Industrial & Applications Paper

ExtremeEarth: Managing Water Availability for Crops Using Earth Observation and Machine Learning

Florian Appel
VISTA Remote Sensing in Geosciences GmbH
Germany
appel@vista-geo.de

Heike Bach
VISTA Remote Sensing in Geosciences GmbH
Germany
bach@vista-geo.de

Silke Migdall
VISTA Remote Sensing in Geosciences GmbH
Germany
migdall@vista-geo.de

Manolis Koubarakis
National and Kapodistrian University of Athens
Greece
koubarak@di.uoa.gr

George Stamoulis
National and Kapodistrian University of Athens
Greece
gstam@di.uoa.gr

Dimitris Bilidas
National and Kapodistrian University of Athens
Greece
d.bilidas@di.uoa.gr

Despina-Athanasia Pantazi
National and Kapodistrian University of Athens
Greece
dpantazi@di.uoa.gr

Lorenzo Bruzzone
University of Trento
Italy
lorenzo.bruzzone@unitn.it

Claudia Paris*
University of Twente
The Netherlands
c.paris@utwente.nl

Giulio Weikmann
University of Trento
Italy
giulio.weikmann@unitn.it

ABSTRACT

Food security, especially in a changing Earth environment, is one of the most challenging issues of this century. Population growth, increased food consumption and the challenges of climate change will extend over the next decades. To deal with these, both regional and global measures are necessary. Biomass production and thus yield will need to be increased in a sustainable way. It is important to minimize the risks of yield loss even under more extreme environmental conditions, while making sure not to deplete or damage the available resources. Two measures are most important for this: irrigation and fertilization. While fertilization relies mainly on industrial goods, irrigation requires reliable water resources in the area that is being farmed, either from groundwater or surface water.

Regarding surface water, a large portion of the world's freshwater is linked to snowfall, snow storage and seasonal release of the water. All these components are subject to increased variability due to climate change and the resulting increase in extreme events. In ExtremeEarth we designed and implemented a workflow that combines Earth Observation data with Deep Learning models to detect water demand and water availability to produce irrigation recommendations for the Danube basin.

1 INTRODUCTION

Food security is the measure of the availability of food and individuals' ability to access it, and is becoming a multi-dimensional problem given the changing Earth environment. Irrigation plays an important role as it requires both ground and surface water. Most of the fresh water comes from snowfall, it is stored in the form of ice or snow and then released when it melts. This cycle of fresh water can me monitored with EO-based products and the production of water availability maps that allow for better planning and support farmers with irrigation.

The goal of the Food Security Use Case in the ExtremeEarth project [9], was to enable the use of Earth Observation (EO) data to develop water availability maps for agricultural areas, offering valuable information for irrigation management [14]. To demonstrate the results of the application, we selected the Danube river basin for the following reasons: (i) variability in water supply due to changing precipitation patterns leading to extremes events (floods and droughts), (ii) significant portion of irrigated agriculture, (iii) significant water supply from water storage by snow in the Alps, (iv) large interest of demo users, and (v) strong economic, environmental and societal value.

To define the requirements for the application, a series of workshops was held in the context of ExtremeEarth. The first workshop was held to drive the design and implementation components, while two more were held towards the end of the project for feedback evaluation from domain experts.

To implement the water availability maps, we need information from: (i) crop type and leaf area index that come from Sentinel-2 images, (ii) soil moisture, biomass, water demand, snow storage, snow run off and groundwater computed offered by the proprietary land surface modelling software PROMET and (iii) snow cover products from the Copernicus CryoLand service, snowmelt data using Sentinel-1 images, snow water equivalent from in-situ sensors.

A core component and one of the main contributions of this work is the crop type map generation from Sentinel-2 images. The

*Work performed while at the University of Trento.

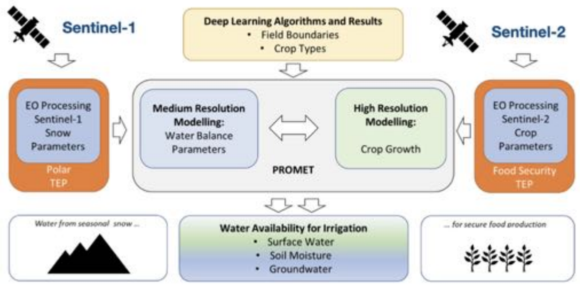


Figure 1: Water Availability for Irrigation

accuracy of the high-resolution crop type maps is fundamental for the correct application of the physically based agro-hydrological PROMET model [6], allowing the simulation of required crop parameters. Thanks to the high spatial resolution of the Sentinel-2, we generate the crop type maps at the required geometrical details (i.e., 10m) for three consecutive years, namely 2018, 2019 and 2020. This temporal span allows for modelling the temporal changes (due to the crop rotation practice several land cover changes will be present on the ground) and studying the impact of the climate on the crop properties.

For the handling of all information regarding water stored as snow, water stored in the soil, the runoff on the one side and the development of the biomass and thus the water demand on the other side, a comprehensive knowledge of the processes and status for the entire catchment area is needed. The PROMET model can integrate EO-derived land information (in this approach: crop types, biomass/leaf area, snow cover), to receive an up-to-date conformability between simulated parameters and observed parameters. Once the crop type maps are available, we utilize the PROMET model to determine water demand and combine this information with water availability sources to produce irrigation recommendations.

The application offers as output field specific irrigation recommendations for agricultural areas in Austria, Hungary and Romania. These consist of recommendations on how much and when to irrigate each crop, and yield forecasts that show crop development with and without the optimized irrigation plans we offer.

The rest of this paper is structured as follows: In Section 2 we present the ExtremeEarth platform that supports the use case. Section 3 presents the different technologies that were designed and implemented for the production of water demand and water availability maps. Finally, Section 4 presents the results of the Food Security use case.

2 THE EXTREMEEARTH PLATFORM

The ExtremeEarth platform [5] brings together the deep learning architectures and the big data technologies required for the development of our application. The ExtremeEarth platform software architecture builds on the integration of ESA Thematic Exploitation Platforms (TEPs), DIASes (Data and Information Access Services), and Hopsworks. The ExtremeEarth infrastructure enables high-performance scalable distributed data processing and deep learning on Copernicus data.

Irrigation recommendation as a task, relies heavily on developing techniques and software that will enable the extraction of information and knowledge from big Copernicus data using deep learning techniques and extreme geospatial analytics. Rapidly increasing volumes of diverse data from distributed sources create challenges for extracting valuable knowledge and commercial value from data. The extraction of useful knowledge is done through deep learning techniques that work at the extreme scale of data expected in Copernicus. As part of the ExtremeEarth platform, Hopsworks is deployed on CREODIAS. The CREODIAS infrastructure is one of five DIAS cloud-based infrastructure platforms built and operated by CloudFerro on behalf of ESA under the Copernicus Program. It provides easy access and processing of petabytes of EO data in a scalable framework.

2.1 Hopsworks

The ExtremeEarth platform utilizes Hopsworks, a data intensive AI platform from Logical Clocks. Hopsworks is an open-source framework for the development and operation of machine learning models. Hopsworks provides unique features that enhance the development of deep learning algorithms using EO data: it provides tools to build end-to-end machine learning pipelines, a feature store, management of machine learning artifacts and assets such as experiments and models, first-class support for popular open-source machine learning frameworks, integration with data science tools, and infrastructure monitoring functionalities. Hopsworks also provides a horizontally scalable platform for deep learning with GPUs and SDKs for hyper-parameter tuning and elastic model serving.

Hopsworks is a platform for both the design and operation of data analytics and ML/DL applications, but also for data engineering, with support for Spark, Flink, and Kafka. HopsFS is used in the storage layer, while Apache Hadoop YARN and Kubernetes are used for resource management. On top of HopsYARN and HopsFS, it supports different services like Spark, Flink, MapReduce, Kafka. Hopsworks then provides an intuitive UI for the services and integrates them into the Project-Dataset model [8].

2.2 Food Security TEP

The Food Security TEP [15] addresses the specific needs of a very wide user community contributing to food security and aiming at sustainable agriculture and aquaculture. The Food Security TEP offers a platform to allow extraction of information from EO data and design services for the food security sector in Europe, Africa Europe and Africa. Thereby it targets to foster smart, data-intensive aquacultural applications in the scientific, private and public domains. The user community of the Food Security TEP is looking for applications that can scale from small farms, to regional and national level, and offer crop development analytics. This community comes from a wide range of fields, including public science, app developers, the finance and insurance sectors, local and national administration and international agencies.

The technical infrastructure is a web-based Platform-as-a-Service, developed by CGI Italy, that leverages the most advanced cloud computing technologies. Food Security TEP provides a single point for access and analysis of data coming from all Copernicus missions, along with additional sources useful to the specific domain. It facilitates implementation of specific services, through the creation of processing algorithms and allows their execution, monitoring and maintenance. The platform is accessible through the Open Expert Interface, that introduces the main functionalities and provides easy access to a variety of tools and datasets. The platform allows data visualization on mobile devices and can offer customized products and services to selected users.

The Food Security TEP is operational and constantly tries to strengthen its federation with cloud platforms.

As with other TEPs implemented by ESA, the idea is to bring the user to the data. The Food Security TEP joins access to EO data, in-situ data, processing tools, computing resources, and hosted processing on one platform. The Food Security TEP is implemented on CREODIAS to make use of optimal EO data access and scalable ICT resources.

2.3 Linked Data Tools

Linked data allows us to create semantic links between the extracted knowledge from Copernicus satellite images and other external sources, such as land cover/use and water products. This extended knowledge graph (KG) along with the semantic web tools were used for validation of the Deep Learning results and assist in the production of services for the Food Security TEP users.

Within the application, a number of spatial information layers are used, ranging from Copernicus satellite data, their direct products, model results and local specific data for the demo and the involved users. The main information is based on satellite data in combination with modelled data to be used as water availability and irrigation information. In order to combine these different resources we used an ontology to model the data in RDF and the linked data tools to answer complex semantic queries. The Food Security ontology is used to translate the results of the ML algorithms in the RDF model using the tool GeoTriples-Spark [12]. The tool JedAI-spatial [17] can then be utilized to discover spatial relations between two different input datasets of linked geospatial data. The data is stored and queried with Strabo2 [1]. The Strabo2 SPARQL endpoint is deployed in Hopsworks and can be accessed by the Food Security TEPs through an HTTP-based RESTful API according to the W3C standard [1]. Finally, the GeoSPARQL query federation engine Semagrow [2] is used for integrating several Strabo2 endpoints under a single GeoSPARQL endpoint.

3 THE EXTREMEEARTH FOOD SECURITY USE CASE

The Food Security Use Case is addressing public institutions (e.g., ministries, regional administrations) as well as private users (farmers, agricultural industry). To make sure that the developed application fits with the requirements of the user community, user requirements gathering was done within the design phase and continued throughout the development phase.

The Food Security Use Case targets the assessment of water demand and water availability for irrigation by combining big data EO analysis with crop growth modelling to provide water availability. It is based on existing technology components that are integrated and adapted for a scalabe approach. Some existing information and products were not planned to be produced fully automatically, but big data and analytics of whole catchments cannot be handled manually. Hence, all aspects of data access, pre-processing and basic product generation have been realized in executable service chains, and can now be used as an element for future developments and enhanced service provisions.

For a successful demonstration of the capabilities of Copernicus EO / big data / analytics solutions, pilot service areas were set up. The functionality of the components and the benefits derived from the tasks within ExtremeEarth have been demonstrated within the Danube Catchment and the Duero Catchment. In both areas, user contacts have been established coming from both public authorities and the agricultural industry. Public authorities are more interested in national or basin-wide information, while the industry users were mostly interested in direct irrigation recommendations.

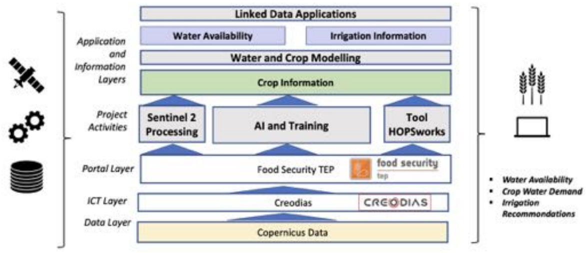


Figure 2: Application Layers

The Food Security Use Case utilizes and adapts components from different layers of the European EO environment. Ranging from Copernicus data to the generation and interlinking of information, including the applications in agricultural and hydrological domains, a very wide range of components has to be taken into account. But in general, the story spans from satellites to benefits for farming and food. By exploiting the existing components and adapting them, impact in technological advances and benefits for the society can be created.

Starting from the data layer, made available by ICT (CREODIAS) and portals (Food Security TEP), the key elements of the Food Security Use Case builds on them and are linked as service chains, to provide water availability and irrigation information. Crop information – giving the crop type for each field and each season - is the central parameter connecting EO efforts, model approaches and information facts.

Refined from the user workshops, the summarized requirements were used to set up all components in an appropriate way. From the functional analysis performed between user requirements definitions and system design, it was indicated to combine and set up a number of technical components to structure the development. Some of the components already existed within the partners' developments and services and had to be combined to the new service chains.

The benefit for the users and impact for food security does thus not come from starting to develop water availability modelling from scratch, but from using the existing experiences on crop growth analyses and water balance modelling, together with high quality crop type classification and state-of-the art machine learning technology. Those components are representing the main building blocks to provide information about water demand, water availability and irrigation recommendations.

Beyond the direct developments in the use case, serving user needs, the exchange of data, integration and interoperability (e.g., linked data approaches, federation of IT resources, onboarding of service elements) play an important role in the context of Food Security. The next sections will give a more detailed overview on the implementation of the technologies and building blocks of the Food Security use case.

3.1 Deep Learning for Crop Type Mapping

The production of accurate crop type maps play an important role in the generation of water availability and water demand maps.

[1] https://www.w3.org/TR/sparql11-http-rdf-update/

Several methodologies have been exploited to classify agricultural areas from time series of high resolution optical multispectral images. In general, deep learning models have shown better performance when compared to shallow classifiers [4, 20, 21]. Long Short-Term Neural Networks [4] in particular showed great performance against the other techniques, allowing to map long time series of information and long-time dependencies. Moreover, recurrent neural networks are robust against missing values in the classification problem [3], allowing the analysis of time series of S2 images with missing information due to the presence of clouds. Convolutional Neural Networks have also been exploited for mono-temporal classification as in [7, 11]. However, for sequential tasks, recurrent networks are generally more suitable, allowing a better characterization of the phenological trend of the classes analyzed.

The main goal of the developed deep learning model was the accurate classification of crop types in our study area (i.e., Danube catchment). In particular, the outputs of the system architecture are the crop type map, and crop boundaries map. To this end, we leverage on the long time series of Sentinel-2 images and publicly available thematic products.

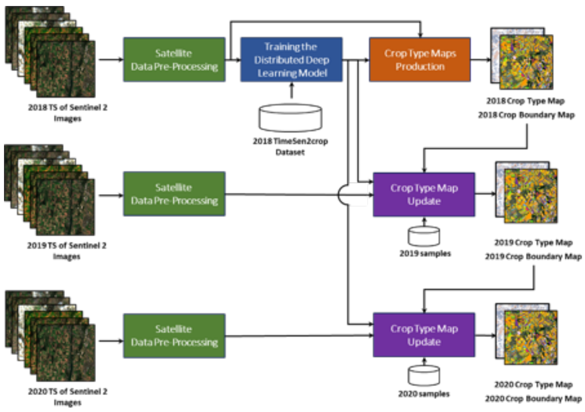


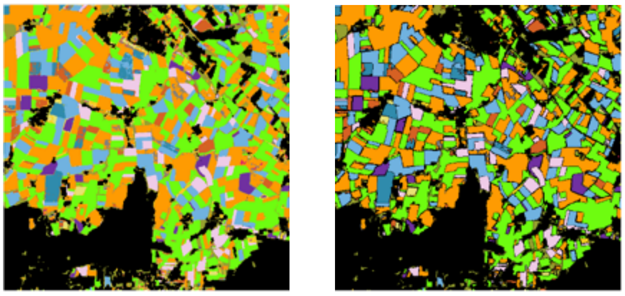
Figure 3: Workflow of the system architecture to generate annual crop type and crop boundaries maps at high spatial resolution

Figure 3 shows the workflow of the system architecture based on a multitemporal deep learning model [18] to produce annual crop type and crop boundaries maps at high spatial resolution. The system has been defined to fully take advantage of the properties of the long and dense Time Series of Sentinel-2 images to achieve accurate crop type mapping results. While a dedicated optical pre-processing step has been defined to deal with the temporal inconsistencies present in Time Series acquired over different tiles, a weighted multitemporal deep learning model is implemented to mitigate the severely imbalanced classification problem. The system architecture consists of four main steps: (i) the optical pre-processing step, (ii) the training of the multitemporal deep learning model, (iii) the crop type maps production, and (iv) the crop type map update.

Although many deep learning algorithms have been proposed for land cover classification [22], most of them, such as the standard Convolution Neural Networks [16] (CNNs), rely on mono temporal images. When using classical CNN for crop type mapping, the temporal information is totally neglected. However, this information is fundamental for the characterization of the phenological parameters of different crop types and thus, for the accurate crop type mapping. Recurrent Neural Network (RNNs) are mainly designed for handling sequential data (i.e., temporal sequences of observations) as the network exploits the previous observation for the classification of the current one. To capture the different phenological characteristics of the crop, we are considering a peculiar kind of RNN, namely the Long short term memory (LSTM) deep network [19, 20]. LSTM neural networks provide long term memory capabilities, as at each observation, information can be stored or retrieved to varying extents. By exploiting the capability of the network of encoding the multitemporal information recorder by the long time series of Sentinel-2 we can accurately model the crop mapping classification task. The LSTM is trained using the training set TimeSen2Crop [24] which has been extracted from publicly available crop type maps and it is composed by more than one million samples. In addition to the single year-mapping, multi-year mapping is considered in order to provide the crop type maps for multiple years.

The trained deep learning model is then used to generate the crop type maps. In 2018, the TimSen2Crop training dataset is contemporary to the Time Series of Sentinel-2 images to be classified. Therefore, the deep learning model simply predicts the crop type maps. To generate the crop type maps of 2019 and 2020, a fine-tuning strategy is adopted to adjust the trained deep learning model to the recent Time Series of images.

To extract the crop type boundaries and reduce the noise at pixel level, the obtained classification maps are post-processed. Indeed, since the LSTM adopted performs a pixel-wise classification, noisy pixels may be present in the maps. Moreover, the crops boundaries may contain mixed spectral signatures belonging to different neighboring crops. However, the PROMET model requires only the pixels belonging to the crop having a pure spectral signature associated to the crop type. For these reasons, a series of morphological operators is applied.



(a) **(b)**

Figure 4: Qualitative examples of crop type map before (a) and after (b) the morphology filtering post processing

Because of the crop rotation practice and changes in both the image acquisition conditions and the crop phenology, the class statistical distributions of image Time Series acquired over different years have significant differences. For this reason, the performance of a deep learning model successfully trained for a specific year can drastically drop if another year is considered [23]. To solve this problem, we took advantage of a well-explored solution typically employed to adapt a pre-trained network to a target dataset, i.e., the fine-tuning approach [10]. Typically, the weights of all the layers in the pre-trained network are frozen except for the latest. Then, the obtained network is re-trained using the training data of the target dataset and after a few epochs the entire network is unfrozen, for the remaining layers to fine-tune.

Although fine-tuning is effective, its accuracy depends on the amount of available recent training data, which is difficult to collect and time consuming. Indeed, from the operational viewpoint, scarce multitemporal training data are typically available. This step aims to update the corresponding crop type map for a different year for which a small amount of reference data is available. To this end, we combine Self-Paced Learning (SPL) and Fine-Tuning (FT) techniques.

In the considered SPL strategy, if at a given iteration a class has enough training samples extracted, the technique focuses on the remaining classes at the next iterations. At the convergence of the iteration in the SPL step, a fine-tuning step is performed. Due to the SPL strategy, the number of samples required by the fine-tuning approach to achieve accurate classification results is much lower than the one needed by the standard approach. Similar to the production of the 2018 crop type maps, the obtained products are post-processed applying the morphological operators. Then, the overlapping areas of the Sentinel-2 tiles are harmonized at crop level according to the posterior values.

3.2 Water Demand Modelling with PROMET

For the handling of all information regarding water stored as snow, water stored in the soil (soil moisture), the runoff on the one side and the development of the biomass and thus the water demand on the other side, a comprehensive knowledge of the processes and status for the entire catchment area is needed. Applying a model, preferably on a physical basis to obtain transferability, is the basis of the provision of the requested information.

The crop growth model PROMET [13] has more than 25 years of development history, and is currently being applied in scientific and service operations of partner VISTA. Modelling with PROMET can integrate EO-derived land information (e.g., crop types, biomass/leaf area, snow cover), to receive an up-to-date conformability between simulated parameters and observed parameters. This assimilation of observed / measured information to physical simulations provides large benefit in information gain. Model internal processes (e.g., phenology, biomass and water uptake from soil), not seen by any satellite, can be used to monitor the land surface dynamics. That information is provided and used for recommendations.

Meteorological data is driving the model and thus the plant development in the simulations. PROMET allows the use of both measured meteorological inputs (station data) and forecast data. For the presented application, several enrichments were performed leading to a multi-stage model approach, combining dedicated model runs for the assessment of the targeted parameters and information. This approach, shown in Figure 5, was chosen based on data availability and handling, and performance issues.

We performed medium resolution model runs (1km spatial resolution) mainly for water availability simulations, including snow, soil, groundwater and rivers and reservoirs. For this level, also the seasonal forecasts have been used to provide a catchment wide information on precipitation, runoff (river and reservoirs) and soil moisture and the derived expected growth of the crops.

The implementation of seasonal forecasts opens a new dimension of information in the context of water availability and irrigation recommendation. Data is released by the various data centers and federated by the European Centre for Medium-Range Weather Forecasts (ECMWF), providing all variables according to PROMET's input needs. New forecasts are released monthly up to 200 days in advance and contain daily and 6h temporal resolution datasets. For the integration in the model, ensemble selection and spatial and temporal scaling needed to be applied. Data access, data download and data conversion have been successfully tested. Early analyses gave indications of unrestricted applicability, with expected uncertainties due to spatial and temporal resolution.

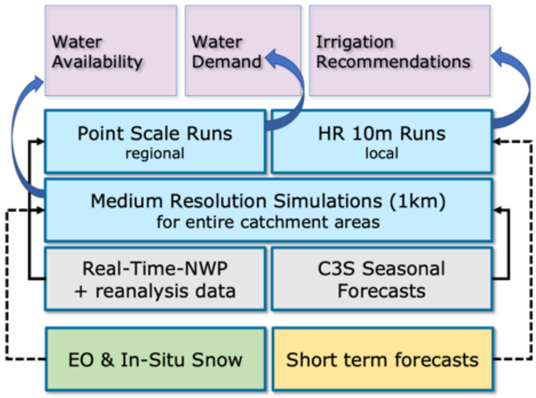


Figure 5: Multi-Stage Model Approach

3.3 Snow Monitoring and Water Availability

Water availability, as in the amount of usable water during the vegetation period for food production, was the focus topic of this use case in ExtremeEarth. Other aspects of water availability, e.g., for domestic and industrial water supply, shipping (touristic and logistic) and energy production (hydropower and cooling of power plants) are also covered by the developments but are not directly addressed in ExtremeEarth.

Within the selected catchments of the Danube and the Douro, a significant portion of the seasonal runoff has its origin in water release from seasonal snow. This melted water, resulting in runoff in the rivers or storage in lakes and reservoirs sums up the uncertainties/dynamics of precipitation and the uncertainties/dynamics of melting conditions (e.g., air temperature, radiation and rain on snow effects). EO-based snow monitoring is an established method to get information over larger areas and in Near-Real-Time / in short order. Nevertheless, the requested information on the water stored in the snow cover (Snow Water Equivalent) is not possible to be retrieved with operational EO methods. The assimilation of observations and use of a model are the appropriate ways of providing all information requested for water resources management.

Within ExtremeEarth, there are three monitoring mechanisms for this application: (i) monitoring of the Snow Covered Area using optical data, (ii) monitoring of the Snow Melt / internal dynamic of the snow cover using SAR data and (iii) measuring the Snow Water Equivalent (SWE) at critical locations using Global Navigation Satellite System (GNSS) based stations.

Monitoring the Snow Covered Area. Using optical satellites is a well-established method to derive the area covered by snow on a daily basis, using medium resolution satellites (e.g., MetOp, Terra/Aqua). Under cloud free conditions, such products can be reliably generated from various satellite data sources with a spectral band in the Short Wave InfraRed (SWIR). Within this task, we enhanced the mechanisms of making use of the existing EU Cryoland services established by an international team within an EU FP7 project [2]. Those data products (500m resolution) have been loaded from the dedicated portal and fed into the snow monitoring service chain. Daily snow cover maps are now available for the last 5 winters.

As a second option, access to the new HR Snow Monitoring Service by Copernicus / EEA became available within the project period. Snow covered area maps (including fractional snow information FSC) with 20m resolution has been made available as part of the Copernicus Land Service [3].

Monitoring of the Snow Melt. To obtain information on the snow melt dynamics, Sentinel-1 observations are the appropriate way to monitor on a regular basis. Independent from cloud and daylight, several observations per day can be analysed. Several snow cover periods have been processed, and mechanisms of data search and batch processing have been tested. Any processing converts Sentinel-1 GRD products (1.6 GB per scene) into terrain corrected backscatter, using SRTM elevation model and radiometric corrections.

Transferring technologies and using the upgrades of the Food Security TEP, e.g. by integration of processing algorithms and processors applied and provided by Sen4CAP [4], efficient Sentinel-1 pre-processing has been made available on the Food Security TEP as well. Continuous and automatic pre-processing of Sentinel-1 data, wet snow mapping and product provision are enabled to be linked with all the other EO products of the Food Security Use Case on platform.

Produced spatial snow melt maps can be used to compare and control the model runs for the Danube, and the Duero catchment. This method was successfully applied in the ESA business application SnowSense [5], then performing a demonstration for the island of Newfoundland / Canada. Within 2021, additional snow monitoring products from the Copernicus Land Services have been made available. In addition to the mentioned Fractional Snow Cover (FSC), also snow state conditions products are provided to Copernicus users [6]. Products are using Sentinel-1 and Sentinel-2 in combination but are partly limited to high mountain areas. Access to the dataset is limited to users of the WEkEO portal and requires a registration for the services.

Measurements of the SWE at critical locations. As mentioned, EO monitoring is not able to derive accurate information on the amount of water stored as snow. To control and update the modelled information, local in-situ measurements are helpful. Beyond the technical aspect of control and updating of the model, local measurements are an important vehicle to convince users and potential customers. Hydrologists and especially agri-economists need technological solutions that fit into their experience of data and information retrieval. The team of VISTA installed a first SnowSense [7] station, using patent pending GNSS technology, in the Tatra Mountains. This location, at the northern border of the Danube catchment, is sparsely equipped with meteorological and hydrological stations. Local SWE information can improve the assessment of water availability information for larger areas.

In collaboration with some local partners and customers, more SnowSense stations have been set up in the Alps (including parts of the Danube catchment). SnowSense will be marketed more intensely in the future. Local Snow Water Equivalent retrieval is available as auxiliary information source and can improve water availability information in future services.

4 RESULTS

Crop Water Demand Mapping in the Danube Basin. Based on the crop type classification, we selected a representative sample of pixels distributed evenly over the Danube basin. To guarantee pure crop specific remote sensing information, we selected our samples to be over a certain distance from field boundaries and roads. In order to derive information for crop water demand in national scale for different crop types, we validated the results with crop type maps and weather data from 2018 for NUTS 2 regions in Austria. This approach revealed regional differences in spring and summer precipitation as shown in Figure 6. These anomalies affected soil moisture development and crop water stress in eastern Austria. Severe crop stress was reported in May and June, that resulted in increased water demand in our simulation. Such differences in the growing periods can result in differing irrigation requirements for each crop type, and showcase the need for accurate and high-resolution crop type maps.

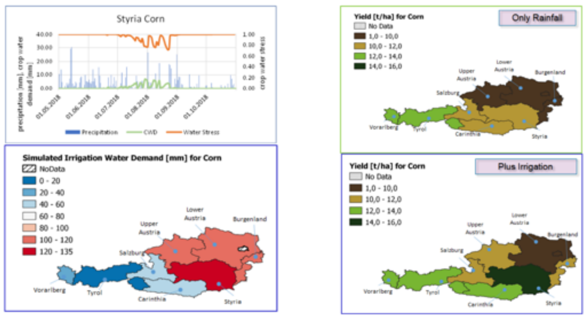


Figure 6: Regional results of water stress, irrigation water demand and the effect on yield of corn, focus on Styria / Austria 2018

In this manner, EO derived crop information can assist in the production of water demand and assist with food security management. Combining water availability from the hydrological cycle of the catchments, predictions for the summer period and our water demand analysis, we can offer irrigation policy advice on local and regional level.

Water Availability Simulations in the Douro Basin. All irrigation measures should happen only when there is sufficient and sustainable water availability in the area or catchment. Considering river runoff and storage in reservoirs, including the storage of water between the seasons in natural reservoirs as snow, soil layers or groundwater, there is need for an universal and extended assessment of the water cycle. Based on the catchment wide simulations with the PROMET model, we selected an evaluation of the basic, local water availability from precipitation and soil storage for the years 2019 and 2020. To present the potential of seasonal forecasts (from Copernicus), we selected the year 2021. To make sure there was comparability and clear evaluation of the different forecast scenarios, the model was set up with a static, homogenised land use and crop map and soil information.

Using the simplified model set up for the area (1 km spatial, 1 hour temporal resolution, reduced crop information) led to new insights for the years 2019 and 2020. We performed our simulations with example crop types (summer wheat and summer

[2] http://www.cryoland.eu/
[3] https://land.copernicus.eu/pan-european/biophysical-parameters/high-resolution-snow-and-ice-monitoring
[4] http://esa-sen4cap.org/
[5] https://business.esa.int/projects/snowsense-dp
[6] https://land.copernicus.eu/pan-european/biophysical-parameters/high-resolution-snow-and-ice- monitoring/snow-products/snow-state-conditions
[7] www.snowsense.de

barley), to make use of the capabilities of the PROMET model to give insights into water stress and water demand. Already from this analysis and model performances, the high local variance and the high difference between the years could be seen.

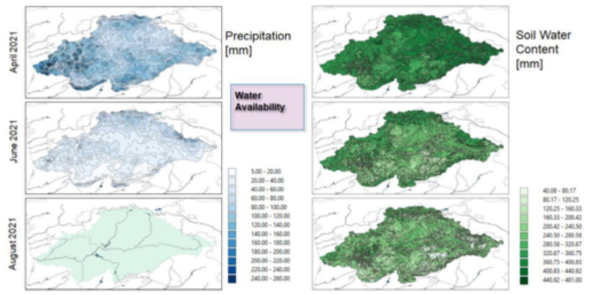


Figure 7: Water Availability in Douro catchment 2021, selected months

Within more detailed simulations the daily/weekly/monthly situation of the water availability has been analysed. As shown in Figure 7, a dry summer, with almost no precipitation in July and August, led to decreasing water availability for farming from the soil. With a vegetation period partly ranging to these months, a high risk of yield loss is given. Based on the simulations of the selected crops, the day-by-day irrigation demand could be evaluated. Together with the already mentioned water availability information/maps, the information could be provided for future use inside or outside of the project team.

As additional analysis, the application of seasonal forecasts – provided by Copernicus – has been evaluated. From the entire bunch of models and ensembles available, we selected the UK model and set-up a preliminary data import and data handling schema. Based on the setting of the water availability and water demand from the observed meteorological inputs, the performance of the seasonal forecasts (with monthly release steps) has been investigated. A forecast of the farming and water demand conditions showed promising result and will be further investigated. A medium-term estimate of the next few months appears to be a beneficial impact for water and farm management.

Irrigation Recommendation. Irrigation Recommendations and forecasts are a promising field of business for EO and model based services. In order to generate this information, in addition to currently available methods and technologies, large amounts of data are required. These are particularly up-to-date and very precisely prepared data, which also have to take into account the technology on the field that is available for irrigation. Using the crop type as well as the leaf area derived from EO data as input to the crop modelling, crop water demand is calculated. Based on this crop water demand, the irrigation advice can then be given in a form that fits the farmer's equipment, e.g. field-wise, sector-wise (e.g. pivot-irrigation) or in a 10x10m raster.

It is important to know exactly the crop type that is being cultivated, as only then can the precise water demand be determined. In existing services of our partner VISTA, this information is given by the farmer itself, but for a wider spread of this application – especially for larger scale assessment, where direct information from farmers are not available, an independent EO based information would be helpful. Here, the dense time series of observations are used to retrieve current crop status. Based on the day-by-day, or week-by-week model simulations, the development of the plant – according to the observed and calculated

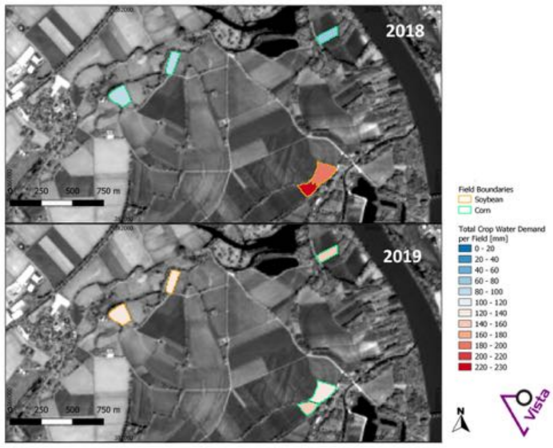


Figure 8: Crop water demand of different soy bean and corn fields in Bavaria (Danube Catchment) in 2018 vs. 2019 (10x10 meter). Due to drought conditions in 2018, the high water demand by the soy beans could not be satisfied by precipitation alone.

phenology – including the specific mechanisms of water use and water loss, can be derived.

In the example in Figure 8, the farm is in an area where traditionally water availability via precipitation was enough to satisfy crop growth. In the past decade, hot dry summers, but also dry spring seasons have changed that and crop stress and resulting yield reductions are becoming more common. The ExtremeEarth system can support farmers by modelling the impact of these conditions on the crops and the resulting yield loss, giving decision support when investing in irrigation systems is meaningful, and giving support when and how much water should be applied to the fields.

On the regional scale, if irrigation water is taken mainly from above-ground water sources like rivers, the ExtremeEarth system can calculate for whole watersheds how much water is available as runoff due to snow melt and precipitation, and how irrigation influences the river runoff. This supports a fair water distribution and helps to make sure that irrigation is ecologically sustainable.

5 CONCLUSION

Continuous monitoring of the water availability as well as the crop development and water demand is needed to react quickly and find the most sustainable solutions for the future. Earth Observation with the Copernicus Program and the recent capabilities of Cloud Computing, enabled ExtremeEarth to develop, enhance and demonstrate the beneficial use and interactions between the data, technical prerequisites and new technical solutions for water management.

We presented our proposed solutions to produce irrigation recommendations with the use of water demand and water availability maps, utilizing EO data and the ExtremeEarth platform, and demonstrated our workflow in the Danube and Douro basins.

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 projects ExtremeEarth (825258) and STELAR (864769). Thank you to Dr. Wolfgang Angermair for whose fields results are shown.

REFERENCES

[1] Dimitris Bilidas, Theofilos Ioannidis, Nikos Mamoulis, and Manolis Koubarakis. 2022. Strabo 2: Distributed Management of Massive Geospatial RDF Datasets. In *The Semantic Web - ISWC 2022 - 21st International Semantic Web Conference, Virtual Event, October 23-27, 2022, Proceedings (Lecture Notes in Computer Science)*, Ulrike Sattler, Aidan Hogan, C. Maria Keet, Valentina Presutti, João Paulo A. Almeida, Hideaki Takeda, Pierre Monnin, Giuseppe Pirrò, and Claudia d'Amato (Eds.), Vol. 13489. Springer, 411–427.

[2] A. Charalambidis, A. Troumpoukis, and S. Konstantopoulos. 2015. SemaGrow: optimizing federated SPARQL queries. In *SEMANTICS*.

[3] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. 2018. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Scientific Reports* 8, 1 (2018), 6085. https://doi.org/10.1038/s41598-018-24271-9

[4] Hugo Crisóstomo de Castro Filho, Osmar Abílio de Carvalho Júnior, Osmar Luiz Ferreira de Carvalho, Pablo Pozzobon de Bem, Rebeca dos Santos de Moura, Anesmar Olino de Albuquerque, Cristiano Rosa Silva, Pedro Henrique Guimarães Ferreira, Renato Fontes Guimarães, and Roberto Arnaldo Trancoso Gomes. 2020. Rice Crop Detection Using LSTM, Bi-LSTM, and Machine Learning Models from Sentinel-1 Time Series. *Remote Sensing* 12, 16 (Aug 2020), 2655. https://doi.org/10.3390/rs12162655

[5] Desta Haileselassie Hagos, Theofilos Kakantousis, Vladimir Vlassov, Sina Sheikholeslami, Tianze Wang, Jim Dowling, Claudia Paris, Daniele Marinelli, Giulio Weikmann, Lorenzo Bruzzone, Salman Khaleghian, Thomas Krmer, Torbjorn Eltoft, Andrea Marinoni, Despina-Athanasia Pantazi, Georgios Stamoulis, Dimitris Bilidas, George Papadakis, George Mandilaras, and Andreas Cziferszky. 2021. ExtremeEarth Meets Satellite Data From Space. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* PP (08 2021), 1–1.

[6] T. Hank, H. Bach, and W. Mauser. 2015. Using a Remote Sensing-Supported Hydro-Agroecological Model for Field-Scale Simulation of Heterogeneous Crop Growth and Yield: Application for Wheat in Central Europe. *Remote Sensing* 7, 4 (2015), 3934–3965.

[7] Fan Hu, Gui-Song Xia, Jingwen Hu, and Liangpei Zhang. 2015. Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sensing* 7, 11 (Nov 2015), 14680–14707. https://doi.org/10.3390/rs71114680

[8] Mahmoud Ismail, Ermias Gebremeskel, Theofilos Kakantousis, Gautier Berthou, and Jim Dowling. 2017. Hopsworks: Improving User Experience and Development on Hadoop with Scalable, Strongly Consistent Metadata. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. 2525–2528.

[9] Manolis Koubarakis, George Stamoulis, Dimitris Bilidas, Theofilos Ioannidis, George Mandilaras, Despina-Athanasia Pantazi, George Papadakis, Vladimir Vlassov, Amir H. Payberah, Tianze Wang, Sina Sheikholeslami, Desta Haileselassie Hagos, Lorenzo Bruzzone, Claudia Paris, Giulio Weikmann, Daniele Marinelli, Torbjorn Eltoft, Andrea Marinoni, Thomas Kraemer, and Andreas Cziferszky. 2021. Artificial Intelligence and Big Data Technologies for Copernicus Data: The ExtremeEarth Project.

[10] Xuan Liu, Mingmin Chi, Yunfeng Zhang, and Yiqing Qin. 2018. Classifying High Resolution Remote Sensing Images by Fine-Tuned VGG Deep Networks. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. 7137–7140.

[11] Konstantinos Makantasis, Konstantinos Karantzalos, Anastasios Doulamis, and Nikolaos Doulamis. 2015. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 4959–4962. https://doi.org/10.1109/IGARSS.2015.7326945

[12] George Mandilaras and Manolis Koubarakis. 2021. Transforming big geospatial data into linked data. In *Forthcoming*.

[13] Wolfram Mauser and Heike Bach. 2009. PROMET -Large scale distributed hydrological modeling to study the impact of climate change on the water flows of mountain watersheds. *Journal of Hydrology* 376 (10 2009), 362–377.

[14] Silke Migdall, Sandra Dotzler, Christian Miesgang, Florian Appel, Markus Muerth, Heike Bach, Giulio Weikmann, Claudia Paris, Daniele Marinelli, and Lorenzo Bruzzone. 2021. Water stress assessment in Austria based on deep learning and crop growth modelling. In *Submitted to BiDS*.

[15] Markus Muerth, Silke Migdall, M Hodrius, Fabian Niggemann, M Holzapfel, Heike Bach, Sven Gilliams, T Roey, Antonio Cuomo, P Harwood, N Pace, I Sanchez, A Dean, J Suwala, A Romeo, I Alonso, P. Mougnaud, and E Volden. 2020. Food Security TEP - Supporting sustainable intensification of food production from Space. *IOP Conference Series: Earth and Environmental Science* 509 (07 2020), 012038. https://doi.org/10.1088/1755-1315/509/1/012038

[16] Rahul Nijhawan, Deepankar Joshi, Naman Narang, Aditya Mittal, and Ankush Mittal. 2019. *A Futuristic Deep Learning Framework Approach for Land Use-Land Cover Classification Using Remote Sensing Imagery*. 87–96.

[17] George Papadakis, Georgios Mandilaras, Nikos Mamoulis, and Manolis Koubarakis. 2021. Progressive, Holistic Geospatial Interlinking. In *The Web Conference*.

[18] Claudia Paris, Giulio Weikmann, and Lorenzo Bruzzone. 2020. Monitoring of agricultural areas by using Sentinel 2 image time series and deep learning techniques. 18.

[19] Marc Rußwurm and Marco Körner. 2017. MULTI-TEMPORAL LAND COVER CLASSIFICATION WITH LONG SHORT-TERM MEMORY NEURAL NETWORKS. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-1/W1 (05 2017), 551–558.

[20] Marc Rußwurm and Marco Körner. 2017. Temporal Vegetation Modelling Using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-spectral Satellite Images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 1496–1504.

[21] Marc Rußwurm, Sébastien Lefèvre, and Marco Körner. 2019. Breizhcrops: A satellite time series dataset for crop type identification. In *International Conference on Machine Learning Time Series Workshop*, Vol. 3.

[22] Yady Tatiana Solano-Correa, Francesca Bovolo, Lorenzo Bruzzone, and Diego Fernández-Prieto. 2020. A Method for the Analysis of Small Crop Fields in Sentinel-2 Dense Time Series. *IEEE Transactions on Geoscience and Remote Sensing* 58, 3 (2020), 2150–2164.

[23] Rei Sonobe, Hiroshi Tani, Xiufeng Wang, Nobuyuki Kobayashi, and Hideki Shimamura. 2014. Parameter tuning in the support vector machine and random forest and their performances in cross- and same-year crop classification using TerraSAR-X. *International Journal of Remote Sensing* 35, 23 (2014), 7898–7909.

[24] Giulio Weikmann, Claudia Paris, and Lorenzo Bruzzone. 2021. TimeSen2Crop: A Million Labeled Samples Dataset of Sentinel 2 Image Time Series for Crop-Type Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), 4699–4708.