


Article

A Bayesian model for annual crop phenological parameter estimation using optical high resolution image time series

Jordi Inglada ¹ ¹ CESBIO, Université de Toulouse, CNES/CNRS/IRD/UPS, Toulouse, France

Abstract: Vegetation status assessment is crucial for agricultural monitoring and management. Vegetation indices derived from high resolution image time series can be used to derive key phenological parameters for annual crops. In this work, we propose a procedure for the estimation of these parameters and their associated uncertainties. The approach uses Bayesian inference through Markov Chain Monte Carlo in order to obtain the full joint posterior distribution of the phenological parameters given the satellite observations. The proposed algorithm is quantitatively validated on synthetic data. Its use on real data is presented together with an application to real-time within season estimation allowing for phenology forecasting.

Keywords: phenology; satellite image time series; vegetation index; Bayesian inference

1. Introduction

Vegetation status assessment is a crucial information for agricultural monitoring and management. In the coming years, climate change [1], increasing population [2] among other factors will rise the level of pressure on agricultural production. Also, agricultural practices are an important source of ecosystem degradation at the global scale [3] and therefore, land use monitoring related to farming is crucial for sustainable land management [4].

Remote sensing imagery in general and, in particular, high temporal and high spatial resolution data as the ones provided by Sentinel-2 [5] constitute a major asset for this kind of application. Indeed, high temporal resolution provides access to phenology monitoring through vegetation status indicators.

These indicators are usually derived from vegetation indices. For instance, the well-known Normalized Difference Vegetation Index (NDVI) is largely used to monitor fractional vegetation cover. The Leaf Area Index (LAI) is preferred when downstream models requiring a physical magnitude are used. The choice between NDVI, LAI or other vegetation indices is application dependent. Bolton and Friedl [6] show that remotely sensed vegetation indices and phenology metrics can be used to forecast agricultural yields. Using Sentinel-2 image time series, NDVI and LAI maps can be produced at 10 meter resolution every 5 days [7].

For most modeling and diagnostic applications a set of metrics describing the temporal evolution of the vegetation is preferred to the raw profiles. These can be for instance the dates of the emergence of the vegetation (start of season), the maturity, the senescence, etc. The extrema values of the vegetation index are also used.

All these metrics can be computed at different temporal and spatial scales depending on the application. Precision agriculture can use pixel-level estimations to deal with heterogeneity within the fields. Field-level estimations are enough for other applications. Evaluations at the end of the season can be useful for long term monitoring, while within season estimations are needed for real-time management and forecasting.

For all applications, the availability of uncertainty estimates is crucial, since many sources of errors are present in the computation of phenological parameters. Indeed, the remote sensing observations can contain different sources of noise: missing acquisitions

due to cloud cover, undetected clouds or cloud shadows, atmospheric correction errors. The chosen metrics assume an underlying model for the vegetation phenology. The accuracy of these assumptions with respect to the observed reality will also introduce errors in the estimates.

In this paper, we propose a methodology for the estimation of phenology parameters for annual crops using high resolution optical satellite image time series which is able to provide uncertainty estimates. The procedure is based on a Bayesian modeling. The approach is illustrated for a specific phenology model for crops (a 6 parameter double logistic function fitted on NDVI), but it is generic enough to be applied to other models.

The paper is organized as follows. In section 2, we review the existing literature on phenology estimation from satellite image time series. The phenological model used in this work is introduced in section 3 and the Bayesian approach is presented in section 4. The data sets used for the qualitative and quantitative assessment of the method are presented in section 5. In section 6 we present the results and we discuss them in section 7. Finally, the conclusion is presented in section 8.

2. Literature review

The characterization of vegetation phenology using remote sensing has been widely addressed in the literature. Most of the methods estimate phenological parameters using either sigmoidal fitting based on the work of Zhang et al. [8] or empirical thresholds on vegetation indices. The most usually extracted markers are start of season and end of season [9], [10], which are also referred as onset and offset [11]. These 2 phenophases are useful in determining the length of the growing season.

In the published literature, the onset of the vegetation appears under many different names (*SoS* for start of season or start of spring, emergence date, etc.). It is also defined in multiple ways which actually represent different phenological stages (e.g., the timing when vegetation starts to green up, the timing when vegetation grows the fastest) [12], [13]. For instance, [8] defines it as the first maxima of the rate of change. An inter-comparison of *SoS* retrieved using 10 satellite methods made by Xu et al. [14] shows that the difference between individual methods can be as much as two months. It also shows that most of the methods have large biases with respect to field data. Other works show disparities up to 100 days [15].

Some studies also derive more than two phenology markers. For example, Zhang et al. computed four parameters including the onset of green-up, maturity, senescence and dormancy [8].

Many papers have compared different approaches on different types of vegetation, as for instance crops [16]; forests, grasslands and shrublands [17]; croplands and forests [18], [19]; and other types of vegetation [20].

All these papers use the full vegetation cycle either for fitting a parametric model or to produce a smoothed profile on which robust estimators can be applied. The key dates are either estimated from the parameters of the fitted models, or by using constraints on the profile derivatives. For instance, Zhang et al. [8] propose the first maximum of the 3rd derivative of the double logistic function, but other versions using the 2nd derivative (the camel-back method [21]) also exist.

The simpler methods using empirical thresholds on vegetation index (VI) profiles [22], [23], need both a smoothing window of several months and the knowledge of the maximum value of the VI. For instance, Reed et al. [22] use 9 14-day composites, which corresponds to more than 4 months. Furthermore, the thresholds are dependent on the vegetation type and the site.

Most of the literature cited above uses medium or low resolution imagery (AVHRR and MODIS) which ensures regular cloud-free data and smooth temporal behaviors. Only recently, have these approaches been compared using high spatial resolution data (30 m) [24], and even though a 2-day revisit cycle was used, the whole season approach was adopted.

A recent review on the topic was done by [25] showing that the choice of the most appropriate model «depends upon the purpose of the study, the growth trajectory to be analyzed and the targeted land cover type(s)».

Another recent review, focused on high spatial and temporal resolution image time series [26] gives interesting discussion about the differences between ground phenology and land surface phenology (satellite). This paper also lists performances of phenology estimation methods and concludes that «more efforts are required for testing the efficiency of Sentinel-2 data in estimating phenophases of crops».

The difference between ground phenology and land surface phenology must be stressed. Ground phenology is defined by very precise stages that can only be observed on the field, like for instance the emergence of 50% of the seeds to the cotyledon stage. This ground phenology is used by agronomists, but it is very costly to monitor and suffers from observer biases. Remote sensing allows to observe a large number of fields without observer bias, but is less accurate and should be considered as a proxy for the ground phenology.

In this work, we choose the most widely used model for land surface phenology, that is, the double logistic for two main reasons. First of all, it has a small number of parameters (6), yet it is flexible enough to describe the phenology of annual crops in temperate areas. Second, as we will show in section 3, it is easy to link the parameters of the mathematical model to the phenological variables of interest, making the estimation of the latter straightforward.

In terms of estimation procedure, we propose an original approach which, to the best of our knowledge is absent from the literature. We use a probabilistic estimation based on Bayesian inference which allows to estimate the full (posterior) probability distribution of all the phenological variables given the satellite observations. The method can operate on irregularly sampled time series (which is the case in presence of cloudy acquisitions) and provides an estimation of the error of the double logistic model with respect to the observed data.

The approach does not need smoothing or gap-filling of the data as it is the case in most published approaches.

It is worth noting that the Bayesian approach proposed here is completely generic and can be applied to phenological models other than the 6-parameter double logistic.

Finally, all prior information about the phenology can easily be incorporated into the model via the prior distributions of the parameters. In this work, we choose rather uninformative priors, yet the results show accurate estimations.

3. The phenological model

As stated in the previous section, we choose to use a 6 parameter double logistic to model the crop phenology as observed through a vegetation index like NDVI. An example of such a model is shown in figure 1. The top plot shows the 2 slopes (growing in green and senescence in red) that will be used to define the phenometrics, and the bottom plot shows the values of those metrics.

The logistic function has the form:

$$f(t) = \frac{1}{1 + e^{\frac{t_0 - t}{t_1}}}, \quad (1)$$

where the t variable represents time in our application.

The double logistic is an affine scaling of the difference of 2 logistics:

$$g(t) = A(f_1(t) - f_2(t)) + B = A \left(\frac{1}{1 + e^{\frac{t_0 - t}{t_1}}} - \frac{1}{1 + e^{\frac{t_2 - t}{t_3}}} \right) + B, \quad (2)$$

where $A + B$ and B are respectively the maximum and the minimum values of $g(t)$. The other 4 parameters control the slopes of growing and senescence phases: t_0 (resp.

t_2) is the date of the maximum increasing (resp. decreasing) slopes and t_1 (resp. t_3) is related to the speed of growth (resp. senescence). The growing and senescence slopes can be obtained from the derivatives of the double logistic. Since

$$\frac{df}{dt}(t) = \frac{e^{\frac{t_0-t}{t_1}}}{t_1 \left(1 + e^{\frac{t_0-t}{t_1}}\right)^2} \quad (3)$$

we have that

$$g'(t) = \frac{dg}{dt}(t) = A \left(\frac{e^{\frac{t_0-t}{t_1}}}{t_1 \left(1 + e^{\frac{t_0-t}{t_1}}\right)^2} - \frac{e^{\frac{t_2-t}{t_3}}}{t_3 \left(1 + e^{\frac{t_2-t}{t_3}}\right)^2} \right). \quad (4)$$

From these expressions, we can derive the 6 phenology metrics of interest:

1. the minimum value of the vegetation index $NDVI_{min}$;
2. the maximum value, $NDVI_{max}$;
3. the start of the season date, SoS ;
4. the maturity date, Mat ;
5. the senescence date, Sen ;
6. the end of the season, EoS .

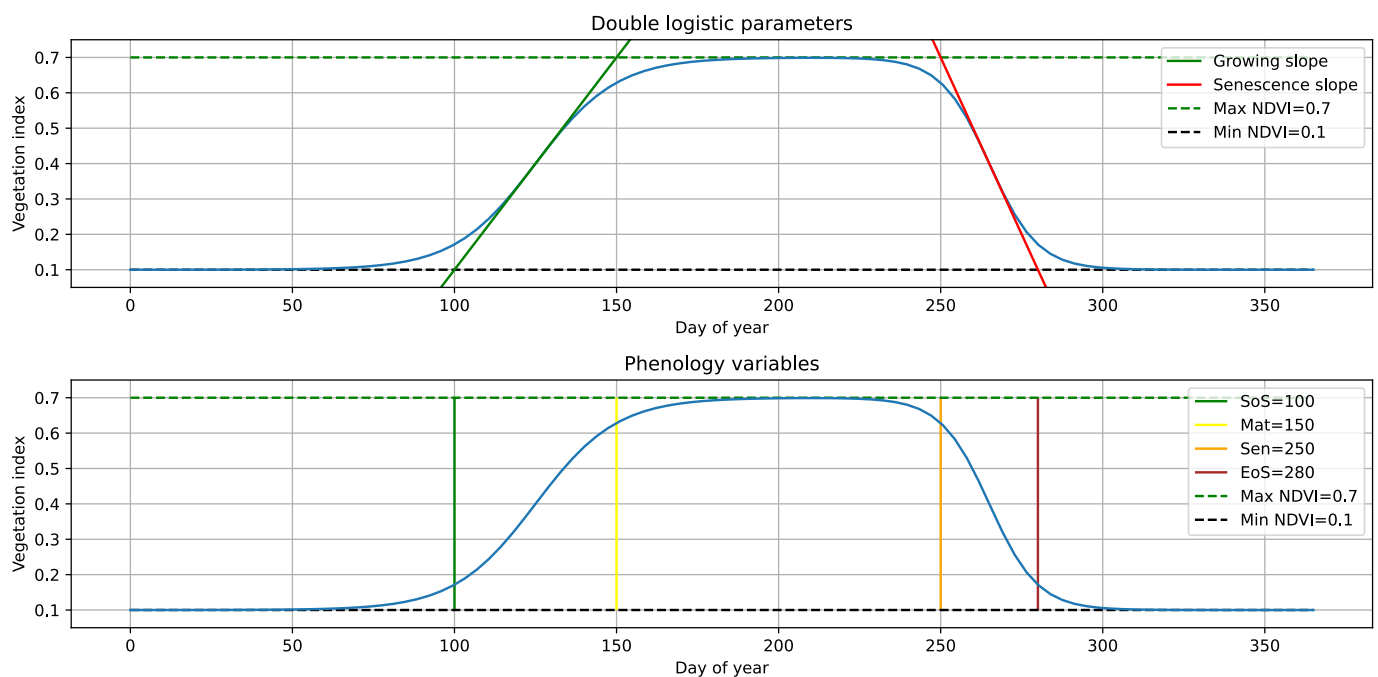


Figure 1. Double logistic function and associated parameters

For a crop with a well defined cycle, $NDVI_{min} \approx B$ and $NDVI_{max} \approx A + B$. We define the SoS (resp. EoS) as the date for which the growing (resp. senescence) slope line (see figure 1 top) crosses the horizontal axis corresponding to the $NDVI_{min}$ value. The dates for Mat and Sen are defined in an analogous way with respect to $NDVI_{max}$.

Let's take the growing slope (green line in figure 1 top) and assume that t_2 is much larger than t_0 . In this case, eq. (4) becomes:

$$g'(t_0) = \frac{dg}{dt}(t_0) = A \left(\frac{e^{\frac{t_0-t_0}{t_1}}}{t_1 \left(1 + e^{\frac{t_0-t_0}{t_1}}\right)^2} - \frac{e^{\frac{t_2-t_0}{t_3}}}{t_3 \left(1 + e^{\frac{t_2-t_0}{t_3}}\right)^2} \right) \approx A \left(\frac{1}{t_1(1+1)^2} - 0 \right), \quad (5)$$

and therefore,

$$g'(t_0) \approx \frac{A}{4t_1}. \quad (6)$$

If we take this derivative as the slope of the line which joins *SoS* and *Mat* in figure 1, we have that

$$g'(t_0) \approx \frac{A}{Mat - SoS}, \quad (7)$$

and therefore,

$$t_1 \approx \frac{Mat - SoS}{4}.$$

We can also assume

$$t_0 \approx \frac{Mat + SoS}{2}.$$

Using the same reasoning, we can approximate

$$t_3 \approx \frac{EoS - Sen}{4}$$

$$t_2 \approx \frac{EoS + Sen}{2}.$$

With these approximations, we have direct relation between the phenological variables and the parameters of the mathematical model. These approximations are inaccurate for very slow growing and senescence vegetation (as for instance fallows), but they are well fit for annual crops in temperate areas.

In the remainder, we will use $m(t; max_{NDVI}, min_{NDVI}, SoS, Mat, Sen, EoS)$ to designate the double logistic function where A, B, t_0, t_1, t_2, t_3 are computed from the approximations above.

4. Bayesian parameter retrieval

As stated in section 2, the differences between ground and land surface phenology as well as the assumptions made by the phenology models used, can lead to inaccurate phenology characterizations. Being able to assess the uncertainty of the estimations is therefore crucial for downstream uses, either by models or by decision makers.

In order to obtain uncertainty estimates of retrieved parameters from an NDVI time profile, we will use a Bayesian approach. The interest of Bayesian estimation is that, given a data generation model (the phenology model described in the previous section), one can estimate the probability distribution of each of the parameters of the model (the phenological variables) given an observation (an NDVI time profile).

The Bayes rule is written as

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} \quad (8)$$

where:

- \mathcal{D} is the observed data, in our case a time profile of NDVI;
- θ are the parameters of the model, that is the $NDVI_{min}, NDVI_{max}, SoS, Mat, Sen, EoS$ in our case;

- $p(\theta|\mathcal{D})$ is the *posterior distribution* of the parameters, that is, the probability distribution of the unknown parameters θ once we have observed some data \mathcal{D} ;
- $p(\mathcal{D}|\theta)$ is called the *likelihood* and corresponds to the probability of observing some data when the parameters are known;
- $p(\theta)$ is the *prior distribution* of the parameters and
- $p(\mathcal{D})$ is the probability of observing some data.

Usually, we would want to find the values of θ which maximize $p(\theta|\mathcal{D})$ for a given observed time profile \mathcal{D} . This is what is called the *maximum a posteriori* (MAP) estimate and for all intents and purposes, is not very different that a direct parameter optimization with respect to the RMSE criterion. Actually, the solutions are equivalent for posterior distributions which are symmetrical with respect to the mode.

What is interesting here, is that we can estimate the complete distribution of the parameters for a given observation, and therefore, have an accurate picture of the uncertainty associated to each parameter.

The question now is how to choose the 3 terms in the right-hand side of equation (8).

4.1. The probability of the observation

The term $p(\mathcal{D})$ is the most difficult to set, since it represents the joint probability of the data points in the time profile. Fortunately, for a given time profile, $p(\mathcal{D})$ is a constant which does not depend on θ . The role of this constant is to ensure that $p(\theta|\mathcal{D})$ has a unit integral. We can therefore ignore this term and normalize $p(\theta|\mathcal{D})$ after the fact.

In other words, we need to estimate

$$p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta) \quad (9)$$

and normalize to ensure

$$\int_{\theta} p(\theta|\mathcal{D})d\theta = 1 \quad (10)$$

4.2. The likelihood of the data

The term $p(\mathcal{D}|\theta)$ represents the process of generation of the data once the parameters of the model are known. The phenological model gives a deterministic link between the parameters and the observation and we need a conditional probability. Since the model is a simplification of the physical reality, we can build a probability distribution to account for the modeling errors. This can be done by modeling the observations as a noisy version of the deterministic model.

Let \mathbf{m} be a deterministic vector such as $[m(t_1|\theta), \dots, m(t_T|\theta)]$ and

$$\theta = [\max_{NDVI}, \min_{NDVI}, SoS, Mat, Sen, EoS, \sigma]$$

We can write the likelihood of the observations as

$$p(\mathcal{D}|\theta) \sim \mathcal{N}(\mathbf{m}, \sigma^2),$$

where σ is the standard deviation of a Gaussian noise and can be estimated as part of the Bayesian inference.

4.3. The prior distribution of the parameters

The term $p(\theta)$ is where we inject our prior knowledge of the parameters. Although this freedom may seem subjective, it is not different than setting initial values for the parameters when using gradient descent or simplex optimization algorithms. The advantage here is that we can provide much richer information than an initial guess.

For instance, let's take the *Mat* parameter for a summer crop. Since it corresponds to the end of the growing stage, we know that it will be between early June and late August (in the case of Southern France). This can be represented by a uniform distribution between day 150 and day 240. If we don't know whether we have a summer or a winter crop, we can widen the distribution to start on day 50. Conversely, if we have knowledge about the climate and the exact type of crop, we can narrow the distribution.

Instead of a uniform distribution, we can use other shapes and introduce a more nuanced information about the probability distribution of each date.

We can also introduce dependencies and constraints between the parameters if needed. For instance, the senescence date is always after the starting date.

However, it is important to point out that the influence of the prior distributions decreases when the number of observation points increases.

4.4. Bayesian inference using Markov Chain Monte Carlo

There exist many techniques for performing Bayesian inference, that is, obtaining the posterior distributions of the model parameters given the observations. Describing them is out of the scope of this paper. For our work, since we do not have a closed expression for the full posterior of the phenological parameters, we will use Markov Chain Monte Carlo (MCMC) to obtain samples from it. More precisely, we use Hamiltonian Monte Carlo as per the NUTS algorithm [27] as implemented in the NumPyro library [28,29].

4.5. A Bayesian model for phenology

As stated above, in order to implement Bayesian inference through MCMC we need to define the likelihood for the observed data. Since we know that the double logistic model is an approximation, we will model the observations as a noisy version of such a model:

$$NDVI(t) = \mathbf{m}(t|\boldsymbol{\theta}) + \varepsilon(t), \quad (11)$$

where $\varepsilon(t)$ is an i.i.d. Gaussian noise. Therefore:

$$NDVI(t) \sim \mathcal{N}(\mu(t), \sigma^2) \quad (12)$$

$$\mu(t) = \mathbf{m}(t, \max_{NDVI}, \min_{NDVI}, SoS, Mat, Sen, EoS), \quad (13)$$

where $\mathbf{m}(t, \max_{NDVI}, \min_{NDVI}, SoS, Mat, Sen, EoS)$ was defined at the end of section 3. With this model, the vector of parameters is

$$\vec{\theta} = (\sigma, \max_{NDVI}, \min_{NDVI}, SoS, Mat, Sen, EoS)$$

and includes the variance of the noise. Given this model, we can simulate NDVI profiles by choosing values for the parameters, obtaining the mean $\mu(t)$ and sampling a Gaussian distribution.

We now need to choose prior distributions for all the parameters to be estimated. We will choose uniform distributions for simplicity. These are rather uninformative priors.

For σ , which represents the standard deviation of the noise in the observations, we will choose a maximum value of 0.1 which corresponds to 10% of the maximum expected range for NDVI values. For \min_{NDVI} , we will define a range between 0 (bare soil) and 0.4 (presence of vegetation). The maximum value of NDVI will be defined relative to the minimum value: we assume that a crop with a typical phenology value will have a \max_{NDVI} at least 0.3 higher than \min_{NDVI} , and that the highest value will not be higher than 1.

The 4 dates for the phenological stages are each one defined in terms of the previous as follows: *EoS* is allowed to be right after *Sen* and up to 90 days later. *Sen* is defined in the same way with respect to *Mat*, and *Mat* follows the same rationale with respect to *SoS*.

For *SoS* we would need to give a very wide prior to take into account winter and summer crops. Instead of doing that, we introduce an additional (latent) variable *SummerLike* which will model the probability of summer crop. This probability is used to adjust the starting point of the interval of prior values for *SoS*. We assume that the earliest *SoS* for a winter crop is on day 30 (end of January) and that the earliest summer crop can have an *SoS* of 120 (late April). It is interesting to note that this latent variable will be available after the inference, and therefore, the algorithm also provides a classification output for winter and summer crops if we threshold this probability.

With the above considerations, the priors for the parameters are set as follows:

$$\sigma \sim \mathcal{U}(0, 0.1) \quad (14)$$

$$\min_{NDVI} \sim \mathcal{U}(0, 0.4) \quad (15)$$

$$\max_{NDVI} \sim \mathcal{U}(\min_{NDVI} + 0.3, 1) \quad (16)$$

$$EoS \sim \mathcal{U}(Sen, Sen + 90) \quad (17)$$

$$Sen \sim \mathcal{U}(Mat, Mat + 90) \quad (18)$$

$$Mat \sim \mathcal{U}(SoS, SoS + 90) \quad (19)$$

$$SoS \sim \mathcal{U}(\Delta_{SoS}, \Delta_{SoS} + 90) \quad (20)$$

$$SummerLike \sim \mathcal{U}(0, 1) \quad (21)$$

$$\Delta_{SoS} = 90 \times SummerLike + 30 \quad (22)$$

4.6. Prior predictive checks

To assess the ability of our model to represent realistic crop NDVI profiles, we can perform simulations. In the case of Bayesian inference, these are called prior predictive checks. Using equations 14 through 22, we can sample values of the parameters, use them in equation 13 to produce an NDVI profile and finally sample an observation using equation 12.

An illustration of such prior predictive checks is given in figure 2. The legends show the values for \max_{NDVI} , \min_{NDVI} , *SoS*, *Mat*, *Sen*, *EoS* and green (resp. red) plots indicate a probability of Summer crop lower (resp. higher) than 0.5.

One can observe a wide variety of behaviors in terms of growing rates, dates, noise level and NDVI dynamics.

5. Data sets

To illustrate the application of phenological parameter extraction using the proposed Bayesian model, 2 different data sets will be used. The first one is a real data set composed of satellite observations of annual crops and will be used for qualitative validation. A second data set, composed of simulated crop NDVI time profiles will be used for quantitative validation of the approach.

5.1. Real data set

The real data set is extracted from the public data sets built for the Sentinel-2 Agriculture project [30]. The data collection procedure and protocol are described in [31].

For this paper, we will use a smaller subset of 25 crop fields of the France MiPy site (near Toulouse in South-West France) for which SPOT Take 5 [32] and Landsat-8 acquisitions were available. Three types of crops are present: winter wheat, sunflower and maize. The number of acquisition dates for one year range between 12 and 18 depending of the field.

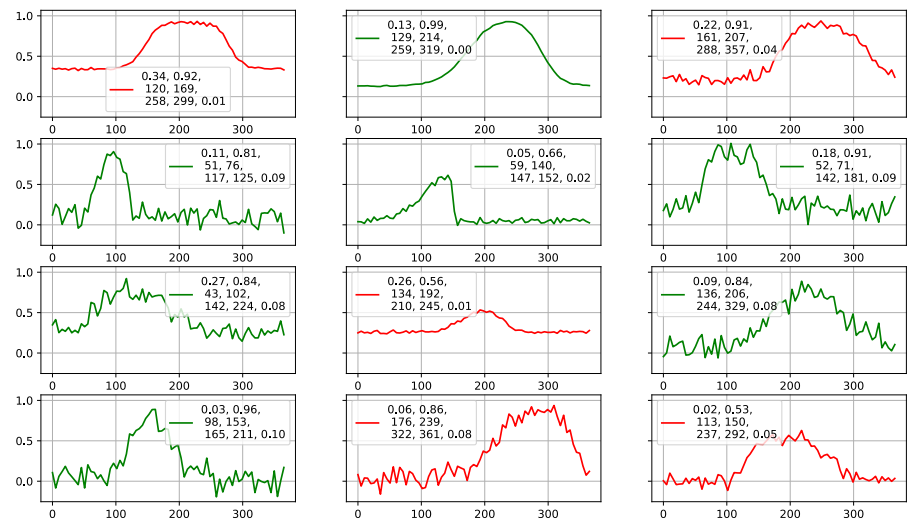


Figure 2. Prior predictive checks of the Bayesian phenology model. The legends show the values for max_{NDVI} , min_{NDVI} , SoS , Mat , Sen , EoS , σ and green (resp. red) plots indicate a probability of Summer crop lower (resp. higher) than 0.5.

Although no reference data for the phenological variables is available for these time series, the results can be visually inspected to assess the usefulness and validity of the Bayesian estimations.

5.2. Simulated data set

The generation of a synthetic but realistic data set is important to provide a quantitative assessment of the performances of the method. If generating NDVI time profiles is straightforward with the model used for the prior predictive checks, an additional step is needed in order to obtain data close to what a satellite may observe. Although having a regular revisit cycle, optical satellites suffer from the presence of clouds. For instance, the 5-day revisit cycle of the Sentinel-2 constellation provides, on average, about 20 clear acquisitions per pixel over Metropolitan France, instead of the 73 theoretically possible (these values are doubled on the areas where 2 orbits overlap).

In order to simulate this irregular acquisition pattern, we will generate NDVI profiles with a 5 day time step and then randomly choose N dates.

Figure 3 illustrates some simulations. The blue solid lines correspond to the full time profiles and the red lines correspond to the sub-sampled profiles. In these examples, a time step of 5 days followed by a random selection of 20 dates is shown. The red dots correspond to the *observed dates*. As one can see, the irregular sampling may lead to missing the phenological key dates and produce distortions on the shape of the underlying double logistic function.

We simulated 3 data sets with the same sampling step of 5 days and different number of observed dates (15, 20 and 30) in order to assess the performances of the method for cases representative of Sentinel-2 acquisitions. Each data set contains 5000 samples.

6. Results

We present the results of the experiments on both the synthetic and the real data sets. For all experiments the NUTS algorithm [27] as implemented in the NumPyro library [28,29] was used with 10000 samples after the generation of 100 warming steps.

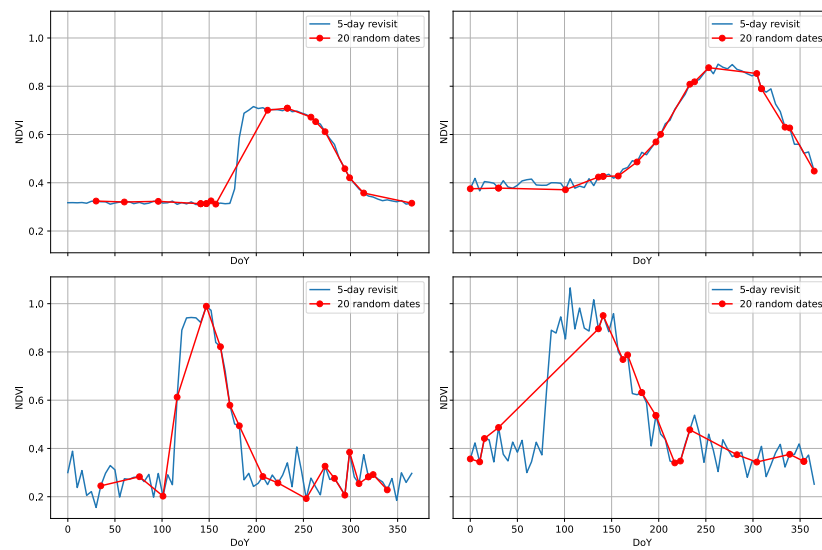


Figure 3. Simulation examples of crop NDVI time profiles using the double logistic model, noisy observations and 20 random dates on a 5 day revisit cycle. The red dots are joined by straight lines to emphasize the missing information due to missing dates.

6.1. Synthetic data set

As stated in section 5.2, 3 simulated data sets corresponding to 15, 20 and 30 available dates on a 5 day revisit cycle are used to give a quantitative assessment of the performances of the methods. For this validation, only the simulations for which the NDVI amplitude (maximum minus minimum) is greater than 0.3 are considered, since smaller amplitudes barely represent the phenology of real crops.

The quantitative assessment is based on 2 metrics. The first one is the mean absolute error (MAE) of a point estimate of each variable of the phenological model. This error is expressed in the units of the variable (a probability for *SummerLike*, an NDVI value for $NDVI_{max}$, $NDVI_{min}$ and σ , and days for the 4 phenological dates *SoS*, *Maturity*, *Senescence* and *EoS*). For this point estimate we choose to use the median of the posterior marginal distributions.

Since one of the interests of using a Bayesian approach is the availability of a full posterior distribution, the second metric evaluates the frequency with which the real value falls within the 5th and the 95th quantiles of the posterior distribution of each variable.

Table 1 provides the results of the quantitative validation on the simulated data set for the 3 temporal samplings and the 2 metrics. As one can observe, the errors on the NDVI values (min and max) are lower than 0.05 and the errors on the dates are always lower than the average revisit cycle, which is of 24 days for 15 dates, 18 days for 20 dates and 12 days for 30 dates. It is worth noting that doubling the number of acquisitions (from 15 to 30) only improves the dates estimations by about 2 days.

In terms of uncertainty estimation, the chosen interval covers the real values in more than 80% of the cases. The coverage increases when the number of observed dates decreases, probably due to the increase of uncertainty with the reduction of observations.

6.2. Real data set

Figures 4 through 8 illustrate the results for the 25 real NDVI time profiles. Each pane in each figure presents the results of the Bayesian inference for an individual profile. For each profile, the inference is performed with 1/4, 1/2, 3/4 of the observations and the full time profile to simulate the inference during the season. This corresponds to each row of each pane. The 4 leftmost columns of each pane present the posterior distributions of the phenological dates (*SoS*, *Mat*, *Sen* and *EoS*). The second rightmost

Table 1: Quantitative performances of the Bayesian inference for 15, 20 and 30 dates. Mean absolute error (MAE) of the point estimate using the median and proportion of estimates falling between the 5th and the 95th quantiles of the posterior marginal distributions. The MAE is measured in probability units for *SummerLike*, in NDVI units for $NDVI_{max}$, $NDVI_{min}$ and σ and in days for *SoS*, *Mat*, *Sen* and *EoS*.

	MAE (15d)	5-95th (15d)	MAE (20d)	5-95th (20d)	MAE (30d)	5-95th (30d)
SummerLike	0.19	88.68	0.18	87.60	0.18	87.72
$NDVI_{max}$	0.05	88.04	0.05	86.32	0.04	84.74
$NDVI_{min}$	0.02	85.28	0.02	83.44	0.01	82.56
<i>SoS</i>	9.93	84.88	8.79	83.38	7.49	81.42
<i>Mat</i>	11.91	84.82	10.72	83.84	9.51	82.52
<i>Sen</i>	12.39	83.22	11.70	81.74	10.26	80.22
<i>EoS</i>	12.65	83.26	11.71	82.02	10.52	79.52
σ	0.01	86.00	0.01	84.92	0.01	81.80

column corresponds to the posterior distribution of the latent variable representing the probability of a Summer crop. The rightmost column shows in blue the observations (the real NDVI data) and the 50% (dark orange) and 95% (light orange) probability intervals of the posterior predictive distributions. The posterior predictive distributions are obtained by running the phenological model with samples drawn from the joint posterior distribution of the model parameters.

In order to help the reader interpret the figures, let's take the profile number 6, at the top of figure 5. On the first row, the inference with only 4 dates early in the year is presented. In this case we observe a wide distribution for all variables: *SoS* is between day 50 and day 200, *Mat* between 70 and 300, *Sen* between 100 and 350 and *EoS* between 100 and 400. The probability of Summer crop is nearly uniformly distributed between 0 and 1, with a slight skew towards Summer, due to the fact that the 4 available dates have a low NDVI value. The posterior predictive highest density intervals (HDI) in the rightmost column show this wide spread of uncertainty. On the second row, 4 additional dates are used for the inference. Since these dates confirm the low NDVI values, the probability of Summer crop increases significantly. As a consequence, the uncertainty on *SoS* is reduced and most of the probability is between day 150 and day 200. The uncertainty on the other 3 phenological dates is also reduced, but this narrowing decreases for dates further in the season, which is expected, since it is more difficult to predict further in the future. The HDI in the posterior predictive distribution are narrowed close to the observed dates and widen afterwards. On the third row, 12 dates are used and *SoS* and *Mat* are fully observed. This greatly reduces the spread of the posterior distributions of these variables as illustrated in the 2 leftmost columns. Almost all probability of Summer crop is above 0.5, which is correct for a Sunflower crop in the study area. The posterior distributions for *Sen* and *SoS* are also narrower, but with a long tail towards later dates, since these dates are not yet observed. The slight decrease in NDVI could be caused by observation noise. This is reflected in the wide intervals on the posterior predictive distribution on the decreasing slope. The fourth row illustrates the inference with 16 dates representing the full season. We see that these additional observations do not reduce the uncertainties on *SoS* and *Mat* which were already fully observed, but the posterior distributions of *Sen* and *SoS* get narrower. It is interesting to note that the uncertainties for *Mat* and *Sen* are significantly larger than for *SoS* and *EoS* due to the fact that the maturity plateau is nearly non existent for this time profile. Also, the spread of *SoS* posterior values is greater than those of *EoS* due to the slight increase of NDVI between days 50 and 100 (probably due to vegetation regrowth before the spring tillage). These events are not taken into account by the simple double logistic model, but the Bayesian inference is able to detect the issue.

For well behaved profiles, the Bayesian inference produces very narrow distributions at the end of the season. This is the case for instance for profile 24 in figure 8, which corresponds to an irrigated maize crop. The reduction of the spread of the posterior

predictive distribution (rightmost column) as a consequence gives nearly a perfect fit of the observed profile, indicating that the double logistic model is a good approximation in this case. This behavior is observed for all maize crops (profiles 17 through 25) except for profiles 17 and 18 which show important vegetation presence at the beginning and the end of the season, resulting in higher uncertainties for *SoS* and *EoS*. Again, this shows the interest of having full posterior distributions of the phenological variables to assess the adequacy of the model to the observed data. The same kind of conclusions can be derived from winter wheat crops (profiles 1 through 5) where the vegetation can emerge before day of year 1, and therefore it is not observed.

7. Discussion

The results presented in section 6.2 illustrate the power of the Bayesian approach. On one hand, the model is defined in terms of the application domain (equations (12) through (14)) in a way that is independent of the inference approach. This allows to introduce all prior knowledge about the parameters of the model. On the other hand, the inference is able to provide full posterior distributions of the parameters given the available observations. The uncertainties are easy to interpret and can be used in downstream models if needed.

There is however one drawback to the inference approach used here. The MCMC technique has a high computational cost, since a Markov chain has to be used for each time profile, which implies generating hundreds of Monte Carlo samples of the joint probability distribution of the parameters. This approach will therefore not scale to an estimation over large geographical areas. However, it can be used at the individual field level in precision agriculture applications.

In order to be able to scale the Bayesian inference, other algorithms than MCMC should be used. One approach would be using variational inference [33,34] where the posterior distributions would be approximated by a parametric family (for instance normal distributions) and gradient descent optimization would be used to estimate the parameters of these distributions (mean and standard deviation in the case of normal distributions). This approach can still be costly, since the optimization has to be run for each time profile. Another approach would be using amortized variational inference, where the optimization is replaced by a regression which is trained offline: for instance, a neural network is used to estimate the parameters of the variational distributions of the parameters (mean and standard deviation) from the observed time profiles. The training data can be generated with the phenological model (synthetic data) or the training can be done using real data and variational auto-encoders [35].

The trade-off between accuracy of the posterior distributions (the variational families are just approximations of the real posteriors) and computational cost has to be analyzed for each application case.

8. Conclusion

In this paper, we have proposed a Bayesian approach for the estimation of crop phenological parameters using vegetation indices derived from high resolution optical satellite image time series. The phenological model is based on a 6 parameter double logistic function which is well suited for annual crops in temperate areas. The interest of such a phenological model is providing a direct link between the mathematical function and the phenological markers. Performing Bayesian inference through Markov Chain Monte Carlo (MCMC), one can obtain samples of the joint posterior distribution of the parameters given the observation. In this way, the full distribution is available for each vegetation index time profile, which allows to perform point estimates as in other estimation procedures, but provides a very rich characterization of the uncertainties.

The proposed approach has been applied to simulated data to provide a quantitative assessment of the expected performances. The results show that the key phenological dates can be estimated with an accuracy higher than the average sampling rate of

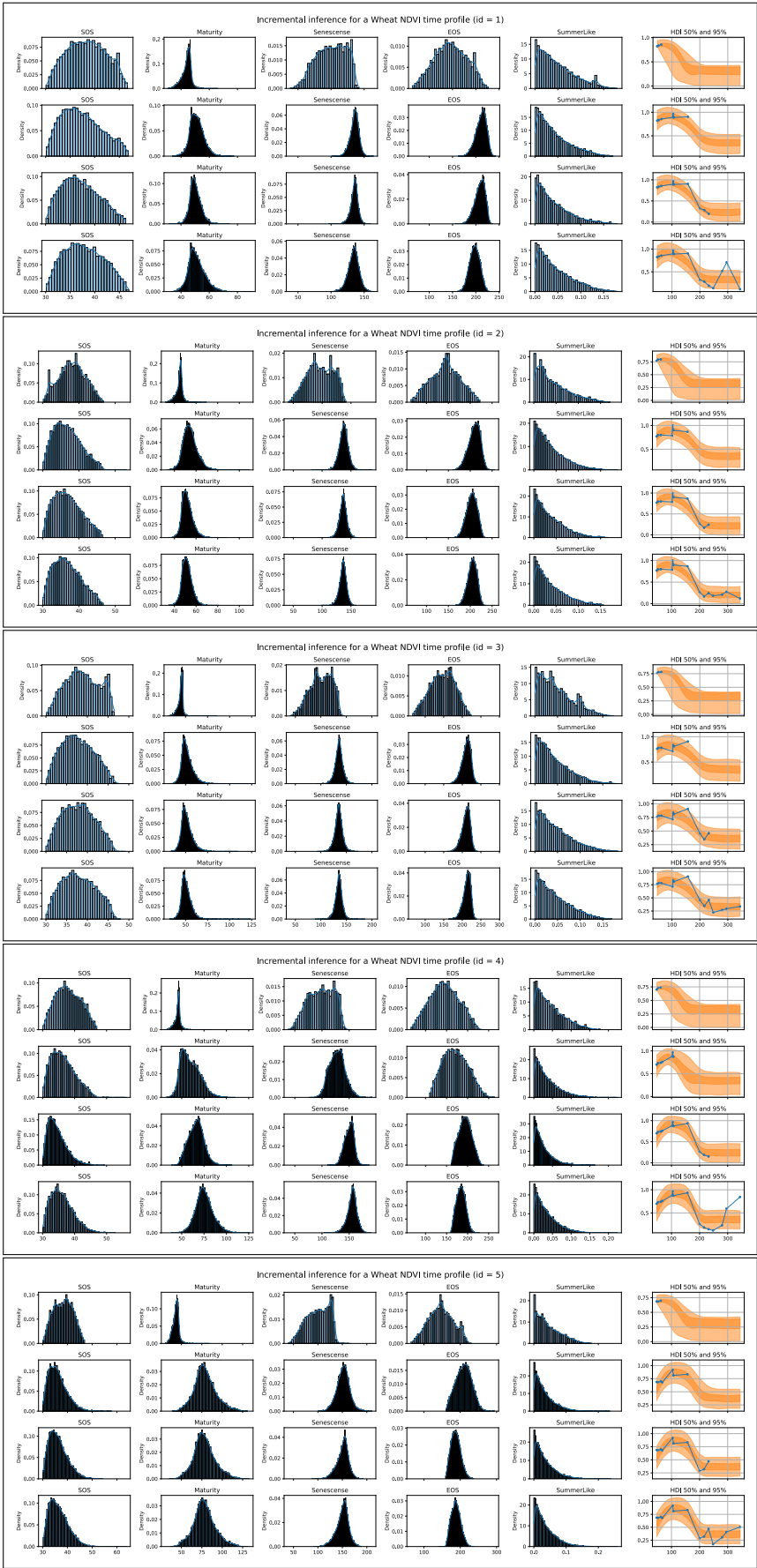


Figure 4. Incremental inference of real profiles 1 - 5

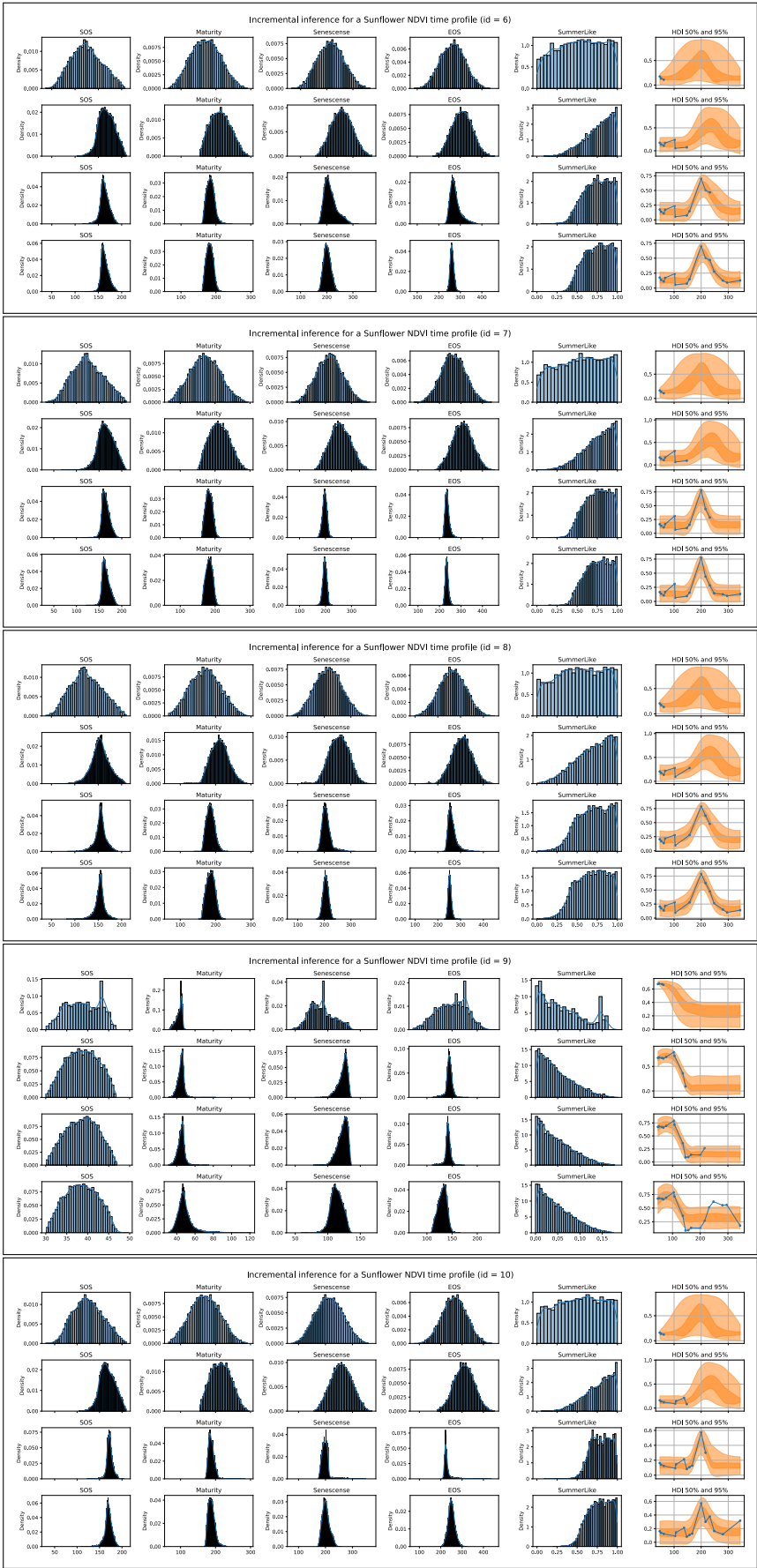


Figure 5. Incremental inference of real profiles 6 - 10

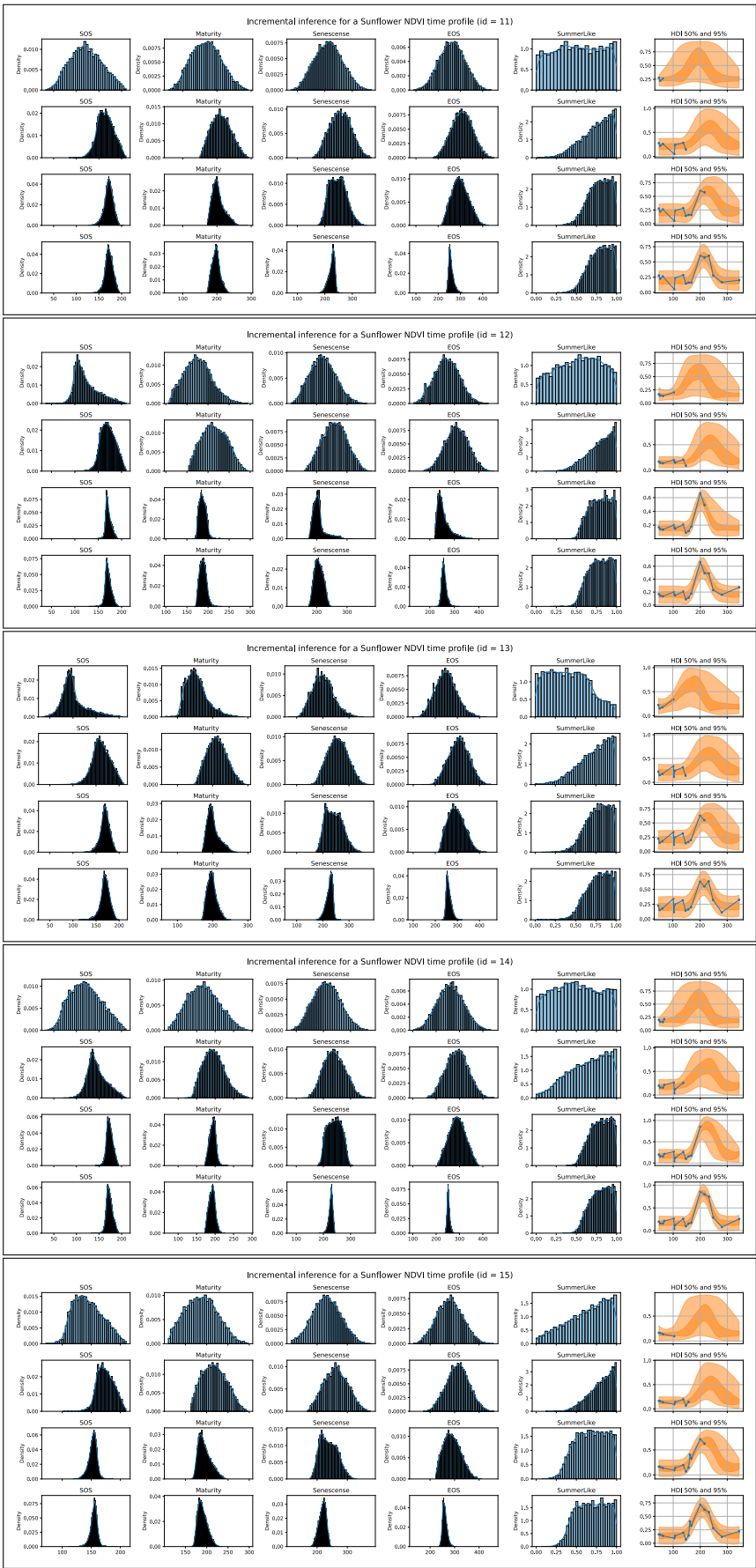


Figure 6. Incremental inference of real profiles 11 - 15

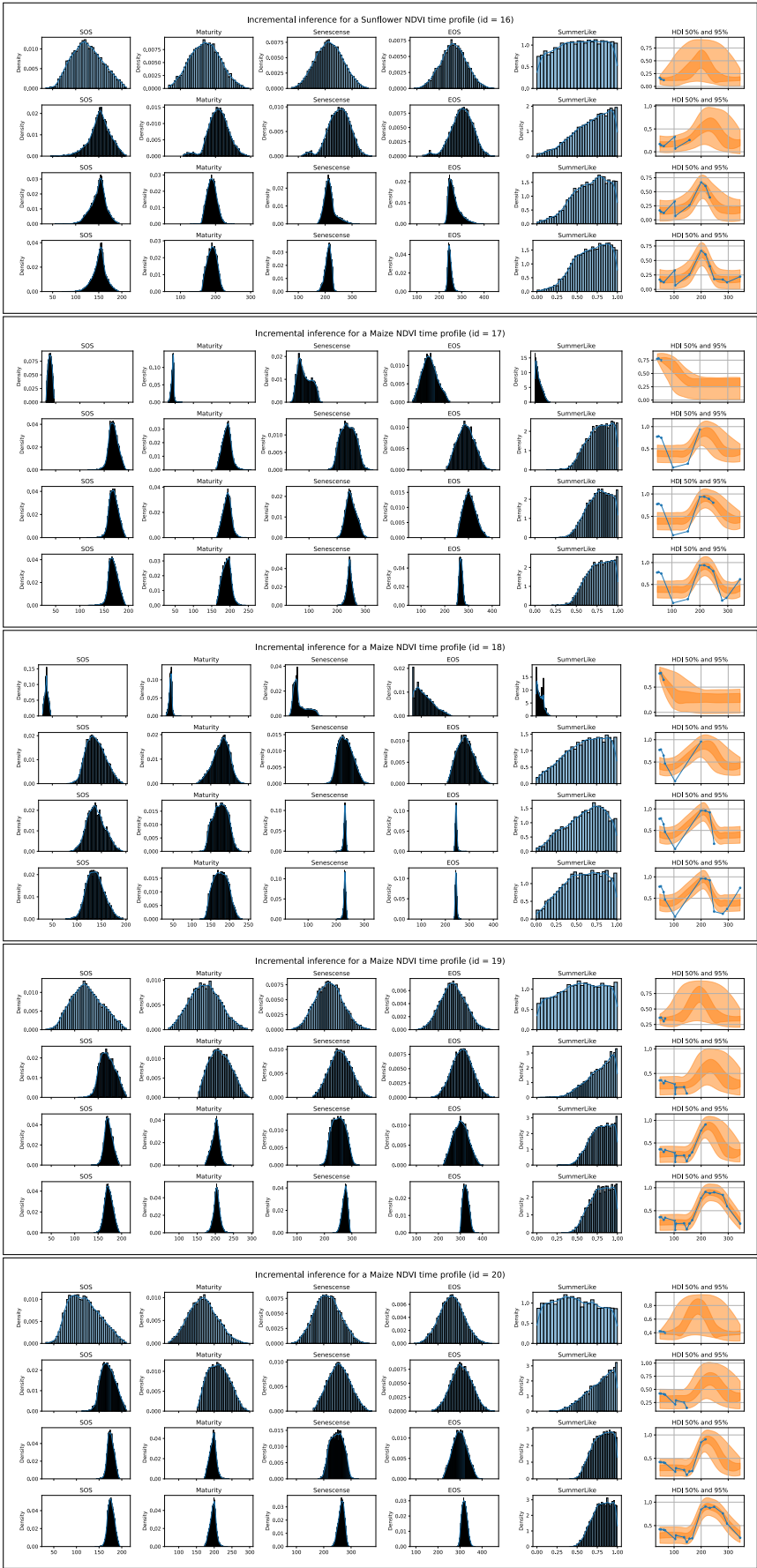


Figure 7. Incremental inference of real profiles 16 - 20

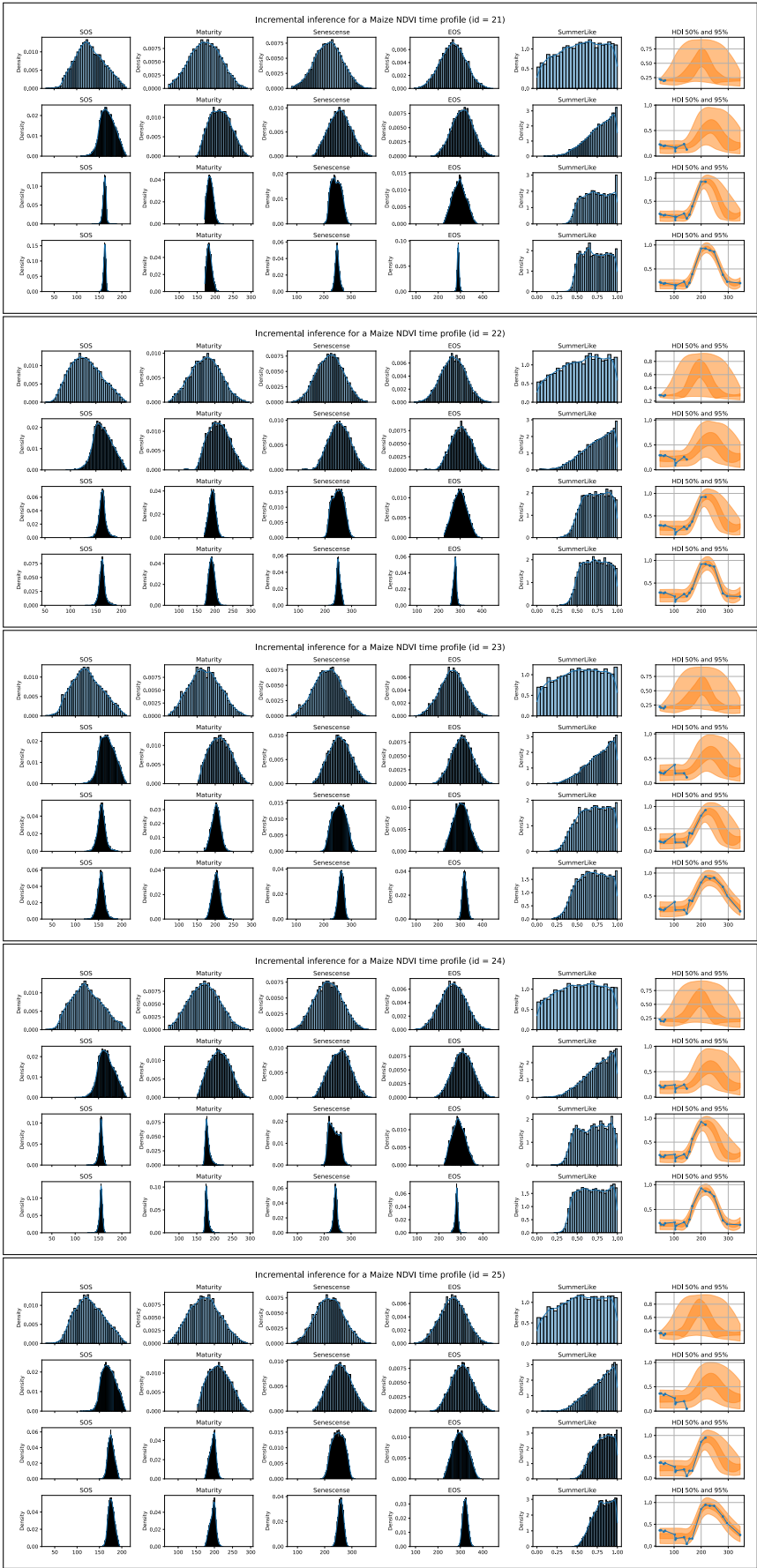


Figure 8. Incremental inference of real profiles 21 - 25

the sensor. The uncertainties provided by the method are highly correlated with the theoretical confidence intervals. The application to real satellite data illustrates the use of the method in real-time situations where the phenology parameters are estimated during the agricultural season. The uncertainties are reduced with the use of consecutive image acquisitions and the posterior predictive distributions can be used to forecast the evolution of the vegetation.

The main limitation of the approach proposed in this paper is its computational cost. Indeed the MCMC needs drawing hundreds of samples per time profile. Amortized variational approaches could be used to scale the procedure to pixel-based inference over large geographical areas.

Data Availability Statement: The source code and the data to reproduce the results presented in this study are openly available in Zenodo at 10.5281/zenodo.5793195.

Acknowledgments: The author would like to thank his colleagues at CESBIO (Valentine Bellet, Jean-François Dejoux, Mathieu Fauvel, Julien Michel and Silvia Valero) for providing insightful comments and corrections to the drafts of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

1. Tilman, D. Forecasting Agriculturally Driven Global Environmental Change. *Science* **2001**, *292*, 281–284. doi:10.1126/science.1057544.
2. Foley, J.A.; Ramankutty, N.; Brauman, K.A.; Cassidy, E.S.; Gerber, J.S.; Johnston, M.; Mueller, N.D.; O’Connell, C.; Ray, D.K.; West, P.C.; Balzer, C.; Bennett, E.M.; Carpenter, S.R.; Hill, J.; Monfreda, C.; Polasky, S.; Rockström, J.; Sheehan, J.; Siebert, S.; Tilman, D.; Zaks, D.P.M. Solutions for a Cultivated Planet. *Nature* **2011**, *478*, 337–342. doi:10.1038/nature10452.
3. Benayas, J.M.R.; Bullock, J.M. Restoration of Biodiversity and Ecosystem Services on Agricultural Land. *Ecosystems* **2012**, *15*, 883–899. doi:10.1007/s10021-012-9552-0.
4. Schwilch, G.; Bestelmeyer, B.; Bunning, S.; Critchley, W.; Herrick, J.; Kellner, K.; Liniger, H.; Nachtergaele, F.; Ritsema, C.; Schuster, B.; Tabo, R.; van Lynden, G.; Winslow, M. Experiences in Monitoring and Assessment of Sustainable Land Management. *Land Degrad. Dev.* **2010**, *22*, 214–225. doi:10.1002/ldr.1040.
5. Drusch, M.; Bello, U.D.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; Meygret, A.; Spoto, F.; Sy, O.; Marchese, F.; Bargellini, P. Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* **2012**, *120*, 25–36. doi:10.1016/j.rse.2011.11.026.
6. Bolton, D.K.; Friedl, M.A. Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics. *Agricultural and Forest Meteorology* **2013**, *173*, 74–84. doi:10.1016/j.agrformet.2013.01.007.
7. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al.. Near Real-Time Agriculture Monitoring At National Scale At Parcel Resolution: Performance Assessment of the Sen2-agri Automated System in Various Cropping Systems Around the World. *Remote Sensing of Environment* **2019**, *221*, 551–568. doi:10.1016/j.rse.2018.11.007.
8. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.C.; Gao, F.; Reed, B.C.; Huete, A. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment* **2003**, *84*, 471–475. doi:10.1016/s0034-4257(02)00135-9.
9. FISHER, J.; MUSTARD, J.; VADEBONCOEUR, M. Green Leaf Phenology At Landsat Resolution: Scaling From the Field To the Satellite. *Remote Sensing of Environment* **2006**, *100*, 265–279. doi:10.1016/j.rse.2005.10.022.
10. Schwartz, M.D.; Reed, B.C. Surface Phenology and Satellite Sensor-Derived Onset of Greenness: An Initial Comparison. *International Journal of Remote Sensing* **1999**, *20*, 3451–3457. doi:10.1080/014311699211499.
11. White, M.A.; Thornton, P.E.; Running, S.W. A Continental Phenology Model for Monitoring Vegetation Responses To Interannual Climatic Variability. *Global Biogeochem. Cycles* **1997**, *11*, 217–234. doi:10.1029/97gb00330.
12. Thayn, J.B.; Price, K.P. Julian Dates and Introduced Temporal Error in Remote Sensing Vegetation Phenology Studies. *International Journal of Remote Sensing* **2008**, *29*, 6045–6049. doi:10.1080/01431160802235829.
13. Fisher, J.I.; Mustard, J.F. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment* **2007**, *109*, 261–273. doi:10.1016/j.rse.2007.01.004.
14. Xu, H.; Twine, T.; Yang, X. Evaluating Remotely Sensed Phenological Metrics in a Dynamic Ecosystem Model. *Remote Sensing* **2014**, *6*, 4660–4686. doi:10.3390/rs6064660.
15. de Beurs, K.M.; Henebry, G.M. Spatio-Temporal Statistical Methods for Modelling Land Surface Phenology. *Phenological Research* **2009**, pp. 177–208. doi:10.1007/978-90-481-3335-2_9.
16. Thayn, J.B. Assessing Vegetation Cover on the Date of Satellite-Derived Start of Spring. *Remote Sensing Letters* **2012**, *3*, 721–728. doi:10.1080/2150704x.2012.674227.
17. Beck, P.S.; Atzberger, C.; Høgda, K.A.; Johansen, B.; Skidmore, A.K. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment* **2006**, *100*, 321–334. doi:10.1016/j.rse.2005.10.021.

18. Schwartz, M.D.; Hanes, J.M. Intercomparing Multiple Measures of the Onset of Spring in Eastern North America. *Int. J. Climatol.* **2009**, *30*, 1614–1626. doi:10.1002/joc.2008.
19. White, K.; Pontius, J.; Schaberg, P. Remote Sensing of Spring Phenology in Northeastern Forests: A Comparison of Methods, Field Metrics and Sources of Uncertainty. *Remote Sensing of Environment* **2014**, *148*, 97–107. doi:10.1016/j.rse.2014.03.017.
20. Cong, N.; Piao, S.; Chen, A.; Wang, X.; Lin, X.; Chen, S.; Han, S.; Zhou, G.; Zhang, X. Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agricultural and Forest Meteorology* **2012**, *165*, 104–113. doi:10.1016/j.agrformet.2012.06.009.
21. Balzter, H.; Gerard, F.; George, C.; Weedon, G.; Grey, W.; Combal, B.; Bartholomé, E.; Bartalev, S.; Los, S. Coupling of Vegetation Growing Season Anomalies and Fire Activity With Hemispheric and Regional-Scale Climate Patterns in Central and East Siberia. *Journal of Climate* **2007**, *20*, 3713–3729. doi:10.1175/jcli4226.
22. Reed, B.C.; Brown, J.F.; VanderZee, D.; Loveland, T.R.; Merchant, J.W.; Ohlen, D.O. Measuring Phenological Variability From Satellite Imagery. *Journal of Vegetation Science* **1994**, *5*, 703–714. doi:10.2307/3235884.
23. Moulin, S.; Kergoat, L.; Viovy, N.; Dedieu, G. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. *Journal of Climate* **1997**, *10*, 1154–1170.
24. Pan, Z.; Huang, J.; Zhou, Q.; Wang, L.; Cheng, Y.; Zhang, H.; Blackburn, G.A.; Yan, J.; Liu, J. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and Geoinformation* **2015**, *34*, 188–197. doi:10.1016/j.jag.2014.08.011.
25. Zeng, L.; Wardlow, B.D.; Xiang, D.; Hu, S.; Li, D. A Review of Vegetation Phenological Metrics Extraction Using Time-Series, Multispectral Satellite Data. *Remote Sensing of Environment* **2020**, *237*, 111511. doi:10.1016/j.rse.2019.111511.
26. Misra, G.; Cawkwell, F.; Wingler, A. Status of Phenological Research Using Sentinel-2 Data: a Review. *Remote Sensing* **2020**, *12*, 2760. doi:10.3390/rs12172760.
27. Hoffman, M.D.; Gelman, A.; others. The No-U-turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **2014**, *15*, 1593–1623.
28. Phan, D.; Pradhan, N.; Jankowiak, M. Composable Effects for Flexible and Accelerated Probabilistic Programming in Numpyro. *CoRR* **2019**, [arXiv:stat.ML/1912.11554].
29. Bingham, E.; Chen, J.P.; Jankowiak, M.; Obermeyer, F.; Pradhan, N.; Karaletsos, T.; Singh, R.; Szerlip, P.; Horsfall, P.; Goodman, N.D. Pyro: Deep Universal Probabilistic Programming. *CoRR* **2018**, [arXiv:cs.LG/1810.09538].
30. Bontemps, S.; Arias, M.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Morin, D.; Rabaute, T.; Savinaud, M.; et al. x201c;sentinel-2 for Agriculturex201d; Supporting Global Agriculture Monitoring. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* **2015**. doi:10.1109/igarss.2015.7326748.
31. Bontemps, S.; Arias, M.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Matton, N.; Morin, D.; Popescu, R.; Rabaute, T.; Savinaud, M.; Sepulcre, G.; Valero, S.; Ahmad, I.; Bégué, A.; Bingfang, W.; de Abelleira, D.; Diarra, A.; Dupuy, S.; French, A.; ul Hassan Akhtar, I.; Kussul, N.; Lebourgeois, V.; Page, M.L.; Newby, T.; Savin, I.; Verón, S.; Koetz, B.; Defourny, P. Building a Data Set Over 12 Globally Distributed Sites To Support the Development of Agriculture Monitoring Applications With Sentinel-2. *Remote Sensing* **2015**, *7*, 16062–16090. doi:10.3390/rs71215815.
32. Hagolle, O.; Sylvander, S.; Huc, M.; Claverie, M.; Clesse, D.; Dechoz, C.; Lonjou, V.; Poulain, V. SPOT4 (Take5): Simulation of Sentinel-2 Time Series on 45 Large sites. *Remote Sensing* **2015**, *7*, 12242–12264. doi:10.3390/rs70912242.
33. Jordan, M.I.; Ghahramani, Z.; Jaakkola, T.S.; Saul, L.K. An Introduction To Variational Methods for Graphical Models. *Learning in Graphical Models* **1998**, pp. 105–161. doi:10.1007/978-94-011-5014-9_5.
34. Ranganath, R.; Gerrish, S.; Blei, D. Black Box Variational Inference. *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*; Kaski, S.; Corander, J., Eds.; PMLR: Reykjavik, Iceland, 2014; Vol. 33, *Proceedings of Machine Learning Research*, pp. 814–822.
35. Kingma, D.P.; Welling, M. Auto-Encoding Variational Bayes. *CoRR* **2013**, [arXiv:stat.ML/1312.6114v10].