



**HAL**  
open science

**SENTINEL-1&2 FOR NEAR REAL TIME CROPPING  
PATTERN MONITORING IN DROUGHT PRONE AREAS.  
APPLICATION TO IRRIGATION WATER NEEDS IN  
TELANGANA, SOUTH-INDIA**

Sylvain Ferrant, A. Selles, Michel Le Page, Al Bitar Ahmad, S. Mermoz, Simon Gascoin, Alexandre Bouvet, S. Ahmed, Yann H. Kerr

► **To cite this version:**

Sylvain Ferrant, A. Selles, Michel Le Page, Al Bitar Ahmad, S. Mermoz, et al.. SENTINEL-1&2 FOR NEAR REAL TIME CROPPING PATTERN MONITORING IN DROUGHT PRONE AREAS. APPLICATION TO IRRIGATION WATER NEEDS IN TELANGANA, SOUTH-INDIA. ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, XLII-3/W6, pp.285-292. <10.5194/isprs-archives-XLII-3-W6-285-2019>. <hal-03279941>

**HAL Id: hal-03279941**

**<https://hal.science/hal-03279941v1>**

Submitted on 8 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

## SENTINEL-1&2 FOR NEAR REAL TIME CROPPING PATTERN MONITORING IN DROUGHT PRONE AREAS. APPLICATION TO IRRIGATION WATER NEEDS IN TELANGANA, SOUTH-INDIA

S. Ferrant<sup>1</sup>, A. Selles<sup>2,3</sup>, M. Le Page<sup>1</sup>, A. AlBitar<sup>1</sup>, S. Mermoz<sup>4</sup>, S. Gascoin<sup>1</sup>, A. Bouvet<sup>1</sup>, S. Ahmed<sup>2,3</sup>, Y. Kerr<sup>1</sup>

<sup>1</sup> Centre d'Etude Spatiale de la BIOSphère-Université Paul Sabatier-Centre National de la Recherche Scientifique, Institut de Recherche pour le Développement-Centre National d'Etudes Spatiales (CESBIO-UPS-CNRS-IRD-CNES), 18 av. Ed. Belin, 31401 Toulouse CEDEX 9, France

(sylvain.ferrant, michel.lepage, ahmad.albitar, simon.gascoin, alexandre.bouvet, yann.kerr)@cesbio.cnes.fr

<sup>2</sup> Bureau de Recherches Géologiques et Minières (BRGM), Université de Montpellier, 1039 rue de Pinville, 34000 Montpellier, France

<sup>3</sup> Indo French Center for Groundwater Research, Bureau de Recherches Géologiques et Minières (BRGM), National Geophysical Research Institute (NGRI), Uppal Road, Hyderabad 50007, India [a.selles@brgm.fr](mailto:a.selles@brgm.fr), [shakeelifcgr@gmail.com](mailto:shakeelifcgr@gmail.com)

<sup>4</sup> Global Earth Observation (GlobEO), av. St Exupery, 31400 Toulouse, [mermoz@globoe.net](mailto:mermoz@globoe.net)

### Commission VI, WG VI/4

**KEY WORDS:** Machine learning, rice detection, Sentinel data fusion, water detection, irrigated water needs estimates

### ABSTRACT:

Indian agriculture relies on monsoon rainfall and irrigation from surface and groundwater. The inter-annual variability of monsoon rainfalls is high, which forces South Indian farmers to adapt their irrigated area extents to local water availability. We are developing and testing an automatic methodology for monitoring spatio-temporal variations of irrigated crops in near real time based on Sentinel-1 and -2 data feed over the Telangana State, South India. These freely available radar and optical data are systematically acquired worldwide, over India since 2016, on a weekly basis. Their high spatial resolution (10–20 m) are well adapted to the small size field crops that is common in India. We have focused first on drought prone areas, North of Hyderabad. Crop fraction remains low and varies widely (from 10 to 60%, ISRO-NRSC, Bhuvan). Those upstream areas, mainly irrigated with groundwater, are composed by less than 20% of irrigated areas during the dry season (Rabi, December to March) and up to 60% of the surface is used for crop production during the Kharif (June to November), which includes rainfed cotton and drip irrigated maize crops and inundated rice. A machine learning algorithm, the Random Forest (RF) method, was automatically used over 6 growing seasons (January to March and July to November, from 2016 to 2018) from the Sentinel-1&2 data stacked for each season, to create crop mapping at 10m resolution over a study area located in the north of Hyderabad (210 by 110km). Six seasonal land cover field surveys were used to train and validate the classifier, with a specific effort on rice and maize field sampling. The lowest irrigated area extents were found for driest conditions in Rabi 2016 and Kharif 2016, accounting for 3.5 and 5% with moderate classification confusion. This confusion decreases with the increase of irrigated crops areas during Rabi 2017. For this season, 22% of rice and 9% of irrigated crops were detected after heavy rainfall events in September 2017, which have filled surface water tanks (3.4% of the surface area) and groundwater (Central Groundwater Board observations). From this abundance situation, the surface water detected for each season decreased regularly to less than 0.3% together with the rice and irrigated area extents respectively from 22 to 11% and 10 to 3%, despite a good monsoon 2017. Groundwater level show similar trends, with a drop from 20 meters depth in October 2016 and 2017 to more than 30 m in June 2018 (more recent available data). The deficit of the monsoon 2018 will certainly bring this situation to a hydrological drought at the beginning of 2019, probably similar to the Rabi 2016 situation. The estimated Irrigated Water Demand (IWD) varies from 51 to 310 mm/season, depending on water availability. This methodology shows the potential of automatically monitoring, in near real time, with standard computers, irrigated area extents presenting fast high resolution variability. As it is based on standard global satellite acquisitions, it is foreseen to be used for other regions, for any studies on farmer's adaptation to climate and hydrological variability, as a proxy to estimate irrigation water needs and water resources availability. In Telangana for instance, it provides an inventory of crop production and irrigation practices before the implementation of mega project infrastructures funded by this new state: - the Kâkâtiya tank restoration program to enhance monsoon runoff capture or the Kaleshwaram project to divert Godavari river water toward upstream Telangana region through tunnels and canals in 20 giant reservoirs.

## 1. INTRODUCTION

### 1.1 General context

To be efficient, environmental decisions from policy makers should be based on evidence and facts, being used to build fair and efficient environmental policies. This objective supposes to understand complex co-evolution of human activities and climatic variability and environmental resources, which implies to multiply environmental observations of continental surface dynamics. For agro-systems, quantifying and predicting supposes to build numerical representation of the cascade of processes and retroaction between climate, soil, water resources

and agricultural practices: cultivation of land, fertilization, irrigation. Agro-hydrological models aims at representing those retroactions in space and time using numerical modelling approaches. These models represent the main processes involved in the crop growth and water and nutrient transfer within catchments: soil water and nutrient uptake, drainage and transfers to aquifer and rivers, as a function of crop calendar, fertilisation, irrigation, etc (Arnold et al. 1998; Breuer, Vaché, et Julich 2008; Engel et al. 1993; Refsgaard et al. 1999; Perrin et al. 2012; S. Ferrant 2009). These modelling objectives require spatially distributed models, where information of soil-crop location within slopes as well as hydrological settings (topography, groundwater storage, reservoir location or

irrigation pumping) are included to provide spatially explicit information on water uses: fertilisation and crop growth parameters (S. Ferrant et al. 2011), soil and groundwater stock and irrigation pumping for instance (Perrin et al. 2012).

It pictures the situation of water resources under human activities, test the efficiency of designed water/agricultural management options (Ferrant et al. 2013) and is used to project the nowadays water consumption under future climate or socio-economical situations (Ferrant et al. 2014).

### 1.2 Limitations: the lack of spatio-temporal variables

These spatially distributed agro-hydrological models are simulating a cascade of many processes, involving numerous parameters to represent the spatial and temporal dynamic of water, nutrient cycle, crop growth, which are calibrated using available observations: water and nutrients stream fluxes at the outlet, average yield (Ferrant et al. 2011; Moreau et al. 2013), aquifer recharge (Perrin et al. 2012) or satellite based evapotranspiration with MODIS at 500 meters (Cheema, Immerzeel, et Bastiaanssen 2014). The lack of observation of soil, crop, and climatic variables leads to a lack of spatial processes calibration which may result in equifinality problems, e.g. more than one parameter set leading to similar results (Beven 2001) or compensation between processes leading to similar stream water fluxes (Ferrant et al. 2011). Uncertainties raised by these modelling approaches at the watershed level are mainly related to the lack of agricultural operations data and soil-crop conditions encountered within the catchment, i.e. saturated conditions within slopes and their feedback on crop productivity. For instance, the timing of fields operations (first fertilization and sowing date) detected by monitoring crop growth using satellite derived Leaf Area Index, has a strong impact on the winter wheat growth simulated in temperate area (Ferrant et al., 2014), with a non negligible impact on soil water shortage estimates in summer (Martin et al., 2016). These crop growth dynamics are crucial input parameters to accurately simulate crop growth dynamics and interaction with water resources (Ferrant et al., 2014).

### 1.3 Contribution of Earth observation to Indian agro-hydrology

The Indian space program provides several high-resolution satellite missions (1 to 56 meter spatial resolution) that have been used to measure seasonal land cover. In 2016, the National Information System for Climate and Environment Studies (NICES) project released a national database of seasonal net sown areas (<http://bhuvan.nrsc.gov.in/data/download/index.php>) which highlights the high spatio-temporal heterogeneity of land cover and its temporal dynamics. This dataset is published as a fraction of crop sown at 5km resolution by the National Remote Sensing Center (NRSC), part of the Indian Space Research Organization (ISRO). A similar program, called Water Bodies Fraction (WBF) and based on the same optical data, produces the areal fraction of surface water every 15 days since 2012. The short-term variations of sown areas show a high temporal variation over recent decades in the Hyderabad region, with 41 to 58% of sown areas in Kharif and 8 to 22% in Rabi between 2005 and 2016. However, the irrigated fraction is an essential environmental variable to be retrieved for water management issues.

The recently launched Sentinel satellites are especially of interest for their acquisition strategies. Sentinel-1 (satellite S1A, launched in 2014 and S1B in 2016) and Sentinel-2 (S2A launched in 2015 and S2B in 2017) are the first generation of operational satellite EO missions for both optical multi-spectral

and radar C-band detection of continental surfaces at a global scale, with high spatial and temporal resolutions (10 to 20 m; 5 to 10 days revisit interval), under a free access licence.

### 1.4 Objectives of the study over Hyderabad region

The aim of this paper is to present the methodology, based only on those new observations from the Sentinel-1 and Sentinel-2 satellites for monitoring the essential agro-hydrological variables of the Indian context. It presents an update of a previous work where synergy and spatial and spectral resolution of Sentinel-1&2 products were evaluated regarding the irrigated crop detection accuracy for 3 growing seasons 2016 and 2017 (Ferrant et al. 2017). Following the conclusions of this study, we used the only 10m resolution Sentinel-1&2 bands and ratios. The main difference in this study is the use of surface reflectance Sentinel-2 produced from the MAJA workflow (MACCS ATCOR joint Algorithm, O. Hagolle, [https://github.com/olivierhagolle/Start\\_maja](https://github.com/olivierhagolle/Start_maja)) which produces accurate cloud mask based on image time series. An automated workflow, presented in this manuscript, used those masks to reject Sentinel-2 images based of haze and cloud presence.

The algorithm has been developed for standard computers to be used by lay users and deliver near real time restitution of Essential Environmental Variables (EEVs) for each season: 1- inundated rice areas, 2- irrigated crops areas, 3- rainfed crop areas, 4- surface water areas. These four variables involve only a few percent of the total area and vary rapidly from season to years with water availability. The seasonal land covers produced in this study have been used to estimate the seasonal Irrigated Water Demand (IWD), based on irrigation practices observed in previous studies (Dewandel et al. 2010; Ferrant et al. 2014; Perrin et al. 2012; Dewandel et al. 2017). This study considers one spatial resolution, 10 meters, and explores the benefit of the synergy between Sentinel-1's radar backscatter advantages (not sensitive to the persistent cloud cover during the Kharif season, highly impacted by crop growth and surface water) and the Sentinel-2's multi-spectral detection, both at an appropriate spatial resolution to the small extent of crop field and water bodies. EEVs variations are discussed regarding contextual data (rainfall and groundwater level) to discuss their interest in a better understanding of farmer's adaptation to climatic and hydrologic variability. We have chosen the Random Forest algorithm (RF) provided with the Orfeo ToolBox (OTB) because it is a fast, open-source processor of high-resolution optical, multispectral and radar images at the terabyte scale. Highly appropriate for building a processing chain, it is suitable for automatically producing agro-hydrological variables at a large scale (Pelletier et al. 2016).

## 2. MATERIAL AND METHODS

### 2.1 Study area

The Kudaliar river catchment (983 km<sup>2</sup>) is shown in Figure 1 and 2. It is located in Telangana State, South India. Figure 1 shows the fraction crop cover in India for two seasons, Kharif 2015 (Figure 1a) and Rabi 2016 (Figure 1b). The catchment is 50 km north of Hyderabad (Figure 1c). This catchment is an upstream watershed unconnected to the perennial regional rivers and is representative of the northern part of the Deccan plateau region in terms of irrigated agriculture, climatic variability and hydrogeological settings. It is characterized by a flat topography (from 430 to 640 meters above sea level) and an absence of perennial streams. The region has a semi-arid climate controlled by the periodicity of the Southwest monsoon: a rainy "Kharif" (local name) season from June to November and a dry "Rabi"

(local name) season from December to March. In contrast with these two seasons, the summer season (April and May) is mainly free of cultivation. Annual precipitation (1980–2000) ranges from 540 mm to 1300 mm with a mean of 879 mm according to the Indian Meteorological Department, 88% of the rain falls during the monsoon season. The annual mean temperature is 26°C, although in summer (April to May) the maximum temperature may reach 45°C.

The catchment is mainly rural and densely populated (300,000 inhabitants in 983 km<sup>2</sup>), with only one small town (Gajwel, 30,000 inhabitants), located in its southwestern portion (see Figure 1c, purple area). Approximately 60% of the watershed is covered by semi-arid natural vegetation (bush). Rain-fed crops (mainly maize and cotton) are sown only during the monsoon period (Kharif, Figure 1a). The irrigated area varies widely with climatic and hydrological conditions since irrigation mainly depends on groundwater availability. The main aquifer is composed by a weathered and fractured Archean granitic unit. Thus, the groundwater availability is controlled by the thickness of the weathered profile and the density of fractures (Boisson et al. 2015). This shallow heterogeneous aquifer is characterized by strong and fast water table fluctuations caused by recharge/pumping cycles (Maréchal et al. 2006). A Water Harvesting System (WHS) consisting of many small dams in a mound of soil materials along the drainage network is maintained by the farmers. They are generally non-permanent lakes: they can be refilled by heavy monsoon rainfall and dry out in a few months through evaporation, infiltration and use for irrigation (Massuel et al. 2014). This area will be used to compute surface area extents from seasonal land covers.

## 2.2 Sentinel-1 and 2 data sets

Sentinel-1 and 2 acquisitions over India started at the end of 2015. They were stacked into a single file for each six growing season to be analyzed by the RF algorithm. The study area is embedded in two Sentinel-2 tiles (44QKE and 44QKF) located north of Hyderabad (Figure 1, orange).

## 2.3 Sentinel-1 and 2 pre-processing

The Sentinel-2 dataset downloaded in this study is the surface reflectance produced by PEPS web platform and Theia land data center (Level-2A). It provides a cloud mask and a surface reflectance corrected from the atmospheric effects, based on the MAJA workflow (Hagolle et al. 2015). In this semi-arid Indian area, this correction is crucial as the seasonal variations of aerosol concentrations in the form of dry dust is high (Sivaprasad et Babu 2014) and alters the reflectance.

The Sentinel-1 dataset comprises Level-1 Ground Range Detected (GRD) data in Interferometric Wide swath mode (IW) consisting of focused SAR data that has been multi-looked and projected to ground range using the WGS84 Earth ellipsoid model. The resulting images in dual polarization (VH and VV) have a dimension of 270x270km and a resolution of 10 m. The terrain correction was then applied to geocode the images by correcting SAR geometric distortions (foreshortening, layover and shadow) using the digital elevation model from the Shuttle Radar Topography Mission, producing ground-projected images. To do this, the Orfeo ToolBox (OTB) was used. Each image was cropped and superposed on the Sentinel-2 tiles. A multi-temporal speckle filter (Bruniquel et Lopes 1997) was preferred to the more classical spatial filter in order to preserve spatial resolution and the fine structure of Sentinel-1 images. As shown previously in (Mermoz et Le Toan 2016; Mermoz et al. 2014), this method produces images with reduced speckle effects from the whole sentinel-1 acquisition time series (30

dates since February 2015) and multi-polarized (VH and VV) images. The speckle filter is expressed as follows:

$$J_k(v) = \frac{\langle I_k(v) \rangle}{N} \sum_{i=1}^N \frac{I_i(v)}{\langle I_i(v) \rangle} \quad \text{with } k=1, \dots, N \quad (\text{eq. 1})$$

where  $J_k(v)$  is the radar intensity of the output image  $k$  at pixel position  $v$ ,  $I_i(v)$  is the radar intensity of the input image  $i$ ,  $\langle I_i(v) \rangle$  is the local average intensity of the input image  $k$  (window size of 5x5) and  $N$  is the number of images ( $N=30$  dates \* 2 polarizations). The resulting theoretical number of looks is 71.

## 2.4 Sentinel Bands and derived indexes selection

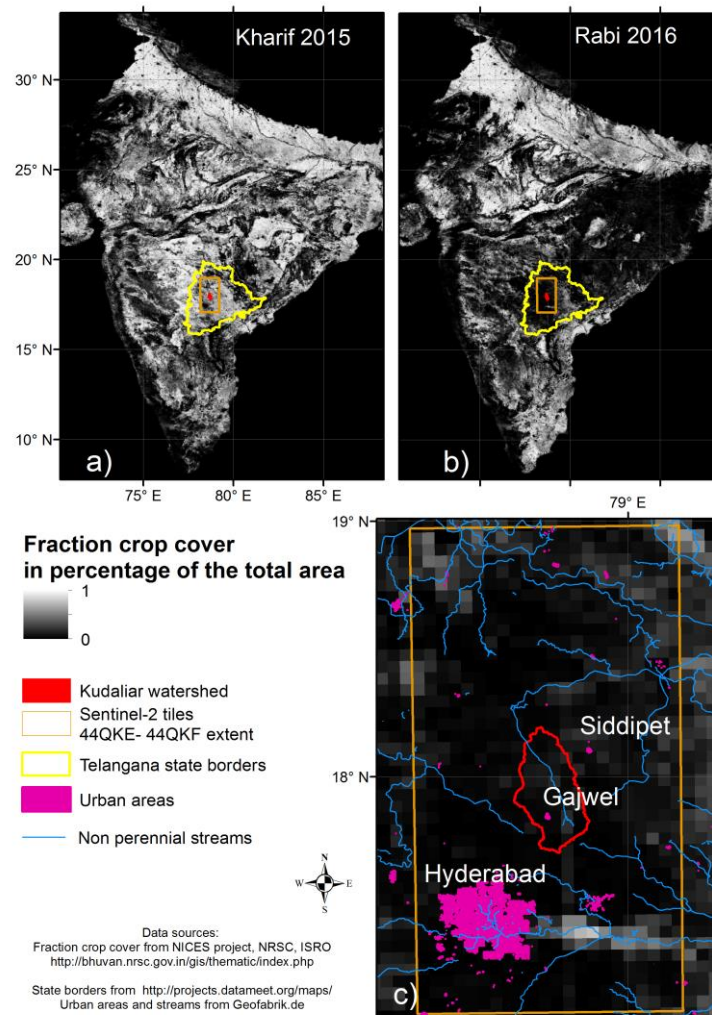
We selected Sentinel-2 bands at 10 m: the blue ( $B_2 = 490$  nm), green ( $B_3 = 560$  nm), red ( $B_4 = 665$  nm), and near infra-red ( $B_8 = 842$  nm) bands and have computed the NDVI for each images (Eq. 2). VV and VH backscatters from Sentinel-1 IW acquisition mode and the ratio based on VV and VH (Eq. 3) were computed at 10 meter resolution and both Sentinel bands and indexes were stacked together by growing seasons. The sizes of these stacks of data obviously vary with the number of images in a season.

$$NDVI_{10m} = \frac{B_8 - B_4}{B_8 + B_4} \quad (\text{Eq. 2}) \quad \text{VV} / \text{VH}_{10m} = \frac{VV}{VH} \quad (\text{Eq. 3})$$

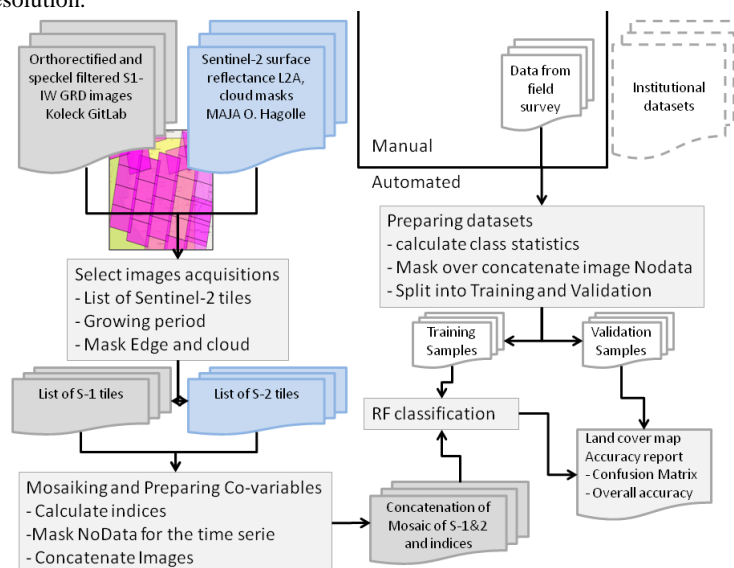
Acquisition dates were selected using a number threshold of valid pixels, based on merge edge mask and cloud mask. Each acquisition dates which are not covering 90% of the pixel zone are not used for the classification. The number of bands and sampling dates varies highly from a season to another, with both the cloud mask between Kharif and Rabi seasons, and with the increase of Sentinel acquisitions from 2016 to now. The workflow for automated selection of data and supervised classification is described in Figure 2.

## 2.5 Land cover surveys for learning samples

Six seasonal surveys of land cover were carried out in February and October, from 2016 to 2018, corresponding to 428 plots of non-irrigated areas: bare ground or natural vegetation cover, rainfed and tree plantations (eucalyptus and teak), 1455 irrigated areas: orchards (mainly mango trees), maize crops and vegetables (beans, lentils) and 192 flooded rice crop areas. The dataset was especially scrutinized for rice as the main water-consuming crop. A USB GPS receiver (G-STAR IV) was connected to GIS software (QGIS) on a laptop to position the operator in the landscape. The operator used the latest Sentinel-2 acquisition to delineate polygons containing both homogeneous land cover and reflectance observed in the field and on the Sentinel-2 image. The dataset created from the field surveys was used to create six seasonal datasets: perennial land cover such as orchard, forest and natural land use are preserved across the three datasets together with the irrigated, flooded and rain-fed crop fields observed during one season. These reference data comprise nine land cover classes: inundated rice, irrigated vegetables, irrigated maize, orchards, forested area, bare ground and natural bushes, urban areas, surface water and rain-fed cotton (only in Kharif). Surface water areas are those always covered by water during the season. Each seasonal data set is then split into two separate training and validation sets of polygons. Seasons are named following the ISRO labelling convention adopted for NICE programs: Rabi 2016 (Jan-Mar 2016), Kharif 2016 (Aug-Nov 2016), respectively R2016, K2016, R2017, K2017, R2018, K2018.



**Figure 1.** Location of the study area in Telangana State, South India. The Kudaliar watershed area is embedded in two Sentinel-2 tiles (orange outline). (a) Fraction crop cover in wet Kharif season 2015 (July to November); (b) Fraction crop cover in dry Rabi season 2016 (December 2015 to March 2016); (c) Fraction crop cover for the study area in Rabi 2016. This product is provided at 5 km resolution from optical remote sensing at 56 m resolution.



**Figure 2.** Automatic seasonal land cover workflow.

## 2.6 Irrigated crop classification methods and strategy

The RF algorithm is a supervised classification method based on ensemble learning theory (Breiman, 2001). It has been widely used to identify land cover using one or more satellite images (Pelletier et al. 2016). The algorithm builds binary decision trees that are decorrelated by using bootstrap samples. The split process also decorrelates the various trees by first randomly selecting at each node a subset of input variables (also called features), and by then performing a variable value test based on the Gini impurity criterion or entropy loss. Ideally, this process is repeated recursively on each derived subset until the node contains very similar samples, or when the splitting no longer adds value to the predictions. During the decision phase, each tree classifies the input data. Then, a majority vote is performed to output the class label.

The RF algorithm requires two main parameters to tune: 1) the number of trees  $K$ , and 2) the number of variables randomly selected to split a node  $mtry$ . In this study, classical parameter values are used:  $K = 100$ ;  $mtry = 25$ .

All the pixels located in the training and validation polygons are used to respectively train the RF learning algorithm and to assess the accuracy of the resulting land-cover maps. The unit of assessment is therefore the pixel for the validation procedure, the number varying with the cropfield polygon, from tens to thousands of pixels. Non-agricultural areas cover 60 to 70% of these rural areas and are sampled using large polygons, implying thousands of learning and validation pixels, whereas rice and irrigated areas consist of hundreds of small polygons with limited numbers of pixels in each.

The accuracy of the land cover map produced is evaluated with the precision, recall and F-score of the targeted irrigated and inundated crop class from the confusion matrices. F-score is the harmonic mean of the user's accuracies and the producer's accuracies. The best value of the F-score is 1, and the worst 0. We use the inundated and irrigated area estimates to compare the methods and seasons. The overall land-cover accuracy is given by the Kappa index for all classes. The size of each satellite time series dataset analyzed is also given, to illustrate the basic informatics settings required by this method.

## 2.7 Estimation of Irrigated Water Demand (IWD)

The water demand corresponding to the inundated and irrigated area extents is estimated from previous estimates based on field work and the DST-GW model developed and calibrated in this area (Dewandel et al. 2010) which were used in the agro-hydrological modeling study (Ferrant et al. 2014). Daily rice and vegetable irrigation rates were estimated from field surveys (measurements of instant well discharges, irrigated field areas, and monitoring of daily pumping durations). Daily irrigation rates for rice are 12 mm/day and 9 mm/day for Rabi and Kharif respectively (Massuel et al. 2008). The IWD will be expressed in mm/season as a water thickness divided by the watershed area, same unit used for rainfall and aquifer recharge.

# 3. RESULTS AND DISCUSSION

## 3.1 Random Forest crop detection accuracy

The scores of the classification are presented in Table 1, for irrigated crops and cotton. As expected, rice detection using Sentinel-2 for Rabi seasons (cloud free period) yields high F-

Scores (Fscores from 0.77 to 0.96) and intermediate F-scores in Kharif (0.35 to 0.83). Very lower F-scores are found for others irrigated crops (maize and vegetables, from 0.03 to 0.68), depending, but not only, on the number of valid acquisition dates. Indeed, the number of Sentinel acquisitions progressively increased over the world from 2016 to now, with the launch of twin satellites. But specific period, if sampled with a Sentinel-2 acquisition, are determinant to avoid confusion between irrigated classes and improve the detection accuracy, especially during Kharif period.

Nevertheless, the K2017 shows really low F-Scores for irrigated crops and high confusions between rice and other crops. As for other Kharif period, only one cloud free Sentinel-2 is available, yet in 2017, this image comes very late in the season, at the end of November. Sentinel-1 time series, used alone for this season, gives the same low F-scores. For this season, Sentinel acquisitions do not sample properly the cropping pattern reflected by the seasonal field data set. On the contrary, the relatively high Kappa indices (from 0.58 to 0.92) are consistent with a good accuracy over natural land use, urban areas and surface water. For instance, Water F-scores are above 0.9 for all seasons, even K2017.

## 3.2 Seasonal variability of inundated and irrigated areas

Table 2 shows the areas estimated for each classification. Rice areas are varying from around 5 to 23% in Rabi and 4 to 11% in Kharif. This is consistent with previous study over this study area (Ferrant et al. 2017, 2014). The high uncertainty behind the K2017 estimates, with a low precision and high confusion between rice and other crops, raise doubts about the high rice area extent estimated (More than 40%). The same classification using only Sentinel-1 time series provides an estimate of 34% of rice irrigated area, with similar classification accuracy (results not shown). We shall notice that the success of irrigated area retrieval is highly dependent on an existing cloud free Sentinel-2 observation within the crop growth period. Another possible cause of this result is the high heterogeneity of rice stages found during the field trip, compare to the two other Kharif. Random forest classifier cannot handle so many heterogeneities for one class with only one optical image at the end of the crop growth period.

## 3.3 Seasonal hydrological variables

As part of landuse classes, surface water areas are retrieved from this seasonal land cover classification. They vary from 0.07 to 3.4% of the catchment area (table 2), distributed in hundreds of small tanks spread along the river network. Table 2 shows cumulative seasonal rainfall measured in the north of Hyderabad, with higher rainfall accumulation than average monsoon for this area (Patancherru, source: ICRISAT). R2016 meets the driest hydrological conditions, with less than 0.1% of surface water in the Kudaliar watershed, consisting in mainly small remainders of muddy pool found in biggest tanks of the area. Aquifer water levels were low (around 35 meters depth) since the previous monsoon rainfalls were low (600 and 700 mm for 2014 and 2015, respectively). The 2016 monsoon came late so July and August were still dry, but there was extremely heavy rainfall in September 2016. Hyderabad received 425.2 mm of rain compared with its monthly mean of 132.9 mm, surpassing the decade's highest rainfall of 266.6 mm (Sept., 2007).

In January and July 2016, farmers have had to restrain rice area extents to a minimum (less than 5%), to face this groundwater

and surface water shortage context. Irrigated rice paddies were located around borewell that were still providing enough irrigated water (around 60 to 50 mm of water at the catchment scale).

Class	Sensor	Sentinel-1&2					
	Season	Rabi 2016	Kharif 2016	Rabi 2017	Kharif 2017	Rabi 2018	Kharif 2018
Flooded Rice	Precision	0.87	0.98	0.96	0.29	0.95	0.95
	Recall	0.95	0.51	0.96	0.44	0.65	0.73
	F-score	0.90	0.68	0.96	0.35	0.77	0.83
Irrigated Maize Vegetables	Precision	0.22	0.45	0.59	0.1	0.71	0.31
	Recall	0.11	0.17	0.74	0.16	0.65	0.15
	F-score	0.15	0.03	0.66	0.12	0.68	0.2
Rainfed Cotton	Precision	-	0.84	-	0.24	-	0.66
	Recall	-	0.72	-	0.28	-	0.88
	F-score	-	0.78	-	0.26	-	0.76
Overall	Kappa	0.82	0.58	0.88	0.84	0.92	0.87
Technical Details	Nb dates	5 S1, 2 S2	12 S1, 1 S2	10 S1, 4 S2	17 S1, 1 S2	13 S1, 9 S2	14 S1, 1 S2
	Nb bands	25	41	50	56	84	47
	Size (Go) for 2 tiles	23	37.8	46.1	51.6	77.4	43.3

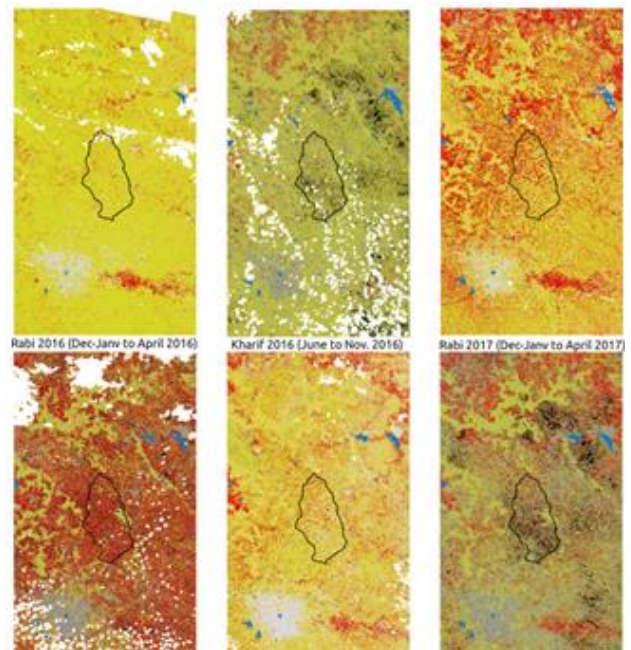
**Table 1.** Results of Random Forest classification using multi-temporal Sentinel-1 and 2 imagery at 10 meters, for five growing seasons: Rabi 2016 (Jan. to March 2016), Kharif 2016 (July to Nov. 2016), Rabi and Kharif 2017 and 2018. Precision and recall respectively stand for the predicted fraction of pixels classified as the appropriate class and the fraction of validation pixels that have been correctly classified. The F-score is the harmonic mean of both metrics and evaluates the relevance of the classifier. The Kappa index is an estimation of this relevance for all nine crop classes considered: rice, vegetable, maize, forest, orchard, bare ground dry areas or bush, urban area, water and cotton (if any). Technical details on the size of images stacks are detailed. Kharif 2017 column is colored in grey, to highlight the low F-scores obtained, and mark the low confidence we have on the irrigated area extents and corresponding IWD estimates.

Season	Rabi 2016	Kharif 2016	Rabi 2017	Kharif 2017	Rabi 2018	Kharif 2018
Flooded (% - ha)	4.95 – 4770	3.75 – 3565	22.73 – 22418	43.7 – 42679	11.2 – 11088	10.9 – 10761
irrigated	1.48 – 1434	2.26 – 2150	9.7 – 9607	3.9 – 3801	11.85 – 11690	3.14 – 30103
Cotton		11.6 – 11053		15.37 – 15004		21.9 – 21684
Seasonal permanent Water surface (% - ha)	0.07 – 73	3.4 – 3234	1.4 – 1413	1 – 977	0.36 – 361	0.27 – 267
Water Demand (Rice - others) mm	62.8 – 3.1	46.9 – 4.7	295.4 – 21.1	562.4 – 8.34	146.1 – 25.6	141.8 – 6.8
Seasonal Rainfall (mm)	207.5	1031.4	18.4	1108.5	47.2	578

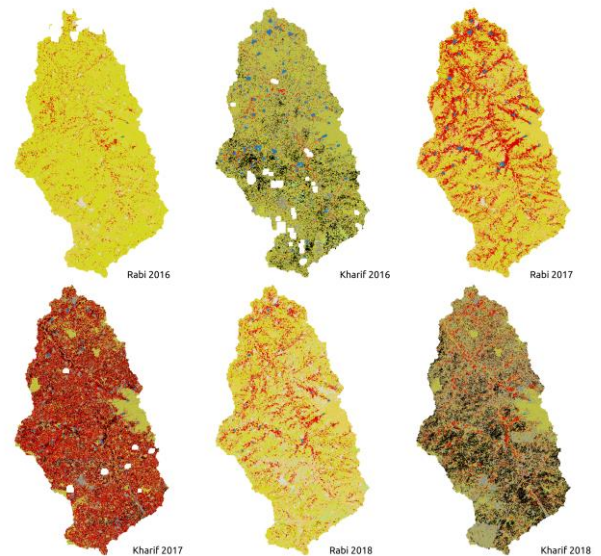
Source: Icrisat

**Table 2.** Areas of irrigated and flooded crops and corresponding Irrigated Water Demand (IWD) inside the watershed (around 1000km<sup>2</sup>). Crop areas are expressed both in % of the total watershed area and in ha. IWD are expressed in mm/season, the corresponding water total over the watershed area. This total should be compared to rainfall, also expressed in mm. Surface water is expressed in percent of the total area.

a)



b)



**Figure 3:** Sentinel-2 Tiles 44QKE and 44QKF Land cover (a) and extracted for the Kudaliar watershed (b), produced for 6 seasons from the workflow presented in figure 2. The black polygon is the watershed where all the statistics have been computed. White areas are nodata where clouds were detected. Hyderabad city is in the south-west corner, rice command area of the Musi river is located at the east of the city. Karimnagar dam is in the north east corner.

The situation changed in R2017, after the heavy rainfall of September. The surface water harvesting system is full, surface water have reached 3.4% of the catchment area, decreasing to 1.4% at the end of April 2017 with the cumulated effect of ground water recharge, evaporation and pumping for irrigation. Rice area reaches the highest extent : 22.73%, table 2, vs 17.11% estimated in Ferrant et al. (2017). Rice paddies are mainly located downstream of the tanks, benefiting from this surface water abundance.

The following monsoon season was abundant, with 1108mm of rainfall, but lower rainfall intensities than the previous monsoon. As a consequence, surface water reservoirs were not

fully refilled, the surface water area reached 1% only, after being reduced from 3.4 to 1.4% since the two previous monsoons. Contrary to the accurate estimates of surface water, confidence in irrigated crop area estimates is low and rice inundated area extent are certainly overestimated, so is the IWD. The next R2018 is composed by 11.2% of rice and 11.8% of irrigated maize and vegetables with limited surface water resources (0.36%). Both irrigated crops are mainly fed with groundwater. With a drop of rice inundated area by 2 compare to R2017, IWD is divided by 2, as an adaptation to the limitation of water availability. This trend brings irrigated crop area to a drastic drop in the next Kharif season, together with a drop of surface water areas. On the contrary, rice irrigated areas are maintained at the same level so as the associated IWD. These three area extents remain higher than those observed in the driest season in R2016, indicating that groundwater is still available. Rice area extent, as the prioritized crop, remains stable, but will start to drop by the next dry season 2019, with the drop of groundwater resources.

### 3.4 Essential Agro-Hydrological variables co-variation

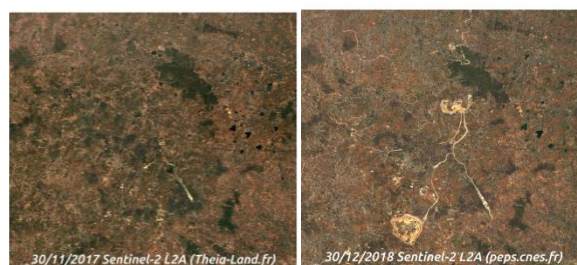
This application of automatic restitution of irrigated/inundated crop areas and surface water fluctuations within heterogeneous agro-systems illustrates the innovative use of emergent combined tools: exploration of big data from high spatial and temporal resolution remote sensing using cloud computing architecture and machine learning based scripts. It is used here to explore the co-variation of climate variables with hydrological and agricultural related variables. It is a way to quantify the short term adaptation of farmers to water availability or the feedback of drought on agriculture over large region.

In this application, we can highlight the farmer memory effect based on groundwater scarcity. The dry situation in July 2016 forced farmers to keep their rice areas to a minimum, to match low availability of water at the time of sowing, and lacking foreknowledge of the heavy rains that occurred the following September. Farmers' decision to massively sow flooded rice and drip-irrigated corn for the R2017 is explained by the abundance of surface and groundwater and a need to recover from low revenue. From this situation of abundance in January 2017 to the end of 2018, surface and groundwater drop down regularly, followed by the rice IWD which decreased from 295 mm (between 208 to 338mm estimated in previous Ferrant et al., 2017) to 146 mm in K2018.

This temporality of the water resources status is fundamental in farmers' decisions regarding the use of irrigated water. A 11 meters drop of groundwater levels, measured by the Central Groundwater Board measurements, occurred between Oct. 2016 and June 2018, reaching 32 meters depth. The low recharge that should have occurred with the rainfall deficit of 2018 monsoon will probably bring those level to the January-March 2016 situation (39 meters depth), in the fractured layers of the crystalline aquifer (Dewandel et al. 2017). In those upstream areas, chosen for not being part of surface water irrigation scheme, irrigation is highly hinging on groundwater and small agricultural surface water. The accurate restitution of rice area extent is also a good proxy of groundwater abundance. This is not the case nearby the Musi River, downstream Hyderabad area for instance, where rice is flooded with sewage water from the river twice a year (Figure 3a).

### 3.5 Other land cover changes detected

Two mega structures in circle shape are observed since 2017 in the Kudaliar river basin surrounding Gajwel town (10 and 17 km<sup>2</sup>, Figure 4). These two future lakes (named Konda Pochoma) are part of the 10 Billions euros Kaleschwaram project. It aims at pumping and bringing Godavari water located 160km away from these upstream areas to a network of big canals up along with 20 mega reservoirs. The construction site consists at both digging the ground for about 20m inside the circle delineated by walls and building this few hundred meter large and 30m tall wall using extracted saprolite, consolidated on the basement with granite rocks, impermeability made with a vertical layer of Black soils 50 to 100 meters thick located in the middle of the walls from the basement to the top. Big canals connecting those reservoirs are dug to bring water from Kaleschwaram town, located 200km north of Gajwel, on the bank of the Godavari River, where the river water will be pumped. Billions of cubic meters will be routed to the North Telangana to dispatch Godavari water into rural areas, for irrigated water, and the Hyderabad urban areas to fulfil its exponential growing domestic water needs. This big change will probably bring deep mutation in local farmer's practices, by certainly increasing the rice cultivation as observed in command irrigated areas like Karimnagar dam (Figure 3a, North east).



**Figure 4:** Konda Pochoma reservoir under construction, Gajwel Town, seen with Sentinel-2 satellites. Big reservoirs connected by canals. Highway surrounding Gajwel town were constructed in 2018.

## 4. CONCLUSION

Flooded rice areas and surface water dynamics are both essential variables in predicting groundwater usage in a small South Indian watershed. An automated workflow, presented in this manuscript, used level-2A Sentinel-2 images with masks of haze and cloud using MAJA algorithm (downloaded in [www.theia.cnes.fr](http://www.theia.cnes.fr)). Less Sentinel-2 were used in comparison with the previous work published over the same area for the 3 first growing seasons (Ferrant et al., 2017). Differences between both studies for those 3 seasons are low compare to the uncertainty of EEVs retrieval presented in Ferrant et al., 2017. Small irrigated areas are detected during the dry conditions of R2016 and K2016, with around 5% of irrigated areas, with moderate classification confusion. A large increase of irrigated areas (around 30%) is estimated with low classification confusion in R2017. This is explained by the replenishment of surface water resources during the 2016 monsoon. From this point, surface water and irrigated area extent, together with groundwater level, drop regularly from respectively 3 to 0.27% and from 20 to 13% in R2018 and K2018, despite a good monsoon 2017. The rainfall deficit of 2018 monsoon (10%) explains a drop in irrigated area extent and surface water extents, even if rice area remains stable at around 10%, the rice production being a priority for farmers. This high seasonal

variability represents the direct impact of the farmers' strategy of adaptation to water storage.

The automatic classification presented here shows a weakness for one season in K2017. Due to a lack of cloud free Sentinel-2 images, the radar data from Sentinel-1, together with a training sample extracted from field work was not enough descriptive to obtain satisfactory classification results. Other strategy will be explored, such as splitting rice samples into different growing stages, and rice detection methods based on radar backscatter threshold (Bouvet, Le Toan, et Lam-Dao 2009; Phan et al. 2018) rather than machine learning algorithms.

#### ACKNOWLEDGEMENTS

This work was carried out in the context of the project "Agro-hydrology from Space," with the support of IRD and NGRI facilities. We thank Wajiddudin for his field expertise and the Indo-French Center for Groundwater Research (IFCGR) team for their help and good welcome. This work was possible thanks to Theia land data services infrastructures. We would like to thank T. Koleck, O. Hagolle for the tools developed for Sentinel download and pre-processing, as well as OrfeoToolBox developer team, used as essential bricks to build the workflow presented in this manuscript.

#### REFERENCES

- Arnold, J. G., R. Srinivasan, R. S. Mutiah, et J. R. Williams. 1998. « Large-area hydrologic modeling and assessment: Part I. Model development ». *Journal of American Water Resource Association* 34: 73–89.
- Beven, K. 2001. « Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology ». *Journal of Hydrology* 249 (1–4): 11–29.
- Boisson, A., N. Guihéneuf, J. Perrin, O. Bour, B. Dewandel, A. Dausse, M. Viossanges, S. Ahmed, et J. C. Maréchal. 2015. « Determining the vertical evolution of hydrodynamic parameters in weathered and fractured south Indian crystalline-rock aquifers: insights from a study on an instrumented site ». *Hydrogeology Journal* 23 (4): 757–773. <https://doi.org/10.1007/s10040-014-1226-x>.
- Bouvet, A., T. Le Toan, et N. Lam-Dao. 2009. « Monitoring of the rice cropping system in the Mekong delta using ENVISAT/ASAR dual polarization data. » *IEEE Transactions on Geoscience and Remote Sensing* 47 (2): 517–26.
- Breiman, Leo. 2001. « Random Forests ». *Machine Learning* 45 (1): 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Breuer, L., K. Vaché, et S.H.-G.F. Julich. 2008. « Current concepts in nitrogen dynamics for mesoscale catchments. » *Hydrological Sciences Journal* 53 (5).
- Bruniquel, J., et A. Lopes. 1997. « Multi-Variate Optimal Speckle Reduction in SAR Imagery ». *International Journal of Remote Sensing* 18 (3): 603–27. <https://doi.org/10.1080/014311697218962>.
- Cheema, M.J.M., W. W. Immerzeel, et W. Bastiaanssen. 2014. « Spatial quantification of groundwater abstraction in the irrigated Indus basin. » *Groundwater* 52 (1): 25–36.
- Dewandel, B., J. Perrin, S. Ahmed, S. Aulong, Z. Hrkal, P. Lachassagne, M. Samad, et S. Massuel. 2010. « Development of a Tool for managing groundwater resources in semi-arid hard-rock regions. Application to a rural watershed in south India ». *Hydrological Processes* 24 (19): 2784–97.
- Dewandel, Benoît, Yvan Caballero, Jérôme Perrin, Alexandre Boisson, Fabrice Dazin, Sylvain Ferrant, Subash Chandra, et Jean-Christophe Maréchal. 2017. « A Methodology for Regionalizing 3-D Effective Porosity at Watershed Scale in Crystalline Aquifers ». *Hydrological Processes* 31 (12): 2277–95. <https://doi.org/10.1002/hyp.11187>.
- Engel, B.A., R. Shrinivasan, J. G. Arnold, C. Rewerts, et S.J. Brown. 1993. « Nonpoint source (NPS) pollution modeling using models integrated with Geographic Information Systems (GIS) ». *Water Science and Technology* 28 (3–5): 685–90.
- Ferrant, S. 2009. « Modélisation agro-hydrologique des transferts de nitrates à l'échelle des bassins versants agricoles gascons ».
- Ferrant, S., Y. Caballero, J. Perrin, S. Gascoin, B. Dewandel, S. Aulong, F. Dazin, S. Ahmed, et J.C. Maréchal. 2014. « Projected impacts of climate change on farmers' extraction of groundwater from crystalline aquifers in South India. » *Scientific Report* 4.
- Ferrant, S., P. Durand, E. Justes, J. L. Probst, et J. M. Sanchez-Perez. 2013. « Simulating the long term impact of nitrogen scenarios in a small agricultural catchment. » *Agricultural Water Management* 124: 85–96.
- Ferrant, S., F. Oehler, P. Durand, L. Ruiz, J. Salmon-Monviola, E. Justes, P. Dugast, A. Probst, J. L. Probst, et J. M. Sanchez-Perez. 2011. « Understanding nitrogen transfer dynamics in a small agricultural catchment: comparison of a distributed (TNT2) and a semi distributed (SWAT) modelling approaches ». *Journal of Hydrology* 406: 1–15.
- Ferrant, Sylvain, Adrien Selles, Michel Le Page, Pierre-Alexis Herrault, Charlotte Pelletier, Ahmad Al-Bitar, Stéphane Mermoz, et al. 2017. « Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India ». *Remote Sensing* 9 (11): 1119. <https://doi.org/10.3390/rs9111119>.
- Hagolle, Olivier, Mireille Huc, David Villa Pascual, et Gerard Dedieu. 2015. « A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENUS and Sentinel-2 Images ». *Remote Sensing* 7 (3): 2668–91. <https://doi.org/10.3390/rs70302668>.
- Maréchal, J. C., B. Dewandel, S. Ahmed, L. Galeazzi, et F. K. Zaidi. 2006. « Combined estimation of specific yield and natural recharge in semi-arid groundwater basin with irrigated agriculture ». *Journal of Hydrology* 329: 281–93.
- Massuel, S., J. Perrin, C. Mascré, W. Mohamed, A. Boisson, et S. Ahmed. 2014. « Managed aquifer recharge in South India: What to expect from small percolation tanks in hard rock? » *Journal of Hydrology* 512 (mai): 157–67. <https://doi.org/10.1016/j.jhydrol.2014.02.062>.
- Massuel, S., J. Perrin, M. Wajid, C. Mascré, et B. Dewandel. 2008. « A simple low-cost method to monitor duration of ground water pumping ». *Ground Water* 47: 141–45.
- Mermoz, Stéphane, et Thuy Le Toan. 2016. « Forest Disturbances and Regrowth Assessment Using ALOS PALSAR Data from 2007 to 2010 in Vietnam, Cambodia and Lao PDR ». *Remote Sensing* 8 (3): 217. <https://doi.org/10.3390/rs8030217>.
- Mermoz, Stéphane, Thuy Le Toan, Ludovic Villard, Maxime Réjou-Méchain, et Joerg Seifert-Granzin. 2014. « Biomass assessment in the Cameroon savanna using ALOS PALSAR data ». *Remote Sensing of Environment* 155 (décembre): 109–19. <https://doi.org/10.1016/j.rse.2014.01.029>.
- Moreau, P., V. Viaud, V. Parnaudeau, J. Salmon-Monviola, et P. Durand. 2013. « An approach for global sensitivity analysis of a complex environmental model to spatial inputs and parameters: A case study of an agro-hydrological model ». *Environmental Modelling and Software* 47: 74–87.
- Pelletier, Charlotte, Silvia Valero, Jordi Inglada, Nicolas Champion, et Gérard Dedieu. 2016. « Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas ». *Remote Sensing of Environment* 187: 156–168.
- Perrin, J., S. Ferrant, S. Massuel, B. Dewandel, J. C. Marechal, S. Aulong, et S. Ahmed. 2012. « Assessing water availability in a semi-arid watershed of southern India using a semi-distributed model ». *Journal of Hydrology* 460–461: 143–55.
- Phan, Hoa, Thuy Le Toan, Alexandre Bouvet, Lam Dao Nguyen, Tien Pham Duy, et Mehrez Zribi. 2018. « Mapping of Rice Varieties and Sowing Date Using X-Band SAR Data ». *Sensors* 18 (1): 316. <https://doi.org/10.3390/s18010316>.
- Refsgaard, J., M. Thorsen, J. Jensen, S. LKleeschulte, et S. Hansen. 1999. « Large scale modelling of groundwater contamination from nitrate leaching. » *Journal of Hydrology* 211: 117–40.
- Sivaprasad, P., et C. A. Babu. 2014. « Seasonal Variation and Classification of Aerosols over an Inland Station in India ». *Meteorological Applications* 21 (2): 241–48. <https://doi.org/10.1002/met.1319>.