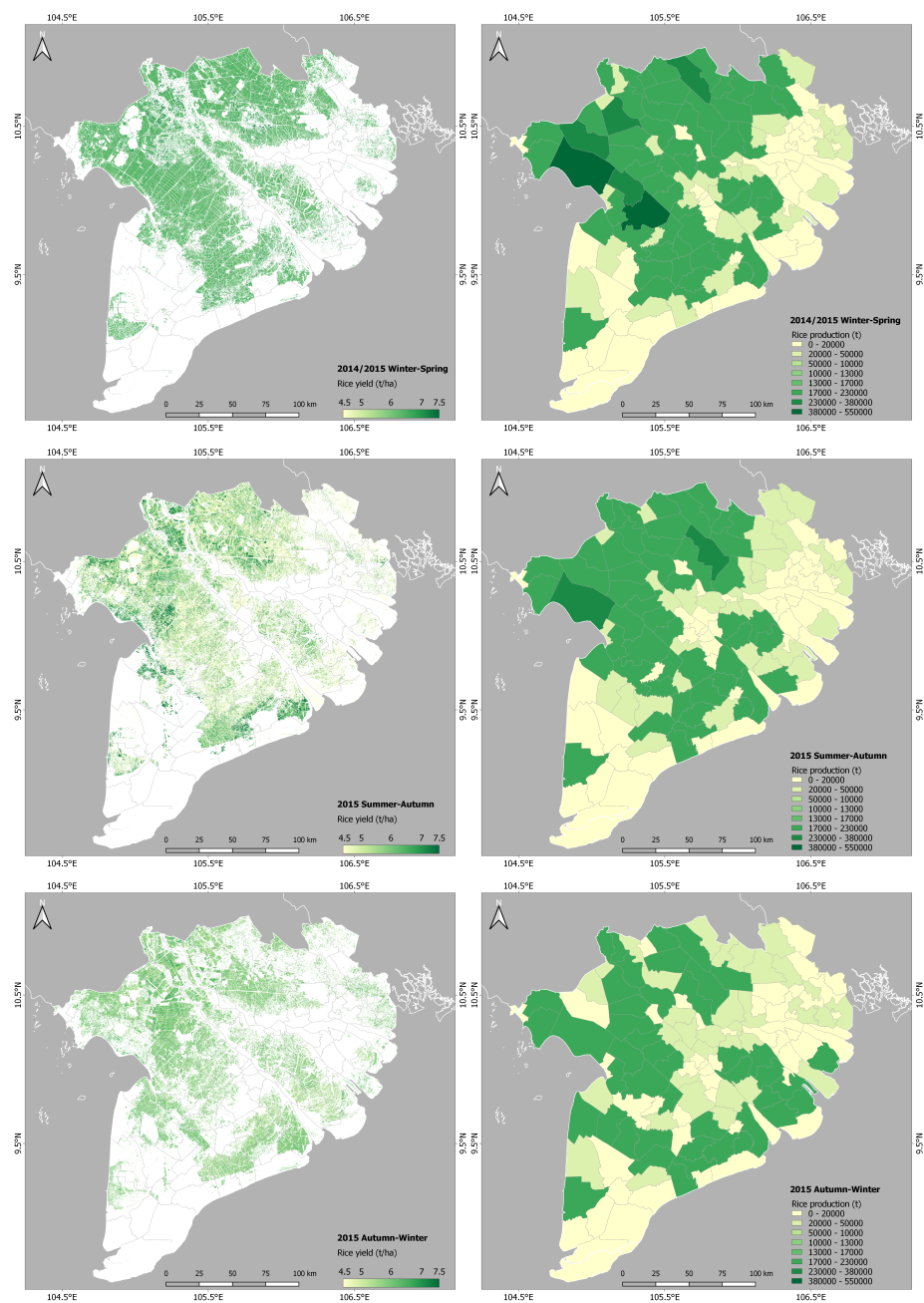


Figure 4: Estimated rice yield and production at district level for the 2015 rice growing seasons in the Mekong Delta, Vietnam.

Our estimated rice production for the year 2015 correlates well with the data published by the statistics offices at the district level with  $R^2$ s of 0.93 for the Winter-Spring, 0.86 for the Summer-Autumn and 0.87 for the Autumn-Winter season. The RMSE in tonnes per season at the district level was 24,846 for Winter-Spring, 25761 for Summer-Autumn and 14,221 for Autumn-Winter.

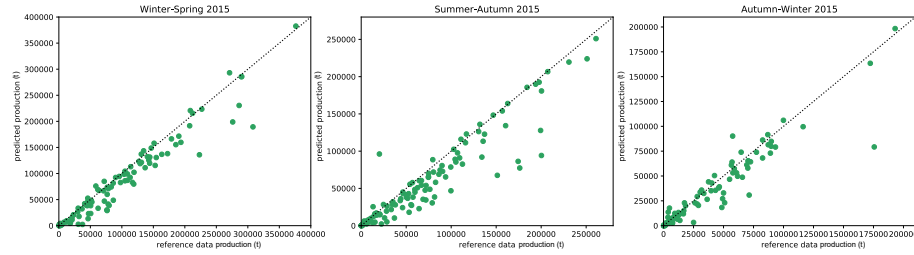


Figure 5: Comparison of estimated rice production at district level and production reference data from province statistics offices for the 2015 rice growing seasons.

Aggregated to the province level the estimated rice production correlates well to the GSO published data for the Winter-Spring, but not the other two seasons. The  $R^2$ s and RMSE in tonnes are 0.92 and 184,332 for Winter-Spring, 0.46 and 521,229 for Summer-Autumn and 0.06 and 447,994 for Autumn-Winter.

Comparing the reference datasets from the GSO and the official statistics at the district level aggregated to the province level we can see good correlation for the Winter-Spring and Autumn-Winter season but significant differences for the other Summer-Autumn season. The correlation coefficient  $R^2$  and RMSE in tonnes for the Winter-Spring season are 0.99 and 25,986, 0.99 and 4,153 for the Autumn-Winter season and 0.65 and 419,414 for the Summer-Autumn season, respectively. Furthermore the GSO reference dataset exhibits more missing data than the district level data and contains no rice production data for 5 of the 13 provinces in the Autumn-Winter growing season (see table 2).

Table 2: Rice production in tonnes reference data collected from the Province Statistics Offices ( $PSO$ ) and General Statistics Office ( $GSO$ ) at the province level for Winter-Spring ( $WS$ ), Summer-Autumn ( $SA$ ) and Autumn-Winter( $AW$ ) season. Missing values are indicated by NaN.

Province	$WS_{PSO}$	$WS_{GSO}$	$SA_{PSO}$	$SA_{GSO}$	$AW_{PSO}$	$AW_{GSO}$
An Giang	1804409	1804400	1277644	2250100	19230	19200
Bac Lieu	339226	339200	327033	326000	390184	380200
Ben Tre	87555	87600	67110	71800	119416	119400
Ca Mau	NaN	NaN	148986	152100	313104	313100
Can Tho	647251	645300	427948	762800	334911	NaN
Dong Thap	1443150	1443200	1168870	1941300	753241	NaN
Hau Giang	618460	618500	436040	674600	235522	NaN
Kien Giang	2173921	2224500	1565481	2135200	283256	283300
Long An	1535960	1569700	1039710	1331000	18018	18000
Soc Trang	892427	952400	1015824	1197800	150689	144500
Tien Giang	541012	541000	624868	803400	178506	NaN
Tra Vinh	NaN	469300	408070	408100	476510	476500
Vinh Long	425545	437000	339892	657700	317794	NaN

The performance of each seasonal regression model using cross-validation is shown in figure 6, comparing surveyed rice yields to predicted yields using the seasonal model on the test split of the reference data. All models show a good performance predicting the yield on the withheld test split. The Winter-Spring model ( $RF_{WS}$ ) has a  $R^2$  of 0.83, a RMSE of 0.58 and a MAE of 0.46. For the Summer-Autumn model ( $RF_{SA}$ ) we report a  $R^2$  of 0.82, a RMSE of 0.72 and a MAE of 0.55. The Autumn-Winter model ( $RF_{AW}$ ) achieved a  $R^2$  of 0.80, a RMSE of 0.46 and a MAE of 0.39.

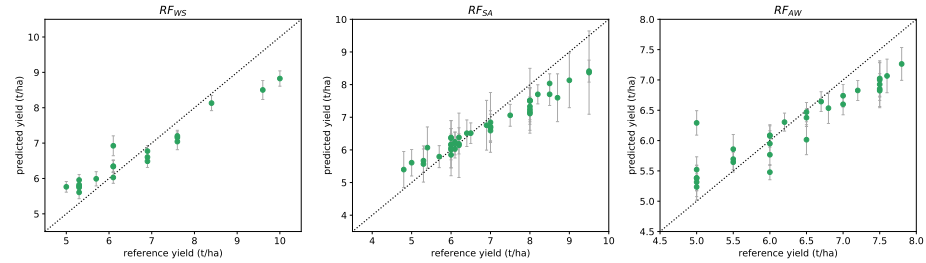


Figure 6: Cross-validation of the random forest regression models comparing reference and predicted yield (t/ha) for the Winter-Spring ( $RF_{WS}$ ), Summer-Autumn ( $RF_{SA}$ ) and Autumn-Winter ( $RF_{AW}$ ) season.

#### 4. Discussion

We presented a study estimating rice production in the Mekong Delta for three rice growing seasons centered around the year 2015. Rice production was estimated using Sentinel-1 time series, random forest regression, extensive field survey data and a

rice area classification method based on superpixel segmentation and a phenological decision tree. The regression models showed good agreement in cross validation and the estimated production correlates well with official data from the province's statistics offices. However, there are a number of potential error sources in this study that need to be discussed to appreciate the performance and transferability of the applied methods - namely the potential of rice area misclassification as a result of the superpixel segmentation or phenological decision tree classifier, potential bias and limitations of the empirical models and reliability of the reference data.

This study estimates rice production as the product of classified rice area and predicted yield per growing season, therefore a misclassification of rice area will influence the final production estimation. Even though the estimated rice production correlates well with the reference datasets it is possible that under- and overestimation of rice area compensate each other within an administrative unit. We aimed to reduce the influence of this error source by using reference data from the smallest administrative unit available to us. There is also the chance, that a misclassification of rice area was compensated by an over- or underestimation of the yield. Given the unavailability of a systematic reference dataset of rice area and yield for the Mekong Delta at the district level or lower we are unable to quantify this potential error source. To our knowledge there is no reference dataset available that would allow the assessment of rice area mapping accuracy in the Mekong Delta for the year 2015, let alone for individual rice cropping seasons. There is also the chance of introducing noise into the backscatter time series by the process of spatial averaging and the segmentation procedure failing to capture field boundaries and rice area omission errors due to key phenological growing stages not being captured in the temporal signal. The former could be reduced by using a higher resolution dataset as the basis for the segmentation, which was unavailable to us at the required scale, and the latter was a potential issue in 2015 but time series density improved significantly from 2016 onwards due to reduced revisit time, with the launch of Sentinel-1B, and increased data downlink (see Clauss et al. (2017) for extended discussion on the rice area classification method).

The random forest regression models performed well compared to the test split of the yield data and enabled rice production estimation with good agreement compared to government data. Apart from the influence of the reliability of the training data these empirical models can be a source of error and should be applied with caution. Firstly, we trained one model for each rice growing season and predicted rice yield of the same season but for a different year. This method worked in the Mekong Delta and we suspect this is in part due to a relatively stable cropping calendar and the application of such a method might fail when applied to different rice growing regions. Secondly, we are unable to assess whether these empirical models are able to predict rice production under adverse conditions such as floods and droughts caused by extreme weather events. The training data does not contain sample points of fields affected by potentially yield reducing conditions and we suspect the models might be unable to accurately predict the rice production for such areas. Outliers in the production estimation for the 2015 Autumn-Winter rice season might be explained partly by the onset of the 2015-2016 El Niño, at the end of 2015, and the resulting extreme temperature events in April 2016 (Thirumalai et al., 2017).

Our production estimation relies on rice yield data collected during field surveys

with questionnaires and rice production data published by the province's statistics offices. Errors in the first will propagate into the random forest regression models, while errors in the latter will influence the validity of the accuracy estimation. The yield data collected during the field surveys was self-reported by the farmers, a method that has been criticized, especially in the health and social sciences, as potentially biased (Hofmann et al., 2005). Performing yield measurements has the potential of providing more reliable yield data at the field level, but comes at the cost of the necessity of multiple field visits, destructive sampling and increased labour. The collection of such a dataset would be of interest for the wider scientific community working in the Mekong Delta but is out of scope of this study. Furthermore, the reliability of the reference data from government sources might have a larger impact on assessing the accuracy of rice production estimations. Rice production data at the province level is published by the GSO by collecting district level data from the province's statistics offices and aggregating it to the province level. We summed up the rice production per season and district, as published by the province's statistics offices, to the corresponding provinces and found disagreement with the province level data published by the GSO, especially in the Autumn-Winter and Summer-Autumn seasons. For Summer-Autumn we found a difference in the reported values, whereas for the Autumn-Winter season many districts reported rice production but the GSO dataset indicates missing data for the Can Tho, Dong Thap, Hau Giang, Tien Giang and Vinh Long provinces (see table 2). We ultimately decided to use district level data to assess the reliability of the rice production estimation to avoid errors of over- and underestimation due to the larger area of the provinces and data at the commune level (third tier) not being available to us. The production estimation shows good agreement with the reference data but we would like to stress the need for more comprehensive and more reliable reference data regarding rice yield, area and production at the district and commune level.

## 5. Conclusions

Our study showed that rice production can be estimated per growing season using Sentinel-1 SAR time series and random forest regression models. A prerequisite to rice production estimation is the ability to accurately classify rice areas and the availability of reliable reference data to train empirical models. The ability to estimate rice production in a timely, transparent and evidence-based manner is important in the context of food security and trade policy and valuable to decision and policy makers as well as to the rice farmers themselves. The reliance on national datasets regarding rice production has long been identified as an issue due to the time delay in their publishing as well as the amount of labour they consume. Remote sensing based production estimation, potentially in conjunction with national reference datasets, can aid in making information about the amount of rice produced available faster and on larger scales.

We showed that time series from the free and open Copernicus Sentinel-1 data archive can be used to map rice areas and estimate rice production. The Copernicus Sentinel-1 constellation alleviates the often cited issue of unavailability of dense time series or the cost of ordering such datasets. Using the co-location architecture of SAR time series data and vast amounts of processing power available through cloud processing initiatives, such as the Google Earth Engine, allows for a comparatively cheap and

fast estimation of rice production at regional or larger scale. Estimating rice production at larger scales is currently limited less by cost or availability of earth observation data or processing infrastructures but the collection of reliable reference data to train empirical models and estimate their performance.

465 The methods applied in this study enabled us to predict rice production for three rice growing seasons in the Mekong Delta for the year 2015 with  $R^2$ s of 0.93 for the Winter-Spring, 0.86 for the Summer-Autumn and 0.87 for the Autumn-Winter season compared to reference data at the district level. Three points were identified that could potentially improve the performance of the rice production estimation: the performance of the  
470 rice area classification, amount and reliability of yield reference data used to train the empirical models and availability of large scale reference data regarding rice production to accurately assess the accuracy of the prediction. Our study has shown that Sentinel-1 SAR time-series have the potential to estimate rice production at the local scale using empirical models by estimating the rice production in the Mekong Delta for the year  
475 2015 using empirical models per growing season trained on 2017 reference data. Further studies need to be conducted to evaluate the performance of such approaches at larger scales and in different study areas, which require the availability of large scale, reliable reference datasets regarding rice area, yield and production.

#### Acknowledgements

480 We would like to express our thanks to Dr. Vo Quoc Tuan and his students from Can Tho University, Vietnam, for their help collecting the reference data.

#### Funding

This work was partially funded by the German Ministry of Education and Research (BMBF) through the Catch-Mekong project.

#### 485 References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34, 2274–2282. URL: <http://ieeexplore.ieee.org/document/6205760/>. doi:10.1109/TPAMI.2012.120.
- 490 Alexandratos, N., & Bruinsma, J. (2012). World Agriculture Towards 2030/2050: the 2012 revision. Technical Report FAO Rome. URL: <http://www.fao.org/docrep/016/ap106e/ap106e.pdf>.
- Aschbacher, J., Pongsrihadulchai, A., Karnchanasutham, S., Rodprom, C., Paudyal, D. R., & Le Toan, T. (1995). Assessment of ERS-1 SAR data for rice crop mapping and monitoring. 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications, 3, 2183–2185. URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=524142>. doi:10.1109/IGARSS.1995.524142.
- 495



- Asilo, S., de Bie, K., Skidmore, A., Nelson, A., Barbieri, M., & Maunahan, A. (2014).  
500 Complementarity of Two Rice Mapping Approaches: Characterizing Strata Mapped  
by Hypertemporal MODIS and Rice Paddy Identification Using Multitemporal SAR.  
*Remote Sensing*, 6, 12789–12814. URL: <http://www.mdpi.com/2072-4292/6/12/12789/>. doi:10.3390/rs61212789.
- Boschetti, M., Busetto, L., Manfron, G., Laborte, A., Asilo, S., Pazhanivelan, S.,  
505 & Nelson, A. (2017). PhenoRice: A method for automatic extraction of spatio-  
temporal information on rice crops using satellite data time series. *Remote Sensing  
of Environment*, 194, 347–365. URL: <http://dx.doi.org/10.1016/j.rse.2017.03.029>.  
doi:10.1016/j.rse.2017.03.029.
- Bouvet, A., & Le Toan, T. (2011). Use of ENVISAT/ASAR wide-swath data for timely  
510 rice fields mapping in the Mekong River Delta. *Remote Sensing of Environment*, 115,  
1090–1101. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425710003603>.  
doi:10.1016/j.rse.2010.12.014.
- Bouvet, A., Le Toan, T., & Lam-Dao, N. (2009). Monitoring of the Rice Crop-  
ping System in the Mekong Delta Using ENVISAT/ASAR Dual Polarization Data.  
515 *IEEE Transactions on Geoscience and Remote Sensing*, 47, 517–526. URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4768722>. doi:10.1109/  
TGRS.2008.2007963.
- Busetto, L., Casteleyn, S., Granell, C., Pepe, M., Barbieri, M., Campos-Taberner, M.,  
Casa, R., Collivignarelli, F., Confalonieri, R., Crema, A., Garcia-Haro, F. J., Gatti,  
520 L., Gitas, I. Z., Gonzalez-Perez, A., Grau-Muedra, G., Guarneri, T., Holecz, F., Kat-  
santonis, D., Minakou, C., Miralles, I., Movedi, E., Nutini, F., Pagani, V., Palombo,  
A., Paola, F. D., Pascucci, S., Pignatti, S., Rampini, A., Ranghetti, L., Ricciardelli,  
E., Romano, F., Stavrakoudis, D. G., Stroppiana, D., Viggiano, M., & Boschetti, M.  
525 (2017). Downstream Services for Rice Crop Monitoring in Europe: From Regional to  
Local Scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Re-  
mote Sensing*, 10, 5423–5441. URL: <http://ieeexplore.ieee.org/document/7898821/>.  
doi:10.1109/JSTARS.2017.2679159.
- Chakraborty, M., Manjunath, K. R., Panigrahy, S., Kundu, N., & Parihar, J. S. (2005).  
530 Rice crop parameter retrieval using multi-temporal, multi-incidence angle Radarsat  
SAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59, 310–322.  
doi:10.1016/j.isprsjprs.2005.05.001.
- Chakraborty, M., Panigrahy, S., & Sharma, S. (1997). Discrimination of rice crop  
grown under different cultural practices using temporal ERS-1 synthetic aperture  
535 radar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52, 183–191.  
doi:10.1016/S0924-2716(97)00009-9.
- Chen, C., & McNairn, H. (2006). A neural network integrated approach for  
rice crop monitoring. *International Journal of Remote Sensing*, 27, 1367–1393.  
URL: <http://www.tandfonline.com/doi/abs/10.1080/01431160500421507>. doi:10.  
1080/01431160500421507.

- 540 Chen, C., Quilang, E., Alosnos, E., & Finnigan, J. (2011). Rice area mapping, yield, and production forecast for the province of Nueva Ecija using RADARSAT imagery. *Canadian Journal of Remote Sensing*, 37, 1–16. URL: <http://www.tandfonline.com/doi/abs/10.5589/m11-024>. doi:10.5589/m11-024.
- Choudhury, I., & Chakraborty, M. (2006). SAR signature investigation of rice crop using RADARSAT data. *International Journal of Remote Sensing*, 27, 519–534. doi:10.1080/01431160500239172.
- Christ, M., Kempa-Liehr, A. W., & Feindt, M. (2016). Distributed and parallel time series feature extraction for industrial big data applications, . URL: <http://arxiv.org/abs/1610.07717>. arXiv:1610.07717.
- 550 Clauss, K., Ottinger, M., & Kuenzer, C. (2017). Mapping rice areas with Sentinel-1 time series and superpixel segmentation. *International Journal of Remote Sensing*, 39, 1399–1420. doi:10.1080/01431161.2017.1404162.
- Corcione, V., Nunziata, F., & Mascolo, L. (2016). A study of the use of COSMO-SkyMed SAR PingPong polarimetric mode for rice growth monitoring. *International Journal of*, 1161. URL: <http://www.tandfonline.com/doi/abs/10.1080/01431161.2015.1131902>. doi:10.1080/01431161.2015.1131902.
- 555
- De Datta, S. (1981). *Principles and Practices of Rice Production*. New York: Wiley.
- Diuk-Wasser, M. A., Dolo, G., Bagayoko, M., Sogoba, N., Toure, M. B., Moghadam, M., Manoukis, N., Rian, S., Traore, S. F., & Taylor, C. E. (2006). Patterns of irrigated rice growth and malaria vector breeding in Mali using multi-temporal ERS-2 synthetic aperture radar. *International Journal of Remote Sensing*, 27, 535–548. URL: <http://www.tandfonline.com/doi/abs/10.1080/01431160500104350>. doi:10.1080/01431160500104350.
- 560
- Dong, J., & Xiao, X. (2016). Evolution of regional to global paddy rice mapping methods: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119, 214–227. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0924271616300995>. doi:10.1016/j.isprsjprs.2016.05.010.
- 565
- Erban, L. E., Gorelick, S. M., & Zebker, H. A. (2014). Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. *Environmental Research Letters*, 9. doi:10.1088/1748-9326/9/8/084010.
- 570
- Fairhurst, T., & Dobermann, A. (2002). Rice in the global food supply. *Better Crops International*, 16, 3–6.
- FAOSTAT (2015). Rice production and trade 2015. URL: <http://www.fao.org/faostat/>.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. URL: <http://www.nature.com/articles/sdata201566>. doi:10.1038/sdata.2015.66.
- 575

- General Statistics Office of Vietnam (2016). Statistical Yearbook of Vietnam 2015.  
Hanoi: Statistical Publishing House.
- General Statistics Office of Vietnam (2017a). Data Portal of the General Statistics Office  
of Vietnam. URL: [http://www.gso.gov.vn/Default\\_en.aspx?tabid=766](http://www.gso.gov.vn/Default_en.aspx?tabid=766).
- General Statistics Office of Vietnam (2017b). Statistical Handbook of Vietnam 2016.  
Hanoi: Statistical Publishing House.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R.  
(2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone.  
*Remote Sensing of Environment*, . URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425717302900>. doi:10.1016/j.rse.2017.06.031.
- Hijmans, R., Kapoor, J., Wieczorek, J., Garcia, N., Maunahan, A., Rala, A., & Mandel,  
A. (2015). Global Administrative Areas, Version 2.0. URL: <http://www.gadm.org/>.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A  
Meta-Analysis on the Correlation Between the Implicit Association Test and Ex-  
plicit Self-Report Measures. *Personality and Social Psychology Bulletin*, 31, 1369–  
1385. URL: <http://journals.sagepub.com/doi/10.1177/0146167205275613>. doi:10.  
1177/0146167205275613.
- Inoue, Y., Kurosu, T., Maeno, H., Uratsuka, S., Kozu, T., Dabrowska-Zielinska, K.,  
& Qi, J. (2002). Season-long daily measurements of multifrequency (Ka, Ku, X,  
C, and L) and full-polarization backscatter signatures over paddy rice field and their  
relationship with biological variables. *Remote Sensing of Environment*, 81, 194–204.  
doi:10.1016/S0034-4257(01)00343-1.
- Jia, M., Tong, L., Chen, Y., Wang, Y., & Zhang, Y. (2013). Rice biomass re-  
trieval from multitemporal ground-based scatterometer data and RADARSAT-2  
images using neural networks. *Journal of Applied Remote Sensing*, 7, 073509.  
URL: [http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JRS.7.](http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JRS.7.073509)  
073509. doi:10.1117/1.JRS.7.073509.
- Karila, K., Nevalainen, O., Krooks, A., Karjalainen, M., & Kaasalainen, S. (2014). Mon-  
itoring Changes in Rice Cultivated Area from SAR and Optical Satellite Images in Ben  
Tre and Tra Vinh Provinces in Mekong Delta, Vietnam. *Remote Sensing*, 6, 4090–  
4108. URL: <http://www.mdpi.com/2072-4292/6/5/4090>. doi:10.3390/rs6054090.
- Khush, G. S. (2005). What it will take to Feed 5.0 Billion Rice consumers in  
2030. *Plant Molecular Biology*, 59, 1–6. URL: <http://link.springer.com/10.1007/s11103-005-2159-5>.  
doi:10.1007/s11103-005-2159-5.
- Kuenzer, C., Campbell, I., Roch, M., Leinenkugel, P., Tuan, V. Q., & Dech, S. (2013a).  
Understanding the impact of hydropower developments in the context of upstream-  
downstream relations in the Mekong river basin. *Sustainability Science*, 8, 565–584.  
doi:10.1007/s11625-012-0195-z.

- Kuenzer, C., Guo, H., Huth, J., Leinenkugel, P., Li, X., & Dech, S. (2013b). Flood mapping and flood dynamics of the mekong delta: ENVISAT-ASAR-WSM based time series analyses. *Remote Sensing*, 5, 687–715. doi:10.3390/rs5020687.
- 620 Kuenzer, C., & Knauer, K. (2013). Remote sensing of rice crop areas. *International Journal of Remote Sensing*, 34, 2101–2139. doi:10.1080/01431161.2012.738946.
- Kurosu, T., Fujita, M., & Chiba, K. (1995). Monitoring of rice crop growth from space using the ERS-1 C-band SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 1092–1096. doi:10.1109/36.406698.
- 625 Le Toan, T., Ribbes, F., Floury, N., Fujita, M., Kurosu, T., Wang, L. F., Floury, N., Ding, K. H., Kong, J. A., Fujita, M., & Kurosu, T. (1997). Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. *IEEE Transactions on Geoscience and Remote Sensing*, 35, 41–56. doi:10.1109/36.551933.
- 630 Li, Y., Liao, Q., Li, X., Liao, S., Chi, G., & Peng, S. (2003). Towards an operational system for regional-scale rice yield estimation using a time-series of Radarsat ScanSAR images. *International Journal of Remote Sensing*, 24, 4207–4220. URL: <http://www.tandfonline.com/doi/abs/10.1080/0143116031000095970>. doi:10.1080/0143116031000095970.
- 635 Liew, S. C., Chen, P., Kam, S.-P., Tuong, T.-P., Minh, V. Q., & Lim, H. (1998a). Rice crops monitoring in the Mekong river delta using combined ERS and RADARSAT synthetic aperture radar. *IGARSS '98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings*. (Cat. No.98CH36174), 5, 2746–2748. doi:10.1109/IGARSS.1998.702338.
- 640 Liew, S. C., Kam, S.-P., Tuong, T.-P., Chen, P., Minh, V. Q., & Lim, H. (1998b). Application of Multitemporal ERS-2 Synthetic Aperture Radar in Delineating Rice Cropping Systems in the Mekong River Delta, vietnam. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1412–1420. doi:10.1109/36.718845.
- 645 Lopez-Sanchez, J. M., Ballester-Berman, J. D., & Hajnsek, I. (2011). First Results of Rice Monitoring Practices in Spain by Means of Time Series of TerraSAR-X Dual-Pol Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4, 412–422. URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5462990>. doi:10.1109/JSTARS.2010.2047634.
- 650 McNairn, H., & Brisco, B. (2004). The application of C-band polarimetric SAR for agriculture: a review. *Canadian Journal of Remote Sensing*, 30, 525–542. URL: <http://www.tandfonline.com/doi/abs/10.5589/m03-069>. doi:10.5589/m03-069.
- Mosleh, M. K., Hassan, Q. K., & Chowdhury, E. H. (2015). Application of Remote Sensors in Mapping Rice Area and Forecasting Its Production: A Review. *Sensors*, 15, 769–791. URL: <http://www.ncbi.nlm.nih.gov/pubmed/25569753>. doi:10.3390/s150100769.

- 655 Muthayya, S., Sugimoto, J. D., Montgomery, S., & Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, 1324, 7–14. URL: <http://doi.wiley.com/10.1111/nyas.12540>. doi:10.1111/nyas.12540.
- 660 Nelson, A., Setiyono, T., Rala, A., Quicho, E., Raviz, J., Abonete, P., Maunahan, A., Garcia, C., Bhatti, H., Villano, L., Thongbai, P., Holecz, F., Barbieri, M., Collivignarelli, F., Gatti, L., Quilang, E., Mabalay, M., Mabalot, P., Barroga, M., Bacong, A., Detoito, N., Berja, G., Varquez, F., Wahyunto, P., Kuntjoro, D., Murdiyati, S., Pazhanivelan, S., Kannan, P., Mary, P., Subramanian, E., Rakwatin, P., Intrman, A., Setapayak, T., Lertna, S., Minh, V. Q., Tuan, V., Duong, T., Quyen, N., Van Kham, D., 665 Hin, S., Veasna, T., Yadav, M., Chin, C., & Ninh, N. (2014). Towards an Operational SAR-Based Rice Monitoring System in Asia: Examples from 13 Demonstration Sites across Asia in the RIICE Project. *Remote Sensing*, 6, 10773–10812. URL: <http://www.mdpi.com/2072-4292/6/11/10773/>. doi:10.3390/rs61110773.
- 670 Nguyen, D. B., Clauss, K., Cao, S., Naeimi, V., Kuenzer, C., & Wagner, W. (2015). Mapping Rice Seasonality in the Mekong Delta with Multi-Year Envisat ASAR WSM Data. *Remote Sensing*, 7, 15868–15893. doi:10.3390/rs71215808.
- 675 Nguyen, D. B., Gruber, A., & Wagner, W. (2016). Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data. *Remote Sensing Letters*, 7, 1209–1218. URL: <https://www.tandfonline.com/doi/full/10.1080/2150704X.2016.1225172>. doi:10.1080/2150704X.2016.1225172.
- Nguyen, T. T. H., De Bie, C. A. J. M., Ali, A., Smaling, E. M. A., & Chu, T. H. (2012). Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam, through hyper-temporal SPOT NDVI image analysis. *International Journal of Remote Sensing*, 33, 415–434. doi:10.1080/01431161.2010.532826.
- 680 Onojeghuo, A. O., Blackburn, G. A., Wang, Q., Atkinson, P. M., Kindred, D., & Miao, Y. (2018). Mapping paddy rice fields by applying machine learning algorithms to multi-temporal Sentinel-1A and Landsat data. *International Journal of Remote Sensing*, 39, 1042–1067. URL: <https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1395969>. doi:10.1080/01431161.2017.1395969.
- 685 Panigrahy, S., Chakraborty, M., Sharma, S. A., Kundu, N., Ghose, S. C., & Pal, M. (1997). Early estimation of rice area using temporal ERS-1 synthetic aperture radar data a case study for the Howrah and Hughly districts of West Bengal, India. *International Journal of Remote Sensing*, 18, 1827–1833. doi:10.1080/014311697218133.
- 690 Panigrahy, S., Manjunath, K. R., Chakraborty, M., Kundu, N., & Parihar, J. S. (1999). Evaluation of RADARSAT Standard Beam data for identification of potato and rice crops in India. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 254–262. doi:10.1016/S0924-2716(99)00020-9.
- 695 Patel, N. K., Medhavy, T. T., Patnaik, C., & Hussain, a. (1995). Multi temporal ERS-1 SAR data for identification of rice crop. *Journal of the Indian Society of Remote Sensing*, 23, 33–39. doi:10.1007/BF03007970.

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830. URL: <http://dl.acm.org/citation.cfm?id=1953048.2078195>.  
700
- Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11, 1633–1644. URL: <http://www.hydrol-earth-syst-sci.net/11/1633/2007/>. doi:10.5194/hess-11-1633-2007.
- 705 Pei, Z., Zhang, S., Guo, L., McNairn, H., Shang, J., & Jiao, X. (2011). Rice identification and change detection using TerraSAR-X data. *Canadian Journal of Remote Sensing*, 37, 151–156. URL: <http://www.tandfonline.com/doi/abs/10.5589/m11-025>. doi:10.5589/m11-025.
- Phan, H., Le Toan, T., Bouvet, A., Nguyen, L., Pham Duy, T., & Zribi, M. (2018). Mapping of Rice Varieties and Sowing Date Using X-Band SAR Data. *Sensors*, 18, 316. URL: <http://www.mdpi.com/1424-8220/18/1/316>. doi:10.3390/s18010316.  
710
- Renaud, F. G., Syvitski, J. P., Sebesvari, Z., Werners, S. E., Kremer, H., Kuenzer, C., Ramesh, R., Jeuken, A., & Friedrich, J. (2013). Tipping from the Holocene to the Anthropocene: How threatened are major world deltas? *Current Opinion in Environmental Sustainability*, 5, 644–654. URL: <http://linkinghub.elsevier.com/retrieve/pii/S1877343513001553>. doi:10.1016/j.cosust.2013.11.007.  
715
- Ribbes, F. (1999). Rice field mapping and monitoring with RADARSAT data. *International Journal of Remote Sensing*, 20, 745–765. URL: <http://www.tandfonline.com/doi/abs/10.1080/014311699213172>. doi:10.1080/014311699213172.
- 720 Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP Climate Forecast System Version 2. *Journal of Climate*, 27, 2185–2208. URL: <http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00823.1>. doi:10.1175/JCLI-D-12-00823.1.
- 725 Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., & Yokozawa, M. (2007). Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environment*, 109, 295–313. doi:10.1016/j.rse.2007.01.011.
- Sakamoto, T., Van Nguyen, N., Ohno, H., Ishitsuka, N., & Yokozawa, M. (2006). Spatio-temporal distribution of rice phenology and cropping systems in the Mekong Delta with special reference to the seasonal water flow of the Mekong and Bassac rivers. *Remote Sensing of Environment*, 100, 1–16. doi:10.1016/j.rse.2005.09.007.  
730
- Salas, W., Boles, S., Li, C., Yeluripati, J. B., Xiao, X., Frolking, S., & Green, P. (2007). Mapping and modelling of greenhouse gas emissions from rice paddies with satellite radar observations and the DNDC biogeochemical model. *Aquatic Conservation*:  
735

- Marine and Freshwater Ecosystems, 17, 319–329. URL: <http://doi.wiley.com/10.1002/aqc.837>. doi:10.1002/aqc.837.
- Sentinel-1 Team (2013). Sentinel-1 User Handbook. Technical Report European Space Agency. URL: [https://sentinel.esa.int/documents/247904/685163/Sentinel-1\\_User\\_Handbook](https://sentinel.esa.int/documents/247904/685163/Sentinel-1_User_Handbook).  
740
- Sentinel-1 Team (2018). Sentinel-1 Mission Status Report 191. Technical Report European Space Agency. URL: [https://sentinel.esa.int/documents/247904/3339966/Sentinel-1-Mission\\_Status\\_Report\\_191-Period-30\\_Jan-05\\_Feb\\_2018.pdf](https://sentinel.esa.int/documents/247904/3339966/Sentinel-1-Mission_Status_Report_191-Period-30_Jan-05_Feb_2018.pdf).
- Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., Brown, R., & Staples, G. (2001). Rice monitoring and production estimation using multitemporal RADARSAT. Remote Sensing of Environment, 76, 310–325. doi:10.1016/S0034-4257(00)00212-1.  
745
- Shen, S., Yang, S., Li, B., Tan, B., Li, Z., & Le Toan, T. (2009). A scheme for regional rice yield estimation using ENVISAT ASAR data. Science in China Series D: Earth Sciences, 52, 1183–1194. URL: <http://link.springer.com/10.1007/s11430-009-0094-z>.  
750 doi:10.1007/s11430-009-0094-z.
- Son, N. T., Chen, C. F., Chen, C. R., Duc, H. N., & Chang, L. Y. (2014). A phenology-based classification of time-series MODIS data for rice crop monitoring in Mekong Delta, Vietnam. Remote Sensing, 6, 135–156. doi:10.3390/rs6010135.
- Son, N. T., Chen, C. F., Chen, C. R., & Minh, V.-Q. (2017). Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. Geocarto International, (pp. 1–15). URL: <https://www.tandfonline.com/doi/full/10.1080/10106049.2017.1289555>. doi:10.1080/10106049.2017.1289555.  
755
- Thirumalai, K., DiNezio, P. N., Okumura, Y., & Deser, C. (2017). Extreme temperatures in Southeast Asia caused by El Niño and worsened by global warming. Nature Communications, 8, 15531. URL: <http://www.nature.com/doi/10.1038/ncomms15531>. doi:10.1038/ncomms15531.  
760
- Torbick, N., Chowdhury, D., Salas, W., & Qi, J. (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. Remote Sensing, 9, 119. doi:10.3390/rs9020119.
- United Nations (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Technical Report United Nations, Department of Economic and Social Affairs, Population Division New York.  
765
- Van Khanh Triet, N., Viet Dung, N., Fujii, H., Kumm, M., Merz, B., & Apel, H. (2017). Has dyke development in the Vietnamese Mekong Delta shifted flood hazard downstream? Hydrology and Earth System Sciences, 21, 3991–4010. doi:10.5194/hess-21-3991-2017.  
770
- Wang, L. F., Kong, J. A., Ding, K. H., Le Toan, T., Ribbes, F., & Floury, N. (2005). Electromagnetic Scattering Model for Rice Canopy Based on Monte Carlo Simulation. Progress In Electromagnetics Research, 52, 153–171. URL: <http://www.jpier.org/PIER/pier.php?paper=0408061>. doi:10.2528/PIER04080601.  
775

- Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J. Y., Salas, W., & Moore, B. (2006). Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment*, 100, 95–113. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425705003433>. doi:10.1016/j.rse.2005.10.004.
- 780 Yang, Z., Li, K., Shao, Y., Brisco, B., & Liu, L. (2016). Estimation of Paddy Rice Variables with a Modified Water Cloud Model and Improved Polarimetric Decomposition Using Multi-Temporal RADARSAT-2 Images. *Remote Sensing*, 8, 878. URL: <http://www.mdpi.com/2072-4292/8/10/878>. doi:10.3390/rs8100878.
- 785 Yonezawa, C., Negishi, M., Azuma, K., Watanabe, M., Ishitsuka, N., Ogawa, S., & Saito, G. (2012). Growth monitoring and classification of rice fields using multitemporal RADARSAT-2 full-polarimetric data. *International Journal of Remote Sensing*, 33, 5696–5711. doi:10.1080/01431161.2012.665194.
- 790 Zhang, X., Zhang, P., Shen, K., & Pei, Z. (2016). Rice identification at the early stage of the rice growth season with single fine quad Radarsat-2 data. In C. M. U. Neale, & A. Maltese (Eds.), *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII* (p. 99981J). Edinburgh: SPIE. URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2240154>. doi:10.1117/12.2240154.
- 795 Zhang, Y., Yang, B., Liu, X., & Wang, C. (2017). Estimation of rice grain yield from dual-polarization Radarsat-2 SAR data by integrating a rice canopy scattering model and a genetic algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 57, 75–85. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0303243416302124>. doi:10.1016/j.jag.2016.12.014.

## Appendix

800 The importance for each calculated time series feature and seasonal regression model is listed in table 3. The shortened name perc refers to the value of the percentile given in the number following it, diff\_perc is the difference between percentile values. lag refers to the time series lagged by the number of days given. stdDev stands for the standard deviation. fft\_imag and fft\_angle refer to the imaginary part and the angle of the fourier transformed time series of which the minimum (min), maximum (max) and mean have been calculated.

805



Table 3: Variables of importance for each feature and seasonal random forest regression model.

Feature	Winter-Spring	Summer-Autumn	Autumn-Winter
stdDev	0.013	0.016	0.005
mean	0.005	0.021	0.005
10perc	0.007	0.018	0.009
25perc	0.025	0.019	0.005
50perc	0.055	0.053	0.023
60perc	0.038	0.051	0.012
75perc	0.027	0.023	0.030
90perc	0.020	0.038	0.028
diff_90perc10perc	0.015	0.035	0.006
diff_75perc25perc	0.013	0.014	0.011
lag10_10perc	0.242	0.018	0.030
lag10_25perc	0.033	0.030	0.013
lag10_50perc	0.028	0.036	0.040
lag10_75perc	0.027	0.055	0.099
lag10_90perc	0.019	0.015	0.028
lag20_10perc	0.014	0.024	0.076
lag20_25perc	0.024	0.021	0.031
lag20_50perc	0.009	0.035	0.017
lag20_75perc	0.052	0.019	0.024
lag20_90perc	0.015	0.012	0.032
lag30_10perc	0.014	0.070	0.021
lag30_25perc	0.028	0.029	0.024
lag30_50perc	0.014	0.020	0.034
lag30_75perc	0.019	0.044	0.030
lag30_90perc	0.034	0.013	0.017
lag40_10perc	0.016	0.016	0.022
lag40_25perc	0.008	0.019	0.025
lag40_50perc	0.004	0.020	0.039
lag40_75perc	0.016	0.022	0.029
lag40_90perc	0.005	0.014	0.015
lag50_10perc	0.029	0.029	0.024
lag50_25perc	0.009	0.026	0.013
lag50_50perc	0.015	0.023	0.048
lag50_75perc	0.009	0.039	0.022
lag50_90perc	0.031	0.018	0.022
fft_angle_mean	0.000	0.003	0.000
fft_angle_min	0.046	0.021	0.047
fft_angle_max	0.000	0.000	0.000
fft_imag_mean	0.000	0.000	0.000
fft_imag_min	0.012	0.013	0.018
fft_imag_max	0.010	0.009	0.025